



UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA

ESCUELA ACADÉMICO PROFESIONAL DE INGENIERÍA DE SISTEMAS

**“DATAMART PARA EL PROCESO DE ANÁLISIS DE VENTAS EN
EL ÁREA DE MARKETING DE LA EMPRESA LABORATORIOS
ROEMMERS S.A”**

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE
INGENIERA DE SISTEMAS**

AUTOR:

ASTETE PASCUAL, MARÍA ISABEL

ASESOR:

MG. AUCCAHUASI AIQUIPA, WILVER

LÍNEA DE INVESTIGACIÓN:

Estratégico y toma de decisiones

LIMA – PERÚ

2015

PÁGINA DEL JURADO

Mg. Iván Martín Pérez Farfán

Mg. Raúl Eduardo Huarote Zegarra

Dr. Santiago Esteban Contreras Aranda

DEDICATORIA

A quienes más amo y debo en este mundo:

A Dios, a quien pido guíe mi camino, le doy gracias por cada día y por las personas que ha puesto en mi vida.

A mi familia: Mis padres; Haydee y José, por su esfuerzo, por las enseñanzas, por los valores y su amor incondicional. A mis hermanos, mis abuelitos y mis tíos Álvaro y Carlos, por ser parte importante en mi vida.

A Jordan por ser mi compañero, mi amigo, mi confidente y mi amor.

AGRADECIMIENTO

En primer lugar a Dios, porque sin él no podría ser ni hacer nada.

A Jorge Linares Villamar, ahora Gerente de Producto en la empresa Laboratorios Roemmers S.A, Sponsor en la implementación de esta tesis.

A los profesores que contribuyeron con sus conocimientos y en la evaluación de expertos, especialmente a los Magísteres: Luis Ramos Mendoza, Orleans Gálvez Tapia e Iván Pérez Farfán.


DECLARATORIA DE AUTENTICIDAD

Yo, María Isabel Astete Pascual, estudiante de la Escuela profesional de Ingeniería de Sistemas, de la universidad Cesar Vallejo, sede/filial Lima Norte; declaro que el trabajo académico titulado: **"DATAMART PARA EL PROCESO DE ANÁLISIS DE VENTAS EN EL ÁREA DE MARKETING DE LA EMPRESA LABORATORIOS ROEMMERS S.A"**, presentada para la obtención del grado académico de Ingeniera de Sistemas, es de mi autoría.

Por lo tanto, declaro lo siguiente:

- He mencionado todas las fuentes empleadas en el presente trabajo de investigación, identificando correctamente toda cita textual o de paráfrasis previamente de otras fuentes, de acuerdo con lo establecido por las normas de elaboración de trabajos académicos.
- No he utilizado ninguna otra fuente distinta de aquella expresamente señaladas en este trabajo
- Este trabajo de investigación no ha sido previamente presentado completa, ni parcialmente para la obtención de otro grado académico o título profesional.
- Soy consciente de que mi trabajo puede ser revisado electrónicamente en búsqueda de plagios.
- De encontrar uso de material intelectual ajeno sin el debido reconocimiento de su fuente o autor, me someto a las sanciones que determinan el procedimiento disciplinario.

Lima, 01 de diciembre 2015



María Isabel Astete Pascual

PRESENTACIÓN

El título de la tesis es “Datamart para el proceso de análisis de ventas en el área de marketing de la empresa Laboratorios Roemmers S.A”.

Se realizó en el área de Marketing de la empresa Laboratorios Roemmers S.A, con el fin de mejorar el proceso de análisis de ventas de dicha empresa debido a que, este análisis era ineficaz ya que los reportes de análisis de ventas realizados por los gerentes de producto y por jefe de marketing, no cumplían con las necesidades de información detectadas, además, se empleaba mucho tiempo en la generación de reportes, y este tiempo resultaba equivalente a un costo laboral total excesivo, por lo que el proceso resultaba ser también ineficiente.

Por lo mencionado el objetivo de la tesis fue determinar la influencia de un datamart en el área de marketing de la empresa Laboratorios Roemmers S.A.

