



**FACULTAD DE CIENCIAS DE GESTIÓN  
ESCUELA DE INGENIERÍA DE SISTEMAS**

**TESIS**

**DESARROLLO DE UNA SOLUCIÓN DE BUSINESS INTELLIGENCE PARA  
MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE  
RENTAS DE LA MUNICIPALIDAD DE LURÍN**

**AUTOR**

**VARGAS CHUMPITAZ, FELIX JULIO**

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERO DE SISTEMAS**

**ASESOR:**

**MG. JOSÉ LUIS HERRERA SALAZAR**

**LIMA – PERÚ**

**2016**

### **DEDICATORIA**

Dedicado a mis padres Adela y Felix, quienes me han dado todo su apoyo durante estos 5 años de estudio, durante toda mi vida y quienes estoy seguro estarán ahí en todo momento hasta que el Señor lo permita.

### **AGRADECIMIENTO**

A Dios, pues nos da las fuerzas para seguir adelante y porque sin Él nada sería posible. A mi familia, quienes me dan su apoyo día a día y me alientan en cada decisión que tomo. A mis profesores Luis Camacho Colán, Javier Gamboa Cruzado y José Luis Herrera Salazar, quienes en su momento me brindaron las pautas necesarias para la realización del proyecto. Al ingeniero Juan Carlos Cervantes Aguirre, por permitirme realizar el proyecto en la municipalidad de Lurín.

## RESUMEN

### **DESARROLLO DE UNA SOLUCIÓN DE BUSINESS INTELLIGENCE PARA MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE RENTAS DE LA MUNICIPALIDAD DE LURÍN**

**FELIX VARGAS CHUMPITAZ**  
felix.vargas369@gmail.com

El presente proyecto de tesis consiste en el desarrollo de una solución de Business Intelligence, herramienta que ayudará a mejorar el proceso de toma de decisiones en el área de rentas de la municipalidad de Lurín.

Si bien es cierto que la gerencia de rentas de la municipalidad de Lurín sí usa la información del sistema transaccional para tomar decisiones, el proceso de obtención de reportes resulta demasiado tardío y demanda mucho esfuerzo por parte del personal de informática.

Ante esta situación fue que se planteó el desarrollo de la solución de Business Intelligence, pues haría que el proceso sea mucho más rápido y con menor esfuerzo invertido. Para el desarrollo se empleó la metodología de Ralph Kimball, la cual se adaptó más a nuestro caso de estudio pues está enfocada únicamente a una parte de la empresa o institución.

La solución final para el usuario consistirá en una serie de reportes acerca del estado actual e histórico de las recaudaciones y deudas realizadas por el área, para que así se puedan tomar decisiones basándose de información real.

**Palabras clave:** Inteligencia de negocios, Data Mart, Ralph Kimball, Pentaho, Software libre, Recaudación, Reportes.

## **ABSTRACT**

### **DEVELOPMENT OF A BUSINESS INTELLIGENCE SOLUTION FOR IMPROVING THE DECISION MAKING PROCESS IN THE RENTS AREA OF THE LURIN MUNICIPALITY**

**FELIX VARGAS CHUMPITAZ**  
felix.vargas369@gmail.com

This thesis project consist of the development of a Business Intelligence solution, tool that will help improve the process of decision making in the rents area of the Lurin municipality.

While it is true that the management of rents of the Lurin municipality use the information of transactional system to make decisions, the process of obtaining reports is too late and demand much effort from IT staff.

Given this situation was that the development of a Business Intelligence solution was raised because it would make the process much faster and with less effort involved. To development, Ralph Kimball methodology was used, which is more adapted to our case study because it is focused only part of the company or institution.

The final solution for the user consist of a series of reports on the current and historical status of the collections and debts made by the area, so that decisions can be made based on real information.

**Key Words:** Business Intelligence, Data Mart, Ralph Kimball, Pentaho, Free Software, Collection, Reports.

## INTRODUCCIÓN

La información va tomando cada vez más relevancia, tanto así que día tras día va tomando más importancia como activo para las empresas e instituciones, ya sean públicas o privadas. Pero la información no se obtiene ya “elaborada”, sino que la conseguimos en forma de datos, los cuales deben ser registrados mediante sistemas transaccionales u hojas de cálculo y posteriormente analizados para así poder obtener información útil para la empresa o institución. Como se mencionó, la información va tomando cada vez más importancia. Así podemos inferir que los datos también son importantes para nosotros, pero dicha importancia no va necesariamente relacionada a la cantidad ya que se puede tener una cantidad descomunal de datos en nuestros sistemas pero si no sabemos analizarlas ni explotarlas no nos servirá, será un activo totalmente inútil.

No saber qué hacer con los datos puede ocurrir en cualquier empresa o institución sin importar el rubro en el cual se desempeñe ya que toda empresa maneja algunos datos ya sea de proveedores, clientes, ventas, producción, etc. No necesariamente se tiene que contar con un sistema transaccional pues los datos pueden estar almacenados también en archivos de Excel, cosa que es totalmente válida. El problema en sí es tener muchos datos y no saber aprovecharlos, es decir no apoyarnos de ellos para obtener ventajas competitivas frente a las demás empresas del mercado. Las ventajas que se pueden obtener son por ejemplo conocer qué sector de la población consume más un determinado producto, qué producto se vende más, qué vendedor es el más eficiente o qué tienda es la que factura más al mes. Estas ventajas no son las únicas que podemos obtener, es solamente un ejemplo. Por otro lado tampoco tienen que ser precisamente éstas, ya que pueden variar de acuerdo al rubro de la empresa. Como se puede apreciar, los ejemplos se refieren a alguna empresa de venta de productos a la cual le puede interesar qué local factura más al mes, pero en el caso de un ente recaudador le interesaría qué local recauda más dinero también al mes. Esta es una muestra de que sin importar el rubro al cual una empresa o institución se dedique, siempre se puede obtener ventajas competitivas gracias al buen uso de los datos obtenidos.

El presente proyecto consiste en el desarrollo de una solución de Business Intelligence para mejorar la toma de decisiones en el área de rentas de la Municipalidad de Lurín. Esto permitirá que la gerencia de rentas pueda conocer cómo va el negocio, es decir le brindará una visión del rumbo que se va tomando. Y más allá de conocer la realidad del área, la solución permitirá tomar decisiones más acertadas.

En el capítulo I se realiza el Planteamiento Metodológico, en el cual se ve la realidad problemática. Así mismo se realiza la definición y enunciado del problema. También se define el tipo y el nivel de la investigación realizada. Acto seguido se dejan bien definidos la justificación, los objetivos (generales y específicos) y la hipótesis. Y se termina esta parte con la definición de las variables e indicadores, las limitaciones que tiene la investigación, el diseño de la investigación, y las técnicas e instrumentos utilizados para la recolección de información.

En el capítulo II se toca el Marco Referencial, empezando por los antecedentes de la investigación, es decir el estudio de trabajos realizados similares a éste, trabajos que abarcaron el tema de la inteligencia de negocios para la solución de problemas. Posteriormente se pasa a definir los términos que se consideraron importantes en el desarrollo del proyecto de investigación.

El capítulo III abarca el Desarrollo de la Solución de Business Intelligence. Empezando por una definición de las generalidades, pasando por el estudio de factibilidad tanto técnica y operativa como económica. Finalmente se toma el desarrollo en sí de la solución. Se plasma fase por fase el camino que se siguió para el desarrollo de nuestra solución basándonos en la metodología de Ralph Kimball, la cual también es conocida como Business Dimensional LifeCycle (Ciclo de Vida Dimensional del Negocio).

En el capítulo IV se lleva a cabo el Análisis de Resultados y la Contrastación de la Hipótesis. Se define la población y la muestra utilizadas para la presente investigación, al igual que el nivel de confianza empleado. Luego se realiza el análisis de la información obtenida en la pre prueba y en la post prueba. Se cierra el capítulo con la contrastación de la hipótesis.

El capítulo final, el capítulo V, consiste en las Conclusiones y Recomendaciones que dejaron el desarrollo de la tesis.

Para terminar se muestran las referencias bibliográficas que sirvieron de ayuda en el transcurso de la elaboración de la tesis, los anexos y apéndices que están ahí como apoyo y un glosario de términos para facilitar el entendimiento de algunas palabras.

*El autor*

## TABLA DE CONTENIDOS

DEDICATORIA .....	i
AGRADECIMIENTO.....	ii
RESUMEN.....	iii
ABSTRACT.....	iv
INTRODUCCIÓN .....	v
ÍNDICE DE FIGURAS.....	xi
ÍNDICE DE TABLAS .....	xv

### CAPÍTULO I: PLANTEAMIENTO METODOLÓGICO

1.1. EL PROBLEMA	
1.1.1. Realidad Problemática.....	2
1.1.2. Definición del Problema.....	5
1.1.3. Enunciado del Problema.....	9
1.2. TIPO Y NIVEL DE LA INVESTIGACIÓN	
1.2.1. Tipo de Investigación .....	9
1.2.2. Nivel de Investigación.....	9
1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN.....	9
1.4. OBJETIVOS DE LA INVESTIGACIÓN	
1.4.1. Objetivo General .....	10
1.4.2. Objetivos Específicos .....	10
1.5. HIPÓTESIS.....	10
1.6. VARIABLES E INDICADORES	
1.6.1. Variable Independiente: Solución de Business Intelligence .....	10
1.6.2. Variable Dependiente: Proceso de Toma de Decisiones del Área de Rentas.....	11
1.7. LIMITACIONES DE LA INVESTIGACIÓN.....	12
1.8. DISEÑO DE LA INVESTIGACIÓN .....	12
1.9. TÉCNICAS E INSTRUMENTOS PARA LA RECOLECCIÓN DE DATOS	
1.9.1. Técnicas.....	13
1.9.2. Instrumentos .....	13

### CAPÍTULO II: MARCO REFERENCIAL

2.1. ANTECEDENTES DE LA INVESTIGACIÓN.....	15
2.2. MARCO TEÓRICO	
2.2.1. Business Intelligence.....	21



2.2.2.	Metodología de Ralph Kimball .....	32
2.2.3.	Metodología de Bill Inmon .....	38
2.2.4.	Toma de Decisiones .....	44
2.2.5.	Software libre .....	47
2.2.6.	Pentaho .....	48
2.2.7.	Recaudación Tributaria Municipal .....	48

### **CAPÍTULO III: DESARROLLO DE LA SOLUCIÓN DE BUSINESS INTELLIGENCE**

3.1.	GENERALIDADES .....	51
3.2.	ESTUDIO DE FACTIBILIDAD	
3.2.1.	Factibilidad Técnica .....	51
3.2.2.	Factibilidad Operativa .....	52
3.2.3.	Factibilidad Económica .....	53
3.3.	PLANIFICACIÓN DEL PROYECTO	
3.3.1.	Descripción del Proyecto .....	54
3.3.2.	Objetivos del Proyecto .....	54
3.3.3.	Alcance del Proyecto .....	54
3.3.4.	Stakeholders .....	55
3.3.5.	Análisis de Riesgos .....	56
3.3.6.	Cronograma de Actividades .....	57
3.4.	DEFINICIÓN DE LOS REQUERIMIENTOS DEL NEGOCIO	
3.4.1.	Proceso de negocio: Toma de Decisiones del Área de Rentas .....	58
3.4.2.	Proceso de Negocio y Temas Analíticos .....	59
3.4.3.	Matriz Procesos/Dimensiones .....	60
3.4.4.	Requerimientos .....	60
3.4.5.	Documentación de los Requerimientos .....	61
3.4.6.	Hoja de Gestión .....	64
3.4.7.	Hoja de Análisis .....	65
3.5.	MODELO DIMENSIONAL	
3.5.1.	Dimensiones .....	65
3.5.2.	Dimensión: Contribuyente .....	66
3.5.3.	Dimensión: Cuota .....	66
3.5.4.	Dimensión: Tipo .....	66
3.5.5.	Dimensión: Zona .....	66

3.5.6.	Dimensión: Tiempo.....	67
3.5.7.	Dimensión: Estado .....	67
3.5.8.	Granularidad.....	67
3.5.9.	Hechos.....	67
3.5.10.	Medidas .....	68
3.5.11.	Diseño del Modelo Estrella.....	68
3.6.	<b>DISEÑO DE LA ARQUITECTURA TÉCNICA</b>	
3.6.1.	Arquitectura.....	69
3.6.2.	Back-room.....	69
3.6.3.	Front-room .....	70
3.7.	<b>SELECCIÓN E INSTALACIÓN DE PRODUCTOS</b>	
3.7.1.	Evaluación de los productos.....	70
3.7.2.	Herramientas seleccionadas .....	71
3.8.	<b>DISEÑO FÍSICO</b>	
3.8.1.	Tablas de Apoyo.....	77
3.8.2.	Tabla Dimensión: DIM_Contribuyente.....	78
3.8.3.	Tabla Dimensión: DIM_Zona .....	78
3.8.4.	Tabla Dimension: DIM_Cuota.....	78
3.8.5.	Tabla Dimensión: DIM_TipoPago.....	79
3.8.6.	Tabla Dimensión: DIM_Tiempo .....	79
3.8.7.	Tabla Dimensión: DIM_Estado.....	79
3.8.8.	Tabla de Hechos: FACT_Recaudación .....	79
3.8.9.	Diseño Modelo Físico .....	80
3.9.	<b>DISEÑO ETL</b>	
3.9.1.	Extracción (Extract) .....	81
3.9.2.	Transformación (Transformation).....	83
3.9.3.	Carga (Load) .....	85
3.10.	<b>DISEÑO DEL CUBO OLAP.....</b>	93
3.11.	<b>ESPECIFICACIÓN DE APLICACIONES BI</b>	
3.11.1.	Lista de Reportes.....	94
3.12.	<b>DESARROLLO DE APLICACIONES BI</b>	
3.12.1.	Elaboración de Reportes.....	101
3.13.	<b>IMPLEMENTACIÓN</b>	
3.13.1.	Verificación de Tecnología .....	117

3.13.2. Manual de Usuario .....	118
3.14. MANTENIMIENTO Y CRECIMIENTO	
3.14.1. Pruebas de Funcionamiento del Sistema .....	118

## **CAPÍTULO IV: ANÁLISIS DE RESULTADOS Y CONTRASTACIÓN DE LA HIPÓTESIS**

4.1. POBLACIÓN Y MUESTRA	
4.1.1. Población.....	120
4.1.2. Muestra.....	120
4.2. NIVEL DE CONFIANZA .....	120
4.3. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	
4.3.1. Resultados Genéricos .....	120
4.3.2. Resultados Específicos .....	121
4.4. CONTRASTACIÓN DE LA HIPÓTESIS .....	133

## **CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES**

5.1. CONCLUSIONES .....	142
5.2. RECOMENDACIONES .....	143
REFERENCIAS BIBLIOGRÁFICAS.....	144
APÉNDICES.....	147
GLOSARIO DE TÉRMINOS.....	173

## ÍNDICE DE FIGURAS

Figura 01. Ubicación de la Municipalidad de Lurín .....	4
Figura 02. Cuadro estadístico de la recaudación de arbitrios en la zona A el año 2014 .....	5
Figura 03. Proceso de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín (AS - IS).....	6
Figura 04. Proceso de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín (TO - BE).....	8
Figura 05. Componentes de Business Intelligence.....	21
Figura 06. Componentes del proceso de DataWarehousing.....	25
Figura 07. Ejemplo de cubo OLAP .....	27
Figura 08. Proceso ETL usando un ODS .....	29
Figura 09. Star Schema (Esquema estrella).....	31
Figura 10. Snowflake Schema (Esquema copo de nieve) .....	32
Figura 11. Ciclo de vida de la metodología de Kimball.....	33
Figura 12. Migración al entorno diseñado – Parte I.....	41
Figura 13. Migración al entorno diseñado – Parte II.....	42
Figura 14. Migración al entorno diseñado – Parte III .....	42
Figura 15. La retroalimentación entre el analista de sistema de soporte a la toma de decisiones y el arquitecto de datos.....	43
Figura 16. Procedimiento para la toma de decisiones .....	44
Figura 17. Cronograma de Actividades.....	57
Figura 18. Proceso de Negocio: Toma de Decisiones en el Área de Rentas.....	58
Figura 19. Determinación de nivel de granularidad .....	67
Figura 20. Modelo Dimensional: Recaudación.....	68
Figura 21. Arquitectura Técnica.....	69
Figura 22. Back - Room .....	69
Figura 23. Front - Room.....	70
Figura 24. Pentaho Data Integration (Kettle) .....	72
Figura 25. Pentaho Schema Workbench .....	72
Figura 26. Pentaho Design Studio.....	73
Figura 27. JPivot .....	73
Figura 28. Pentaho Report Designer .....	74
Figura 29. Pentaho BI Server .....	74
Figura 30. CDE .....	75
Figura 31. IvyDC .....	75

Figura 32. PostgreSQL 9.3.....	76
Figura 33. Oracle VM VirtualBox .....	76
Figura 34. Modelo Físico: Recaudación.....	80
Figura 35. Extracción para: Tabla apoyo_cta_aica .....	81
Figura 36. Extracción para: Tabla apoyo_predios_ur .....	81
Figura 37. Extracción para: Tabla apoyo_microzona .....	82
Figura 38. Extracción para: Tabla Dim_Contribuyente .....	82
Figura 39. Extracción para: Tabla Dim_Microzona.....	82
Figura 40. Extracción para: Tabla Dim_Cuota .....	83
Figura 41. Extracción para: Tabla Dim_Tiempo.....	83
Figura 42. Transformación para: Tabla apoyo_cta_aica.....	83
Figura 43. Transformación para: Tabla apoyo_predios_ur .....	84
Figura 44. Transformación para: Tabla apoyo_microzona .....	84
Figura 45. Transformación para: Tabla Dim_Contribuyente .....	84
Figura 46. Transformación para: Tabla Dim_Zona.....	84
Figura 47. Transformación para: Tabla Dim_Cuota .....	84
Figura 48. Transformación para: Tabla Dim_TipoPago .....	84
Figura 49. Transformación para: Tabla Dim_Estado .....	85
Figura 50. Transformación para: Tabla Dim_Tiempo .....	85
Figura 51. Transformación para: Tabla FACT_Recaudacion .....	85
Figura 52. Trabajo: Trabajo ETL parte 1 .....	86
Figura 53. Transformación: Carga tabla apoyo cta aica.....	86
Figura 54. Transformación: Carga tabla apoyo predios ur.....	87
Figura 55. Transformación: Carga tabla apoyo microzona.....	87
Figura 56. Trabajo: Trabajo ETL parte 2 .....	88
Figura 57. Transformación: Carga Dimensión Contribuyente .....	88
Figura 58. Transformación: Carga Dimensión Zona.....	89
Figura 59. Transformación: Carga Dimensión Tipo .....	89
Figura 60. Transformación: Carga Dimensión Cuota .....	90
Figura 61. Transformación: Carga Dimensión Tiempo .....	90
Figura 62. Transformación: Carga Dimensión Estado .....	91
Figura 63. Transformación: Carga Hecho Recaudación .....	91
Figura 64. Action Sequences con Pentaho Design Studio .....	92
Figura 65. Action Sequences en la consola de usuario Pentaho BI Server .....	92
Figura 66. Programación de los action sequences.....	93

Figura 67. Creación del Cubo en Pentaho Schema Workbench.....	93
Figura 68. Publicación del cubo en Pentaho BI Server .....	94
Figura 69. Prototipo: Reporte Impuesto Predial por Cuotas .....	94
Figura 70. Prototipo: Reporte Arbitrios Municipales por Cuotas .....	95
Figura 71. Prototipo: Reporte Arbitrios Municipales por Zonas.....	95
Figura 72. Prototipo: Reporte Impuesto Predial – Arbitrios Municipales por Fechas .....	96
Figura 73. Prototipo: Dashboard Top Contribuyentes y Recaudación por Zonas.....	96
Figura 74. Prototipo: Dashboard No Top Contribuyentes y Deudas por Zonas .....	97
Figura 75. Prototipo: Dashboard Recaudación por Microzonas, por Años y por Cuotas .....	97
Figura 76. Prototipo: Dashboard Deudas por Microzonas, por Años y por Cuotas .....	98
Figura 77. Prototipo: Dashboard Cantidad Contribuyentes que Pagan.....	98
Figura 78. Prototipo: Dashboard KPI de Impuesto Predial.....	99
Figura 79. Prototipo: Dashboard KPI de Arbitrios Municipales.....	99
Figura 80. Prototipo: Cubo de Tipos de Pagos por Cuotas .....	100
Figura 81. Prototipo: Cubo de Arbitrios Municipales por Zonas.....	100
Figura 82. Prototipo: Cubo de Tipos de Pagos por Fechas .....	101
Figura 83. Reporte Recaudación Impuesto Predial por Cuotas.....	102
Figura 84. Reporte Deudas Impuesto Predial por Cuotas .....	103
Figura 85. Reporte Recaudación Arbitrios Municipales por Cuotas.....	104
Figura 86. Reporte Deudas Arbitrios Municipales por Cuotas .....	105
Figura 87. Reporte Recaudación Arbitrios Municipales por Zonas.....	106
Figura 88. Reporte Deudas Arbitrios Municipales por Zonas.....	107
Figura 89. Reporte Impuesto Predial – Arbitrios. Municipales por Fechas .....	108
Figura 90. Dashboard Top Contribuyentes y Recaudación por Zonas.....	109
Figura 91. Dashboard Recaudación por Microzonas, por Años y por Cuotas .....	110
Figura 92. Dashboard No Top Contribuyentes y Deudas por Zonas .....	111
Figura 93. Dashboard Deudas por Microzonas, por Años y por Cuotas .....	112
Figura 94. Dashboard Cantidad Contribuyentes que Pagan.....	113
Figura 95. Dashboard KPI de Impuesto Predial.....	113
Figura 96. Dashboard KPI de Arbitrios Municipales.....	114
Figura 97. Cubo Recaudación Tipos de Pagos por Cuotas .....	114
Figura 98. Cubo Recaudación Arbitrios Municipales por Zonas .....	115
Figura 99. Cubo Deudas Tipos de Pagos por Cuotas .....	115
Figura 100. Cubo Deudas Arbitrios Municipales por Zonas.....	116
Figura 101. Cubo Recaudación por Fechas.....	116

Figura 102. Estadística Descriptiva para $KPI_1$ (Post).....	124
Figura 103. Estadística Descriptiva para $KPI_2$ (Post).....	126
Figura 104. Estadística Descriptiva para $KPI_3$ (Post).....	128
Figura 105. Estadística Descriptiva para $KPI_4$ (Post).....	130
Figura 106. Distribución de Probabilidad $KPI_1$ .....	134
Figura 107. Distribución de Probabilidad $KPI_2$ .....	136
Figura 108. Distribución de Probabilidad $KPI_3$ .....	137
Figura 109. Distribución de Probabilidad $KPI_4$ .....	139

## ÍNDICE DE TABLAS

Tabla 01. Valores actuales de los indicadores.....	7
Tabla 02. Reporte Recaudación por Zonas .....	7
Tabla 03. Variable independiente: Indicadores.....	10
Tabla 04. Variable independiente: Índices .....	11
Tabla 05. Variable dependiente: Indicadores .....	11
Tabla 06. Variable dependiente: Índices .....	11
Tabla 07. Técnicas para la recolección de datos .....	13
Tabla 08. Instrumentos para la recolección de datos.....	13
Tabla 09. Temas analíticos.....	35
Tabla 10. Matriz de procesos / dimensiones .....	35
Tabla 11. Hardware disponible .....	51
Tabla 12. Software Disponible.....	52
Tabla 13. Costos del Desarrollo de la Solución .....	53
Tabla 14. Principales Stakeholders del Proyecto .....	55
Tabla 15. Funciones de Equipo de Trabajo.....	55
Tabla 16. Matriz de Riesgo .....	56
Tabla 17. Temas Analíticos basados en entrevistas .....	59
Tabla 18. Procesos de Negocio basados en entrevistas.....	60
Tabla 19. Matriz Procesos/Dimensiones.....	60
Tabla 20. Lista de Requerimientos.....	60
Tabla 21. Hoja de Gestión: Recaudación .....	64
Tabla 22. Hoja de Análisis: Recaudación .....	65
Tabla 23. Dimensiones.....	65
Tabla 24. Dimensión Contribuyente .....	66
Tabla 25. Dimensión Cuota.....	66
Tabla 26. Dimensión Tipo.....	66
Tabla 27. Dimensión Zona.....	66
Tabla 28. Dimensión Tiempo.....	67
Tabla 29. Dimensión Estado .....	67
Tabla 30. Tablas de Hechos: Recaudación.....	68
Tabla 31. Medidas .....	68
Tabla 32. Matriz de Evaluación .....	71
Tabla 33. Diseño Físico: Apoyo Cta Aica.....	77
Tabla 34. Diseño Físico: Apoyo Microzona .....	77



Tabla 35. Diseño Físico: Apoyo Predios UR .....	78
Tabla 36. Diseño Físico: Dimensión Contribuyente .....	78
Tabla 37. Diseño Físico: Dimensión Microzona.....	78
Tabla 38. Diseño Físico: Dimensión Cuota .....	78
Tabla 39. Diseño Físico Dimensión TipoPago.....	79
Tabla 40. Diseño Físico: Dimensión Tiempo.....	79
Tabla 41. Diseño Físico: Dimensión Estado .....	79
Tabla 42. Diseño Físico: Hecho Recaudación.....	79
Tabla 43. Verificación de la tecnología.....	117
Tabla 44. Pruebas de Funcionamiento del Sistema.....	118
Tabla 45. Resultados de la Pre-Prueba y Post-Prueba para todos los KPI. ....	122
Tabla 46. Resultados de Pre-Prueba y Post-Prueba para el KPI <sub>1</sub> .....	123
Tabla 47. Resultados de Pre-Prueba y Post-Prueba para el KPI <sub>2</sub> .....	125
Tabla 48. Resultados de Pre-Prueba y Post-Prueba para el KPI <sub>3</sub> .....	127
Tabla 49. Resultados de Pre-Prueba y Post-Prueba para el KPI <sub>4</sub> .....	129
Tabla 50. Indicadores para la Contrastación de la Hipótesis.....	133
Tabla 51. Resumen de Prueba t student para el KPI <sub>1</sub> .....	134
Tabla 52. Resumen de Prueba t student para el KPI <sub>2</sub> .....	136
Tabla 53. Resumen de Prueba t student para el KPI <sub>3</sub> .....	138
Tabla 54. Resumen de Prueba t student para el KPI <sub>4</sub> .....	139

**CAPÍTULO I:**  
**PLANTEAMIENTO METODOLÓGICO**

*Hay una fuerza motriz más poderosa que  
el vapor, la electricidad y la energía  
atómica. Esa fuerza es la voluntad.*

**Albert Einstein**

**1879 - 1955**

## 1.1. EL PROBLEMA

### 1.1.1. Realidad Problemática

Sabemos que tomar una decisión consiste en elegir entre dos o más alternativas. Y en efecto, es una acción que realizamos diariamente ya que elegimos por ejemplo con qué ropa nos vestiremos o qué es lo que comeremos en el almuerzo. Sin embargo, cuando trasladamos esto a un ámbito empresarial las cosas ya no son tan sencillas. Y decimos esto porque <sup>1</sup>el proceso de toma de decisiones, que continuamente ha sido catalogado como el motor que le da vida a los negocios, es el proceso de donde depende el triunfo de cualquier organización a través de la correcta elección de alternativas.

Vemos así, que si como directivos tomamos decisiones acertadas, lograremos el éxito en nuestra organización. Pero si por el contrario, tomamos decisiones erróneas, nos veremos ante una situación muy desfavorable. Con esto vemos la gran importancia que tiene el realizar una correcta toma de decisiones.

A lo anteriormente dicho, podemos agregar que <sup>2</sup>en el ámbito organizacional, la mayoría de las decisiones significativas se realizan mediante el juicio. Lo cual indica una falta de apoyo en la información.

Acá podemos hacer énfasis en la importancia de la información en el proceso de toma de decisiones. Y ante esto podemos preguntarnos, ¿qué papel cumple aquí la información? La respuesta que podemos dar es que <sup>3</sup>la información es la materia prima, el input de la decisión, y una vez tratada dentro del proceso de toma de decisiones se obtiene como output de la acción a ejecutar. Esto, visto desde un plano sistémico, deja en evidencia que la información va más allá de ser considerado como algo importante, sino más bien como algo necesario para el proceso de toma de decisiones. Cosa que es apoyada por Villanueva<sup>4</sup> al considerar que la toma de decisiones que se lleva a cabo dentro de las organizaciones debe cumplir con ciertas características, una de las cuales es ser fundamentada en información concreta.

Con todo esto vemos que el problema a la hora de tomar decisiones es no basarse en información para elegir la mejor alternativa. Y como podemos darnos cuenta, todo sigue una cadena, es decir comenzamos diciendo que el proceso de toma de decisiones es importante porque determina el éxito de la empresa, pero a su vez algo importante (y necesario) para realizar el proceso de manera óptima es la información. Aquí debemos recalcar que la información debe ser, primero que nada, verídica, exacta y precisa. Y por supuesto también debe ser oportuna, o sea que se encuentre accesible para quien la requiera y en el momento en que lo requiera.

---

<sup>1</sup> Villanueva, L., La Toma de Decisiones en la Organización y el Gran Valor del Profesional de la..., 2015

<sup>2</sup> Chacín, L., Tecnología de Información como Apoyo al Proceso..., Revista Espacios, 2010, Vol. 31, p. 13

<sup>3</sup> VV.AA., Dirección Estratégica, Ed. Vértice, España, 2007, p. 63

<sup>4</sup> Villanueva, L., Op. Cit.

## RECAUDACIÓN TRIBUTARIA MUNICIPAL EN EL PERÚ

En el Perú existen 1829 municipalidades, de ellas 194 son provinciales y 1635 distritales. Todas cuentan con grandes diferencias en cuanto al tamaño poblacional y territorial. Sin embargo, todas tienen a su cargo los mismos tributos municipales, los cuales son:

- Impuesto predial: este impuesto grava el valor de los predios urbanos y rústicos, como los terrenos, edificaciones, instalaciones fijas y permanentes que constituyan parte integrante del mismo, que no puedan ser separadas sin alterar o destruir la edificación.
- Impuesto al patrimonio vehicular: Impuesto que grava la propiedad de los vehículos, automóviles, camionetas, station wagon, camiones, buses y omnibuses, con una antigüedad no mayor de 3 años contados a partir de su primera inscripción en el Registro de Propiedad Vehicular.
- Impuesto de alcabala: Impuesto de realización inmediata que grava las transferencias de propiedad de bienes inmuebles urbanos o rústicos a título oneroso o gratuito, cualquiera sea su forma o modalidad, inclusive las ventas con reserva de dominio.
- Impuesto a los juegos: Es un tributo que grava la realización de actividades relacionadas con los juegos, tales como loterías, bingos y rifas, así como la obtención de premios en juegos de azar.
- Impuesto a las apuestas: Impuesto que grava los ingresos de las entidades organizadoras de eventos hípicas y similares, en las que se realice apuestas.
- Impuesto a los espectáculos públicos no deportivos: Impuesto que grava el monto que se abona por concepto de ingreso a espectáculos públicos no deportivos en locales y parques cerrados.
- Arbitrios municipales: Son una contraprestación por la prestación de un servicio público por parte de las municipalidades; es decir; son aquellas tasas que se pagan por la prestación, mantenimiento del servicio público, individualizado en el contribuyente. Los arbitrios son por recolección de residuos sólidos, barrido de calles, parques y jardines públicos, y serenazgo.

Otra cosa que caracteriza a la recaudación tributaria municipal en el Perú es que<sup>5</sup> es generalmente percibido como un trabajo poco grato, pues las personas evitan con frecuencia el pago de los impuestos. Pocas son las personas que ven a la recaudación tributaria municipal como una oportunidad de contribuir a generar recursos para promover el desarrollo en la localidad y transformar el entorno en el cual viven.

---

<sup>5</sup> Dirección General de Presupuesto Público, Manual para la Mejora de la Recaudación del Impuesto..., 2011

Esto, que podríamos considerarlo como una cultura de no pago, es indirectamente apoyada por las mismas gestiones municipales ya que son éstas las que promueven y decretan las famosas amnistías, las cuales consisten en la exoneración total o parcial de las deudas e intereses. Así, las personas no se preocupan en cancelar sus impuestos según el cronograma de pagos establecido, pues si bien es cierto en teoría el no pagar a la fecha genera intereses, en la práctica éstos no se llegan a cobrar debido a las amnistías. De esta manera las personas pueden pagar sus impuestos a destiempo y no verse afectados económicamente por ello.

## MUNICIPALIDAD DE LURÍN

Es la institución del estado, facultada para ejercer el gobierno del distrito de Lurín. Se encarga de la adecuada y transparente gestión de los recursos disponibles en base a las necesidades de la población. La gestión municipal está conformada por tres componentes clave:

- Capacidad de respuesta: Hacer lo correcto. Brindar servicios y obras que respondan a las necesidades de la población.
- Gestión responsable de recursos: Hacerlo bien. Diseñar y ejecutar proyectos con un costo realista y en un tiempo razonable.
- Rendición social de cuentas: Rendir cuentas por lo hecho. Dar cuenta de los montos invertidos y de los resultados alcanzados en favor de la mejor calidad de vida de población y el desarrollo local.

La municipalidad de Lurín se encuentra ubicada en Plaza de Armas s/n, Lurín, Lima – Perú. (Ver Figura 01).



**Figura 01.** Ubicación de la Municipalidad de Lurín

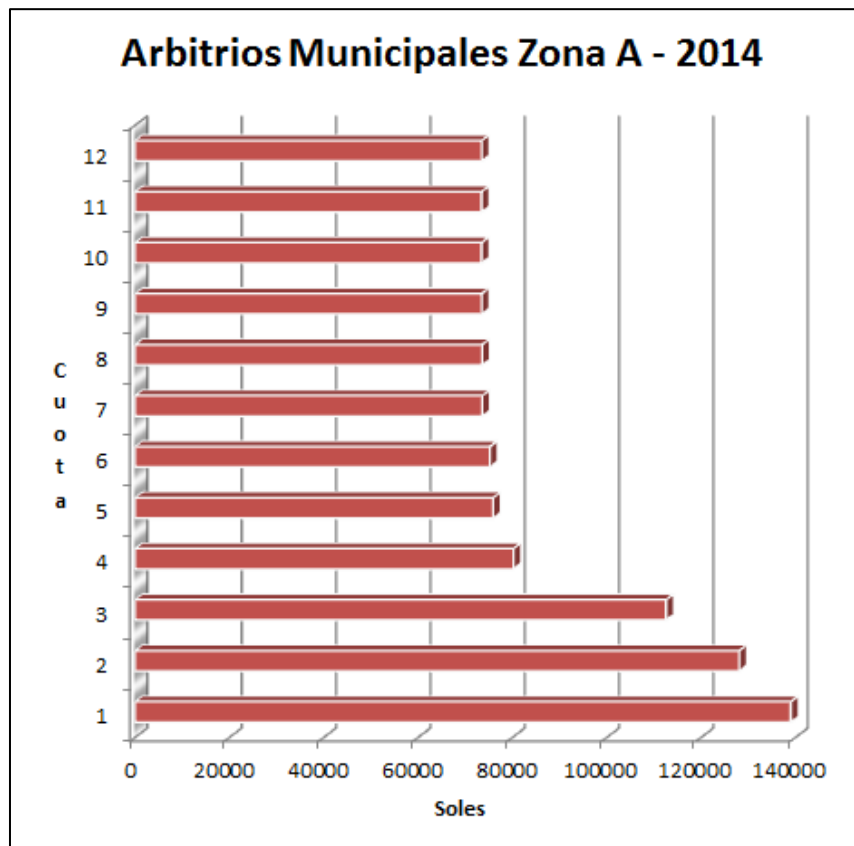
Fuente: <http://google.com/maps>

### 1.1.2. Definición del Problema

El problema en cuestión se centra en el área de rentas de la Municipalidad de Lurín, en cuya gerencia el proceso de toma decisiones muchas veces no hace uso de la información actual del estado de las recaudaciones de impuestos y arbitrios. Esto es debido a que el sistema transaccional (el llamado sistema de recaudación) no permite conocer por ejemplo en qué zonas o microzonas del distrito se recaudan los mayores montos por impuestos prediales y arbitrios municipales, o quiénes son los contribuyentes que pagan los mayores montos (los PRICOS, principales contribuyentes).

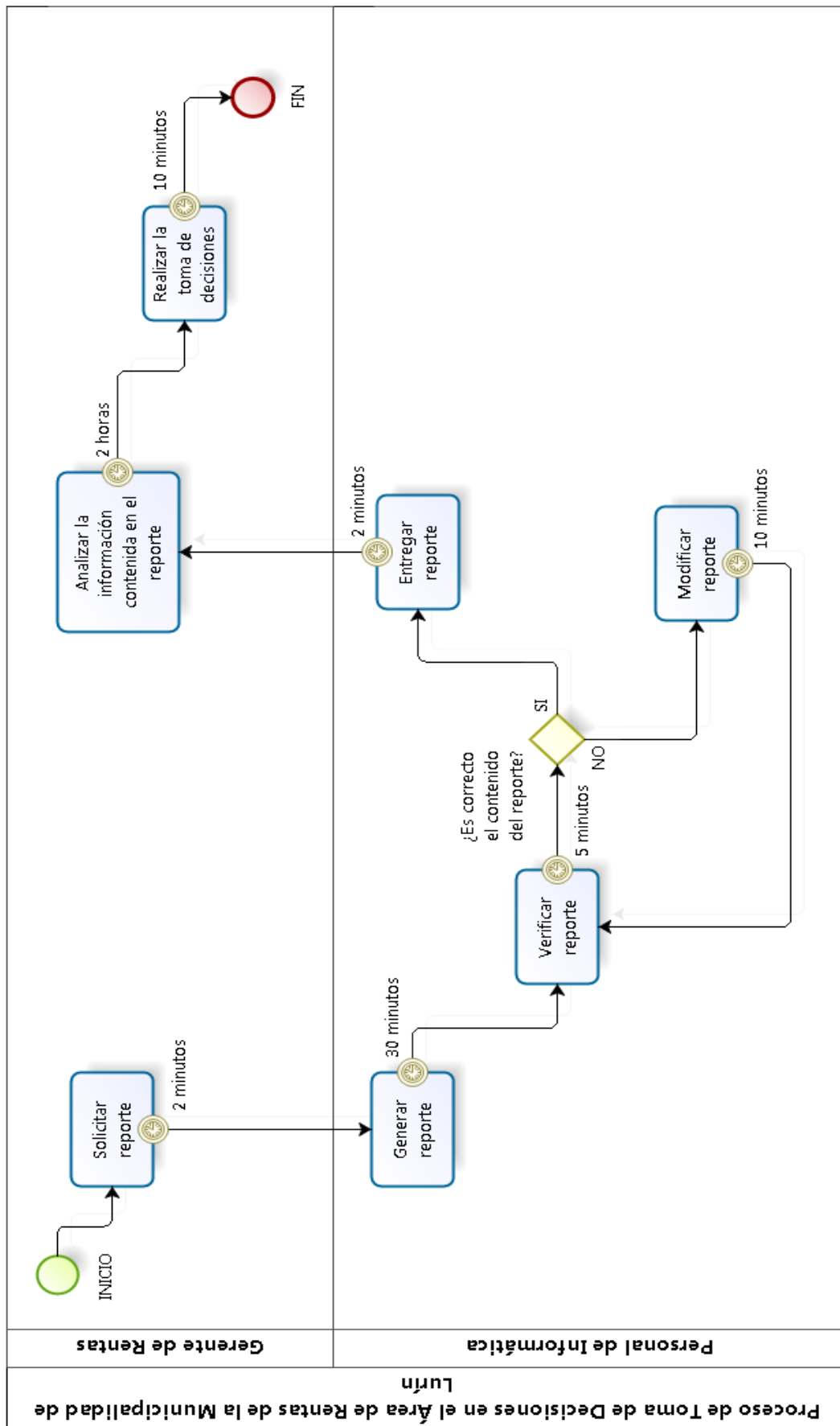
Es por esto que para obtener dicha información, la gerencia de rentas debe solicitar reportes al área de informática, cuyo personal los genera con datos obtenidos por medio de sentencias SQL aplicadas directamente a la base de datos, lo cual demanda mucho tiempo y esfuerzo.

Aun así, los reportes obtenidos no muestran la información solicitada, sino que la gerencia misma debe realizar un ordenamiento de los datos y luego recién mediante gráficos, visualizar el estado de las recaudaciones. Es decir, los reportes no son nada amigables y sobre todo no nos muestran información basada en los indicadores claves de desempeño (KPI).



**Figura 02.** Cuadro estadístico de la recaudación de arbitrios en la zona A el año 2014

Fuente: Área de Informática



**Figura 03.** Proceso de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín (AS - IS)

Fuente: Elaboración Propia

Del proceso expuesto podemos ver e inferir problemas con:

- El tiempo empleado en la generación de reportes.
- El tiempo que el usuario emplea en el análisis de la información.
- El número de veces que el usuario accede a la información al día.
- El porcentaje de exactitud de la información.
- El nivel de satisfacción del usuario.

**Tabla 01.** Valores actuales de los indicadores

Fuente: Elaboración Propia

Indicador	Datos de Pre Prueba (Promedio)
Tiempo empleado en la generación de reportes	30 minutos
Tiempo que el usuario emplea en el análisis de la información	2 horas
Número de veces que el usuario accede a la información al día	0
Porcentaje de exactitud de la información	82%
Nivel de satisfacción del usuario	Bajo

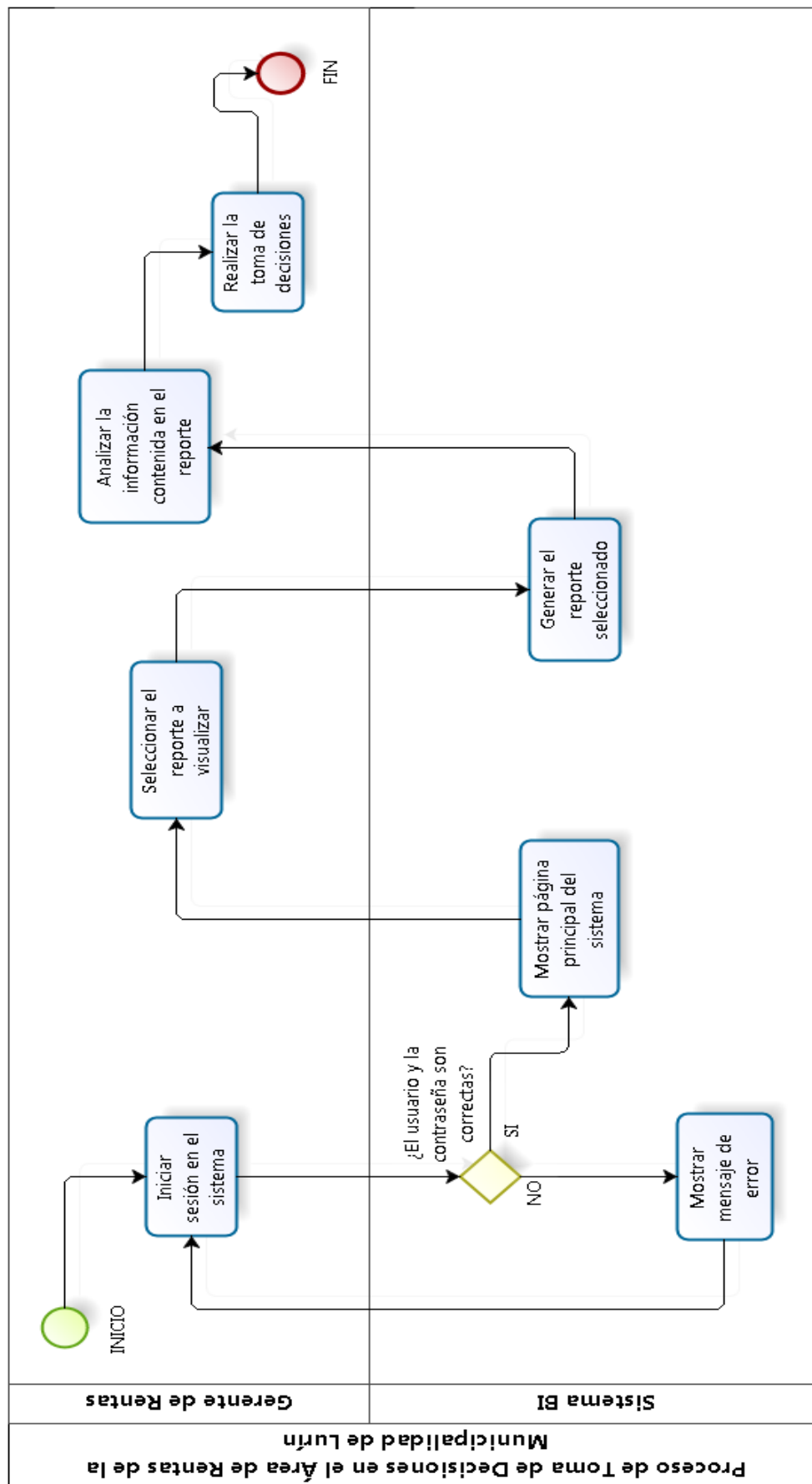
En la Tabla 02 podemos apreciar un reporte de los montos totales recaudados por zonas. Con este reporte el gerente de rentas puede observar cómo van las recaudaciones de arbitrios municipales por zonas, y medir el desempeño de los sectoristas, que son las personas encargadas de la recaudación por cada una de las zonas que conforman el distrito de Lurín.

**Tabla 02.** Reporte Recaudación por Zonas

Fuente: Área de Informática

ZONA	ARBITRIOS MUNICIPALES RECAUDADOS 2013
Zona A	S/. 2 127 509
Zona B	S/. 968 651
Zona C	S/. 294 495
Zona D	S/. 292 790
Zona E	S/. 147 513
<b>TOTAL</b>	<b>S/. 3 830 958</b>





**Figura 04.** Proceso de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín (TO - BE)

Fuente: Elaboración Propia

### **1.1.3. Enunciado del Problema**

¿En qué medida el uso de Business Intelligence mejorará el Proceso de Toma de Decisiones en el área de Rentas de la Municipalidad de Lurín?

## **1.2. TIPO Y NIVEL DE LA INVESTIGACIÓN**

### **1.2.1. Tipo de Investigación**

Cuantitativa: Porque se investiga por medio de técnicas estadísticas.

Aplicada: Porque aplica teorías especializadas con el tema de investigación.

### **1.2.2. Nivel de Investigación**

Correlacional: Porque nos permite conocer cómo se comporta una variable conociendo el comportamiento de otra.

## **1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN**

### **Científica**

La teoría que se emplea es la metodología llamada Ciclo de Vida Dimensional del Negocio<sup>6</sup>, la cual fue un trabajo de muchos años por parte de Ralph Kimball. Dicha metodología defiende el enfoque bottom-up, es decir un enfoque ascendente para diseñar un almacén de datos o Data Warehouse, la misma que Kimball definió como “la unión de todos los Data Mart de una entidad”.

La importancia que tiene en este proyecto es que gracias a esta metodología se podrá elaborar la solución de Business Intelligence desde cero, en otras palabras servirá de guía o manual para el desarrollo del producto, el cual tiene como fin mejorar el proceso de toma de decisiones.

### **Institucional**

Actualmente, en el área de rentas de la Municipalidad de Lurín se necesita conocer los montos de dinero que pagan y adeudan los contribuyentes, ya sea por cuotas, zonas, microzonas o fechas. De igual modo se necesita saber qué significan mencionadas cifras para así tomar decisiones acertadas.

Ante esta situación, el desarrollo de la solución de Business Intelligence busca una optimización del proceso de toma de decisiones del área de Rentas.

### **Social**

El beneficio social que brinda este proyecto es que al solucionar el problema en la institución, puede servir de ejemplo o modelo para otras instituciones que adolezcan del mismo problema. Abriendo de este modo la posibilidad de que en un futuro puedan aplicarlo también, y por ende obtener una solución viable.

---

<sup>6</sup> Kimball, R., et Al., The Data Warehouse Lifecycle..., Ed. Wiley, EE.UU., 2008, Traducción propia, p. 2.2

Por otro lado podemos decir que el beneficio social también va muy ligado a un beneficio académico, pues el presente proyecto también servirá de guía para los ciclos posteriores que realicen proyectos de tesis relacionados al tema del Business Intelligence.

## 1.4. OBJETIVOS DE LA INVESTIGACIÓN

### 1.4.1. Objetivo General

Desarrollar una solución de Business Intelligence para mejorar el proceso de Toma de Decisiones en el área de rentas de la Municipalidad de Lurín.

### 1.4.2. Objetivos Específicos

- Disminuir el tiempo empleado en la generación de reportes para el área de rentas de la Municipalidad de Lurín.
- Disminuir el tiempo que el usuario emplea en el análisis de la información contenida en los reportes.
- Aumentar el número de veces al día que el usuario accede a la información sobre las recaudaciones.
- Aumentar el porcentaje de exactitud de la información contenida en los reportes.
- Aumentar el nivel de satisfacción que tiene el usuario frente a la obtención de reportes.

## 1.5. HIPÓTESIS

El desarrollo de una solución de Business Intelligence mejorará el proceso de Toma de Decisiones del área de Rentas de la Municipalidad de Lurín

## 1.6. VARIABLES E INDICADORES

### 1.6.1. Variable Independiente: Solución de Business Intelligence

#### a) INDICADORES

**Tabla 03.** Variable independiente: Indicadores

Fuente: Elaboración Propia

#### INDICADOR: Presencia - Ausencia

Cuando indique NO, es porque no ha sido desarrollada la solución de Business Intelligence en el área de rentas de la municipalidad de Lurín y aún tenemos la situación problemática. Y cuando indique SI, es porque ya se ha desarrollado la solución de Business Intelligence en el área, esperando mejorar el proceso de toma de decisiones.

## b) ÍNDICE

Tabla 04. Variable independiente: Índices

Fuente: Elaboración Propia

INDICADOR	ÍNDICE
Presencia - Ausencia	NO, SI

## 1.6.2. Variable Dependiente: Proceso de Toma de Decisiones del Área de Rentas

## a) INDICADORES

Tabla 05. Variable dependiente: Indicadores

Fuente: Elaboración Propia

INDICADOR	DESCRIPCIÓN
Tiempo empleado en la generación de reportes	Tiempo utilizado para la generación de reportes.
Tiempo que el usuario emplea en el análisis de la información	Tiempo que emplea el usuario para analizar la información contenida en los reportes.
Número de veces que el usuario accede a la información al día	Cantidad de veces que el usuario hace uso de la información en un día.
Porcentaje de exactitud de la información	Porcentaje de exactitud de la información contenida en los reportes.
Nivel de satisfacción del usuario	Grado en que el usuario está conforme con el contenido de los reportes

## b) ÍNDICE

Tabla 06. Variable dependiente: Índices

Fuente: Elaboración Propia

INDICADOR	ÍNDICE	UNIDAD DE MEDIDA	UNIDAD DE OBSERVACIÓN
Tiempo empleado en la generación de reportes	[1...30]	Minutos	Reloj
Tiempo que el usuario emplea en el análisis de la información	[1...120]	Minutos	Reloj
Número de veces que el usuario accede a la información al día	[1...5]	Número de veces	Usuario
Porcentaje de exactitud de la información	[1...100]	Porcentaje	Reportes generados
Nivel de satisfacción del usuario	[Bajo, Regular, Alto]	-	Usuario

### 1.7. LIMITACIONES DE LA INVESTIGACIÓN

- ✓ La presente investigación abarca solamente el área de rentas de la Municipalidad de Lurín.
- ✓ Se comprende el periodo de tiempo desde abril a diciembre del año 2014.
- ✓ La información de la empresa debe ser usada solamente con fines educativos y a favor de la investigación.
- ✓ Se tiene solamente las mañanas y las tardes para tener acceso al centro de información.
- ✓ Solamente se trabajará con la infraestructura tecnológica con la que cuenta la subgerencia de informática.

### 1.8. DISEÑO DE LA INVESTIGACIÓN

Pre Experimental: Porque demostrará la hipótesis a través de métodos experimentales. No tiene un Grupo Control para comparación de resultados. Sólo se hará en una entidad específica.

**Ge O<sub>1</sub> X O<sub>2</sub>**

#### **Dónde:**

**Ge:** Grupo experimental conformado por un número de procesos de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín.

**O<sub>1</sub>:** Los valores de los indicadores de la variable dependiente durante la Pre-Prueba, es decir antes del desarrollo de la solución de Business Intelligence.

**X:** El desarrollo de la solución de Business Intelligence.

**O<sub>2</sub>:** Los valores de los indicadores de la variable dependiente durante la Post-Prueba, es decir después del desarrollo de la solución de Business Intelligence.

#### **Descripción**

Consiste en la conformación de un grupo experimental (Ge) formado por un número de procesos de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín, a cuyos indicadores se les realiza una Pre-Prueba (O<sub>1</sub>), luego se realiza el desarrollo de la solución de Business Intelligence (X) para solucionar el problema con el proceso y finalmente se realiza la Post-Prueba (O<sub>2</sub>) para ver los resultados.

## 1.9. TÉCNICAS E INSTRUMENTOS PARA LA RECOLECCIÓN DE DATOS

### 1.9.1. Técnicas

**Tabla 07.** Técnicas para la recolección de datos

Fuente: Elaboración Propia

TÉCNICA	USO
Observación	Se observó cómo es el proceso de Toma de Decisiones en el área de Rentas. Se observó la participación tanto del gerente de rentas como del subgerente de informática. Todo esto con el fin de obtener un punto de vista propio acerca del proceso.
Entrevista	Se entrevistó al gerente de rentas, y al subgerente de informática.
Revisión documentaria	Se revisó tesis que guarden relación al tema. Igualmente libros que apoyan con conceptos utilizados.

### 1.9.2. Instrumentos

**Tabla 08.** Instrumentos para la recolección de datos

Fuente: Elaboración Propia

INSTRUMENTO	USO
Cuestionario	Se empleó dos tipos de cuestionarios, uno para el gerente de rentas y el otro para el subgerente de informática. Esto con el fin de obtener información tanto por parte de quien solicita la información como de parte de quien realiza los reportes.
Libreta de notas	Se utilizó una libreta de notas común y corriente para anotar pequeños detalles relacionados al proceso de obtención de reportes. Así como también para anotar los datos técnicos de la infraestructura que se utiliza en la institución.

**CAPÍTULO II:**  
**MARCO REFERENCIAL**

*Quien sólo vive para sí, está  
muerto para los demás.*

**Publio Siro**

**85 a. C. – 43 a. C.**

## 2.1. ANTECEDENTES DE LA INVESTIGACIÓN

**A) Autor:** William David Bazán Vera

**Título:** Desarrollo de un Data Mart con indicadores financieros como soporte para la toma de decisiones en el departamento financiero del Gobierno Autónomo Descentralizado Municipal San Francisco de Milagro<sup>7</sup>.

**Resumen:**

El presente trabajo muestra el diseño de un modelo lógico multidimensional para la implementación de un Data Mart, que sirve como base para la toma de decisiones del área financiera del gobierno autónomo descentralizado municipal San Francisco de Milagro.

Este trabajo intenta dar una idea bastante acertada sobre el trabajo con sistemas Data Mart, presentando la teoría fundamental que sustenta este tipo de representación y procesamiento de los datos en la actualidad: El Modelo Dimensional (DFM: Dimension Fact Model).

Además brinda ejemplos prácticos de cómo llevar a cabo tareas primordiales de un Data Mart, como son: La extracción de los datos desde otros sistemas o formatos. Mostrando así las facilidades que brinda la herramienta de transformación de datos de Microsoft SQL Server 2008 (DTS: Data Transformation Service).

En otros apartados posteriores trata de resumir el tema de la programación distribuida, las diferentes técnicas de distribución de los datos orientado principalmente a Microsoft SQL Server 2008, las diferentes estrategias para mejorar el hardware de un Data Mart Financiero.

El Data Mart Financiero que se implementará, es un sistema orientado a la consulta, en el que se producen procesos de carga de datos. Es consultado mediante herramientas OLAP (OnLine Analytical Processing) proporcionan una alternativa a los sistemas transaccionales, ofreciendo una visión de los datos orientada hacia el análisis, una rápida y flexible navegación por estos.

---

<sup>7</sup> Bazán, W., Desarrollo de un Datamart con..., Universidad Estatal de Milagro, Ecuador, 2011, pp. 7-14



**B) Autor:** Melsi Ocas Terrones

**Título:** Desarrollo de un Data Mart en el área de administración y finanzas de la Municipalidad Distrital de los Baños del Inca<sup>8</sup>.

**Resumen:**

El presente trabajo de tesis desarrolla un Data Mart para el apoyo al proceso de toma de decisiones del área de administración y finanzas de la Municipalidad Distrital de los Baños del Inca.

Debido a que sus sistemas actuales no soportan el manejo adecuado de grandes volúmenes de información, tienen el problema de utilizar su información para emplearla en la toma de decisiones de la institución.

Para llevar adelante el desarrollo del Data Mart se utilizó la metodología de Kimball, conformada por las siguientes etapas:

- ▶ **Planificación:** En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.
- ▶ **Análisis de requerimientos:** Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo.
- ▶ **Modelado Dimensional:** El proceso de diseño comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados.
- ▶ **Diseño del sistema de Extracción, Transformación y Carga (ETL):** El sistema de Extracción, Transformación y Carga (ETL) es la base sobre la cual se alimenta el Data Warehouse. Si el sistema ETL se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos
- ▶ **Especificación y desarrollo de aplicaciones de BI:** Las aplicaciones de BI son la cara visible de la inteligencia de negocios: los informes y aplicaciones de análisis proporcionan información útil a los usuarios.

---

<sup>8</sup> Ocas, M., Desarrollo de un Data Mart en el Área..., Universidad Privada del Norte, Perú, 2012, p. viii

**C) Autor:** Fiorelly Shirley Guillén Rodríguez

**Título:** Desarrollo de un Data Mart para mejorar la toma de decisiones en el área de tesorería de la Municipalidad Provincial de Cajamarca<sup>9</sup>.

**Resumen:**

El presente proyecto de tesis tiene como objetivo desarrollar un Data Mart para el apoyo en la toma de decisiones del área de Tesorería de la Municipalidad Provincial de Cajamarca.

La Municipalidad Provincial de Cajamarca es una institución que maneja una gran cantidad de información día a día. Sin embargo, no saben cómo administrar adecuadamente todas las cantidades recaudadas, debido a que su sistema actual no soporta el manejo adecuado de grandes volúmenes de información. Así la Municipalidad Provincial de Cajamarca tiene el problema de utilizar su información para emplearla en la toma de decisiones.

El objetivo principal es proveer una solución de Inteligencia de Negocios que de soporte a las necesidades de información hacia los usuarios finales, que vienen a ser las demás áreas de la Municipalidad Provincial de Cajamarca, del área de Tesorería.

La solución consiste en brindar una gama de reportes de análisis gerenciales que permitan visualizar las recaudaciones que obtienen para las demás áreas existentes dentro de la Municipalidad Provincial de Cajamarca. Los resultados obtenidos son los reportes para la toma de decisiones de forma rápida y confiable.

Se decide usar herramientas de software libre para la base de datos y para los procesos de extracción y explotación por sus bajos recursos.

En base a las características propias de la empresa el uso de la metodología de Ralph Kimball resulta una solución eficaz en tiempo y recursos debido a que abarca la solución al problema en un corto plazo.

---

<sup>9</sup> Guillén, F., Desarrollo de un Data Mart para..., Universidad Privada del Norte, Perú, 2012, pp. viii-119

**D) Autor:** Miriam Elizabeth Chirán Enríquez

**Título:** Modelo para la implementación inteligencia de negocios que apoyen a la toma de decisiones en instituciones públicas de protección social<sup>10</sup>.

**Resumen:**

El presente trabajo constituye una referencia para la implementación de un modelo de Inteligencia de Negocios que apoye en la toma de decisiones en instituciones públicas de protección social y así mejorar su efectividad al construir una solución de Inteligencia de Negocios. Se presenta por un lado la investigación sobre los procesos e información del MIES a través de un diagnóstico del sistema y por otro lado el análisis, metodología, diseño e implementación del caso de aplicación seleccionado: “Modelo para la Implementación de Inteligencia de Negocios que apoyen a la toma de decisiones en Instituciones Públicas de Protección Social”. Logrando con esto el análisis de los impactos que causa la implementación del modelo a través de un sistema piloto de inteligencia de negocios, desarrollado para la Red de Protección Solidaria.

