



FACULTAD DE INGENIERIA

CARRERA DE INGENIERIA DE SISTEMAS

“IMPLEMENTACIÓN DE UN DATA MART PARA EL AREA
COMERCIAL DE LA EMPRESA DE TRANSPORTES
ITTSABUS S.R.L.”

Tesis para optar el título profesional de:

Ingeniero de Sistemas

Autor:

Br. Eduardo Javier Díaz Roncal

Asesor:

Dr. Alberto Carlos Mendoza de los Santos

Trujillo – Perú

2018

APROBACIÓN DE LA TESIS

El (La) asesor(a) y los miembros del jurado evaluador asignados, **APRUEBAN** la tesis desarrollada por el(la) Bachiller **Eduardo Javier Díaz Roncal**, denominada:

**“IMPLEMENTACIÓN DE UN DATA MART PARA EL AREA COMERCIAL DE LA
EMPRESA DE TRANSPORTES ITTSABUS S.R.L.”**

Dr. Ing. Alberto Carlos Mendoza de los Santos
ASESOR

Mg. Ing. Rolando Javier Berrú Beltrán
JURADO
PRESIDENTE

Mg. Ing. Víctor Enemesio Dávila Rodríguez
JURADO

Ing. Luis Mauricio Gutiérrez Magán
JURADO

DEDICATORIA

Dedicado a mis padres, amigos y familiares que siempre estuvieron conmigo apoyándome no solo en esta tesis sino en toda mi vida, en especial a mi esposa y a mi querido hijo que han dado tanto por mí. Con cariño esta tesis se las dedico a todos ustedes.

AGRADECIMIENTO

Agradezco a Dios por darme la fortaleza de poder seguir viviendo y poder crecer en el ámbito profesional.

A mi esposa por nunca haber dejado de creer en mí y siempre haberme brindado todo su apoyo incondicional por todo el tiempo dedicado a la realización de esta tesis.

A mi pequeño hijo que siempre me motiva a ser mejor y a no claudicar en todo lo que me propongo.

A mi madre por su paciencia, confianza y apoyo incondicional en todas las etapas de mi vida, gracias por ser una excelente madre.

A mi padre que, a pesar de no haber estado muy presente en mi vida por ser un hombre de negocios, siempre me dio todo el apoyo en mis estudios universitarios para poder así trabajar en lo que más me gusta.

A cada una de las personas que de alguna manera u otra aportaron su granito de arena para poder lograr esto.

ÍNDICE DE CONTENIDOS

APROBACIÓN DE LA TESIS	ii
DEDICATORIA.....	iii
AGRADECIMIENTO.....	iv
ÍNDICE DE CONTENIDOS.....	v
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS	x
RESUMEN.....	xiii
ABSTRACT	xiv
CAPÍTULO 1. INTRODUCCIÓN.....	15
1.1. Realidad problemática	15
1.2. Formulación del problema	16
1.3. Justificación	16
1.4. Limitaciones.....	16
1.5. Objetivos	17
1.5.1. <i>Objetivo General</i>	17
1.5.2. <i>Objetivos Específicos</i>	17
CAPÍTULO 2. MARCO TEÓRICO.....	18
2.1. Antecedentes.....	18
2.2. Bases Teóricas.....	19
2.2.1. <i>¿Qué es Business Intelligence?</i>	19
2.2.2. <i>Componentes de Business Intelligence</i>	25
2.2.3. <i>Data Warehouse</i>	25
2.2.4. <i>Data Mart</i>	29
2.2.5. <i>Sistemas OLTP – Online Transactional Processing</i>	32
2.2.6. <i>Sistemas OLAP – Online Analytical Processing</i>	33
2.2.7. <i>Proceso ETL (Extract – Transform - Load)</i>	35
2.2.8. <i>Modelo Multidimensional</i>	37
2.2.9. <i>Metodología Ralph Kimball</i>	44
2.2.9.1. <i>Planificación del Proyecto</i>	46
2.2.9.2. <i>Definición de Requerimientos del Negocio</i>	46
2.2.9.3. <i>Modelado Dimensional</i>	47
2.2.9.4. <i>Diseño Físico</i>	48
2.2.9.5. <i>Diseño e Implementación del subsistema de Extracción, Transformación y Carga (ETL)</i>	49
2.2.9.6. <i>Implementación</i>	49
2.2.9.7. <i>Mantenimiento y Crecimiento del Data Warehouse</i>	49

2.2.9.8.	<i>Especificación de aplicaciones de BI</i>	50
2.2.9.9.	<i>Diseño de la Arquitectura Técnica</i>	51
2.2.10.	<i>Metodología Hefesto</i>	51
2.2.10.1.	<i>Análisis de Requerimientos</i>	52
2.2.10.2.	<i>Análisis de los OLTP</i>	53
2.2.10.3.	<i>Modelo Lógico del Data Warehouse</i>	54
2.2.10.4.	<i>Integración de Datos</i>	56
2.2.11.	<i>Metodología Bill Inmon</i>	57
2.2.11.1.	<i>Arquitectura Externa</i>	59
2.2.11.2.	<i>Arquitectura Interna</i>	60
2.3.	Definición de términos básicos.....	60
CAPÍTULO 3. HIPÓTESIS		62
3.1.	Formulación de la hipótesis.....	62
3.2.	Variables.....	62
3.3.	Operacionalización de variables.....	62
CAPÍTULO 4. DESARROLLO		63
4.1.	Desarrollo del Proyecto de Business Intelligence.....	63
4.1.1.	<i>Materiales</i>	63
4.1.1.1.	<i>Recursos Humanos</i>	63
4.1.1.2.	<i>Hardware</i>	63
4.1.1.3.	<i>Software</i>	64
4.1.1.4.	<i>Cronograma</i>	66
4.1.1.5.	<i>Costos</i>	67
4.1.1.6.	<i>Evaluación Comparativa de Metodologías BI</i>	70
4.1.2.	<i>Planificación del Proyecto</i>	71
4.1.3.	<i>Definición de los Requerimientos del Negocio</i>	72
4.1.4.	<i>Arquitectura de la Solución</i>	77
4.1.5.	<i>Modelo Dimensional</i>	78
4.1.6.	<i>Diseño Físico</i>	81
4.1.7.	<i>Procesos ETL</i>	81
4.1.8.	<i>Implementación del Data Mart</i>	106
CAPÍTULO 5. MATERIAL Y MÉTODOS		110
5.1.	Tipo de diseño de investigación.....	110
5.1.1.	<i>Tipo de Investigación</i>	110
5.1.2.	<i>Diseño de Investigación</i>	110
5.2.	Material.....	110
5.2.1.	<i>Unidad de estudio</i>	110
5.2.2.	<i>Población</i>	110
5.2.3.	<i>Muestra</i>	110
5.3.	Métodos.....	111
5.3.1.	<i>Encuesta</i>	111
5.3.2.	<i>Técnicas de Procesamiento y Análisis de Datos</i>	111

CAPÍTULO 6. RESULTADOS.....	112
6.1. Indicador Nivel de Satisfacción.....	112
6.2. Indicador Tiempo del Proceso de Toma de Decisiones.....	129
6.3. Indicador Cantidad de Reportes con Errores.....	131
CAPÍTULO 7. DISCUSIÓN	134
CONCLUSIONES.....	135
RECOMENDACIONES.....	136
REFERENCIAS.....	137
ANEXOS	139

ÍNDICE DE TABLAS

Tabla N° 1. Diferencias entre Data Warehouse y Data Mart	31
Tabla N° 2. Diferencias entre OLAP y OLTP	34
Tabla N° 3. Operacionalización de Variables	62
Tabla N° 4. Roles del Proyecto	63
Tabla N° 5. Recursos de Hardware	64
Tabla N° 6. Recursos de Software	64
Tabla N° 7. Bienes	67
Tabla N° 8. Costo Recursos Humanos.....	68
Tabla N° 9. Costos de Software	68
Tabla N° 10. Costos de Hardware	69
Tabla N° 11. Costos Totales.....	69
Tabla N° 12. Evaluación Comparativa.....	70
Tabla N° 13. Evaluación de Riesgos	72
Tabla N° 14. Funcionalidades requeridas.....	73
Tabla N° 15. Métricas requeridas	74
Tabla N° 16. Dimensiones y jerarquías	80
Tabla N° 17. Pre-Test Nivel de Satisfacción.....	112
Tabla N° 18. Post-Test Nivel de Satisfacción	112
Tabla N° 19. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según siente que es productivo en cuanto a las actividades que realiza a diario.....	112
Tabla N° 20. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según La información que les brindan los reportes del sistema actual le es suficiente	113
Tabla N° 21. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según tiene disponible la información necesaria para realizar su trabajo	114
Tabla N° 22. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Está a gusto con el tiempo que toma obtener la información comercial	115
Tabla N° 23. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información de ocupabilidad de los buses	117
Tabla N° 24. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información detallada de los canales de venta.	118
Tabla N° 25. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Puede realizar su trabajo con rapidez.	119
Tabla N° 26. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Tiene disponibilidad de la información comercial las 24hrs del día	120
Tabla N° 27. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Cree usted que actualmente puede analizar la información de manera adecuada. ..	121

Tabla N° 28. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Cree tener información de ventas exacta	122
Tabla N° 29. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Está conforme con el tiempo que involucra la toma de decisiones.....	123
Tabla N° 30. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información necesaria de los clientes	124
Tabla N° 31. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según La información actual que manejan les ayuda a predecir las ventas.....	125
Tabla N° 32. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Manejan información de indicadores para la toma de decisiones	126
Tabla N° 33. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Considera los reportes que maneja fiables	127
Tabla N° 34. Resumen Pre y Post Nivel de Satisfacción	128
Tabla N° 35. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Nivel de Satisfacción.....	128
Tabla N° 36. Datos Tabulados Tiempo Proceso de Toma de Decisiones	130
Tabla N° 37. Datos Tabulados Cantidad de Reportes Con Errores.....	131

ÍNDICE DE FIGURAS

Fig. N° 1. Transformación de datos hacia la toma de decisiones	22
Fig. N° 2. Ejemplo Ticket de Supermercado.....	23
Fig. N° 3. Estructura del Data Warehouse.....	27
Fig. N° 4. Data warehouse, fuentes de datos, preparación de datos y análisis	28
Fig. N° 5. Conjunto de Datamarts.....	30
Fig. N° 6. Data Marts Dependientes	30
Fig. N° 7. Data Marts Independientes.....	31
Fig. N° 8. Sistemas OLTP	33
Fig. N° 9. Sistemas OLAP	34
Fig. N° 10. Proceso ETL.....	37
Fig. N° 11. Cubo Multidimensional	38
Fig. N° 12. Cubo con slicing	39
Fig. N° 13. Cubo con dicing.....	40
Fig. N° 14. Cubo con roll-up	40
Fig. N° 15. Cubo con drill-down.....	41
Fig. N° 16. Ejemplo Esquema Estrella.....	43
Fig. N° 17. Ejemplo Esquema Copo de Nieve	44
Fig. N° 18. Diagrama de bloques de metodología Kimball	45
Fig. N° 19. Pasos de Metodología Hefesto.....	51
Fig. N° 20. Modelo Conceptual Ampliado.....	54
Fig. N° 21. Diseño de Tablas de Dimensiones	55
Fig. N° 22. Tabla de Hechos	56
Fig. N° 23. Arquitectura de Bill Inmon.....	59
Fig. N° 24. Componentes Internos de un Data Warehouse	60
Fig. N° 25. Cronograma de Actividades	66
Fig. N° 26. Diagrama Star Net Viajes	75
Fig. N° 27. Diagrama Star Net Forma de Pago	75
Fig. N° 28. Diagrama Star Net Ventas.....	76
Fig. N° 29. Diagrama Star Net Ocupabilidad	76
Fig. N° 30. Arquitectura de la Solución.....	77
Fig. N° 31. Diagrama Físico de Base de Datos Transaccional	79
Fig. N° 32. Diagrama Entidad-Relación.....	80
Fig. N° 33. Diagrama Físico	81
Fig. N° 34. Flujo de Datos ETL Completo.....	82
Fig. N° 35. Limpieza Total	83
Fig. N° 36. Carga Dimensión Tiempo	84

Fig. N° 37. Extraer Tipo de Servicio	86
Fig. N° 38. Extraer Oficina.....	88
Fig. N° 39. Extraer Canal de Venta	90
Fig. N° 40. Extraer Forma de Pago	92
Fig. N° 41. Extraer Destino.....	94
Fig. N° 42. Extraer Cliente.....	96
Fig. N° 43. Extraer Rango de Edad	98
Fig. N° 44. Cargar Fact Ocupabilidad.....	100
Fig. N° 45. Cargar Fact Viajes.....	102
Fig. N° 46. Cargar Fact Pagos	104
Fig. N° 47. Cargar Fact Ventas	106
Fig. N° 48. Base de Datos Data Mart	107
Fig. N° 49. Dashboard Ventas.....	108
Fig. N° 50. Dashboard Ocupabilidad	108
Fig. N° 51. Dashboard Forma de Pago	109
Fig. N° 52. Dashboard Cliente.....	109
Fig. N° 53. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según siente que es productivo en cuanto a las actividades que realiza a diario.....	113
Fig. N° 54. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según la información que les brindan los reportes del sistema actual le es suficiente	114
Fig. N° 55. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según tiene disponible la información necesaria para realizar su trabajo	115
Fig. N° 56. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Está a gusto con el tiempo que toma obtener la información comercial.....	116
Fig. N° 57. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información de ocupabilidad de los buses	117
Fig. N° 58. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información detallada de los canales de venta	118
Fig. N° 59. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Puede realizar su trabajo con rapidez.....	119
Fig. N° 60. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Tiene disponibilidad de la información comercial las 24hrs del día	120
Fig. N° 61. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Cree usted que actualmente puede analizar la información de manera adecuada.	121
Fig. N° 62. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Cree tener información de ventas exacta	122
Fig. N° 63. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Está conforme con el tiempo que involucra la toma de decisiones	123

Fig. N° 64. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información necesaria de los clientes	124
Fig. N° 65. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según La información actual que manejan les apoya en predecir las ventas.....	125
Fig. N° 66. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Manejan información de indicadores para la toma de decisiones	126
Fig. N° 67. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Considera los reportes que maneja fiables	127
Fig. N° 68. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Nivel de Satisfacción.....	129
Fig. N° 69. Prueba T de student Tiempo de Proceso de Toma de Decisiones	131
Fig. N° 70. Prueba T de student Cantidad de Reportes Con Errores	133

RESUMEN

En la presente investigación se implementa un data mart para el área comercial de la empresa de transportes ITTSABUS S.R.L., con el objetivo de mejorar el proceso de toma de decisiones. El tipo de estudio es de descriptivo experimental, ya que el propósito de la investigación es saber si la implementación de un Data Mart mejora el proceso de toma de decisiones, para este caso variable independiente (Data Mart) y variable dependiente (Proceso de Toma de Decisiones). La variable independiente tiene las siguientes dimensiones: información clave y disponibilidad de información. La variable dependiente tiene las siguientes dimensiones: tiempo, actores con acceso a la información comercial y calidad de la información.