Para ello se realizó la implementación del datamart y se obtuvo que este, incrementó el nivel de eficacia en el cumplimiento de las necesidades de información, disminuye el tiempo de generación de reportes y el costo laboral total en la generación de dichos reportes.

La presente Investigación se divide en siete capítulos: En el capítulo I, se presenta la introducción de la investigación en donde se expone la realidad problemática, trabajos, previos, teorías relacionadas al tema, formulación del problema justificación del estudio, hipótesis y objetivos. En el capítulo II, se detalla el método, el cual involucra el diseño de investigación, variables, población y muestra, técnicas e instrumentos de recolección de datos, métodos de análisis de datos y aspectos éticos. En el capítulo III se demuestran los resultados mediante el análisis descriptivo, comparativo, inferencial y las pruebas de hipótesis. En el capítulo IV se realiza la discusión de los resultados obtenidos, así mismo se hace comparación con los resultados obtenidos por los trabajos previos. En el capítulo V se exponen las conclusiones a las que llegó el investigador. En el capítulo VI se exponen las sugerencias también brindadas por el investigador. En el capítulo VII Se detallan las referencias bibliográficas utilizadas en la investigación. La tesis finaliza con los anexos correspondientes.

ÍNDICE

PÁGINA DEL JURADO	ii
DEDICATORIA	iii
AGRADECIMIENTO.....	iv
DECLARATORIA DE AUTENTICIDAD	v
PRESENTACIÓN.....	vi
RESUMEN	xiv
ABSTRACT	15
I. INTRODUCCIÓN	17
1.1. Realidad problemática	18
1.2. Trabajos previos.....	23
1.3. Teorías relacionadas al tema	30
A. Datamart.....	30
B. Proceso Análisis de Ventas	45
C. Metodología de desarrollo – Datamart.....	53
D. Metodología para la herramienta de presentación al usuario final.....	58
1.4. Formulación del problema	60
A. Problema Principal	60
B. Problemas Específicos.....	60
1.5. Justificación del estudio	60
A. Justificación Tecnológica.....	60
B. Justificación Económica	61
C. Justificación Institucional.....	61
D. Justificación Operativa.....	62
1.6. Hipótesis.....	63
A. Hipótesis General	63
B. Hipótesis Específicas	63
1.7. Objetivos.....	63
A. Objetivo General	63
B. Objetivos Específicos.....	63
II. MÉTODO	65
2.1. Diseño de investigación.....	66
2.2. Variables, operacionalización.....	66

A.	VARIABLES.....	66
B.	OPERACIONALIZACIÓN DE VARIABLES.....	68
2.3.	<i>Población y muestra</i>	68
A.	Población.....	68
B.	Muestra.....	69
2.4.	<i>Técnicas e instrumentos de recolección de datos, validez y confiabilidad</i>	71
A.	Técnicas:.....	71
B.	Instrumentos.....	72
C.	Validez y confiabilidad	73
2.5.	<i>Métodos de análisis de datos</i>	74
2.6.	<i>Aspectos éticos</i>	81
III.	RESULTADOS	83
A.	<i>Análisis Descriptivos</i>	83
B.	<i>Análisis Comparativo</i>	89
C.	<i>Análisis Inferencial</i>	92
D.	<i>Pruebas de Hipótesis</i>	95
IV.	DISCUSIÓN	104
V.	CONCLUSIONES	107
VI.	RECOMENDACIONES	109
VII.	REFERENCIAS	110
	ANEXOS	116
	<i>Anexo 01: Matriz de consistencia</i>	<i>117</i>
	<i>Anexo 02: ENTREVISTA</i>	<i>119</i>
	<i>Anexo 03: Aprobación del documento de Análisis de Requerimientos</i>	<i>121</i>
	<i>Anexo 04: Juicio de Expertos - Validación del Instrumento para el indicador “Nivel de eficacia en el cumplimiento de las necesidades de información”</i>	<i>122</i>
	<i>Anexo 05: Juicio de Expertos - Validación del Instrumento para el indicador: “Tiempo de generación de reportes”</i>	<i>125</i>
	<i>Anexo 06: Juicio de Expertos - Validación del Instrumento para el indicador: “Costo laboral total de generación de reportes”</i>	<i>128</i>
	<i>Anexo 07: Juicio de Expertos – Elección de la metodología para el desarrollo del Datamart.</i>	<i>131</i>
	<i>Anexo 08: Estudio Remuneración de ejecutivos según Transearch 2013</i>	<i>134</i>
	<i>Anexo 09: Ficha de registro 1 – Nivel de eficacia en el cumplimiento de las necesidades de Información PreTest</i>	<i>140</i>