Considerando el tamaño y la complejidad de los datos actuales que los responsables de las decisiones tienen que manejar, no es sorprendente que el sentido común en los procesos directivos y de gestión a menudo resulten inadecuados o contraproducentes para lograr los resultados esperados, teniendo en cuenta que pueden influenciar en la vida de las personas, motivo por el cual las instituciones públicas enfocadas a la protección y el desarrollo social deben tener disponible la información necesaria para poder tomar decisiones utilizando herramientas tecnológicas que faciliten este proceso, con el fin de cumplir los objetivos de estas entidades. Este es un factor que de una u otra manera afecta a la sociedad. Por esta razón y tomando en cuenta la gran necesidad de la toma de decisiones oportunas en el momento adecuado, puede ahorrar costos, tiempo, logística entre otros. Este trabajo está orientado a investigar las metodologías y los procesos que involucran a los diferentes institutos y programas que conforman el MIES, tomando atención especial al PPS “Programa de Protección Social” y al departamento de la Red de Protección Solidaria, con el fin de conocer todos los datos que influyen directamente al momento de tomar una decisión, de manera especial los requerimientos en base a indicadores que se deben tomar en cuenta. Todo esto, a través de la implementación de Inteligencia de Negocios.

---

<sup>10</sup> Chirán, M., Modelo para la Implementación..., Universidad Central del Ecuador, Ecuador, 2013, pp. 1-2

**E) Autor:** Gustavo David Vega Vivas

**Título:** Inteligencia de negocios. Aplicación en la administración del presupuesto en una empresa del sector público<sup>11</sup>.

**Resumen:**

El presente trabajo consistió en el diseño e implementación de una herramienta de software usando tecnología de inteligencia de negocios, que cuenta con las características de confiabilidad y oportunidad, así como las últimas tecnologías de punta en materia de informática, para el apoyo a la toma de decisiones en los niveles estratégicos de la Secretaría de Comunicaciones y Transportes.

Para ello, se realizó primeramente una identificación de la información y las fuentes de datos que utilizan las áreas involucradas para la elaboración de informes. Después se diseñó una base de datos multidimensional para consolidar esa información y se construyeron las interfaces de usuario necesarias, que sirven para mostrar los parámetros de control que los usuarios requieran evaluar.

Para empezar se describe el significado, características y beneficios de los Sistemas de Información Ejecutivos, así como los conceptos de almacén de datos y de mercado de datos. También se habla sobre modelos de datos multidimensionales, cubos de información y minería de datos.

Posteriormente se describen las actividades que se desarrollan en las áreas involucradas y la razón de querer implementar un Sistema de Información Ejecutivo en la organización.

Finalmente, se presenta la propuesta de solución y se hace una medición del impacto que tuvo su implementación, la cual resultó ser muy favorable en términos de confiabilidad y oportunidad a nivel estratégico y operacional.

Se puede concluir que en una empresa de características tan grandes y variables como lo es una empresa pública, si se tiene al personal capacitado y con la madurez suficiente para tomar decisiones en base al análisis de la información disponible, entonces la inversión en tecnología será aprovechada en su totalidad

---

<sup>11</sup> Vega, G., Inteligencia de Negocios. Aplicación en..., Instituto Policlínico Nacional, México, 2010, pp. iv-154

**F) Autor:** Álvaro Villanueva Ojeda

**Título:** Análisis, diseño e implementación de un Data Warehouse de soporte de decisiones para un hospital del sistema de salud público<sup>12</sup>.

**Resumen:**

Las entidades de salud del sector público deben de tomar decisiones orientadas a satisfacer la demanda de servicios de los pacientes que acuden a los centros de salud y es por ello muy importante buscar mejorar los sistemas de información ligados a estos procesos de decisión. El presente tema de tesis propone la construcción de un Data Warehouse que servirá de apoyo en el proceso de toma de decisiones del directorio del hospital, el cual, decidirá en base a datos históricos y cuadros generados en línea.

Un sistema de este tipo permitirá reducir carga de pabellones, optimizar el uso del personal, mejorar la atención al paciente, mejorar la calidad de servicio otorgada, brindar un servicio especializado a los pacientes, gestionar recursos, conocer el estado actual de los pacientes, identificar fallas en los procesos, realizar auditorías y realizar notificaciones en tiempo real, entre otras cosas.

Los datos almacenados por la entidad no tienen utilidad si es que no se transforman en información que sirva como base para tomar decisiones. Es por ello que es necesario que todos los datos históricos sean sometidos a un proceso de limpieza para poder garantizar su confiabilidad. Este sistema se encargará de hacer una limpieza de los datos almacenados para poder generar con ellos reportes que ayuden al directorio a la toma de decisiones.

Para la realización del actual tema de tesis, se está optando por utilizar la suite de Inteligencia de Negocios proporcionada por Pentaho, la cual es una herramienta libre y completa. Con el uso de esta herramienta se garantiza que la entidad de salud pública no tendrá que destinar costos adicionales por licencias de software. Sin embargo, la dificultad en implementar con esta herramienta viene dada por su misma naturaleza libre (open source) y su poco tiempo en producción. Por esta razón, el presente proyecto dará pautas para la utilización e instalación de esta suite, lo cual servirá de base para proyectos similares que deseen implementar proyectos con ella.

Para implementar este proyecto de tesis se realizarán todos los pasos de un proyecto de Inteligencia de Negocios: diseño y construcción del Data Warehouse y los Data Marts, creación y programación de los procesos ETL, creación de los cubos, creación de los informes, y finalmente implementación de la plataforma BI (Web).

---

<sup>12</sup> Villanueva, A., Análisis, Diseño e Implementación..., Pontificia Universidad Católica..., Perú, 2008, pp. 2-3

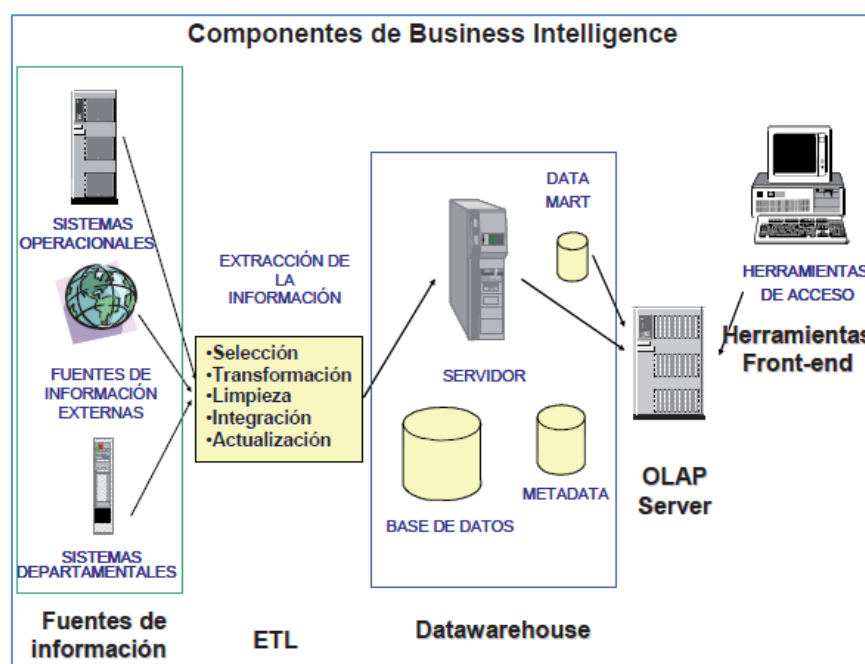
## 2.2. MARCO TEÓRICO

### 2.2.1. Business Intelligence

El término Business Intelligence (que traducido al español sería Inteligencia de Negocios) es definido como un conjunto de estrategias y tecnologías que ayudan a convertir los datos en información de calidad, y dicha información en conocimiento que nos permita una toma de decisiones más acertada y nos ayude así a mejorar nuestra competitividad<sup>13</sup>.

La definición es apoyada por Cano, quien nos dice que “el objetivo básico de la Business Intelligence es apoyar de forma sostenible y continuada a las organizaciones para mejorar su competitividad, facilitando la información necesaria para la toma de decisiones [...]. Mediante el uso de tecnologías y metodologías de Business Intelligence pretendemos convertir datos en información y a partir de la información ser capaces de descubrir conocimiento”<sup>14</sup>.

El Business Intelligence cuenta con componentes, los cuales son: (también pueden ser apreciados gráficamente en la Figura 05)



**Figura 05.** Componentes de Business Intelligence

Fuente: Cano, J.

- Fuentes de información, de las cuales partiremos para alimentar de información el Data Warehouse.
- Proceso ETL de extracción, transformación y carga de los datos en el Data Warehouse. Antes de almacenar los datos en un Data Warehouse, éstos deben

<sup>13</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, p. 9

<sup>14</sup> Cano, J., Business Intelligence: Competir con información, Ed. Banesto, España, 2007, pp. 22 - 23

ser transformados, limpiados, filtrados y redefinidos. Normalmente, la información que tenemos en los sistemas transaccionales no está preparada para la toma de decisiones.

- El propio Data Warehouse o almacén de datos, con la metadata o diccionario de datos. Se busca almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración.
- El motor OLAP, que nos debe proveer capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de escenarios en grandes volúmenes de datos.
- Las herramientas de visualización, que nos permitirán el análisis y la navegación a través de los mismos<sup>15</sup>.

### **Data Mart**

Según el propio Kimball, un Data Mart “es un subconjunto lógico de un Data Warehouse [...]. Un Data Warehouse es formado por la unión de todos los Data Marts [...]. Muchas veces se ve al Data Mart como la restricción del Data Warehouse a un solo proceso de negocio”<sup>16</sup>.

La diferencia de un Data Mart con respecto a un Data Warehouse es solamente en cuanto al alcance. Mientras que un Data Warehouse es un sistema centralizado con datos globales de la empresa y de todos sus procesos operacionales, un Data Mart es un subconjunto temático de datos, orientado a un proceso o un área de negocio específica. Debe tener una estructura óptima desde todas las perspectivas que afecten a los procesos de dicha área. Es más, según Ralph Kimball, cada Data Mart debe estar orientado a un proceso determinado dentro de la organización, por ejemplo, a pedidos de clientes, a compras, a inventario de almacén, a envío de materiales, etc. Para Ralph Kimball el conjunto de Data Marts forma el Data Warehouse<sup>17</sup>.

El trabajo de construir un Data Warehouse corporativo puede generar inflexibilidades, o ser costoso y requerir plazos de tiempo que las organizaciones no están dispuestos a aceptar. En parte, estas razones originaron la aparición de los Data Mart. Los Data Mart están dirigidos a una comunidad de usuarios dentro de la organización, que puede estar formada por los miembros de un departamento, o por los usuarios de un determinado nivel organizativo, o por un grupo de trabajo multidisciplinar con objetivos comunes.

Los Data Mart almacenan información de un número limitado de áreas; por ejemplo, pueden ser de marketing y ventas o de producción. Normalmente se definen para responder a usos muy concretos.

---

<sup>15</sup> Cano, J., Business Intelligence: Competir con información, Ed. Banesto, España, 2007, pp. 93 - 94

<sup>16</sup> Kimball, R., et Al., The Data Warehouse Lifecycle..., Ed. Wiley, EE.UU., 2008, Traducción propia, p. 1.4

<sup>17</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, p. 12

Normalmente, los Data Mart son más pequeños que los Data Warehouses. Tienen menos cantidad de información, menos modelos de negocio y son utilizados por un número inferior de usuarios<sup>18</sup>.

### **Data Warehouse**

Un Data Warehouse es una base de datos corporativa en la que se integra información depurada de las diversas fuentes que hay en la organización. Dicha información debe ser homogénea y fiable, se almacena de forma que permita su análisis desde muy diversas perspectivas, y que a su vez dé unos tiempos de respuesta óptimos. Para ello la información se encuentra altamente desnormalizada y modelada de una forma bastante diferente a los sistemas transaccionales. Principalmente se utilizan los modelos en estrella (star schema) y en copo de nieve (snowflake schema) [...] Adicionalmente estos almacenes contienen metadatos (datos sobre los datos), que aportan un valor adicional, permitiendo tener información sobre su procedencia (sobre todo cuando tenemos múltiples fuentes), la periodicidad con la que han sido introducidos, la fiabilidad que nos ofrecen, etc. Todo ello nos aporta una ayuda adicional, tanto al usuario final como a los técnicos responsables de su mantenimiento, ayudando a estos últimos en aspectos como su auditoría y administración<sup>19</sup>.

Cuando queremos analizar un problema empresarial, normalmente la información que necesitamos proviene de distintos sistemas, pero nosotros la requerimos en un mismo entorno para facilitar su análisis. Normalmente, en los sistemas transaccionales no la tenemos preparada para ser analizada: sólo tenemos la información de las transacciones actuales, pero no la de los periodos anteriores o la de las previsiones. Si queremos estudiar la evolución de las ventas necesitamos saber:

- Ventas actuales.
- Ventas de los periodos anteriores.
- Presupuesto de ventas del ejercicio.

Si seguimos con la tendencia actual, incluso quizá sea preciso las ventas estimadas hasta el final del período.

Inicialmente, podemos sentir la tentación de recuperar esa información, introducirla en una hoja de cálculo y a partir de ahí comenzar el análisis. ¿Es este el camino correcto? Obviamente no, pero vamos a justificar las razones por las que no es la mejor opción:

- Al introducir la información proveniente de distintos sistemas podemos cometer errores.

---

<sup>18</sup> Cano, J., Business Intelligence: Competir con información, Ed. Banesto, España, 2007, p. 117

<sup>19</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, pp. 11 - 12



- Debemos invertir una cantidad de tiempo considerable en la introducción de la información.
- Cada vez que queramos hacer el análisis de ventas deberemos repetir el proceso.
- Si alguien necesita más detalle de las ventas, probablemente no dispondremos del mismo y quizás respondamos que nos llevaría mucho tiempo introducir toda la información a nivel de detalle.
- Cuando hagamos las consultas para extraer la información del entorno transaccional, penalizaremos el rendimiento de las aplicaciones.
- Cuando se produzcan modificaciones en el sistema transaccional, deberemos actualizar nuestra hoja de cálculo.
- Cada uno de los usuarios de información querrá los informes con un diseño determinado: Pueden aparecer “entornos paralelos” de información.
- Si no hay un acuerdo común, es posible que tengamos distintas versiones de una misma realidad: La cifra de ventas de uno de los informes igual no coincide con la de otro.
- Los usuarios finales no siempre saben dónde reside la información que necesitan.

Podríamos aducir muchas más razones, pero será mejor que analicemos las causas que las provocan y veamos cómo solucionarlas.

La aparición de los Data Warehouse o Almacenes de datos son la respuesta a las necesidades de los usuarios que necesitan información consistente, integrada, histórica y preparada para ser analizada para poder tomar decisiones.

Al recuperar la información de los distintos sistemas, tanto transaccionales como departamentales o externos, y almacenándolos en un entorno integrado de información diseñado por los usuarios, el Data Warehouse nos permitirá analizar la información contextualmente y relacionada dentro de la organización<sup>20</sup>.

### **DataWarehousing**

Según Wolff<sup>21</sup>, es el proceso de convertir los datos operacionales de una organización en una herramienta competitiva, con el objetivo de satisfacer los requerimientos de información interna de la empresa para una mejor gestión.

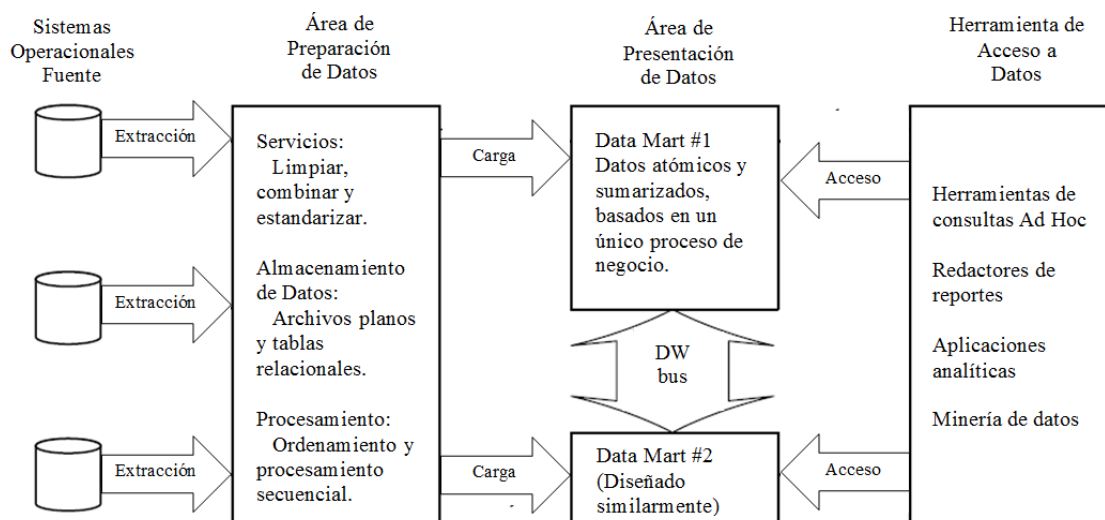
Podemos entender entonces al DataWarehousing como el proceso que produce un almacén de datos o Data Warehouse. Ahora bien, recordando que un Data Mart es un como un Data Warehouse pero que abarca solamente un proceso de negocio en específico, podemos decir que para la creación de un Data Mart también se aplica el proceso de DataWarehousing, pero orientado a un determinado proceso del negocio.

---

<sup>20</sup> Cano, J., Business Intelligence: Competir con información, Ed. Banesto, España, 2007, pp. 111 - 113

<sup>21</sup> Wolff, C., La tecnología datawarehousing, Revista Ingeniería Informática, 1999, Edición 3, p. 1

Kimball<sup>22</sup> considera que el proceso consta de cuatro componentes: sistemas operacionales fuente, área de preparación de datos, área de presentación de datos y herramientas de acceso a datos.



**Figura 06.** Componentes del proceso de DataWarehousing

Fuente: Kimball, R.

- **Sistemas Operacionales Fuente:** Sistemas de registro que capturan las transacciones del negocio. Los sistemas fuente deben ser considerados como parte fuera del Data Warehouse ya que se tiene poco o ningún control sobre su contenido y el formato de los datos que contienen.
- **Área de Preparación de Datos:** Es a la vez un área de almacenamiento y un conjunto de procesos comúnmente conocido como extracción, transformación y carga (ETL). El área de preparación de datos es todo lo que existe entre los sistemas operacionales fuente y el área de presentación de datos. Es algo análogo a la cocina de un restaurante, donde los productos alimenticios crudos son transformados en buena comida. En el data warehouse, datos operacionales crudos son transformados en un almacén entregable para las consultas de usuario.
- **Área de Presentación de Datos:** Es donde los datos son organizados, almacenados y puestos a disposición para consultas directas por los usuarios, redactores de informes y otras aplicaciones analíticas. Es todo lo que la comunidad del negocio ve y toca por medio de herramientas de acceso a datos.
- **Herramientas de Acceso a Datos:** Es el componente principal final del entorno. Usamos el término herramientas para referirnos a la variedad de capacidades que pueden ser proporcionadas a los usuarios del negocio para aprovechar el área de presentación para la toma de decisiones. Por definición, todas las herramientas de acceso a datos consultan los datos existentes en el área de presentación.

<sup>22</sup> Kimball, R., et Al., The Data Warehouse Toolkit. The..., Ed. John Wiley & Son, EE.UU., 2002, pp. 7 - 13

Con la Figura 06 podemos apreciar la relación entre los mencionados componentes.

### **OLTP (OnLine Transaction Processing)**

Los sistemas OLTP están diseñados para gestionar un gran número de peticiones concurrentes sobre sus bases de datos, y que los usuarios puedan insertar, modificar, borrar y consultar dichos datos. Están enfocados a que cada operación (transacción) trabaje con pequeñas cantidades de filas, y a que ofrezcan una respuesta rápida. Habitualmente utilizan sistemas de bases de datos relacionales para gestionar los datos, y suelen estar altamente normalizados [...] Como ejemplo de este tipo de sistemas, podemos citar un ERP, un CRM, aplicaciones departamentales, aplicaciones de ventas, de comercio electrónico, y un largo etcétera. En general cualquier aplicación con la que el usuario interactúa para introducir datos operacionales al sistema<sup>23</sup>.

### **OLAP (OnLine Analytical Processing)**

El Procesamiento analítico en línea (OLAP) tiene como objetivo agilizar la consulta de grandes volúmenes de información. Para ello utiliza estructuras multidimensionales, conocidas como cubos OLAP, que contienen datos pre calculados y agregados. Estos sistemas tienen una velocidad de respuesta muy superior a los sistemas OLTP. Un cubo OLAP es un vector multidimensional, de N dimensiones, aunque por su nombre puede hacernos creer inicialmente que sólo tiene tres dimensiones. En él la información se almacena en cada una de estas dimensiones, de forma ordenada y jerarquizada, lo que nos ayuda a realizar un rápido análisis de su contenido. Una base de datos multidimensional puede contener varios de estos cubos OLAP. Los usuarios piensan de forma multidimensional, queriendo analizar la información desde diferentes perspectivas (dimensiones), haciéndose preguntas como las siguientes:

- ¿Cuáles son las ventas actuales, comparadas con las del mismo periodo del año anterior? Quiero esta información desglosada por zona, por cliente, por vendedor, y por familia de producto.
- ¿Cuál es nuestra rentabilidad por cliente? ¿Y por producto?
- ¿Cuáles son los pedidos pendientes por cliente, por tiempo y por producto?

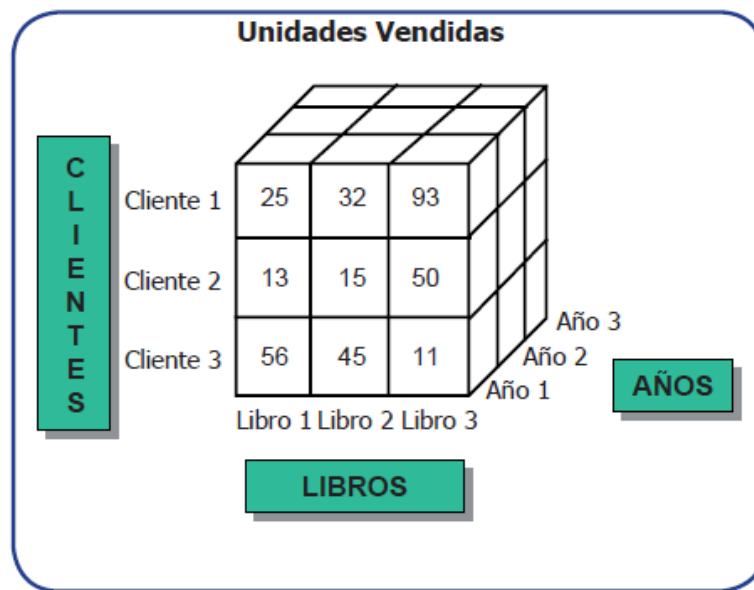
Y por, y por... Cada uno de estos “y por ...” sería una dimensión, mientras que las unidades, cantidades, importes, beneficios de... ventas, compras, pedidos... serían las medidas. Por tanto, un cubo OLAP está estructurado en dimensiones, que son las diferentes perspectivas desde las que queremos analizar la información, y en medidas, que son los diferentes hechos con valores concretos que solicita el usuario<sup>24</sup>.

---

<sup>23</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, pp. 10 - 11

<sup>24</sup> Idem, pp. 16 - 17

Según Cano<sup>25</sup> la representación gráfica del OLAP son los cubos.



**Figura 07.** Ejemplo de cubo OLAP

Fuente: Cano, J.

En la Figura 07 podemos ver que en el cubo tenemos las unidades vendidas de cada uno de los libros, para los distintos clientes y en los distintos años. Este es el concepto de multidimensionalidad. Disponemos de las unidades vendidas de cada uno de los libros para cada uno de los clientes y en cada uno de los años: el contenido de un cubo individual son las ventas de un libro a un cliente en un año. Los contenidos de cada uno de los cubos individuales del cubo recogen lo que llamamos “hechos” (en nuestro ejemplo las unidades vendidas). En la actualidad, las soluciones OLAP permiten que cada una de los cubos individuales pueda contener más de un hecho.

Las herramientas OLAP cuentan con 3 tipos, los cuales se diferencian en la forma de cómo acceden a los datos:

- **ROLAP: Relational OLAP**

Las capacidades OLAP acceden directamente a la base de datos relacional. Se accede por tanto a una base de datos relacional (RDBMS). Accede habitualmente sobre un modelo “estrella”. La principal ventaja es que no tiene limitaciones en cuanto al tamaño, pero es más lento que el MOLAP, aunque algunos productos comerciales nos permiten cargar cubos virtuales para acelerar los tiempos de acceso.

- **MOLAP: Multidimensional OLAP**

La implementación OLAP accede directamente sobre una base de datos multidimensional. La ventaja principal de esta alternativa es que es muy rápida en los tiempos de respuesta y la principal desventaja es que, si queremos cambiar las dimensiones, debemos cargar de nuevo el cubo.

<sup>25</sup> Cano, J., Business Intelligence: Competir con información, Ed. Banesto, España, 2007, pp. 127 - 130

- **HOLAP: Hybrid OLAP**  
Accede a los datos de alto nivel en una base de datos multidimensional y a los atómicos directamente sobre la base de datos relacional. En esencia utiliza las ventajas del ROLAP y del MOLAP.

### **ETL (Extract, Transform and Load)**<sup>26</sup>

Un Data Warehouse o un Data Mart, se cargan periódicamente, y en él se unifica información procedente de múltiples fuentes, creando una base de datos que cumple una serie de características descritas anteriormente. Esto implica que deben existir una serie de procesos que leen los datos de las diferentes fuentes, los transforman y adaptan al modelo que hayamos definido, los depuran y limpian, y los introducen en esta base de datos de destino. Esto es lo que se conoce como procesos ETL, procesos de Extracción, Transformación y Carga (Load).

#### Extracción de los datos

Uno de los aspectos fundamentales que debemos considerar a la hora de diseñar es si optamos por una actualización completa, que es mucho más simple, o si optamos por una actualización incremental, que es lo más conveniente. Si optamos por la primera alternativa el proceso consiste en borrar los datos y volver a cargarlos, pero si optamos por la segunda, habrá establecer una serie de controles y técnicas, entre las que destacamos las siguientes:

- Borrado sólo de parte de los datos, por ejemplo desde una fecha, y carga de los nuevos datos.
- Comparar los datos de origen y de destino, actualizando sólo los cambios.
- Uso de triggers u otras técnicas de replicación.
- Captura de datos mediante una aplicación diseñada específicamente para ello.

Lo habitual es que en nuestros procesos combinemos varias de estas técnicas, según la casuística que tengamos y diseño del proceso que hayamos realizado en cada caso.

#### Transformación de los datos

En los procesos de transformación, es preciso asegurarnos de que los datos sean válidos, de su integridad y de su utilidad, lo que suele incluir realizar cálculos y generar nuevos valores. Los datos deben ser depurados para eliminar inconsistencias, discrepancias y duplicidades.

Estas transformaciones suelen conllevar cambios con respecto a la estructura de origen para adaptarla al destino, cambios en el contenido de los valores de origen y creación de nuevos valores en las filas de destino.

---

<sup>26</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, pp. 14 - 16

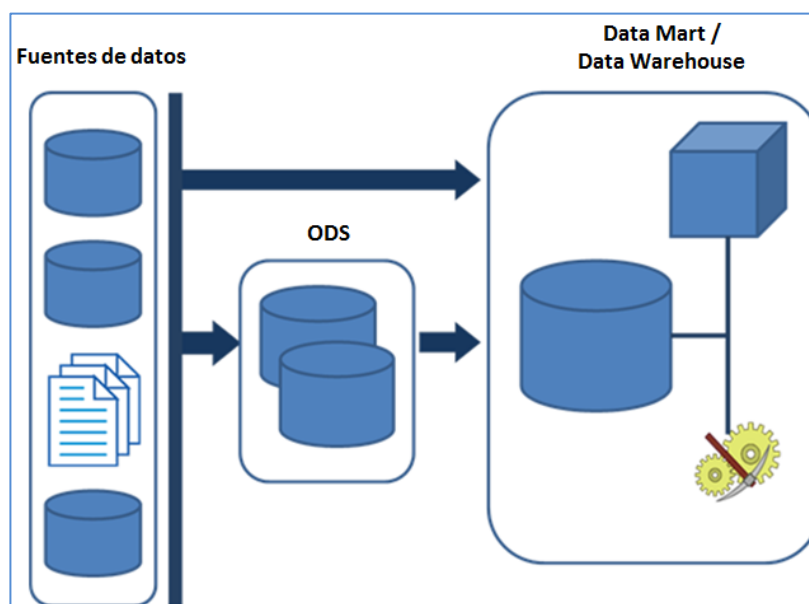
### Carga en el Data Warehouse

Como último paso debemos realizar el proceso de incorporar los datos al Data Warehouse y/o a los diferentes Data Marts, y a los cubos OLAP. Todo ello según la presentación, formatos definidos y periodicidad de actualización propuesta.

Es muy importante diseñar un buen proceso ETL, en él se deben reconciliar todos los datos de las diferentes fuentes, realizar los cálculos necesarios, mejorar la calidad de los datos, y por supuesto, adaptarlos al nuevo modelo físico y almacenarlos en él.

En muchas ocasiones la información no pasa directamente de las fuentes al Data Mart o Data Warehouse, sino que lo hace a través de unas bases de datos intermedias, que son necesarias en muchas ocasiones dada la complejidad y disparidad de las fuentes. Habitualmente, los datos, antes de entrar en el Data Mart o Data Warehouse, se almacenan en un área de staging y/o un ODS (Operational Data Store).

La Figura 08 nos ayuda a visualizar la secuencia que sigue el proceso de ETL.



**Figura 08.** Proceso ETL usando un ODS

Fuente: Ramos, S.

### ODS (Operational Data Store)

Un área de staging es un área temporal que se encuentra en el flujo de datos entre las fuentes y el Data Mart o Data Warehouse con el fin de facilitar la extracción de datos, de realizar tareas de limpieza (data cleansing), de mejorar la calidad de los datos, e incluso de ser utilizada como caché de los datos

operacionales o acceder a un nivel de detalle de los datos y de los cambios no almacenados en el Data Mart o Data Warehouse.

Un almacén operacional de datos (ODS Operational Data Store) es un área que va a dar soporte a los sistemas transaccionales, desde los que se alimenta con una periodicidad muy baja, y sirve como base de datos de consulta, a la que se conectan herramientas de reporting con el fin de que el sistema transaccional tenga una menor carga de trabajo. Se encuentra normalizado, y no es algo específico de un sistema de BI, puede existir o no, y es independiente de que exista o no un Data Mart o un Data Warehouse. Pero si existe, puede ser más apropiado obtener de él la información para alimentar el Data Mart o Data Warehouse que hacerlo desde el sistema transaccional, para así también quitar estas lecturas al sistema transaccional<sup>27</sup>.

### **Tabla de dimensiones (Dimension tables)**

Son las que almacenan la información de las dimensiones. Una dimensión contiene una serie de atributos o características, por las cuales podemos agrupar, rebanar o filtrar la información. A veces estos atributos están organizados en jerarquías que permiten analizar los datos de forma agrupada, dicha agrupación se realiza mediante relaciones uno a muchos (1:N). Por ejemplo, en una dimensión Fecha es fácil que encontremos una jerarquía formada por los atributos Año, Mes y Día, otra por Año, Semana y Día; en una dimensión Producto podemos encontrarnos una jerarquía formada por los atributos Categoría, Subcategoría y Producto. Como se ha podido comprobar en los ejemplos, se puede dar el caso de que exista más de una jerarquía para una misma dimensión<sup>28</sup>.

### **Tabla de hechos (Fact tables)**

Son tablas que representan un proceso de negocio, por ejemplo, las ventas, las compras, los pagos, los apuntes contables, los clics sobre nuestro sitio web, etc. Están formadas por los siguientes elementos:

- **Clave principal:** identifica de forma única cada fila. Al igual que en los sistemas transaccionales toda tabla debe tener una clave principal, en una tabla de hechos puede tenerla o no, y esto tiene sus pros y sus contras, pero ambas posturas son defendibles.
- **Claves externas (Foreign Keys):** apuntan hacia las claves principales (claves subrogadas) de cada una de las dimensiones que tienen relación con dicha tabla de hechos.

---

<sup>27</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, pp. 14 - 15

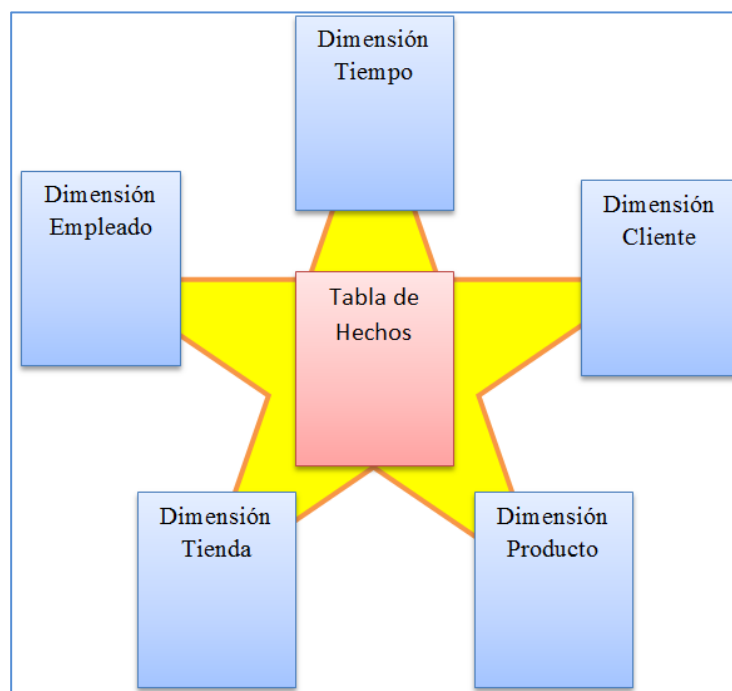
<sup>28</sup> Idem, p. 28

- Medidas (Measures): representan columnas que contienen datos cuantificables, numéricos, que se pueden agregar. Por ejemplo, cantidad, importe, precio, margen, número de operaciones, etc.
- Metadatos y linaje: nos permite obtener información adicional sobre la fila, como por ejemplo, que día se incorporó al Data Warehouse, de qué origen proviene (si tenemos varias fuentes), etc. No es necesario para el usuario de negocio, pero es interesante analizar en cada tabla de hechos qué nos aporta y si merece pena introducir algunas columnas de este tipo<sup>29</sup>.

### Esquema en estrella (Star schema)

Como nos dice Ramos<sup>30</sup>, es el esquema en el que hay una única tabla central, la tabla de hechos, que contiene todas las medidas y una tabla adicional por cada una de las perspectivas desde las que queremos analizar dicha información, es decir por cada una de las dimensiones.

En la Figura 09 podemos apreciar una representación gráfica del esquema en estrella.



**Figura 09.** Star Schema (Esquema estrella)

Fuente: Elaboración Propia

<sup>29</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, p. 32

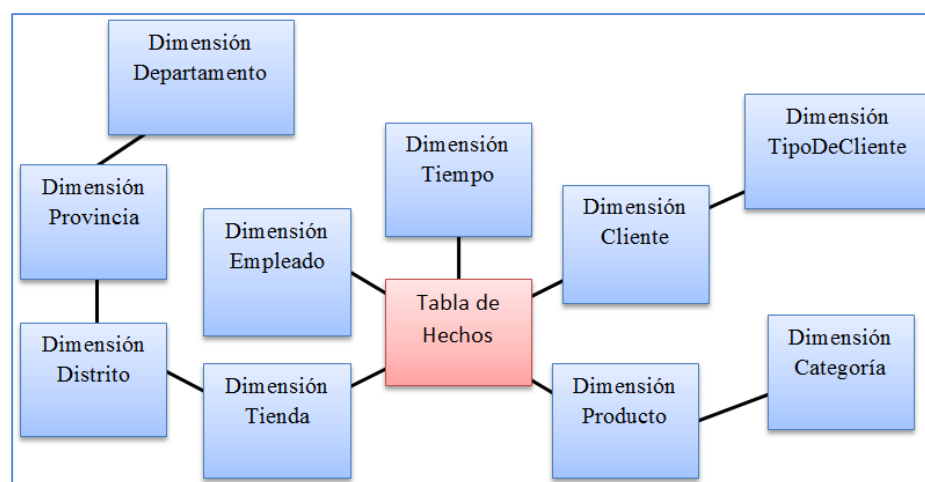
<sup>30</sup> Idem, p. 12



### Esquema en copo de nieve (Snowflake schema)

Según Ramos<sup>31</sup>, este tipo de esquema tiene una estructura más compleja que el esquema en estrella. La diferencia es que algunas de las dimensiones no están relacionadas directamente con la tabla de hechos, sino que se relacionan con ella a través de otras dimensiones. En este caso también tenemos una tabla de hechos, situada en el centro, que contiene todas las medidas y una o varias tablas adicionales, con un mayor nivel de normalización.

En la Figura 10 podemos apreciar una representación gráfica del esquema en copo de nieve.



**Figura 10.** Snowflake Schema (Esquema copo de nieve)

Fuente: Elaboración Propia

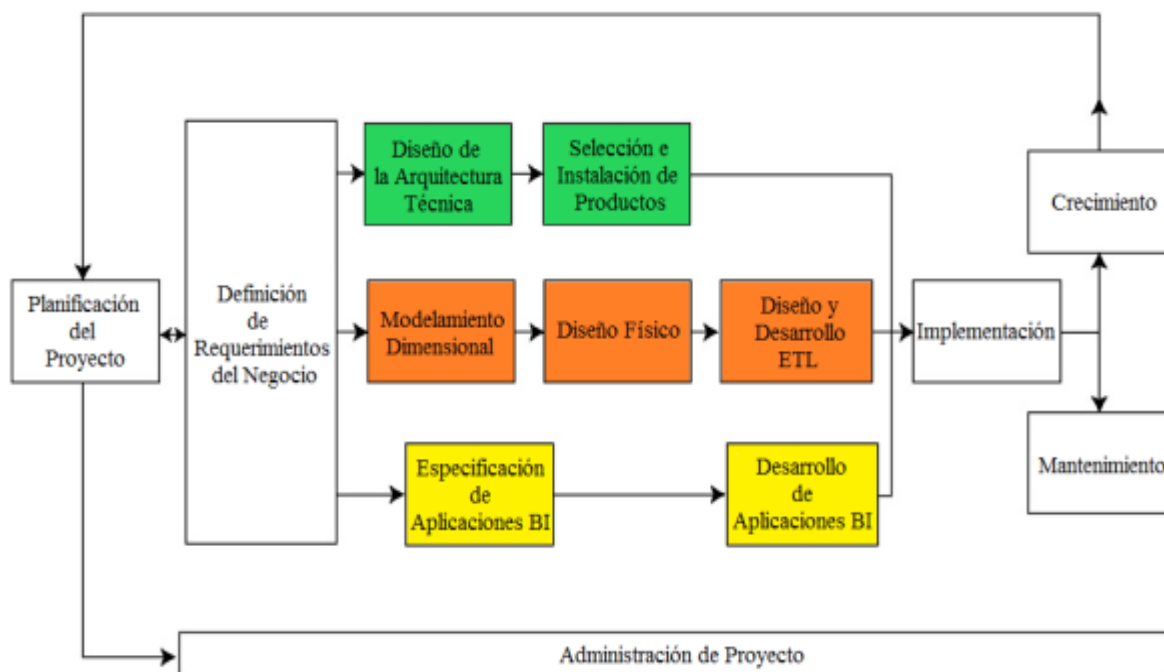
#### 2.2.2. Metodología de Ralph Kimball

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida tiene un enfoque basado en cuatro principios básicos:

- Centrarse en el negocio: Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado. Usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- Construir una estructura de información: Diseñar una base de información única, integrada, fácil de usar y de alto rendimiento que cumpla con la amplia gama de requisitos de negocio que ha identificado en la empresa.
- Realizar entregas en incrementos significativos: crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos.
- Ofrecer la solución completa: proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y

<sup>31</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, p. 13

accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación<sup>32</sup>.



**Figura 11.** Ciclo de vida de la metodología de Kimball

Fuente: Mundy, J.

Según Mundy<sup>33</sup>, la construcción de una solución de DW/BI (Data Warehouse / Business Intelligence) es una tarea compleja, es por esto que esta metodología nos ayuda a simplificar esa complejidad. Las fases de esta metodología (ciclo de vida) se muestran en la Figura 11.

De ahí, podemos tomar dos observaciones. Primero, hay que notar el rol central de la definición de requerimientos del negocio. Los requerimientos del negocio son el inicio de los tres caminos que le siguen. Y en segundo lugar podemos ver tres caminos que se concentran en tres áreas separadas:

- Camino tecnológico (Camino Superior). Cuyas actividades nos ayudan a determinar qué herramientas usaremos, cómo instalarlas y cómo configurarlas.
- Camino de datos (Camino del medio). Aquí diseñaremos el modelo dimensional, y desarrollaremos el subsistema de Extracción, Transformación y Carga (Extract, Transformation, and Load) para poblar el DW.
- Camino de aplicaciones (Camino Inferior). En esta parte es donde llevaremos a cabo las actividades de diseño y desarrollo de las aplicaciones BI para los usuarios finales.

<sup>32</sup> Mundy, J., et Al., The Microsoft Data Warehouse..., Ed. Wiley, EE. UU., 2006, Traducción Propia, p. xxxv

<sup>33</sup> Idem, pp. xxxvi - xxxvii

### 1. Planificación del Proyecto<sup>34</sup>

En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información. En la visión de programas y proyectos de Kimball, Proyecto, se refiere a una iteración simple del KLC (Kimball LifeCycle), desde el lanzamiento hasta el despliegue. Esta tarea incluye las siguientes acciones típicas de un plan de proyecto:

- Definir el alcance (entender los requerimientos del negocio).
- Identificar las tareas
- Programar las tareas
- Planificar el uso de los recursos.
- Asignar la carga de trabajo a los recursos
- Elaboración de un documento final que representa un plan del proyecto.

Además en esta parte definimos cómo realizar la administración o gestión de esta subfase que es todo un proyecto en sí mismo, con las siguientes actividades:

- Monitoreo del estado de los procesos y actividades.
- Rastreo de problemas
- Desarrollo de un plan de comunicación comprensiva que dirija la empresa y las áreas de TI.

### 2. Definición de Requerimientos del Negocio<sup>35</sup>

La definición de los requerimientos es en gran medida un proceso de entrevistar al personal de negocio y técnico, pero siempre conviene tener un poco de preparación previa. Se debe aprender tanto como se pueda sobre el negocio, los competidores, la industria y los clientes del mismo. Hay que leer todos los informes posibles de la organización; rastrear los documentos de estrategia interna; entrevistar a los empleados, analizar lo que se dice en la prensa acerca de la organización, la competencia y la industria. Se deben conocer los términos y la terminología del negocio. Parte del proceso de preparación es averiguar a quién se debe realmente entrevistar. Esto normalmente implica examinar cuidadosamente el organigrama de la organización. Hay básicamente cuatro grupos de personas con las que hablar desde el principio: el directivo responsable de tomar las decisiones estratégicas; los administradores intermedios y de negocio responsables de explorar alternativas estratégicas y aplicar decisiones; personal de sistemas, si existen, la gente que realmente sabe qué tipos de problemas informáticos y de datos existen; y por último, la gente que se necesita entrevistar por razones políticas. A partir de las entrevistas,

---

<sup>34</sup> Rivadera, G., La Metodología de Kimball para el Diseño de Almacenes de Datos (Data Warehouse)., 2010

<sup>35</sup> Idem

podemos identificar temas analíticos y procesos de negocio. Los temas analíticos agrupan requerimientos comunes en un tema común (ver Tabla 09).

**Tabla 09.** Temas analíticos

Fuente: Mundy, J.

Tema Analítico	Análisis o Requerimiento Inferido o Pedido	Proceso de Negocio de Soporte	Comentarios
Planificación de Ventas	Análisis histórico de órdenes de revendedores	Órdenes de compras	Por cliente, por país, por región de ventas
	Proyección de ventas	Órdenes de compras	La proyección es un proceso de negocio que usa las órdenes como entradas

Por otra parte, a partir del análisis se puede construir una herramienta de la metodología denominada matriz de procesos/dimensiones (Bus Matrix en inglés). Una dimensión es una forma o vista o criterio por medio de cual se pueden resumir, cruzar o cortar datos numéricos a analizar, datos que se denominan medidas (measures en inglés). Esta matriz tiene en sus filas los procesos de negocio identificados, y en las columnas, las dimensiones identificadas. Un ejemplo de esta matriz se puede observar en la Tabla 10. Cada X en la intersección de las filas y columnas significa que en el proceso de negocio de la fila seleccionada se identifican las dimensiones propuestas.

**Tabla 10.** Matriz de procesos / dimensiones

Fuente: Mundy, J.

Proceso de Negocio	Dimensiones				
	Tiempo	Producto	Empleado	Cliente (Revendedor)	Territorio de ventas
Proyección de ventas	x	x	x	x	x
Compras	x	x	x	x	x
Control de llamadas	x	x	x	x	x

### 3. Modelado Dimensional<sup>36</sup>

La creación de un modelo dimensional es un proceso dinámico y altamente iterativo. El proceso de diseño comienza con un modelo dimensional de alto

<sup>36</sup> Rivadera, G., La Metodología de Kimball para el Diseño de Almacenes de Datos (Data Warehouse)., 2010

nivel obtenido a partir de los procesos priorizados de la matriz descrita en el punto anterior. El proceso iterativo consiste en cuatro pasos:

1. Elegir el proceso de negocio.
2. Establecer el nivel de granularidad.
3. Elegir las dimensiones.
4. Identificar medidas y las tablas de hechos.

### 3.1 Elegir el proceso de negocio

El primer paso es elegir el área a modelar. Esta es una decisión de la dirección, y depende fundamentalmente del análisis de requerimientos y de los temas analíticos anotados en la etapa anterior.

### 3.2. Establecer el nivel de granularidad

La granularidad significa especificar el nivel de detalle. La elección de la granularidad depende de los requerimientos del negocio y lo que es posible a partir de los datos actuales. La sugerencia general es comenzar a diseñar el DW al mayor nivel de detalle posible, ya que se podría luego realizar agrupamientos al nivel deseado.

### 3.3. Elegir las dimensiones

Las dimensiones surgen naturalmente de las discusiones del equipo, y facilitadas por la elección del nivel de granularidad y de la matriz de procesos/dimensiones. Las tablas de dimensiones tienen un conjunto de atributos (generalmente textuales) que brindan una perspectiva o forma de análisis sobre una medida en una tabla hechos. Una forma de identificar las tablas de dimensiones es que sus atributos son posibles candidatos para ser encabezado en los informes, tablas pivot, cubos, o cualquier forma de visualización, unidimensional o multidimensional.

### 3.4. Identificar las tablas de hechos y medidas

El último paso consiste en identificar las medidas que surgen de los procesos de negocios. Una medida es un atributo (campo) de una tabla que se desea analizar, sumalizando o agrupando sus datos, usando los criterios de corte conocidos como dimensiones. Las medidas se encuentran en tablas que denominamos tablas de hechos (fact en inglés).

Cada tabla de hechos tiene como atributos una o más medidas de un proceso organizacional, de acuerdo a los requerimientos. Un registro contiene una medida expresada en números, como ser cantidad, tiempo, dinero, etc., sobre la cual se desea realizar una operación de agregación (promedio, conteo, suma,

etc.) en función de una o más dimensiones. La granularidad es el nivel de detalle que posee cada registro de una tabla de hechos.

#### 4. Diseño Físico

El diseño físico de la base de datos se centra en la definición de las estructuras físicas necesarias para soportar el diseño lógico de la base de datos. Los elementos principales de este proceso incluyen la definición de los estándares de nomenclatura y la configuración del entorno de base de datos<sup>37</sup>.

#### 5. Diseño y Desarrollo ETL

El sistema de Extracción, Transformación y Carga (ETL) es la base sobre la cual se alimenta el Data Warehouse. Si el sistema ETL se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar (grabar) la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis<sup>38</sup>.

Las herramientas de extracción, transformación y carga (ETL), son las piezas de software responsables de la extracción de datos de varias fuentes, limpiarlas, personalizarlas, integrarlas e insertarlas en el Data Warehouse. Construir el proceso ETL es uno de las más grandes tareas en la construcción del almacén de datos; es complejo, consume tiempo, esfuerzo, costos y recursos del proyecto. Construir un Data Warehouse requiere centrarse en entender tres áreas principales: el área fuente, el área de destino y el mapeo del área (proceso ETL)<sup>39</sup>.

#### 6. Diseño de la arquitectura técnica

Los entornos del almacén de datos requieren la integración de numerosas tecnologías. El diseño de la arquitectura técnica establece el marco y la visión de la arquitectura. Se necesita considerar tres factores: los requerimientos de negocio, el actual entorno técnico, y las instrucciones técnicas estratégicas planificadas, para establecer el diseño de arquitectura técnica del Data Warehouse.

#### 7. Selección e instalación de productos

Usando el diseño de arquitectura técnica, específicos componentes de estructura como plataforma de hardware, sistema de gestión de base de datos, herramientas de data staging o herramientas de acceso de datos, necesitan ser evaluados y seleccionados. Una vez que los productos han sido evaluados y seleccionados,

---

<sup>37</sup> Kimball, R., et Al., *The Data Warehouse Lifecycle...*, Ed. Wiley, EE.UU., 2008, Traducción propia, p. 2.4

<sup>38</sup> Rivadera, G., *La Metodología de Kimball para el Diseño de Almacenes de Datos (Data Warehouse)*, 2010

<sup>39</sup> Shaker, H., et Al., *A Proposed Model...*, Revista de la Universidad de King Saud, 2011, Vol. 23, p. 91

entonces son instalados y probados a fondo para asegurar la debida integración de extremo a extremo dentro del entorno del almacén de datos.

### 8. Especificación de aplicaciones BI

Las especificaciones de las aplicaciones describen la plantilla de los reportes, los parámetros controlados por el usuario, y los cálculos requeridos. Estas especificaciones aseguran que el equipo de desarrollo y los usuarios del negocio tengan un común entendimiento de las aplicaciones que serán entregadas.

### 9. Desarrollo de aplicaciones BI

Después de la especificación de las aplicaciones, el desarrollo de aplicaciones de usuario final implica la configuración de la herramienta de metadatos y la construcción de los reportes especificados. De manera óptima, estas aplicaciones son construidas utilizando una herramienta avanzada de acceso a datos que proporciona importantes aumentos en la productividad para el equipo de desarrollo de aplicaciones. Además, esto ofrece un potente mecanismo para que los usuarios de negocio puedan modificar fácilmente las plantillas de reporte existentes.

### 10. Implementación

La implementación representa la unión de la tecnología, los datos y las aplicaciones de usuario final, siendo así esta unión accesible desde el escritorio de los usuarios del negocio. Se requiere una amplia planificación para asegurar que estas piezas del rompecabezas encajen correctamente. Además, la asistencia al usuario y estrategias de comunicación o de retroalimentación deben ser establecidos antes de que cualquier tipo de usuario del negocio tenga acceso al almacén de datos.

### 11. Mantenimiento y crecimiento

Mucho trabajo queda después de la implementación inicial del almacén de datos. Es necesario a continuar centrándose en sus usuarios del negocio, proporcionándoles apoyo y capacitación. También es necesario centrar la atención en el back-room, lo que garantiza que los procesos y procedimientos están en su lugar para un funcionamiento eficaz y continuo del Data Warehouse. Métricas de aceptación y rendimiento del Data Warehouse se deben medir a través del tiempo y registrados. Por último, el plan de mantenimiento debe incluir una estrategia de comunicación de amplio alcance<sup>40</sup>.

#### **2.2.3. Metodología de Bill Inmon**

Bill Inmon nos presenta dos descripciones del camino a seguir para la construcción de un Data Warehouse. Un plan de migración y una metodología.

---

<sup>40</sup> Kimball, R., et Al., The Data Warehouse ..., Ed. Wiley, EE.UU., 2008, Traducción propia, pp. 2.4 – 2.5

La metodología que propone difiere del plan de migración en varias maneras. El plan de migración describe actividades generales dinámicamente. La metodología describe actividades específicas, los resultados de esas actividades y el orden de las actividades, pero las dinámicas iterativas de crear un Data Warehouse no son descritas. En otras palabras el plan de migración describe un plan impreciso en tres dimensiones, mientras la metodología describe un plan detallado en una dimensión. Juntos ellos forman un retrato completo de lo que se requiere para construir el Data Warehouse<sup>41</sup>.

### **Data-Driven Development Methodology**<sup>42</sup>

Una metodología dirigida por los datos tiene cuando menos dos características distintivas:

1. Se enfoca en utilizar el código y los datos que han sido elaborados previamente como base para la nueva construcción en lugar de construir alrededor de ellos. Para realizar esto, es imperativo el reconocimiento de aspectos comunes en los datos y el procesamiento, la clave para este reconocimiento es el modelo de datos.
2. Hay un énfasis en el almacén central de datos –El Data Warehouse- como la base para el procesamiento de soporte a la toma de decisiones, lo que reconoce que el procesamiento de soporte a la toma de decisiones tiene un ciclo de vida de desarrollo sumamente diferente a los sistemas operacionales.

El desarrollo de sistemas operacionales se forma alrededor de un ciclo de vida de desarrollo que inicia con los requerimientos y termina con el código, lo que se diferencia grandemente del procesamiento de soporte a la toma de decisiones, el cual inicia con los datos y termina con los requerimientos.

La metodología propuesta por Inmon expone los resultados que se deben obtener y describe el orden en que las cosas deben ser hechas. Cómo son alcanzados los resultados requeridos se deja enteramente al criterio del desarrollador.

La metodología la dividió en tres partes componentes, incluye las actividades clásicas del desarrollo de sistemas operacionales para que sea una metodología dirigida por los datos.

### **Migration Plan (Plan de migración)**

El punto inicial para el plan de migración es el modelo corporativo de datos. Este modelo representa la información que necesita la organización. Hay que tener en cuenta que representa lo que se necesita, no necesariamente lo que se tiene.

---

<sup>41</sup> Inmon, W., *Building The Data Warehouse*, Ed. Wiley, EE. UU., 2003, Traducción propia, p. 290.

<sup>42</sup> Idem, pp. 278 - 287.



El modelo de datos necesita identificar al menos lo siguiente:

- Los principales temas de la organización
- Las relaciones entre los principales temas.
- El agrupamiento de las llaves y atributos que más representen a los principales temas, incluyendo lo siguiente:
  - Atributos de los principales temas.
  - Las llaves de los principales temas.
  - Los grupos repetitivos de atributos y llaves
- Conectores entre las áreas de los principales temas.
- Relaciones de subtipos.

Luego que el modelo de datos está en su lugar, la siguiente actividad es definir el sistema de registros. Éste es definido según los sistemas existentes de la organización. El sistema de registros es la identificación de la mejor información que tiene la organización.

La mejor fuente de datos existentes es determinada por los siguientes criterios:

- ¿Qué datos existentes son los más completos?
- ¿Qué datos existentes son los más oportunos?
- ¿Qué datos existentes son los más precisos?
- ¿Qué datos existentes son los más cercanos a la fuente de entrada?
- ¿Qué datos existentes son los más cercanos a la estructura del modelo de datos?

Luego que el sistema de datos es definido, el siguiente paso es diseñar el data warehouse. Si el modelo de datos ha sido elaborado correctamente, el diseño del Data Warehouse es simple. Sólo unos pocos elementos del modelo de datos necesitan ser cambiados para convertir el modelo de datos en un un diseño de data warehouse. Principalmente necesitamos hacer lo siguiente:

- Un elemento de tiempo necesita ser agregado a la estructura clave.
- Todos los datos puramente operacionales necesitan ser eliminados.
- Las relaciones de integridad referencial necesitan ser convertidas en artefactos
- Los datos derivados frecuentemente necesitados son agregados al diseño

La estructura de datos necesita ser alterada cuando sea apropiado para:

- Agregar arreglos de datos
- Agregar datos redundantes
- Promover datos separados bajo condiciones adecuadas
- Mezcla de tablas cuando sea apropiado
- Fusionar tablas cuando sea apropiado.

Luego que el data warehouse es diseñado, el siguiente paso es diseñar y construir las interfaces entre el sistema de registros y el data warehouse. Las interfaces son las que pueblan el data warehouse.

A primera vista, las interfaces parecen ser simplemente un proceso de extracción. Pero muchas más actividades ocurren en este punto:

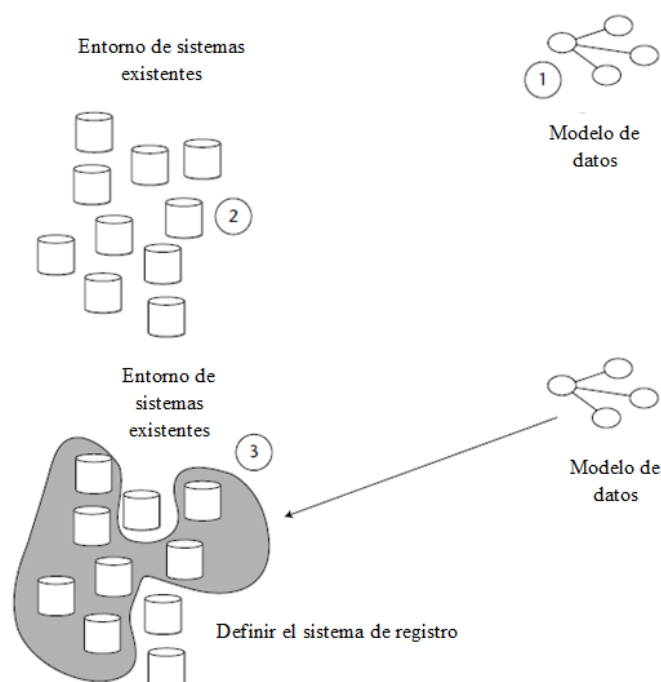
- Integración de los datos del entorno operacional.
- Alteración de las bases de tiempo de los datos
- Condensación de datos
- Exploración eficiente del entorno de los sistemas existentes.

La siguiente actividad es iniciar la población de la primera área temática. Hay muchas razones para poblar solo una fracción de los datos requeridos en el data warehouse en este punto. Poblar solo una cantidad pequeña de datos significa que los cambios pueden ser realizados fácil y rápidamente. Poblar una gran cantidad de datos disminuye la flexibilidad del data warehouse.

Una vez que el usuario final ha tenido oportunidad de ver los datos y proporcionar una retroalimentación al arquitecto de datos, entonces es seguro poblar mayores volúmenes de datos.

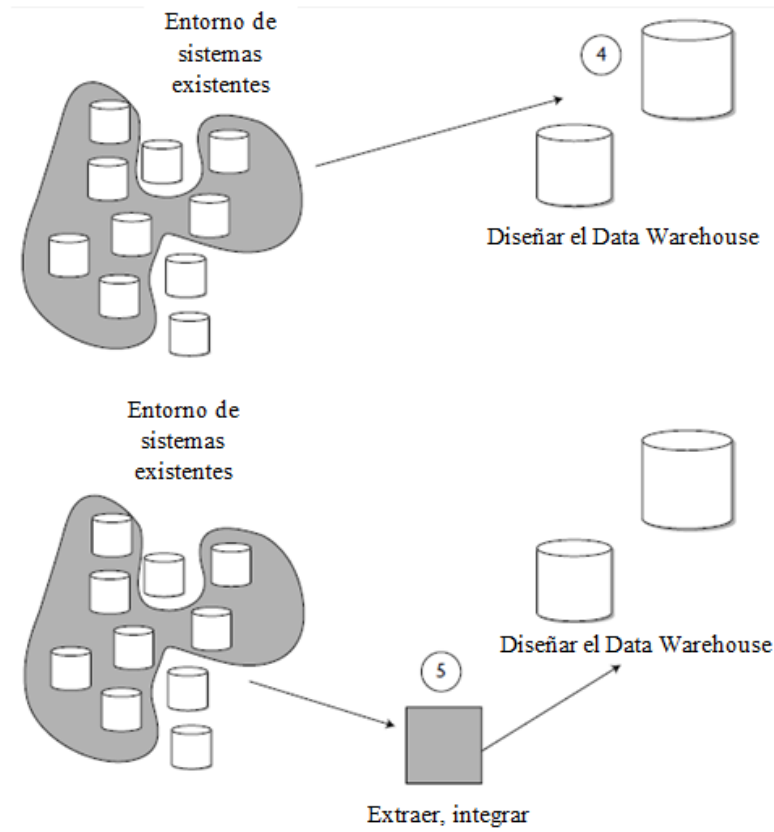
No hay que olvidar que las actividades y asuntos del entorno del data warehouse son completamente independientes de las actividades y asuntos del entorno de los sistemas operacionales existentes.