Los métodos usados para la recolección de datos en la presente investigación fueron las encuestas cuyo instrumento fue el cuestionario y la observación cuyo instrumento fue la ficha de observación.

Los indicadores de los cuales se obtuvieron mejoras fueron: para la variable dependiente Proceso de Toma de Decisiones y su indicador Porcentaje de Reducción del Tiempo del Proceso de Toma de Decisiones, se obtuvo un resultado de reducción de tiempo de un 86.85%, para su otro indicador Cantidad de Reportes Con Errores, se obtuvo un resultado de reducción de reportes con errores de un 100.00% y para su último indicador Nivel de Satisfacción se obtuvo como resultado que los 4 ejecutivos del área comercial quedaron satisfechos. Estos Resultados fueron contrastados mediante la estadística dándonos como resultado la aceptación de la Hipótesis Alternativa (Ha) que nos dice que la implementación de un Data Mart para el área comercial de la empresa de transportes ITTSABUS S.R.L., si mejora el proceso de toma de decisiones de la misma, brindando una mejora significativa al área comercial en lo que respecta al proceso de toma de decisiones.

ABSTRACT

In the present investigation a data mart is implemented for the commercial area of the transport company ITTSABUS S.R.L., with the aim of improving the decision-making process. The type of study is experimental descriptive, since the purpose of the research is to know if the implementation of a Data Mart improves the decision-making process, for this case independent variable (Data Mart) and dependent variable (Process of Taking Decisions). The independent variable has the following dimensions: key information and availability of information. The dependent variable has the following dimensions: time, actors with access to commercial information and quality of information.

The methods used for the data collection in the present investigation were the surveys whose instrument was the questionnaire and the observation whose instrument was the observation card.

The indicators from which improvements were obtained were: for the dependent variable Process of Decision Making and its indicator Percentage of Time Reduction of the Decision Making Process, a result of reduction of time of 86.85% was obtained, for its other Indicator Amount of Reports With Errors, a result of reduction of reports with errors of 100.00% was obtained and for its last indicator Level of Satisfaction was obtained as a result that the 4 executives of the commercial area were satisfied. These results were contrasted by statistics, giving us as a result the acceptance of the Alternative Hypothesis (H_a) that tells us that the implementation of a Data Mart for the commercial area of the transport company ITTSABUS SRL, if it improves the decision-making process of the same, providing a significant improvement to the commercial area in regard to the decision-making process.

CAPÍTULO 1. INTRODUCCIÓN

1.1. Realidad problemática

Vivimos en la sociedad de la información. Gracias a internet y al desarrollo de los sistemas de información en las empresas, sus directivos pueden acceder a mucha más información de calidad y con mayor rapidez. La capacidad para tomar decisiones de negocio precisas y de forma rápida se ha convertido en un factor muy importante en las empresas exitosas. Sin embargo, los directivos cada vez tienen más información y menos tiempo para analizarla, a su vez los sistemas tradicionales suelen presentar una estructura no adecuada para llevar a cabo este fin. Si bien es cierto que los sistemas se adaptan en una menor o mayor medida para poder manejar los datos de la empresa, estos no permiten el análisis de la información adecuado que nos lleve a tomar las decisiones correctas en el día a día.

Si bien se cuenta con información, la misma se encuentra muy desordenada. El proceso de extracción y consolidación de los datos para la realización de los reportes se realiza manualmente y sistemáticamente, provocando retraso en la atención de requerimientos haciendo que en muchos casos la entrega de información no sea oportuna. Por otro lado, algunos reportes requieren mayor conocimiento en el manejo de las herramientas informáticas, por lo que son derivados al área de sistemas, ocasionando más pérdida de tiempo

Debido a esto es que ITTSABUS necesita una herramienta que les brinde la información necesaria de una manera rápida y que cumpla con los requerimientos de información del área comercial para poder realizar un correcto análisis y así puedan tomar las decisiones correctas y precisas que se necesitan en este rubro tan competitivo.

Con esta investigación se espera obtener una mejora en la rapidez para la obtención de la información por parte de los usuarios del área comercial, reduciendo la carga operativa y dependencia del área de sistemas y un mejor monitoreo de los indicadores, que permita a la gerencia comercial identificar patrones en el

comportamiento de las ventas, permitiendo obtener respuestas más acertadas basado en la demanda del mercado, para la toma de decisiones.

1.2. Formulación del problema

¿De qué manera la implementación de un data mart influye en la mejora del proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L?

1.3. Justificación

En lo económico, con esta investigación se busca mejorar el proceso de toma de decisiones, lo cual mejorará la visibilidad de las ventas en el día a día de tal manera que les permita de una manera proactiva generar nuevos servicios a demanda lo cual tendrá un impacto directo en el incremento de las ventas ya que no se perderán clientes. En lo tecnológico la implementación del data mart en el área comercial permitirá agilizar los procesos de obtención de información lo cual hará que la empresa sea más competitiva con las herramientas de explotación que serán implementadas. En lo académico la presente investigación contribuirá directamente con los gerentes y jefes en temas relacionados al campo de la gestión en el ámbito estratégico, si bien la inteligencia de negocios no es algo nuevo, el rubro al que fue aplicado en la presente investigación si lo es, lo cual permitirá generar conocimiento para que sirva como base de estudios futuros. En lo social la presente investigación busca incrementar los ingresos de la empresa ITTSABUS S.R.L., esto se verá reflejado en la evolución de las ventas lo cual tendrá un impacto positivo en la satisfacción del cliente, ya que se optimizan recursos con eficiencia y eficacia para poder tomar mejores decisiones. En lo ambiental la presente investigación ayudará a tener un mejor uso de los recursos, lo que generará menos desperdicio de materiales.

1.4. Limitaciones

- No se aprobó presupuesto para la compra de un nuevo servidor para el data mart, por lo cual se usará el mismo servidor del sistema transaccional para ahorrar costos.

- Disponibilidad del personal del área comercial, por lo cual se elaboró el cronograma de actividades con la firma del gerente general y se lo envió a todos los involucrados en el proyecto a manera de comprometerlos.

1.5. Objetivos

1.5.1. Objetivo General

Implementar un data mart para la toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

1.5.2. Objetivos Específicos

- Determinar en cuanto disminuye el tiempo del proceso de toma de decisiones del área comercial.
- Determinar en cuanto disminuye la cantidad de reportes con errores del área comercial.
- Determinar el nivel de satisfacción del personal del área comercial.

CAPÍTULO 2. MARCO TEÓRICO

2.1. Antecedentes

Alex Jesus Durand Mendoza (2014). En la investigación titulada. “*Desarrollo de un datamart para mejorar la toma de decisiones en el área de ventas de la corporación Furukawa*”, tuvo como objetivo la implementación de un data mart debido a que se tiene la necesidad de analizar las ventas realizadas por las diferentes áreas de la corporación, a su vez tomar decisiones de continuidad, de expansión o de absorción de las mismas, pero la información que se tiene está desordenada por lo que no se logra tomar decisiones correctas. La investigación me permitió conocer sobre la metodología de Ralph Kimball, la cual se basa en la creación de data marts por cada unidad de negocio para posteriormente consolidar toda la información en un data warehouse. Esto nos brinda beneficios como son: menor tiempo de desarrollo de la metodología y costos no muy elevados. La investigación concluye estableciendo una herramienta para la gerencia de ventas con la cual pueden realizar su análisis de ventas, también logra una reducción en los tiempos de generación de los reportes en un 65%, incrementando también el nivel de satisfacción de los usuarios de la herramienta.

Martha Luz Tuñoque Julcas, Oswaldo Vilchez Zapata (2016). En la investigación titulada “*Aplicación de inteligencia de negocios haciendo uso del data Warehouse 2.0 en la empresa constructora Beaver para mejorar el proceso de control de información de los centros de costos*”. La empresa no tiene una herramienta que permita analizar el historial de las liquidaciones, actividades de obra in situ, Resumen de Presupuestos de cada uno de los Centros de Costos que hay a nivel nacional. A esto se le suma el desconocimiento de liquidez con que cuentan actualmente los Centros de costos y en forma general cuánto dinero se está manejando y gastando. No se tienen estadísticas de las operaciones de obra que se realizan en los centros de costos. Por lo cual el principal objetivo de la investigación es implementar un datamart para mejorar el proceso de control de información de los centros de costos de la empresa. La investigación concluye con una reducción en horas para generar los reportes de 8 horas a 10 segundos, incrementa el número de reportes exitosos de 5

a 50 y disminuye el tiempo del proceso de generación de información de los centros de costos de 4.1 horas a 10 segundos. La investigación me permitió conocer más a detalle las herramientas que se usan para los procesos de extracción, transformación y carga de un data warehouse. Por lo cual me decidí aplicar Integration Services para diseñar los procesos ETL de Business Intelligence.

Jubitza Lisbeth Salazar Tataje (2017). En la investigación titulada *“Implementacion de Inteligencia de Negocios para el Área Comercial de la Empresa Azaleia – Basado en Metodología Agil Scrum”*. La implementación parte desde el análisis realizado al proceso del área de ventas de la empresa Azaleia del Perú, dónde se pudo evidenciar varios puntos importantes de los cuales resaltan tres. Como primer punto, el manejo de diferentes sistemas que contienen información del área, genera para el usuario carga operativa en la obtención y consolidación de esta. Como segundo punto, se encuentra la dependencia generada con el área de sistemas, debido al mantenimiento que realizan en la estructura de la base de datos para poder obtener la información solicitada por los diferentes usuarios generando un cuello de botella lo cual impacta en tiempo de respuesta en las ventas. Y como tercer punto, la integridad en la información como soporte a la toma de decisiones de la gerencia general. El objetivo general era la implementación de una plataforma de inteligencia de negocios que permita a la empresa Azaleia Perú tener un repositorio de información centralizado para poder acceder a los datos en línea. La empresa anteriormente dedicaba entre 3 y 4 días en la elaboración de reportes y gráficas que viene a ser el 50% del tiempo laboral que emplean. Luego de la implementación de la herramienta BI se redujo a solo 4 horas, también hubo un incremento en las ventas y se eliminó la dependencia del área de sistemas. La herramienta de generación para los dashboards fue Qlikview, es por esto que decidí usar la misma herramienta para realizar el dashboard con la información necesaria para el área comercial.

2.2. Bases Teóricas

2.2.1. ¿Qué es Business Intelligence?

Para definir BI, partiremos de la definición del glosario de términos de Gartner (2012):

*“BI es un **proceso interactivo** para **explorar y analizar información estructurada** sobre un **área** (normalmente almacenada en un **datawarehouse**), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones. El proceso de Business Intelligence incluye la **comunicación** de los descubrimientos y **efectuar** los cambios. Las áreas incluyen clientes, proveedores, productos, servicios y competidores.”*

Cano (2007), descompone detalladamente esta definición en lo siguiente:

- ✓ **Proceso interactivo:** cuando nos referimos a BI estamos suponiendo que se trata de un análisis de información continuo en el tiempo, no sólo en un momento puntual. Aunque evidentemente este último tipo de análisis nos puede aportar valor, es incomparable con lo que nos puede aportar un proceso continuo de análisis de información, en el que por ejemplo podemos ver tendencias, cambios, variabilidades, etc.
- ✓ **Explorar:** En toda aplicación de BI hay una etapa inicial en el que por primera vez accedemos a información que nos facilita su interpretación. En esta primera fase, lo que hacemos es “explorar” para comprender qué sucede en nuestro negocio; es posible incluso que descubramos nuevas relaciones que hasta el momento desconocíamos.
- ✓ **Analizar:** Pretendemos descubrir relaciones entre variables, tendencias, es decir, cuál puede ser la evolución de la variable, o patrones. Por ejemplo, si un cliente tiene una serie de características, cuál es la probabilidad que otro con similares características actúe igual que el anterior.
- ✓ **Información estructurada y “data warehouse”:** La información que utilizamos en BI está almacenada en tablas relacionadas entre ellas. Las tablas tienen registros y cada uno de los registros tiene distintos valores para cada uno de los atributos. Estas tablas están almacenadas en lo que conocemos como data warehouse o almacén de datos.
- ✓ **Área de análisis:** Toda aplicación de BI debe tener un objeto de análisis concreto. Nos podemos centrar en los clientes, los productos, los proveedores, los resultados de una localización, etc. Que pretendemos analizar con detalle y con un objetivo concreto: por ejemplo, la reducción de costes, el incremento de

ventas, el aumento de la participación de mercado, el ajuste de previsiones de venta, el cumplimiento los objetivos de venta presupuestados, etc.

- ✓ **Comunicar los resultados y efectuar los cambios:** Un objetivo fundamental de la BI es que, una vez descubierto algo, sea comunicado a aquellas personas que tengan que realizar los cambios pertinentes en la organización para mejorar nuestra competitividad y obtener con ello una ventaja competitiva con respecto a las demás organizaciones.

Vitt (2002), nos propone otra definición:

“El BI es usado por diferentes usuarios y desarrolladores de software para distinguir un amplio rango de tecnologías, plataformas de software, aplicaciones específicas y procesos. Se utiliza este término desde tres diferentes perspectivas:

- ✓ *Tomar mejores decisiones rápidamente.*
- ✓ *Convertir los datos en información.*
- ✓ *Utilizar un método razonable para la gestión empresarial.”*

El objetivo primario del Business Intelligence es ayudar a las personas a tomar decisiones de tal manera que le permita a una organización dirigir de la mejor forma, decidir, medir, gestionar y optimizar su eficiencia y los resultados financieros, todo esto siempre alineado a los objetivos del negocio. Para llegar a esto, las organizaciones hacen significativas inversiones en desarrollar sistemas Business Intelligence para convertir los datos originales en información de utilidad. (A. Gómez, 2013)

Según A. Gómez (2013), el interés por adoptar el Business Intelligence tiene las siguientes características:

- ✓ Buscar hechos (datos) que se puedan medir cuantitativamente acerca del negocio.
- ✓ Usar métodos organizados y tecnologías para analizar los hechos.