Anexo 10: Ficha de registro 2 – Nivel de eficacia en el cumplimiento de las necesidades de Información PostTest	143
Anexo 11: Ficha de observación 1 – Tiempo de generación de reportes PreTest	146
Anexo 12: Ficha de observación 2 – Tiempo de generación de reportes PostTest	147
Anexo 13: Ficha de registro 3 – Costo laboral total de generación de reportes PreTest	148
Anexo 14: Ficha de registro 4 – Costo laboral total de generación de reportes PostTest	149
Anexo 15: Constancia de implementación	150
ANEXO 16: Desarrollo de la Implementacion del datamart	151
ANEXO 17: METODOLOGÍA HEFESTO	152
ANEXO 18: CUBO REALIZADO EN ANÁLISIS SERVICES (SSAS).....	209
ANEXO 19: METODOLOGÍA ICONIX PARA LA HERRAMIENTA DE PRESENTACIÓN PARA EL USUARIO FINAL	213
ANEXO 20: CONSULTA DE DATOS UTILIZANDO MICROSOFT EXCEL	233
ANEXO 21: CONSULTA DE DATOS UTILIZANDO REPORTING SERVICES	238
ANEXO 22: CONSULTA DE DATOS UTILIZANDO PENTAHO REPORT DESIGNER	249
ANEXO 22: TABLA COMPARATIVA DE LAS HERRAMIENTAS UTILIZADAS POR EL INVESTIGADOR PARA LA CONSULTA DE DATOS (REPORTE).....	259

INDICE DE FIGURAS

Figura 1: Canales de venta	19
Figura 2: Distribución de ventas por colocación Roemmers	19
Figura 3: Realidad problemática - Nivel de eficacia en el cumplimiento de las necesidades de información.....	20
Figura 4: Realidad problemática - Tiempo de generación de reportes	22
Figura 5: Costo Laboral de generación de Reportes	23
Figura 6: Inteligencia de Negocios	30
Figura 7: Arquitectura del datamart	35
Figura 8: ROLAP VS MOLAP	38
Figura 9: Principio 80:20	46
Figura 10: Proceso de Análisis de Ventas.....	47
Figura 11: Desiciones clave para el proceso de análisis de Ventas	49
Figura 12: Metodologías de desarrollo - datamart.....	53
Figura 13: Metodología HEFESTO	56
Figura 14: Metodología ICONIX.....	59
Figura 15: Distribución normal.....	78
Figura 16: Tabla de distribución T-Student.....	79
Figura 17: Histograma del Nivel de eficacia en el cumplimiento de las necesidades de la información (Pre-Test).....	84