La secuencia del plan de migración la podemos apreciar gráficamente en las siguientes figuras (Figura 12, Figura 13 y Figura 14).



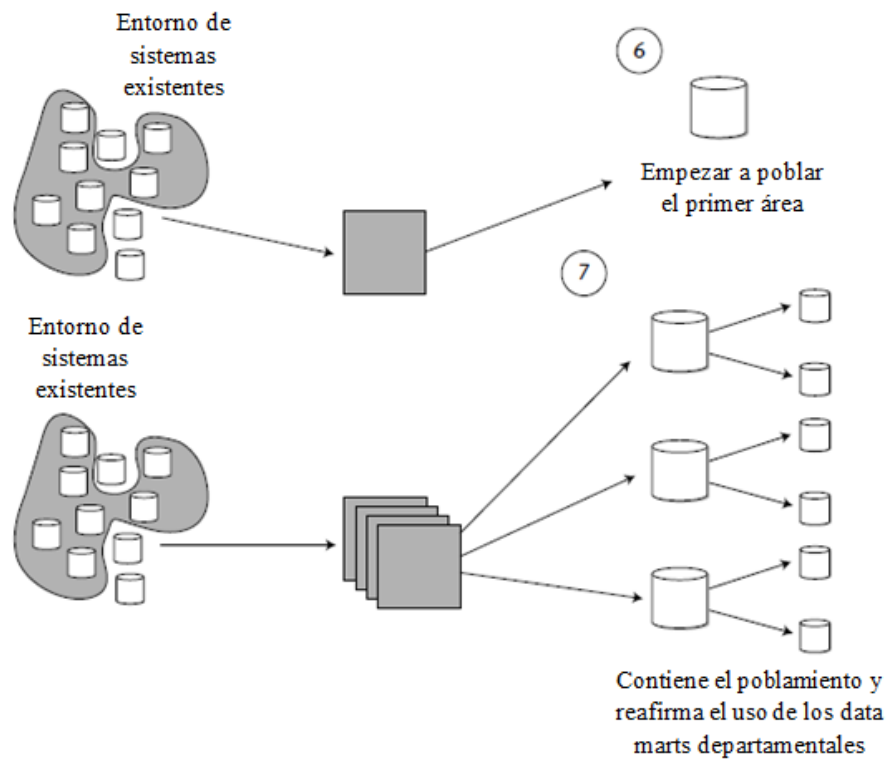
**Figura 12.** Migración al entorno diseñado – Parte I

Fuente: Inmon, W.



**Figura 13.** Migración al entorno diseñado – Parte II

Fuente: Inmon, W.



**Figura 14.** Migración al entorno diseñado – Parte III

Fuente: Inmon, W.

### El bucle de Retroalimentación

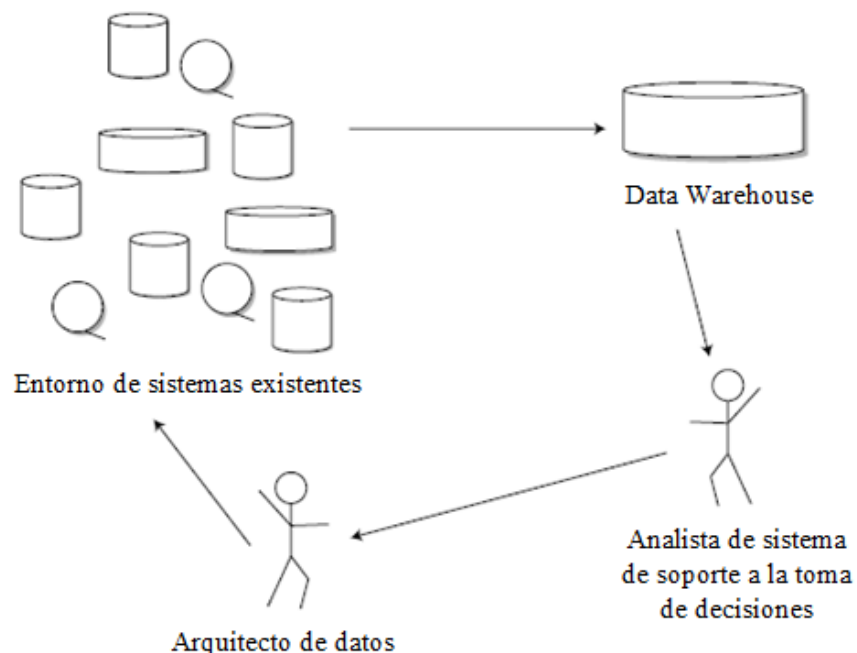
El corazón del éxito en el desarrollo del data warehouse es el bucle de retroalimentación entre el arquitecto de datos y el analista de sistemas de soporte a la toma de decisiones.

El data warehouse es poblado desde los sistemas existentes. El analista de sistemas de soporte a la toma de decisiones usa el data warehouse como base para sus análisis. Al encontrar nuevas oportunidades, el analista de sistemas de soporte a la toma de decisiones transmite esos requerimientos al arquitecto de datos, quien realiza los ajustes apropiados.

Unas pocas observaciones acerca del bucle de retroalimentación son de vital importancia para el éxito del ambiente del data warehouse:

1. El analista de sistemas de soporte a la toma de decisiones opera en un “dame lo que deseo, entonces podré decirte lo que en realidad deseo”.
2. A más corto sea el ciclo del bucle de retroalimentación, mayor es el éxito del esfuerzo del data warehouse.
3. A más grande sea el volumen de datos que tengan que ser cambiados, más largo es el bucle de retroalimentación.

La secuencia de retroalimentación puede ser apreciada en la Figura 15.

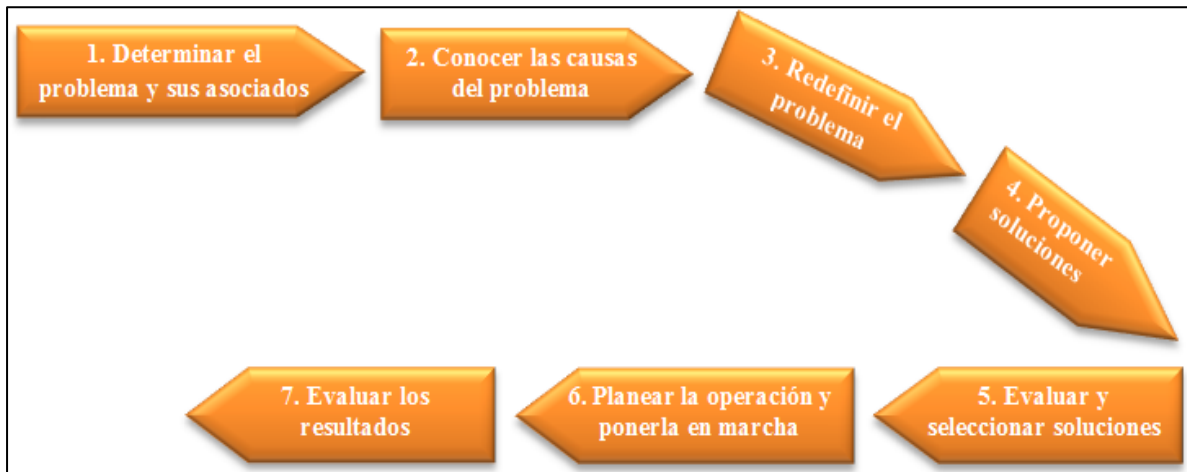


**Figura 15.** La retroalimentación entre el analista de sistema de soporte a la toma de decisiones y el arquitecto de datos

Fuente: Inmon, W.

### 2.2.4. Toma de Decisiones

A palabras simples, la toma de decisiones es el proceso de elegir entre dos o más opciones, o como nos dice Umanzor, “es el proceso de aprendizaje natural o estructurado mediante el cual se elige entre dos o más alternativas, opciones o formas para resolver diferentes situaciones o conflictos de la vida, la familia, empresa, organización”<sup>43</sup>.



**Figura 16.** Procedimiento para la toma de decisiones

Fuente: Elaboración Propia

Para Amaya<sup>44</sup> la toma de decisiones es algo fundamental para cualquier actividad humana. Todas las personas somos tomadores de decisiones. Sin embargo, tomar una correcta decisión implica un proceso de razonamiento constante. Los gerentes de las empresas deben tomar diversas decisiones cada día. Algunas de estas decisiones son rutinarias pero otras ocasionar considerables consecuencias para la empresa. Estas decisiones podrían involucrar tanto la ganancia como la pérdida de grandes sumas de dinero, o el cumplimiento o no de las metas trazadas por la empresa.

Para la toma de decisiones, Espíndola<sup>45</sup> propone un procedimiento que consta de 7 pasos, y que puede ser gráficamente en la Figura 16.

1) Determinar el problema y sus asociados: Primero que nada, debemos contar con una descripción completa del problema y las circunstancias que lo rodean. Además tenemos que tener en cuenta que los problemas por lo general no vienen solos, sino que generan otros problemas o están asociados a otros.

2) Conocimiento de las causas del problema: Frecuentemente los problemas tienen una causa principal, aunque esto no siempre se cumple. Para poder

<sup>43</sup> Umanzor, C., ¿Cómo enseñar a tomar decisiones acertadas?, 2011

<sup>44</sup> Amaya, J., Toma de Decisiones Gerenciales. Métodos..., Ed. Ecoe Ediciones, Colombia, 2010, pp. 3 - 4

<sup>45</sup> Espíndola, C., Análisis de problemas y toma de..., Ed. Pearson Educación, México, 2005, pp. 133-135

obtener información acerca de lo que está generando el problema tenemos que valernos de diferentes instrumentos:

- ✓ Entrevistas personales.
- ✓ Encuestas de opinión.
- ✓ Observación directa.

3) Redefinición del problema: Una vez que ya tenemos conocimiento de las causas del problema, debemos redefinir el problema. En un momento dado, podemos darnos cuenta que tenemos que resolver más de un problema. Hay que toma en cuenta que un problema, en vez de hacerse más específico, podría volverse más amplio y general.

4) Proponer soluciones: En esta fase se generan, mediante una lluvia de ideas, la mayor cantidad de soluciones posible. En este proceso la creatividad es un factor importante y no debe descartarse ninguna idea por más absurda que parezca. En muchas ocasiones este tipo de ideas, acompañadas de algunos cambios, pueden convertirse en excelentes opciones.

5) Evaluar y seleccionar soluciones: Una vez que contamos con las ideas o alternativas de solución, debemos proceder a la evaluación de las mejores. Para esto, es necesario analizar y prever las consecuencias tanto buenas como malas que nos generaría aplicar las soluciones aportadas. Además debemos tomar en cuenta los recursos disponibles y la trascendencia de la solución. Debemos tener presente que algunas soluciones no sólo podrían resolver el problema que buscamos solucionar sino que además podrían ayudarnos a resolver muchos otros.

6) Planeación operativa y puesta en marcha: En esta parte se debe establecer cómo se realizará la implantación de la solución:

- Las fases que deberán hacerse; es decir, la secuencia de ejecuciones que se harán en determinado tiempo.
- Establecer quiénes serán los responsables de ejecutarlas y de supervisar los trabajos.
- Los recursos humanos y materiales que serán necesarios en cada etapa.
- La calendarización de cada actividad que va a ejecutarse.

7) Evaluar los resultados obtenidos: Para la fase final existen dos tipos de evaluación: formativa y sumativa. La formativa vigila que cada fase del proceso establecido se lleve a cabo tal como se planeó. Y la evaluación sumativa, en cambio, evalúa si los resultados finales de todo el proyecto cumplen con las expectativas esperadas.

## **Tipos de Toma de Decisiones<sup>46</sup>**

Todas las decisiones no son iguales, ni producen las mismas consecuencias, ni tampoco su adopción es de idéntica relevancia, por ello existen distintos tipos de decisiones, para su clasificación destacaremos las más representativas.

### **1. Tipología por niveles**

Está conectada con el concepto de estructura organizativa y la idea de jerarquía que se deriva de la misma, las decisiones se clasifican en función de la posición jerárquica o nivel administrativo ocupado por el decisor. Desde este planteamiento distinguiremos.

- a) Decisiones estratégicas (o de planificación), son decisiones adoptadas por decisores situados en el ápice de la pirámide jerárquica o altos directivos. Estas se refieren a las relaciones entre la organización o empresa y su entorno. Son decisiones de una gran trascendencia puesto que definen los fines y objetivos generales que afectan a la totalidad de la organización; son decisiones singulares a largo plazo y no repetitivas, por lo que la información es escasa y sus efectos son difícilmente reversibles; los errores pueden comprometer el desarrollo de la empresa y en determinados casos su supervivencia, por lo que requieren un alto grado de reflexión y juicio.
- b) Decisiones tácticas o de pilotaje; son decisiones tomadas por directivos intermedios. Estas decisiones pueden ser repetitivas y el grado de repetición es suficiente para confiar en precedentes, los errores no implican sanciones muy fuertes a no ser que se vayan acumulando.
- c) Decisiones operativas, adoptadas por ejecutivos que se sitúan en el nivel más inferior. Son las relacionadas con las actividades corrientes de la empresa. El grado de repetitividad es elevado: se traducen a menudo en rutinas y procedimientos automáticos, por lo que la información es disponible. Los errores se pueden corregir rápidamente ya que el plazo al que afecta es a corto y las sanciones son mínimas.

### **2. Tipología por métodos**

Esta clasificación se basa en la similitud de los métodos empleados para la toma de decisiones, independientemente de los niveles de decisión. Así distingue una serie continua de decisiones en cuyos extremos están las decisiones programadas y no programadas.

- a) Se entiende por decisiones programadas aquellas que son repetitivas y rutinarias, cuando se ha definido un procedimiento o se ha establecido un

---

<sup>46</sup> Gutiérrez, G., Teoría de la Toma de Decisiones. Definición, Etapas y Tipos., 2014

criterio (o regla de decisión) que facilita hacerles frente, permitiendo el no ser tratadas de nuevo cada vez que se debe tomar una decisión.

- b) Las decisiones no programadas son aquellas que resultan nuevas para la empresa, no estructuradas e importantes en sí mismas. No existe ningún método establecido para manejar el problema porque este no haya surgido antes o porque su naturaleza o estructura son complejas, o porque es tan importante que merece un tratamiento hecho a la medida. También se utiliza para problemas que puedan ocurrir periódicamente pero quizá requiera de enfoques modificados debido a cambios en las condiciones internas o externas. Los directivos de alto nivel se enfrentan a decisiones no programadas, puesto que son problemas sin estructurar y a medida que se desciende en la jerarquía organizacional, mas estructurados o comprensibles resultan los problemas y por tanto, más programadas resultarán las decisiones.

### 2.2.5. Software libre

Es común identificar el software libre con una determinada tecnología, por lo que se suele sostener que el software libre es Linux y, por lo tanto, los impulsores del software libre están a favor de Linux y en contra de otros sistemas operativos. Contrariamente a esta idea, es importante aclarar que software libre no es una determinada tecnología, no es un tipo de programa de computadora ni un sistema operativo<sup>47</sup>.

Con software libre nos referimos a la libertad de los usuarios para ejecutar, copiar, distribuir, estudiar, modificar y mejorar el software. Nos referimos especialmente a cuatro clases de libertad para los usuarios de software:

- Libertad 0: la libertad para ejecutar el programa sea cual sea nuestro propósito.
- Libertad 1: la libertad para estudiar el funcionamiento del programa y adaptarlo a tus necesidades —el acceso al código fuente es condición indispensable para esto.
- Libertad 2: la libertad para redistribuir copias y ayudar así a tu vecino.
- Libertad 3: la libertad para mejorar el programa y luego publicarlo para el bien de toda la comunidad —el acceso al código fuente es condición indispensable para esto.

Software libre es cualquier programa cuyos usuarios gocen de estas libertades. De modo que deberías ser libre de redistribuir copias con o sin modificaciones, de forma gratuita o cobrando por su distribución, a cualquiera y en cualquier

---

<sup>47</sup> Da Rosa, F., Heinz, F., Guía práctica sobre software libre, Ed. Mastergraf, Uruguay, 2007, p. 29



lugar. Gozar de esta libertad significa, entre otras cosas, no tener que pedir permiso ni pagar para ello.

Asimismo, deberías ser libre para introducir modificaciones y utilizarlas de forma privada, ya sea en tu trabajo o en tu tiempo libre, sin siquiera tener que mencionar su existencia. Si decidieras publicar estos cambios, no deberías estar obligado a notificarlo de ninguna forma ni a nadie en particular<sup>48</sup>.

### **2.2.6. Pentaho**

Pentaho se define a sí mismo como una plataforma de BI “orientada a la solución” y “centrada en procesos” que incluye todos los principales componentes requeridos para implementar soluciones basados en procesos y ha sido concebido desde el principio para estar basada en procesos.

Las soluciones que Pentaho pretende ofrecer se componen fundamentalmente de una infraestructura de herramientas de análisis e informes integrado con un motor de workflow de procesos de negocio. La plataforma será capaz de ejecutar las reglas de negocio necesarias, expresadas en forma de procesos y actividades y de presentar y entregar la información adecuada en el momento adecuado.

La plataforma ha sido desarrollada bajo el lenguaje de programación Java y tiene un ambiente de implementación también basado en Java, haciendo así que Pentaho sea una solución muy flexible al cubrir una alta gama de necesidades empresariales<sup>49</sup>.

Para Bouman<sup>50</sup>, una forma de categorizar los componentes de Pentaho es por su funcionalidad. Por funcionalidad quiere decir la tarea o tareas para los que un programa fue diseñado. Desde la perspectiva del usuario la funcionalidad es lo que define el propósito del programa. Algunos componentes de Pentaho ofrecen típicas funcionalidades de BI, tales como el proceso ETL, generación de reportes y OLAP. Pero además hay funcionalidades ofrecidas por la plataforma Pentaho que no son específicas de BI, tales como autenticación y autorización de usuarios, conexiones de base de datos y ejecución de tareas programadas.

### **2.2.7. Recaudación Tributaria Municipal**

Recaudar impuestos es una función vital para cualquier Estado, aunque no siempre recaudar impuestos ha sido bien visto a lo largo del tiempo. Una buena recaudación no sólo permite que el aparato estatal funcione y sea percibido como eficiente sino que también posibilita que los recursos sean invertidos en los sectores más necesitados. A nivel municipal, esta importancia no deja de ser

---

<sup>48</sup> Stallman, R., Software libre para una sociedad libre, Ed. Traficantes de Sueños, España, 2004, pp. 59 - 60

<sup>49</sup> Enciclopedia Cubana, Plataforma Pentaho, 2011

<sup>50</sup> Bouman, R., et Al., Pentaho solutions. Business intelligence and data..., Ed. Wiley, EE. UU., 2009, p. 65.

relevante pues la municipalidad es la representación más cercana y visible que los ciudadanos tienen del Estado en su localidad<sup>51</sup>.

Los tributos cuya recaudación está a cargo de las municipalidades son: impuesto predial (impuesto a los terrenos y edificaciones), impuesto al patrimonio vehicular (impuesto a la propiedad de vehículos), impuesto de alcabala (impuesto a la transferencia de bienes inmuebles), impuesto a los juegos (impuesto a la realización de loterías, bingos y rifas), impuesto a las apuestas (impuesto a los ingresos de eventos hípicas y similares), impuesto a los espectáculos públicos no deportivos, arbitrios municipales (pagos por los servicios de recolección de residuos sólidos, barrido de calles, parques y jardines públicos, y serenazgo).

---

<sup>51</sup> Ciudadanos al Día, Eficiencia en la Recaudación Municipal: El caso de los SATs., 2004

**CAPÍTULO III:**  
**DESARROLLO DE LA SOLUCIÓN DE BUSINESS**  
**INTELLIGENCE**

*Hay tres grupos de personas: los  
que hacen que las cosas pasen; los  
que miran las cosas que pasan y los  
que se preguntan ¿Qué pasó?*

**Nicholas Murray Butler**

**1862 - 1947**

### 3.1. GENERALIDADES

La solución de Business Intelligence para mejorar el proceso de Toma de Decisiones en el Área de Rentas de la Municipalidad de Lurín consta de una serie de reportes con información útil y concisa acerca del estado de las recaudaciones y deudas de arbitrios municipales e impuesto predial de los contribuyentes del distrito. Esto con el fin de que las decisiones que tome la gerencia sean siempre basadas en información verídica.

La metodología a utilizar para el desarrollo será la metodología de Ralph Kimball, la cual nos guía paso a paso en la elaboración de la solución de Business Intelligence. Esto debido a que solamente abarcaremos un área en específico.

Con el fin de disminuir costos se hará uso de software libre, más exactamente la suite Pentaho, la cual nos proporciona todas las herramientas necesarias para el desarrollo de una solución de Business Intelligence de principio a fin.

### 3.2. ESTUDIO DE FACTIBILIDAD

#### 3.2.1. Factibilidad Técnica

El estudio de factibilidad técnica consistió en una evaluación de la tecnología con la que cuenta la institución actualmente, esto debido a que necesitamos saber si contamos con lo necesario para el desarrollo de la solución.

En cuanto al hardware, lo que se necesita es un servidor físico, dos PC de escritorio y una red local.

La institución cuenta con el hardware especificado en la Tabla 11.

**Tabla 11.** Hardware disponible

Fuente: Elaboración Propia

HARDWARE	DESCRIPCIÓN	COMENTARIOS
Servidor	<ul style="list-style-type: none"> <li>✓ Microprocesador: Intel Xeon E5320 1.86ghz (x64)</li> <li>✓ Memoria RAM: 16GB</li> <li>✓ Disco Duro: 4TB</li> <li>✓ RJ45: 1 entrada</li> <li>✓ Monitor VGA</li> <li>✓ Mouse</li> <li>✓ Teclado</li> </ul>	-
PC	<ul style="list-style-type: none"> <li>✓ Microprocesador: Intel Dual Core G620 2.60GHz (x64)</li> <li>✓ Memoria RAM: 4GB</li> <li>✓ Disco Duro: 500GB</li> <li>✓ RJ45: 1 entrada</li> </ul>	Para el usuario final.

	<ul style="list-style-type: none"> <li>✓ Monitor VGA</li> <li>✓ Mouse</li> <li>✓ Teclado</li> </ul>	
PC	<ul style="list-style-type: none"> <li>✓ Microprocesador: Intel Dual Core G620 2.60GHz (x64)</li> <li>✓ Memoria RAM: 4GB</li> <li>✓ Disco Duro: 500GB</li> <li>✓ RJ45: 1 entrada</li> <li>✓ Monitor VGA</li> <li>✓ Mouse</li> <li>✓ Teclado</li> </ul>	Para el desarrollo de la solución.
Switch	<ul style="list-style-type: none"> <li>✓ Marca: D-Link</li> <li>✓ Modelo: Des -1210-52</li> <li>✓ N° Puertos: 48</li> </ul>	Para la red local.

En cuanto al software, se necesita herramientas de ofimática, sistema operativo Windows 7, Windows server 2008, un navegador y una suite BI.

La institución cuenta con el software especificado en la Tabla 12.

**Tabla 12.** Software Disponible

Fuente: Elaboración Propia

SOFTWARE	ESTADO	COMENTARIOS
Office 2010 Standard	Licenciado	Para las PCs de la institución
Windows 7	Licenciado	Para las PCs de la institución
Windows Server 2008	Licenciado	Para el servidor de base de datos
Mozilla Firefox	Libre	Para las PCs con acceso a internet

Luego del estudio de factibilidad técnica, se determinó que la institución cuenta con todo el hardware necesario para el desarrollo de la solución BI. Respecto al software, lo único con lo que no cuenta es con la suite BI. Sin embargo, se piensa usar software libre. Esto hace que en el aspecto técnico, el proyecto sea factible.

### 3.2.2. Factibilidad Operativa

El estudio de factibilidad operativa nos permite saber si se pondrá en marcha la solución que se pretende desarrollar, beneficiando así a los usuarios involucrados.

Debido al elevado costo de tiempo y esfuerzo que demanda la obtención de reportes del estado de las recaudaciones, tanto la gerencia de rentas como la

subgerencia de informática han aceptado la solución BI, la cual mejorará el proceso. Así pues, ninguno de los usuarios presentó oposición al cambio.

Por el lado de la gerencia de rentas, el gerente se ha comprometido al uso de la solución pues lo considera necesario. Y por el lado de la subgerencia de informática, el personal se ha comprometido a administrar las herramientas y a dar soporte en caso sea necesario.

### 3.2.3. Factibilidad Económica

Desde un punto de vista financiero, toda inversión cuenta con un costo y una ganancia. Además siempre está presente la pregunta de ¿en cuánto tiempo se recuperará la inversión inicial? Sin embargo en el caso de soluciones BI, los beneficios que nos otorgan son beneficios intangibles. Incluso muchas veces el BI es considerado un responsable secundario de las mejoras pues si bien es cierto detecta los problemas y evidencia los puntos a corregir, no es precisamente quien hace las correcciones.

Podemos mencionar los beneficios intangibles que nos brindará la solución BI:

- ✓ Aumentar la velocidad del proceso, lo cual permite aprovechar ese tiempo en otras actividades.
- ✓ Facilitar el manejo de grandes cantidades de información, lo cual hará mucho más sencillo su análisis.
- ✓ Generar información confiable, que apoye a la toma de decisiones.
- ✓ Aprovechar al máximo los recursos tecnológicos con los que cuenta la institución.

Costos

A pesar de todo cabe especificar los costos que generará el desarrollo de la solución BI, los mismos que se muestran en la Tabla 13.

**Tabla 13.** Costos del Desarrollo de la Solución

Fuente: Elaboración Propia

TIPO	RECURSO	CANTIDAD	COSTO (S./)
Recurso Humano	Personal	1	5000
Software	Suite BI	1	0
Materiales de Oficina	Papel bond	500	25.0
	Lapicero	5	2.5
	Lápiz	5	2.5
	Tinta para impresora	4	50.0
Otros gastos	Movilidad	-	50.0
<b>TOTAL</b>			<b>5130.0</b>

Luego de este estudio se hace evidente que el desarrollo de la propuesta sí es económicamente factible.

### **3.3. PLANIFICACIÓN DEL PROYECTO**

#### **3.3.1. Descripción del Proyecto**

La solución propuesta para el área de rentas de la municipalidad de Lurín le permitirá a la gerencia tener conocimiento acerca del estado actual e histórico de las recaudaciones y deudas (de arbitrios municipales e impuestos prediales) de los contribuyentes del distrito por medio de reportes fáciles de entender que contendrán datos confiables. Con esto, las diversas acciones que se realicen (por ejemplo dar facilidades de pago a las zonas que deben más) serán siempre basadas en información real y en sus respectivos indicadores. Para esto se desarrollará una solución de Business Intelligence para el proceso de recaudación del área de rentas. Se efectuará un análisis dimensional, el cual dará paso a la obtención de cubos dimensionales u OLAP. Finalmente se llevará a cabo el diseño de los reportes para la gerencia.

#### **3.3.2. Objetivos del Proyecto**

Reducir el tiempo y el esfuerzo empleado en la generación de los reportes relacionados a la recaudación, debido a que actualmente esta actividad es bastante tardía y demanda mucho esfuerzo.

Brindar a la gerencia de rentas una forma de visualización del estado actual e histórico de las recaudaciones realizadas y deudas ocurridas, de manera sencilla y confiable, basado en indicadores.

#### **3.3.3. Alcance del Proyecto**

- Desarrollo de un Data Mart para el área de rentas de la municipalidad de Lurín con los datos obtenidos del sistema transaccional de recaudación.
- Elaboración del cubo OLAP mediante la base de datos dimensional.
- Diseño de reportes:
  - ✓ Montos recaudados de impuesto predial por cuotas, fechas.
  - ✓ Montos recaudados de arbitrios municipales por cuotas, fechas y zonas.
  - ✓ Montos adeudados de impuesto predial por cuotas.
  - ✓ Montos adeudados de arbitrios municipales por cuotas y zonas.
  - ✓ Contribuyentes que pagan los mayores montos.
  - ✓ Contribuyentes que adeudan los mayores montos.
  - ✓ Cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial.

- ✓ Estado de la recaudación según KPI.

### 3.3.4. Stakeholders

Los principales interesados del proyecto vienen a ser el gerente de rentas, el subgerente de informática y el autor, cuyas funciones se especifican en la Tabla 14. Por otro lado, las funciones del equipo de trabajo, que llevará a cabo la realización del producto final se encuentran en la Tabla 15.

**Tabla 14.** Principales Stakeholders del Proyecto

Fuente: Elaboración Propia

STAKEHOLDER	CARGO	FUNCIÓN
Felix Vargas Chumpitaz	Jefe de Proyecto	Es la persona que se encarga de la planificación y desarrollo del proyecto, y quien se hace responsable de que el mismo se lleve a cabo de manera óptima.
Juan Carlos Cervantes Aguirre	Subgerente de Informática	Tiene la labor de asegurar el correcto funcionamiento de los componentes de cómputo de propiedad de la Municipalidad de Lurín, así como también asesorar a los distintos usuarios.
Pedro Rueda Torrejón	Gerente de Rentas	Cumple la función de dirigir el entero proceso de recaudación de impuestos y arbitrios (de todo tipo) de la Municipalidad.

**Tabla 15.** Funciones de Equipo de Trabajo

Fuente: Elaboración Propia

NOMBRE	CARGO	FUNCIÓN
Felix Vargas Chumpitaz	Analista de Dimensiones	Analizar las dimensiones, hechos y medidas.
Felix Vargas Chumpitaz	Experto en ETL	Realizar los procesos de extracción, transformación y carga.
Felix Vargas Chumpitaz	Experto en Cubos OLAP	Diseñar y elaborar los cubos multidimensionales.
Felix Vargas Chumpitaz	Desarrollador de Aplicaciones BI	Desarrollar la aplicación web para el usuario final.



### 3.3.5. Análisis de Riesgos

Tabla 16. Matriz de Riesgo

Fuente: Elaboración Propia

TIPO	RIESGO	PROBABILIDAD	IMPACTO	MITIGACIÓN	RESPONSABLE
Tecnológico	Falla en el equipo que alojará al Data Mart	Baja	Medio	Tener disponible otra máquina virtual en caso de emergencia.	Felix Vargas Chumpitaz
	Mal funcionamiento del servidor de base de datos del sistema transaccional	Baja	Alto	Tener disponible otro equipo para uso en caso de emergencia. Realizar backups diariamente.	Felix Vargas Chumpitaz
	Presencia de virus informáticos	Baja	Medio	Actualizar diariamente el antivirus. Controlar la descarga de archivos y uso de USB.	Felix Vargas Chumpitaz
	Equipos disponibles para trabajar con bajos niveles de recursos (disco duro, memoria RAM)	Baja	Medio	Tener disponible una laptop de óptima capacidad en caso de emergencia.	Felix Vargas Chumpitaz
	Falla en los componentes de red (router, switch, cables)	Baja	Alto	Contar con routers, switches y cables de repuesto.	Felix Vargas Chumpitaz
Organizacional	Poco o nulo apoyo por parte de la institución	Baja	Medio	Realizar periódicamente visitas a la institución.	Felix Vargas Chumpitaz
	La institución no facilita la realización de las pruebas de funcionamiento	Media	Medio	Organizar con anticipación los días de las pruebas.	Felix Vargas Chumpitaz
	Cambio de opinión de la gerencia respecto a los productos finales	Media	Medio	Mostrar los avances con regularidad.	Felix Vargas Chumpitaz
De Proyecto	Tiempo de culminación del proyecto mal calculado	Media	Media	Reorganizar las actividades, aumentando esfuerzos si es necesario	Felix Vargas Chumpitaz
	Paralización del proyecto por motivos de fuerza mayor	Baja	Alto	Realizar una adecuada planificación. Documentar todos los avances.	Felix Vargas Chumpitaz

### 3.3.6. Cronograma de Actividades

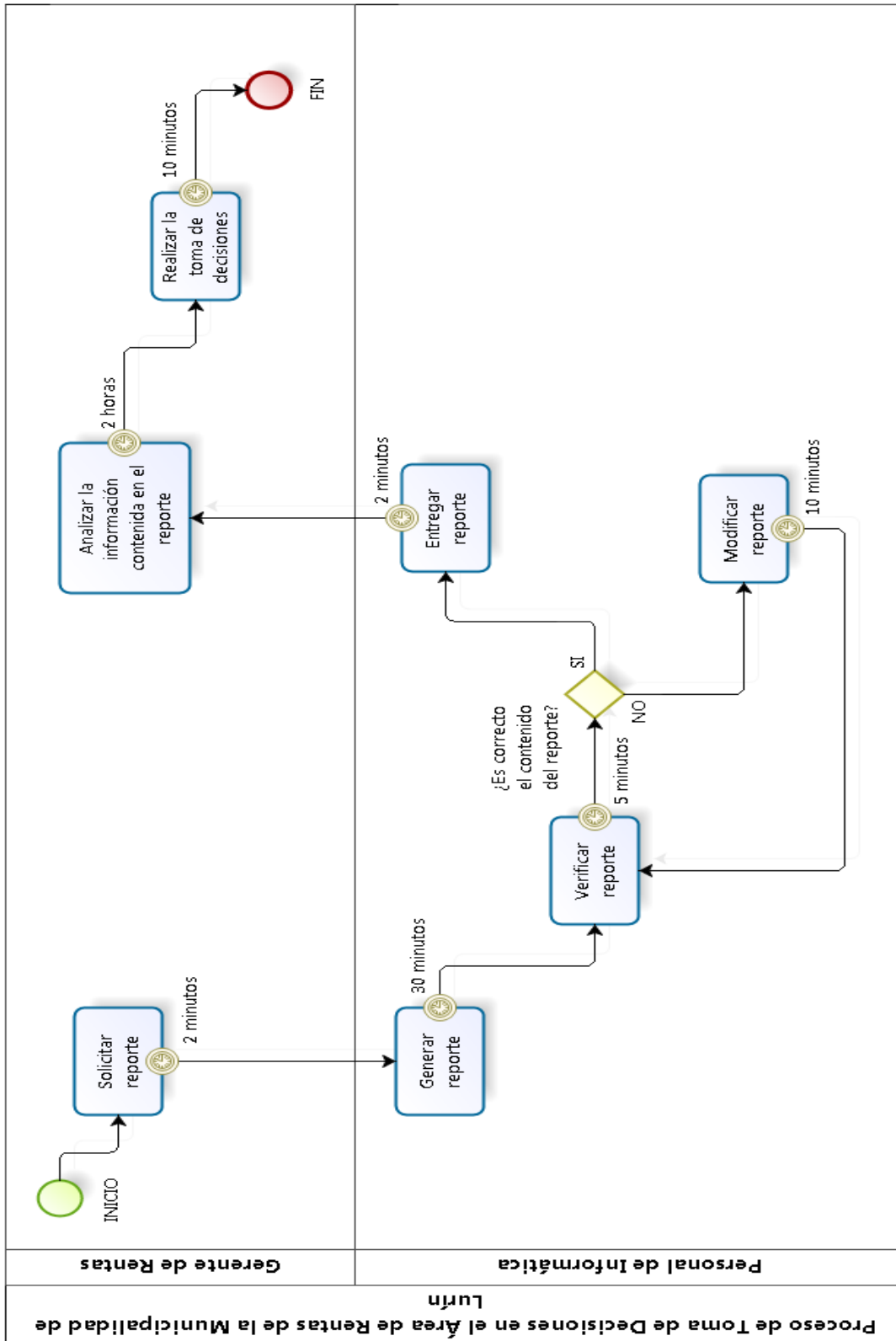
	Nombre de tarea	Duración	Comienzo	Fin	eces	Nombres de los recursos
1	<b>DESARROLLO DE UNA SOLUCIÓN DE BUSINESS INTELLIGENCE PARA MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE RENTAS DE LA MUNICIPALIDAD DE LURÍN</b>	<b>66 días</b>	<b>lun 15/09/14</b>	<b>mié 19/11/14</b>		
2	<b>Planificación del Proyecto</b>	<b>5 días</b>	<b>lun 15/09/14</b>	<b>vie 19/09/14</b>		
3	Descripción del proyecto	0,5 días	lun 15/09/14	lun 15/09/14		Felix Vargas Chumpitaz
4	Objetivos del proyecto	0,5 días	lun 15/09/14	lun 15/09/14	3	Felix Vargas Chumpitaz
5	Alcance del proyecto	0,5 días	mar 16/09/14	mar 16/09/14	4	Felix Vargas Chumpitaz
6	Stakeholder	0,5 días	mar 16/09/14	mar 16/09/14	5	Felix Vargas Chumpitaz
7	Análisis de riesgos	1 día	mié 17/09/14	mié 17/09/14	6	Felix Vargas Chumpitaz
8	Cronograma de actividades	2 días	jue 18/09/14	vie 19/09/14	7	Felix Vargas Chumpitaz
9	<b>Definición de los Requerimientos del Negocio</b>	<b>5 días</b>	<b>sáb 20/09/14</b>	<b>mié 24/09/14</b>	<b>8</b>	
10	Proceso de negocio: Toma de decisiones del área de rentas	1 día	sáb 20/09/14	sáb 20/09/14		Felix Vargas Chumpitaz
11	Proceso de negocio y temas analíticos	1 día	dom 21/09/14	dom 21/09/14	10	Felix Vargas Chumpitaz
12	Matriz procesos/dimensiones	1 día	lun 22/09/14	lun 22/09/14	11	Felix Vargas Chumpitaz
13	Requerimientos	0,5 días	mar 23/09/14	mar 23/09/14	12	Felix Vargas Chumpitaz
14	Documentación de los requerimientos	0,5 días	mar 23/09/14	mar 23/09/14	13	Felix Vargas Chumpitaz
15	Hoja de gestión	0,5 días	mié 24/09/14	mié 24/09/14	14	Felix Vargas Chumpitaz
16	Hoja de análisis	0,5 días	mié 24/09/14	mié 24/09/14	15	Felix Vargas Chumpitaz
17	<b>Modelado Dimensional</b>	<b>6 días</b>	<b>jue 25/09/14</b>	<b>mar 30/09/14</b>	<b>16</b>	
18	Dimensiones	2 días	jue 25/09/14	vie 26/09/14		Felix Vargas Chumpitaz
19	Granularidad	1 día	sáb 27/09/14	sáb 27/09/14	18	Felix Vargas Chumpitaz
20	Hechos	1 día	dom 28/09/14	dom 28/09/14	19	Felix Vargas Chumpitaz
21	Medidas	1 día	lun 29/09/14	lun 29/09/14	20	Felix Vargas Chumpitaz
22	Diseño del modelo estrella	1 día	mar 30/09/14	mar 30/09/14	21	Felix Vargas Chumpitaz
23	<b>Diseño de la Arquitectura Técnica</b>	<b>3 días</b>	<b>mié 01/10/14</b>	<b>vie 03/10/14</b>	<b>22</b>	
24	Arquitectura	1 día	mié 01/10/14	mié 01/10/14		Felix Vargas Chumpitaz
25	Back-room	1 día	jue 02/10/14	jue 02/10/14	24	Felix Vargas Chumpitaz
26	Front-room	1 día	vie 03/10/14	vie 03/10/14	25	Felix Vargas Chumpitaz
27	<b>Selección e instalación de productos</b>	<b>3 días</b>	<b>sáb 04/10/14</b>	<b>lun 06/10/14</b>	<b>26</b>	
28	Evaluación de productos	2 días	sáb 04/10/14	dom 05/10/14		Felix Vargas Chumpitaz
29	Herramientas seleccionadas	1 día	lun 06/10/14	lun 06/10/14	28	Felix Vargas Chumpitaz
30	<b>Diseño Físico</b>	<b>5 días</b>	<b>mar 07/10/14</b>	<b>sáb 11/10/14</b>	<b>29</b>	
31	Tablas de apoyo	0,5 días	mar 07/10/14	mar 07/10/14		
32	Tabla dimensión: Tiempo	0,5 días	mar 07/10/14	mar 07/10/14	31	Felix Vargas Chumpitaz
33	Tabla dimensión: Contribuyente	0,5 días	mié 08/10/14	mié 08/10/14	32	Felix Vargas Chumpitaz
34	Tabla dimensión: TipoPago	0,5 días	mié 08/10/14	mié 08/10/14	33	Felix Vargas Chumpitaz
35	Tabla dimensión: Zona	0,5 días	jue 09/10/14	jue 09/10/14	34	Felix Vargas Chumpitaz
36	Tabla dimensión: Estado	0,5 días	jue 09/10/14	jue 09/10/14	35	Felix Vargas Chumpitaz
37	Tabla dimensión: Cuota	0,5 días	vie 10/10/14	vie 10/10/14	36	Felix Vargas Chumpitaz
38	Tabla de hechos: Recaudación	0,5 días	vie 10/10/14	vie 10/10/14	37	Felix Vargas Chumpitaz
39	Diseño modelo físico	1 día	sáb 11/10/14	sáb 11/10/14	38	Felix Vargas Chumpitaz
40	<b>Diseño ETL</b>	<b>19 días</b>	<b>dom 12/10/14</b>	<b>jue 30/10/14</b>	<b>39</b>	
41	Extracción	4 días	dom 12/10/14	mié 15/10/14		Felix Vargas Chumpitaz
42	Transformación	7 días	jue 16/10/14	mié 22/10/14	41	Felix Vargas Chumpitaz
43	Carga	5 días	jue 23/10/14	lun 27/10/14	42	Felix Vargas Chumpitaz
44	Automatización del proceso	3 días	mar 28/10/14	jue 30/10/14	43	Felix Vargas Chumpitaz
45	<b>Diseño del cubo OLAP</b>	<b>4 días</b>	<b>vie 31/10/14</b>	<b>lun 03/11/14</b>	<b>44</b>	
46	Creación del cubo	4 días	vie 31/10/14	lun 03/11/14	44	Felix Vargas Chumpitaz
47	<b>Especificación de Aplicaciones BI</b>	<b>2 días</b>	<b>mar 04/11/14</b>	<b>mié 05/11/14</b>	<b>46</b>	
48	Lista de reportes	2 días	mar 04/11/14	mié 05/11/14	46	Felix Vargas Chumpitaz
49	<b>Desarrollo de aplicaciones BI</b>	<b>7 días</b>	<b>jue 06/11/14</b>	<b>mié 12/11/14</b>	<b>48</b>	
50	Elaboración de reportes	7 días	jue 06/11/14	mié 12/11/14		Felix Vargas Chumpitaz
51	<b>Implementación</b>	<b>3 días</b>	<b>jue 13/11/14</b>	<b>sáb 15/11/14</b>	<b>50</b>	
52	Verificación de tecnología	1 día	jue 13/11/14	jue 13/11/14		Felix Vargas Chumpitaz
53	Manual de usuario	2 días	vie 14/11/14	sáb 15/11/14	52	Felix Vargas Chumpitaz
54	<b>Mantenimiento y Crecimiento</b>	<b>4 días</b>	<b>dom 16/11/14</b>	<b>mié 19/11/14</b>	<b>53</b>	
55	Pruebas de funcionamiento del sistema	4 días	dom 16/11/14	mié 19/11/14	53	Felix Vargas Chumpitaz

Figura 17. Cronograma de Actividades

Fuente: Elaboración Propia

### 3.4. DEFINICIÓN DE LOS REQUERIMIENTOS DEL NEGOCIO

#### 3.4.1. Proceso de negocio: Toma de Decisiones del Área de Rentas



**Figura 18.** Proceso de Negocio: Toma de Decisiones en el Área de Rentas

Fuente: Elaboración Propia

### 3.4.2. Proceso de Negocio y Temas Analíticos

**Tabla 17.** Temas Analíticos basados en entrevistas

Fuente: Elaboración Propia

TEMAS ANALÍTICOS	ANÁLISIS SOLICITADOS O INFERIDOS	PROCESO DE NEGOCIO COMPATIBLE	COMENTARIOS
<b>Planificación de Recaudación</b>	Información histórica de la recaudación de impuesto predial según cuotas	Recaudación de impuesto predial	Por cuotas (4) con su respectivo año
	Información histórica de la recaudación de arbitrios municipales según cuotas	Recaudación de arbitrios municipales	Por cuotas (12) con su respectivo año.
	Información histórica de la recaudación de arbitrios municipales según zonas	Recaudación de arbitrios municipales	Por microzonas y zonas
	Información de las deudas de impuesto predial por cuotas	Recaudación de impuesto predial	Por cuotas (4) con su respectivo año
	Información de las deudas de arbitrios municipales por cuotas	Recaudación de arbitrios municipales	Por cuotas (12) con su respectivo año.
	Información de las deudas de arbitrios municipales por zonas	Recaudación de arbitrios municipales	Por microzonas y zonas
	Información histórica de las recaudaciones por fechas.	Recaudación de arbitrios e impuestos	Por años, trimestres y meses.
<b>Reporte de Recaudación</b>	Contribuyentes que pagan mayores/menores montos	Recaudación de arbitrios e impuestos	Por cuotas y años.
	Cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial	Recaudación de arbitrios e impuestos	Por cuotas y fechas.
	Estado de las recaudaciones según KPIs	Recaudación de arbitrios e impuestos	Por cuotas y años.

**Tabla 18.** Procesos de Negocio basados en entrevistas

Fuente: Elaboración Propia

LETRA	PROCESO DE NEGOCIO	TEMAS ANALÍTICOS COMPATIBLES
A	Recaudación de impuesto predial y arbitrios municipales	Planificación de la recaudación, Reporte de la recaudación, Análisis de la recaudación

### 3.4.3. Matriz Procesos/Dimensiones

**Tabla 19.** Matriz Procesos/Dimensiones

Fuente: Elaboración Propia

PROCESO DE NEGOCIO	DIMENSIONES					
	Contribuyente	Cuota	Tipo	Zona	Tiempo	Estado
Recaudación de impuesto predial y arbitrios municipales	X	X	X	X	X	X

### 3.4.4. Requerimientos

**Tabla 20.** Lista de Requerimientos

Fuente: Elaboración Propia

CÓDIGO	REQUERIMIENTO
REQ01	Mostrar los montos de impuesto predial recaudados por cuotas y fechas.
REQ02	Mostrar los montos de arbitrios municipales recaudados por cuotas y fechas.
REQ03	Mostrar los montos de arbitrios municipales recaudados por zonas y microzonas.
REQ04	Mostrar los montos de impuesto predial adeudados por cuotas.
REQ05	Mostrar los montos de arbitrios municipales adeudados por cuotas.
REQ06	Mostrar los montos de arbitrios municipales adeudados por zonas y microzonas.

<b>REQ07</b>	Mostrar los contribuyentes que pagan las mayores cantidades de dinero por año fiscal.
<b>REQ08</b>	Mostrar los contribuyentes que adeudan las mayores cantidades de dinero por año fiscal.
<b>REQ09</b>	Mostrar la cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial por cuotas y fechas.
<b>REQ10</b>	Mostrar el estado de la recaudación de los arbitrios municipales e impuesto predial según los KPIs.

### 3.4.5. Documentación de los Requerimientos

- **REQ01:** Mostrar los montos de impuesto predial recaudados por cuotas y fechas.

Se desea obtener la información real actual e histórica de las recaudaciones de impuesto predial realizadas, mediante el filtrado por cuotas (ejemplo: “cuánto dinero se recaudó por la cuota 1 del impuesto predial del año 2010”) y fechas (ejemplo: “cuánto dinero se recaudó por impuesto predial el mes de marzo del año 2014”). De igual manera se podrá ver las cantidades recaudadas por la suma de todas las cuotas de un año y todos los meses y trimestres de un año (ejemplo: “cuánto dinero se recaudó por todas las cuotas del impuesto predial del año 2012”).

Responsable: Felix Vargas Chumpitaz.

- **REQ02:** Mostrar los montos de arbitrios municipales recaudados por cuotas y fechas.

Se desea obtener la información real actual e histórica de las recaudaciones de impuesto predial realizadas, mediante el filtrado por cuotas (ejemplo: “cuánto dinero se recaudó por la cuota 5 del arbitrio municipal del año 2013”) y fechas (ejemplo: “cuánto dinero se recaudó por impuesto predial el mes de marzo del año 2014”). De igual manera se podrá ver las cantidades recaudadas por la suma de todas las cuotas de un año y todos los meses y trimestres de un año (ejemplo: “cuánto dinero se recaudó por todas las cuotas de los arbitrios municipales del año 2010”).

Responsable: Felix Vargas Chumpitaz.

- **REQ03:** Mostrar los montos de arbitrios municipales recaudados por zonas y microzonas.

Se desea obtener la información real actual e histórica de las recaudaciones de arbitrios municipales realizadas, mediante el filtrado por zonas (ejemplo: “cuánto dinero se recaudó por los arbitrios municipales del 2011 en la zona A”) y por microzonas (ejemplo: “cuánto dinero se recaudó por los arbitrios municipales del 2009 en la microzona Lurín Cercado”).

Responsable: Felix Vargas Chumpitaz.

- **REQ04:** Mostrar los montos de impuesto predial adeudados por cuotas.

Se desea obtener la información real actual e histórica de las deudas de impuesto predial, mediante el filtrado por cuotas (ejemplo: “cuánto dinero se adeuda por la cuota 4 del impuesto predial del año 2013”). De igual manera se podrá ver los montos adeudados por la suma de todas las cuotas de un año (ejemplo: “cuánto dinero se adeuda por todas las cuotas del impuesto predial del año 2011”).

Responsable: Felix Vargas Chumpitaz.

- **REQ05:** Mostrar los montos de arbitrios municipales adeudados por cuotas.

Se desea obtener la información real actual e histórica de las deudas de arbitrios municipales, mediante el filtrado por cuotas (ejemplo: “cuánto dinero se adeuda por la cuota 10 del arbitrio municipal del año 2009”). De igual manera se podrá ver los montos adeudados por la suma de todas las cuotas de un año (ejemplo: “cuánto dinero se adeuda por todas las cuotas del año 2012”).

Responsable: Felix Vargas Chumpitaz.

- **REQ06:** Mostrar los montos de arbitrios municipales adeudados por zonas y microzonas.

Se desea obtener la información real actual e histórica de las deudas de arbitrios municipales, mediante el filtrado por zonas (ejemplo: “cuánto dinero se adeuda por los arbitrios municipales del 2010 en la zona E”) y por microzonas (ejemplo: “cuánto dinero se adeuda por los arbitrios municipales del 2009 en la microzona Huertos de Lurín”).

Responsable: Felix Vargas Chumpitaz.

- **REQ07:** Mostrar los contribuyentes que pagan las mayores cantidades de dinero por año fiscal.

Se pretende saber cuáles son los contribuyentes que ocupan los primeros puestos en lo que a pagar arbitrios municipales e impuesto predial se refiere. Es decir los contribuyentes que pagan las mayores cantidades de dinero. Aquí se emplea el filtro por años (ejemplo: “cuáles fueron los contribuyentes que más pagaron por las cuotas del año 2012”).

Responsable: Felix Vargas Chumpitaz.

- **REQ08:** Mostrar los contribuyentes que adeudan las mayores cantidades de dinero por año fiscal.

Se pretende saber cuáles son los contribuyentes que ocupan los primeros puestos en lo que a adeudar pagos de arbitrios municipales e impuesto predial se refiere. Es decir los contribuyentes que deben las mayores cantidades de dinero. Aquí también se emplea el filtro por años (ejemplo: “cuáles son los contribuyentes que más montos adeudan por las cuotas del año 2014”).

Responsable: Felix Vargas Chumpitaz.

- **REQ09:** Mostrar la cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial por cuotas y fechas.

Se desea obtener información sobre la cantidad de contribuyentes que cancelan los pagos de arbitrios municipales e impuesto predial. Se busca conocer la cantidad de contribuyentes que pagaron una determinada cuota de un año en específico, y también la cantidad de contribuyentes que pagaron en un determinado lapso de tiempo. Para esto se utilizará el filtro por cuotas (ejemplo: “cuántos contribuyentes realizaron sus pagos por la cuota 01 del año 2013”) y por años (ejemplo: “cuántos contribuyentes realizaron sus pagos el primer trimestre del año 2011”).

Responsable: Felix Vargas Chumpitaz.

- **REQ10:** Mostrar el estado de la recaudación de los arbitrios municipales e impuesto predial según los KPIs.

Se desea obtener información sobre el estado actual de la recaudación de los arbitrios municipales y del impuesto predial basándonos en los indicadores clave de rendimiento. Se utilizará solo el filtro por años (ejemplo: “cómo va la recaudación del impuesto predial el presente año”).

Responsable: Felix Vargas Chumpitaz.



### 3.4.6. Hoja de Gestión

**Tabla 21.** Hoja de Gestión: Recaudación

Fuente: Elaboración Propia

Hoja de Gestión				
<b>Proceso</b>	Recaudación de Impuesto Predial y Arbitrios Municipales			
<b>Objetivo</b>	Obtener recursos monetarios a partir de la recaudación del impuesto predial y de los arbitrios municipales a los contribuyentes a fin de solventar los gastos de las actividades que realice la municipalidad en favor del distrito.			
<b>Estrategia</b>	Información en lugares concurridos sobre los modos de pago. Beneficios por pronto pago. Amnistías. Campañas tributarias los fines de semana.			
<b>Indicador</b>	<b>Indicadores</b>	<b>Medidas</b>	<b>Estados</b>	
	Montos por importe recaudados	sum(importe)	>90%	
			51% - 89%	
			<50%	
	Montos por derecho de emisión recaudados	sum(deremi)	>90%	
			51% - 89%	
			<50%	
	Montos por mora recaudados	sum(morcan)	>90%	
			51% - 89%	
			<50%	
	Montos totales recaudados	sum(importe+deremi+morcan)	>90%	
			51% - 89%	
<50%				

### 3.4.7. Hoja de Análisis

**Tabla 22.** Hoja de Análisis: Recaudación

Fuente: Elaboración Propia

Hoja de Análisis					
<b>Proceso</b>	Recaudación de Impuesto Predial y Arbitrios Municipales				
<b>Medidas</b>	Importe recaudado Derecho de emisión recaudado Mora recaudado Total recaudado				
<b>Interrogante</b>	<b>Dimensión</b>	<b>Maneras de analizar la dimensión</b>			
<b>¿Qué?</b>	Tipo Pago	<i>código</i>	<i>descripción</i>		
<b>¿Quién?</b>	Contribuyente	<i>nombre</i>	<i>código</i>		
<b>¿Cómo?</b>	Cuota	<i>año</i>	<i>cuota</i>	<i>descripción</i>	
	Estado	<i>código</i>	<i>descripción</i>		
<b>¿Dónde?</b>	Zona	<i>zona</i>	<i>microzona</i>		
<b>¿Cuándo?</b>	Tiempo	<i>año</i>	<i>trimestre</i>	<i>mes</i>	<i>fecha</i>

## 3.5. MODELO DIMENSIONAL

### 3.5.1. Dimensiones

**Tabla 23.** Dimensiones

Fuente: Elaboración Propia

DIMENSIÓN	DESCRIPCIÓN
Contribuyente	Almacena a todos los contribuyentes del distrito de Lurín registrados en el sistema transaccional.
Cuota	Almacena los números de las cuotas de los pagos de arbitrios municipales e impuesto predial y sus años.
Tipo Pago	Almacena los tipos de pago con los que se desea trabajar (Impuesto predial y Arbitrios municipales).
Zona	Almacena todas las microzonas del distrito de Lurín registradas en el sistema transaccional con la zona a la cual pertenece.
Tiempo	Almacena las fechas en las que se realizaron los pagos de impuesto predial y arbitrios municipales.
Estado	Almacena los estados de los pagos de arbitrios municipales e impuesto predial.

### 3.5.2. Dimensión: Contribuyente

**Tabla 24.** Dimensión Contribuyente

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Nombre	Nombre completo de un contribuyente.	Vanesa Soto Vásquez
Código Interno	Código interno del contribuyente en el sistema de recaudación	2186214, 2196010

### 3.5.3. Dimensión: Cuota

**Tabla 25.** Dimensión Cuota

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Descripción	Cuota ligada al año al que pertenece	201301
Cuota	Cuota a la que pertenece el pago	01, 09
Año	Año correspondiente a la cuota	2012, 2006

### 3.5.4. Dimensión: Tipo

**Tabla 26.** Dimensión Tipo

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Nombre	Nombre literal del tipo de pago.	Impuesto Predial, Arbitrios Municipales
Código del pago	Código interno del tipo de pago en el sistema de recaudación	001, 004

### 3.5.5. Dimensión: Zona

**Tabla 27.** Dimensión Zona

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Cód. Predio	Código del predio de la microzona	0154827
Nombre	Nombre completo de la microzona.	AA:HH José Olaya, Lurín Cercado
Zona	Nombre de la zona a la cual pertenece la microzona	Zona A, Zona D

### 3.5.6. Dimensión: Tiempo

Tabla 28. Dimensión Tiempo

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Fecha	Fecha escrita mediante números.	2014-03-06
Mes	Nombre completo del mes en español.	Marzo, Julio
Trimestre	Representación del trimestre con números romanos.	II, IV
Año	Representación numérica del año correspondiente a una fecha	2012, 2009

### 3.5.7. Dimensión: Estado

Tabla 29. Dimensión Estado

Fuente: Elaboración Propia

ATRIBUTO	DESCRIPCIÓN	EJEMPLO
Cód. Estado	Código con el que se identifica un determinado estado.	C, P
Descripción	Nombre que recibe un determinado estado.	Cancelado, Pendiente

### 3.5.8. Granularidad

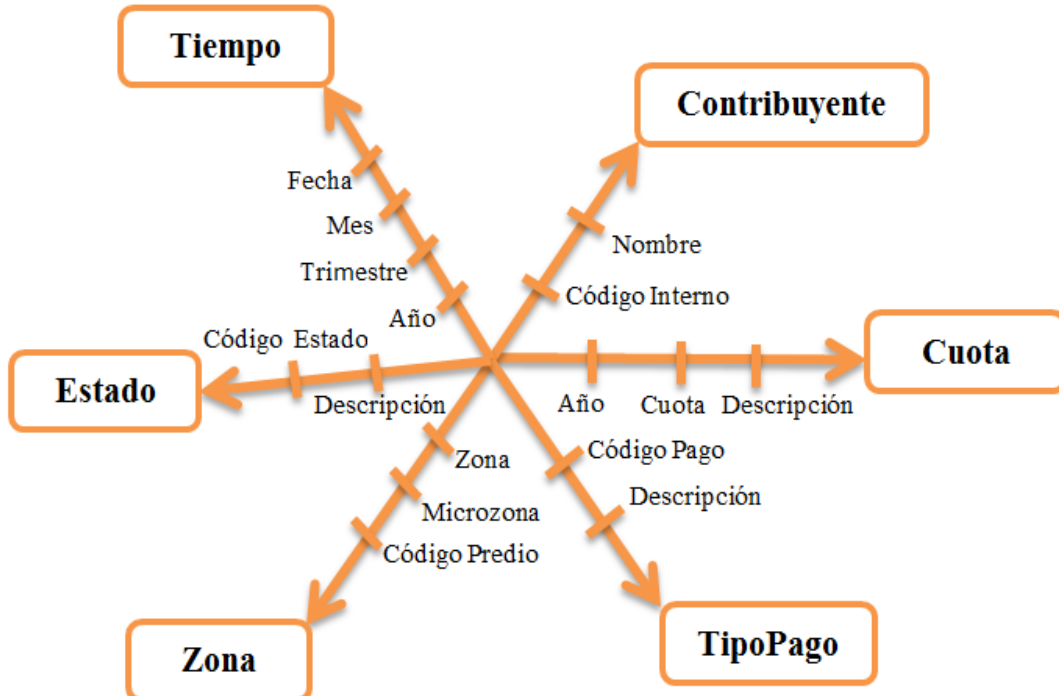


Figura 19. Determinación de nivel de granularidad

Fuente: Elaboración Propia

**Tabla 30.** Tablas de Hechos: Recaudación

Fuente: Elaboración Propia

TABLA DE HECHOS	DESCRIPCIÓN
Recaudación de Impuesto Predial y Arbitrios Municipales	Tabla de hechos referida a los montos de impuesto predial y arbitrios municipales recaudados y por recaudar a los contribuyentes del distrito de Lurín.