- ✓ Inventar o compartir modelos que expliquen las relaciones de causa y efecto entre las decisiones operativas y los efectos que éstas tienen en alcanzar los objetivos de negocio.
- ✓ Experimentar con métodos alternos y supervisar con retroalimentación sobre los resultados.
- ✓ Gestión de la empresa (decisiones e iniciativas) basadas en todas estas características.

Una definición más amplia es la que proponen en The data warehouse Institute:

“Business Intelligence (BI) es un término paraguas que abarca los procesos, las herramientas, y las tecnologías para convertir datos en información, información en conocimiento y planes para conducir de forma eficaz las actividades de los negocios. BI abarca las tecnologías de data warehousing los procesos en el ‘back end’, consultas, informes, análisis y las herramientas para mostrar información (estas son las herramientas de BI) y los procesos en el ‘front end’.”



Fig. N° 1. Transformación de datos hacia la toma de decisiones

(Víctor Armenta, 2017)

(Mendoza De los Santos, 2015)

Veamos un ejemplo para que nos ayude a comprender el verdadero significado de esta definición. Imaginemos un supermercado, la información del cual disponemos son los tickets de venta. Supongamos un sistema de información simple que está basado en la información que recogemos de las cajas registradoras. (Cano, 2007)

La información que contiene un ticket de venta es: su número, la fecha, la hora, el código de cajero/a, el código de supermercado, los códigos de los artículos vendidos, la descripción de los artículos, las unidades, el precio unitario, el total por artículo, el total del ticket y la forma de pago. Como podemos ver en la imagen: (Cano, 2007)

N° de ticket: 99999

Fecha: dd/mm/aaaa

Hora: hh:mm:ss

Código cajero: 999

Código supermercado: 999

UNIDADES	COD. ARTICULO	DESCRIPCIÓN	PRECIO UD.	TOTAL
XX	XXXXXX	aaaaaaaaa	xxxx.xx	xx,xxx.xx
XX	XXXXXX	bbbbbbbbb	xxx.xx	xx,xxx.xx

Forma de pago: AA

Total ticket: xx,xxx.xx

Fig. N° 2. Ejemplo Ticket de Supermercado

(Cano, 2007)

A partir de la información de los tickets podemos saber:

- ✓ Importe total de las ventas del día.
- ✓ Número de tickets por hora o fracción de tiempo.
- ✓ Número de tickets atendidos por un cajero/a.
- ✓ Ventas por artículo en unidades e importe.

- ✓ Número de tickets por día.
- ✓ Importe cobrado mediante efectivo o tarjetas de crédito.
- ✓ Importe del ticket medio.
- ✓ Número medio de tickets por día, hora, cajero/a.

Según Cano (2007), toda esta información es de tipo operativo, pero a este nivel nos facilita la toma de decisiones tales como:

- ✓ Reponer las existencias, acumulando la cantidad de ventas por artículo.
- ✓ Asignar los turnos de los cajeros/as, en función del número de tickets vendidos por hora.
- ✓ Ver cuáles han sido los productos más vendidos.
- ✓ Ver cuál es el medio de pago utilizado por nuestros clientes.

Según Cano (2007), si se produce una disminución de las ventas, y previamente habíamos presupuestado el número de tickets y el importe del ticket medio, podremos analizar qué ha sucedido:

- ✓ Disminución del número de tickets.
- ✓ Disminución del ticket medio.
- ✓ Una combinación de ambas.

Las respuestas nos dirán si tenemos un problema de afluencia a nuestro supermercado, o si es que los clientes nos están comprando menos de lo esperado cada vez que vienen. Las acciones a tomar son absolutamente distintas en cada caso: en el primero, deberán estar relacionadas con la promoción de nuestro supermercado para atraer clientes y en el segundo, deberemos intentar que nos compren más productos. Esta información tiene mucho más valor, ya que nos permite tomar decisiones estratégicas. Si hacemos un análisis por producto, podemos descubrir que están bajando sus ventas y, en el supuesto de que tengamos existencias con caducidad, debemos decidir rápidamente qué haremos con ellas. Si analizamos los tickets, quizás descubramos que hay relaciones entre productos: cuando un cliente compra un paquete de fideos, ¿cuál es la probabilidad de que compre una lata de tomate? Esta información es muy útil para las

promociones o para la ubicación de los productos en las estanterías de los lineales. Si hemos decidido llevar a cabo una promoción, nos interesa saber cuál ha sido su efectividad y el porqué; este aprendizaje nos permitirá plantear mejores promociones en el futuro, e indirectamente servir mejor a nuestros clientes. Y así plantear decisiones estratégicas que generen ventajas competitivas para la organización (Cano, 2007).

2.2.2. Componentes de Business Intelligence

Según Elsa Corral (2013), todas las soluciones de Business Intelligence tienen funciones parecidas, pero deben de reunir al menos los siguientes componentes:

- ✓ **Multidimensionalidad:** se refiere a toda la información que se puede encontrar en fuentes como las bases de datos, hojas de cálculo, etc. Una herramienta de Business Intelligence debe ser capaz de reunir toda la información dispersa de la empresa de tal manera que pueda proporcionar a los distintos departamentos la accesibilidad, poder y flexibilidad que necesitan para poder analizar la información.
- ✓ **Data Mining:** es un proceso que nos permite descubrir patrones en grandes volúmenes de datos.
- ✓ **Agentes:** son programas que piensan, los cuales pueden realizar tareas a un nivel muy básico sin necesidad de la intervención humana.
- ✓ **Data Warehouse:** es la respuesta de la tecnología de información a la descentralización de la toma de decisiones en la forma de una base de datos corporativa que se caracteriza por integrar y depurar información de todas las áreas funcionales y/o sistemas (ERP, CRM, aplicaciones a medida, archivos de texto, cubos, etc.).

2.2.3. Data Warehouse

Un data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización, independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios, con las propiedades siguientes: estable, coherente, fiable y con información histórica. (J. Curto, 2011)

Según el artículo “Business Intelligence” (s.f.), el término data warehouse fue acuñado por primera vez por Bill Inmon, y se traduce literalmente como almacén de datos. No obstante, y como cabe suponer, es mucho más que eso. Según definió el propio Bill Inmon, un datawarehouse se caracteriza por ser:

- ✓ **Integrado:** los datos almacenados en el datawarehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas transaccionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- ✓ **Temático:** sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno transaccional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del data warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
- ✓ **Histórico:** el tiempo es parte implícita de la información contenida en un data warehouse. En los sistemas transaccionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el data warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el data warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- ✓ **No volátil:** el almacén de información de un data warehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del data warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

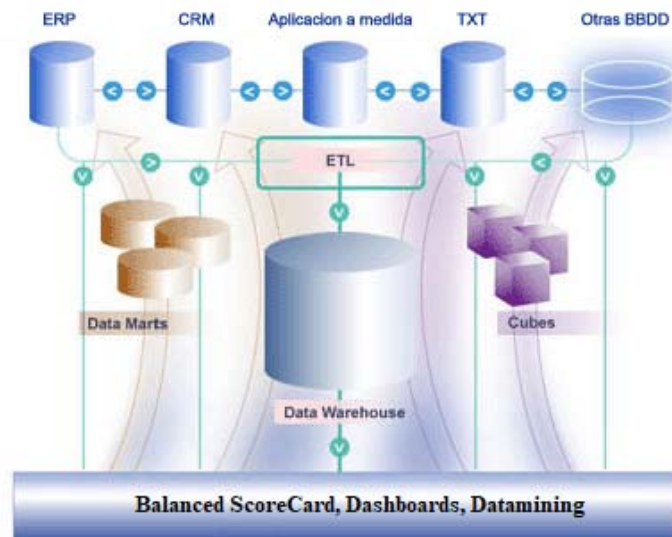


Fig. N° 3. Estructura del Data Warehouse

(sinnexus.com)

Según el artículo “Business Intelligence” (s.f.), otra característica del data warehouse es que contiene metadatos, es decir, datos sobre los datos. Los metadatos permiten saber la procedencia de la información, su periodicidad de refresco, su fiabilidad, forma de cálculo, etc.

Los metadatos serán los que permiten simplificar y automatizar la obtención de la información desde los sistemas transaccionales a los sistemas informacionales (Business Intelligence).

Según el artículo “Business Intelligence” (s.f.), los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, son:

- ✓ **Dar soporte al usuario final:** ayudándole a acceder al data warehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de Business Intelligence como DSS, EIS o CMI.
- ✓ **Dar soporte a los responsables técnicos del data warehouse en aspectos de auditoría:** gestión de la información histórica, administración del data warehouse, elaboración de programas de extracción de la información,

especificación de las interfaces para la realimentación a los sistemas transaccionales de los resultados obtenidos, etc.

Por último, destacar que para comprender íntegramente el concepto de data warehouse, es importante entender cuál es el proceso de construcción del mismo, denominado ETL (Extracción, Transformación y Carga), a partir de los sistemas transaccionales de una compañía:

- ✓ **Extracción:** obtención de información de las distintas fuentes tanto internas como externas.
- ✓ **Transformación:** filtrado, limpieza, depuración, homogeneización y agrupación de la información.
- ✓ **Carga:** organización y actualización de los datos y los metadatos en la base de datos.



Fig. N° 4. Data warehouse, fuentes de datos, preparación de datos y análisis
(Elaboración Propia)

Según el artículo “Business Intelligence” (s.f.), una de las claves del éxito en la construcción de un data warehouse es el desarrollo de forma gradual, seleccionando a un departamento usuario como piloto y expandiendo progresivamente el almacén de datos a los demás usuarios. Por ello es importante elegir este usuario inicial o piloto, siendo importante que sea un departamento con pocos usuarios, en el que la necesidad de este tipo de sistemas es muy alta y se puedan obtener y medir resultados a corto plazo.

Según el artículo “Business Intelligence” (s.f.), las principales aportaciones de un data warehouse:

- ✓ Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.
- ✓ Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- ✓ Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- ✓ Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- ✓ Supone una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares.

2.2.4. Data Mart

Es un subconjunto de los datos del data warehouse cuyo objetivo es responder a un determinado análisis, función o necesidad, con una población de usuarios específica. Al igual que en un data warehouse, los datos están estructurados en modelos de estrella o copo de nieve, y un data mart puede ser dependiente o independiente de un data warehouse. Por ejemplo, un posible uso sería para la minería de datos o para la información de marketing y ventas, producción, etc. El data mart está pensado para cubrir las necesidades de un grupo de trabajo o de un determinado departamento dentro de la organización. (J. Curto, 2011)

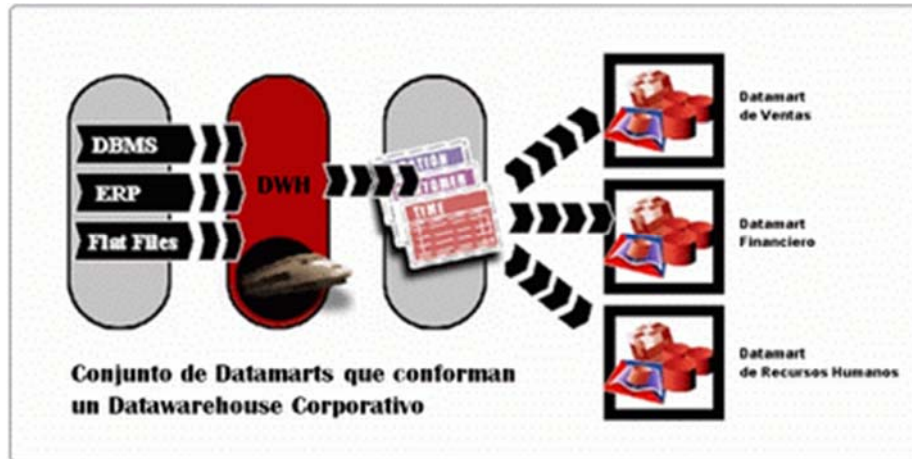


Fig. N° 5. Conjunto de Datamarts

(sinnexus.com)

Los Data Mart pueden ser independientes o dependientes. Los primeros son alimentados directamente de los orígenes de información, mientras que los segundos se alimentan desde el data warehouse corporativo. Los Data Mart independientes pueden perpetuar el problema de los “silos de información” y en su evolución pueden llegar a generar inconsistencias con otros Data Mart. (Cano, 2007)

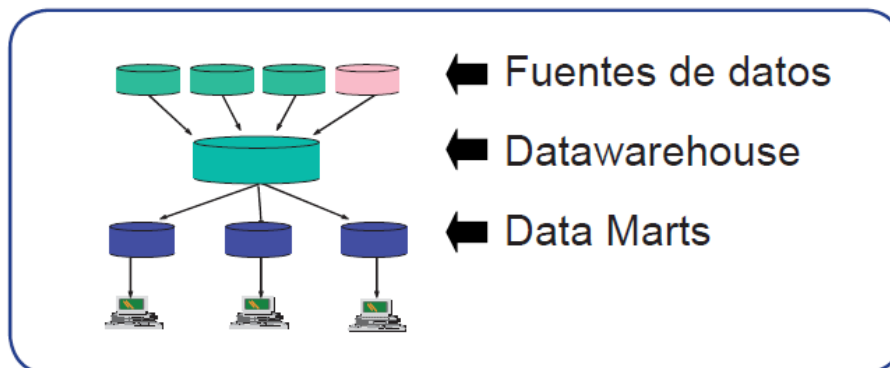


Fig. N° 6. Data Marts Dependientes

(Cano, 2007)

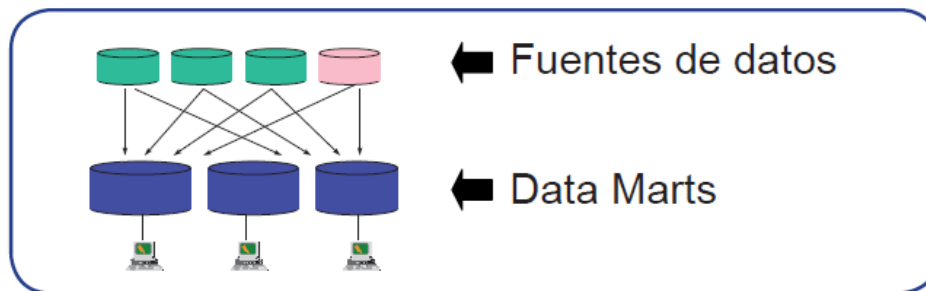


Fig. N° 7. Data Marts Independientes

(Cano, 2007)

Según Cano (2007), existen varios tipos de data mart, entre los cuales destacan:

- ✓ **Data Mart OLAP:** está basado en los cubos OLAP, que se construyen agregando, según los requisitos de cada área funcional, las dimensiones y los indicadores necesarios de cada cubo relacional.
- ✓ **Data Mart OLTP:** pueden basarse en un simple extracto del Data Warehouse, sin embargo, lo normal es introducir mejoras en su rendimiento, aprovechando las características particulares de cada área de la organización. Los data marts que están dotados con estas estructuras óptimas de análisis presentan las siguientes ventajas:
 - Poco volumen de datos.
 - Mayor rapidez de consulta.
 - Validación directa de la información.
 - Facilidad para la clasificación de datos históricos.