Figura 18: Histograma de Nivel de eficacia en el cumplimiento de las necesidades de la información (Post - Test)	85
Figura 19: Histograma Tiempo de generación de reportes (Pre-Test)	86
Figura 20: Histograma tiempo de generación de reportes (Post-Test)	87
Figura 21: Histograma costo laboral total de reportes (Pre-Test)	88
Figura 22: Histograma tiempo de generación de reportes (Post-Test)	89
Figura 23: Comparativa del nivel de eficacia en el cumplimiento de las necesidades de información.....	90
Figura 24: Comparativa del tiempo de generación de Reportes.....	91
Figura 25: Comparativa del costo laboral total de generación de reportes.....	92
Figura 26: Distribución T - Nivel de eficacia en el cumplimiento de las necesidades de información.....	97
Figura 27: Distribución T - Tiempo de generación de reportes	99
Figura 28: Distribución T - Costo total de generación de reportes.....	102
Figura 29: Modelo Solución Implementado	151
Figura 30: Modelo Conceptual Inicial - Análisis de Ventas	163
Figura 31: Modelo Conceptual Inicial - Total Unidades Vendidas	164
Figura 32: Modelo Conceptual Inicial - Total de Ventas en Soles.....	165
Figura 33: Modelo Conceptual Inicial - Ranking de Ventas	166
Figura 34: Modelo Conceptual Inicial - Tasa de Crecimiento	167
Figura 35: Modelo Conceptual Inicial - Precio Promedio	168
Figura 36: Modelo Conceptual Inicial - YTD	169
Figura 37: Modelo de datos OLTP (Herramienta SQL).....	174
Figura 38: Establecer Correspondencias.....	179
Figura 39: Modelo Conceptual Ampliado	184
Figura 40: Dimensión Cliente	188
Figura 41: Dimensión Producto	189
Figura 42: Dimensión Geografía.....	189
Figura 43: Dimensión Distribuidora	190
Figura 44: Dimensión Tiempo	190
Figura 45: Hecho Venta	192
Figura 46: Modelo Lógico del DATAMART	193
Figura 47: Secuencia ETL.....	200
Figura 48: Carga Dimensión Geografía	201
Figura 49: Carga dimensión Producto.....	203
Figura 50: Carga dimensión Distribuidora	204
Figura 51: Carga dimensión Cliente	205
Figura 52: Carga dimensión Tiempo	206
Figura 53: Carga HECHO_VENTA.....	207
Figura 54: Modelo de Cubo en SSAS	209
Figura 55: Proyecto en SSAS.....	209
Figura 56: Jerarquias para la dimension Cliente	210
Figura 57: Jerarquias para la dimensión Geografía	210
Figura 58: Jerarquias para la dimensión Producto	210
Figura 59: Jerarquias para la dimensión Tiempo	211
Figura 60: SSAS - Indicadores Principales	211

Figura 61: SSAS - Indicadores Calculados.....	211
Figura 62: SSAS - Estructura final del cubo	212
Figura 63: Modelo de dominio.....	213
Figura 64: Prototipo Formulario Login.....	214
Figura 65: Prototipo Inicio.....	214
Figura 66: Prototipo Formulario ReportesBase	215
Figura 67: Prototipo Formulario MisReportes	215
Figura 68: Modelo de Casos de Uso.....	216
Figura 69: CUS1 - Login.....	216
Figura 70: CUS2 - DescargarReporteBase	218
Figura 71: CUS3: VerMisReportes	219
Figura 72: Diagrama de Robustes Login.....	221
Figura 73: Diagrama de Robustes DescargarReportesBase.....	222
Figura 74: Diagrama de Robustes de VerMisReportes _ SubirReporte	223
Figura 75: Diagrama de Robustez VerMisReportes _ DescargarReporte	224
Figura 76: Diagrama de Secuencia Login	225
Figura 77: Diagrama de Secuencia: ReportesBase.....	226
Figura 78: Diagrama de Secuencia VerMisReportes -SubirReporte	227
Figura 79: Diagrama de Secuencia Diagrama de Secuencia VerMisReportes - DescargarReporte.....	228
Figura 80: Diagrama de base de datos.....	229
Figura 81: Excel y SSAS	233
Figura 82: Seleccionar SSAS y Cubo	233
Figura 83: Verificar conexion	234
Figura 84: Reporte de Unidades Vendidas.....	234
Figura 85: Reporte Monto total de ventas	235
Figura 86: Reporte Tasa de crecimiento	235
Figura 87: Reporte Precio Promedio.....	236
Figura 88: Reporte YTD	236
Figura 89: Reporte Ranking de Ventas.....	237
Figura 90: Creación de un nuevo Report Project Server	238
Figura 91: creación de un nuevo reporte.....	239
Figura 92: Plantilla del reporte.....	239
Figura 93: toolbox de Reporting Services	240
Figura 94: DataSource	241
Figura 95: Origen de datos (SSAS).....	241
Figura 96: Query Designer	242
Figura 97: Consulta MDX.....	242
Figura 98: Campos requeridos en el reporte	243
Figura 99: Preview del reporte.....	243
Figura 100: modo reporte en drill.....	244
Figura 101: Reporte en drill	244
Figura 102: Iniciar como administrador	245
Figura 103: Web Service Url.....	245
Figura 104: Report Manager Url	246
Figura 105: Configuración properties	246