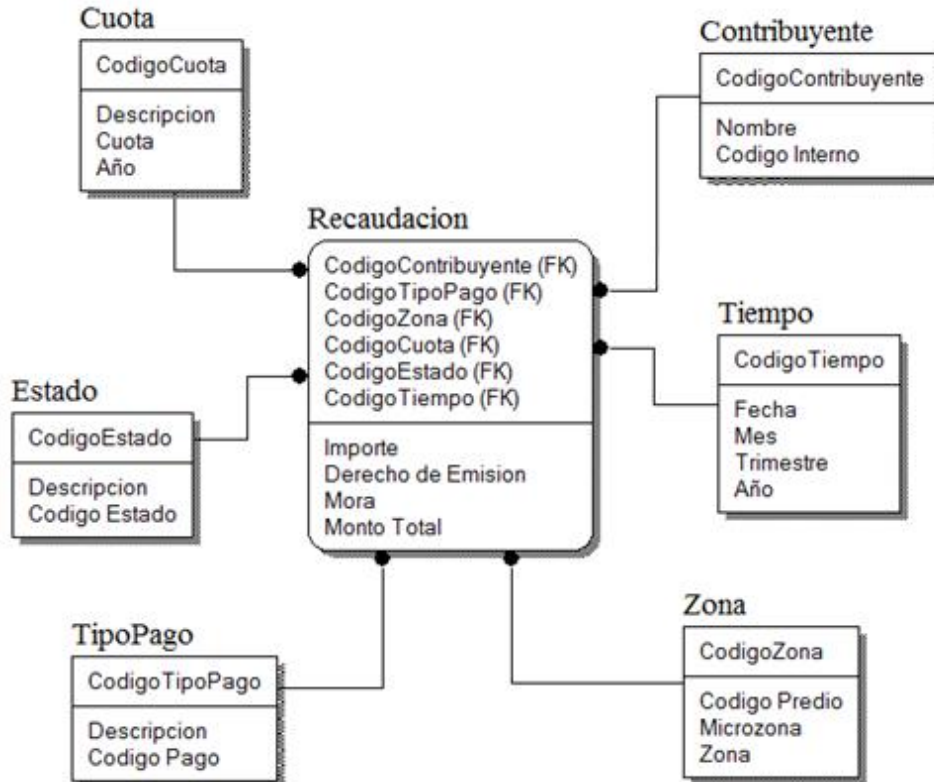
**3.5.10. Medidas**

**Tabla 31.** Medidas

Fuente: Elaboración Propia

MEDIDA	FÓRMULA
Importe	$\sum(\text{importe})$
Derecho de Emisión	$\sum(\text{derecho de emisión})$
Mora	$\sum(\text{mora})$
Monto Total	$\sum(\text{importe} + \text{derecho de emisión} + \text{mora})$

**3.5.11. Diseño del Modelo Estrella**



**Figura 20.** Modelo Dimensional: Recaudación

Fuente: Elaboración Propia

### 3.6. DISEÑO DE LA ARQUITECTURA TÉCNICA

#### 3.6.1. Arquitectura

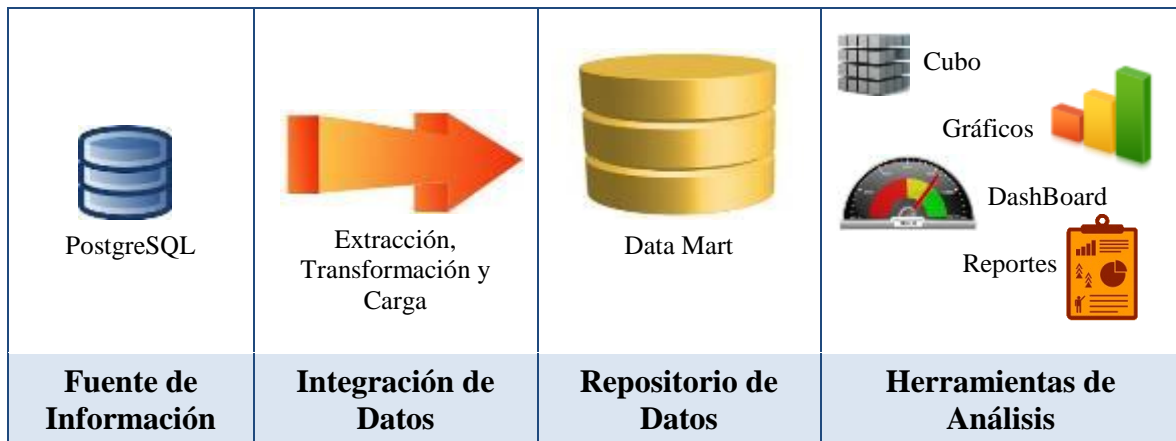


Figura 21. Arquitectura Técnica

Fuente: Elaboración Propia

La solución solo podrá ser accedida desde la red local (LAN). Como podemos apreciar en la Figura 21, la arquitectura técnica de la solución consta de 4 partes:

- Fuente de información: Base de datos del sistema de recaudación.
- Integración de Datos: Proceso de extracción, transformación y carga de los datos de la BD del sistema de recaudación hacia el repositorio de datos.
- Repositorio de Datos: Data Mart propuesto para el área de rentas, el cual almacenará los datos extraídos del sistema de recaudación.
- Herramientas de Análisis: Cubo dimensional, gráficos, dashboard y reportes.

#### 3.6.2. Back-room

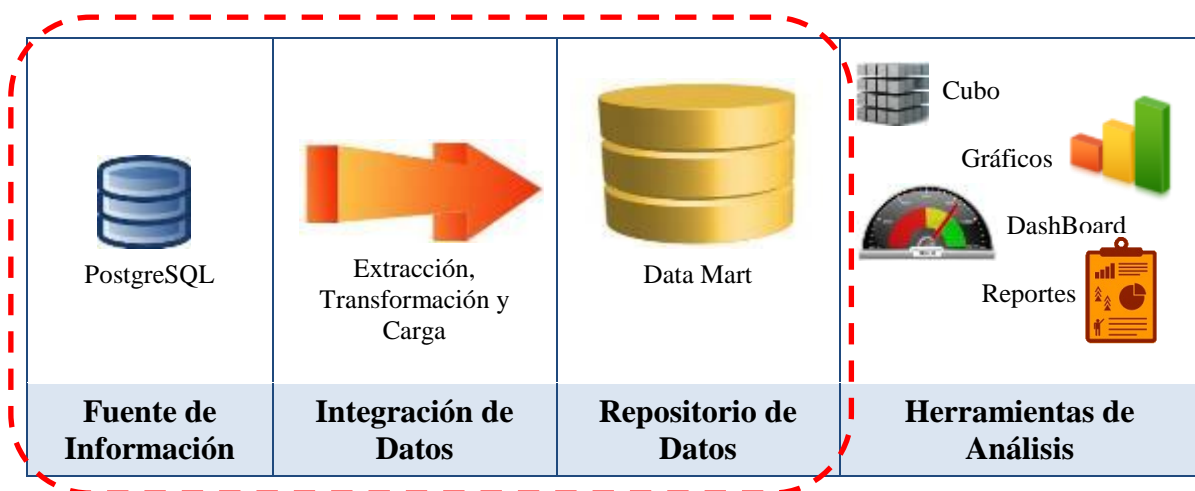


Figura 22. Back - Room

Fuente: Elaboración Propia

- ◆ El servidor de base de datos del sistema de recaudación es un servidor físico HP-PC, Intel Xeon E5320 1.86ghz (x64), con 16gb de RAM y con 4 terabytes de disco duro. Sistema operativo Windows Server 2008 y SGBD PostgreSQL 9.3.
- ◆ El servidor de base de datos que alojará al Data Mart será un servidor virtualizado dentro del servidor físico de base de datos, y tendrá disponible 4gb de RAM y 500 gigabytes de espacio de almacenamiento dinámico. Sistema operativo Windows Server 2008 y SGBD PostgreSQL 9.3.

### 3.6.3. Front-room

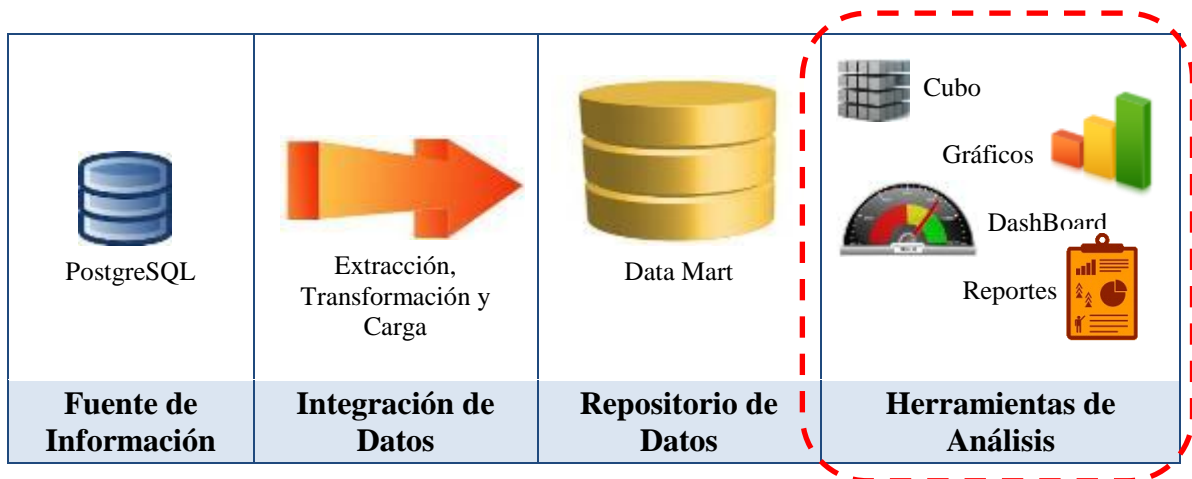


Figura 23. Front - Room

Fuente: Elaboración Propia

- ◆ La estación de trabajo desde la cual el gerente de rentas podrá acceder a las herramientas de análisis (cubo dimensional, gráficos, dashboards y reportes) es una PC Intel Dual Core G620 2.60GHz (x64), con 4gb de RAM y 500gb de disco duro. Sistema operativo Windows 7.

## 3.7. SELECCIÓN E INSTALACIÓN DE PRODUCTOS

### 3.7.1. Evaluación de los productos

Para empezar, se seleccionan los productos a evaluarse. En nuestro caso contaremos con 6 opciones, 3 de ellas gratuitas y 3 de ellas comerciales, las cuales son:

- SQL Server Business Intelligence
- IBM Cognos Business Intelligence
- MicroStrategy

- Pentaho
- JasperSoft Business Intelligence Suite
- Jedox Suite

Una vez observadas las alternativas de productos a utilizarse para el desarrollo de la solución, se procede a realizar una matriz en la cual se evalúa cada una de ellas. Para esto se realiza una comparación a fin de determinar cuál se ellas se adapta más a nuestras necesidades. La matriz de evaluación puede ser observada en la Tabla 32.

**Tabla 32.** Matriz de Evaluación

Fuente: Elaboración Propia

	SQL Server Business Intelligence	IBM Cognos Business Intelligence	MicroStrategy	Pentaho	JasperSoft Business Intelligence Suite	Jedox Suite
Suite BI completa	X	X	X	X	X	X
Software Libre				X	X	X
Independencia de la base de datos	X	X	X	X	X	X
Reportes en entorno web	X	X	X	X	X	X
Autenticación de usuarios			X	X	X	X
Alto nivel de madurez	X	X	X	X		

### 3.7.2. Herramientas seleccionadas

Luego de la evaluación de las alternativas de herramientas a usarse para el desarrollo de la solución BI, se vio que la mejor opción para nuestro caso es Pentaho, sobre todo porque es una suite completa para desarrollar soluciones de inteligencia de negocios y cuenta con una versión gratuita, la cual no tiene funciones que se encuentren limitadas y/o restringidas.

Pentaho cuenta con muchas herramientas para el desarrollo de soluciones BI (muchas de ellas realizan las mismas acciones), las cuales pueden instalarse por separado y de acuerdo a las necesidades del usuario. De igual manera cuenta con plugins para aumentar las funcionalidades del servidor BI.

A continuación se mostrarán las herramientas y plugins a utilizarse para el desarrollo de nuestra solución.



### Pentaho Data Integration (Kettle) 5.2.0.0

Herramienta que nos permite realizar el proceso ETL, es decir extraer la información de una base de datos OLTP, transformar dicha información de acuerdo a nuestras necesidades y posteriormente cargarla en una base de datos destino (almacén de datos), para que pueda ser consultada y analizada mediante herramientas analíticas y de desarrollo de reportes..

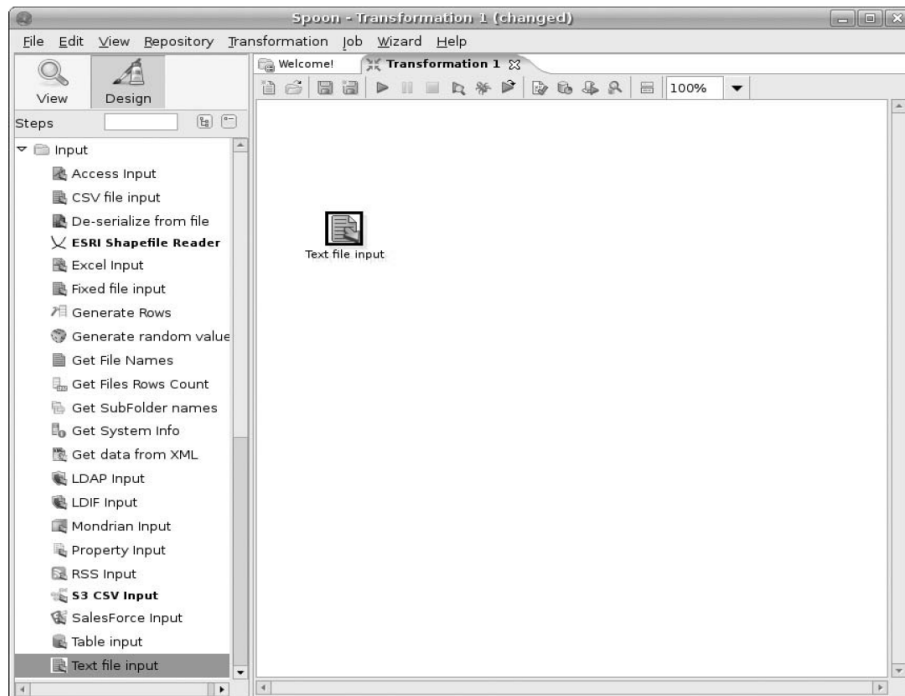


Figura 24. Pentaho Data Integration (Kettle)

Fuente: Bouman, R.

### Pentaho Schema Workbench 3.8.0.0

Herramienta que nos facilita la creación de cubos OLAP ya que nos permite elaborarlos en un entorno gráfico. Además podemos publicar nuestros cubos directamente en el Pentaho BI Server.

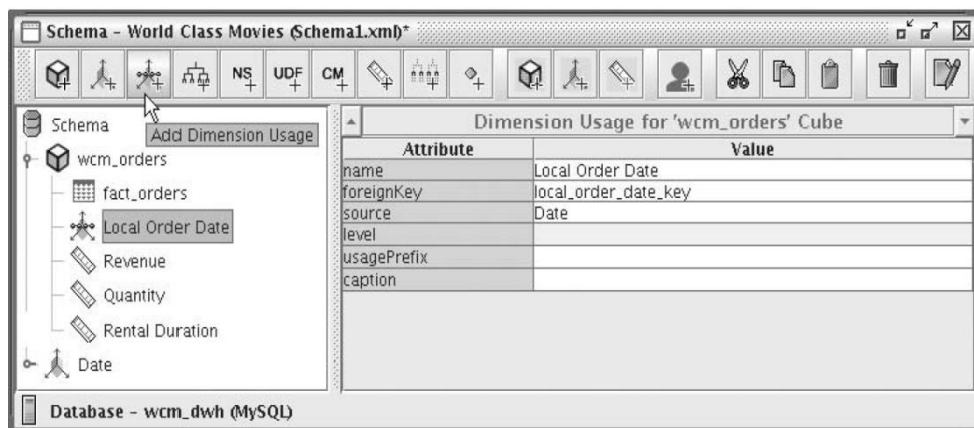


Figura 25. Pentaho Schema Workbench

Fuente: Bouman, R.

### Pentaho Design Studio 4.0.0

Es una herramienta, basada en el IDE Eclipse, que se encarga de la creación de Xactions, las cuales son un conjunto de acciones ejecutadas en el Pentaho BI Server ya sea a petición de un usuario o automáticamente previa programación.

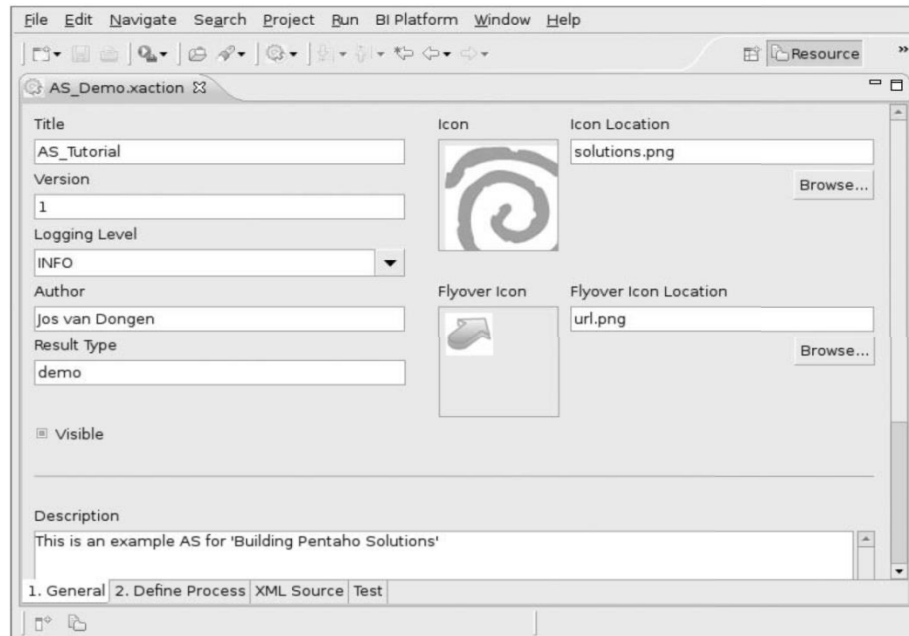


Figura 26. Pentaho Design Studio

Fuente: Bouman, R.

### JPivot

Es un plugin integrado en el Pentaho BI server que nos permite navegar a través de nuestros cubos OLAP mostrándonos su información mediante tablas y gráficos, y con funciones de expandir y contraer, ordenar, mostrar y ocultar, etc. Además nos da la opción de exportación a hojas de cálculo.

Date	DVD	Customer	Measures
All Date.Months	All DVDs	All Customers	• Revenue
+ 2000	+ All DVDs	+ All Customers	≈10,921,326.03
+ 2001	+ All DVDs	+ All Customers	≈2,985.87
+ 2002	+ All DVDs	+ All Customers	≈24,030.58
+ 2003	+ All DVDs	+ All Customers	≈67,190.34
+ 2004	+ All DVDs	+ All Customers	≈124,674.01
			≈272,594.35

Figura 27. JPivot

Fuente: Bouman, R.

### Pentaho Report Designer 5.2.0.0

Es una herramienta que nos permite diseñar nuestros reportes desde una interfaz de usuario de fácil entendimiento para el usuario. Además nos permite publicar nuestros reportes directamente en el Pentaho BI Server.

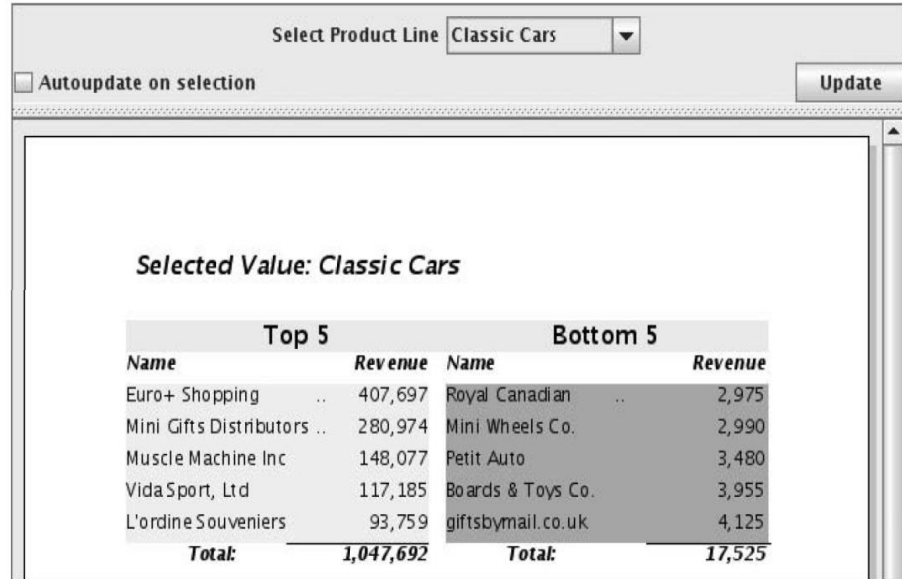


Figura 28. Pentaho Report Designer

Fuente: Bouman, R.

### Pentaho BI Server 5.2.0.0

Es el componente principal de toda la suite Pentaho. Es la que se encarga de administrar las cuentas de usuario de la aplicación web, los cubos OLAP, los reportes generados, los Xaction o secuencia de acciones, ect. También cuenta con su propio diseñador de dashboards. Además nos permite aumentar sus funcionalidades ya que soporta la instalación de diversos plugins como el J4Pivot, IvyDC, entre otros.



Figura 29. Pentaho BI Server

Fuente: Elaboración Propia

### Community Dashboard Editor (CDE) 5.2.0.0

Plugin que nos permite la creación, edición y publicación de tableros de control (dashboards). Podemos realizar gráficos de barras, gráficos circulares, lineales, etc.

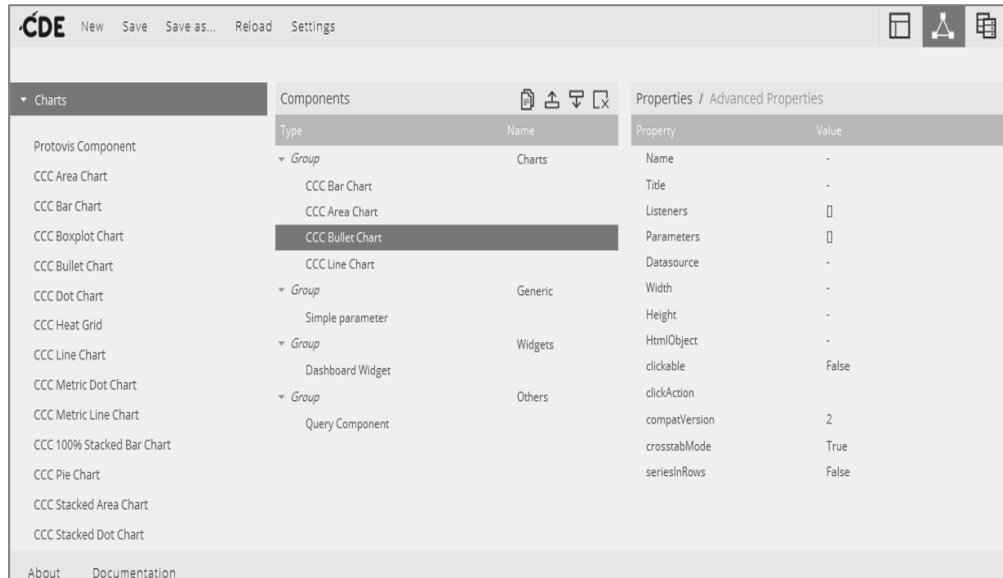


Figura 30. CDE

Fuente: Elaboración Propia

### Ivy Dashboard Components (IvyDC) 0.0.4

Plugin del Pentaho BI Server que agrega nuevos componentes al Community Dashboard Editor, es decir nos permite enriquecer nuestros dashboards.

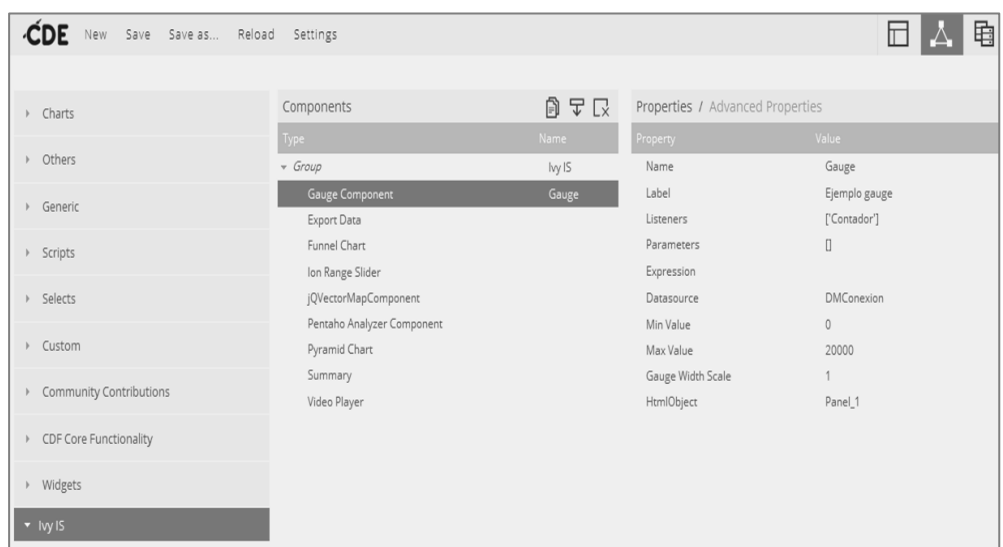


Figura 31. IvyDC

Fuente: Elaboración Propia

### PostgreSQL 9.3

Es un sistema de gestión de bases de datos objeto-relacional con código fuente libre desde el cual podemos crear y administrar nuestras bases de datos. Cuenta con una interfaz gráfica llamada pgAdmin III.

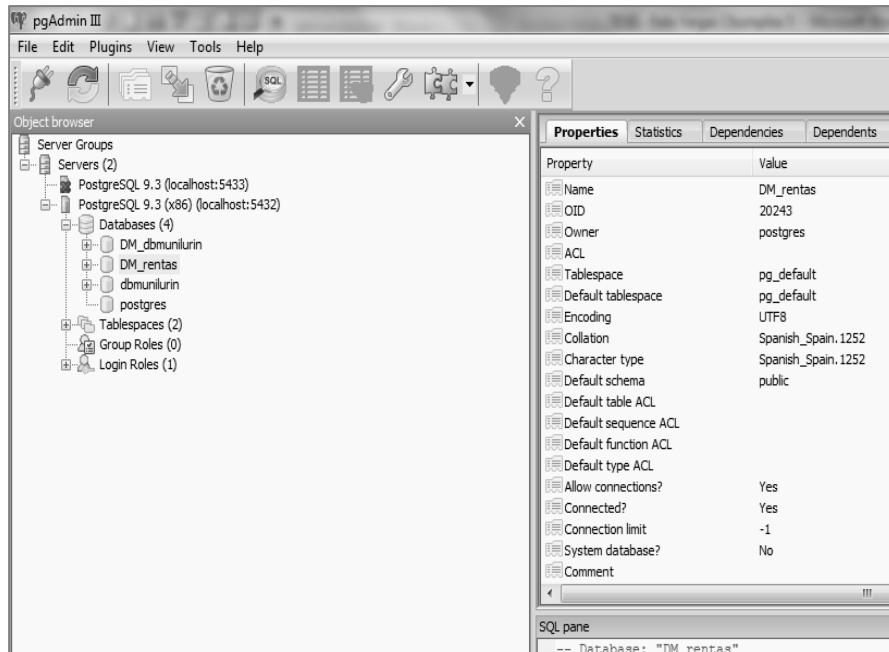


Figura 32. PostgreSQL 9.3

Fuente: Elaboración Propia

### Oracle VM VirtualBox 4.3.12

Software gratuito que nos permite la creación de máquinas virtuales dentro de un servidor/pc físico, es decir podemos correr múltiples sistemas operativos al mismo tiempo dentro de una misma computadora.



Figura 33. Oracle VM VirtualBox

Fuente: Elaboración Propia

### 3.8. DISEÑO FÍSICO

#### 3.8.1. Tablas de Apoyo

Para llenar la tabla de hechos se necesita cruzar datos contenidos tanto en la base de datos dimensional como en la base de datos transaccional. PostgreSQL no permite realizar este tipo de consultas de una manera normal. Para lograrlo, se necesita usar la extensión DBlink.

Sin embargo, al usar la extensión DBlink en nuestras consultas, éstas demoran más en ejecutarse. Es así que se decidió usar una tabla de apoyo (apoyo\_cta\_aica) para hacer posible nuestro llenado. Además debido a que las consultas son algo complejas y por ende algo tardías, se optó por crear dos tablas de apoyo más (apoyo\_microzona, apoyo\_predios\_ur), con lo cual el tiempo que demoran en ejecutarse dichas consultas se redujeron.

Para llegar a estas conclusiones se tuvo que realizar distintas pruebas. Esto con el objetivo de encontrar el mejor camino para llevar a cabo el llenado de la base de datos dimensional.

##### 3.8.1.1. Tabla Apoyo: apoyo\_cta\_aica

**Tabla 33.** Diseño Físico: Apoyo Cta Aica

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
codcli	char(7)	NO	NO
codpre	char(7)	NO	NO
cuenta	char(3)	NO	NO
actualcuota	char(4)	NO	NO
estado	char(1)	NO	NO
feccan	date	NO	NO
importe	numeric(12,2)	NO	NO
deremi	numeric(12,2)	NO	NO
morcan	numeric(12,2)	NO	NO
total	numeric(12,2)	NO	NO

##### 3.8.1.2. Tabla Apoyo: apoyo\_microzona

**Tabla 34.** Diseño Físico: Apoyo Microzona

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
desurb	varchar(80)	NO	NO
deszon	varchar(10)	NO	NO

**3.8.1.3. Tabla Apoyo: apoyo\_predios\_ur****Tabla 35.** Diseño Físico: Apoyo Predios UR

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
codpred	char(7)	NO	NO
desurb	varchar(80)	NO	NO

**3.8.2. Tabla Dimensión: DIM\_Contribuyente****Tabla 36.** Diseño Físico: Dimensión Contribuyente

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
contribuyente_key	serial	SI	NO
nombre	varchar(100)	NO	NO
cod_interno	char(7)	NO	NO

**3.8.3. Tabla Dimensión: DIM\_Zona****Tabla 37.** Diseño Físico: Dimensión Microzona

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
microzona_key	serial	SI	NO
codpred	char(7)	NO	NO
microzona	varchar(80)	NO	NO
zona	varchar(10)	NO	NO

**3.8.4. Tabla Dimension: DIM\_Cuota****Tabla 38.** Diseño Físico: Dimensión Cuota

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
cuota_key	serial	SI	NO
descripción	varchar(6)	NO	NO
cuota	varchar(2)	NO	NO
anio	varchar(4)	NO	NO

### 3.8.5. Tabla Dimensión: DIM\_TipoPago

Tabla 39. Diseño Físico Dimensión TipoPago

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
tipo_key	serial	SI	NO
descripción	varchar(20)	NO	NO
codtipo	char(3)	NO	NO

### 3.8.6. Tabla Dimensión: DIM\_Tiempo

Tabla 40. Diseño Físico: Dimensión Tiempo

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
tiempo_key	serial	SI	NO
fecha	date	NO	NO
mes	varchar(10)	NO	NO
trimestre	varchar(3)	NO	NO
anio	char(4)	NO	NO

### 3.8.7. Tabla Dimensión: DIM\_Estado

Tabla 41. Diseño Físico: Dimensión Estado

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
tipo_key	serial	SI	NO
descripción	varchar(20)	NO	NO
codtipo	char(3)	NO	NO

### 3.8.8. Tabla de Hechos: FACT\_Recaudación

Tabla 42. Diseño Físico: Hecho Recaudación

Fuente: Elaboración Propia

Columna	Tipo de Dato	Es Llave Primaria	Nulo
tipo_key	integer	SI	NO
cuota_key	integer	SI	NO



contribuyente_key	integer	SI	NO
microzona_key	integer	SI	NO
tiempo_key	integer	SI	NO
estado_key	integer	SI	NO
importe	numeric(12,2)	NO	NO
derecho_emision	numeric(12,2)	NO	NO
mora	numeric(12,2)	NO	NO
monto_total	numeric(12,2)	NO	NO

### 3.8.9. Diseño Modelo Físico

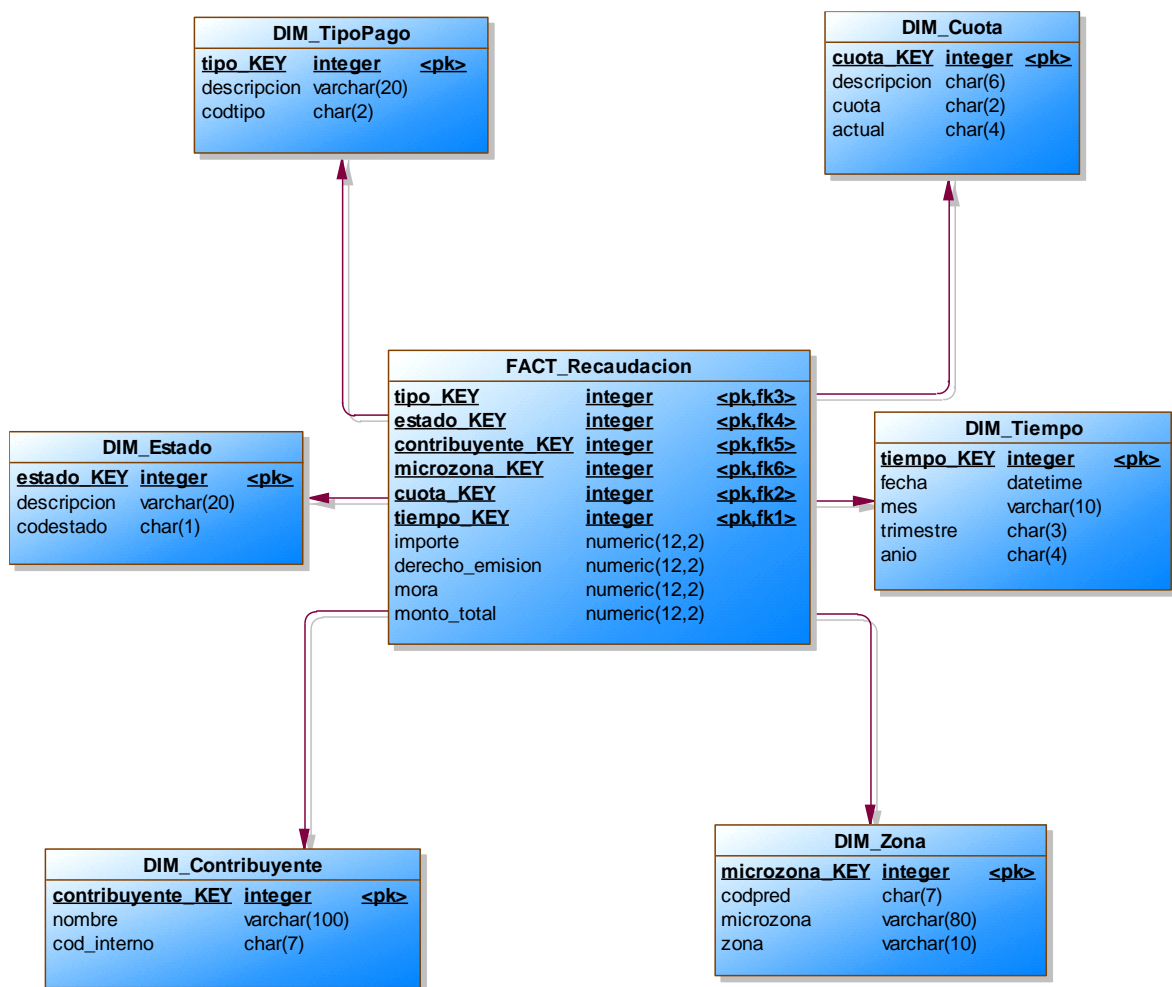
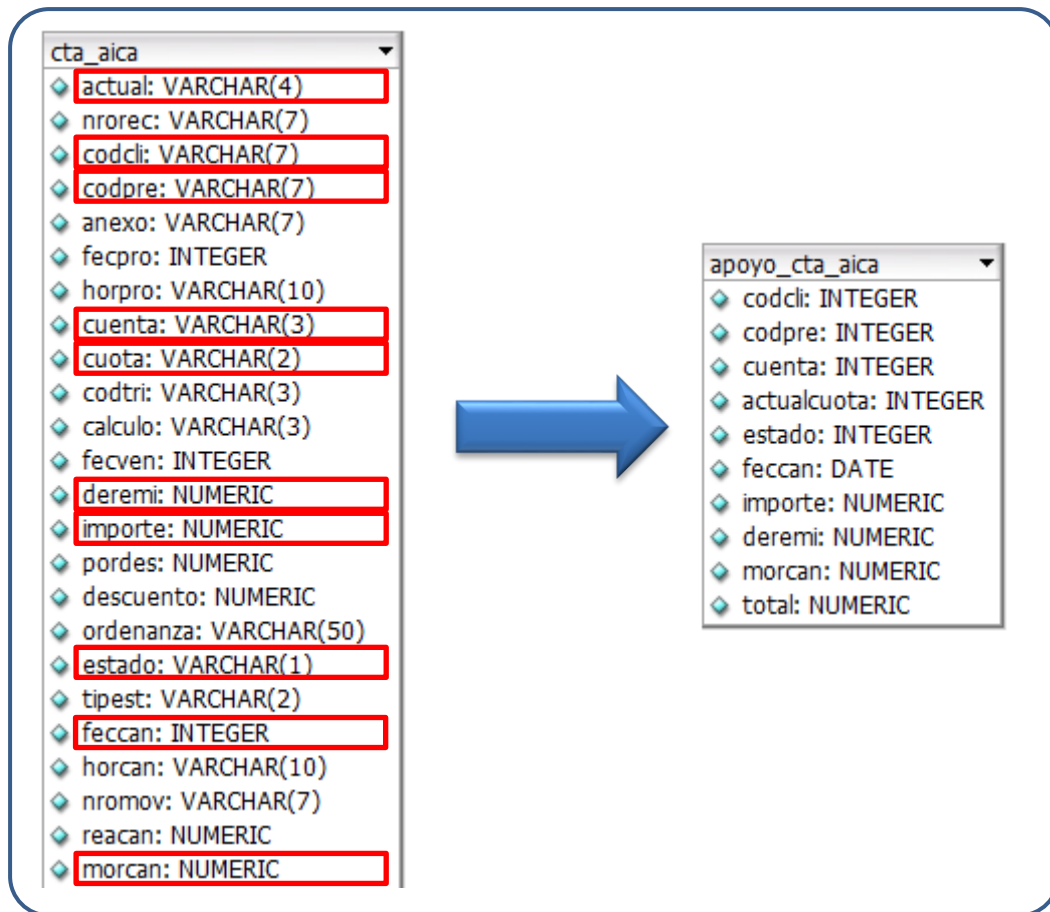


Figura 34. Modelo Físico: Recaudación

Fuente: Elaboración Propia

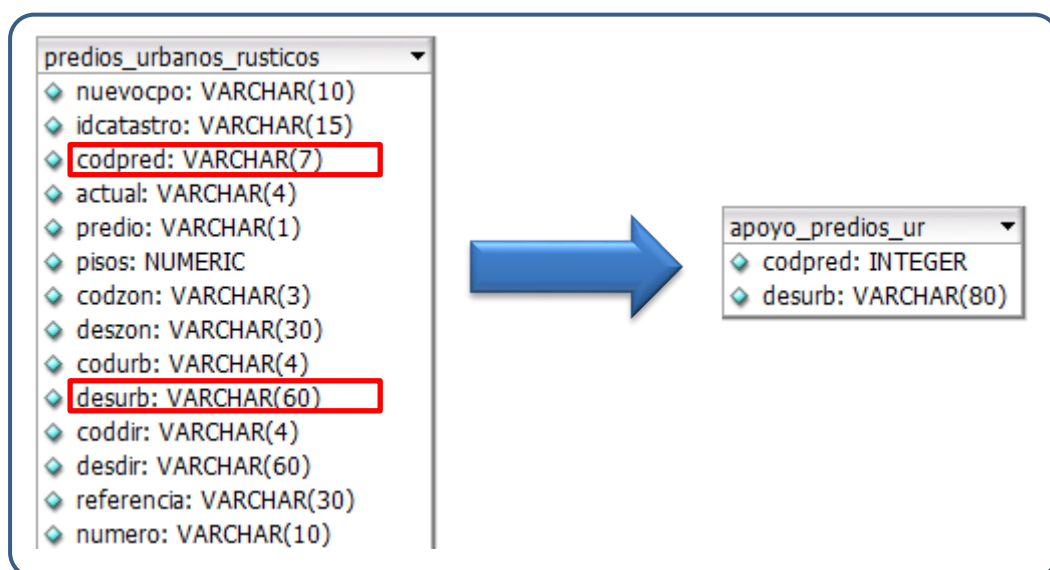
### 3.9. DISEÑO ETL

#### 3.9.1. Extracción (Extract)



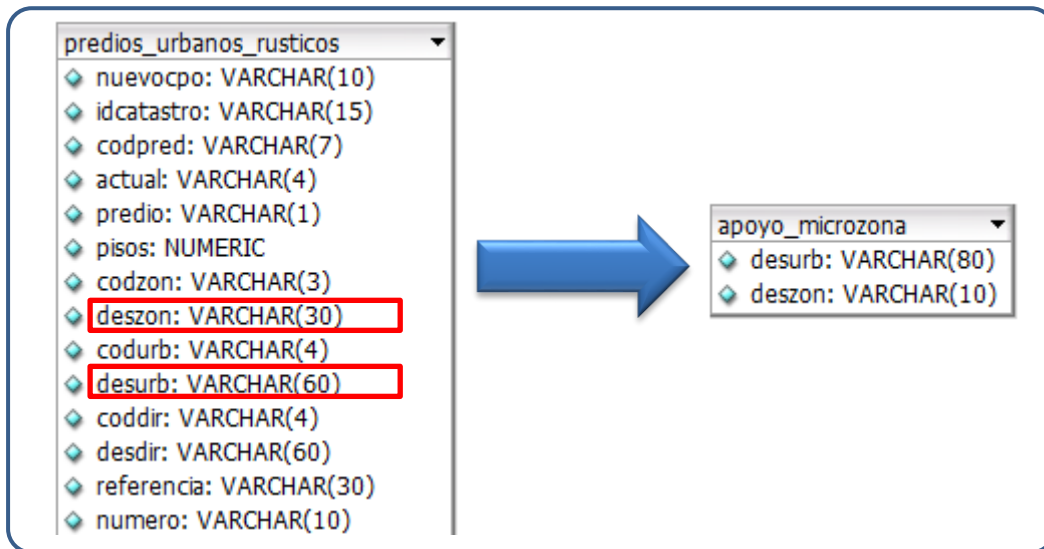
**Figura 35.** Extracción para: Tabla apoyo\_cta\_aica

Fuente: Elaboración Propia

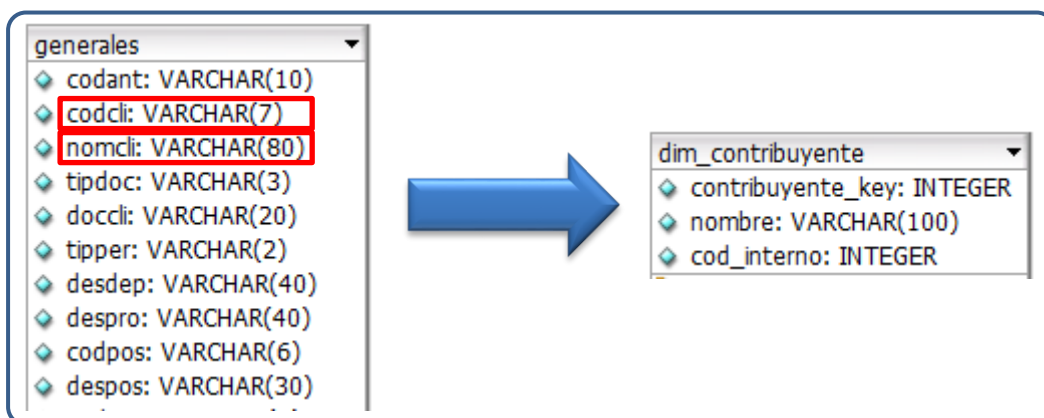


**Figura 36.** Extracción para: Tabla apoyo\_predios\_ur

Fuente: Elaboración Propia



**Figura 37.** Extracción para: Tabla apoyo\_microzona Fuente: Elaboración Propia



**Figura 38.** Extracción para: Tabla Dim\_Contribuyente Fuente: Elaboración Propia



**Figura 39.** Extracción para: Tabla Dim\_Microzona Fuente: Elaboración Propia

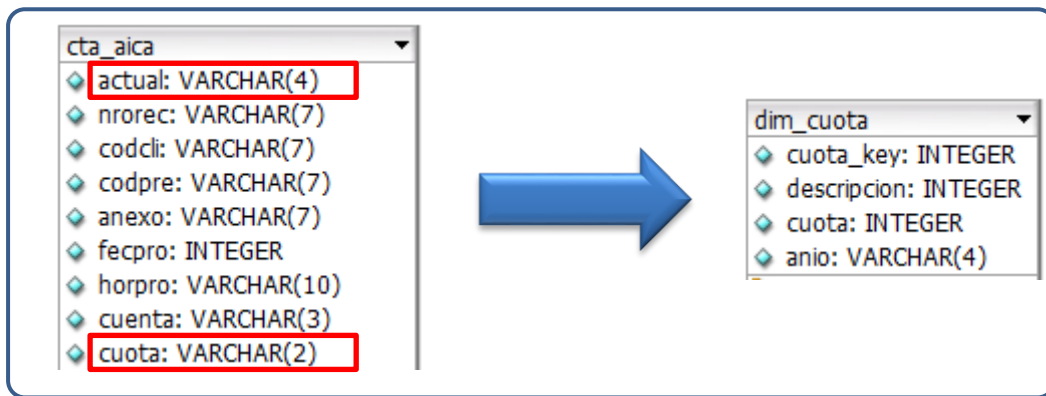


Figura 40. Extracción para: Tabla Dim\_Cuota

Fuente: Elaboración Propia

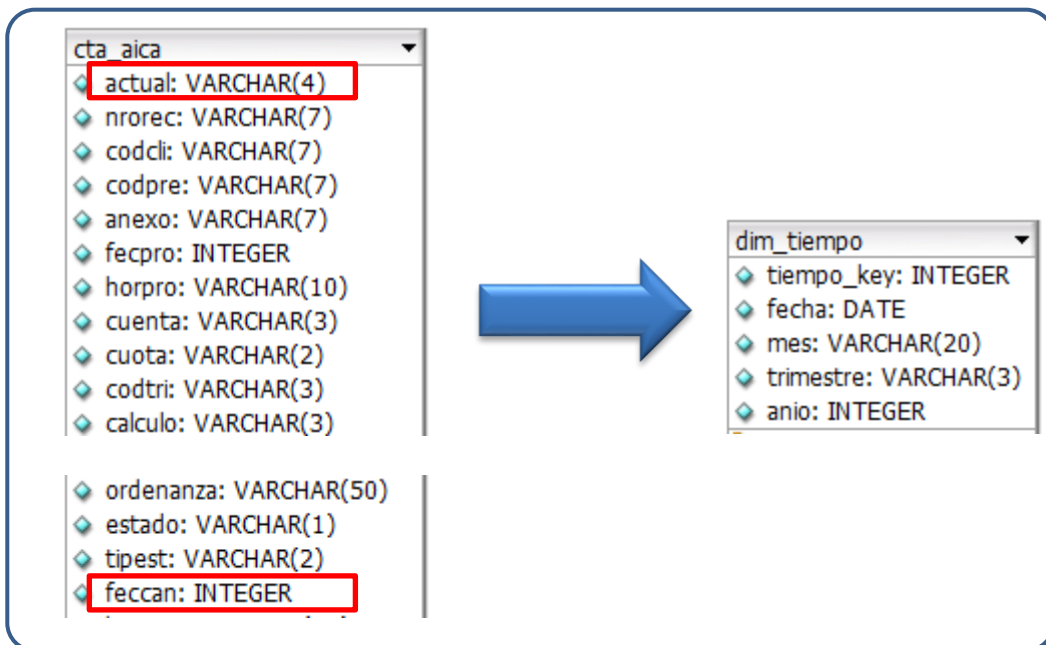


Figura 41. Extracción para: Tabla Dim\_Tiempo

Fuente: Elaboración Propia

### 3.9.2. Transformación (Transformation)

```

select codcli,codpre,cuenta,actualcuota,estado,feccan,sum(deremi)
as deremi,sum(importe) as importe,sum(morcan) as morcan,sum(total)
as total from
(
  select cta.codcli,codpre,cuenta,actual||cuota as actualcuota,cta.estado,
  case when feccan is null then '0001-01-01'::date else feccan::date end
  as feccan,sum(case cuota when '01' then deremi else 0.00 end) as deremi,
  sum(importe) as importe,sum(morcan) as morcan,
  sum(importe+deremi+morcan) as total
  from cta_aica cta
  inner join generales gene on gene.codcli=cta.codcli
  where (actual between '2003' and '2014') and cuenta in ('001','004')
  and (cuota between '01' and '12') and (cta.estado = 'C' or
  (cta.estado in ('E','P') and fecven::date < now()::date))
  group by cta.codcli,codpre,cuenta,actual,cuota,cta.estado,feccan
) as ccc
group by codcli,codpre,cuenta,actualcuota,estado,feccan
    
```

Figura 42. Transformación para: Tabla apoyo\_cta\_aica

Fuente: Elaboración Propia

```

select codpred,trim(desurb) as desurb from predios_urbanos_rusticos
where actual='2014'
--Se agrega una fila para que no se pierdan registros
--en los INNER JOIN
INSERT INTO apoyo_predios_ur(codpred,desurb)
values ('0000000','LIBRE')

```

**Figura 43.** Transformación para: Tabla apoyo\_predios\_ur Fuente: Elaboración Propia

```

select desurb,deszon from(
  select trim(deszon) as deszon,trim(desurb) as desurb
  from predios_urbanos_rusticos
  where codzon in ('001','002','003','004','005')
  and actual='2014' and desurb not in ('SIN SECTOR')
  group by desurb,deszon
) as ccc group by desurb,deszon
--Se agrega una fila para que no se pierdan
--registros en los INNER JOIN
INSERT INTO apoyo_microzona(desurb,deszon)
values ('LIBRE','LIBRE')

```

**Figura 44.** Transformación para: Tabla apoyo\_microzona Fuente: Elaboración Propia

```

SELECT nomcli,codcli FROM generales
where nomcli not in ('') order by codcli

```

**Figura 45.** Transformación para: Tabla Dim\_Contribuyente Fuente: Elaboración Propia

```

select apur.codpred, apur.desurb, am.deszon
from apoyo_predios_ur apur
inner join apoyo_microzona am on am.desurb = apur.desurb

```

**Figura 46.** Transformación para: Tabla Dim\_Zona Fuente: Elaboración Propia

```

SELECT DISTINCT cuota,actual,actual||cuota as descripcion
FROM cta_aica
WHERE cuenta in ('001','004') AND codcli='0000005'
AND actual::integer between 2003 and 2014
order by actual,cuota

```

**Figura 47.** Transformación para: Tabla Dim\_Cuota Fuente: Elaboración Propia

```

INSERT INTO dim_tipoPago(descripcion,codtipo) values
('IMP. PREDIAL','001'),
('ARB. MUNICIPALES','004')

```

**Figura 48.** Transformación para: Tabla Dim\_TipoPago Fuente: Elaboración Propia

```
INSERT INTO dim_estado(codestado,descripcion) values
('C','Cancelado'),
('P','Pendiente'),
('E','Saldo')
```

Figura 49. Transformación para: Tabla Dim\_Estado

Fuente: Elaboración Propia

```
select feccan as fecha,
lpad(date_part('month',feccan)::varchar,2,'0')||'-'||to_char(feccan,'TMonth')
as mes, case when date_part('month', feccan) between 1 and 3 then 'I'
when date_part('month', feccan) between 4 and 6 then 'II'
when date_part('month', feccan) between 7 and 9 then 'III' else 'IV' end
as trimestre, date_part('year',feccan) as anio
from
(
select distinct feccan::date from cta_aica where feccan is not null
) as ccc
--Se agrega una fila para que no se pierdan
--registros en los INNER JOIN
INSERT INTO dim_tiempo(fecha,mes,trimestre,anio)
VALUES ('0001-01-01','-----','-','----')
```

Figura 50. Transformación para: Tabla Dim\_Tiempo

Fuente: Elaboración Propia

```
select tipo.tipo_key, cuota.cuota_key, contri.contribuyente_key,
micro.microzona_key, est.estado_key, time.tiempo_key,
importe, deremi, morcan, total
from apoyo_cta_aica apoyo
inner join dim_tipopago tipo on tipo.codtipo=apoyo.cuenta
inner join dim_cuota cuota on cuota.descripcion=apoyo.actualcuota
inner join dim_contribuyente contri on contri.cod_interno=apoyo.codcli
inner join dim_zona micro on micro.codpred=apoyo.codpre
inner join dim_estado est on est.codestado=apoyo.estado
inner join dim_tiempo time on time.fecha=apoyo.feccan
```

Figura 51. Transformación para: Tabla FACT\_Recaudacion

Fuente: Elaboración Propia

### 3.9.3. Carga (Load)

Para reducir el tiempo de carga se hizo una división de todo el proceso en dos partes, así, tenemos dos trabajos y cada uno de ellos con sus respectivas transformaciones y scripts:

- ✓ Trabajo ETL parte 1
- ✓ Trabajo ETL parte 2

A continuación se pasará a detallar cada una de estas partes para un mejor entendimiento.

#### Trabajo ETL parte 1

La primera parte del proceso ETL es la que se encarga de realizar el llenado de las tablas de apoyo que permiten agilizar el llenado total del data mart. En la Figura 52 se puede apreciar el recorrido que lleva esta primera parte.

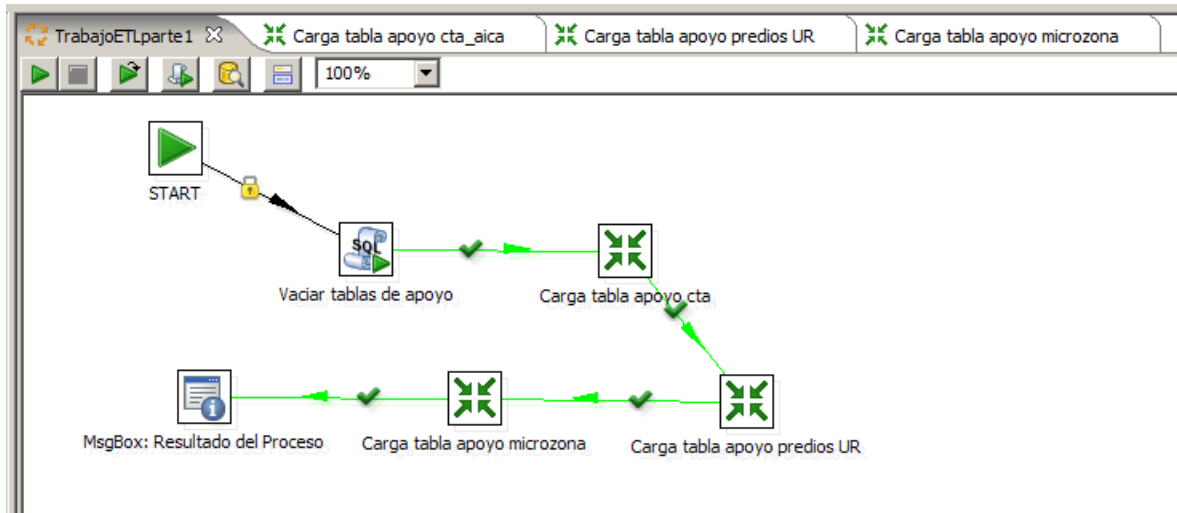


Figura 52. Trabajo: Trabajo ETL parte 1

Fuente: Elaboración Propia

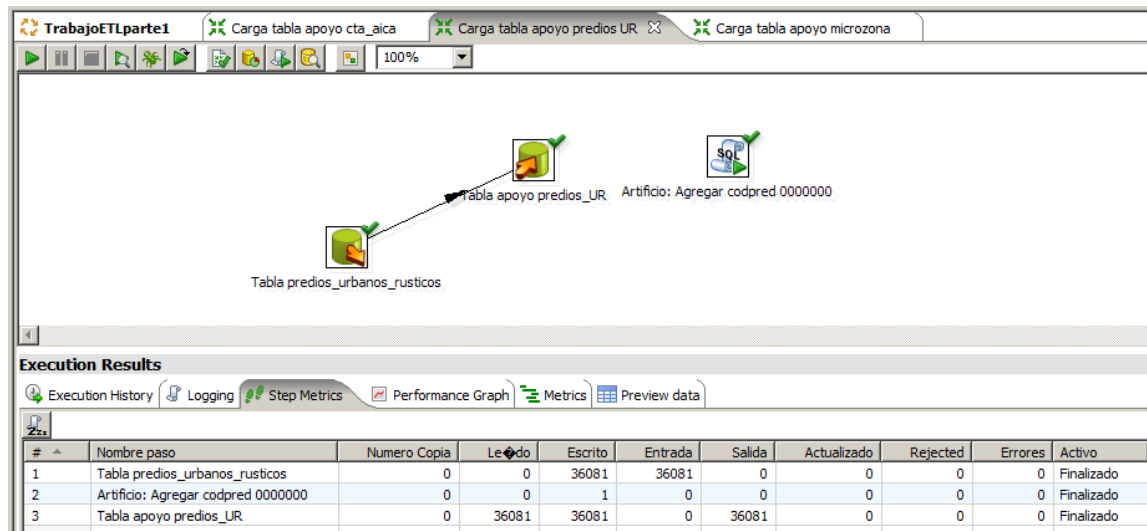
La transformación Carga tabla apoyo cta\_aica traslada los datos de la tabla cta\_aica de la base de datos transaccional, cumpliendo algunas condiciones estipuladas, hacia nuestra tabla de apoyo ubicada en nuestra base de datos dimensional (apoyo\_cta\_aica). Para más detalle se puede acudir a la Figura 53.

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Tabla cta_aica	0	0	3009545	3009545	0	0	0	0	Finalizado
2	Tabla Apoyo cta	0	3009545	3009545	0	3009545	0	0	0	Finalizado

Figura 53. Transformación: Carga tabla apoyo cta aica

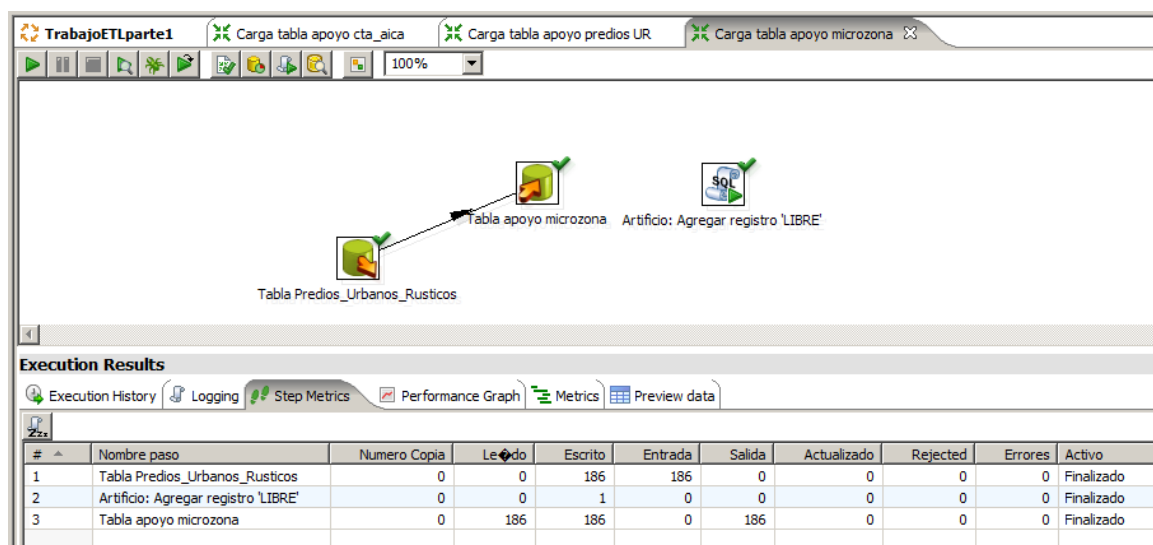
Fuente: Elaboración Propia

La transformación Carga tabla apoyo predios UR nos llena una tabla de apoyo en el data mart para almacenar una serie de datos que nos servirá a la hora de llenar la tabla dimensión microzona. Como se puede ver en la Figura 54, se agrega un registro adicional para no perder datos.