Las diferencias entre un Data Warehouse y un Data Mart:

Tabla N° 1. Diferencias entre Data Warehouse y Data Mart

Data Warehouse	Data Mart
----------------	-----------

Alcance	Construido para satisfacer las necesidades de toda la organización.	Construido para satisfacer las necesidades de un área de negocio específica.
Objetivo	Diseñado para optimizar la integración y la administración de los datos fuente.	Diseñado para optimizar la entrega de información para la toma de decisiones.
Características de los datos	Administra grandes cantidades de datos históricos a nivel atómico.	Se concentra en administrar resúmenes y/o datos totalizados.
Pertenencia	Pertenece a toda la organización.	Pertenece al área de negocio al cual está orientado.
Administración	Es administrado por la unidad de sistemas de la organización.	Es administrado por el personal de sistemas de la unidad propietaria del Data Mart.

Nota: Por (Cano, 2007)

2.2.5. Sistemas OLTP – Online Transactional Processing

Según el artículo “Business Intelligence” (s.f.), los sistemas OLTP son bases de datos orientadas al procesamiento de transacciones. Una transacción genera un proceso atómico (que debe ser validado con un *commit*, o invalidado con un *rollback*), y que puede involucrar operaciones de inserción, modificación y borrado de datos. El proceso transaccional es típico de las bases de datos transaccionales.

- ✓ El acceso a los datos está optimizado para tareas frecuentes de lectura y escritura. (Por ejemplo, la enorme cantidad de transacciones que tienen que soportar las bases de datos de bancos o hipermercados diariamente).
- ✓ Los datos se estructuran según el nivel aplicación (programa de gestión a medida, ERP o CRM implantado, sistema de información departamental, etc.).

- ✓ Los formatos de los datos no son necesariamente uniformes en los diferentes departamentos (es común la falta de compatibilidad y la existencia de islas de datos).
- ✓ El historial de datos suele limitarse a los datos actuales o recientes.

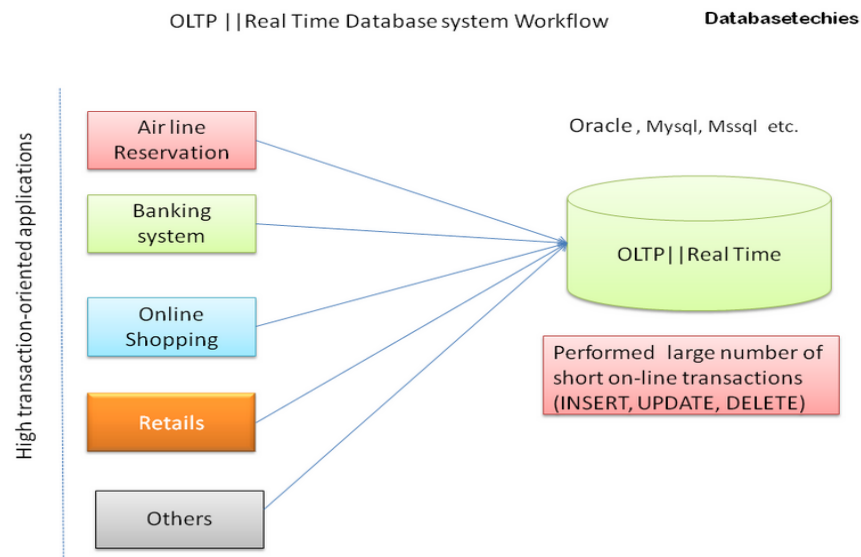


Fig. N° 8. Sistemas OLTP

(Databasetechies.com)

2.2.6. Sistemas OLAP – Online Analytical Processing

Según el artículo “Business Intelligence” (s.f.), los sistemas OLAP son bases de datos orientadas al procesamiento analítico. Este análisis suele implicar, generalmente, la lectura de grandes cantidades de datos para llegar a extraer algún tipo de información útil: tendencias de ventas, patrones de comportamiento de los consumidores, elaboración de informes complejos... etc. Este sistema es típico de los data marts.

- ✓ El acceso a los datos suele ser de sólo lectura. La acción más común es la consulta, con muy pocas inserciones, actualizaciones o eliminaciones.
- ✓ Los datos se estructuran según las áreas de negocio, y los formatos de los datos están integrados de manera uniforme en toda la organización.
- ✓ El historial de datos es a largo plazo, normalmente de dos a cinco años.

- ✓ Las bases de datos OLAP se suelen alimentar de información procedente de los sistemas operacionales existentes, mediante un proceso de extracción, transformación y carga (ETL).

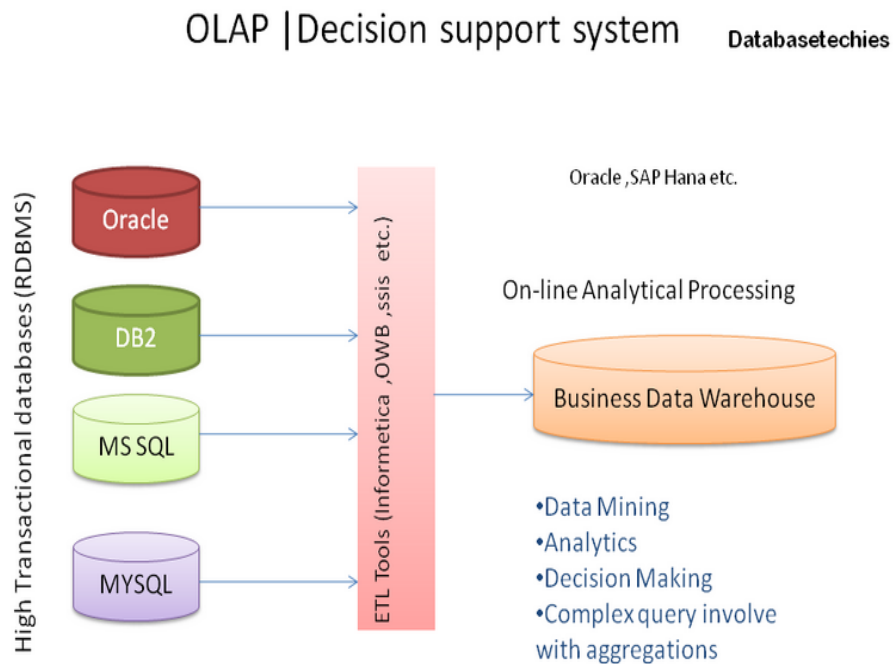


Fig. N° 9. Sistemas OLAP
(Databasetechies.com)

Las diferencias más resaltantes entre OLAP y OLTP son las siguientes:

Tabla N° 2. Diferencias entre OLAP y OLTP

Transaccional (OLTP)	Informacional (OLAP)
Almacén de datos actuales.	Almacén de datos históricos.
Almacena datos al detalle.	Almacena datos al detalle y datos agregados a distintos niveles.

Bases de datos medianas.	Bases de datos grandes.
Los datos son actuales.	Los datos son estáticos.
Los procesos son repetitivos.	Los procesos no son previsibles.
Tiempo de respuesta pequeña.	Tiempo de respuesta variable.
Toma de decisiones diarias.	Toma de decisiones estratégicas.

Nota: por (Jorge Soro, 2013)

2.2.7. Proceso ETL (Extract – Transform - Load)

El proceso trata de recuperar los datos de las fuentes de información y alimentar el data warehouse. (Cano, 2007):

El proceso de ETL consume entre el 60% y el 80% del tiempo de un proyecto de Business Intelligence, por lo que es un proceso clave en la vida de todo proyecto.

Esta parte del proceso de construcción del data warehouse es costosa y consume una parte significativa de todo el proceso, por ello requiere recursos, estrategia, habilidades especializadas y tecnologías. (Cano, 2007)

La extracción, transformación y carga (el proceso ETL) es necesario para transportar los datos de las fuentes de información hacia el data warehouse.

Según Cano (2007), el proceso ETL se divide en 5 subprocesos:

- ✓ **Extracción:** Este proceso recupera los datos físicamente de las distintas fuentes de información. En este momento disponemos de los datos en bruto. El principal objetivo de la extracción es extraer tan sólo aquellos datos de los sistemas transaccionales que son necesarios y prepararlos para el resto de los subprocesos de ETL.
- ✓ **Limpieza:** Este proceso recupera los datos en bruto y comprueba su calidad, elimina los duplicados y, cuando es posible, corrige los valores erróneos y completa los valores vacíos, es decir se transforman los datos, siempre que sea posible, para reducir los errores de carga. En este momento disponemos de datos limpios y de alta calidad. La limpieza de datos se divide en distintas etapas, que son:

- **Depurar los valores (Parsing):** Este proceso localiza e identifica los elementos individuales de información en las fuentes de datos y los aísla en los ficheros destino. Por ejemplo: separar el nombre completo en nombre, primer apellido, segundo apellido, o la dirección en: calle, numero, piso, etcétera.
 - **Corregir (Correcting):** Este proceso corrige los valores individuales de los atributos usando algoritmos de corrección y fuentes de datos externas. Por ejemplo: comprueba una dirección y el código postal correspondiente.
 - **Estandarizar (Standardizing):** Este proceso aplica rutinas de conversión para transformar valores en formatos definidos (y consistentes) aplicando procedimientos de estandarización y definidos por las reglas del negocio. Por ejemplo: trato de Sr., Sra., etc. o sustituyendo los diminutivos de nombres por los nombres correspondientes.
 - **Relacionar (Matching):** Este proceso busca y relaciona los valores de los registros, corrigiéndolos y estandarizándolos, basándose en reglas de negocio para eliminar duplicados. Por ejemplo: identificando nombres y direcciones similares.
 - **Consolidar (Consolidating):** Este proceso analiza e identifica relaciones entre registros relacionados y los junta en una sola representación.
-
- ✓ **Transformación:** Este proceso recupera los datos limpios y de alta calidad y los estructura y resume en los distintos modelos de análisis. El resultado de este proceso es la obtención de datos limpios, consistentes, resumidos y útiles. En este proceso también ajustamos el nivel de granularidad o detalle.
 - ✓ **Integración:** Este proceso valida que los datos que cargamos en el data warehouse son consistentes con las definiciones y formatos del data warehouse; los integra en los distintos modelos de las distintas áreas de negocio que hemos definido en el mismo. Estos procesos pueden ser complejos y es fundamental comprobar que se ha desarrollado correctamente, ya que en caso contrario pueden llevar a decisiones erróneas a los usuarios.

- ✓ **Actualización:** Este proceso es el que nos permite añadir los nuevos datos al data warehouse, determina la periodicidad con el que haremos nuevas cargas de datos al data warehouse.

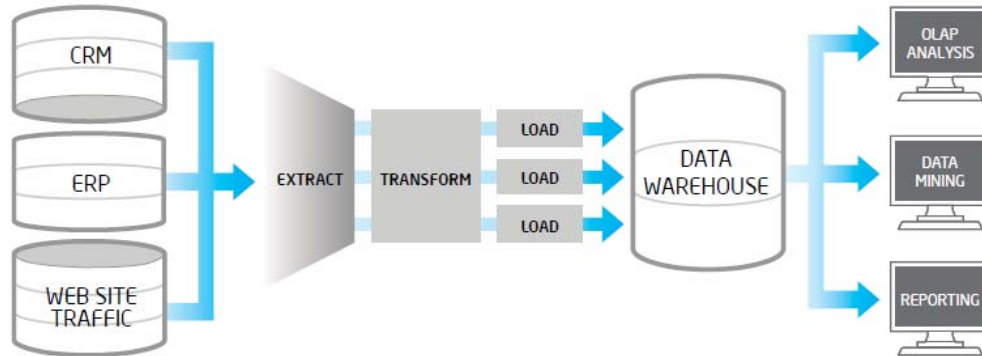


Fig. N° 10. Proceso ETL

(dataprix.com)

2.2.8. Modelo Multidimensional

Los usuarios necesitan analizar información a distintos niveles de agregación y sobre múltiples dimensiones: Por ejemplo, ventas de productos por zona de ventas, por tiempo, por clientes o tipo de cliente y por región geográfica. Los usuarios pueden hacer este análisis al máximo nivel de agregación o al máximo nivel de detalle. OLAP provee de estas funcionalidades y algunas más, con la flexibilidad necesaria para descubrir las relaciones y las tendencias que otras herramientas menos flexibles no pueden aportar. (Cano, 2007)

A estos tipos de análisis les llamamos multidimensionales, porque nos facilitan el análisis de un hecho desde distintas perspectivas o dimensiones. Esta es la forma natural que se aplica para analizar la información por parte de los tomadores de decisiones, ya que los modelos de negocio normalmente son multidimensionales. (Cano, 2007)

Según Cano (2007), los datos en un DW se modelan en data cubes (“cubos de datos”, sería su traducción literal), estructuras multidimensionales cuyas operaciones más comunes son:

- **Roll up:** incremento en el nivel de agregación de los datos.
- **Drill down:** incremento en el nivel de detalle, opuesto a roll up.
- **Slice:** reducción de la dimensionalidad de los datos mediante selección.
- **Dice:** reducción de la dimensionalidad de los datos mediante proyección.
- **Pivotaje o rotación:** reorientación de la visión multidimensional de los datos.