Figura 106: Reporte deployado en la web	247
Figura 107: Reportes Online en la web del proyecto	247
Figura 108: Visores de los reportes online en la web del proyecto.....	248
Figura 109: Reporte en la web del proyecto.....	248
Figura 110: Creación de carpeta OLAP.....	249
Figura 111: Archivos a copiar en carpeta OLAP	250
Figura 112: configuración de archivo msmdpump	250
Figura 113: Configuración nuevo grupo de aplicaciones	251
Figura 114: Configuración nuevo site	251
Figura 115: Configuración nueva aplicación	252
Figura 116: Autenticación	252
Figura 117: asignación de control.....	253
Figura 118: consifuración de asignacion del controlador.....	253
Figura 119: Restricciones de Solicitud	254
Figura 120: Respuesta XML.....	254
Figura 121: Instalación Pentaho.....	255
Figura 122: Inicio pentaho	255
Figura 123: Visor Pentaho.....	256
Figura 124: Entorno Pentaho	256
Figura 125: Datasource tipo OLAP4J.....	257
Figura 126: configuración del datasource.....	257
Figura 127: Consulta MDX.....	258
Figura 128: Reporte Obtenido con pentaho	258

INDICE DE TABLAS

Tabla 1: Comparación entre sistema transaccional y datawarehouse/datamart	33
Tabla 2: Comparación modelos dimensionales	39
Tabla 3 : Cuadro comparativo de herramientas ETL.....	41
Tabla 4: Comparativa Mondrian Pentaho Vs SSAS	43
Tabla 5: Comparativa Herramientas de Reporting	44
Tabla 6: Comparación de metologías para la elaboración de un datamart	54
Tabla 7: Cuadro de puntuaciones - Metologías / Evaluación de expertos	55
Tabla 8: Operacionalización de variables.....	68
Tabla 9: Determinación de la muestra.....	70
Tabla 10: Indicador - Técnicas - Instrumentos.....	74
Tabla 11: Tabla de niveles de confianza	80
Tabla 12: Análisis Descriptivo de nivel de eficacia en el cumplimiento de las necesidades de información (Pre-Test).....	83
Tabla 13: Nivel de eficacia en el cumplimiento de las necesidades de información (Post-Test)	84
Tabla 14: Análisis Descriptivo de Tiempo de generación de reportes (Pre-Test).....	85
Tabla 15: Análisis descriptivo Tiempo de generación de reportes (Post-Test)	86
Tabla 16: Análisis Descriptivo Costo laboral total de generación de reportes (Pre-Test) ..	87
Tabla 17: Análisis descriptivo costo laboral total de generación de reportes (Post-Test) .	88
Tabla 18: Comparación de medias del nivel de eficacia en el cumplimiento de las	

necesidades de información.....	89
Tabla 19: Comparación de medias del tiempo de generación de reportes	90
Tabla 20: Comparación de medias del costo laboral total de generación de reportes.....	91
Tabla 21: Prueba Shapiro - Wilk – Nivel de eficacia en el cumplimiento de las necesidades de Información	93
Tabla 22: Prueba Shapiro Wilk -Tiempo de generación de Reportes	93
Tabla 23: Prueba Shapiro Wilk para el costo laboral total de generación de reportes.....	94
Tabla 24: Prueba de rangos de Wilcoxon – Nivel de eficacia en el cumplimiento de las necesidades de información.....	96
Tabla 25: Estadísticos de contraste Wilcoxon – Nivel de eficacia en el cumplimiento de las necesidades de información.....	96
Tabla 26: T-STUDENT Pruebas muestras relacionadas - Tiempo de generación de reportes	98
Tabla 27: T-STUDENT Pruebas muestras relacionadas - Costo Laboral Total de generación de reportes	101
Tabla 28: Identificación de Indicadores y perspectivas	160
Tabla 29: Diccionario de datos base de datos transaccional (OLTP)	175
Tabla 30: Correspondencias Perspectivas– OLTP	180
Tabla 31: Correspondencias Indicadores - OLTP.....	181
Tabla 32: Nivel de granularidad	181
Tabla 33: Tablas OLAP - Diccionario de Datos	194
Tabla 34: Campos Tablas OLAP - Diccionario de Datos	195
Tabla 35: Descripción CUS1: Login	217
Tabla 36: Descripción CUS2: DescargarReportesBase	218
Tabla 37: Descripción CUS3: VerMisReportes	220
Tabla 38: Diccionario de Datos	230