**Figura 54.** Transformación: Carga tabla apoyo predios ur Fuente: Elaboración Propia

La transformación Carga tabla apoyo microzona traslada datos desde la tabla predios\_urbanos\_rusticos de la base de datos transaccional a una tabla que servirá de apoyo para el llenado de la tabla dimensión microzona. Como podemos observar en la Figura 55, se hace añade un registro, como en el caso anterior, para no perder datos.



**Figura 55.** Transformación: Carga tabla apoyo microzona Fuente: Elaboración Propia

### Trabajo ETL parte 2

Es la segunda y última parte de nuestro proceso de carga, la cual se encarga de realizar el llenado de las tablas de dimensiones y de hechos. En la Figura 56 se puede apreciar el recorrido que lleva esta parte del proceso.



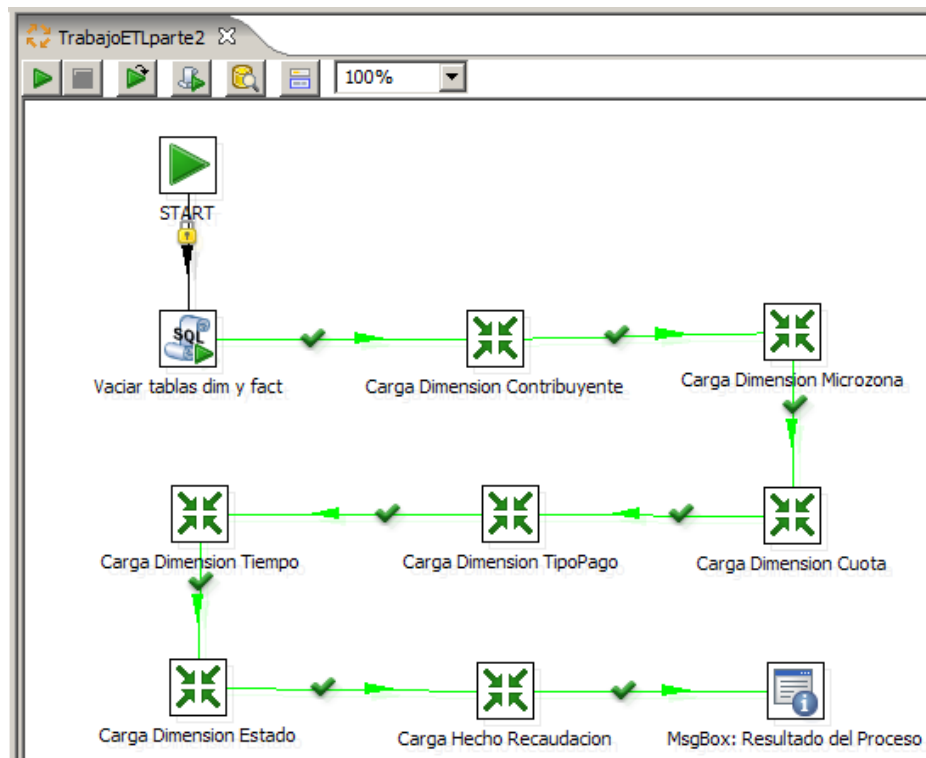


Figura 56. Trabajo: Trabajo ETL parte 2

Fuente: Elaboración Propia

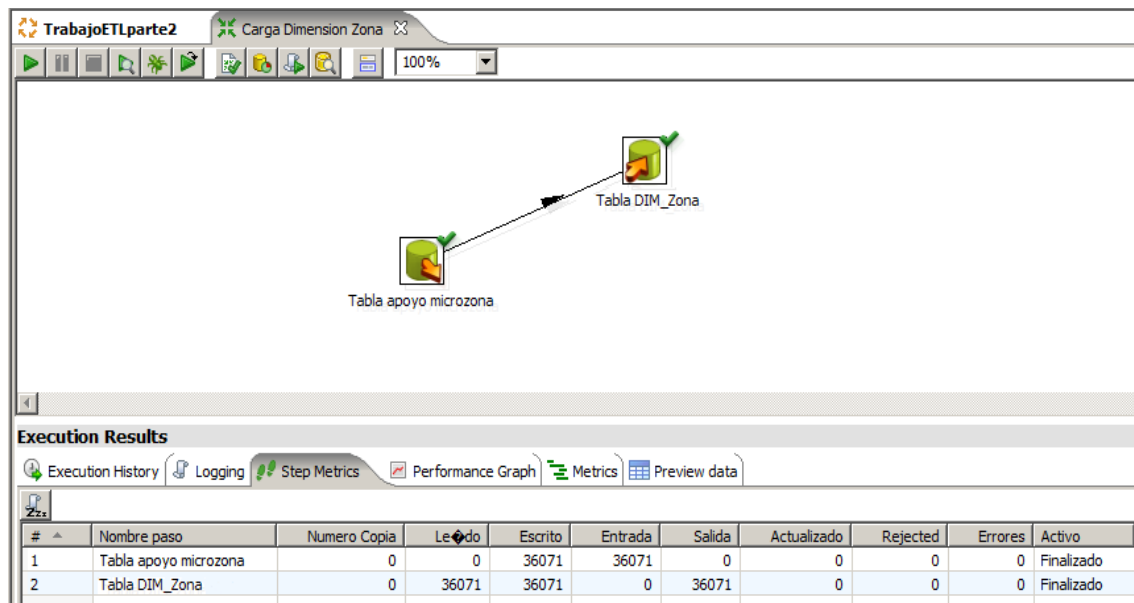
La transformación Carga Dimensión Contribuyente (Figura 57) traslada los datos de la tabla generales de la base de datos transaccional a nuestra tabla dimensional.

#	Nombre paso	Numero Copia	Leído	Escrito	Entrada	Salida	Actualizado	Rejected	Errores	Activo
1	Tabla generales	0	0	26903	26903	0	0	0	0	Finalizado
2	Tabla DIM_Contribuyente	0	26903	26903	0	26903	0	0	0	Finalizado

Figura 57. Transformación: Carga Dimensión Contribuyente

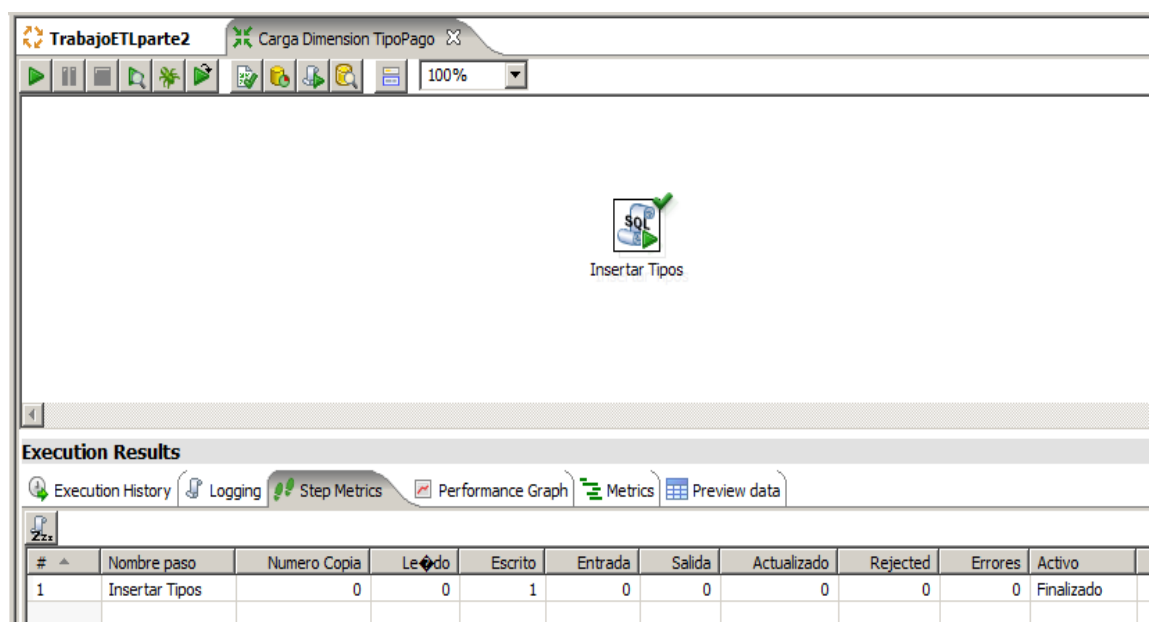
Fuente: Elaboración Propia

La transformación Carga Dimensión Zona (Figura 58) traslada los datos de las tablas de apoyo microzona y predios UR, los cuales fueron cargados previamente en nuestra base de datos dimensional.



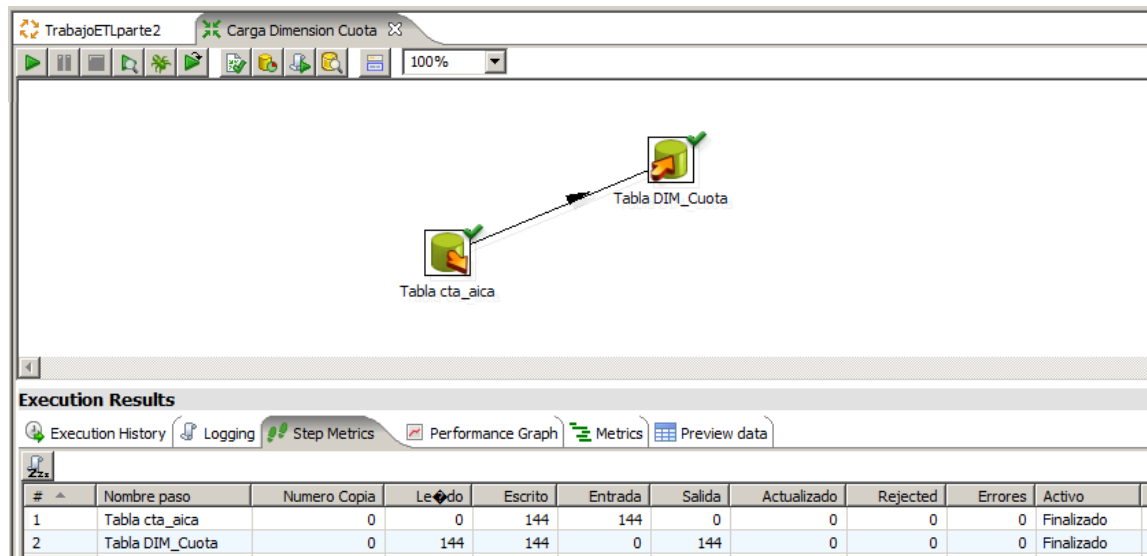
**Figura 58.** Transformación: Carga Dimensión Zona Fuente: Elaboración Propia

La transformación Carga Dimensión Tipo (Figura 59) no traslada los datos desde una tabla de la base de datos transaccional, sino que crea los registros a utilizarse, tal como se pudo apreciar en la Figura 48.



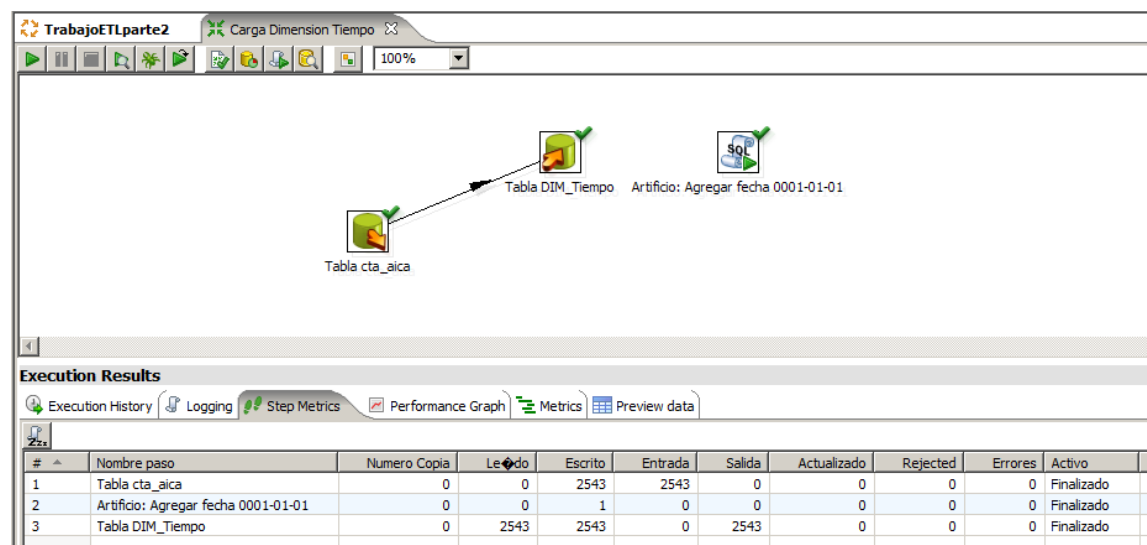
**Figura 59.** Transformación: Carga Dimensión Tipo Fuente: Elaboración Propia

La transformación Carga Dimensión Cuota (Figura 60) traslada los datos relacionados a las cuotas y sus respectivos años almacenados en la tabla cta\_aica de la base de datos transaccional hasta nuestra tabla dimensional.



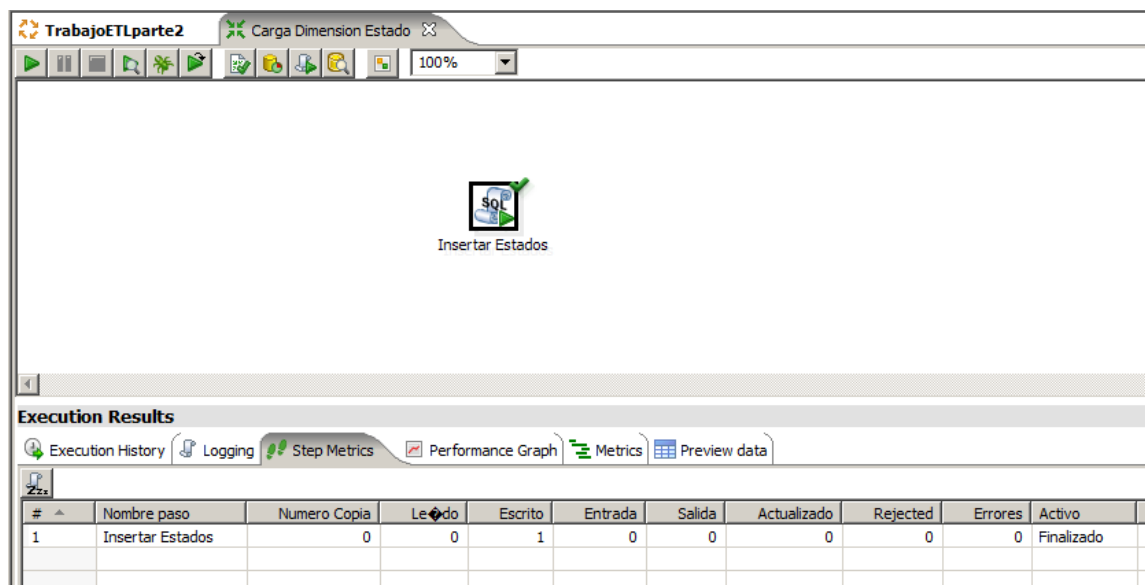
**Figura 60.** Transformación: Carga Dimensión Cuota Fuente: Elaboración Propia

La transformación Carga Dimensión Tiempo (Figura 61) traslada los datos de las fechas en las que se efectuaron recaudaciones de la tabla cta\_aica de la base de datos transaccional, hacia nuestra tabla dimensional. Además se crea un registro adicional para no perder datos.



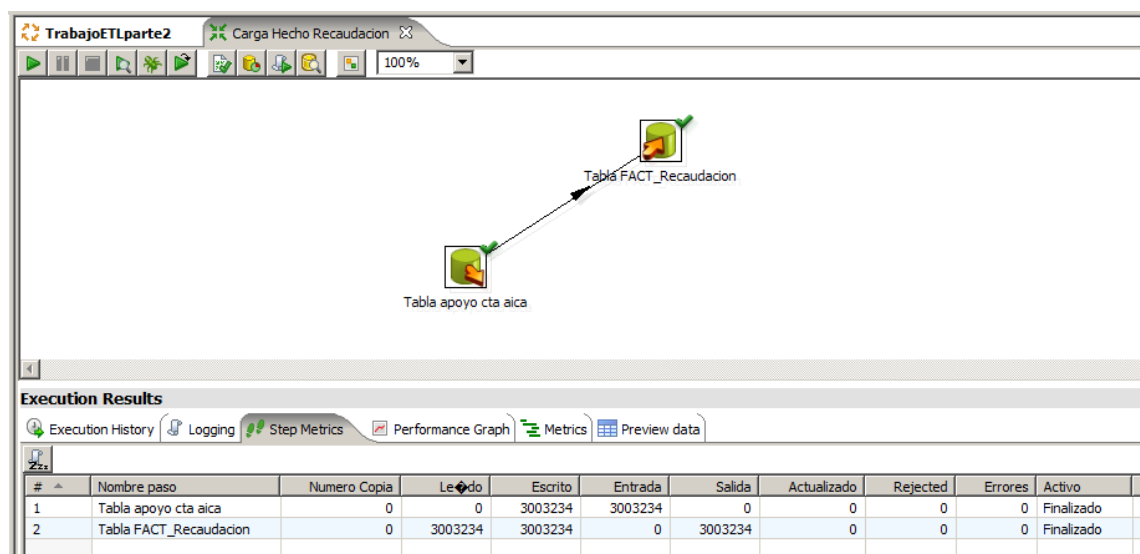
**Figura 61.** Transformación: Carga Dimensión Tiempo Fuente: Elaboración Propia

La transformación Carga Dimensión Estado (Figura 62) no traslada los datos desde una tabla de la base de datos transaccional, sino que crea los registros a utilizarse, tal como se pudo apreciar en la Figura 49.



**Figura 62.** Transformación: Carga Dimension Estado Fuente: Elaboración Propia

La transformación Carga Hecho Recaudación (Figura 63) traslada los datos de la tabla de apoyo cta\_aica y de las tablas de dimensiones hacia nuestra tabla de hechos Recaudación.

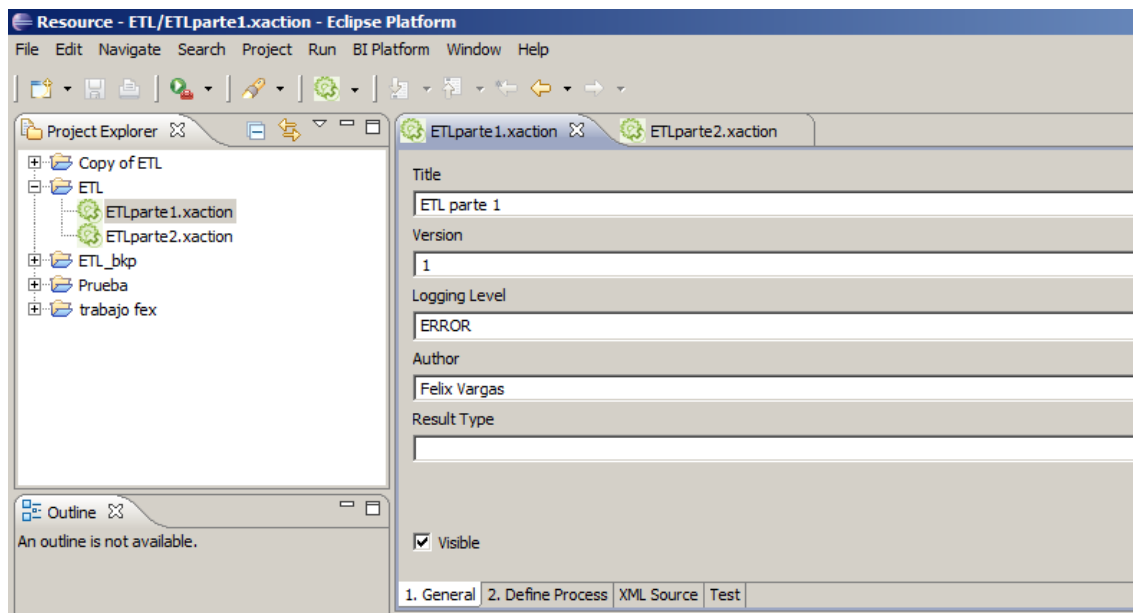


**Figura 63.** Transformación: Carga Hecho Recaudación Fuente: Elaboración Propia

### 3.9.4. Automatización del Proceso

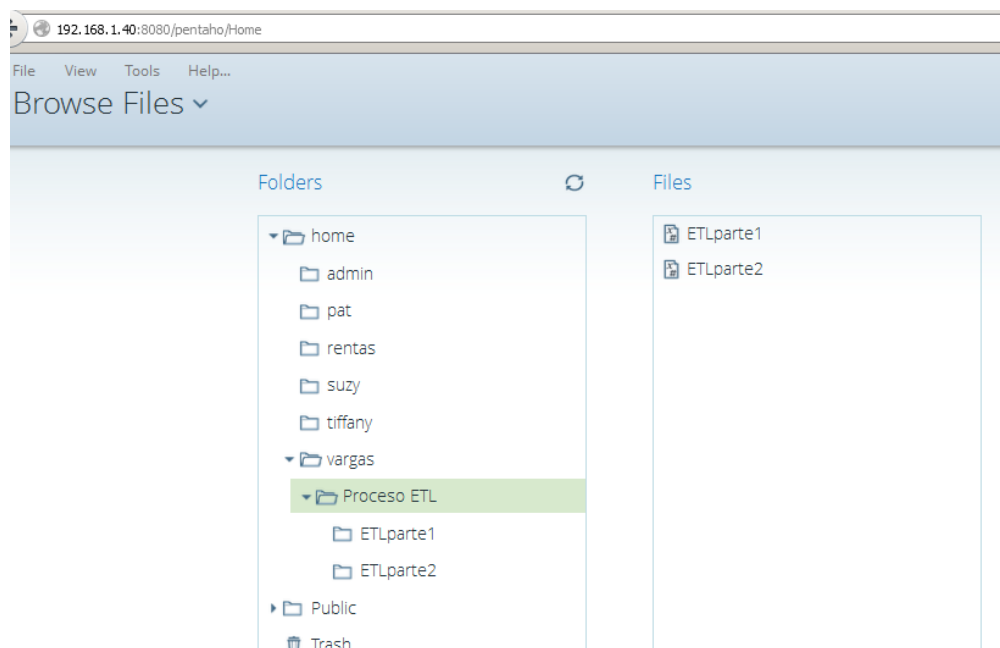
Si bien es cierto que tanto los trabajos como las transformaciones se pueden realizar de forma manual con el PDI o con archivos batch, se consideró que era mejor que la ejecución del proceso sea de forma automática. Para ello se utilizó la herramienta Pentaho Design Studio para la creación de action

sequences (Figura 64), las cuales permiten realizar determinadas acciones desde la consola de usuario de Pentaho BI server (Figura 65).



**Figura 64.** Action Sequences con Pentaho Design Studio

Fuente: Elaboración Propia



**Figura 65.** Action Sequences en la consola de usuario Pentaho BI Server

Fuente: Elaboración Propia

Incluso echo esto, para que el proceso ETL sea ejecutado se necesitaría que el mismo usuario haga la solicitud. Es por eso que los action sequences fueron programados en la consola de usuario (Figura 66). Esto con el fin de que se ejecuten automáticamente a una hora en la que el sistema no se encuentra en uso.

Schedule Name	Repeats	Source File	Output Location	Last Run	Next Run	Created By	Status
PentahoSystemVersionC	Every day at 14:43:23	PentahoSystemVersionCheck	-	2014 Nov 24 14:43:23	2014 Nov 25 14:43:23	admin	Normal
ETLparte1	Every day at 01:00:00	/home/vargas/Proceso ETL/ETLparte1	/home/vargas/Proceso ETL/Logs	-	2014 Nov 25 01:00:00	vargas	Normal
ETLparte2	Every day at 02:00:00	/home/vargas/Proceso ETL/ETLparte2	/home/vargas/Proceso ETL/Logs	-	2014 Nov 25 02:00:00	vargas	Normal

Figura 66. Programación de los action sequences

Fuente: Elaboración Propia

### 3.10. DISEÑO DEL CUBO OLAP

Para la creación del cubo OLAP se utilizó la herramienta Pentaho Schema Workbench, la cual nos permite crearla en un entorno gráfico y muy comprensible. Se creó el cubo Recaudación, y se le agregó las dimensiones y medidas tal como se puede observar en la Figura 67. Finalmente el cubo fue publicado en el servidor BI (Figura 68).

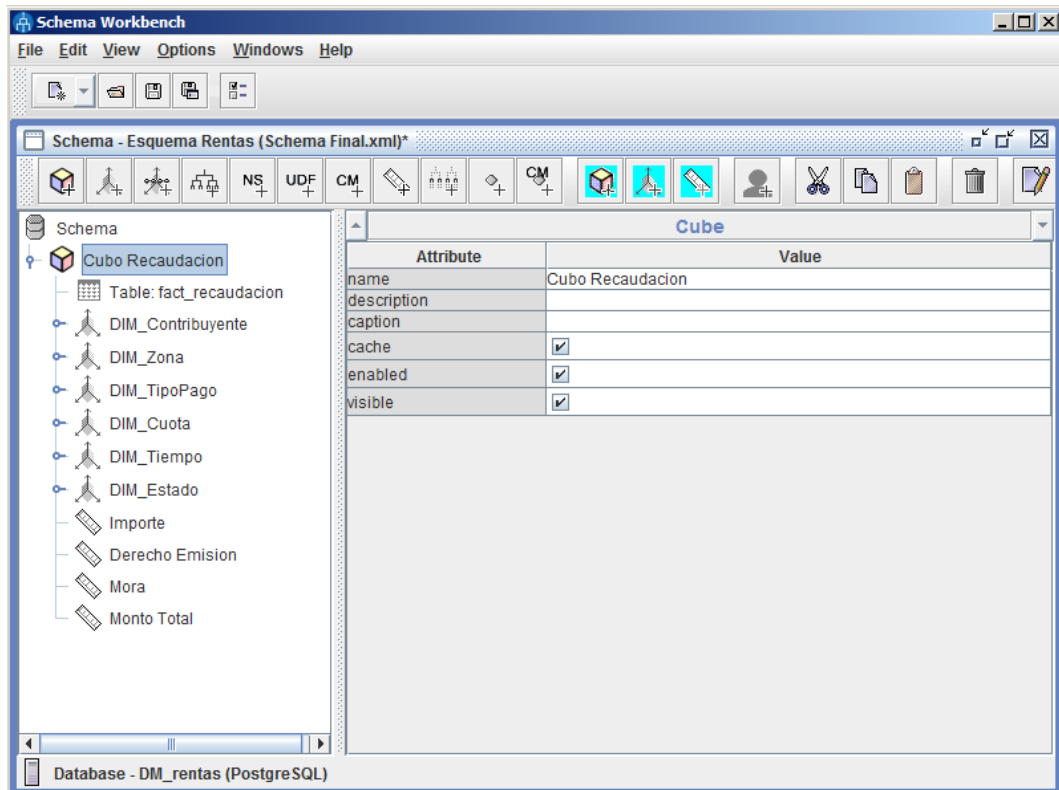


Figura 67. Creación del Cubo en Pentaho Schema Workbench

Fuente: Elaboración Propia

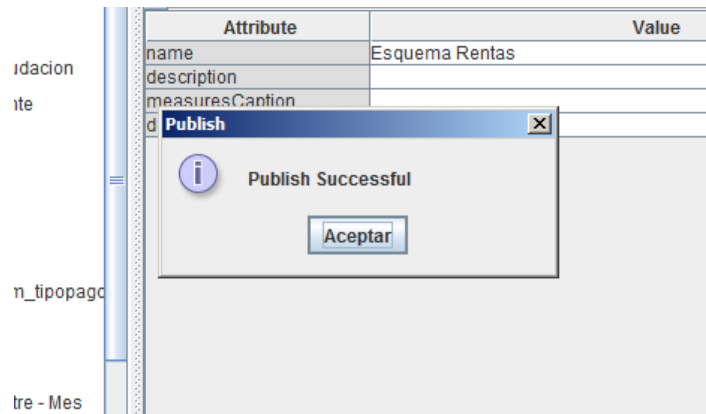


Figura 68. Publicación del cubo en Pentaho BI Server

Fuente: Elaboración Propia

### 3.11.ESPECIFICACIÓN DE APLICACIONES BI

#### 3.11.1. Lista de Reportes

Para la especificación de los reportes se desarrollaron prototipos para poder darnos una idea de cómo serían los reportes una vez culminados, es decir qué información contendrían y cuál sería su distribución. Los prototipos de los reportes fueron:

- Reporte Recaudación / Deuda de impuesto predial por cuotas: Este prototipo sirve tanto para las recaudaciones como para las deudas del impuesto predial ya que la plantilla es la misma para ambos casos y lo que cambia es el título y los valores de las cifras. El usuario podrá apreciar la recaudación o la deuda (según sea el caso) del importe, derecho de emisión, mora y monto total del año fiscal que se elija en el combo. Las cifras estarán agrupadas en las 4 cuotas del impuesto predial. (Figura 69)

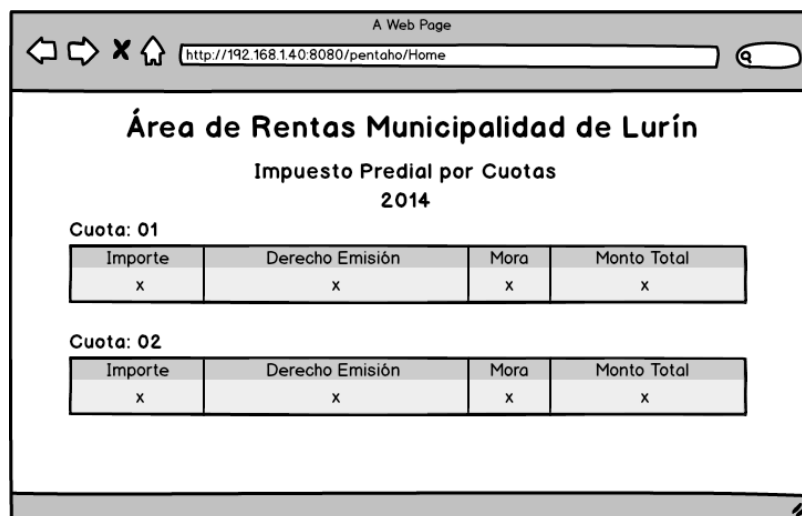


Figura 69. Prototipo: Reporte Impuesto Predial por Cuotas

Fuente: Elaboración Propia

- Reporte Recaudación / Deuda de arbitrios municipales por cuotas: Este prototipo es utilizado tanto para las recaudaciones como para las deudas de los arbitrios municipales ya que la plantilla es la misma para ambos casos. El usuario podrá apreciar la recaudación o la deuda (según el caso) del importe, derecho de emisión, mora y monto total del año fiscal que se elija en el combo. Las cifras estarán agrupadas en las 12 cuotas de los arbitrios municipales. (Figura 70)

**Área de Rentas Municipalidad de Lurín**  
Arbitrios Municipales por Cuotas  
2014

**Cuota: 01**

Importe	Derecho Emisión	Mora	Monto Total
x	x	x	x

**Cuota: 02**

Importe	Derecho Emisión	Mora	Monto Total
x	x	x	x

**Figura 70.** Prototipo: Reporte Arbitrios Municipales por Cuotas Fuente: Elaboración Propia

- Reporte Recaudación / Deuda de arbitrios municipales por zonas: Este prototipo sirve tanto para las recaudaciones como para las deudas de los arbitrios municipales ya que la plantilla es la misma para ambos casos y lo que cambia es el título y los valores de las cifras. El usuario podrá apreciar el número de la cuota, la recaudación o la deuda (según sea el caso) del importe, derecho de emisión, mora y monto total del año fiscal que se elija en el combo. Las cifras estarán agrupadas en las 5 zonas del distrito de Lurín. (Figura 71)

**Área de Rentas Municipalidad de Lurín**  
Arbitrios Municipales por Años y Zonas  
Año: 2014

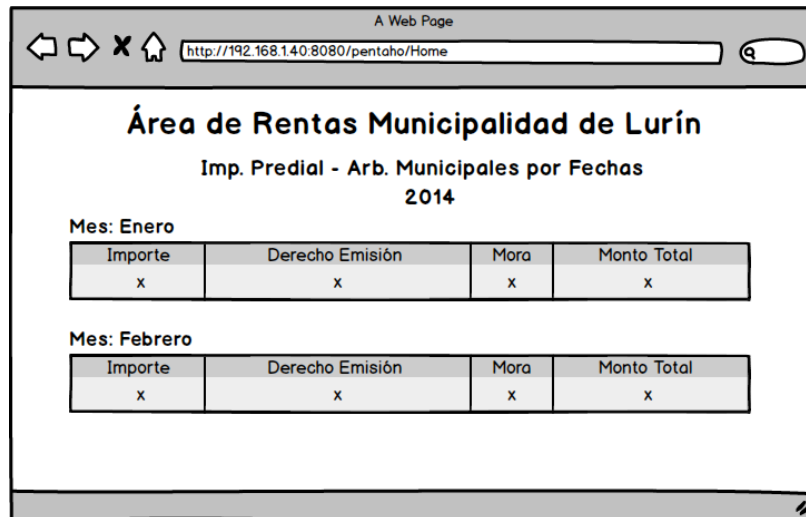
**Zona: ZONA A**

Cuota	Importe	Derecho Emisión	Mora	Monto Total
01	x	x	x	x
02	x	x	x	x
03	x	x	x	x
04	x	x	x	x
05	x	x	x	x
06	x	x	x	x

**Figura 71.** Prototipo: Reporte Arbitrios Municipales por Zonas Fuente: Elaboración Propia

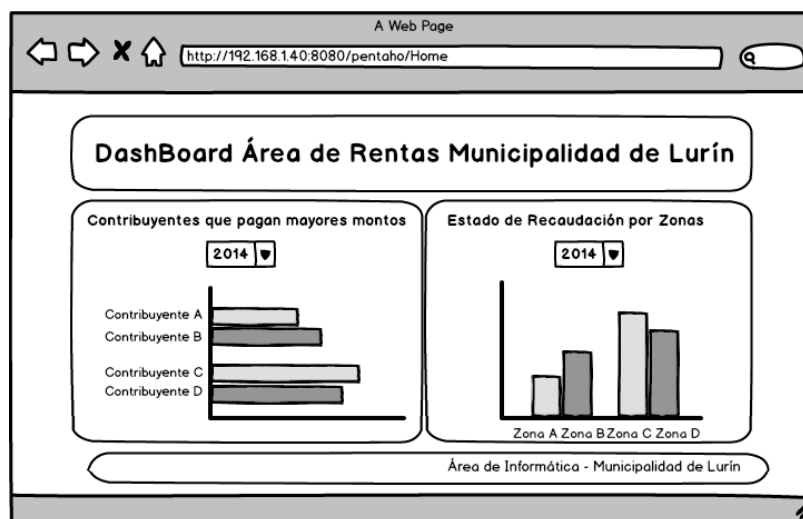


- Reporte Recaudación Imp. Predial y Arb. Municipales por fechas: Este prototipo está referido a las recaudaciones realizadas según el año y el tipo de pago (impuesto predial o arbitrios municipales) que el usuario elija en los combos. Se podrá apreciar el importe, derecho de emisión, mora y monto total. Las cifras estarán agrupadas en los 12 meses del año. (Figura 72)



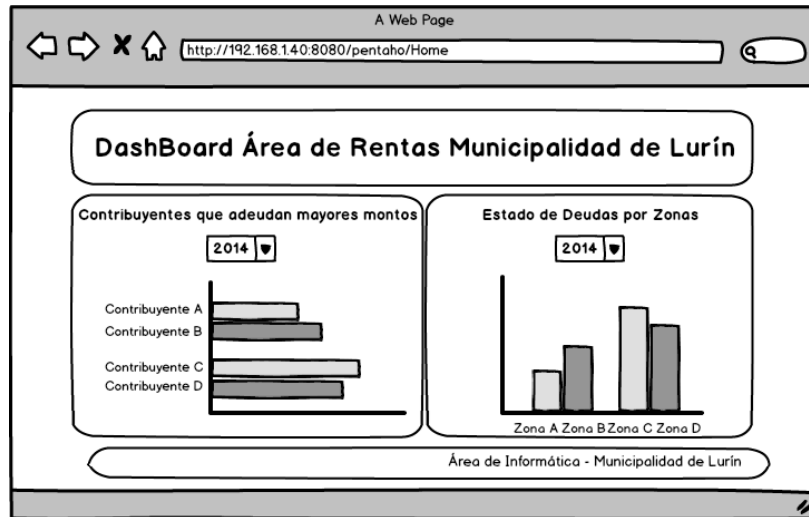
**Figura 72.** Prototipo: Reporte Impuesto Predial – Arbitrios Municipales por Fechas Fuente: Elaboración Propia

- Dashboard de contribuyentes que pagan mayores montos y estado de la recaudación por zonas: Este prototipo de dashboard consta de dos partes. En la parte izquierda el usuario podrá ver un gráfico de barras de los contribuyentes que han pagado las mayores cantidades de dinero (de impuesto predial y arbitrios municipales) en el año fiscal que se elija. Y en la parte derecha se podrá ver el estado de las recaudaciones de impuesto predial y arbitrios municipales de cada una de las zonas del distrito, según el año que se escoja en el combo. (Figura 73)



**Figura 73.** Prototipo: Dashboard Top Contribuyentes y Recaudación por Zonas Fuente: Elaboración Propia

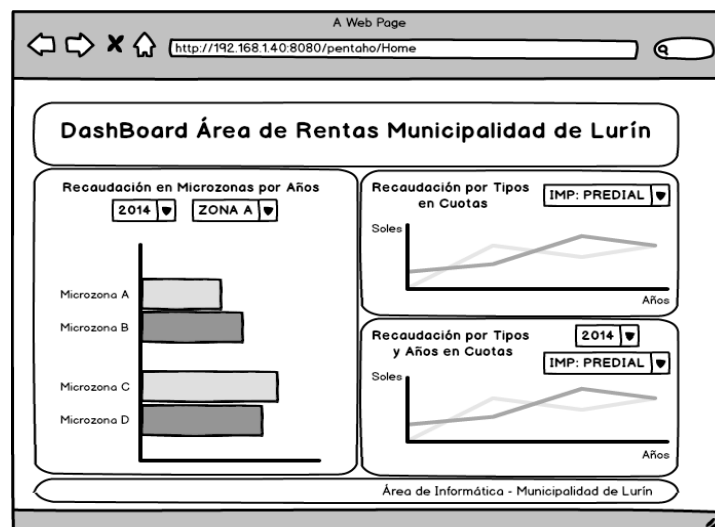
- Dashboard de contribuyentes que adeudan mayores montos y estado de las deudas por zonas: En la parte izquierda el usuario podrá ver un gráfico de barras de los contribuyentes que se encuentran debiendo las mayores cantidades de dinero en el año fiscal que se elija. Y en la parte derecha se podrá ver el estado de las deudas de cada una de las zonas del distrito, según el año que se escoja en el combo. (Figura 74)



**Figura 74.** Prototipo: Dashboard No Top Contribuyentes y Deudas por Zonas

Fuente: Elaboración Propia

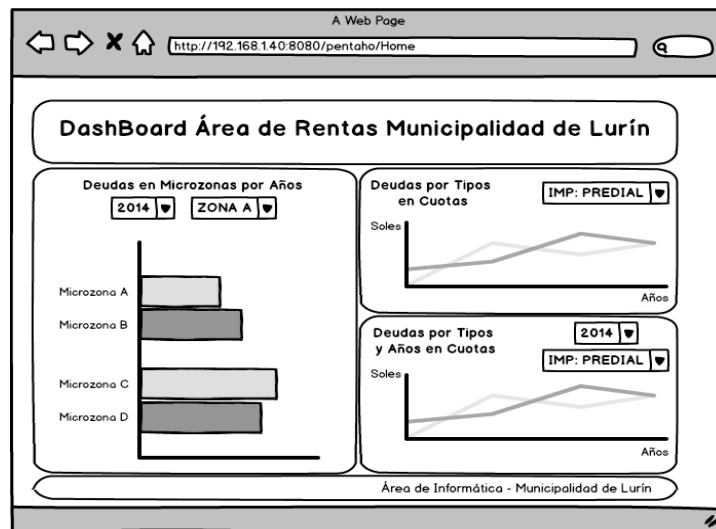
- Dashboard de recaudación por microzonas, por años y por cuotas: En la parte izquierda el usuario podrá ver un gráfico de barras de los montos recaudados en las microzonas pertenecientes a la zona escogida y al año fiscal seleccionado. En la parte derecha superior se podrá ver cómo ha ido la recaudación a lo largo de los años fiscales, según el tipo de pago que se escoja. Y en la parte derecha inferior se verá cómo ha ido la recaudación en el transcurso de las cuotas de un año fiscal escogido, y según un determinado tipo de pago. (Figura 75)



**Figura 75.** Prototipo: Dashboard Recaudación por Microzonas, por Años y por Cuotas

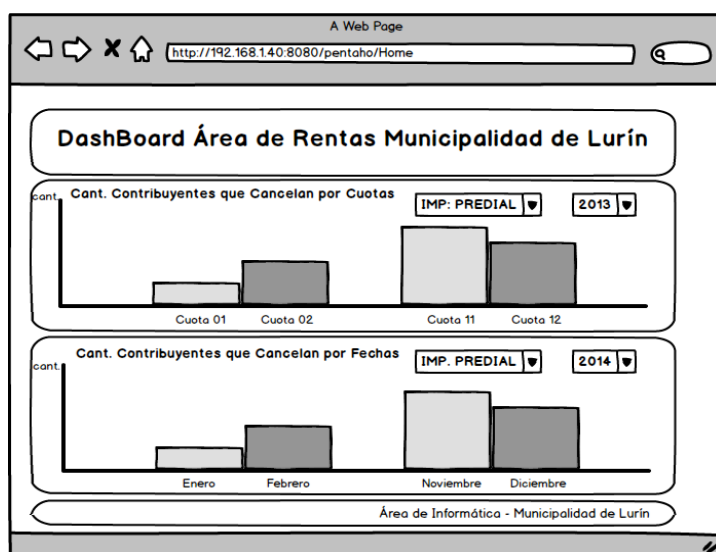
Fuente: Elaboración Propia

- Dashboard de las deudas por microzonas, por años y por cuotas: En la parte izquierda el usuario podrá ver un gráfico de barras de los montos que deben en las microzonas pertenecientes a la zona escogida y al año fiscal seleccionado. En la parte derecha superior se podrá ver cómo han ido las deudas a lo largo de los años fiscales, según el tipo de pago que se escoja. Y en la parte derecha inferior se verá cómo han ido las deudas en el transcurso de las cuotas de un año fiscal escogido, y según un determinado tipo de pago. (Figura 76)



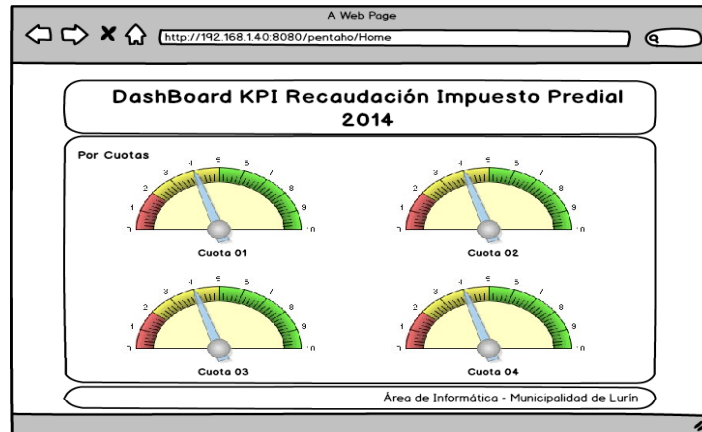
**Figura 76.** Prototipo: Dashboard Deudas por Microzonas, por Años y por Cuotas Fuente: Elaboración Propia

- Dashboard cantidad de contribuyentes que realizan pagos: En la parte superior se podrá visualizar la cantidad de contribuyentes que cancelan las cuotas de un determinado año fiscal elegido y según un tipo de pago. Mientras que en la parte inferior el usuario verá también la cantidad de contribuyentes que cancelan pero por cada mes. (Figura 77)



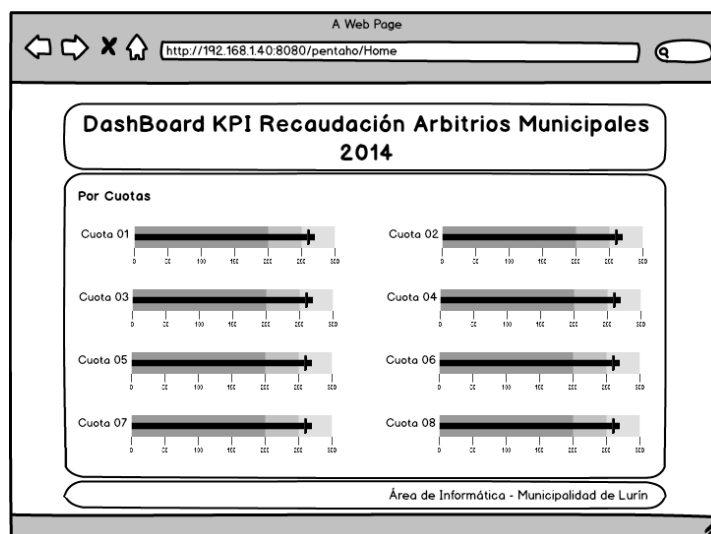
**Figura 77.** Prototipo: Dashboard Cantidad Contribuyentes que Pagan Fuente: Elaboración Propia

- Dashboard de KPI de impuesto predial: Este prototipo de dashboard está referido a los indicadores clave de desempeño (KPI) de la recaudación del impuesto predial de un año fiscal. La información se mostrará mediante velocímetros, los mismos que indicarán el estado de la recaudación distinguiéndola por colores (rojo mal, amarillo regular, verde óptimo). Debido a que el impuesto predial consta de 4 cuotas, habrán 4 velocímetros, es decir uno para cada cuota. (Figura 78)



**Figura 78.** Prototipo: Dashboard KPI de Impuesto Predial Fuente: Elaboración Propia

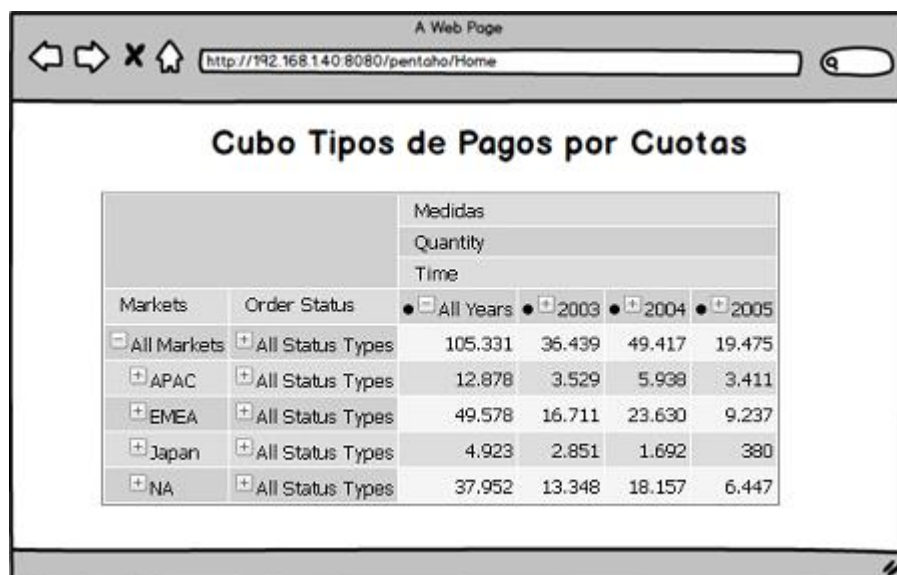
- Dashboard de KPI de arbitrios municipales: Este prototipo de dashboard está referido a los indicadores clave de desempeño (KPI) de la recaudación de los arbitrios municipales de un año fiscal. La información se mostrará mediante bullet charts, los mismos que indicarán el estado de la recaudación distinguiéndola por la tonalidad de los colores (gris oscuro mal, gris claro regular, blanco óptimo). Los bullet chart no ocupan tanto espacio como los velocímetros, por lo que habrá uno por cada una de las 12 cuotas de los arbitrios. (Figura 79)



**Figura 79.** Prototipo: Dashboard KPI de Arbitrios Municipales

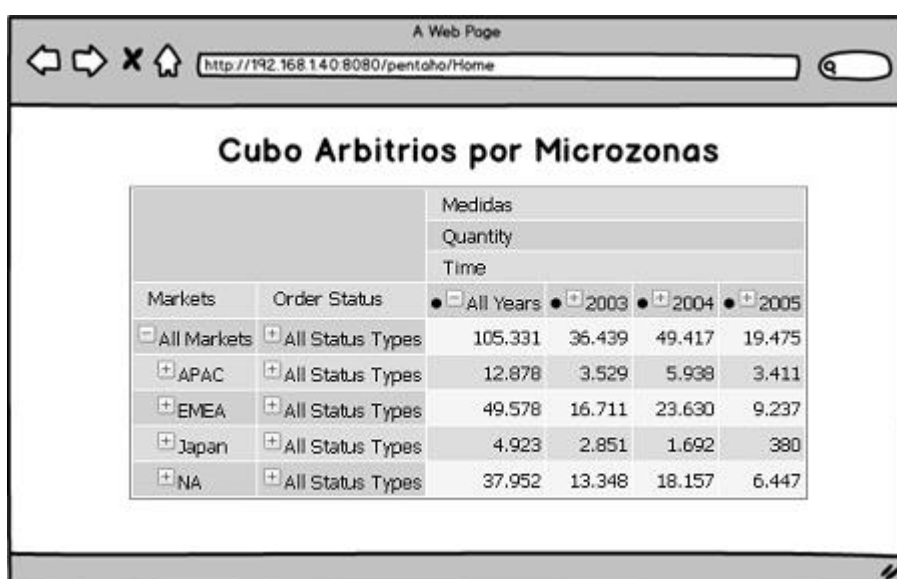
Fuente: Elaboración Propia

- Cubo Recaudación / Deudas de tipo de pagos por cuotas: Este prototipo sirve a la vez para el visor jpivot de recaudación y de deuda, los cuales tendrán en las columnas los tipos de pago (impuesto predial y arbitrios municipales), en las filas los años fiscales con sus cuotas y en las medidas estarán el importe, derecho de emisión, mora y monto total. (Figura 80)



**Figura 80.** Prototipo: Cubo de Tipos de Pagos por Cuotas Fuente: Elaboración Propia

- Cubo Recaudación / Deudas de arbitrios municipales por zonas: Este prototipo sirve a la vez para el visor jpivot de recaudación y de deuda, los cuales tendrán en las columnas los tipos de pago y el año fiscal (con sus cuotas), en las filas las zonas con sus microzonas y en las medidas estarán el importe, derecho de emisión, mora y monto total. (Figura 81)



**Figura 81.** Prototipo: Cubo de Arbitrios Municipales por Zonas Fuente: Elaboración Propia

- Cubo Recaudación de tipo de pagos por fechas: Este prototipo está referido al visor jpivot de recaudación, el cual tendrá en el lado de las columnas los tipos de pago (impuesto predial y arbitrios municipales), mientras que en el lado de las filas tendrá los años con sus respectivos trimestres (y éstos últimos vendrán con sus respectivos meses). Finalmente en el lado de las medidas estarán el importe, derecho de emisión, mora y monto total. (Figura 82)

		Medidas			
		Quantity			
		Time			
Markets	Order Status	<input type="radio"/> All Years	<input type="radio"/> 2003	<input type="radio"/> 2004	<input type="radio"/> 2005
<input type="checkbox"/> All Markets	<input type="checkbox"/> All Status Types	105.331	36.439	49.417	19.475
<input type="checkbox"/> APAC	<input type="checkbox"/> All Status Types	12.878	3.529	5.938	3.411
<input type="checkbox"/> EMEA	<input type="checkbox"/> All Status Types	49.578	16.711	23.630	9.237
<input type="checkbox"/> Japan	<input type="checkbox"/> All Status Types	4.923	2.851	1.692	380
<input type="checkbox"/> NA	<input type="checkbox"/> All Status Types	37.952	13.348	18.157	6.447

**Figura 82.** Prototipo: Cubo de Tipos de Pagos por Fechas

Fuente: Elaboración Propia


### 3.12. DESARROLLO DE APLICACIONES BI

#### 3.12.1. Elaboración de Reportes

Para la elaboración de los reportes a medida se utilizó la herramienta Pentaho Report Designer, mientras que para los dashboards el Community Dashboard Editor (plugin de Pentaho BI Server) y finalmente para los cubos se empleó el visor del JPivot (otro plugin de Pentaho BI Server). Todos fueron elaborados de acuerdo a los prototipos mostrados en el punto anterior.

Cabe recordar que en algunos prototipos se indicó que servirían tanto para recaudación como para deudas, por lo tanto, a partir de un solo prototipo salieron 2 reportes.


Desde la Figura 83 a la Figura 101 se muestran imágenes reales de los reportes ya en funcionamiento, con valores de los combos seleccionados aleatoriamente.