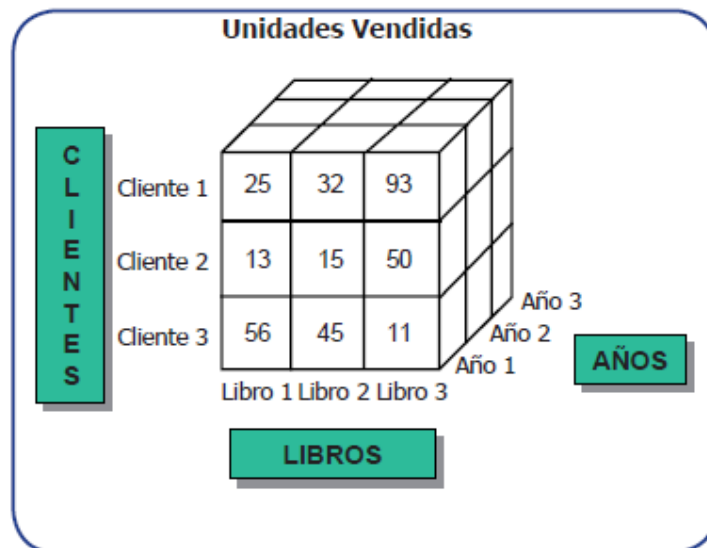


Fig. N° 11. Cubo Multidimensional

(Cano, 2007)

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. (Cano, 2007)

Las herramientas OLAP nos permiten “rotar” (en inglés “slicing”) los cubos, es decir, cambiar el orden de las distintas dimensiones: En lugar de analizar por clientes, como en el caso anterior, quizás estamos interesados en analizarlo por libros, ya que los usuarios que lo quieren consultar son distintos y tienen distintas necesidades. (Cano, 2007)

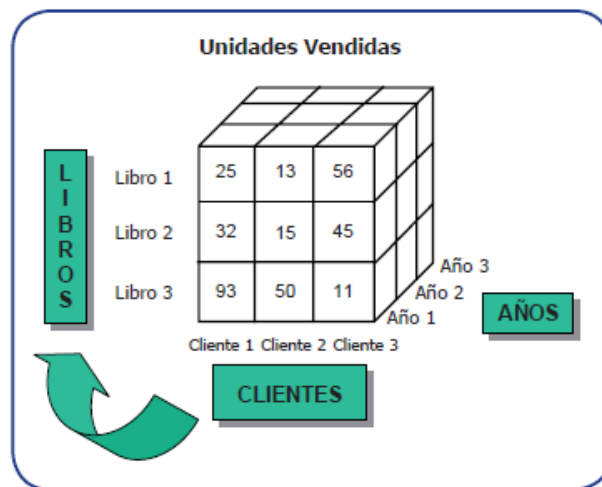


Fig. N° 12. Cubo con slicing

(Cano, 2007)

Como vemos en el ejemplo anterior, hemos cambiado la dimensión “clientes” por la de “libros”.

También podemos seleccionar (en inglés “dicing”) sólo algunas de las celdas, por ejemplo: ¿Cuáles son las ventas al cliente 2, de los libros 1 y 2, en el año 1?

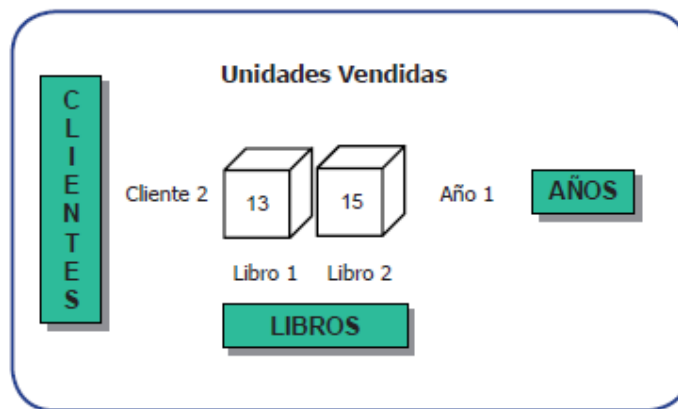


Fig. N° 13. Cubo con dicing
 (Cano, 2007)

O lo que nos puede interesar es el total de libros, máximo nivel de agregación (en inglés "roll-up"):

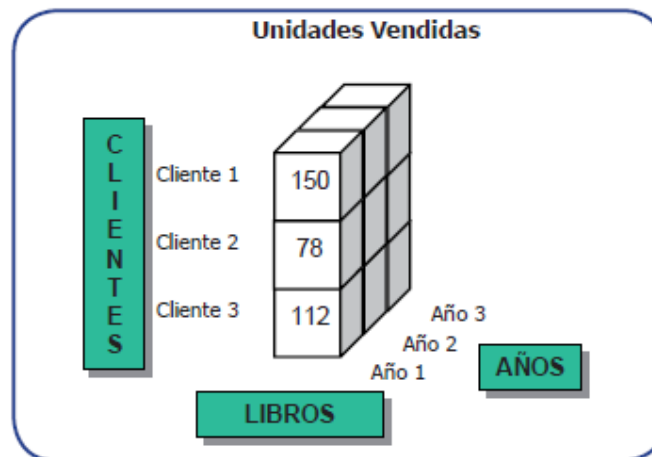


Fig. N° 14. Cubo con roll-up
 (Cano, 2007)

Imaginemos que tenemos libros de dos materias distintas: El libro 1 y el libro 2 son de la materia A y el libro 3 de la materia B. Partiendo del cubo anterior de las ventas agregadas, bajamos a más detalle (en inglés "drill-down") a través de la jerarquía "materias". En ese caso obtendríamos:

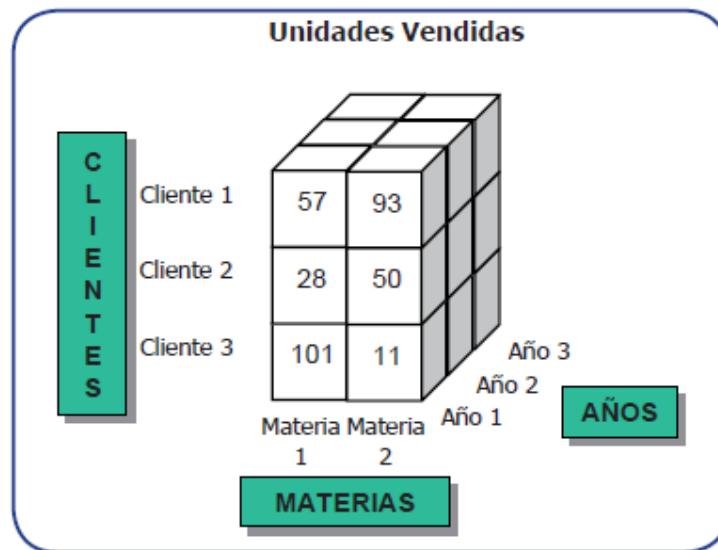


Fig. N° 15. Cubo con drill-down

(Cano,2007)

Según Cano (2007), existen distintos tipos de herramientas OLAP. La diferencia entra ellas, básicamente, depende 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 (MDDDB). 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.
- ✓ **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.

Según Cano (2007), las características y/o elementos de un modelo multidimensional son los siguientes:

- ✓ **Dimensiones:** perspectivas o entidades respecto a las cuales una organización quiere mantener sus datos organizados (p.ej. tiempo, localización, clientes, proveedores, etc.).
- ✓ **Miembros:** nombres o identificadores que marcan una posición dentro de la dimensión.
- ✓ **Jerarquías:** los miembros de las dimensiones se suelen organizar en forma de jerarquías.
- ✓ **Hechos:** colecciones de datos relacionados compuestas por medidas y un contexto.
 - Las dimensiones determinan el contexto de los hechos.
 - Cada hecho particular está asociado a un miembro de cada dimensión.
- ✓ **Medidas:** atributos numéricos asociados a los hechos (lo que realmente se mide). Ejemplos: Volumen de las ventas, coste asociado a un producto, número de transacciones efectuadas, porcentaje de beneficios.

Según Cano (2007), las características de un Esquema Estrella son:

- ✓ Una tabla de hechos que contiene los datos sin redundancias.
- ✓ Una sola tabla por dimensión.
- ✓ La tabla de hechos (*Fact table*) tiene un atributo columna que forma la clave de cada dimensión.
- ✓ Cada tabla de dimensión (*Dimension table*) es una tabla simple desnormalizada.

Cuando unimos distintos esquemas “estrella” que tienen distintas tablas de hechos, pero comparten las de las dimensiones, hablamos de constelaciones de hechos; algunos autores hablan incluso de esquema “galaxia”.

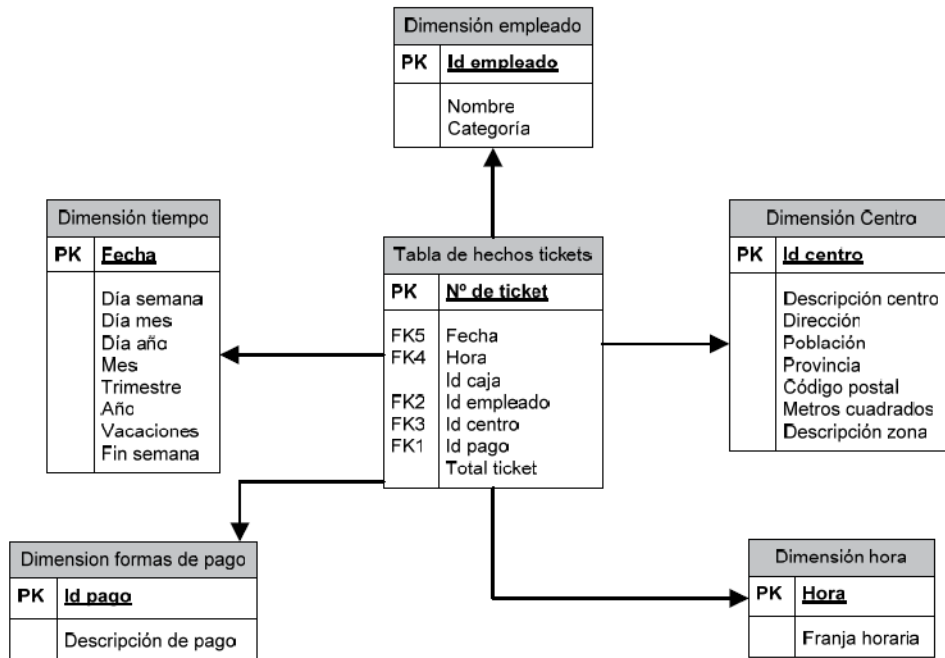


Fig. N° 16. Ejemplo Esquema Estrella

(Cano, 2007)

Según Cano (2007), las características de un Esquema Copo de Nieve son:

El esquema “estrella” no está totalmente normalizado, ya que en la tabla de la dimensión Centro tenemos una redundancia que es “Descripción zona”: Se repetirá tantas veces la zona como centros existan en la misma. El esquema “copo de nieve” o “snowflake” soluciona este problema. El esquema “copo de nieve” del ejemplo líneas más arriba sería el siguiente:

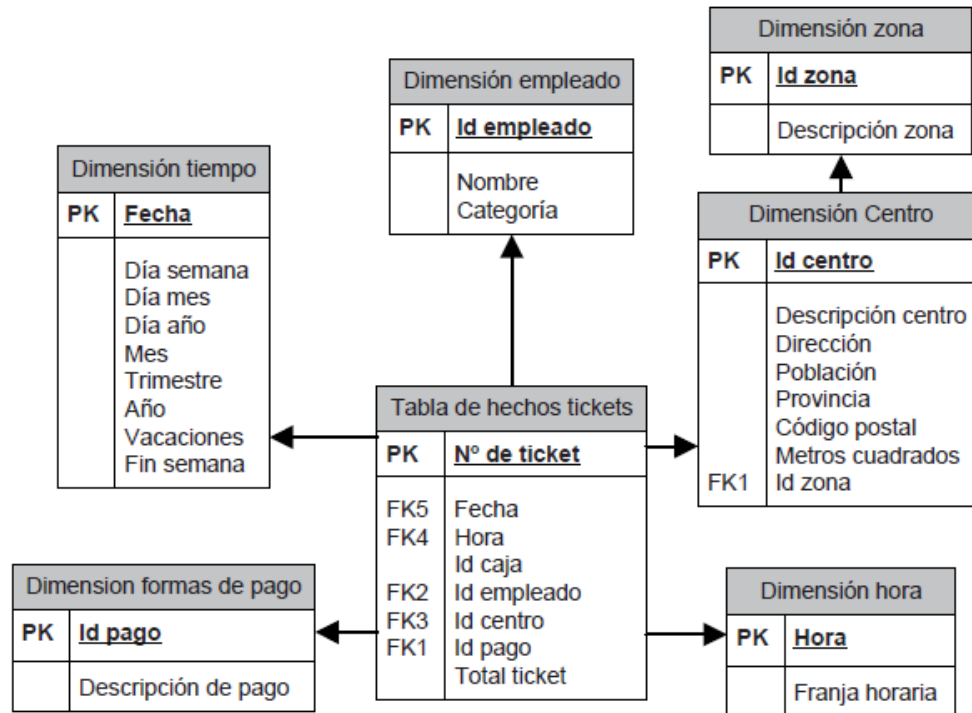


Fig. N° 17. Ejemplo Esquema Copo de Nieve

(Cano, 2007)

Como vemos, en el esquema “copo de nieve” aparecen relaciones entre las tablas de dimensiones, mientras que en el esquema “estrella” sólo hay relaciones entre la tabla de hechos y las de dimensiones. En este caso, las tablas de dimensiones están totalmente normalizadas, lo que reduce el espacio que ocupan, aunque en algunos casos esta diferencia no es significativa. (Cano, 2007)

2.2.9. Metodología Ralph Kimball

Según el artículo “Inteligencia de Negocios” (2014), la Metodología Kimball, es una metodología empleada para la construcción de un data warehouse que no es más que, una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza.

La metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida del proyecto de Data Warehouse, está basado en cuatro principios básicos:

- Centrarse en el negocio.
- Construir una infraestructura de información adecuada.
- Realizar entregas en incrementos significativos (este principio consiste en crear el Data Warehouse en incrementos entregables en plazos de 6 a 12 meses, en este punto, la metodología se parece a las metodologías ágiles de construcción de software).
- Ofrecer la solución completa (En este se punto proporcionan todos los elementos necesarios para entregar valor a los usuarios de negocios, para esto ya se debe tener un almacén de datos bien diseñado, se deberán entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación).