RESUMEN

Un datamart es una base de datos enfocada al almacenamiento de datos con fines análisis en un área específica de negocio. Debido a la estructura que esta base de datos posee, se obtiene información a detalle y en diferentes perspectivas de los procesos involucrados.

El objetivo principal de la tesis fue determinar la influencia de un datamart en el proceso de análisis de ventas en el área de Marketing de la empresa Laboratorios Roemmers S.A. Lo cual involucró la investigación, desarrollo e implementación de dicho datamart.

El tipo de estudio fue experimental y aplicado, el diseño de investigación fue pre-experimental. El método de investigación fue hipotético deductivo. La metodología de desarrollo del datamart usada fue HEFESTO. Así mismo, se usó el gestor de base de datos Microsoft SQL Server 2008 con la suite de BI de Microsoft (SSIS y SSAS), para mostrar los reportes se realizó una pequeña aplicación web bajo la metodología Iconix, que permite a los usuarios descargar plantillas de reportes correspondientes, y trabajar con tablas dinámicas en Excel.

La población fue de 8 reportes generados en el mes de noviembre del año 2014 en el área de Marketing de la empresa Laboratorios Roemmers S.A. La muestra fue considerada la misma que la población, por ser pequeña, para los 3 indicadores evaluados en el proceso (Nivel de eficacia en el cumplimiento de las necesidades de información, tiempo de generación de reportes y costo laboral total de generación de reportes).

Los resultados señalan que el cumplimiento de las necesidades de información antes de implementar el datamart era de 39%, y luego de su implementación aumento a 100%. En cuanto al tiempo de generación de reportes, era de 11.52 horas y luego de la implementación disminuyó a 0.61 horas. Así mismo el costo laboral total por reporte era de 735.32 soles y con la implementación del datamart disminuyó a 41.55 soles. Finalmente, después de haber obtenido resultados satisfactorios de los indicadores de estudio, se concluye que un datamart permite que la información se transforme en conocimiento, lo cual mejora el proceso de análisis de ventas en la empresa Laboratorios Roemmers S.A.

PALABRAS CLAVES: Datamart, análisis de ventas, BI, Hefesto.

ABSTRACT

A data mart is a database focused on the storage of data for analysis in a specific area of business. Because of the structure that this database has information is obtained to detail and different perspectives of the processes involved.

The main aim of the thesis was to determine the influence of a datamart in the process of analyzing sales in the area of Marketing Company S.A. Laboratories Roemmers Which involved research, development and implementation of the datamart.

The type of study was experimental and applied research design was pre-experimental. The research method was deductive hypothetical. The datamart development methodology used was HEFESTO. Likewise, the manager database Microsoft SQL Server 2008 was used with the suit Microsoft BI (SSIS and SSAS), to display reports a small web application under the Iconix methodology, which allows users to download templates are made corresponding reports, and work with PivotTables in Excel.

The population was 8 reports generated in November 2014 in the area of Marketing Company S.A. Laboratories Roemmers the sample was considered the same as the population, being small for the 3 indicators assessed in the process (Level of effectiveness in fulfilling information needs, time reporting and total labor cost of generating reports).

The results indicate that compliance with the information requirements before implementing the datamart was 39%, and after implementation increased to 100%. As for the time of reporting, it was 11.52 hours and after implementation decreased to 0.61 hours. Also the total labor cost per report was 735.32 soles and the implementation of datamart decreased to 41.55 soles. Finally, after obtaining satisfactory results of the indicators of study, it is concluded that a datamart allows information to be transformed into knowledge, which improves the process of analyzing sales Laboratories Roemmers S.A.

KEYWORDS: Datamart, sales analysis, BI, Hephaestus.