 <b>Área de Rentas Municipalidad de Lurín</b>			
<b>Recaudación Impuesto Predial por Años y Cuotas</b>			
<b>Año: 2014</b>			
<b>Cuota: 01</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
2.314.319,64	58.797	701,82	2.373.818,46
<b>Cuota: 02</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
2.146.080,85	0	0,75	2.146.081,6
<b>Cuota: 03</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
2.338.352,33	0	0,34	2.338.352,67
<b>Cuota: 04</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
2.184.828,42	0	0	2.184.828,42

Área de Informática - Municipalidad de Lurín

**Figura 83.** Reporte Recaudación Impuesto Predial por Cuotas

Fuente: Elaboración Propia


 <b>Área de Rentas Municipalidad de Lurín</b>			
<b>Deuda Impuesto Predial por Años y Cuotas</b>			
<b>Año: 2014</b>			
<b>Cuota: 01</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
973.951,65	94.446	0	1.068.397,65
<b>Cuota: 02</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
1.141.909,84	0	0	1.141.906,84
<b>Cuota: 03</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
949.672,44	0	0	949.672,44
<b>Cuota: 04</b>			
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
1.103.203,25	0	0	1.103.203,25

Área de Informática - Municipalidad de Lurín

**Figura 84.** Reporte Deudas Impuesto Predial por Cuotas


Fuente: Elaboración Propia



 <b>Área de Rentas Municipalidad de Lurín</b>				
<b>Recaudación Arbitrios Municipales por Años y Cuotas</b>				
<b>Año: 2014</b>				
<b>Cuota: 01</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
448.215,5	48.135	322,15	496.672,65	
<b>Cuota: 02</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
456.632,42	0	16,69	456.649,11	
<b>Cuota: 03</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
439.924,9	0	45,46	439.970,36	
<b>Cuota: 04</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
394.978,52	0	0	394.978,52	
<b>Cuota: 05</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
448.172,44	0	0	448.172,44	
<b>Cuota: 06</b>				
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>	
445.827,54	0	0	445.827,54	
Área de Informática - Municipalidad de Lurín				


**Figura 85.** Reporte Recaudación Arbitrios Municipales por Cuotas

Fuente: Elaboración Propia

 <b>Área de Rentas Municipalidad de Lurín</b>			
<b>Deuda Arbitrios Municipales por Años y Cuotas</b>			
<b>Año: 2014</b>			
<b>Cuota: 01</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
329.112,02	102.957	0	432.069,02
<b>Cuota: 02</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
320.677,56	0	0	320.677,56
<b>Cuota: 03</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
337.130,93	0	0	337.130,93
<b>Cuota: 04</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
382.021,32	0	0	382.021,32
<b>Cuota: 05</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
328.785,47	0	0	328.785,47
<b>Cuota: 06</b>			
<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
331.121,2	0	0	331.121,2
Área de Informática - Municipalidad de Lurín			

**Figura 86.** Reporte Deudas Arbitrios Municipales por Cuotas


Fuente: Elaboración Propia

 <b>Área de Rentas Municipalidad de Lurín</b>				
<b>Recaudación Arbitrios Municipales por Zonas</b> <b>Año: 2014</b>				
Zona: ZONA A				
Cuota	Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)
01	173.928,47	19.485	208,04	193.621,51
02	178.752,64	0	9,94	178.762,58
03	166.604,36	0	20,98	166.625,34
04	156.537,51	0	0	156.537,51
05	168.085,3	0	0	168.085,3
06	166.808,82	0	0	166.808,82
07	167.553,12	0	0	167.553,12
08	166.461,07	0	0	166.461,07
09	159.112,89	0	1,63	159.114,52
10	157.305,95	0	0	157.305,95
11	156.823,93	0	0	156.823,93
12	146.104,2	0	0	146.104,2

Área de Informática - Municipalidad de Lurín

**Figura 87.** Reporte Recaudación Arbitrios Municipales por Zonas


Fuente: Elaboración Propia

 <b>Área de Rentas Municipalidad de Lurín</b>				
<b>Deuda Arbitrios Municipales por Zonas</b>				
<b>Año: 2014</b>				
<b>Zona: ZONA A</b>				
<b>Cuota</b>	<b>Importe (S/.)</b>	<b>Derecho de Emision (S/.)</b>	<b>Mora (S/.)</b>	<b>Monto Total (S/.)</b>
01	108.735,41	34.404	0	143.139,41
02	103.904,09	0	0	103.904,09
03	116.007,15	0	0	116.007,15
04	126.066,37	0	0	126.066,37
05	114.499,4	0	0	114.499,4
06	115.776,91	0	0	115.776,91
07	115.030,36	0	0	115.030,36
08	116.119,77	0	0	116.119,77
09	123.465,86	0	0	123.465,86
10	125.255,96	0	0	125.255,96
11	125.737,98	0	0	125.737,98
12	136.457,71	0	0	136.457,71

Área de Informática - Municipalidad de Lurín

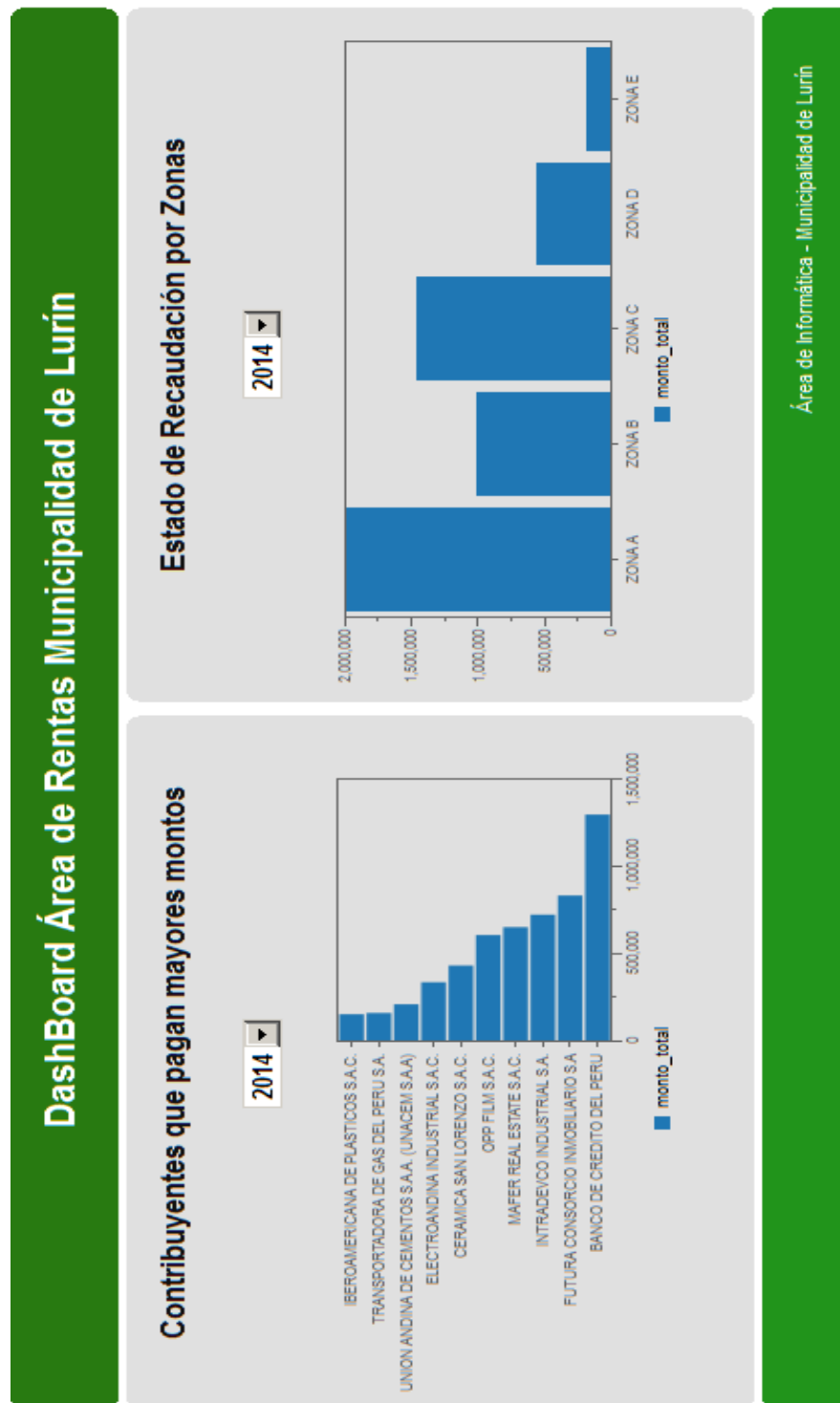
**Figura 88.** Reporte Deudas Arbitrios Municipales por Zonas

Fuente: Elaboración Propia

 <b>Área de Rentas Municipalidad de Lurín</b>				
<b>Recaudación ARB. MUNICIPALES</b> <b>por Fechas</b> <b>Año: 2014</b>				
<b>Mes: 01-Enero</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
135.713,61	3.619	4.226,8	143.559,41	
<b>Mes: 02-Febrero</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
1.000.438,68	20.041,5	5.314,85	1.025.795,03	
<b>Mes: 03-Marzo</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
1.428.143,41	16.903,5	6.003,03	1.451.049,94	
<b>Mes: 04-Abril</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
379.836,58	5.548,5	7.334,79	392.719,87	
<b>Mes: 05-Mayo</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
461.901,56	2.394	0	464.295,56	
<b>Mes: 06-Junio</b>				
Importe (S/.)	Derecho de Emision (S/.)	Mora (S/.)	Monto Total (S/.)	
161.449,5	2.388	0	163.837,5	
Área de Informática - Municipalidad de Lurín				

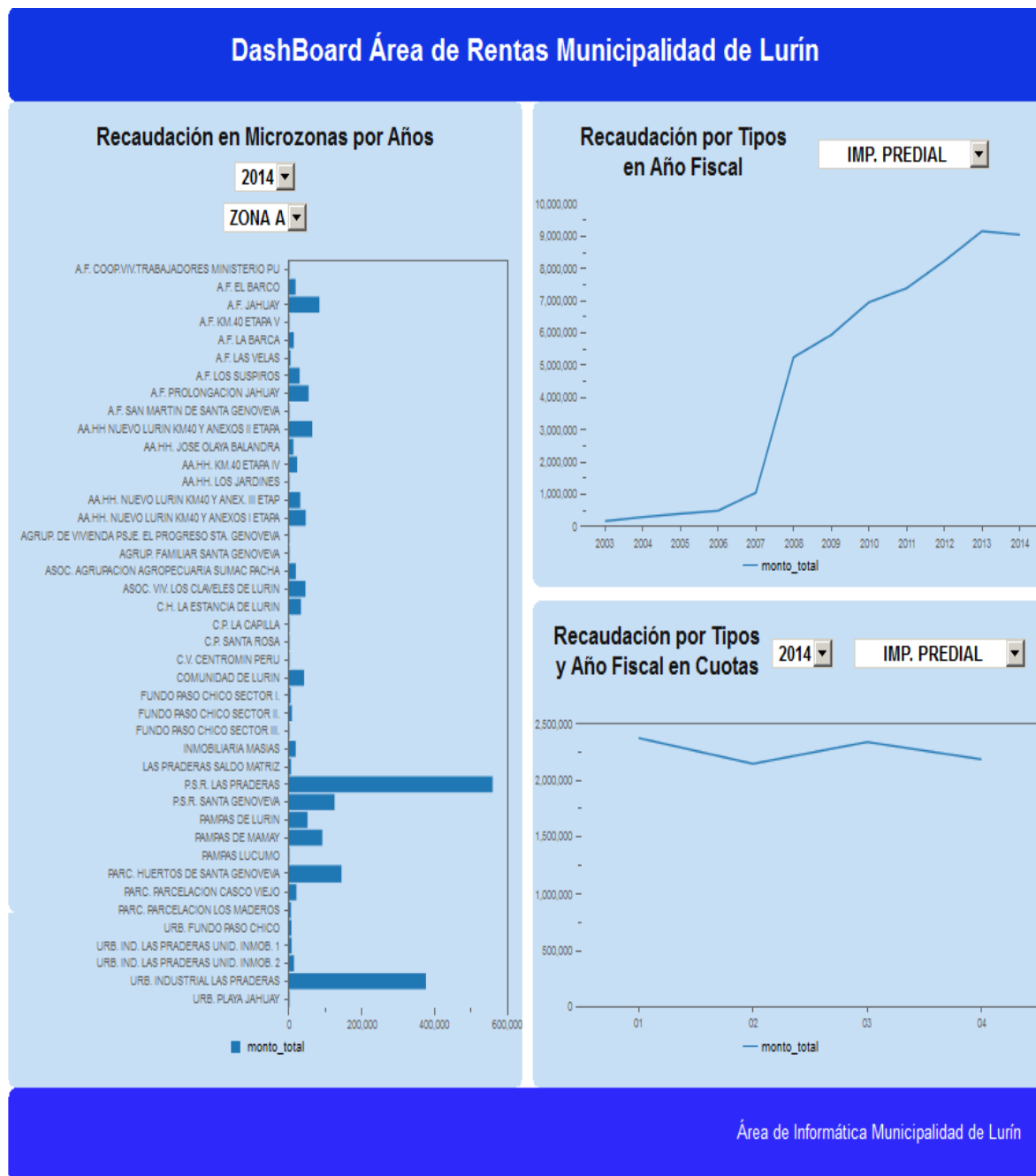
**Figura 89.** Reporte Impuesto Predial – Arbitrios.  
Municipales por Fechas

Fuente: Elaboración Propia



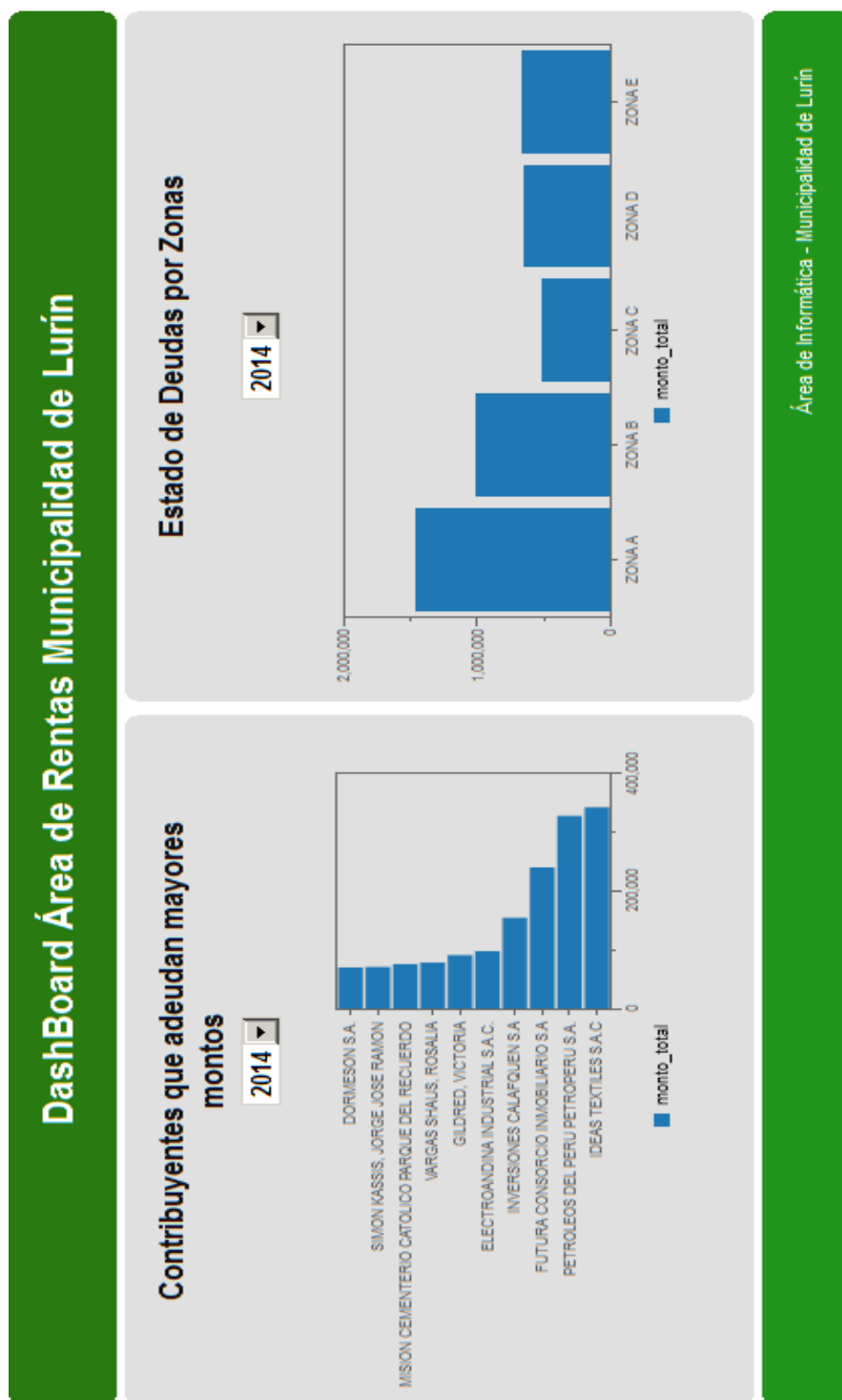
**Figura 90.** Dashboard Top Contribuyentes y Recaudación por Zonas

Fuente: Elaboración Propia



**Figura 91.** Dashboard Recaudación por Microzonas, por Años y por Cuotas

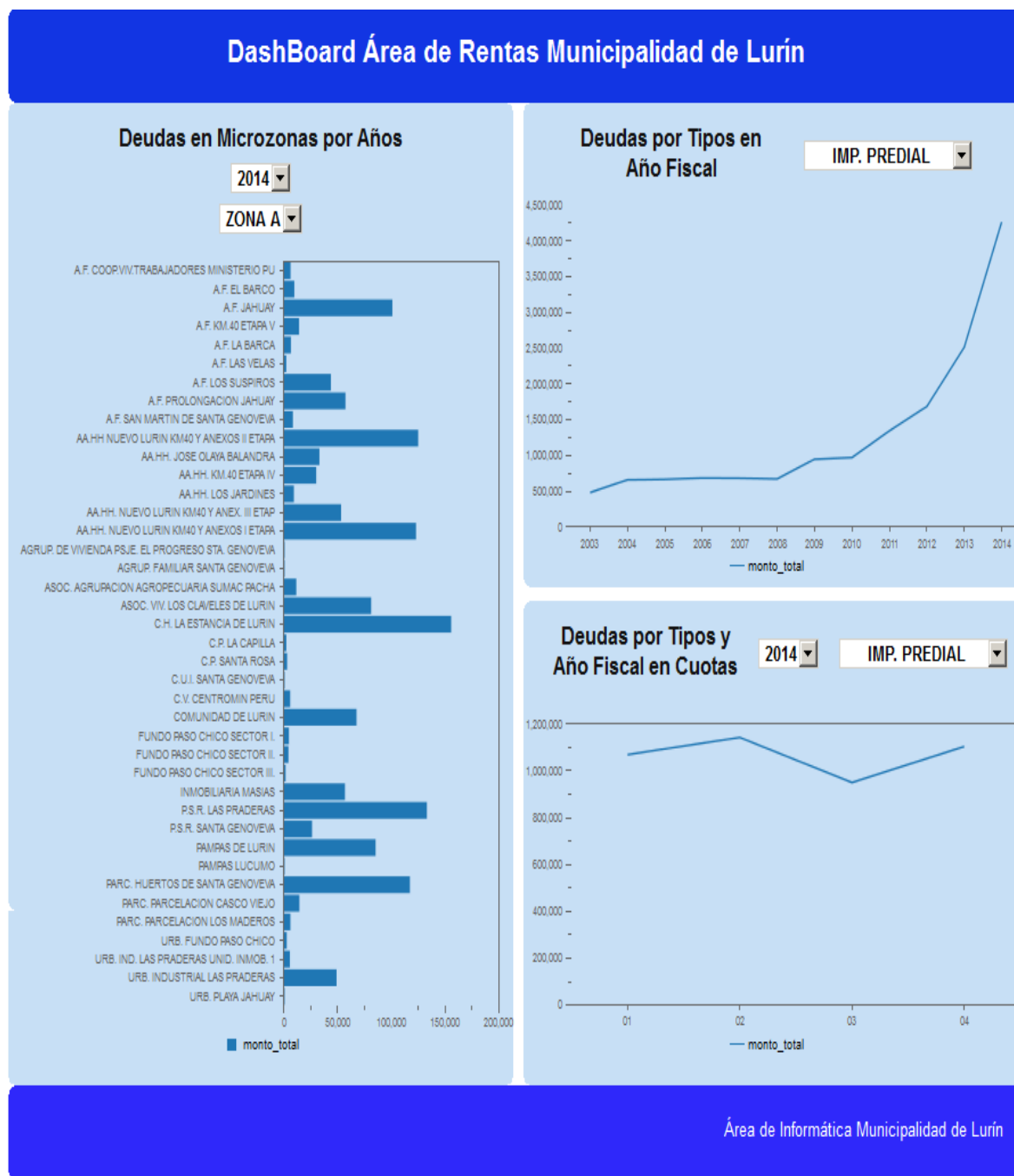
Fuente: Elaboración Propia



**Figura 92.** Dashboard No Top Contribuyentes y Deudas por Zonas

Fuente: Elaboración Propia





**Figura 93.** Dashboard Deudas por Microzonas, por Años y por Cuotas

Fuente: Elaboración Propia

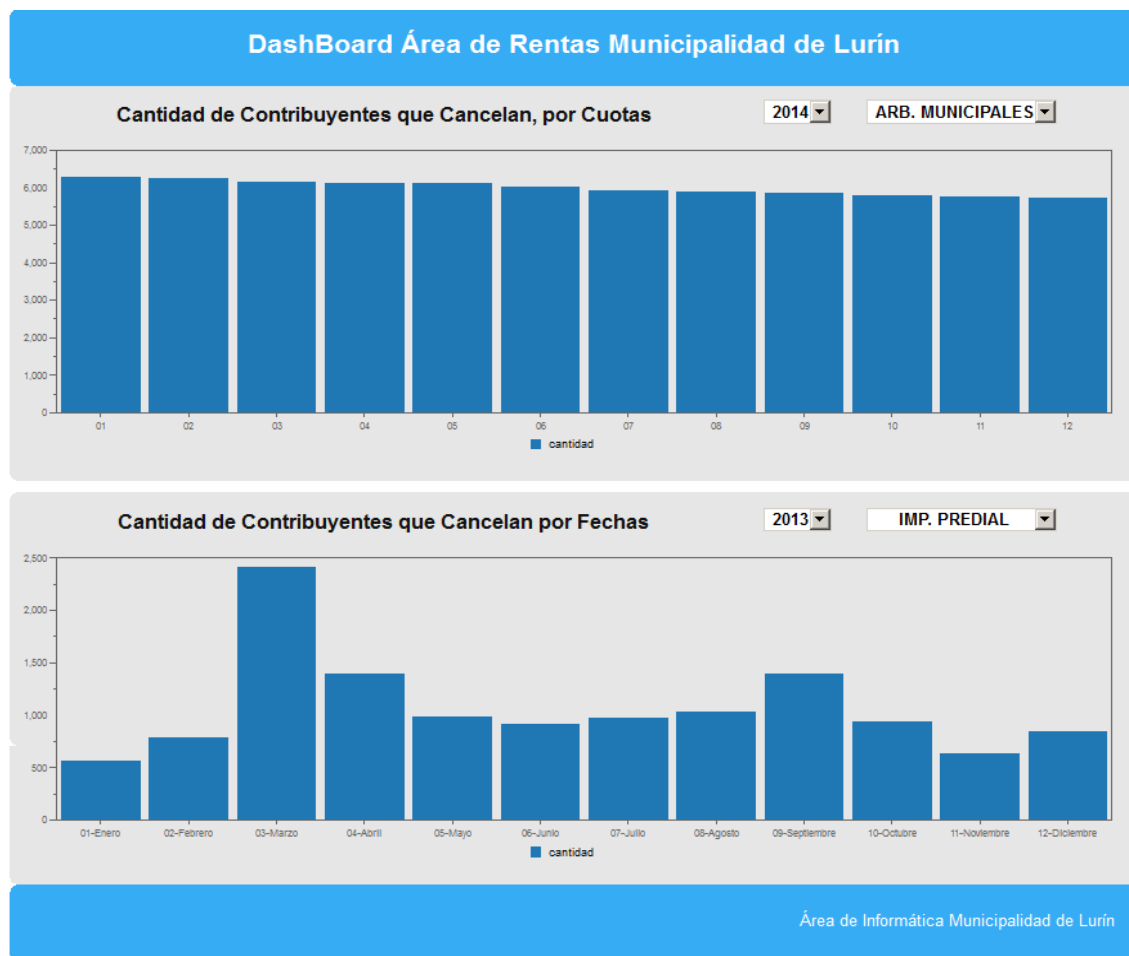


Figura 94. Dashboard Cantidad Contribuyentes que Pagan Fuente: Elaboración Propia

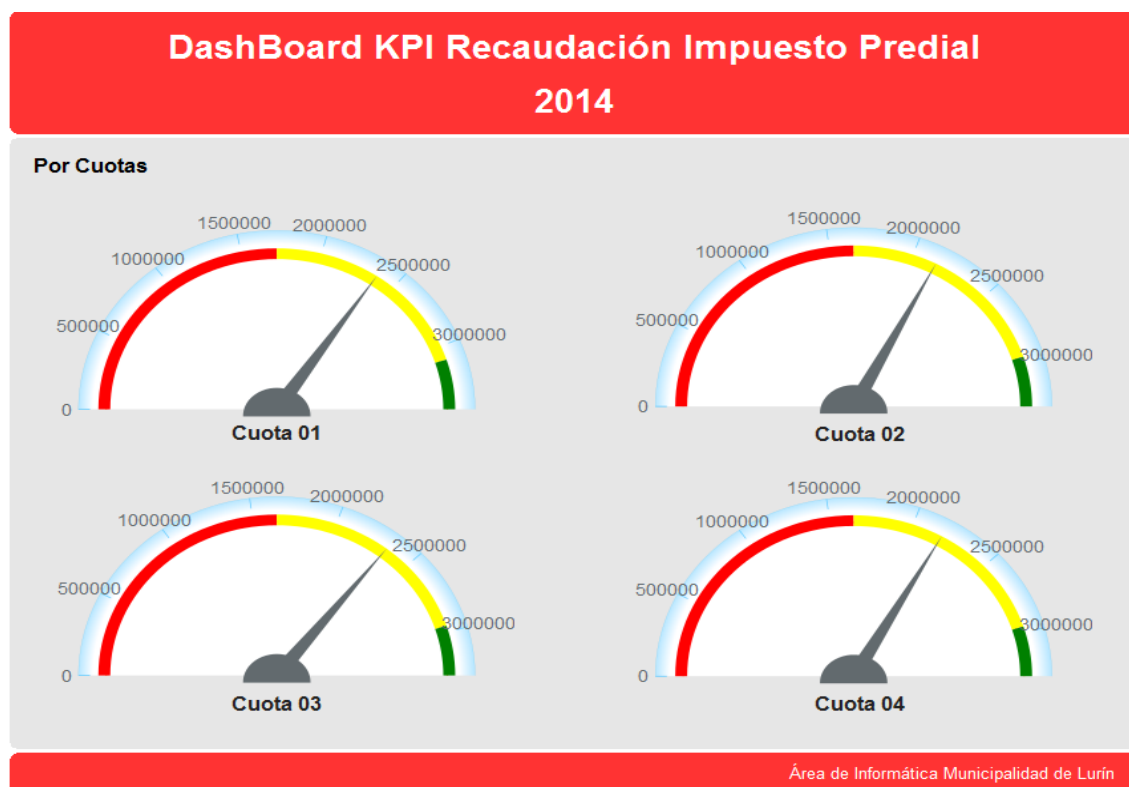


Figura 95. Dashboard KPI de Impuesto Predial Fuente: Elaboración Propia

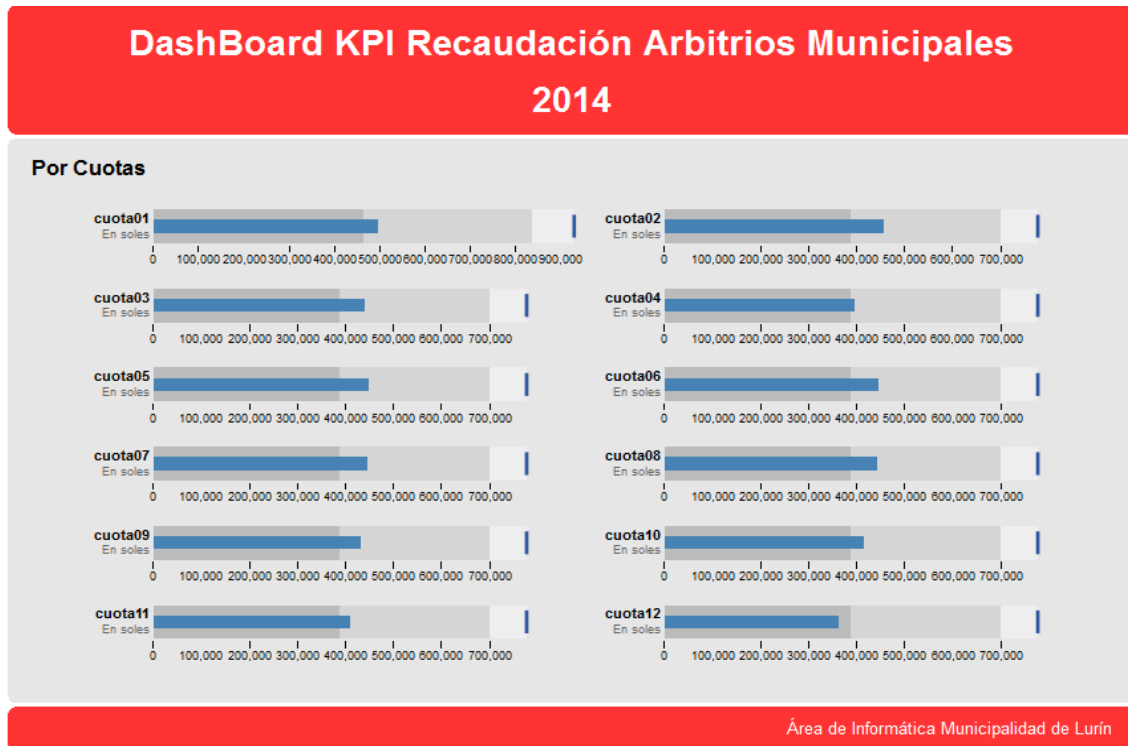


Figura 96. Dashboard KPI de Arbitrios Municipales

Fuente: Elaboración Propia

		ARB. MUNICIPALES							
		IMP. PREDIAL				ARB. MUNICIPALES			
		Medidas				Medidas			
Año - Cuota	TipoPago	• Importe	• Derecho Emision	• Mora	• Monto Total	• Importe	• Derecho Emision	• Mora	• Monto Total
2003		141.720,52	12.310,99	11.977,53	166.009,04	109.536,47	6.619,7	6.742,75	122.898,92
01		34.004,32	12.310,99	3.153,32	49.468,63	9.228,68	6.619,7	2.184	18.032,38
02		33.788,83	0	3.179,66	36.968,49	9.194,11	0	552,26	9.746,37
03		37.591,28	0	3.045,62	40.636,9	9.190,57	0	431,22	9.621,79
04		36.336,09	0	2.598,93	38.935,02	9.226,59	0	481,31	9.707,9
05						8.991,94	0	413,43	9.405,37
06						8.998,42	0	405,92	9.404,34
07						9.037,4	0	401,33	9.438,73
08						9.029,48	0	387,99	9.417,47
09						9.067,04	0	382,3	9.449,34
10						9.132,85	0	374,39	9.507,24
11						9.213,27	0	368,44	9.581,71
12						9.226,12	0	360,16	9.586,28
2004		249.981,02	26.800,96	13.745,68	290.527,66	273.107,29	9.761,64	12.283,1	295.152,03
2005		346.103,8	30.125,7	19.075,17	395.304,67	353.931,32	11.518	12.904,5	378.353,82
2006		438.162,83	34.284,2	16.637,23	489.084,26	471.496,24	25.224	15.907,84	512.628,08
2007		954.984,68	51.770,3	39.580,7	1.046.335,68	733.395,57	46.402,19	32.005	811.880,76
2008		5.079.009,84	86.664	78.898,86	5.244.573,7	3.111.417,61	80.601	63.988,62	3.256.025,23
2009		5.739.824,29	87.536	110.846,32	5.938.206,61	3.531.511,25	83.742	58.712,89	3.674.032,14
2010		6.819.739,88	41.196	89.930,66	6.950.914,54	4.246.899,48	34.917	70.269,44	4.352.178,92
2011		7.204.739,52	88.891,26	93.001,44	7.386.614,22	4.545.711,31	85.542	42.050,75	4.673.610,06
2012		8.083.145,46	84.951	63.659,02	8.231.239,48	5.688.556,8	82.149	40.899,02	5.811.604,82
2013		9.041.383,31	72.708	42.352,6	9.156.443,91	5.565.333,17	67.743	31.527,45	5.664.603,62
2014		8.983.581,24	58.797	702,91	9.043.081,15	5.138.764,19	48.135	385,93	5.187.285,12

Figura 97. Cubo Recaudación Tipos de Pagos por Cuotas

Fuente: Elaboración Propia

Tubo Recaudación Arbitrios Municipales									
Tubo Recaudación Arbitrios Municipales									
ARB. MUNICIPALES									
Año - Cuota									
2014									
2013									
2012									
Zona - Microzona	Importe	Derecho Emision	Mora	Monto Total	Importe	Derecho Emision	Mora	Monto Total	Importe
ZONA A	1.964.078,26	19.485	240,59	1.983.803,85	2.143.803,66	30.942	13.737,56	2.188.483,22	2.126.804,93
ZONA B	992.428,87	12.513	62,42	1.005.004,29	1.128.354,97	15.987	5.744,39	1.150.086,36	1.205.227,43
ZONA C	1.456.011,57	4.890	51,05	1.460.952,62	1.457.708,68	6.033	7.147,94	1.470.889,62	1.476.322,22
ZONA D	551.041,64	5.265	13,91	556.320,55	638.991,49	6.570	2.755,64	648.317,13	655.912,39
ZONA E	175.203,85	5.982	17,96	181.203,81	196.474,37	8.211	2.141,92	206.827,29	224.289,83

Figura 98. Cubo Recaudación Arbitrios Municipales por Zonas

Fuente: Elaboración Propia

Cubo Deudas Tipos de Pagos por Cuotas									
Cubo Deudas Tipos de Pagos por Cuotas									
IMP. PREDIAL									
ARB. MUNICIPALES									
Año - Cuota									
Medidas									
Importe	Derecho Emision	Mora	Monto Total	Importe	Derecho Emision	Mora	Monto Total	Importe	Monto Total
2003	464.989,54	15.288,03	88,38	480.365,95	263.021,02	11.362	0	274.383,02	
01	114.028,07	15.288,03	35,27	129.351,37	21.984,2	11.362	0	33.346,2	
02	114.519,53	0	49,65	114.569,18	21.883,3	0	0	21.883,3	
03	121.405,82	0	1,73	121.407,55	21.887,12	0	0	21.887,12	
04	115.036,12	0	1,73	115.037,85	21.910,06	0	0	21.910,06	
05					21.886,32	0	0	21.886,32	
06					21.941,37	0	0	21.941,37	
07					21.944,29	0	0	21.944,29	
08					21.938,03	0	0	21.938,03	
09					21.903,61	0	0	21.903,61	
10					21.909,61	0	0	21.909,61	
11					21.919,57	0	0	21.919,57	
12					21.913,54	0	0	21.913,54	
2004	631.214,92	26.838,5	0	658.053,42	491.746,54	13.181,5	0	504.928,04	
2005	637.671,98	27.721,7	0	665.393,68	584.081,45	14.900,5	0	598.981,95	
2006	655.424,95	28.615,7	0	684.040,65	669.301,45	31.488	0	700.789,45	
2007	647.314,19	33.471,4	6,72	680.792,31	494.788,39	38.689	21,55	533.510,94	
2008	640.076,58	30.858	334,49	671.275,07	683.338,62	42.009	2,74	725.362,36	
2009	912.061,82	33.543	841,32	946.446,14	835.336,65	42.768	6,9	878.123,55	
2010	958.662,84	10.890	528,98	970.114,82	1.061.380,45	15.834	46,23	1.077.290,68	
2011	1.297.465,19	45.267	433,42	1.343.141,61	1.084.259,7	46.509	4,28	1.130.904,98	
2012	1.629.495,84	53.574	195,4	1.683.112,24	1.618.110,17	58.764	24,83	1.676.899	
2013	2.432.400,83	78.717	370,96	2.511.485,79	2.952.545,01	86.538	41,14	3.039.124,15	
2014	4.168.737,18	94.446	0	4.263.180,18	4.185.397,87	102.957	0	4.288.354,87	

Figura 99. Cubo Deudas Tipos de Pagos por Cuotas

Fuente: Elaboración Propia

Tubo Deudas Arbitrios Municipales por Zonas									
	Tubo Deudas Arbitrios Municipales por Zonas								
	ARB. MUNICIPALES								
	Año - Cuota								
	2014				2013				2012
	Medidas								
Zona - Microzona	Importe	Derecho Emision	Mora	Monto Total	Importe	Derecho Emision	Mora	Monto Total	Importe
ZONA A	1.427.056,97	34.404	0	1.461.460,97	949.700,24	27.249	15,58	976.964,82	461.771,29
ZONA B	980.764	25.032	0	1.005.796	749.400,51	21.150	8,06	770.558,57	435.688,92
ZONA C	505.878,57	8.043	0	513.921,57	273.024,6	6.480	6,88	279.511,48	134.539,36
ZONA D	632.415,28	15.384	0	647.799,28	464.708,04	13.518	8,73	478.234,77	272.560,14
ZONA E	639.283,05	20.094	0	659.377,05	515.711,62	18.141	1,89	533.854,51	313.550,46

**Figura 100.** Cubo Deudas Arbitrios Municipales por Zonas Fuente: Elaboración Propia

Cubo Recaudación por Fechas									
	Cubo Recaudación por Fechas								
	IMP. PREDIAL				ARB. MUNICIPALES				
	Medidas								
Año - Trimestre - Mes	Importe	Derecho Emision	Mora	Monto Total	Importe	Derecho Emision	Mora	Monto Total	
2003	6.029,67	159,5	0	6.189,17	2.067,18	109	0	2.176,18	
2004	36.116,42	655	0	36.771,42	754,63	55	0	809,63	
2005	80.258,3	1.287	0	81.545,3	7.040,12	208,5	0	7.248,62	
2006	71.541,29	1.808,5	0	73.349,79	12.270,75	496,5	0	12.767,25	
2007	156.450,74	3.241	53,85	159.745,59	48.835,69	1.463	217,73	50.516,42	
2008	5.044.058,32	111.187,66	38.799,96	5.194.045,94	3.113.992,91	84.622,5	35.543,47	3.234.224,88	
2009	5.530.497,34	99.829,49	62.946,52	5.693.273,35	3.488.383,08	88.258	51.183,9	3.627.824,98	
201	126,78	0	0	126,78	304,56	6	0	310,56	
2010	6.277.696,93	68.433	43.759	6.389.924,93	3.818.939,5	52.526,19	33.737,16	3.905.211,85	
2011	7.370.240,41	93.585,8	102.204,74	7.566.043,95	4.702.190,06	79.109,34	112.706,34	4.894.377,74	
2012	8.778.517,62	97.313,96	67.198,08	8.943.020,66	5.929.656,11	83.139,5	58.711,12	6.071.596,73	
2013	10.141.716,51	109.045	121.802,72	10.372.045,23	6.952.999,63	100.895	72.677,61	7.126.596,24	
I	2.930.875,37	30.209,5	8.187,12	2.968.770,99	2.174.569,05	26.519	7.826,19	2.208.914,24	
01-Enero	108.722	4.089	2.645,73	115.456,73	261.738,1	3.895,5	1.587,29	267.220,89	
02-Febrero	835.954,84	7.226,5	3.417,15	846.097,49	513.944,01	6.113	3.149,47	523.206,48	
03-Marzo	1.986.198,53	18.894	2.124,24	2.007.216,77	1.398.886,94	16.510,5	3.089,43	1.418.486,87	
II	2.481.490,43	26.739	19.113,16	2.527.342,59	1.576.122,67	21.546,5	14.824,21	1.612.493,38	
III	2.393.579,51	33.148,5	38.958,45	2.465.674,46	1.774.191,37	30.653,5	14.565,71	1.819.410,58	
IV	2.335.771,2	18.948	55.543,99	2.410.257,19	1.428.116,54	22.176	35.461,5	1.485.778,04	
2014	9.454.220,92	82.457	143.621,35	9.680.293,27	5.450.464,61	70.721,5	22.879,47	5.544.065,58	

**Figura 101.** Cubo Recaudación por Fechas Fuente: Elaboración Propia

### 3.13.IMPLEMENTACIÓN

#### 3.13.1. Verificación de Tecnología

En fases anteriores se definió cuáles serían los recursos hardware y software a utilizarse para el proyecto. El hardware mencionado estaba presente en ese momento, es decir, el área de informática ya contaba con este. Pero en cuanto al software, aún no se lo descargaba.

Llegados a esta fase de implementación, deberemos antes que nada verificar si contamos realmente con los recursos antes planeados. En la Tabla 43 se puede apreciar la verificación de los recursos.

**Tabla 43.** Verificación de la tecnología

Fuente: Elaboración Propia

TIPO	RECURSO	¿SE TIENE EL RECURSO?	COMENTARIOS
Hardware	Servidor físico HP	SI	-
	PC Intel	SI	-
Software	Oracle VM VirtualBox 4.3.12	SI	-
	Servidor virtualizado	SI	-
	PostgreSQL 9.3	SI	-
	Pentaho Data Integration (Kettle) 5.2.0.0	SI	-
	Pentaho Design Studio 4.0.0	SI	-
	Pentaho Schema Workbench 3.8.0.0	SI	-
	JPivot	SI	-
	Pentaho Report Designer 5.2.0.0	SI	-
	Pentaho BI Server 5.2.0.0	SI	-
	Community Dashboard Editor (CDE) 5.2.0.0	SI	-
	Ivy Dashboard Components (IvyDC) 0.0.4	SI	Fue descargado e instalado pero al final no fue utilizado.

### 3.13.2. Manual de Usuario

El manual de usuario para Administrador, quien se encarga de la gestión de usuarios, de la ejecución del proceso ETL y de la creación de dashboards KPI para nuevos años, podemos verlo en el Apéndice V. Mientras que el manual para el Gerente, quien únicamente puede acceder a los reportes y visualizarlos, podemos verlo en el Apéndice VI.

## 3.14. MANTENIMIENTO Y CRECIMIENTO

### 3.14.1. Pruebas de Funcionamiento del Sistema

Una vez implementado el sistema, es necesario probarlo tanto para ver el si se da el funcionamiento esperado como para descubrir errores. Así pues, se realizó una prueba por cada funcionalidad del producto. Los resultados podemos observarlos en la Tabla 44.

**Tabla 44.** Pruebas de Funcionamiento del Sistema

Fuente: Elaboración Propia

<b>FUNCIONALIDAD</b>	<b>RESULTADO</b>	<b>TIEMPO DE DEMORA (seg)</b>	<b>COMENTARIOS</b>
Inicio de servidor BI	Satisfactorio	43	Solo es necesario ejecutarlo si se apaga el servidor virtual.
Autenticación de usuario en el sistema	Satisfactorio	10	-
Ejecución del proceso ETL (parte1)	Satisfactorio	420	Desde la consola de usuario.
Ejecución del proceso ETL (parte2)	Satisfactorio	1620	Desde la consola de usuario.
Ejecución de los reportes a medida	Satisfactorio	45	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.
Ejecución de los dashboards	Satisfactorio	25	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.
Ejecución de los cubos	Satisfactorio	35	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.

**CAPÍTULO IV:**  
**ANÁLISIS DE RESULTADOS Y CONTRASTACIÓN DE LA**  
**HIPÓTESIS**

*Valor es lo que se necesita para  
levantarse y hablar, pero  
también es lo que se requiere  
para sentarse y escuchar.*

**Winston Churchill**

**1874 - 1965**



## **4.1. POBLACIÓN Y MUESTRA**

### **4.1.1. Población**

Las unidades a ser analizadas son todas las tomas de decisiones de la gerencia de rentas de la Municipalidad de Lurín.

### **4.1.2. Muestra**

Para la investigación se realizará un muestreo intencional<sup>52</sup> (no aleatorio) y se tomará una muestra de 30 tomas de decisiones al considerarse una cantidad razonable y manejable.

## **4.2. NIVEL DE CONFIANZA**

El nivel de confianza seleccionado será del 95%.

## **4.3. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS**

### **4.3.1. Resultados Genéricos**

#### **A) Planificación del proyecto**

- ✓ Descripción del proyecto.
- ✓ Objetivos del proyecto.
- ✓ Alcance del proyecto.
- ✓ Stakeholders.
- ✓ Análisis de riesgos.
- ✓ Plan de proyecto.

#### **B) Definición de los Requerimientos del Negocio**

- ✓ Proceso de negocio.
- ✓ Proceso de negocio y temas analíticos.
- ✓ Matriz procesos / dimensiones.
- ✓ Requerimientos.
- ✓ Documentación de los requerimientos.
- ✓ Hoja de gestión.
- ✓ Hoja de análisis.

#### **C) Modelo Dimensional**

- ✓ Dimensiones.
- ✓ Granularidad.
- ✓ Hechos.
- ✓ Medidas.
- ✓ Diseño del modelo dimensional.

#### **D) Diseño de la Arquitectura Técnica**

---

<sup>52</sup> Ávila, H., Introducción a la Metodología de la Investigación, Edición Electrónica, España, 2006, p. 89

- ✓ Arquitectura.
- ✓ Back – room
- ✓ Front - room

#### **E) Selección e Instalación de Productos**

- ✓ Evaluación de los productos.
- ✓ Herramientas seleccionadas.

#### **F) Diseño Físico**

- ✓ Tablas de apoyo.
- ✓ Tablas de dimensiones.
- ✓ Tabla de hechos.
- ✓ Diseño modelo físico.
- ✓ Diseño modelo físico de tablas de apoyo.

#### **G) Diseño ETL**

- ✓ Extracción
- ✓ Transformación
- ✓ Carga
- ✓ Automatización del proceso.

#### **H) Diseño del Cubo OLAP**

#### **I) Especificación de Aplicaciones BI**

- ✓ Lista de reportes.

#### **J) Desarrollo de Aplicaciones BI**

- ✓ Elaboración de reportes.

#### **K) Implementación**

- ✓ Verificación de tecnología
- ✓ Manual de usuario.

#### **L) Mantenimiento y Crecimiento**

- ✓ Pruebas de funcionamiento del sistema.
- ✓ Nuevos reportes.

### **4.3.2. Resultados Específicos**

En la siguiente tabla se muestran las medidas de los indicadores tanto de la Pre-Prueba como de la Post-Prueba.

**Tabla 45.** Resultados de la Pre-Prueba y Post-Prueba para todos los KPI.

Fuente: Elaboración Propia

N°	KPI <sub>1</sub> : Tiempo empleado en la generación de reportes (min.)		KPI <sub>2</sub> : Tiempo que el usuario emplea en el análisis de la información (min.)		KPI <sub>3</sub> : Número de veces que el usuario accede a la información al día (#)		KPI <sub>4</sub> : Porcentaje de exactitud de la información (%)		KPI <sub>5</sub> : Nivel de satisfacción del usuario	
	Pre-Prueba	Post-Prueba	Pre-Prueba	Post-Prueba	Pre-Prueba	Post-Prueba	Pre-Prueba	Post-Prueba	Pre-Prueba	Post-Prueba
1	32	1	121	23	1	2	85	97	Bajo	Regular
2	28	2	120	20	0	1	81	96	Regular	Alto
3	30	2	119	19	0	0	87	95	Regular	Regular
4	34	1	119	22	0	3	75	97	Bajo	Alto
5	29	1	125	20	1	2	82	96	Bajo	Regular
6	33	1	118	20	0	3	85	95	Regular	Alto
7	30	1	120	18	0	0	83	95	Bajo	Alto
8	28	2	123	24	0	0	82	97	Regular	Alto
9	34	1	123	21	0	2	86	97	Bajo	Alto
10	30	2	119	18	2	3	79	95	Regular	Regular
11	28	1	117	19	0	0	87	95	Bajo	Alto
12	34	1	120	19	0	1	81	98	Regular	Alto
13	30	1	122	23	0	0	76	95	Bajo	Regular
14	28	2	120	22	1	4	86	98	Regular	Alto
15	27	1	120	20	2	2	76	98	Bajo	Regular
16	31	2	120	20	0	0	86	98	Bajo	Regular
17	33	2	118	19	0	2	85	97	Regular	Alto
18	29	1	123	18	0	1	78	98	Bajo	Alto
19	30	1	121	20	0	3	89	98	Bajo	Alto
20	27	1	118	23	0	0	80	96	Bajo	Alto
21	34	1	119	21	1	1	78	98	Regular	Alto
22	30	1	118	23	0	2	89	97	Regular	Regular
23	27	1	118	20	0	0	86	98	Bajo	Alto
24	32	2	122	19	0	3	81	98	Bajo	Regular
25	30	2	124	17	0	0	84	97	Regular	Alto
26	28	1	117	19	2	3	86	97	Bajo	Alto
27	31	1	124	21	1	2	83	96	Bajo	Alto
28	34	2	121	23	0	3	75	97	Regular	Regular
29	33	2	119	20	0	0	79	97	Bajo	Alto
30	30	1	123	21	0	2	84	98	Bajo	Alto

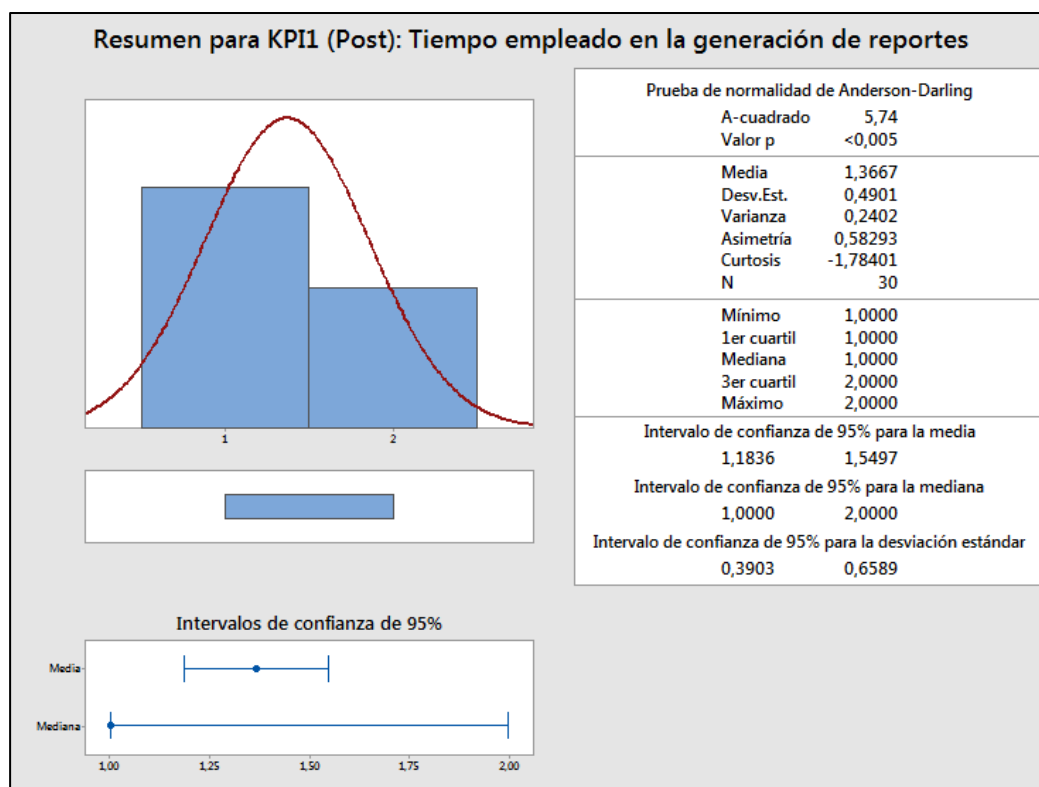
**A. KPI<sub>1</sub>:** Tiempo empleado en la generación de reportes.

**Tabla 46.** Resultados de Pre-Prueba y Post-Prueba para el KPI<sub>1</sub> Fuente: Elaboración Propia

	Pre-Prueba	Post-Prueba		
	32	1	1	1
	28	2	2	2
	30	2	2	2
	34	1	1	1
	29	1	1	1
	33	1	1	1
	30	1	1	1
	28	2	2	2
	34	1	1	1
	30	2	2	2
	28	1	1	1
	34	1	1	1
	30	1	1	1
	28	2	2	2
	27	1	1	1
	31	2	2	2
	33	2	2	2
	29	1	1	1
	30	1	1	1
	27	1	1	1
	34	1	1	1
	30	1	1	1
	27	1	1	1
	32	2	2	2
	30	2	2	2
	28	1	1	1
	31	1	1	1
	34	2	2	2
	33	2	2	2
	30	1	1	1
<b>Promedio</b>	<b>30.47</b>	<b>1.37</b>		
<b>Meta planteada</b>		<b>5</b>		
<b>Nº menor a promedio</b>		<b>19</b>	<b>30</b>	<b>30</b>
<b>% menor a promedio</b>		<b>63.3</b>	<b>100</b>	<b>100</b>

- El 63.3% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que su tiempo promedio.
- El 100.0% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que la meta planteada.
- El 100.0% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que el tiempo promedio en la Pre-Prueba.

## Estadística Descriptiva



**Figura 102.** Estadística Descriptiva para KPI<sub>1</sub> (Post)

Fuente: Elaboración Propia

- La distancia "promedio" de las observaciones individuales de los Tiempos Empleados en la Generación de Reportes con respecto a la media es de 0.49 minutos.
- Alrededor del 95% de los Tiempos Empleados en la Generación de Reportes están dentro de 2 desviaciones estándar de la media, es decir, entre 1.18 y 1.55 minutos.
- La Curtosis = -1.78 indica que tenemos datos de tiempos con picos muy bajos.
- La Asimetría = 0.58 indica que la mayoría de los Tiempos Empleado en la Generación de Reportes son bajos.
- El 1er Cuartil (Q1) = 1.000 minutos, indica que el 25% de los Tiempos Empleados en la Generación de Reportes son menores que o igual a este valor.
- El 3er Cuartil (Q3) = 2.000 minutos, indica que el 75% de los Tiempos Empleados en la Generación de Reportes son menores que o igual a este valor.

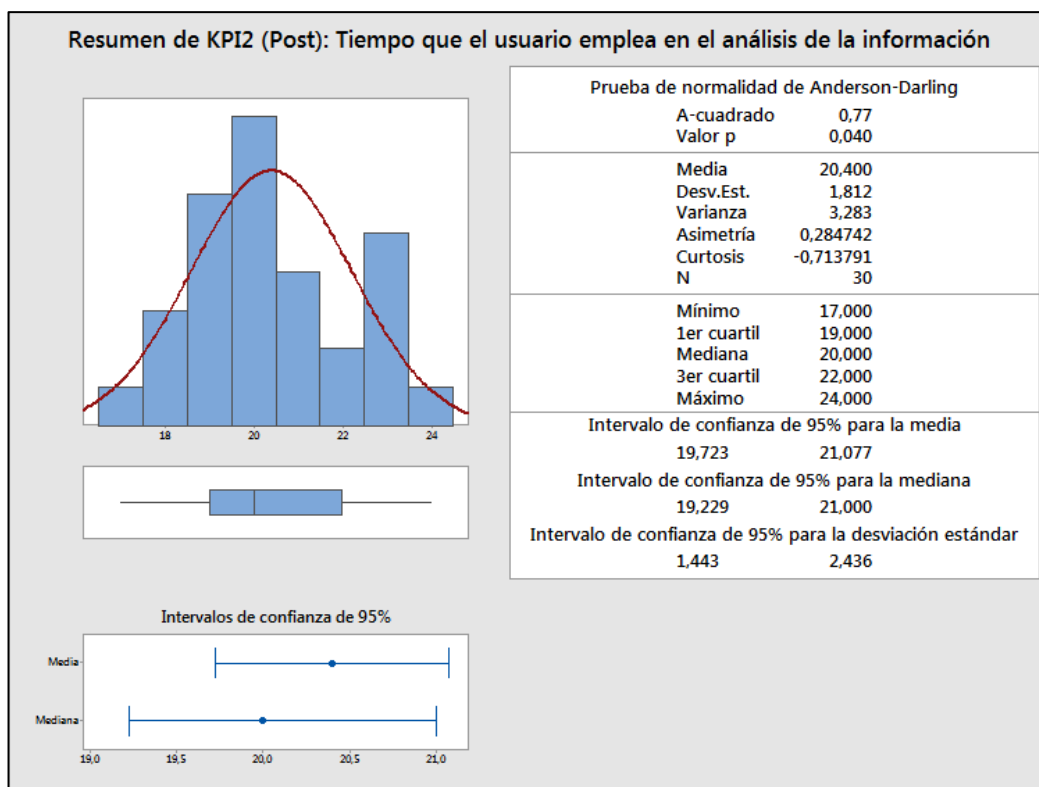
**B. KPI<sub>2</sub>:** Tiempo que el usuario emplea en el análisis de la información.

**Tabla 47.** Resultados de Pre-Prueba y Post-Prueba para el KPI<sub>2</sub> Fuente: Elaboración Propia

	<b>Pre-Prueba</b>	<b>Post-Prueba</b>	
	121	23	23
	120	20	20
	119	19	19
	119	22	22
	125	20	20
	118	20	20
	120	18	18
	123	24	24
	123	21	21
	119	18	18
	117	19	19
	120	19	19
	122	23	23
	120	22	22
	120	20	20
	120	20	20
	118	19	19
	123	18	18
	121	20	20
	118	23	23
	119	21	21
	118	23	23
	118	20	20
	122	19	19
	124	17	17
	117	19	19
	124	21	21
	121	23	23
	119	20	20
	123	21	21
<b>Promedio</b>	<b>120.37</b>	<b>20.4</b>	
<b>Meta planteada</b>		<b>25</b>	
<b>Nº menor a promedio</b>		<b>18</b>	<b>30</b>
<b>% menor a promedio</b>		<b>60</b>	<b>100</b>

- El 60.0% de los Tiempos Empleados en el Análisis de Reportes en la Post-Prueba fueron menores que su tiempo promedio.
- El 100.0% de los Tiempos Empleados en el Análisis de Reportes en la Post-Prueba fueron menores que la meta planteada.
- El 100.0% de los Tiempos Empleados en el Análisis de Reportes en la Post-Prueba fueron menores que el tiempo promedio en la Pre-Prueba.

### Estadística Descriptiva



**Figura 103.** Estadística Descriptiva para KPI<sub>2</sub> (Post)

Fuente: Elaboración Propia

- La distancia "promedio" de las observaciones individuales de los Tiempos que el Usuario Emplea en el Análisis de la Información con respecto a la media es de 1.81 minutos.
- Alrededor del 95% de los Tiempos que el Usuario Emplea en el Análisis de la Información están dentro de 2 desviaciones estándar de la media, es decir, entre 19.72 y 21.08 minutos.
- La Curtosis = -0.71 indica que tenemos datos de tiempos con picos muy bajos.
- La Asimetría = 0.28 indica que la mayoría de los Tiempos que el Usuario Emplea en el Análisis de la Información son bajos.
- El 1er Cuartil (Q1) = 19.000 minutos, indica que el 25% de los Tiempos que el Usuario Emplea en el Análisis de la Información son menores que o igual a este valor.
- El 3er Cuartil (Q3) = 22.000 minutos, indica que el 75% de los Tiempos que el Usuario Emplea en el Análisis de la Información son menores que o igual a este valor.

**C. KPI<sub>3</sub>:** Número de veces que el usuario accede a la información al día.

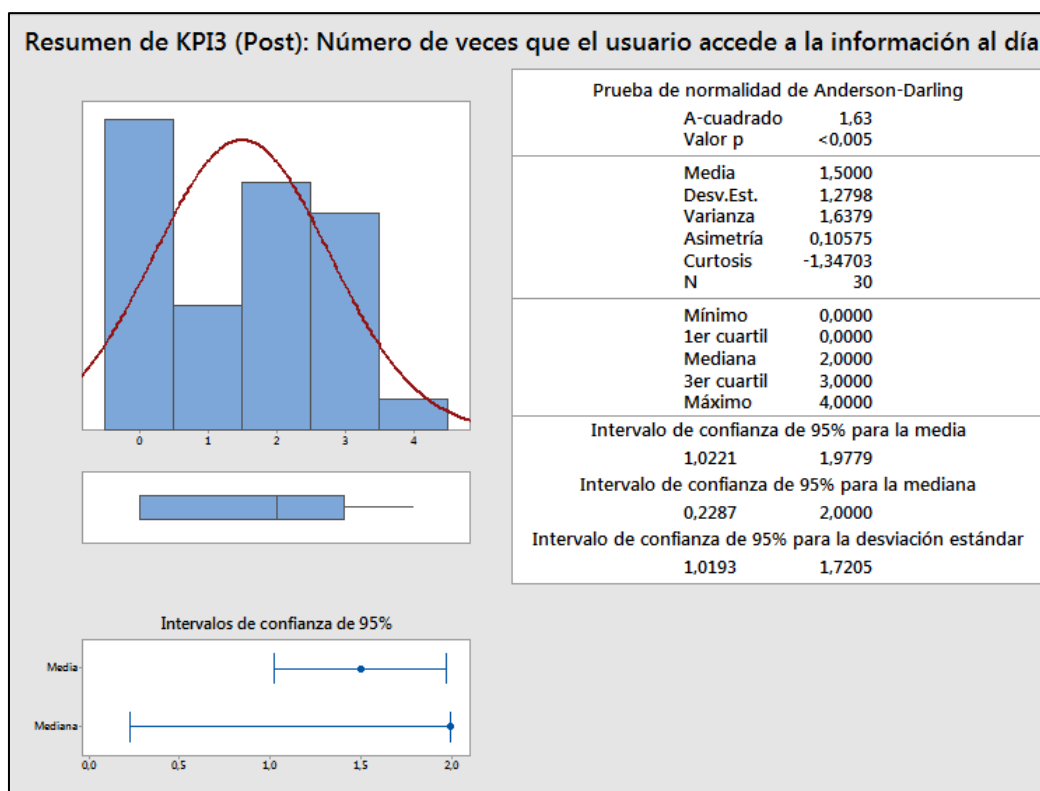
**Tabla 48.** Resultados de Pre-Prueba y Post-Prueba para el KPI<sub>3</sub> Fuente: Elaboración Propia

Pre-Prueba	Post-Prueba		
1	2	2	2
0	1	1	1
0	0	0	0
0	3	3	3
1	2	2	2
0	3	3	3
0	0	0	0
0	0	0	0
0	2	2	2
2	3	3	3
0	0	0	0
0	1	1	1
0	0	0	0
1	4	4	4
2	2	2	2
0	0	0	0
0	2	2	2
0	1	1	1
0	3	3	3
0	0	0	0
1	1	1	1
0	2	2	2
0	0	0	0
0	3	3	3
0	0	0	0
2	3	3	3
1	2	2	2
0	3	3	3
0	0	0	0
0	2	2	2
<b>Promedio</b>	<b>0.37</b>	<b>1.5</b>	
<b>Meta planteada</b>		<b>1</b>	
<b>Nº mayor a promedio</b>		<b>16</b>	<b>20</b>
<b>% mayor a promedio</b>		<b>53.3</b>	<b>66.7</b>

- El 53.3% de los Números de Reportes Generados a la Semana en la Post-Prueba fueron mayores que su número promedio.
- El 53.3% de los Números de Reportes Generados a la Semana en la Post-Prueba fueron mayores que la meta planteada.
- El 66.7% de los Números de Reportes Generados a la Semana en la Post-Prueba fueron mayores que el número promedio en la Pre-Prueba.



### Estadística Descriptiva



**Figura 104.** Estadística Descriptiva para KPI<sub>3</sub> (Post)

Fuente: Elaboración Propia

- La distancia "promedio" de las observaciones individuales de los Números de Veces que el Usuario Accede a la Información al Día con respecto a la media es de 1.28 veces.
- Alrededor del 95% de los Números de Veces que el Usuario Accede a la Información al Día están dentro de 2 desviaciones estándar de la media, es decir, entre 1.02 y 1.98 veces.
- La Curtosis = -1.35 indica que tenemos datos de números de veces con picos muy bajos.
- La Asimetría = 0.11 indica que la mayoría de los Números de Veces que el Usuario Accede a la Información al Día son altos.
- El 1er Cuartil (Q1) = 0.000 veces, indica que el 25% de los Números de Veces que el Usuario Accede a la Información al Día son menores que o igual a este valor.
- El 3er Cuartil (Q3) = 3.000 veces, indica que el 75% de los Números de Veces que el Usuario Accede a la Información al Día son menores que o igual a este valor.

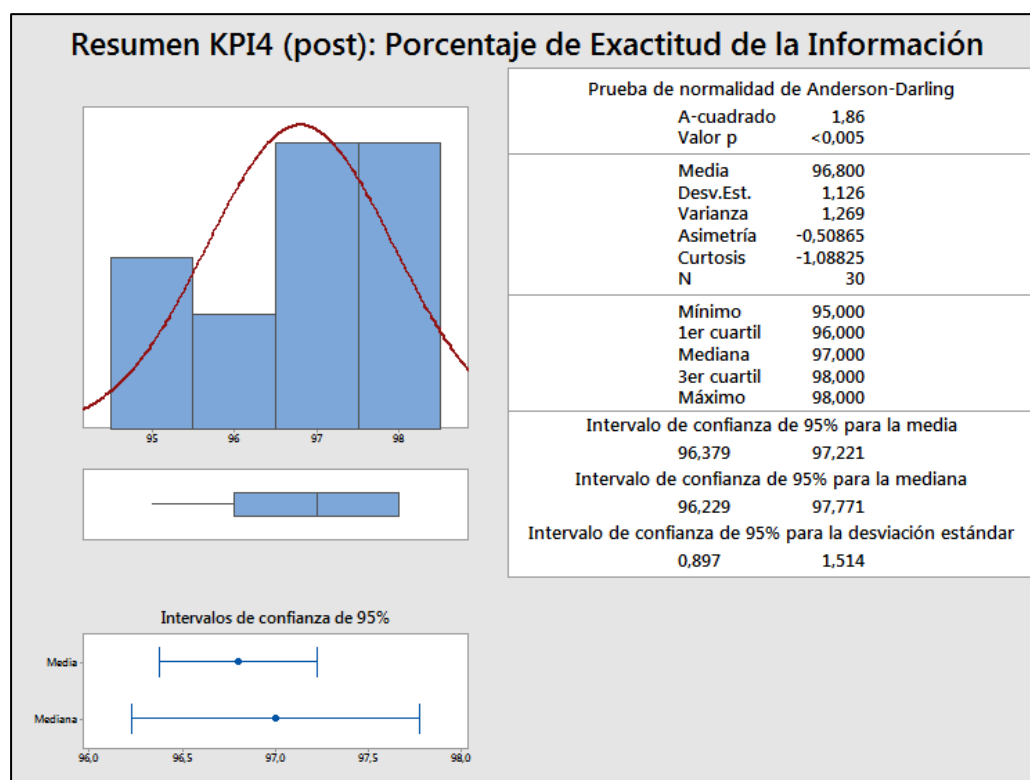
**D. KPI4: Porcentaje de Exactitud de la Información.**

**Tabla 49.** Resultados de Pre-Prueba y Post-Prueba para el KPI4 Fuente: Elaboración Propia

	<b>Pre-Prueba</b>	<b>Post-Prueba</b>	
	85	97	97
	81	96	96
	87	95	95
	75	97	97
	82	96	96
	85	95	95
	83	95	95
	82	97	97
	86	97	97
	79	95	95
	87	95	95
	81	98	98
	76	95	95
	86	98	98
	76	98	98
	86	98	98
	85	97	97
	78	98	98
	89	98	98
	80	96	96
	78	98	98
	89	97	97
	86	98	98
	81	98	98
	84	97	97
	86	97	97
	83	96	96
	75	97	97
	79	97	97
	84	98	98
<b>Promedio</b>	<b>82.47</b>	<b>96.8</b>	
<b>Meta planteada</b>		<b>95</b>	
<b>Nº mayor a promedio</b>		<b>20</b>	<b>30</b>
<b>% mayor a promedio</b>		<b>66.67</b>	<b>100.00</b>

- El 66.67% de los Porcentajes de Exactitud de la Información en la Post-Prueba fueron mayores que su número promedio.
- El 80.00% de los Porcentajes de Exactitud de la Información en la Post-Prueba fueron mayores que la meta planteada.
- El 100.00% de los Porcentajes de Exactitud de la Información en la Post-Prueba fueron mayores que el número promedio en la Pre-Prueba.