La construcción de una solución de Data warehouse/Business Intelligence es sumamente compleja, y Kimball nos propone una metodología que nos ayuda a simplificar esa complejidad. Las tareas de esta metodología (ciclo de vida) se describen a continuación:

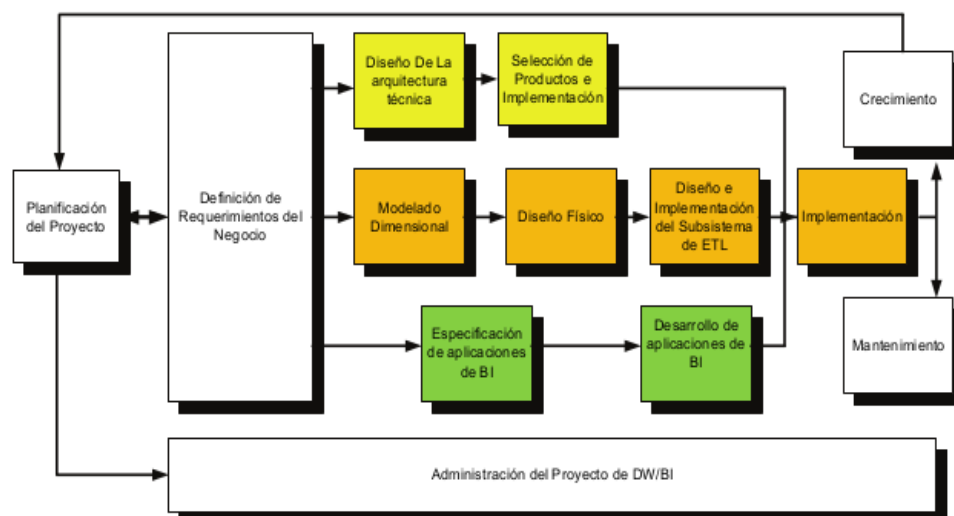


Fig. N° 18. Diagrama de bloques de metodología Kimball
 (inteligenciadenegociosval.blogspot.com)

2.2.9.1. Planificación del Proyecto

Según el artículo “Inteligencia de Negocios” (2014), en este proceso se determina el propósito del proyecto de Data Warehouse/Business Intelligence, sus objetivos específicos y el alcance del mismo, los principales riesgos y una aproximación inicial a las necesidades de información.

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.

2.2.9.2. Definición de Requerimientos del Negocio

Según el artículo “Inteligencia de Negocios” (2014), la definición de requerimientos, es un proceso de entrevistar al personal de negocio y técnico, aunque siempre conviene, tener un poco de preparación previa. En esta tarea, se debe aprender sobre el negocio, los competidores, la industria y los clientes del mismo. Se debe dar una revisión a 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 y se deben conocer los términos y la terminología del negocio.

Se sugiere entrevistar al personal que se encuentra en los cuatro grupos que se mencionan a continuación:

- El directivo responsable de tomar las decisiones estratégicas.

- Los administradores intermedios y de negocio responsables de explorar alternativas estratégicas y aplicar decisiones
- El personal de sistemas, si existe (estas son las personas que realmente saben qué tipos de problemas informáticos y de datos existen en la organización)
- El personal que se entrevista por razones políticas.

2.2.9.3. Modelado Dimensional

Según el artículo “Inteligencia de Negocios” (2014), es un proceso dinámico y altamente iterativo. Comienza con un modelo dimensional de alto nivel obtenido a partir de los procesos priorizados y descritos en la tarea anterior, y el proceso iterativo consiste en cuatro pasos:

- a) Elegir el proceso de negocio:** que consiste en, elegir el área a modelizar. Esta es una decisión de la dirección, y depende fundamentalmente del análisis de requerimientos y de los temas analíticos anotados en la etapa anterior.
- b) 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 Data Warehouse al mayor nivel de detalle posible, ya que se podrían realizar agrupamientos posteriores, al nivel deseado.
- c) 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 (que se realiza en la tarea 2.2.9.2) 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.

d) Identificar medidas y las tablas de hechos: Este 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, sumando o agrupando sus datos y usando los criterios de corte conocidos como dimensiones. Las medidas habitualmente se vinculan con el nivel de granularidad del punto 2, y se encuentran en tablas que denominamos tablas de hechos. 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, en este punto, es el nivel de detalle que posee cada registro de una tabla de hechos.

2.2.9.4. Diseño Físico

Según el artículo “Inteligencia de Negocios” (2014), en esta tarea, se contestan las siguientes preguntas:

¿Cómo puede determinar cuán grande será el sistema de Data Warehouse/Business Intelligence?

¿Cuáles son los factores de uso que llevarán a una configuración más grande y más compleja?

¿Cómo se debe configurar el sistema?

¿Cuánta memoria y servidores se necesitan? ¿Qué tipo de almacenamiento y procesadores?

¿Cómo instalar el software en los servidores de desarrollo, prueba y producción?

¿Qué necesitan instalar los diferentes miembros del equipo de Data Warehouse/Business Intelligence en sus estaciones de trabajo?

¿Cómo convertir el modelo de datos lógico en un modelo de datos físicos en la base de datos relacional?

¿Cómo conseguir un plan de indexación inicial?

¿Debe usarse la partición en las tablas relacionales?

2.2.9.5. Diseño e Implementación del subsistema de Extracción, Transformación y Carga (ETL)

Según el artículo “Inteligencia de Negocios” (2014), el subsistema de Extracción, Transformación y Carga (ETL) es la base sobre la cual se alimenta el Data Warehouse. Si 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.

2.2.9.6. Implementación

Según el artículo “Inteligencia de Negocios” (2014), a implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Existen varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y las estrategias de feedback.

2.2.9.7. Mantenimiento y Crecimiento del Data Warehouse

Según el artículo “Inteligencia de Negocios” (2014), para administrar el entorno del Data Warehouse existente es importante enfocarse en los usuarios de negocio, los cuales son el motivo de su existencia, además de gestionar adecuadamente las operaciones del Data Warehouse, medir y proyectar su éxito y comunicarse constantemente con los usuarios para establecer un flujo de retroalimentación, En esto consiste el Mantenimiento. Finalmente, es importante sentar las bases para el crecimiento y evolución del Data Warehouse en donde el aspecto clave es manejar el crecimiento y evolución de forma iterativa utilizando el Ciclo de Vida propuesto, y establecer las oportunidades de crecimiento y evolución en orden por nivel prioridad.

Si avanzamos por el camino inferior del diagrama, encontramos las tareas asociadas al área Aplicaciones de Inteligencia de Negocios, en esta ruta

se encuentran tareas en las que diseñamos y desarrollamos las aplicaciones de negocios para los usuarios finales. Las tareas pertenecientes al área, se describen a continuación.

2.2.9.8. Especificación de aplicaciones de BI

Según el artículo “Inteligencia de Negocios” (2014), en esta tarea se proporciona, a una gran comunidad de usuarios una forma más estructurada y, por lo tanto, más fácil, de acceder al almacén de datos. Se proporciona este acceso estructurado a través de lo que llamamos, Business Intelligence Applications. Las aplicaciones de Business Intelligence son la cara visible de la inteligencia de negocios: los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de Business Intelligence incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo, a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama:

- a) **Informes estándar:** son informes relativamente simples, de formato predefinido, y parámetros de consulta fijos, proporcionan a los usuarios un conjunto básico de información acerca de lo que está sucediendo en un área determinada de la empresa y se utilizan día a día.
- b) **Aplicaciones analíticas:** Son más complejas que los informes estándar. Estas aplicaciones pueden incluir algoritmos y modelos de minería de datos, que ayudan a identificar oportunidades o cuestiones subyacentes en los datos, y el usuario puede pedir cambios en los sistemas transaccionales basándose en los conocimientos obtenidos del uso de la aplicación de Business Intelligence. Algunas aplicaciones analíticas comunes incluyen:
 - Análisis de la eficacia de las promociones.
 - Análisis de rutas de acceso en un sitio Web.
 - Análisis de afinidad de programas.

- Planificación del espacio en espacios comerciales.
- Detección de fraudes.
- Administración y manejo de categorías de productos.

2.2.9.9. Diseño de la Arquitectura Técnica

Según el artículo “Inteligencia de Negocios” (2014), el área de arquitectura técnica cubre los procesos y herramientas que se aplican a los datos. En el área técnica existen dos conjuntos que tienen distintos requerimientos, brindan sus propios servicios y componentes de almacenaje de datos, por lo que se consideran cada uno aparte: El back room (habitación trasera) y el front room (habitación frontal). El back room es el responsable de la obtención y preparación de los datos, por lo que también se conoce como adquisición de datos y el front room es responsable de entregar los datos a la comunidad de usuario y también se le conoce como acceso de datos.

2.2.10. Metodología Hefesto

Según Darío en su artículo “Hefesto” (2010), nos indica que la metodología Hefesto puede resumirse en el siguiente gráfico:



Fig. N° 19. Pasos de Metodología Hefesto

(Darío, 2010)

Como se puede apreciar, se comienza recolectando las necesidades de información de los usuarios y se obtienen las preguntas claves del negocio. Luego, se deben identificar los indicadores resultantes de los interrogativos y sus respectivas perspectivas de análisis, mediante las cuales se construirá el modelo conceptual de datos del DW.

Después, se analizarán los OLTP para determinar cómo se construirán los indicadores, señalar las correspondencias con los datos fuentes y para seleccionar los campos de estudio de cada perspectiva.

Una vez hecho esto, se pasará a la construcción del modelo lógico del depósito, en donde se definirá cuál será el tipo de esquema que se implementará. Seguidamente, se confeccionarán las tablas de dimensiones y las tablas de hechos, para luego efectuar sus respectivas uniones.

Por último, utilizando técnicas de limpieza y calidad de datos, procesos ETL, etc, se definirán políticas y estrategias para la Carga Inicial del DW y su respectiva actualización.

2.2.10.1. Análisis de Requerimientos

Según Darío en su artículo “Hefesto” (2010), Lo primero que se hará será identificar los requerimientos de los usuarios a través de preguntas que expliciten los objetivos de su organización. Luego, se analizarán estas preguntas a fin de identificar cuáles serán los indicadores y perspectivas que serán tomadas en cuenta para la construcción del DW. Finalmente se confeccionará un modelo conceptual en donde se podrá visualizar el resultado obtenido en este primer paso

a) Identificar Preguntas

El primer paso comienza con el acopio de las necesidades de información, el cual puede llevarse a cabo a través de muy variadas y diferentes técnicas, cada una de las cuales poseen características inherentes y específicas, como por ejemplo entrevistas, cuestionarios, observaciones, etc.

b) Identificar Indicadores y Perspectivas

Una vez que se han establecido las preguntas de negocio, se debe proceder a su descomposición para descubrir los

indicadores que se utilizarán y las perspectivas de análisis que intervendrán.

c) **Modelo Conceptual**

En esta etapa, se construirá un modelo conceptual a partir de los indicadores y perspectivas obtenidas en el paso anterior.

A través de este modelo, se podrá observar con claridad cuáles son los alcances del proyecto, para luego poder trabajar sobre ellos, además al poseer un alto nivel de definición de los datos, permite que pueda ser presentado ante los usuarios y explicado con facilidad. (Darío, 2010)

2.2.10.2. **Análisis de los OLTP**

Según Darío en su artículo “Hefesto” (2010), nos dice que seguidamente, se analizarán las fuentes OLTP para determinar cómo serán calculados los indicadores y para establecer las respectivas correspondencias entre el modelo conceptual creado en el paso anterior y las fuentes de datos. Luego, se definirán qué campos se incluirán en cada perspectiva. Finalmente, se ampliará el modelo conceptual con la información obtenida en este paso.

a) **Conformar Indicadores**

En este paso se deberán explicitar cómo se calcularán los indicadores, definiendo los siguientes conceptos para cada uno de ellos:

- Hecho/s que lo componen, con su respectiva fórmula de cálculo. Por ejemplo: Hecho1 + Hecho2.
- Función de sumalización que se utilizará para su agregación. Por ejemplo: SUM,AVG, COUNT, etc.

b) **Establecer Correspondencias**

El objetivo de este paso, es el de examinar los OLTP disponibles que contengan la información requerida, como así también sus características, para poder identificar las correspondencias entre el modelo conceptual y las fuentes de datos.

La idea es, que todos los elementos del modelo conceptual estén correspondidos en los OLTP.

c) Nivel de Granularidad

Una vez que se han establecido las relaciones con los OLTP, se deben seleccionar los campos que contendrá cada perspectiva, ya que será a través de estos por los que se examinarán y filtrarán los indicadores. Para ello, basándose en las correspondencias establecidas en el paso anterior, se debe presentar a los usuarios los datos de análisis disponibles para cada perspectiva. Es muy importante conocer en detalle que significa cada campo y/o valor de los datos encontrados en los OLTP, por lo cual, es conveniente investigar su sentido, ya sea a través de diccionarios de datos, reuniones con los encargados del sistema, análisis de los datos propiamente dichos, etc.

d) Modelo Conceptual Ampliado

En este paso, y con el fin de graficar los resultados obtenidos en los pasos anteriores, se ampliará el modelo conceptual, colocando bajo cada perspectiva los campos seleccionados y bajo cada indicador su respectiva fórmula de cálculo. (Darío, 2010)

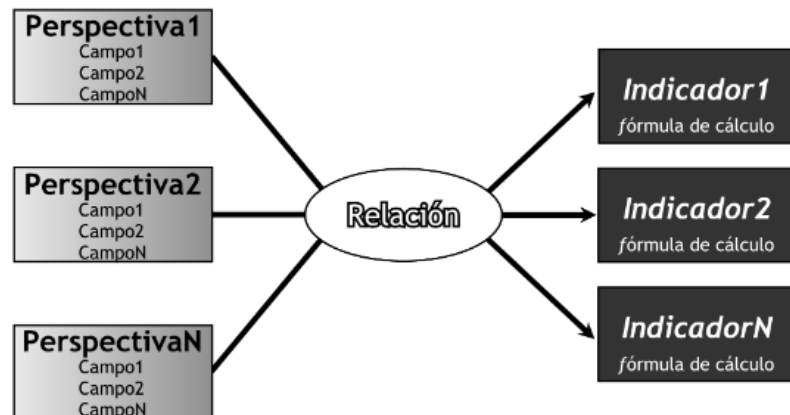


Fig. N° 20. Modelo Conceptual Ampliado

(Darío, 2010)

2.2.10.3. Modelo Lógico del Data Warehouse

Según Darío en su artículo “Hefesto” (2010), nos dice que a continuación, se confeccionará el modelo lógico de la estructura del DW, teniendo como base el modelo conceptual que ya ha sido creado. Para ello, primero se definirá el tipo de modelo que se utilizará y luego se llevarán a cabo las

acciones propias al caso, para diseñar las tablas de dimensiones y de hechos. Finalmente, se realizarán las uniones pertinentes entre estas tablas.

a) Tipo de Modelo Lógico del Data Warehouse

Se debe seleccionar cuál será el tipo de esquema que se utilizará para contener la estructura del depósito de datos, que se adapte mejor a los requerimientos y necesidades de los usuarios. Es muy importante definir objetivamente si se empleará un esquema en estrella, constelación o copo de nieve, ya que esta decisión afectará considerablemente la elaboración del modelo lógico.

b) Tablas de Dimensiones

En este paso se deben diseñar las tablas de dimensiones que formaran parte del DW.