## Estadística Descriptiva



**Figura 105.** Estadística Descriptiva para KPI<sub>4</sub> (Post)

Fuente: Elaboración Propia

- La distancia "promedio" de las observaciones individuales de los Porcentajes de Exactitud de la Información con respecto a la media es de 1.13 decisiones.
- Alrededor del 95% de los Porcentajes de Exactitud de la Información están dentro de 2 desviaciones estándar de la media, es decir, entre 96.38 y 97.22 decisiones.
- La Curtosis = -1.09 indica que tenemos datos de números de decisiones con picos muy bajos.
- La Asimetría = -0.51 indica que la mayoría de los Porcentajes de Exactitud de la Información son altos.
- El 1er Cuartil (Q1) = 96 por ciento, indica que el 25% de los Porcentajes de Exactitud de la Información son menores que o igual a este valor.
- El 3er Cuartil (Q3) = 98 por ciento, indica que el 75% de los Porcentajes de Exactitud de la Información son menores que o igual a este valor.

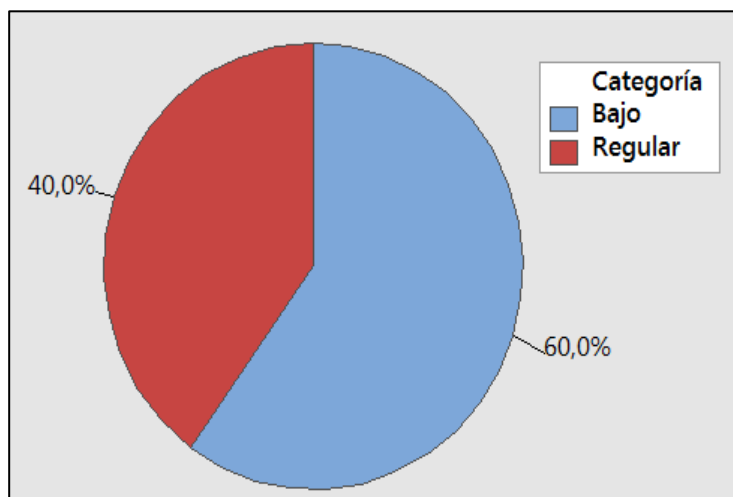
**E. KPI5: Nivel de satisfacción del usuario.**

**Valores de la Pre-Prueba**

Nro. Medición	1	2	3	4	5	6	7	8	9	10
Valor	Baj	Reg	Reg	Baj	Baj	Reg	Baj	Reg	Baj	Reg
	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
	Baj	Reg	Baj	Reg	Baj	Baj	Reg	Baj	Baj	Baj
	<b>21</b>	<b>22</b>	<b>23</b>	<b>24</b>	<b>25</b>	<b>26</b>	<b>27</b>	<b>28</b>	<b>29</b>	<b>30</b>
	Reg	Reg	Baj	Baj	Reg	Baj	Baj	Reg	Baj	Baj

Estado	Frecuencia
Alto	0
Regular	12
Bajo	18

Estado	Frecuencia
Bueno	0
Malo	30



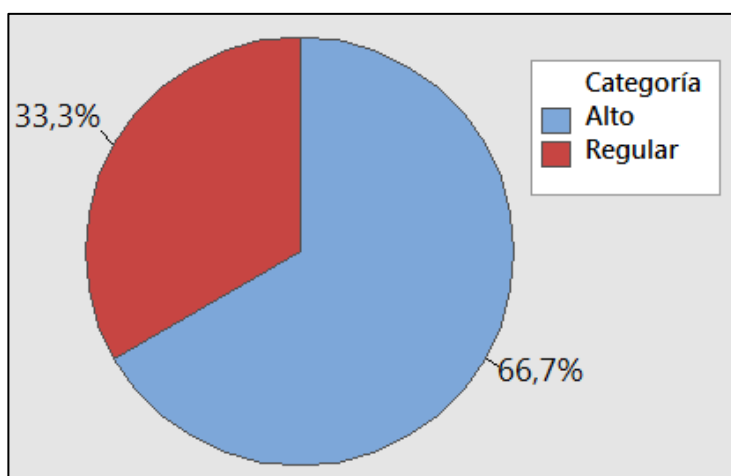
- El 60% de las veces el Nivel de Satisfacción fue catalogado como Bajo por el gerente.
- El 40% de las veces el Nivel de Satisfacción fue catalogado como Regular por el gerente.
- El 0% de las veces el Nivel de Satisfacción fue catalogado como Alto por el gerente.
- Se determina que el 0% de las veces el Nivel de Satisfacción es Bueno.
- Se determina que el 100% de las veces el Nivel de Satisfacción es Malo.

**Valores de la Post-Prueba**

Nro. Medición	1	2	3	4	5	6	7	8	9	10
Valor	Reg	Alt	Reg	Alt	Reg	Alt	Alt	Alt	Alt	Reg
	11	12	13	14	15	16	17	18	19	20
	Alt	Alt	Reg	Alt	Reg	Reg	Alt	Alt	Alt	Alt
	21	22	23	24	25	26	27	28	29	30
	Alt	Reg	Alt	Reg	Alt	Alt	Alt	Reg	Alt	Alt

Estado	Frecuencia
Alto	20
Regular	10
Bajo	0

Estado	Frecuencia
Bueno	20
Malo	10



- El 0% de las veces el Nivel de Satisfacción fue catalogado como Bajo por el gerente.
- Sólo el 33.3% de las veces el Nivel de Satisfacción fue catalogado como Regular por el gerente.
- Ahora el 66.7% de las veces el Nivel de Satisfacción fue catalogado como Alto por el gerente.
- Se determina ahora que el 66.7% de las veces el Nivel de es Bueno.
- Se determina ahora que sólo el 33.3% de las veces el Nivel de Satisfacción es Malo.

### 4.4. CONTRASTACIÓN DE LA HIPÓTESIS

Para nuestro proyecto de investigación se presentaron cinco indicadores:

**Tabla 50.** Indicadores para la Contrastación de la Hipótesis Fuente: Elaboración Propia

Indicador	Pre-Prueba (Media: x1)	Post-Prueba (Media: x2)	Comentario
Tiempo empleado en la generación de reportes	30.47 min	1.37 min	---
Tiempo que el usuario emplea en el análisis de la información	120.37 min	20.4 min	---
Números de veces que el usuario accede a la información al día	0.37 veces	1.5 veces	---
Porcentaje de exactitud de la información	82.47 %	96.8%	---
Nivel de satisfacción del usuario	---	---	No contrastado. Indicador cualitativo.

#### A. Contrastación para Tiempo Empleado en la Generación de Reportes

Se debe validar el impacto que tiene el desarrollo de la solución de Business Intelligence en el Tiempo Empleado en la Generación de Reportes del proceso de toma de decisiones de rentas, llevado a cabo en la muestra. Se realiza una medición antes del desarrollo de la solución de BI (Pre-Prueba) y otra después del desarrollo de la solución de BI (Post-Prueba). La siguiente tabla contiene los Tiempos Empleados en la Generación de Reportes para ambas muestras:

Pre - Prueba	32	28	30	34	29	33	30	28	34	30	28	34	30	28	27
	31	33	29	30	27	34	30	27	32	30	28	31	34	33	30

Post - Prueba	1	2	2	1	1	1	1	2	1	2	1	1	1	2	1
	2	2	1	1	1	1	1	1	2	2	1	1	2	2	1

**Hi:** El desarrollo de una solución de Business Intelligence disminuye el Tiempo Empleado en la Generación de Reportes (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

**Solución:**

**a) Planteamiento de la Hipótesis**

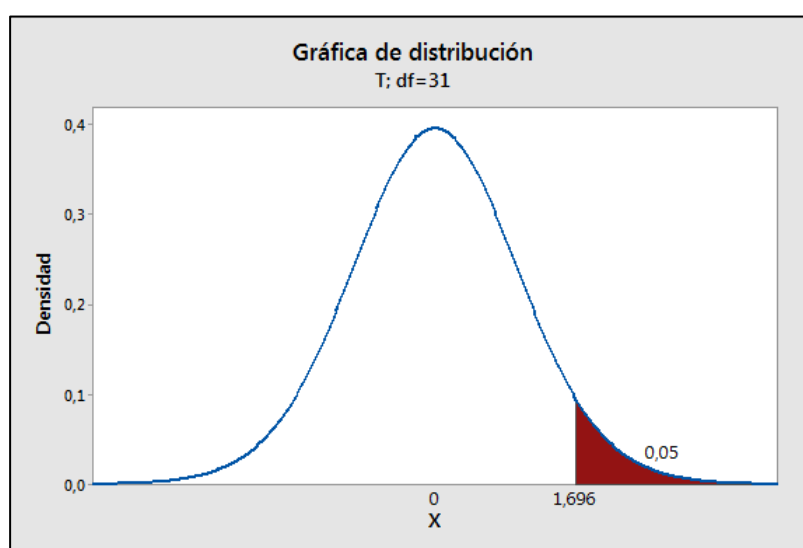
$\mu_1$  = Media del Tiempo Empleado en la Generación de Reportes en la Pre-Prueba.

$\mu_2$  = Media del Tiempo Empleado en la Generación de Reportes en la Post-Prueba.

$$H_0: \mu_1 \leq \mu_2$$

$$H_a: \mu_1 > \mu_2$$

**b) Criterios de Decisión**



**Figura 106.** Distribución de Probabilidad  $KPI_1$

Fuente: Elaboración Propia

**Tabla 51.** Resumen de Prueba t student para el  $KPI_1$

Fuente: Elaboración Propia

	Pre - Prueba	Post - Prueba
<b>Media (x)</b>	30.47	1.37
<b>Desviación Estándar (S)</b>	2.34	0.49
<b>Observaciones</b>	30	30
<b>Diferencia Hipotética de las Medias</b>	0	
<b>t_calculado: <math>t_c</math></b>	66.53	
<b>p-valor (una cola)</b>	0.000	
<b>Valor Crítico de una <math>t_{\alpha/2}</math> (una cola): <math>t_t</math></b>	1.696	

**c) Decisión Estadística**

Puesto que el valor-p = 0 <  $\alpha$  = 0.05, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula ( $H_0$ ) y considerar que la hipótesis alterna ( $H_a$ ) es cierta. La prueba resultó ser significativa.

**B. Contrastación para Tiempo que el Usuario Emplea en el Análisis de la Información**

Se debe validar el impacto que tiene el desarrollo de la solución de Business Intelligence en el Tiempo que el Usuario Emplea en el Análisis de la Información del proceso de toma de decisiones de rentas, llevado a cabo en la muestra. Se realiza una medición antes del desarrollo de la solución de BI (Pre-Prueba) y otra después del desarrollo de la solución de BI (Post-Prueba). La siguiente tabla contiene los Tiempos Empleados en el Análisis de Reportes para ambas muestras:

<b>Pre - Prueba</b>	121	120	119	119	125	118	120	123	123	119	117	120	122	120	120
	120	118	123	121	118	119	118	118	122	124	117	124	121	119	123

<b>Post - Prueba</b>	23	20	19	22	20	20	18	24	21	18	19	19	23	22	20
	20	19	18	20	23	21	23	20	19	17	19	21	23	20	21

**H<sub>i</sub>:** El desarrollo de una solución de Business Intelligence disminuye el Tiempo que el Usuario Emplea en el Análisis de la Información (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

**Solución:**

**a) Planteamiento de la Hipótesis**

$\mu_1$  = Media del Tiempo que el Usuario Emplea en el Análisis de la Información en la Pre-Prueba.

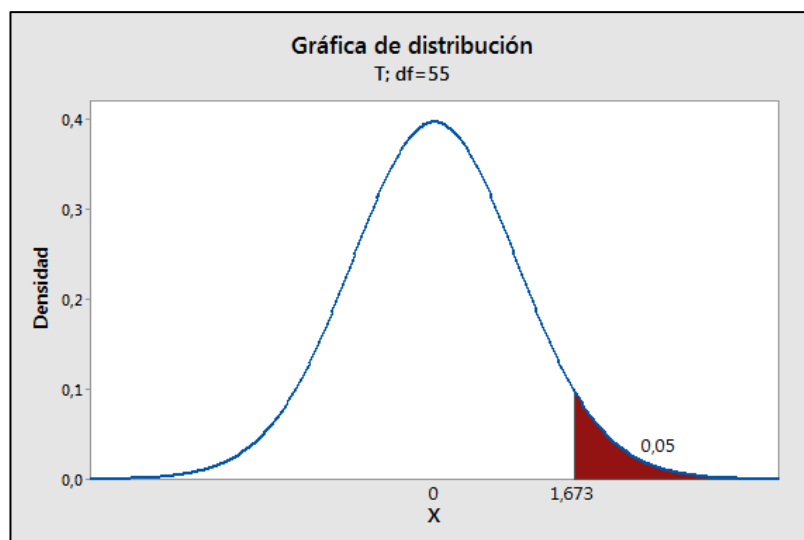
$\mu_2$  = Media del Tiempo que el Usuario Emplea en el Análisis de la Información en la Post-Prueba.

$$H_0: \mu_1 \leq \mu_2$$

$$H_a: \mu_1 > \mu_2$$

**b) Criterios de Decisión**





**Figura 107.** Distribución de Probabilidad KPI<sub>2</sub>

Fuente: Elaboración Propia

**Tabla 52.** Resumen de Prueba t student para el KPI<sub>2</sub>

Fuente: Elaboración Propia

	Pre - Prueba	Post - Prueba
<b>Media (x)</b>	120.37	20.4
<b>Desviación Estándar (S)</b>	2.22	1.81
<b>Observaciones</b>	30	30
<b>Diferencia Hipotética de las Medias</b>	0	
<b>t_calculado: t<sub>c</sub></b>	191.06	
<b>p-valor (una cola)</b>	0.000	
<b>Valor Crítico de una t<sub>α/2</sub> (una cola): t<sub>t</sub></b>	1.673	

### c) Decisión Estadística

Puesto que el valor-p = 0 < α = 0.05, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H<sub>0</sub>) y considerar que la hipótesis alterna (H<sub>a</sub>) es cierta. La prueba resultó ser significativa.

### C. Contrastación para Número de Veces que el Usuario Accede a la Información al Día

Se debe validar el impacto que tiene el desarrollo de la solución de Business Intelligence en el Número de Veces que el Usuario Accede a la Información al Día, llevado a cabo en la muestra. Se realiza una medición antes del desarrollo de la solución de BI (Pre-Prueba) y otra después del desarrollo de la solución de BI

(Post-Prueba). La siguiente tabla contiene los Números de Veces que el Usuario Accede a la Información al Día para ambas muestras:

<b>Pre - Prueba</b>	1	0	0	0	1	0	0	0	0	2	0	0	0	1	2
	0	0	0	0	0	1	0	0	0	0	2	1	0	0	0

<b>Post - Prueba</b>	2	1	0	3	2	3	0	0	2	3	0	1	0	4	2
	0	2	1	3	0	1	2	0	3	0	3	2	3	0	2

**Hi:** El desarrollo de una solución de Business Intelligence aumenta el Número de Veces que el Usuario Accede a la Información al Día (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

**Solución:**

**a) Planteamiento de la Hipótesis**

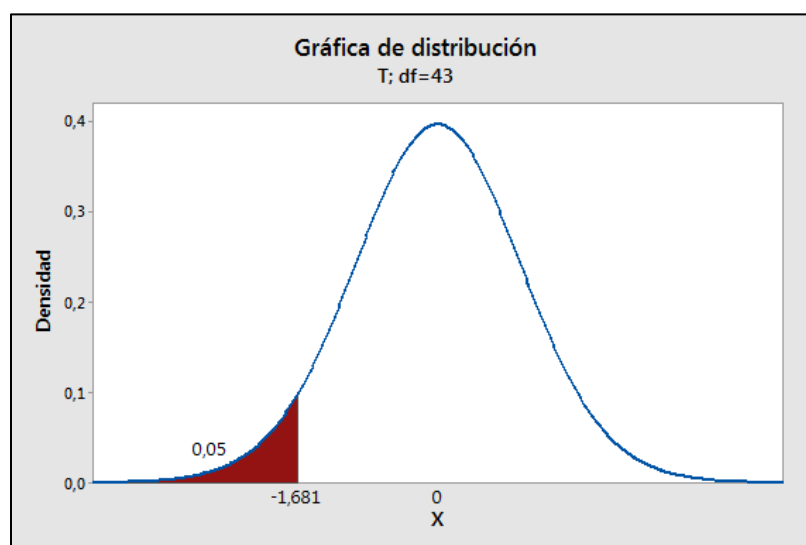
$\mu_1$  = Media del Núm. de Veces que el Usuario Accede a la Información al Día en la Pre-Prueba.

$\mu_2$  = Media del Núm. de Veces que el Usuario Accede a la Información al Día en la Post-Prueba.

$$H_0: \mu_1 \geq \mu_2$$

$$H_a: \mu_1 < \mu_2$$

**b) Criterios de Decisión**



**Figura 108.** Distribución de Probabilidad  $KPI_3$

Fuente: Elaboración Propia

**Tabla 53.** Resumen de Prueba t student para el KPI<sub>3</sub>

Fuente: Elaboración Propia

	Pre - Prueba	Post - Prueba
<b>Media (x)</b>	0.37	1.5
<b>Desviación Estándar (S)</b>	0.67	1.28
<b>Observaciones</b>	30	30
<b>Diferencia Hipotética de las Medias</b>	0	
<b>t_calculado: t<sub>c</sub></b>	-4.3	
<b>p-valor (una cola)</b>	0.000	
<b>Valor Crítico de una t<sub>α/2</sub> (una cola): t<sub>t</sub></b>	-1.681	

**c) Decisión Estadística**

Puesto que el valor-p = 0 < α = 0.05, los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula (H<sub>0</sub>) y considerar que la hipótesis alterna (H<sub>a</sub>) es cierta. La prueba resultó ser significativa.

**D. Contrastación para Porcentaje de Exactitud de la Información**

Se debe validar el impacto que tiene el desarrollo de la solución de Business Intelligence en el Porcentaje de Exactitud de la Información, llevado a cabo en la muestra. Se realiza una medición antes del desarrollo de la solución de BI (Pre-Prueba) y otra después del desarrollo de la solución de BI (Post-Prueba). La siguiente tabla contiene los Porcentajes de Exactitud de la Información para ambas muestras:

Pre - Prueba	85	81	87	75	82	85	83	82	86	79	87	81	76	86	76
	86	85	78	89	80	78	89	86	81	84	86	83	75	79	84

Post - Prueba	97	96	95	97	96	95	95	97	97	95	95	98	95	98	98
	98	97	98	98	96	98	97	98	98	97	97	96	97	97	98

**Hi:** El desarrollo de una solución de Business Intelligence aumenta el Porcentaje de Exactitud de la Información (Post-Prueba) con respecto a la muestra a la que no se aplicó (Pre-Prueba).

**Solución:**

**a) Planteamiento de la Hipótesis**

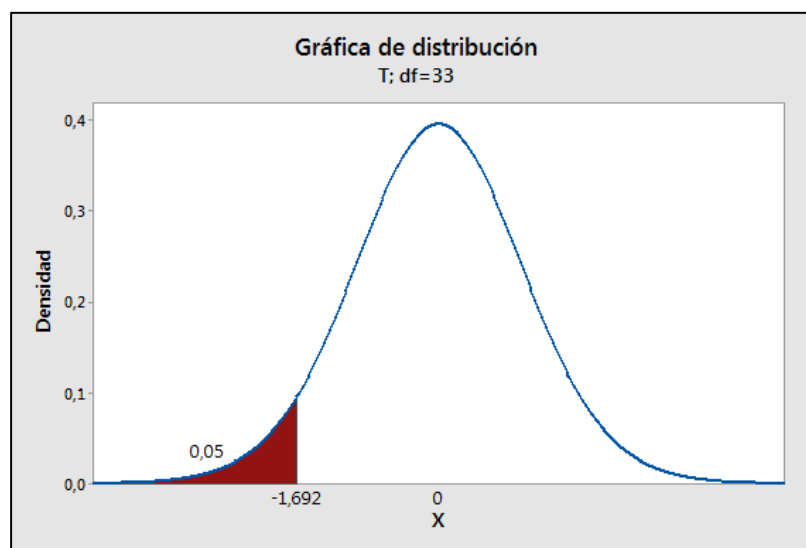
$\mu_1$  = Media del Porcentaje de Exactitud de la Información en la Pre-Prueba.

$\mu_2$  = Media del Porcentaje de Exactitud de la Información en la Post-Prueba.

$$H_0: \mu_1 \geq \mu_2$$

$$H_a: \mu_1 < \mu_2$$

**b) Criterios de Decisión**



**Figura 109.** Distribución de Probabilidad KPI<sub>4</sub>

Fuente: Elaboración Propia

**Tabla 54.** Resumen de Prueba t student para el KPI<sub>4</sub>

Fuente: Elaboración Propia

	Pre - Prueba	Post - Prueba
<b>Media (x)</b>	82.47	96.8
<b>Desviación Estándar (S)</b>	4.09	1.13
<b>Observaciones</b>	30	30
<b>Diferencia Hipotética de las Medias</b>	0	
<b>t_calculado: t<sub>c</sub></b>	-18.5	
<b>p-valor (una cola)</b>	0.000	
<b>Valor Crítico de una t<sub>α/2</sub> (una cola): t<sub>t</sub></b>	-1.692	

**c) Decisión Estadística**

Puesto que el valor- $p = 0 < \alpha = 0.05$ , los resultados proporcionan suficiente evidencia para rechazar la hipótesis nula ( $H_0$ ) y considerar que la hipótesis alterna ( $H_a$ ) es cierta. La prueba resultó ser significativa.

**CAPÍTULO V:**  
**CONCLUSIONES Y RECOMENDACIONES**

*Sólo una cosa convierte en imposible  
un sueño: el miedo a fracasar.*

**Paulo Coelho**

**1947**

## 5.1. CONCLUSIONES

- a) El desarrollo de la solución de Business Intelligence ayudó a mejorar el proceso de Toma de Decisiones del área de Rentas de la Municipalidad de Lurín.
- b) Se logró disminuir el tiempo empleado en la generación de reportes para el área de informática de la Municipalidad de Lurín, lo cual hizo que los tiempos de espera ya no fueran tan pesados.
- c) El tiempo que el usuario de la gerencia de rentas empleaba en el análisis de la información contenida en los reportes disminuyó, haciendo que las tomas de decisiones fueran más rápidas.
- d) El número de veces que el usuario de la gerencia de rentas accedía a la información logró aumentarse. Con esto se pudo aprovechar más la información que se tiene de las recaudaciones y deudas a la hora de tomar decisiones.
- e) El porcentaje de exactitud de la información contenida en los reportes aumentó. Esto significó realizar una mejor toma de decisiones, debido a que al ser más exacta la información mejor visión de la realidad se logra tener.
- f) Se logró aumentar el nivel de satisfacción que tenía el usuario respecto a la obtención de los reportes, haciendo así que se sienta más cómodo a la hora de solicitar la información que considera necesaria para tomar decisiones.

## 5.2. RECOMENDACIONES

- a) Antes de empezar con el desarrollo, es de suma importancia conocer muy bien el negocio y sobre todo el proceso con el que trabajamos. Con esto nos evitamos problemas a medio camino, los cuales demandan mucho tiempo en su solución y/o corrección.
- b) Seguir una determinada metodología aumenta las probabilidades de éxito de nuestro proyecto de desarrollo de una solución de Business Intelligence, pues nos brinda una guía paso a paso de las actividades a realizar. Así mismo cabe resaltar que esos pasos han sido ya utilizados y probados no solo a nivel nacional sino internacional.
- c) Resulta una buena idea apoyarse en tablas adicionales (tablas de apoyo) en caso el software gestor de base de datos no permita realizar consultas entre tablas de distintas BD y si no se cuenta con los recursos necesarios para adquirir otro software gestor de base de datos que sí lo permita. Además que esto permite agilizar el proceso de carga.
- d) Utilizar software libre es una forma de economizar a la hora de realizar un proyecto de BI, pues los programas son gratuitos para todos. Eso es una gran ventaja a comparación del software privado que por lo general cuesta una suma elevada de dinero. Incluso programas libres como la suite Pentaho han demostrado ya su robustez e integridad al momento de implementarse, dejando claro que es una herramienta totalmente confiable.



## REFERENCIAS BIBLIOGRÁFICAS

### Tesis:

1. Bazán, W. (2011). *Desarrollo de un Data Mart con Indicadores Financieros como soporte para la Toma de Decisiones en el departamento financiero del Gobierno Autónomo Descentralizado Municipal San Francisco de Milagro*. Tesis Post-Grado, Universidad Estatal de Milagro, San Francisco de Milagro, Ecuador.
2. Ocas, M. (2012). *Desarrollo de un Data Mart en el área de Administración y Finanzas de la Municipalidad Distrital de Baños del Inca*. Tesis Pre-Grado, Universidad Privada del Norte, Cajamarca, Perú.
3. Guillén, F. (2012). *Desarrollo de un Data Mart para mejorar la Toma de Decisiones en el área de Tesorería de la Municipalidad Provincial de Cajamarca*. Tesis de Pre-Grado, Universidad Privada del Norte, Cajamarca, Perú.
4. Chirán, M. (2013). *Modelo para la implementación de Inteligencia de Negocios que apoyen a la Toma de Decisiones en instituciones públicas de Protección Social*. Tesis Post-Grado, Universidad Central del Ecuador, Quito, Ecuador.
5. Vega, G. (2010). *Inteligencia de Negocios. Aplicación en la administración del Presupuesto en una empresa del Sector Público*. Tesis Post-Grado, Instituto Politécnico Nacional, México D.F., México.
6. Villanueva, A. (2008). *Análisis, Diseño e Implementación de un Data Warehouse de Soporte de Decisiones para un Hospital del Sistema de Salud Público*. Tesis Pre-Grado, Pontificia Universidad Católica del Perú, Lima, Perú.

### Libros:

1. VV. AA. (2007). *Dirección Estratégica*. España: Editorial Vértice.
2. Ramos, S. (2011). *Microsoft Business Intelligence: Vea el cubo medio lleno*. España: SolidQ.
3. Cano, J. (2007). *Business Intelligence: Competir con información*. España: Banesto.
4. Amaya, J. (2010). *Toma de Decisiones Gerenciales. Métodos cuantitativos para la administración*. Segunda edición. Colombia: Ecoe Ediciones.
5. Espíndola, C. (2005). *Análisis de Problemas y Toma de Decisiones*. Tercera Edición. México: Pearson Educación.

6. Kimball, R., Reeves, L., Ross, M., Thornthwaite, W. (2008). *The Data Warehouse Lifecycle Toolkit*. Estados Unidos: Wiley.
7. Kimball, R., Ross, M. (2002). *The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling*. Segunda Edición. Estados Unidos: John Wiley & Sons, Inc.
8. Mundy, J., Thornthwaite, W., Kimball, R. (2006). *The Microsoft Data Warehouse Toolkit : With SQL Server 2005 and the Microsoft Business Intelligence Toolset*. Estados Unidos: John Wiley & Sons.
9. Inmon, W. (2003). *Building the Data Warehouse*. Tercera edición. Estados Unidos: John Wiley & Sons, Inc.
10. Da Rosa, F., Heinz, F. (2007). *Guía práctica sobre software libre. Su selección y aplicación local en América Latina y El Caribe*. Uruguay: Mastergraf.
11. Stallman, R. (2004). *Software Libre para una Sociedad Libre*. Traducido por: Jaron Rowan, Diego Sanz Paratcha, Laura Trinidad. España: Traficantes de Sueños.
12. Bouman, R., Dongen, J. (2009). *Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL*. Estados Unidos: Wiley Publishing, Inc.
13. Ávila, H. (2006). *Introducción a la Metodología de la Investigación*. España: Edición Electrónica.

#### **Artículos de Revistas Electrónicas:**

1. Chacín, L. (2010). *Tecnología de la información como apoyo al proceso para la toma de decisiones gerenciales en organizaciones del sector eléctrico de Venezuela*. Revista Espacios. Volumen 31. Obtenida el 15 de febrero de 2015, de <http://www.revistaespacios.com/a10v31n02/10310233.html>
2. Wolff, C. (1999). *La Tecnología Datawarehousing*. Revista Ingeniería Informática. Edición número 3. Obtenida el 8 de setiembre de 2014, de <http://www.inf.udec.cl/~revista/ediciones/edicion3/cwolff.PDF>
3. Shaker, H., Abdeltawad, M., Ali, H. (2011) *A proposed model for data warehouse ETL processes*. Revista de la Universidad de King Saud – Computación y Ciencias de la Información. Volumen 23. Obtenida el 6 de junio del 2014, de <http://www-users.cs.umn.edu/~hendawi/materials/A%20proposed%20model%20for%20data%20warehouse%20ETL%20processes.pdf>

## Monografías de Internet:

1. Villanueva, L. (2015). *La toma de decisiones en la organización y el gran valor del profesional de la información en su desarrollo*. Obtenida el 15 de febrero de 2015, de <http://www.infotecarios.com/la-toma-de-decisiones-en-la-organizacion-y-el-gran-valor-del-profesional-de-la-informacion-en-su-desarrollo/>
2. Dirección General de Presupuesto Público, Ministerio de Economía y Finanzas. (2011). *Manual para la mejora de la recaudación del impuesto predial en las municipalidades*. Obtenida el 15 de febrero de 2015, de [http://www2.congreso.gob.pe/sicr/cendocbib/con3\\_uibd.nsf/3BB2BE2F7ACB49DD052578F700568A63/\\$FILE/0441957001313598606.pdf](http://www2.congreso.gob.pe/sicr/cendocbib/con3_uibd.nsf/3BB2BE2F7ACB49DD052578F700568A63/$FILE/0441957001313598606.pdf)
3. Umanzor, C. (2011). *¿Cómo enseñar a tomar decisiones acertadas?*. Obtenida el 9 de junio de 2014, de [http://www.fundesyram.info/document/PDFPUB/TOMA\\_DECISIONES.pdf](http://www.fundesyram.info/document/PDFPUB/TOMA_DECISIONES.pdf)
4. Rivadera, G. (2010). *La metodología de Kimball para el diseño de almacenes de datos (Data Warehouses)*. Obtenida el 9 de junio de 2014, de <http://www.ucasal.edu.ar/hm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>
5. Gutiérrez, G. (2014). *Teoría de la toma de decisiones. Definición, etapas y tipos*. Obtenida el 15 de febrero del 2015, de <http://www.gestiopolis.com/teoria-de-la-toma-de-decisiones-definicion-etapas-y-tipos/>
6. Enciclopedia Cubana (2011). *Plataforma Pentaho*. Obtenida el 9 de junio de 2014, de [http://www.ecured.cu/index.php/Plataforma\\_Pentaho](http://www.ecured.cu/index.php/Plataforma_Pentaho)
7. Ciudadanos al Día. (2004). *Eficiencia en la Recaudación Municipal: El Caso de los SATs*. Obtenida el 11 de noviembre de 2014, de [http://www.ciudadanosaldia.org/images/investigacion\\_y\\_publicaciones/informes\\_cad/InformeCAD\\_35\\_Eficiencia\\_SATs\\_20Dic04\\_1215h.pdf](http://www.ciudadanosaldia.org/images/investigacion_y_publicaciones/informes_cad/InformeCAD_35_Eficiencia_SATs_20Dic04_1215h.pdf)

## APÉNDICES

### APÉNDICE I: MATRIZ DE CONSISTENCIA

PROBLEMA	OBJETIVO GENERAL	HIPÓTESIS	VARIABLES	INDICADORES	ÍNDICES	METODOLOGÍA
¿En qué medida el uso de Business Intelligence mejorará el Proceso de Toma de Decisiones en el área de Rentas de la Municipalidad de Lurín?	Desarrollar una solución de Business Intelligence para mejorar el proceso de Toma de Decisiones en el área de rentas de la Municipalidad de Lurín.	El desarrollo de una solución de Business Intelligence mejorará el proceso de Toma de Decisiones del área de Rentas de la Municipalidad de Lurín.	Variable Independiente	Presencia - Ausencia	No, Si	TIPO DE INVESTIGACION:  - Cuantitativa  - Aplicada  NIVEL DE INVESTIGACIÓN:  - Correlacional  DISEÑO DE INVESTIGACIÓN: Pre Experimental  POBLACIÓN:  Todos los Procesos de Toma de Decisiones del Área de Rentas la
			Variable Dependiente			
				• Solución de Business Intelligence	Tiempo que el usuario emplea en el análisis de la información.	
• Proceso de Toma de Decisiones del Área de Rentas de la Municipalidad de Lurín.	Número de veces que el usuario accede a la información al día	[1...5]				

				Porcentaje de exactitud de la información	[1...100]	Municipalidad de Lurín.
				Nivel de satisfacción del usuario	[Bajo, Regular, Alto]	MUESTRA: Se trabajará con un muestreo intencional (no aleatorio) y se tomará una muestra de 30 tomas de decisiones al considerarse una cantidad razonable y manejable.

## APÉNDICE II: ARTÍCULO CIENTÍFICO

---

### DESARROLLO DE UNA SOLUCIÓN DE BUSINESS INTELLIGENCE PARA MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE RENTAS DE LA MUNICIPALIDAD DE LURÍN

---

#### DEVELOPMENT OF A BUSINESS INTELLIGENCE SOLUTION FOR IMPROVING THE DECISION MAKING PROCESS IN THE RENTS AREA OF THE LURIN MUNICIPALITY

---

**Felix Julio Vargas Chumpitaz**  
Universidad Autónoma del Perú  
Facultad de Ciencias de Gestión  
[felix.vargas369@gmail.com](mailto:felix.vargas369@gmail.com)

---

#### **RESUMEN**

El presente proyecto de tesis consiste en el desarrollo de una solución de Business Intelligence, herramienta que ayudará a mejorar el proceso de toma de decisiones en el área de rentas de la municipalidad de Lurín.

Si bien es cierto que la gerencia de rentas de la municipalidad de Lurín sí usa la información del sistema transaccional para tomar decisiones, el proceso de obtención de reportes resulta demasiado tardío y demanda mucho esfuerzo por parte del personal de informática.

Ante esta situación fue que se planteó el desarrollo de la solución de Business Intelligence, pues haría que el proceso sea mucho más rápido y con menor esfuerzo invertido. Para el desarrollo se empleó la metodología de Ralph Kimball, la cual se adaptó más a nuestro caso de estudio pues está enfocada únicamente a una parte de la empresa o institución.

La solución final para el usuario consistirá en una serie de reportes acerca del estado actual e histórico de las recaudaciones y deudas realizadas por el área, para que así se puedan tomar decisiones basándose de información real.

**Palabras clave:** Inteligencia de negocios, Data Mart, Ralph Kimball, Pentaho, Software libre, Recaudación, Reportes.

#### **ABSTRACT**

This thesis project consist of the development of a Business Intelligence solution, tool that will help improve the process of decision making in the rents area of the Lurin municipality.

While it is true that the management of rents of the Lurin municipality use the information of transactional system to make decisions, the process of obtaining reports is too late and demand much effort from IT staff.

Given this situation was that the development of a Business Intelligence solution was raised because it would make the process much faster and with less effort involved. To development, Ralph Kimball methodology was used, which is more adapted to our case study because it is focused only part of the company or institution.

The final solution for the user consist of a series of reports on the current and historical status of the collections and debts made by the area, so that decisions can be made based on real information.

**Key Words:** Business Intelligence, Data Mart, Ralph Kimball, Pentaho, Free Software, Collection, Reports.

## 1. INTRODUCCIÓN

La información va tomando cada vez más relevancia, tanto así que día tras día va tomando más importancia como activo para las empresas e instituciones, ya sean públicas o privadas. Pero la información no se obtiene ya “elaborada”, sino que la conseguimos en forma de datos, los cuales deben ser registrados mediante sistemas transaccionales u hojas de cálculo y posteriormente analizados para así poder obtener información útil para la empresa o institución. Como se mencionó, la información va tomando cada vez más importancia. Así podemos inferir que los datos también son importantes para nosotros, pero dicha importancia no va necesariamente relacionada a la cantidad ya que se puede tener una cantidad descomunal de datos en nuestros sistemas pero si no sabemos analizarlas ni explotarlas no nos servirá, será un activo totalmente inútil.

No saber qué hacer con los datos puede ocurrir en cualquier empresa o institución sin importar el rubro en el cual se desempeñe ya que toda empresa maneja algunos datos ya sea de proveedores, clientes, ventas, producción, etc. El problema en sí es tener muchos datos y no saber aprovecharlos, es decir no apoyarnos de ellos para obtener ventajas competitivas frente a las demás empresas del mercado. Las ventajas que se pueden obtener son por ejemplo conocer qué sector de la población consume más un determinado producto, qué producto se vende más, qué vendedor es el más eficiente o qué tienda es la que factura más al mes. Estas ventajas no son las únicas que podemos obtener, es solamente un ejemplo. Por otro lado tampoco tienen que ser precisamente éstas, ya que pueden variar de acuerdo al rubro de la empresa. Como se puede apreciar, los ejemplos se refieren a alguna empresa de venta de productos a la cual le puede interesar qué local factura más al mes, pero en el caso de un ente recaudador le interesaría qué local recauda más dinero también al mes. Esta es una muestra de que sin importar el rubro al cual una empresa o institución se dedique, siempre se puede obtener ventajas competitivas gracias al buen uso de los datos obtenidos.

El presente proyecto consiste en el desarrollo de una solución de business intelligence para mejorar la toma de decisiones en el área de rentas de la municipalidad de Lurín. Esto permitirá que la gerencia de rentas pueda conocer cómo va el negocio, es decir le brindará una visión del rumbo que se va tomando. Y más allá de conocer la realidad del área, la solución permitirá tomar decisiones más acertadas

## 2. CONTENIDO

Se ha integrado teorías referentes a Business Intelligence, Metodología de Ralph Kimball. Además teoría sobre el proceso de toma de decisiones y los tributos municipales.

### 2.1 Fundamentación Teórica

#### ¿Qué es Business Intelligence?

El término Business Intelligence (que traducido al español sería Inteligencia de Negocios) es definido como un conjunto de estrategias y tecnologías que ayudan a convertir los datos en información de calidad, y dicha información en conocimiento que nos permita una toma de decisiones más acertada y nos ayude así a mejorar nuestra competitividad<sup>1</sup>.

El business intelligence cuenta con componentes, los cuales son:

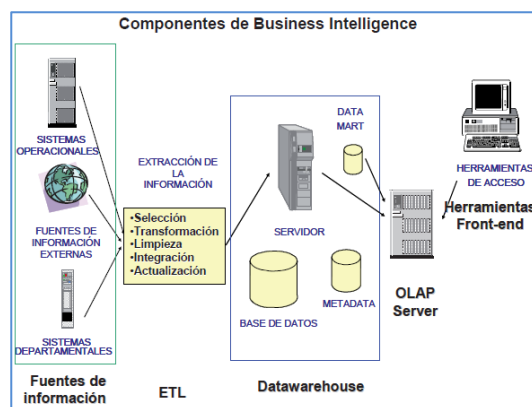


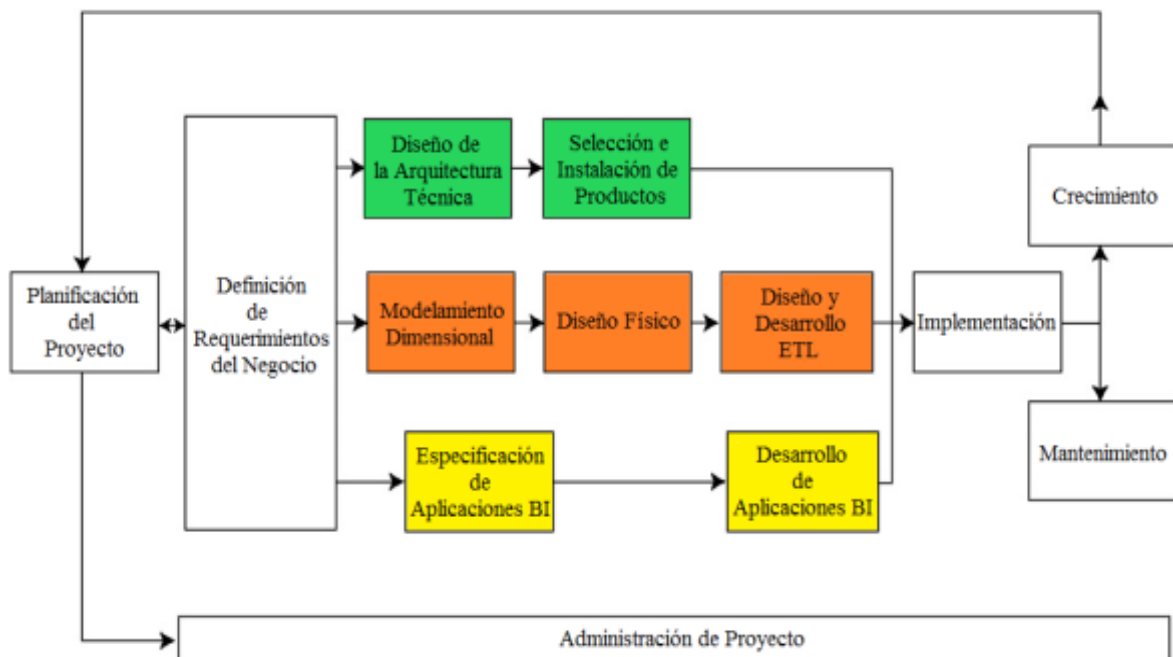
Figura 01. Componentes de Business Intelligence

#### Metodología de Ralph Kimball

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida tiene un enfoque basado en cuatro principios básicos:

- Centrarse en el negocio: Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado. Usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- Construir una estructura de información: Diseñar una base de información única, integrada, fácil de usar y de alto rendimiento que cumpla con la amplia gama de requisitos de negocio que ha identificado en la empresa.

<sup>1</sup> Ramos, S., Microsoft Business Intelligence: Vea el cubo medio lleno, Ed. SolidQ, España, 2011, p. 9



**Figura 02.** Ciclo de Vida Metodología de Kimball

- Realizar entregas en incrementos significativos: crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos.
- Ofrecer la solución completa: proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación<sup>2</sup>. Las fases de esta metodología (ciclo de vida) se muestran en la Figura 01.

## 2.2 Pasos de la Metodología

- **Planificación del Proyecto:** En este proceso se determina el propósito del proyecto de DW/BI, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.
- **Definición de los Requerimientos del Negocio:** La definición de los requerimientos es en gran medida un proceso de entrevistar al personal de negocio y técnico.
- **Modelado Dimensional:** La creación de un modelo dimensional es un proceso dinámico y altamente iterativo. El proceso de diseño

comienza con un modelo dimensional de alto nivel.

- **Diseño Físico:** El diseño físico de la base de datos se centra en la definición de las estructuras físicas necesarias para soportar el diseño lógico de la base de datos.
- **Diseño y Desarrollo ETL:** Las herramientas de extracción, transformación y carga (ETL), son las piezas de software responsables de la extracción de datos de varias fuentes, limpiarlas, personalizarlas, integrarlas e insertarlas en el Data Warehouse.
- **Diseño de la Arquitectura Técnica:** Los entornos del almacén de datos requieren la integración de numerosas tecnologías. El diseño de la arquitectura técnica establece el marco y la visión de la arquitectura.
- **Selección e Instalación de Productos:** Usando el diseño de arquitectura técnica, específicos componentes de estructura como plataforma de hardware, sistema de gestión de base de datos, herramientas de data staging o herramientas de acceso de datos, necesitan ser evaluados y seleccionados.
- **Especificación de Aplicaciones BI:** Las especificaciones de las aplicaciones describen la plantilla de los reportes, los parámetros controlados por el usuario, y los cálculos requeridos.
- **Desarrollo de Aplicaciones BI:** Después de la especificación de las aplicaciones, el desarrollo de aplicaciones de usuario final implica la configuración de la herramienta de metadatos y la construcción de los reportes especificados.
- **Implementación:** La implementación representa la unión de la tecnología, los datos y las aplicaciones de usuario final, siendo así esta

<sup>2</sup> Mundy, J., et Al., The Microsoft Data Warehouse Toolkit. Ed. Wiley, EE. UU., 2011, p. xxxv



unión accesible desde el escritorio de los usuarios del negocio.

- **Mantenimiento y Crecimiento:** Mucho trabajo queda después de la implementación inicial del almacén de datos. Es necesario a continuar centrándose en sus usuarios del negocio, proporcionándoles apoyo y capacitación.

### 3. DESARROLLO DE LA SOLUCIÓN DE BUSINESS INTELLIGENCE

#### 3.1 Planificación del Proyecto

##### 3.1.1 Descripción del Proyecto

La solución propuesta para el área de rentas de la municipalidad de Lurín le permitirá a la gerencia tener conocimiento acerca del estado actual e histórico de las recaudaciones y deudas (de arbitrios municipales e impuestos prediales) de los contribuyentes del distrito por medio de reportes fáciles de entender que contendrán datos confiables. Con esto, las diversas acciones que se realicen (por ejemplo dar facilidades de pago a las zonas que deben más) serán siempre basadas en información real y en sus respectivos indicadores. Para esto se desarrollará una solución de Business Intelligence para el proceso de recaudación del área de rentas, luego se efectuará un análisis dimensional, el cual dará paso a la obtención de cubos dimensionales u OLAP. Finalmente se llevará a cabo el diseño de los reportes para la gerencia

##### 3.1.2 Objetivos del Proyecto

Reducir el tiempo y el esfuerzo empleado en la generación de los reportes relacionados a la recaudación, debido a que actualmente esta actividad es bastante tardía y demanda mucho esfuerzo.

Brindar a la gerencia de rentas una forma de visualización del estado actual e histórico de las recaudaciones realizadas y deudas ocurridas, de manera sencilla y confiable, basado en indicadores.

##### 3.1.3 Alcance del Proyecto

- Desarrollo de un Data Mart para el área de rentas de la municipalidad de Lurín con los datos obtenidos del sistema transaccional de recaudación.

- Elaboración del cubo OLAP mediante la base de datos dimensional.

- Diseño de reportes:

- ✓ Montos recaudados de impuesto predial por cuotas - fechas.
- ✓ Montos recaudados de arbitrios municipales por cuotas - fechas y zonas.
- ✓ Montos adeudados de impuesto predial por cuotas.
- ✓ Montos adeudados de arbitrios municipales por cuotas y zonas..

- ✓ Contribuyentes que pagan los mayores montos.
- ✓ Contribuyentes que adeudan los mayores montos.
- ✓ Cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial.
- ✓ Estado de la recaudación según KPI

#### 3.1.4 Stakeholders

Los principales interesados del proyecto vienen a ser el gerente de rentas, el subgerente de informática y el autor.

**Tabla 01.** Principales Stakeholders del Proyecto

STAKEHOLDER	CARGO	FUNCIÓN
Felix Vargas Chumpitaz	Jefe de Proyecto	Es la persona que se encarga de la planificación y desarrollo del proyecto, y quien se hace responsable de que el mismo se lleve a cabo de manera óptima.
Juan Carlos Cervantes Aguirre	Subgerente de Informática	Tiene la labor de asegurar el correcto funcionamiento de los componentes de cómputo de propiedad de la Municipalidad de Lurín, así como también asesorar a los distintos usuarios.
Pedro Rueda Torrejón	Gerente de Rentas	Cumple la función de dirigir el entero proceso de recaudación de impuestos y arbitrios (de todo tipo) de la Municipalidad.

#### 3.2 Definición de los Requerimientos del Negocio

##### 3.2.1 Proceso de Negocio y Temas Analíticos

**Tabla 02.** Temas Analíticos basados en Entrevistas

TEMAS ANALÍTICOS	ANÁLISIS SOLICITADOS O INFERIDOS	PROCESO DE NEGOCIO COMPATIBLE	COMENTARIOS
Planificación de Recaudación	Información histórica de la recaudación de impuesto predial según cuotas	Recaudación de impuesto predial	Por cuotas (4) con su respectivo año
	Información histórica de la recaudación de arbitrios municipales según cuotas	Recaudación de arbitrios municipales	Por cuotas (12) con su respectivo año.
	Información histórica de la recaudación de arbitrios municipales según zonas	Recaudación de arbitrios municipales	Por microzonas y zonas
	Información de las deudas de impuesto predial por cuotas	Recaudación de impuesto predial	Por cuotas (4) con su respectivo año
	Información de las deudas de arbitrios municipales por cuotas	Recaudación de arbitrios municipales	Por cuotas (12) con su respectivo año.
	Información de las deudas de arbitrios municipales por zonas	Recaudación de arbitrios municipales	Por microzonas y zonas
	Información histórica de las recaudaciones por fechas.	Recaudación de arbitrios e impuestos	Por años, trimestres y meses.
Reporte de Recaudación	Contribuyentes que pagan mayores/menores montos	Recaudación de arbitrios e impuestos	Por cuotas y años.
	Cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial	Recaudación de arbitrios e impuestos	Por cuotas y fechas.
	Estado de las recaudaciones según KPIs	Recaudación de arbitrios e impuestos	Por cuotas y años.

**Tabla 03.** Procesos de Negocio basados en Entrevistas

LETRA	PROCESO DE NEGOCIO	TEMAS ANALÍTICOS COMPATIBLES
A	Recaudación de impuesto predial y arbitrios municipales	Planificación de la recaudación, Reporte de la recaudación, Análisis de la recaudación

**3.2.2 Requerimientos**

**Tabla 04.** Requerimientos

CÓDIGO	REQUERIMIENTO
REQ01	Mostrar los montos de impuesto predial recaudados por cuotas y fechas.
REQ02	Mostrar los montos de arbitrios municipales recaudados por cuotas y fechas.
REQ03	Mostrar los montos de arbitrios municipales recaudados por zonas y microzonas.
REQ04	Mostrar los montos de impuesto predial adeudados por cuotas.
REQ05	Mostrar los montos de arbitrios municipales adeudados por cuotas.
REQ06	Mostrar los montos de arbitrios municipales adeudados por zonas y microzonas.
REQ07	Mostrar los contribuyentes que pagan las mayores cantidades de dinero por año fiscal.
REQ08	Mostrar los contribuyentes que adeudan las mayores cantidades de dinero por año fiscal.
REQ09	Mostrar la cantidad de contribuyentes que cancelan los arbitrios municipales e impuesto predial por cuotas y fechas.
REQ10	Mostrar el estado de la recaudación de los arbitrios municipales e impuesto predial según los KPIs.

**3.2.3 Hoja de Gestión**

**Tabla 05.** Hoja de Gestión

Hoja de Gestión			
Proceso	Recaudación de Impuesto Predial y Arbitrios Municipales		
Objetivo	Obtener recursos monetarios a partir de la recaudación del impuesto predial y de los arbitrios municipales a los contribuyentes a fin de solventar los gastos de las actividades que realice la municipalidad en favor del distrito.		
Estrategia	Información en lugares concurridos sobre los modos de pago. Beneficios por pronto pago. Amnistías. Campañas tributarias los fines de semana.		
Indicador	Indicadores	Medidas	Estados
	Montos por importe recaudados	sum(importe)	>90% 51% - 89% <50%
	Montos por derecho de emisión recaudados	sum(deremi)	>90% 51% - 89% <50%
	Montos por mora recaudados	sum(morcan)	>90% 51% - 89% <50%
	Montos totales recaudados	sum(importe+deremi+morcan)	>90% 51% - 89% <50%

**3.3 Modelado Dimensional**

**3.3.1 Dimensiones**

**Tabla 06.** Dimensiones

DIMENSIÓN	DESCRIPCIÓN
Contribuyente	Almacena a todos los contribuyentes del distrito de Lurín registrados en el sistema transaccional.
Cuota	Almacena los números de las cuotas de los pagos de arbitrios municipales e impuesto predial y sus años.
Tipo Pago	Almacena los tipos de pago con los que se desea trabajar (Impuesto predial y Arbitrios municipales).
Zona	Almacena todas las microzonas del distrito de Lurín registradas en el sistema transaccional con la zona a la cual pertenece.
Tiempo	Almacena las fechas en las que se realizaron los pagos de impuesto predial y arbitrios municipales.
Estado	Almacena los estados de los pagos de arbitrios municipales e impuesto predial.

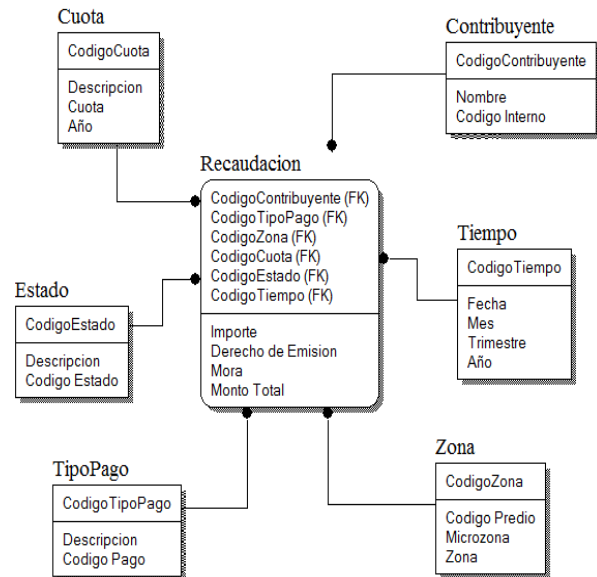
**3.3.2 Hechos**

TABLA DE HECHOS	DESCRIPCIÓN
Recaudación de Impuesto Predial y Arbitrios Municipales	Tabla de hechos referida a los montos de impuesto predial y arbitrios municipales recaudados y por recaudar a los contribuyentes del distrito de Lurín.

**3.3.3 Medidas**

MEDIDA	FÓRMULA
Importe	$\sum(\text{importe})$
Derecho de Emisión	$\sum(\text{derecho de emisión})$
Mora	$\sum(\text{mora})$
Monto Total	$\sum(\text{importe} + \text{derecho de emisión} + \text{mora})$

**3.3.4 Diseño del Modelo Estrella**



**Figura03.** Diseño Modelo Estrella

### 3.4 Diseño de la Arquitectura Técnica

#### 3.4.1 Arquitectura Técnica

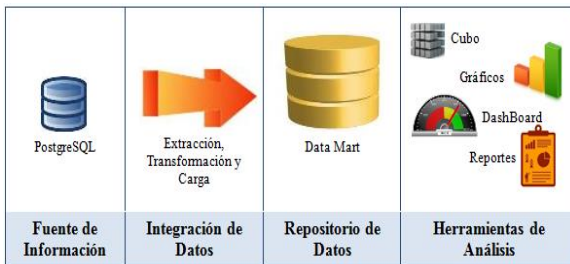


Figura 04. Arquitectura Técnica

### 3.5 Selección e Instalación de Productos

#### 3.5.1 Herramientas Seleccionadas

Luego de la evaluación de las alternativas de herramientas a usarse para el desarrollo de la solución BI, se vio que la mejor opción para nuestro caso es Pentaho, sobre todo porque es una suite completa para desarrollar soluciones de inteligencia de negocios, es software libre y ya tiene antecedentes de casos de éxito a la hora de su uso.

Pentaho cuenta con muchas herramientas para el desarrollo de soluciones BI (muchas de las cuales realizan las mismas acciones), las cuales pueden instalarse por separado y de acuerdo a las necesidades del usuario. De igual manera cuenta con plugins para aumentar las funcionalidades del servidor BI.