Para los tres tipos de esquemas, cada perspectiva definida en el modelo conceptual constituirá una tabla de dimensión. Para ello deberá tomarse cada perspectiva con sus campos relacionados y realizarse el siguiente proceso:

- Se elegirá un nombre que identifique la tabla de dimensión.
- Se añadirá un campo que represente su clave principal.
- Se redefinirán los nombres de los campos si es que no son lo suficientemente intuitivos.

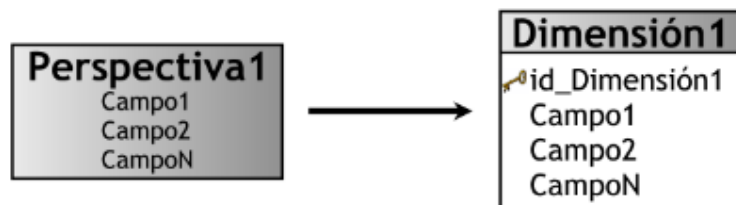


Fig. N° 21. Diseño de Tablas de Dimensiones

(Darío, 2010)

c) Tablas de Hechos

En este paso, se definirán las tablas de hechos, que son las que contendrán los hechos a través de los cuales se construirán los indicadores de estudio.

Para los esquemas en estrella y copo de nieve, se realizará lo siguiente:

- Se le deberá asignar un nombre a la tabla de hechos que represente la información analizada, área de investigación, negocio enfocado, etc.
- Se definirá su clave primaria, que se compone de la combinación de las claves primarias de cada tabla de dimensión relacionada.
- Se crearán tantos campos de hechos como indicadores se hayan definido en el modelo conceptual y se les asignará los mismos nombres que estos. En caso que se prefiera, podrán ser nombrados de cualquier otro modo.

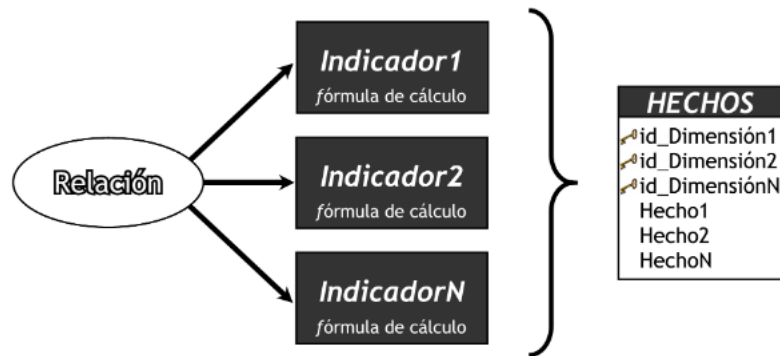


Fig. N° 22. Tabla de Hechos

(Darío, 2010)

d) Uniones

Para todos los tipos de esquemas, se realizarán las uniones correspondientes entre sus dimensiones y sus tablas de hechos.

(Darío, 2010)

2.2.10.4. Integración de Datos

Según Darío en su artículo “Hefesto” (2010), nos dice que una vez construido el modelo lógico, se deberá proceder a poblarlo con datos, utilizando técnicas de limpieza y calidad de datos, procesos ETL, etc.;

luego se definirán las reglas y políticas para su respectiva actualización, así como también los procesos que la llevarán a cabo.

a) Carga Inicial

Debemos en este paso realizar la Carga Inicial al DW, poblando el modelo de datos que hemos construido anteriormente. Para lo cual debemos llevar adelante una serie de tareas básicas, tales como limpieza de datos, calidad de datos, procesos ETL, etc. Concretamente, en este paso se deberá registrar en detalle las acciones llevadas a cabo con los diferentes softwares. Por ejemplo, es muy común que sistemas ETL trabajen con "pasos" y "relaciones", en donde cada "paso" realiza una tarea en particular del proceso ETL y cada "relación" indica hacia donde debe dirigirse el flujo de datos. En este caso lo que se debe hacer es explicar que hace el proceso en general y luego que hace cada "paso" y/o "relación". Es decir, se partirá de lo más general y se irá a lo más específico, para obtener de esta manera una visión general y detallada de todo el proceso.

b) Actualización

Cuando se haya cargado en su totalidad el DW, se deben establecer sus políticas y estrategias de actualización o refresco de datos.

Una vez realizado esto, se tendrán que llevar a cabo las siguientes acciones:

- Especificar las tareas de limpieza de datos, calidad de datos, procesos ETL, etc., que deberán realizarse para actualizar los datos del DW.
- Especificar de forma general y detallada las acciones que deberá realizar cada software. (Darío, 2010)

2.2.11. Metodología Bill Inmon

Según Barrera, Artola y Díaz en su Tesis "Análisis de Inteligencia de Negocios desde la perspectiva de los creadores del concepto" (2016), mencionan que Bill Inmon, propone que un Data Warehouse es el conjunto de datos orientados a temas específicos de las organizaciones, es decir deberán ir cambiando con el transcurso

del tiempo, pero dichos datos, no deben ser volátiles (no es posible eliminar, ni modificar la información).

Al momento de generar un cambio en los datos, se debe de efectuar de una manera que, quede reflejado el cambio que se efectuó y mantener siempre la integridad de la información, se espera que todo resultado del proceso y transformación de la información por medio de herramientas de Business Intelligence, sea un firme apoyo para los tomadores de decisiones al momento de orientar una estrategia de negocio.

Bill Inmon, considera que debido al exceso y sobrecarga de información que se está manejando en un Data Warehouse, estas bases de datos deben estar aisladas y solo manipularse para procesos de BI.

Toda información que se almacena en un Data Warehouse o Data Mart deberá ser procesada y normalizada antes.

La recomendación sobre la información, deberá estar a su máximo detalle posible, con lo cual se logre cubrir todas las necesidades de cada departamento, la importancia a la necesidad de transferir la información de los diferentes OLTP (Sistemas Transaccionales) de las organizaciones a un lugar centralizado donde los datos sean utilizados para el análisis.

Consta de las siguientes características:

- Orientado a temas: Los datos en la base de datos están estructurados de forma que los datos estén relacionados.
- Integrado: La base de datos contiene todos los datos de los sistemas operacionales de la organización, con la característica de ser consistentes.
- No volátil: Los datos no se modifica ni se elimina, almacenado los datos, éste se convierte en información de sólo lectura.
- Variante en el tiempo: Los datos que sufren un cambio a través del tiempo, deben ser registrados para que los informes generados reflejen esas variaciones.



Fig. N° 23. Arquitectura de Bill Inmon

(Barrera, Artola y Díaz, 2016)

2.2.11.1. Arquitectura Externa

Según Barrera, Artola y Díaz en su Tesis “Análisis de Inteligencia de Negocios desde la perspectiva de los creadores del concepto” (2016), nos dicen que la arquitectura que plantea Bill Inmon consta de las siguientes partes:

- a) **Fuente de la Información:** Inicia el proceso de creación de un Data Warehouse, conociendo la información que se necesita de todas las herramientas de las que se tengan acceso, ir a las necesidades de información que se necesitan con la finalidad de un resultado para crear un Data Warehouse.
- b) **Data Warehouse:** La necesidad de normalizar toda la información extraída para ser almacenada en un Data Warehouse, los cuales serán procesados y consultados por un Data Mart.
- c) **Data Marts:** Se crean un subconjunto de los datos de un Data Warehouse con el objetivo de responder a un determinado análisis o necesidad de una población, de un departamento en específico.
- d) **Explotación de los datos:** Se refiere a la manera de presentación de la información para ser consultada y analizada por las áreas.

2.2.11.2. Arquitectura Interna

Según Barrera, Artola y Díaz en su Tesis “Análisis de Inteligencia de Negocios desde la perspectiva de los creadores del concepto” (2016), nos dicen que en cuanto a la arquitectura interna de un Data Warehouse, Bill Inmon considera las siguientes características:

- a) Normalización: el Data Warehouse debe ser basado y diseñado, conforme al diseño de las bases de datos transaccionales con las que se esté interactuando.
- b) Tercera Forma normal: la prioridad es que el modelo de datos esté construido en TFN con lo cual se tenga mayor relación entre los objetos de la base de datos.

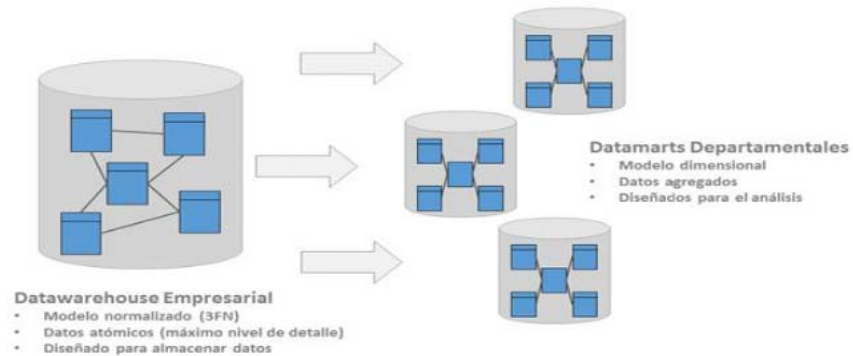


Fig. N° 24. Componentes Internos de un Data Warehouse

(Barrera, Artola y Díaz, 2016)

2.3. Definición de términos básicos

- **Data Mart:** Almacén de datos limitados compuesta por un área y/o departamento de una organización.
- **Data Warehouse:** Repositorio centralizado que almacena la información de todas las áreas de la organización.
- **OLTP (Online Transactional Processing):** Es un sistema o proceso que facilita la administración de información para aplicaciones de tipo transaccionales.
- **OLAP (Online Analytical Processing):** Es un sistema o proceso utilizado para el procesamiento de grandes volúmenes de información actualmente manejada en el campo de Business Intelligence. OLAP tiene

como objetivo principal agilizar el procesamiento de información mediante cubos que contienen resúmenes de información o sistemas transaccionales OLTP.

- **Métrica:** Unidades de medida para el análisis de información.
- **Dashboard:** Interfaz gráfica que muestra información clave para la toma de decisiones del negocio.
- **Esquema estrella:** Es la arquitectura de almacén de datos más simple. En este diseño del almacén de datos la tabla de Variables (Hechos) está rodeada por Dimensiones y juntos forman una estructura que permite implementar mecanismos básicos para poder utilizarla con una herramienta de consultas OLAP.
- **Esquema copo de nieve:** consta de una tabla de hechos que está conectada a muchas tablas de dimensiones, que pueden estar conectadas a otras tablas de dimensiones a través de una relación de muchos a uno, se encuentra en un alto nivel de normalización.
- **Enfoque “Top-down”:** Establece que el desarrollo de una solución de Business Intelligence debe partir por el desarrollo del Data Warehouse y luego este se deriva en Data Mart’s.
- **Enfoque “Bottom-up”:** Establece que el desarrollo de una solución de Business Intelligence debe partir por el desarrollo de los Data Mart’s y que todos en conjunto se constituyen en el Data Warehouse.

CAPÍTULO 3. HIPÓTESIS

3.1. Formulación de la hipótesis

Hipótesis Nula

La implementación de un data mart no mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

Hipótesis Alternativa

La implementación de un data mart mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

3.2. Variables

- Dependiente
Proceso de toma de decisiones.
- Independiente
Data mart.

3.3. Operacionalización de variables

Tabla N° 3. Operacionalización de Variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DIMENSIONES	INDICADORES
Data Mart	Conjunto de datos estructurados que mediante herramientas BI se pueden procesar, integrar y mostrar como información clave de una manera clara y rápida	Información clave	Número de reportes gerenciales
		Disponibilidad de información	Cantidad de horas diarias que el data mart estará disponible
Proceso de toma de decisiones	Proceso que involucra una toma de decisión de alto nivel partiendo de un análisis de información clave	Acceso a la información comercial	Nivel de satisfacción de personal ejecutivo
		Tiempo	Tiempo del proceso de toma de decisiones
		Calidad de la información	Cantidad de Reportes con errores

Nota: Elaboración Propia

CAPÍTULO 4. DESARROLLO

El desarrollo de la presente investigación titulada *“Implementación de un Data Mart para el área comercial de la empresa de transportes IttsaBus S.R.L.”*, empieza con una definición de los materiales requeridos para la investigación, seguido de los costos involucrados, luego se hace un comparativo de las diferentes metodologías usadas en proyectos de Business Intelligence y por último se realiza el desarrollo de todo el proyecto aplicando todas las fases de la metodología seleccionada.

4.1. Desarrollo del Proyecto de Business Intelligence

4.1.1. Materiales

En esta sección se describen los recursos requeridos para el inicio del proyecto. Cabe considerar que se indicó que se usen recursos propios de la empresa. Es decir, no se debía de adquirir licencia de hardware/software ni tampoco algún servicio de consultoría de algún tercero que implique desembolso de dinero adicional. Pero para este punto se aprobó la compra de la licencia de la herramienta QlikView por parte de Gerencia General.

4.1.1.1. Recursos Humanos

Para el desarrollo del proyecto el recurso principal fue el tesista que realizó esta investigación (ver tabla N° 4).

Tabla N° 4. Roles del Proyecto

ROL	NOMBRE
Líder del Proyecto	Eduardo Javier Díaz Roncal
Analista BI	Eduardo Javier Díaz Roncal
DBA	Eduardo Javier Díaz Roncal

Nota: Elaboración Propia

4.1.1.2. Hardware

Para el desarrollo del proyecto se usaron elementos de hardware preexistente (ver tabla N° 5).