### 3.6 Diseño Físico

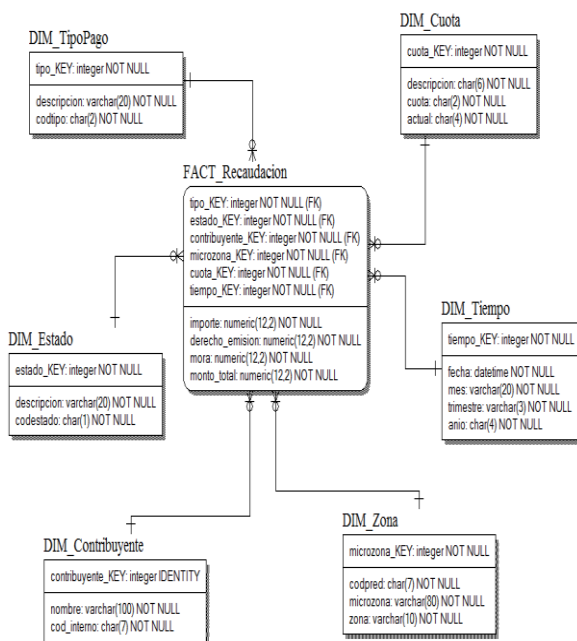


Figura 05. Diseño Modelo Físico

### 3.7 Diseño y Desarrollo ETL

#### 3.7.1 Extracción

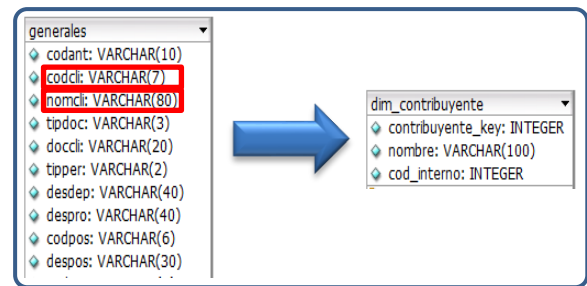


Figura 06. Extracción: dim\_contribuyente

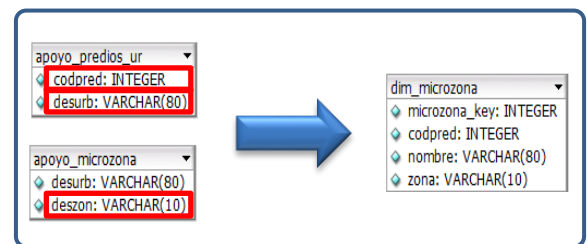


Figura 07. Extracción: dim\_microzona

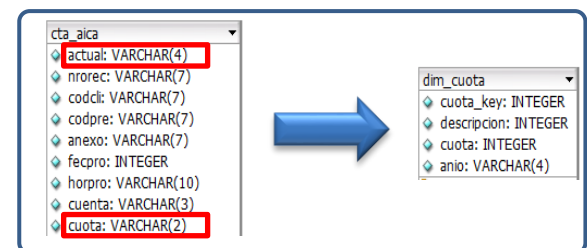


Figura 08. Extracción: dim\_cuota

#### 3.7.2 Transformación

```
SELECT nomcli, codcli FROM generales
where nomcli not in ('') order by codcli
```

```
select apur.codpred, apur.desurb, am.deszon
from apoyo_predios_ur apur
inner join apoyo_microzona am on am.desurb = apur.desurb
```

#### 3.7.3 Carga

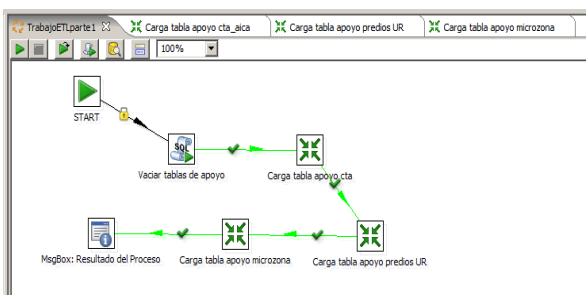


Figura 09. ETL parte 1

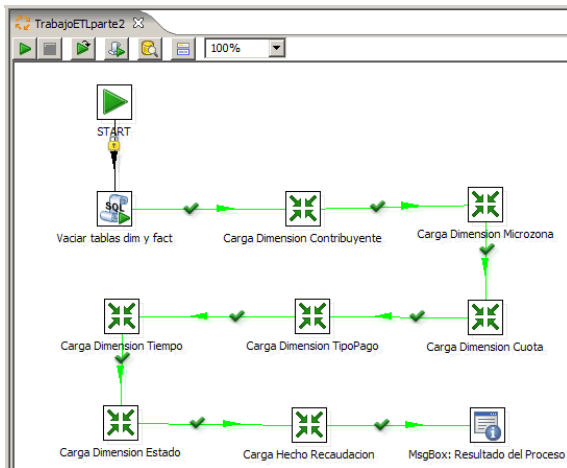


Figura 10. ETL parte 2

### 3.9 Diseño Cubo OLAP

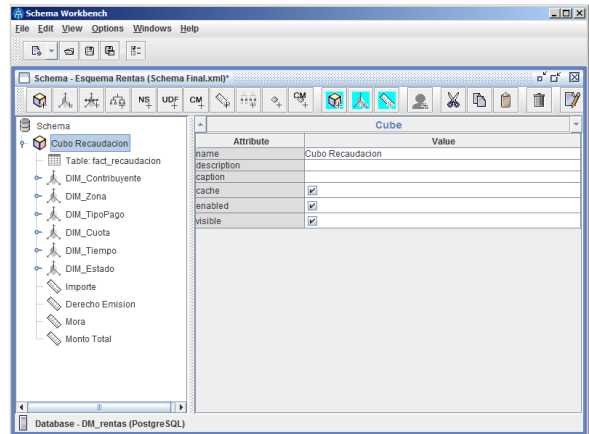


Figura 14. Cubo OLAP

### 3.8 Especificación de Aplicaciones BI

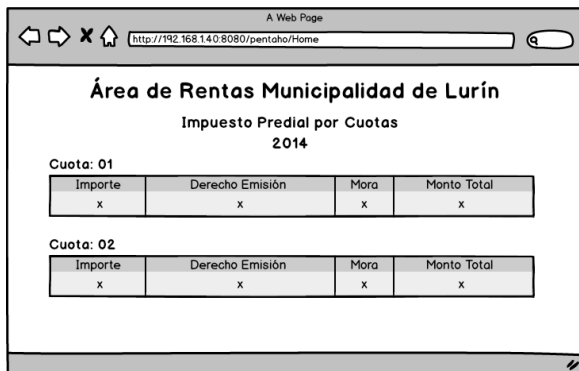


Figura 11. Prototipo Reporte a Medida

### 3.10 Desarrollo de Aplicaciones BI

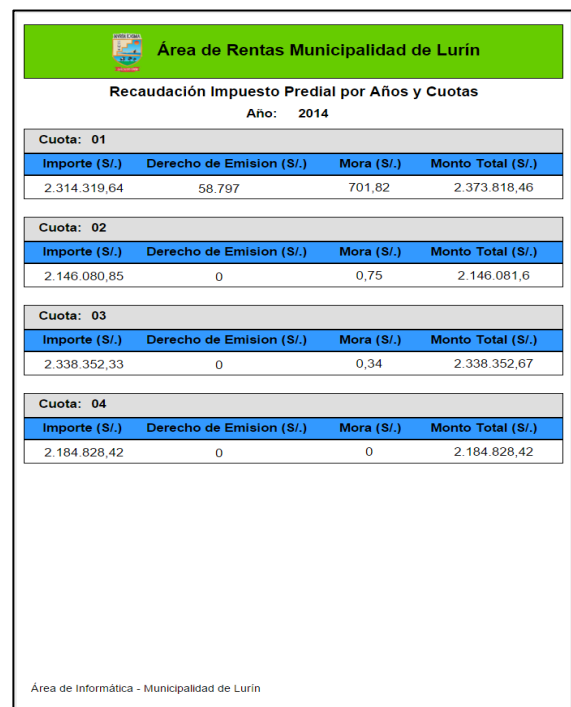


Figura 15. Reporte Recaudación Impuesto Predial

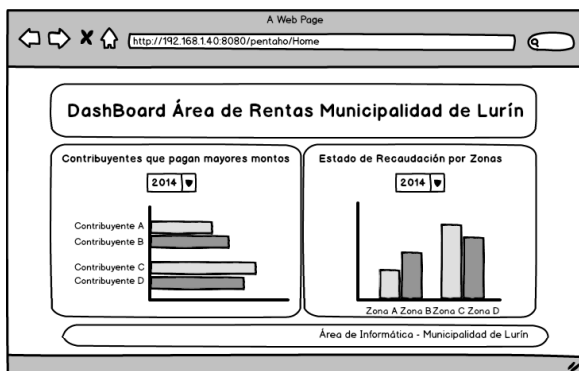


Figura 12. Prototipo Dashboard



Figura 16. Dashboard Recaudación

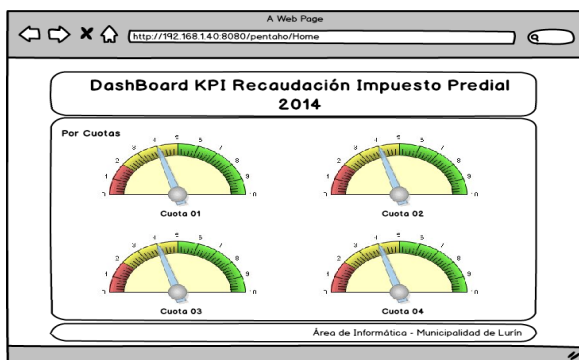


Figura 13. Prototipo Dashboard KPI

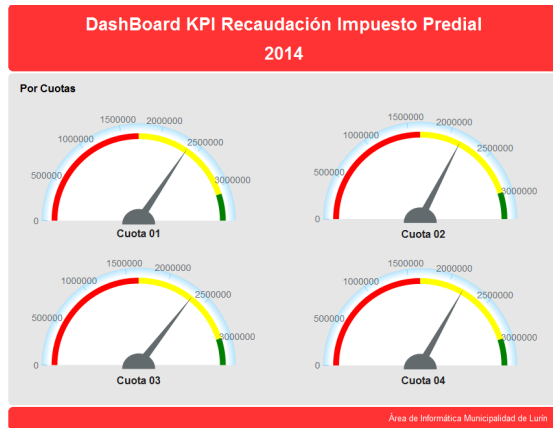


Figura 17. Dashboard KPI

### 3.11 Implementación

Tabla 07. Verificación de la Tecnología

TIPO	RECURSO	¿SE TIENE EL RECURSO?	COMENTARIOS
Hardware	Servidor físico HP	SI	-
	PC Intel	SI	-
Software	Oracle VM VirtualBox 4.3.12	SI	-
	Servidor virtualizado	SI	-
	PostgreSQL 9.3	SI	-
	Pentaho Data Integration (Kettle) 5.2.0.0	SI	-
	Pentaho Design Studio 4.0.0	SI	-
	Pentaho Schema Workbench 3.8.0.0	SI	-
	JPivot	SI	-
	Pentaho Report Designer 5.2.0.0	SI	-
	Pentaho BI Server 5.2.0.0	SI	-
	Community Dashboard Editor (CDE) 5.2.0.0	SI	-
	Ivy Dashboard Components (IvyDC) 0.0.4	SI	Fue descargado e instalado pero al final no fue utilizado.

### 3.12 Mantenimiento y Crecimiento

Tabla 08. Pruebas de Funcionamiento del Sistema

FUNCIONALIDAD	RESULTADO	TIEMPO DE DEMORA (s)	COMENTARIOS
Inicio de servidor BI	Satisfactorio	43	Solo es necesario ejecutarlo si se apaga el servidor virtual.
Autenticación de usuario en el sistema	Satisfactorio	10	-
Ejecución del proceso ETL (parte1)	Satisfactorio	420	Desde la consola de usuario.
Ejecución del proceso ETL (parte2)	Satisfactorio	1620	Desde la consola de usuario.
Ejecución de los reportes a medida	Satisfactorio	45	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.
Ejecución de los dashboards	Satisfactorio	25	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.
Ejecución de los cubos	Satisfactorio	35	Solo la primera vez que se ejecuta. Después solo 1 a 2 segundos.

## 4. MÉTODOS Y MATERIALES

### 4.1 Métodos

Se utilizó: la Observación Directa, la revisión de documentos (en papel y digitalizados), y se conversó con el gerente y personal del área de informática

### 4.2 Materiales

Se tomó como muestra para la investigación a 30 procesos de toma de decisiones en el área de rentas (muestreo intencional, no aleatorio). Se tomaron medidas durante el año 2014.

Se utiliza un tipo especial de investigación, diseñado por el investigador, que considera aspectos de los diseños experimentales y no experimentales. Se trata del **Diseño Experimental Verdadero tipo Panel sin Grupo de Control**, el cual se aplica y explica en detalle en la contrastación de la hipótesis.

$$G_e \quad O_1 \quad X \quad O_2$$

Donde:

$G_e$  → Grupo Experimental (la Muestra).

$O_1$  → Medición de los valores de los Indicadores de la Variable Dependiente antes del desarrollo de la solución BI.

$X$  → Desarrollo de la solución BI en el área de rentas.

$O_2$  → Medición de los valores de los Indicadores de la Variable Dependiente después del desarrollo de la solución BI.

## 5. RESULTADOS

A continuación se presentan las medias de los KPIs para la PrePrueba y PostPrueba: Resultados numéricos.

Tabla 09. Resultados de los Indicadores

Indicador	Pre-Prueba (Media: $x_1$ )	Post-Prueba (Media: $x_2$ )	Comentario
Tiempo empleado en la generación de reportes	30.47 min	1.37 min	---
Tiempo que el usuario emplea en el análisis de la información	120.37 min	20.4 min	---
Números de veces que el usuario accede a la información al día	0.37 veces	1.5 veces	---
Porcentaje de exactitud de la información	82.47 %	96.8%	---
Nivel de satisfacción del usuario	---	---	No contrastado. Indicador cualitativo.

## KPI<sub>1</sub> Tiempo empleado en la generación de reportes

Tabla 10. KPI Tiempo de Generación de Reportes

	Pre-Prueba	Post-Prueba		
32	1	1	1	1
28	2	2	2	2
30	2	2	2	2
34	1	1	1	1
29	1	1	1	1
33	1	1	1	1
30	1	1	1	1
28	2	2	2	2
34	1	1	1	1
30	2	2	2	2
28	1	1	1	1
34	1	1	1	1
30	1	1	1	1
28	2	2	2	2
27	1	1	1	1
31	2	2	2	2
33	2	2	2	2
29	1	1	1	1
30	1	1	1	1
27	1	1	1	1
34	1	1	1	1
30	1	1	1	1
27	1	1	1	1
32	2	2	2	2
30	2	2	2	2
28	1	1	1	1
31	1	1	1	1
34	2	2	2	2
33	2	2	2	2
30	1	1	1	1
<b>Promedio</b>	30,47	1,37		
<b>Meta planteada</b>		5		
<b>Nº menor a promedio</b>		19	30	30
<b>% menor a promedio</b>		63.3	100	100

- El 63.3% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que su tiempo promedio.
- El 100.0% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que la meta planteada.
- El 100.0% de los Tiempos Empleados en la Generación de Reportes en la Post-Prueba fueron menores que el tiempo promedio en la Pre-Prueba.

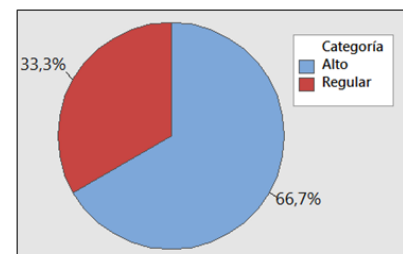
- La distancia "promedio" de las observaciones individuales de los Tiempos Empleados en la Generación de Reportes con respecto a la media es de 0.49 minutos.
- La Curtosis = -1.78 indica que tenemos datos de tiempos con picos muy bajos.
- La Asimetría = 0.58 indica que la mayoría de los Tiempos Empleados en la Generación de Reportes son bajos.
- El 1er Cuartil (Q1) = 1.000 minutos, indica que el 25% de los Tiempos Empleados en la Generación de Reportes son menores que o igual a este valor.
- El 3er Cuartil (Q3) = 2.000 minutos, indica que el 75% de los Tiempos Empleados en la Generación de Reportes son menores que o igual a este valor.

## KPI<sub>5</sub> Nivel de entendimiento de los reportes

Nro. Medición	1	2	3	4	5	6	7	8	9	10
Valor	Reg	Alt	Reg	Alt	Reg	Alt	Alt	Alt	Alt	Reg
	11	12	13	14	15	16	17	18	19	20
	Alt	Alt	Reg	Alt	Reg	Reg	Alt	Alt	Alt	Alt
	21	22	23	24	25	26	27	28	29	30
	Alt	Reg	Alt	Reg	Alt	Alt	Alt	Reg	Alt	Alt

Estado	Frecuencia
Alto	20
Regular	10
Bajo	0

Estado	Frecuencia
Bueno	20
Malo	10



## Con Estadística Descriptiva

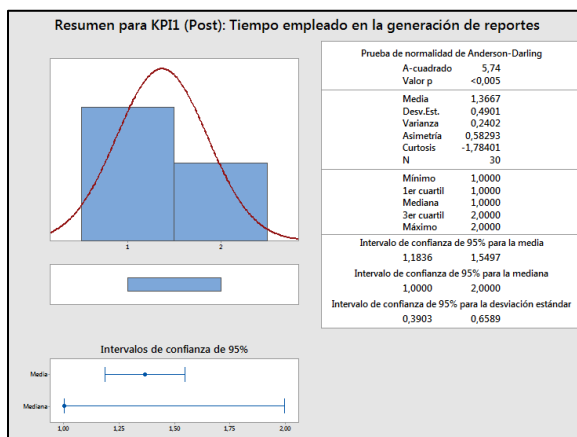


Figura 18. Estadística Descriptiva

- Sólo el 33.3% de las veces el Nivel de Entendimiento de los Reportes fue catalogado como Regular por el gerente.
- Ahora el 66.7% de las veces el Nivel de Entendimiento de los Reportes fue catalogado como Alto por el gerente.
- Se determina ahora que el 66.7% de las veces el Nivel Entendimiento de los Reportes es Bueno.
- Se determina ahora que sólo el 33.3% de las veces el Nivel de Entendimiento de los Reportes es Malo

## 6. DISCUSIÓN

- Al terminar el proyecto de investigación se vio que los valores de los KPIs con los que se trabajó tuvieron mejoras.

- Gracias al desarrollo de la solución de Business Intelligence se logró solucionar los problemas que aquejaban al proceso de toma de decisiones.
- Debido al éxito logrado se tiene en mente aplicar la inteligencia de negocios en otras áreas de la Municipalidad de Lurín.

## 7. CONCLUSIONES

- El desarrollo de la solución de Business Intelligence ayudó a mejorar el proceso de Toma de Decisiones del área de Rentas de la Municipalidad de Lurín.
- La realización de un plan de proyecto al inicio del desarrollo de la solución de Business Intelligence ayudó a tener bien definido lo que se quería realizar, los involucrados, los recursos a utilizarse, los riesgos existentes y las acciones para mitigar los mismos, y las actividades a llevarse a cabo.
- Obtener los requerimientos de parte de la gerencia de rentas y tenerlos definidos desde el principio contribuyó a tener en claro en todo momento lo que el cliente quería como producto final.
- Realizar el modelado dimensional luego de la obtención de los requerimientos, ayudó a presentar la actividad de estudio (recaudación) y los aspectos que intervienen en el (las dimensiones: contribuyente, cuota, tipo de recaudación y microzonas).
- La selección y especificación del hardware y software a utilizarse hizo que se tenga bien en claro lo que necesitaríamos para el desarrollo y el funcionamiento del producto final.
- Una vez definidas las herramientas a utilizarse, la realización del diseño físico de la base de datos permitió tener el almacén donde alojar los datos extraídos desde la base de datos transaccional.
- La determinación del proceso ETL permitió que se tenga bien definido las tablas de donde extraer los datos, las reglas de transformación de los datos para que se adapten a nuestro almacén y la manera cómo se cargan.
- La elaboración del cubo OLAP hizo que el estado actual e histórico de las recaudaciones y deudas puedan analizarse desde todos y cada uno de sus aspectos (dimensiones).
- Tener ya elaborados los reportes permitió que cada vez que se los solicite solamente se actualice la información de las recaudaciones, evitando así la elaboración desde cero.

- Realizar la implementación del Data Mart permitió que el gerente de rentas pueda visualizar de manera comprensible el estado actual e histórico de las recaudaciones realizadas desde su propio escritorio.
- La ejecución de las pruebas de funcionamiento ayudó a constatar el correcto desempeño de la solución final implementada.

## Agradecimientos

Debo expresar mi gratitud al área de informática y al área de rentas de la municipalidad de Lurín por permitirme realizar el presente proyecto de investigación.

## 8. REFERENCIAS BIBLIOGRÁFICAS

- [1] Bazán, W. (2011). *Desarrollo de un Data Mart con Indicadores Financieros como soporte para la Toma de Decisiones en el departamento financiero del Gobierno Autónomo Descentralizado Municipal San Francisco de Milagro*. Tesis Post-Grado, Universidad Estatal de Milagro, San Francisco de Milagro, Ecuador.
- [2] Ocas, M. (2012). *Desarrollo de un Data Mart en el área de Administración y Finanzas de la Municipalidad Distrital de Baños del Inca*. Tesis Pre-Grado, Universidad Privada del Norte, Cajamarca, Perú.
- [3] Guillén, F. (2012). *Desarrollo de un Data Mart para mejorar la Toma de Decisiones en el área de Tesorería de la Municipalidad Provincial de Cajamarca*. Tesis de Pre-Grado, Universidad Privada del Norte, Cajamarca, Perú.
- [4] Chirán, M. (2013). *Modelo para la implementación de Inteligencia de Negocios que apoyen a la Toma de Decisiones en instituciones públicas de Protección Social*. Tesis Post-Grado, Universidad Central del Ecuador, Quito, Ecuador.
- [5] Vega, G. (2010). *Inteligencia de Negocios. Aplicación en la administración del Presupuesto en una empresa del Sector Público*. Tesis Post-Grado, Instituto Politécnico Nacional, México D.F., México.
- [6] Villanueva, A. (2008). *Análisis, Diseño e Implementación de un Data Warehouse de Soporte de Decisiones para un Hospital del Sistema de Salud Público*. Tesis Pre-Grado, Pontificia Universidad Católica del Perú, Lima, Perú.

- [7] Ramos, S. (2011). *Microsoft Business Intelligence: Vea el cubo medio lleno*. España: SolidQ.
- [8] Cano, J. (2007). *Business Intelligence: Competir con información*. España: Banesto.
- [9] Amaya, J. (2010). *Toma de Decisiones Gerenciales. Métodos cuantitativos para la administración*. Segunda edición. Colombia: Ecoe Ediciones.
- [10] Espíndola, C. (2005). *Análisis de Problemas y Toma de Decisiones*. Tercera Edición. México: Pearson Educación.
- [11] Kimball, R., Reeves, L., Ross, M., Thornthwaite, W. (2008). *The Data Warehouse Lifecycle Toolkit*. Estados Unidos: Wiley.
- [12] Kimball, R., Ross, M. (2002). *The Data Warehouse Toolkit. The Complete Guide to Dimensional Modeling*. Segunda Edición. Estados Unidos: John Wiley & Sons, Inc.
- [13] Mundy J., Thornthwaite W., Kimball R. (2006). *The Microsoft Data Warehouse Toolkit : With SQL Server 2005 and the Microsoft Business Intelligence Toolset*. Estados Unidos: John Wiley & Sons.
- [14] Inmon, W. (2003). *Building the Data Warehouse*. Tercera edición. Estados Unidos: John Wiley & Sons, Inc.
- [15] Da Rosa F., Heinz F. (2007). *Guía práctica sobre software libre. Su selección y aplicación local en América Latina y El Caribe*. Uruguay: Mastergraf.
- [16] Stallman, R. (2004). *Software Libre para una Sociedad Libre*. Traducido por: Jaron Rowan, Diego Sanz Paratcha, Laura Trinidad. España: Traficantes de Sueños.
- [17] Bouman R., Dongen J. (2009). *Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL*. Estados Unidos: Wiley Publishing, Inc.



### APÉNDICE III: INSTRUMENTO DE RECOLECCIÓN DE DATOS

<b>ENCUESTA PARA GERENTE DE RENTAS</b>	
<b>NOMBRE:</b>	<b>FECHA:</b>
<b>POR:</b> Felix Vargas Chumpitaz	
1. ¿Al mes cuántos reportes solicita en promedio? Marque una alternativa.	
a) De 0 a 5                      b) De 6 a 10                      c) De 11 a más	
2. ¿Cuánto tiempo se demoran en entregarle los reportes que usted solicita? Marque una alternativa.	
a) De 5 a 30 minutos                      b) De 31 a 60 minutos c) De 1 a 2 horas                      d) Más de 2 horas	
3. ¿Son estos reportes siempre confiables? Marque una alternativa.	
a) Sí    b) Regularmente    c) Pocas veces    d) Nunca	
4. ¿Le ayudan los reportes a tomar la mejor decisión? Marque una alternativa.	
a) Sí    b) Regularmente    c) Pocas veces    d) Nunca	
5. ¿Qué problemas ha observado que ocurre frecuentemente? Complete.	
6. Del 1 al 10 ¿Cuál es el nivel de satisfacción que tiene respecto a la obtención de los reportes?	
a) 1    b) 2    c) 3    d) 4    e) 5 f) 6    g) 7    h) 8    i) 9    j) 10	

<b>ENCUESTA PARA SUBGERENTE DE INFORMÁTICA</b>	
<b>NOMBRE:</b>	<b>FECHA:</b>
<b>POR:</b> Felix Vargas Chumpitaz	
1. La gerencia de rentas ¿Cada cuánto tiempo le solicitan reportes?	
a) Diariamente c) De 2 a 5 veces al mes	b) De 2 a 4 veces por semana d) Casi nunca
2. ¿Cuánto tiempo demora en generar dichos reportes?	
a) De 5 a 30 minutos c) De 1 a 2 horas	b) De 31 a 60 minutos d) Más de 2 horas
3. ¿Cree usted que se satisface ese requerimiento de la mejor manera?	
a) Sí    b) Frecuentemente    c) Pocas veces    d) Casi nunca	
4. ¿Cuál es el proceso para generar los reportes actualmente?	
5. ¿Qué grado de prioridad tienen estos reportes?	
a) Muy Alta    b) Alta    c) Media    d) Baja    e) Muy Baja	
6. ¿Qué tanto esfuerzo le demanda la generación de los reportes?	
a) Mucho Esfuerzo    b) Mediano Esfuerzo    c) Poco Esfuerzo	
7. ¿Se cuenta con las herramientas para conseguir la optimización del proceso?	
a) Sí    b) No	
8. Del 1 al 10 ¿Cuál es el nivel de satisfacción que usted cree que tiene la gerencia de rentas con el proceso de generación de dichos reportes?	
a) 1    b) 2    c) 3    d) 4    e) 5 f) 6    g) 7    h) 8    i) 9    j) 10	

## APÉNDICE IV: COMPARACIÓN DE METODOLOGÍAS

### Comparación Cualitativa









Criterio	Inmon	Kimball
	Enfoque Top-Down	Enfoque Bottom-Up
Alcance	Toda la empresa	Un área es específico
Complejidad	Muy complejo	Bastante simple
Orientación	Basado en datos	Orientado a procesos
Herramientas	Diagrama Entidad-Relación	Modelado dimensional
Accesibilidad del usuario final	Bajo	Alto
Tiempo de desarrollo	Largo	Corto
Costo	Alto costo inicial. Luego va disminuyendo	Bajo costo inicial. Luego va aumentando.

### Comparación Cuantitativa

Criterio	Inmon	Kimball	
	Enfoque Top-Down	Enfoque Bottom-Up	
Complejidad	3	2	
Entendimiento de los requerimientos	3	3	
Entendimiento del problema	3	3	
Tiempo de desarrollo	3	1	
Costo	3	1	
Calidad	3	3	
	1= Bajo	2=Medio	3=Alto

Luego de ambas comparaciones podemos apreciar que la metodología de Ralph Kimball es la más apropiada para nuestro caso, debido a la menor complejidad y al menor costo. Pero la razón más importante para la elección es que nosotros solamente necesitamos abarcar un proceso en específico (recaudación, área de rentas), lo que se ajusta a la metodología de Kimball.

## APÉNDICE V: MATRIZ DE LA SOLUCIÓN DEL PROBLEMA

I. Datos Generales	II. Estado del Arte	III. Tecnologías	IV. Arquitectura de la Solución
<p><b>1. Área de Investigación</b> Business Intelligence</p> <p><b>2. Línea de Investigación</b> Data Mart</p> <p><b>3. Títulos de la tesis</b> Desarrollo de una solución de Business Intelligence para mejorar el proceso de toma de decisiones en el área de rentas de la Municipalidad de Lurín.</p> <p><b>4. Variable Independiente</b> Solución de Business Intelligence</p> <p><b>5. Variable Dependiente</b> Proceso de toma de decisiones</p>	<p><b>1. Modelo de referencia</b> Modelo herramienta Business Intelligence</p> <p><b>1.1. Justificación</b> El desarrollo de la solución de Business Intelligence optimizará el proceso de toma de decisiones en el área de rentas, mostrando el estado de las recaudaciones mediante indicadores y sin demandar mucho tiempo ni esfuerzo.</p> <p><b>2. Metodologías</b> <b>2.1. Metodología de Desarrollo del Proyecto</b> -Planteamiento Metodológico -Marco Referencial -Desarrollo de la Solución -Análisis de Resultados y Contrastación de la Hipótesis -Conclusiones y Recomendaciones</p> <p><b>2.2. Metodología de Desarrollo de la solución</b> Metodología de Ralph Kimball: Ciclo de Vida Dimensional del Negocio</p> <p><b>3. Método Seleccionado</b> No hay método seleccionado</p>	<p><b>1. Plataforma</b> HP-PC, Intel Xeon E5320 1.86ghz (x64),</p> <p><b>2. Sistemas Operativos</b> Windows Server 2008 (Servidor físico y virtual), Windows 7 (Cliente) <b>Versión propuesta:</b> Windows Server 2008</p> <p><b>3. Lenguaje de Programación</b> Java (Suite Pentaho)</p> <p><b>4. Base de datos</b> PostgreSQL <b>Versión propuesta:</b> 9.3</p> <p><b>5. Framework</b> Suite Pentaho <b>Versión propuesta:</b> Community Edition 5.2</p>	 <p><b>Fuente de Información</b> PostgreSQL </p> <p><b>Integración de Datos</b> Extracción, Transformación y Carga ETL </p> <p><b>Repositorio de Datos</b>  Almacén de Datos</p> <p><b>Análisis</b> OLAP  Dashboard  Gráficos  Reportes </p> <p>Modelo: Herramienta de Business Intelligence</p>

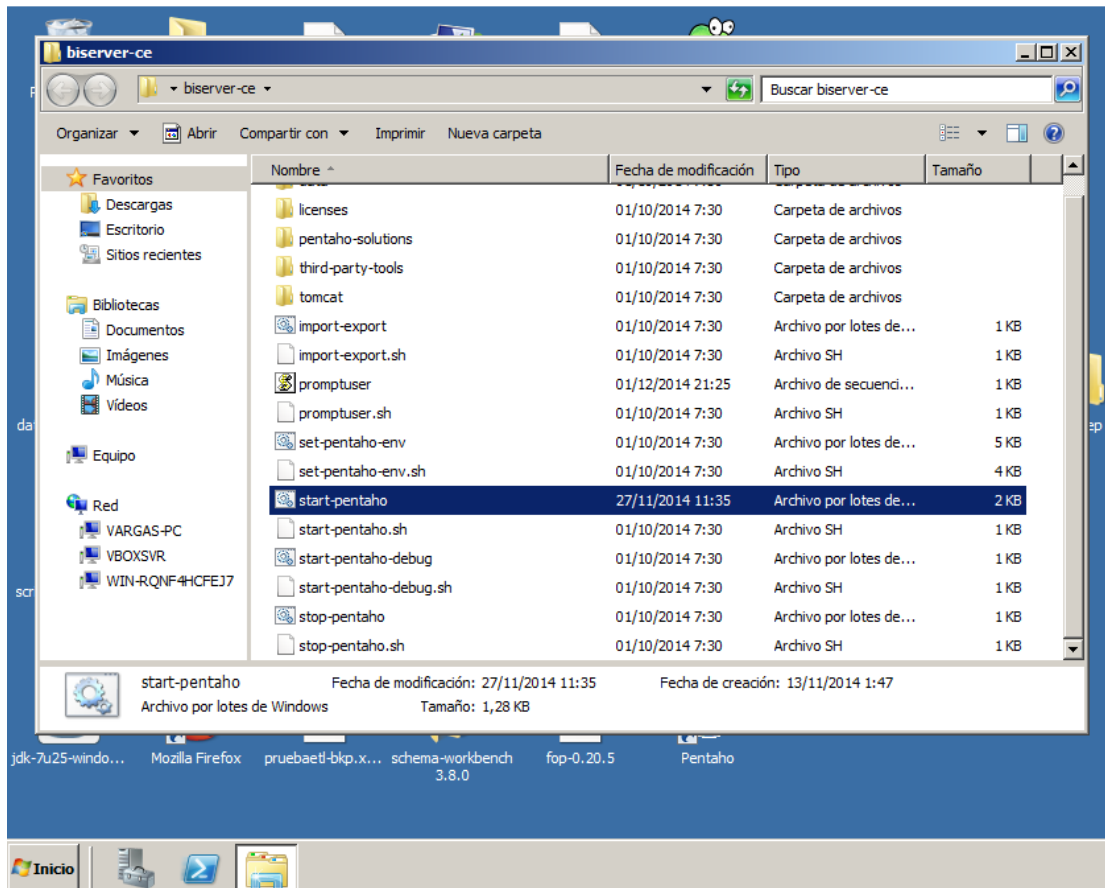
	<p><b>4. Técnica Seleccionada</b></p> <p>No hay método seleccionado</p> <p><b>5. Algoritmo Seleccionado</b></p> <p>No hay Algoritmo Seleccionado</p>	<p><b>6. Otros Requerimientos Tecnológicos</b></p> <p>JDK (Java Development Kit)</p>	<p><b>Descripción:</b></p> <p>La <b>arquitectura</b> de una herramienta de Business Intelligence está basada en una o más <b>fuentes de información</b>, que actúa como <b>entrada de datos</b>. Los datos obtenidos desde las fuentes de información son transformadas, a fin de adaptarse a nuestras necesidades, es decir que los datos se adecuen al modelo físico de nuestro <b>repositorio de datos</b>. Es en este lugar donde los datos son cargados (almacenados). Una vez aquí, se procede a crear el cubo dimensional OLAP, el cual contiene las dimensiones (características) y los hechos (medidas). Finalmente, con esto es posible crear reportes, en modo de gráficos y cuadros de mando (dashboards), con la información solicitada.</p>
--	--	--	---

## APÉNDICE VI: MANUAL DE USUARIO PARA ADMINISTRADOR

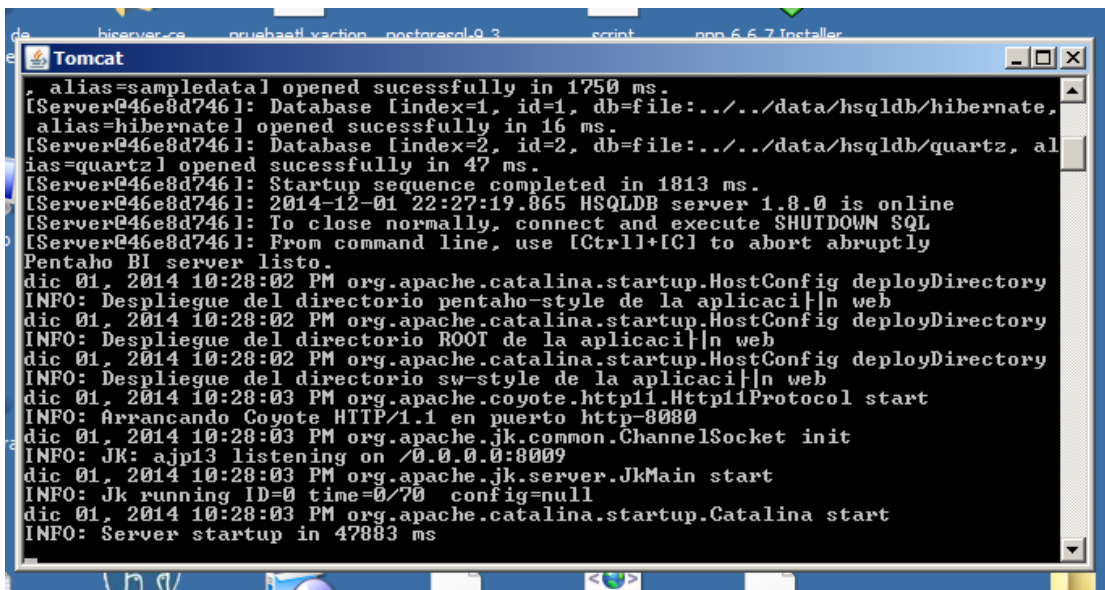
### MANUAL DE USUARIO PARA ADMINISTRADOR

#### INICIAR EL SERVIDOR BI

1. Dentro de la carpeta que contiene el bi-server, ejecutar el archivo start-pentaho.bat.

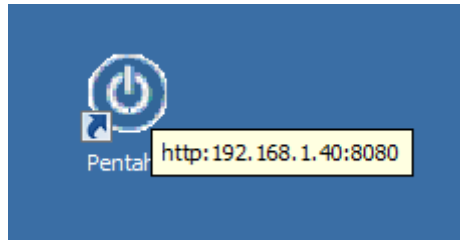


2. Esperamos hasta que cargue el servidor.



## INICIAR SESIÓN EN LA CONSOLA DE USUARIO

1. Iniciar la consola de usuario por medio del acceso directo creado o en su defecto dirigirse a la ruta `http:192.168.1.40:8080` (la IP es del servidor BI).



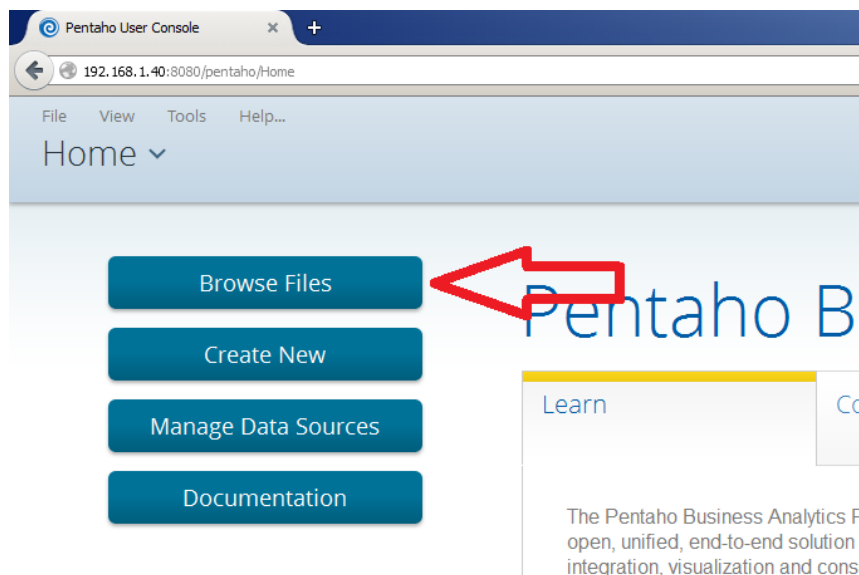
2. Iniciamos sesión con nuestra cuenta de administrador.



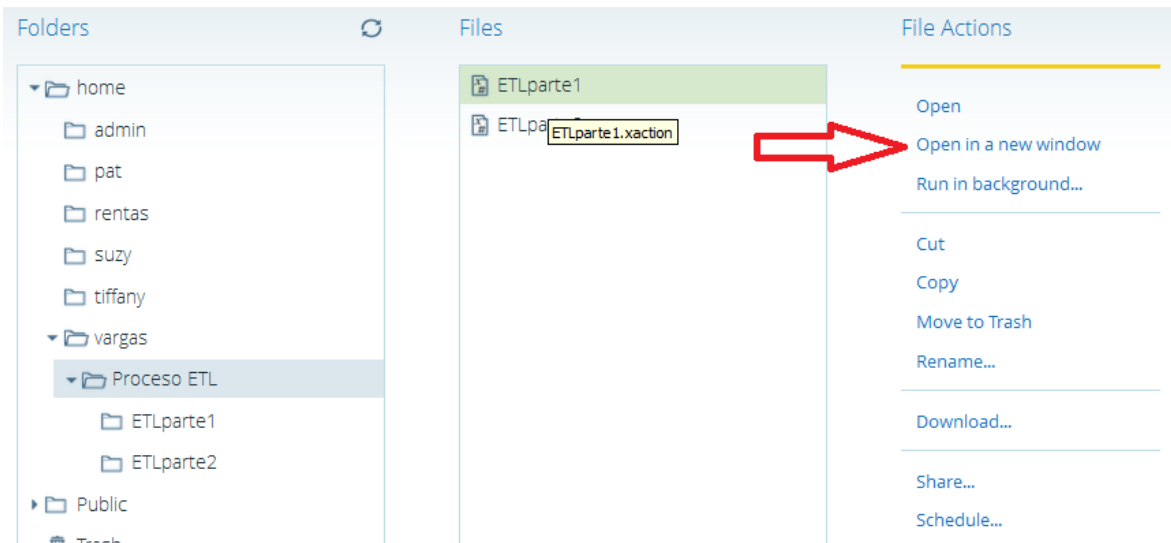
## EJECUCIÓN MANUAL DEL PROCESO ETL

NOTA: El proceso ETL está programado para ejecutarse por sí solo en altas horas de la noche, es por esto que los siguientes pasos solo deben realizarse en caso de ser necesario.

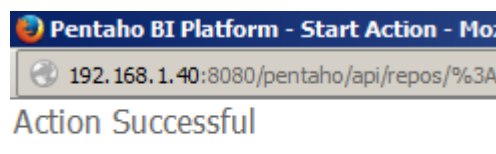
1. Iniciamos sesión y seleccionamos Browse Files.



2. Nos dirigimos a nuestra carpeta de usuario / carpeta Proceso ETL y seleccionamos el archivo ETLparte1.xaction, el cual es un action sequence. Finalmente damos click en Open in a New Window.



3. Esperar hasta que en la nueva ventana aparezca Action Successful, lo cual indica que el proceso fue finalizado con éxito.



Finalmente se realiza la acción con el action sequence llamado ETLpart2.xaction. Una vez terminada ambas partes, el proceso ETL ha terminado.

## OBTENCIÓN DE LOS KPI

1. Nos dirigimos a la base de datos DM\_rentas del servidor y ejecutamos la siguiente sentencia SQL.

```
select tip.codtipo, cuo.anio, cuo.cuota, sum(importe) as importe,
sum(derecho_emision) as derecho_emision, sum(mora) as mora,
sum(importe+derecho_emision+mora) as monto_total
from fact_recaudacion reca
inner join dim_cuota cuo on cuo.cuota_key=reca.cuota_key
inner join dim_contribuyente cont on cont.contribuyente_key=reca.contribuyente_key
inner join dim_zona micro on micro.microzona_key=reca.microzona_key
inner join dim_tipopago tip on tip.tipo_key=reca.tipo_key
inner join dim_estado est on est.estado_key=reca.estado_key
group by tip.codtipo, cuo.anio, cuo.cuota
order by tip.codtipo, cuo.anio, cuo.cuota
```



2. Luego de unos segundos de espera, obtendremos el siguiente resultado:

Output pane							
Data Output							
	codtipo character(3)	anio character varying(4)	cuota character(2)	importe numeric	derecho_emision numeric	mora numeric	monto_total numeric
181	004	2014	01	777327.52	151092	322.15	928741.67
182	004	2014	02	777309.98	0	16.69	777326.67
183	004	2014	03	777055.83	0	45.46	777101.29
184	004	2014	04	776999.84	0	0.00	776999.84
185	004	2014	05	776957.91	0	0.00	776957.91
186	004	2014	06	776948.74	0	0.00	776948.74
187	004	2014	07	776954.91	0	0.00	776954.91
188	004	2014	08	776952.27	0	0.00	776952.27
189	004	2014	09	776928.10	0	1.63	776929.73
190	004	2014	10	776915.04	0	0.00	776915.04
191	004	2014	11	776906.91	0	0.00	776906.91
192	004	2014	12	776905.01	0	0.00	776905.01

OK. Unix Ln 1, Col 1, Ch 1

(codtipo: 004 = Arbitrios Municipales, 001 = Impuesto Predial)

3. Para evitar estar calculando y sacando porcentajes, se creó un archivo en Excel llamado Obtención de KPI, en la cual se llenan valores (ya sean 4 cuotas o 12 cuotas) y automáticamente se calcula los valores.

	A	B	C	D	E
1	<b>CUOTA</b>	<b>INGRESAR:</b>	<b>BIEN / META</b>	<b>REGULAR</b>	<b>MALO</b>
2	1	3442216.11	3442216	3097994	1721108
3	2	3287988.44	3287988	2959189	1643994
4	3	3288025.11	3288025	2959223	1644013
5	4	3288031.67	3288032	2959229	1644016
6	5		0	0	0
7	6		0	0	0
8	7		0	0	0
9	8		0	0	0
10	9		0	0	0
11	10		0	0	0
12	11		0	0	0
13	12		0	0	0
14					

## APÉNDICE VII: MANUAL DE USUARIO PARA GERENTE

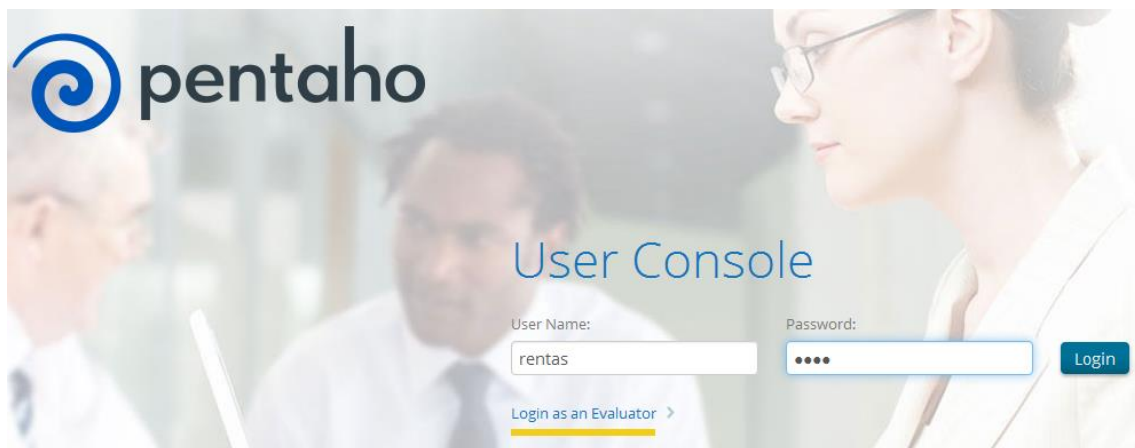
### MANUAL DE USUARIO PARA GERENTE

#### INICIAR SESIÓN EN LA CONSOLA DE USUARIO

1. Ejecutar el ícono ubicado en el escritorio.

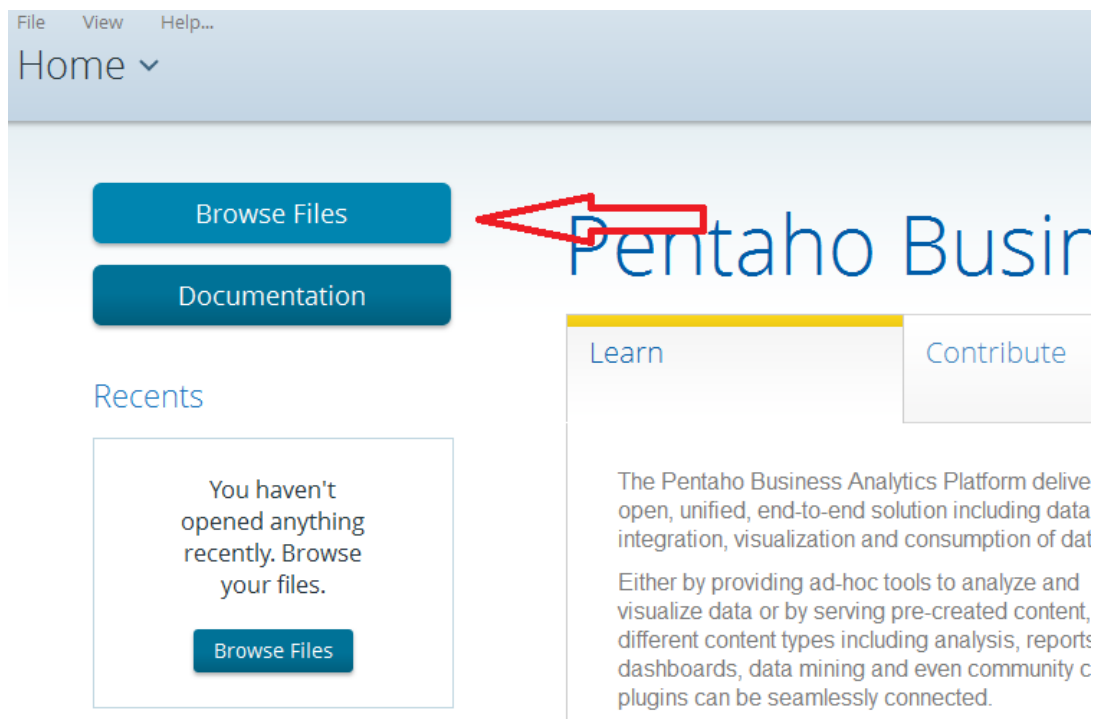


2. Ingresar nuestro usuario y contraseña.



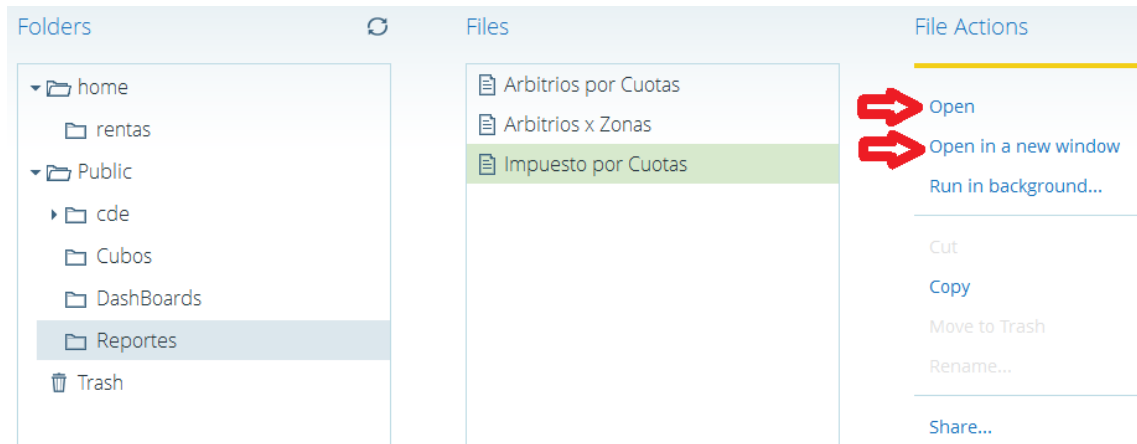
#### VISUALIZAR LOS REPORTEES

1. Una vez iniciada la sesión, damos click en Browse Files.

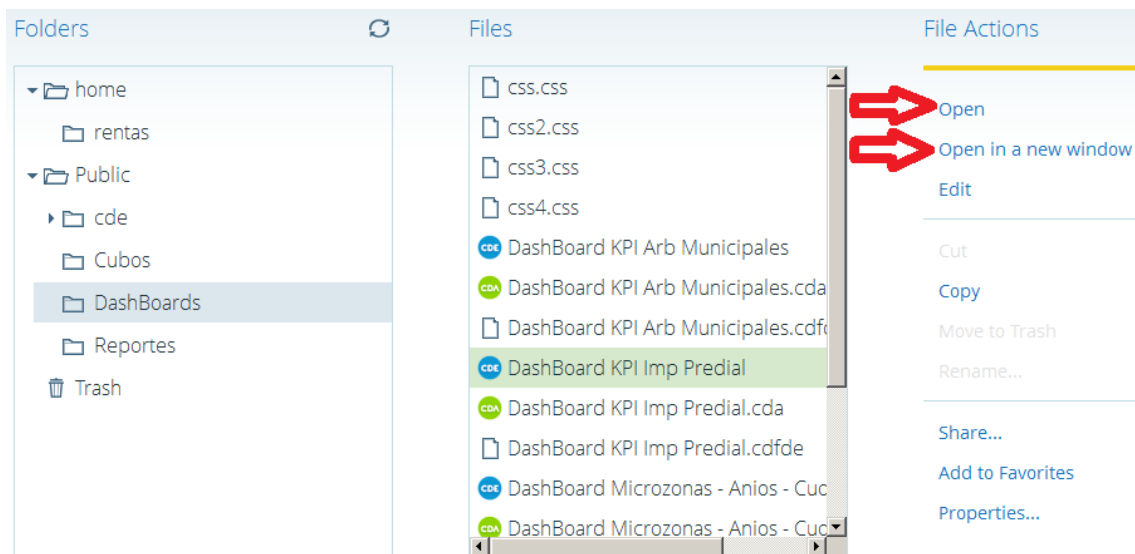


## 2. Entramos a la carpeta Public.

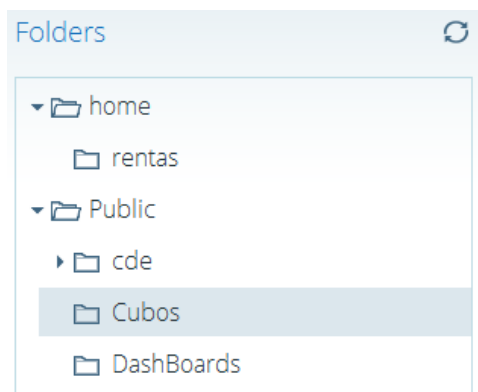
a) Para ver los reportes a medida entramos a la carpeta Reportes, seleccionamos el que deseemos y damos click en Open o en Open in a new window.



b) Para ver los dashboards entramos a la carpeta Dashboards, seleccionamos el que deseemos (los que tienen el ícono azul) y damos click en Open o en Open a new window.



c) Para ver los cubos es el mismo procedimiento, solo que entramos a la carpeta Cubos.



## APÉNDICE VIII: SENTENCIAS SQL PARA LA CREACIÓN DE LAS TABLAS

### Tabla apoyo\_cta\_aica

```
CREATE TABLE apoyo_cta_aica
(
  codcli character(7) NOT NULL,
  codpre character(7) NOT NULL,
  cuenta character(3) NOT NULL,
  actualcuota character(6) NOT NULL,
  estado char(1) NOT NULL,
  feccan date NOT NULL,
  importe numeric(12,2) NOT NULL,
  deremi numeric(12,2) NOT NULL,
  morcan numeric(12,2) NOT NULL,
  total numeric(12,2) NOT NULL
);
```

### Tabla apoyo\_microzona

```
CREATE TABLE apoyo_microzona
(
  desurb character varying(80) NOT NULL,
  deszon character varying(10) NOT NULL
);
```

### Tabla predios\_ur

```
CREATE TABLE apoyo_predios_ur
(
  codpred character(7) NOT NULL,
  desurb character varying(80) NOT NULL
);
```

### Tabla dim\_contribuyente

```
CREATE TABLE dim_contribuyente
(
  contribuyente_key serial PRIMARY KEY,
  nombre character varying(100) NOT NULL,
  cod_interno character(7) NOT NULL
);
```

### Tabla dim\_cuota

```
CREATE TABLE dim_cuota
(
  cuota_key serial PRIMARY KEY,
  descripcion character(6) NOT NULL,
  cuota character(2) NOT NULL,
  anio character varying(4) NOT NULL
);
```

### Tabla dim\_zona

```
CREATE TABLE dim_zona
(
  microzona_key serial PRIMARY KEY,
  codpred character(7) NOT NULL,
  microzona character varying(80) NOT NULL,
  zona character varying(10) NOT NULL
);
```

### Tabla dim\_tiempo

```
CREATE TABLE dim_tiempo
(
  tiempo_key serial PRIMARY KEY,
  fecha date NOT NULL,
  mes character varying(10) NOT NULL,
  trimestre varchar(3) NOT NULL,
  anio character(4) NOT NULL
);
```

### Tabla dim\_tipoPago

```
CREATE TABLE dim_tipoPago
(
  tipo_key serial PRIMARY KEY,
  descripcion character varying(20) NOT NULL,
  codtipo character(3) NOT NULL
);
```

### Tabla dim\_estado

```
CREATE TABLE dim_estado
(
  estado_key serial PRIMARY KEY,
  codestado character(1) NOT NULL,
  descripcion character varying(20) NOT NULL
);
```

### Tabla fact\_recaudacion

```
CREATE TABLE fact_recaudacion
(
  tipo_key integer NOT NULL,
  cuota_key integer NOT NULL,
  contribuyente_key integer NOT NULL,
  microzona_key integer NOT NULL,
  tiempo_key integer NOT NULL,
  estado_key integer NOT NULL,
  importe numeric(12,2) NOT NULL,
  derecho_emision numeric(12,2) NOT NULL,
  mora numeric(12,2) NOT NULL,
  monto_total numeric(12,2) NOT NULL,
  CONSTRAINT tipo_fkey FOREIGN KEY (tipo_key) REFERENCES dim_tipoPago(tipo_key),
  CONSTRAINT cuota_fkey FOREIGN KEY (cuota_key) REFERENCES dim_cuota(cuota_key),
  CONSTRAINT contribuyente_fkey FOREIGN KEY (contribuyente_key)
  REFERENCES dim_contribuyente(contribuyente_key),
  CONSTRAINT microzona_fkey FOREIGN KEY (microzona_key) REFERENCES dim_zona(microzona_key),
  CONSTRAINT tiempo_fkey FOREIGN KEY (tiempo_key) REFERENCES dim_tiempo(tiempo_key),
  CONSTRAINT estado_fkey FOREIGN KEY (estado_key) REFERENCES dim_estado(estado_key),
  CONSTRAINT fact_recaudacion_pkey PRIMARY KEY
  (cuota_key, contribuyente_key, microzona_key, tipo_key, tiempo_key, estado_key)
);
```

## GLOSARIO DE TÉRMINOS

### A

- ❖ **Action Sequence:** Archivo que contiene una secuencia de acciones. Son programadas desde el software Pentaho Design Studio, y almacenadas en el Pentaho BI Server para su posterior ejecución manual y/o automática. Tienen la extensión “.xaction”.
- ❖ **Arbitrios Municipales:** Son una contraprestación por la prestación de un servicio público por parte de las Municipalidades, es decir, son aquellas tasas que se pagan por la prestación, mantenimiento del servicio público, individualizado en el contribuyente.

### B

- ❖ **Backup:** Una copia de seguridad, copia de respaldo o backup (su nombre en inglés) en tecnologías de la información e informática es una copia de los datos originales que se realiza con el fin de disponer de un medio para recuperarlos en caso de su pérdida
- ❖ **Batch:** Un archivo de texto con extensión .bat que contiene una secuencia de órdenes para ser ejecutadas en sistemas DOS, OS/2 y Microsoft Windows.
- ❖ **BI:** Abreviatura de Business Intelligence.
- ❖ **Business Intelligence:** Técnicas y herramientas para obtener información a partir de los datos, y conocimiento a partir de la información.

### C

- ❖ **Contribuyente:** Aquella persona natural o jurídica obligada a cumplir y hacer cumplir los deberes formales según lo establece las leyes tributarias del país. Es decir, son aquella persona física o jurídica con derechos y obligaciones, frente a un ente público, derivados de los tributos. Es quien está obligado a soportar patrimonialmente el pago de los tributos
- ❖ **CRM:** (Customer Relationship Management) es la sigla que se utiliza para definir una estrategia de negocio enfocada al cliente, en la que el objetivo es reunir la mayor cantidad posible de información sobre los clientes para generar relaciones a largo plazo y aumentar así su grado de satisfacción.

### D

- ❖ **Dashboard:** En el mundo de los negocios, es una herramienta mediante la cual se despliega en tiempo real información de la empresa extraída de varias fuentes o bases de datos. Su característica de tiempo real otorga a los usuarios un conocimiento completo sobre la marcha de la empresa y permite hacer análisis instantáneos e inteligencia de negocios.
- ❖ **Data Mart:** Es una base de datos similar al DataWarehouse, con la diferencia que solo se centra en un proceso en específico. Se suele considerar que un conjunto de Data Marts forman un Data Warehouse.

- ❖ **Data Warehouse:** Es una base de datos corporativa que almacena información procesada de diversas fuentes dentro de la empresa. Reúne la información de todas las áreas y las mantiene disponibles para su consulta.
- ❖ **Derecho de Emisión:** En el ámbito de las recaudaciones, es un importe que pagan los contribuyentes por los costos de los talonarios.
- ❖ **Diccionario de Datos:** Conjunto de metadatos que contiene las características lógicas y puntuales de los datos que se van a utilizar en el sistema que se programa, incluyendo nombre, descripción, alias, contenido y organización.
- ❖ **Drill-down:** Es una técnica de resolución de problemas consistente en subdividir el problema en subconjuntos cada uno de los cuales puede resolverse más fácilmente.

## I

- ❖ **Impuesto Predial:** Es un tributo de periodicidad anual que grava el valor de los predios urbanos y rústicos. Para efectos del Impuesto se considera predios a los terrenos, incluyendo los terrenos ganados al mar, a los ríos y a otros espejos de agua, así como las edificaciones e instalaciones fijas y permanentes que constituyan partes integrantes de dichos predios, que no pudieran ser separadas sin alterar, deteriorar o destruir la edificación.

## K

- ❖ **KPI:** Del inglés Key Performance Indicator, conocido como Indicador clave de desempeño, (o también Indicador clave de rendimiento) es una medida del nivel del desempeño de un proceso; el valor del indicador está directamente relacionado con un objetivo fijado de antemano. Normalmente se expresa en porcentaje.

## M

- ❖ **Minería de Datos:** Es un campo de las ciencias de la computación referido al proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos.

## P

- ❖ **Plugin:** También llamado complemento. Es una aplicación que se relaciona con otra para aportarle una función nueva y generalmente muy específica. Esta aplicación adicional es ejecutada por la aplicación principal e interactúan por medio de la API.

## R

- ❖ **RDBMS:** Un sistema de gestión de bases de datos relacionales es aquel que sigue el modelo relacional.

## S

- ❖ **SGBD:** Un Sistema de Gestión de Bases de Datos es un conjunto de programas que permiten el almacenamiento, modificación y extracción de la información en una

base de datos, además de proporcionar herramientas para añadir, borrar, modificar y analizar los datos.

- ❖ **Sistema Transaccional:** Es un tipo de sistema de información diseñado para recolectar, almacenar, modificar y recuperar todo tipo de información que es generada por las transacciones en una organización. Una transacción es un evento o proceso que genera o modifica la información que se encuentran eventualmente almacenados en un sistema de información.
- ❖ **SQL:** El lenguaje de consulta estructurado o SQL (por sus siglas en inglés Structured Query Language) es un lenguaje declarativo de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones en ellas. Una de sus características es el manejo del álgebra y el cálculo relacional que permiten efectuar consultas con el fin de recuperar de forma sencilla información de interés de bases de datos, así como hacer cambios en ellas
- ❖ **Stakeholder:** El término agrupa a trabajadores, organizaciones sociales, accionistas y proveedores, entre muchos otros actores clave que se ven afectados por las decisiones de una empresa