Tabla N° 5. Recursos de Hardware

SERVICIO	DESCRIPCIÓN	CARACTERÍSTICAS TECNICAS
Database Server	1 servidor de base de datos relacional SQL Server 2014.	- 2 procesadores Intel Xeon Silver 4114 2.20 ghz.
ETL Server	1 servidor para el proceso ETL con SQL Server Integration Services.	- 64GB Ram. - SAS 2.4TB 12Gbps. - S.O. Windows Server 2012
CORE Ventas	1 servidor Linux para el servicio de ventas.	- 1 procesador Intel Xeon Silver 4108 1.80 ghz. - 16GB Ram. - SAS 900GB 12Gbps. - S.O. Linux

Nota: Elaboración Propia

4.1.1.3. Software

La infraestructura de software utilizada para la implementación del data mart ya estaba disponible a excepción de la herramienta de explotación de información QlikView lo cual involucro la compra de la licencia respectiva (ver tabla N° 6).

Tabla N° 6. Recursos de Software

SERVICIO	DESCRIPCIÓN	PRODUCTO
----------	-------------	----------

Database Server	1 servidor de base de datos relacional SQL Server 2014.	SQL Server 2014
ETL Server	1 servidor para el proceso ETL con SQL Server Integration Services.	v14.0.1000.169
CORE Ventas	1 servidor Linux para el servicio de ventas.	Linux Ubuntu v18.00
QlikView SBE	1 servidor para la herramienta de BI.	QlikView SBE 12.0
Windows Server	1 servidor para el S.O.	2012 R2 Standard

Nota: Elaboración Propia

4.1.1.4. Cronograma

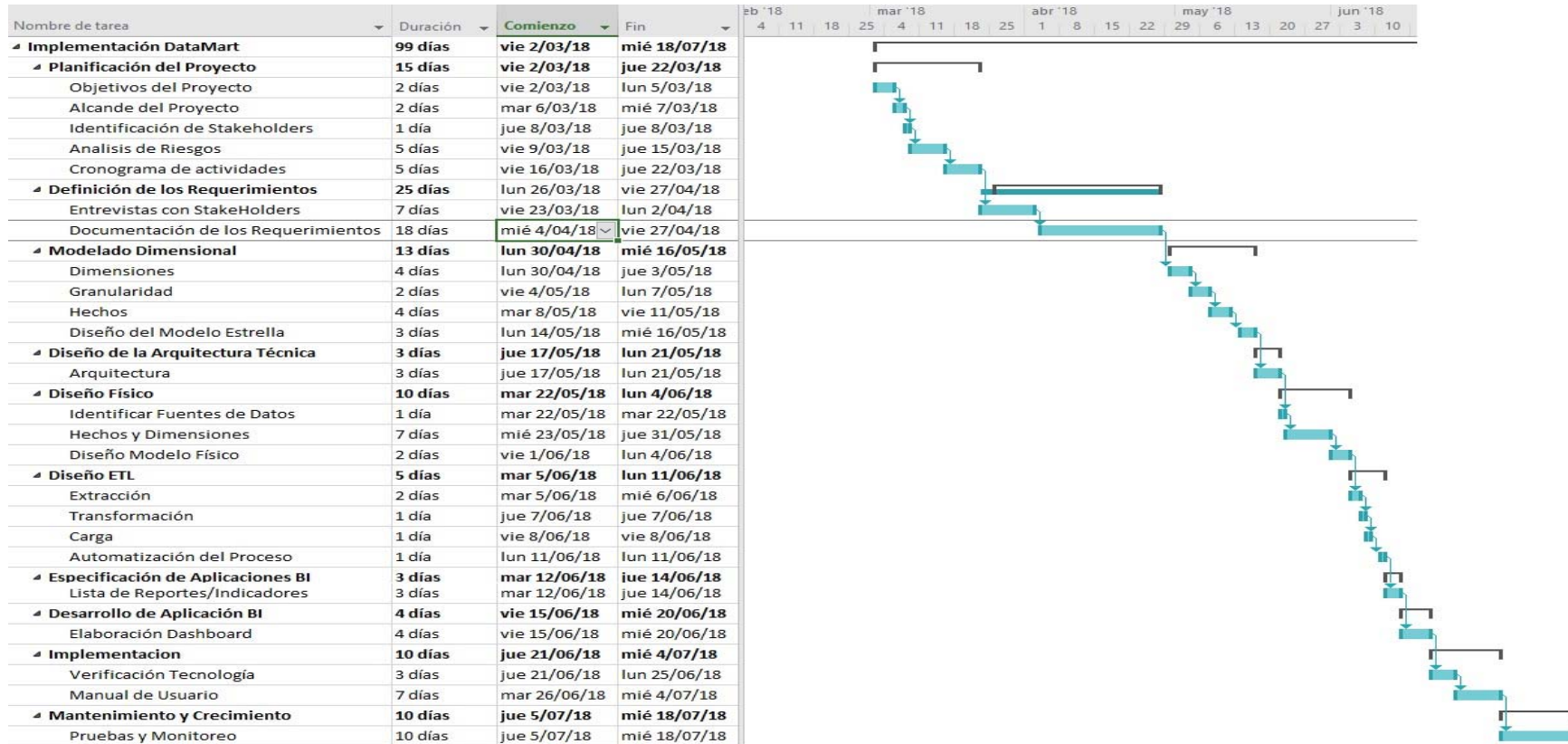


Fig. N° 25. Cronograma de Actividades

(Elaboración propia)

4.1.1.5. Costos

Costos de inversión de la presente investigación.

Bienes

La tabla N° 7 muestra los costos de bienes incurridos en la presente investigación.

Tabla N° 7. Bienes

DESCRIPCIÓN	CANTIDAD	PRECIO	TOTALES
		UNITARIO (S/.)	(S/.)
½ Millar Hoja	2	8.00	16.00
Bond			
Lapicero	3	1.00	3.00
Copias	30	0.50	15.00
Impresiones	1000	0.40	240.00
CD-ROM	1	1.00	1.00
Anillado	5	8.00	40.00
Folder	10	1.00	10.00
Costo Total (S/.)			325.00

Nota: Elaboración Propia

Inversión

Los costos de inversión son los costos de personal, software y hardware requeridos para la implementación del data mart.

Para el costo del personal se consideró un estimado del costo horas hombre y el tiempo que se utilizó para el desarrollo del proyecto (ver tabla N° 8).

Tabla N° 8. Costo Recursos Humanos

ROL	DURACIÓN		COSTO	TOTAL (S/.)
	NRO. DÍAS	HRS POR DÍA	POR HORA (S/.)	
Líder de Proyecto	84	8	21.00	14,112.00
Analista BI	38	8	17.00	5,168.00
DBA	20	8	15.00	2,400.00
Costo Total (S/.)				21,680.00

Nota: Elaboración Propia

Para los costos de software se han considerado las licencias preexistentes (sin costo) más la licencia de QlikView por la cual si se generó un costo adicional (ver tabla N° 9).

Tabla N° 9. Costos de Software

DESCRIPCIÓN	CANTIDAD	PRECIO UNITARIO (S/.)	TOTALES (S/.)
Microsoft SQL Server 2014	1	75,572.00	75,572.00
Windows Server 2012 R2 Standard	1	17,420.17	17,420.17
Linux Ubuntu	1	0.00	0.00
QlikView SBE 12	1	8,400.00	8,400.00
Costo Total (S/.)			101,932.17

Nota: Elaboración Propia, el único costo adicional incurrido fue el de QlikView

Para los costos de hardware se ha considerado el equipamiento preexistente. No se adquirieron equipos adicionales para la implementación (ver tabla N° 10).

Tabla N° 10. Costos de Hardware

DESCRIPCIÓN	EQUIPO	PRECIO UNITARIO (S/.)	TOTALES (S/.)
Database Server			
ETL Server	Dell PowerEdge		
Windows Server	R640	23,425.61	23,425.61
QlikView SBE 12			
CORE Ventas	Dell PowerEdge R640	10,065.97	10,065.97
Costo Total (S/.)			33,491.58

Nota: Elaboración Propia.

Resumen de Costos

Se presenta el resumen de costos de la presente investigación (ver tabla N° 11).

Tabla N° 11. Costos Totales

DESCRIPCIÓN	TOTALES (S/.)
Costos de Bienes	325.00
Costos de Recursos Humanos	21,680.00
Costos de Software	101,932.17
Costos de Hardware	33,491.58
Costo Total (S/.)	157,428.75

Nota: Elaboración Propia.

4.1.1.6. Evaluación Comparativa de Metodologías BI

En esta sección se realiza una evaluación comparativa asignando puntajes del 1 al 3 para algunos factores de análisis que se eligieron para este proyecto de investigación.

Tabla N° 12. Evaluación Comparativa

Factores de Análisis			
(1: No Aplicable, 2: Casi Aplicable, 3: Aplicable)	Ralph Kimball	Bill Inmon	Hefesto
Enfoque	Formar data marts para luego formar el data warehouse. (3)	Formar data warehouse para obtener data marts. (1)	Formar data warehouse para obtener data marts. (1)
Construcción de Data Warehouse	Toma menos tiempo. (3)	Toma mucho tiempo. (1)	Toma mucho tiempo. (1)
Mantenimiento	Sujeta a constantes revisiones. (2)	Revisiones sencillas. (3)	Redundante y sujeta a revisiones. (1)
Costos	Bajo costo. (3)	Alto costo. (1)	Alto costo. (1)
Tiempo de Desarrollo	Corto tiempo. (3)	Mayor tiempo. (1)	Mayor tiempo. (1)
Requisitos de Integración de datos	Áreas funcionales independientes. (3)	Toda la empresa. (1)	Toda la empresa. (1)
Afinidad con el sistema actual	Alto. (3)	Bajo. (1)	Bajo. (1)
Guías y practicas se aplican a SQL	Sí. (3)	Sí. (3)	Sí. (3)

Fácil entendimiento	Alto. (3)	Alto. (3)	Medio. (2)
TOTAL	26	15	12

Nota: Elaboración Propia

Luego de la respectiva evaluación de los factores de análisis tomados en cuenta, concluimos que la metodología de Ralph Kimball es la que más se adecua a este proyecto de investigación.

4.1.2. Planificación del Proyecto

El propósito de implementar el data mart para el área comercial de la empresa de transportes ITTSABUS S.R.L., tiene por objetivo proporcionar una herramienta de Business Intelligence que permita mejorar la toma de decisiones de los ejecutivos del área comercial.

Los stakeholders o interesados en la implementación de este proyecto el Gerente General, Gerente Comercial, Jefe Comercial, Jefe de Marketing y Jefe de Sistemas.

En esta actividad se definió como alcance funcional que el Data Mart contendrá las ventas efectuadas en los respectivos puntos de venta y el nivel de granularidad de la información será de las ventas efectuadas por día (esto incluye la ocupabilidad, canales de venta, forma de pago, servicio más vendido, viajero frecuente).

Se definió la infraestructura de hardware, descrita en el punto 4.1.1.2, y la infraestructura de software, descrita en el punto 4.1.1.3., también se definió el cronograma de actividades (Gantt), descrito en el punto 4.1.1.4.

Se realizó una evaluación de riesgos, de los cuales se identificaron aquellos que tienen impacto directo en el desarrollo del proyecto. La tabla N° 13 muestra los riesgos identificados y el impacto de estos. Los riesgos se encuentran clasificados en una escala de probabilidad entre el 1% y 100%. Se puede concluir que el riesgo R1-03 es el de mayor impacto, esto debido a que un retraso en el desarrollo del

proyecto afectará económicamente al mismo, así como también retrasará los entregables.

Tabla N° 13. Evaluación de Riesgos

CODIGO	RIESGO	PROBABLIDAD	IMPACTO		MITIGACIÓN
			MIN:1 MAX:5	PUNTAJE	
R1-01	Inadecuado levantamiento de información.	20%	5	1	Considerar todos los factores posibles durante la etapa de levantamiento de información.
R1-02	Mala definición del alcance.	10%	3	0.3	Definir el alcance de acuerdo a los requerimientos del negocio.
R1-03	Tiempo mayor al programado.	80%	5	4	Definir holguras, monitorear desviaciones y establecer contingencias.
R1-04	Elección de herramientas de explotación inadecuada	30%	4	1.2	Evaluar con el usuario la elección de la herramienta de explotación.

Nota: Elaboración Propia

4.1.3. Definición de los Requerimientos del Negocio

Los requerimientos generales solicitados por los usuarios para la implementación del data mart fueron los siguientes:

- a) Poder manejar la información a nivel de cliente, de tal manera que se obtengan esos datos de la transacción realizada en cada punto de venta.
- b) Tener dashboards que muestren toda esa información con un nivel de detalle adecuado.
- c) Contar con datos confiables y seguros que permitan tomar decisiones de negocio sin necesidad de realizar verificaciones adicionales para verificar su certeza.

- d) Poder realizar planificaciones y tendencias con el análisis de la información, por lo cual se debe la capacidad de almacenamiento histórico.
- e) Minimizar la dependencia con el área de sistemas de la organización.

Se identificaron las funcionalidades a ser implementadas en el data mart (ver tabla N° 14).

Tabla N° 14. Funcionalidades requeridas

FUNCIONALIDAD	DESCRIPCIÓN
Consultas por oficina.	Se requiere mostrar la información de ventas por oficina.
Consultas por fecha de transacción.	Se requiere mostrar la información de ventas por fecha de operación. Año, trimestre, mes y día.
Consultas por forma de pago.	Se requiere mostrar la información de ventas por forma de pago. Efectivo, mixto, nota de cobranza, transferencia bancaria, tarjeta.
Consultas por tipo de servicio.	Se requiere mostrar la información de ventas por los tipos de servicio. MIXTO 56, MIXTO 53, SOFA 42, SOFA 39, CAMA 32, SOFA 43, SOFA 40, MIXTO ESPECIAL 56, MIXTO ECONOMICO.
Consultas por canal de venta.	Se requiere mostrar la información de ventas por canal de venta. Call Center, Online, Ventanilla.
Consulta ocupabilidad por destinos.	Se requiere mostrar la información de ocupabilidad por todos los destinos (Origen-Destino)
Consulta viajes de clientes.	Se requiere mostrar la información de los clientes que tengan más de 80 viajes en el año.

Consulta rango de edad de clientes.	Se requiere mostrar la información de los clientes por rangos de edad. [18-24], [25-34], [35-44], [Mayor]. Para los que no tengan fecha de nacimiento mostrar [Edad Inválida].
-------------------------------------	--

Nota: Elaboración Propia

Las funcionalidades deben analizarse a través de métricas o medidas. Se identificaron métricas durante las reuniones con los usuarios (ver tabla N° 15).

Tabla N° 15. Métricas requeridas

MÉTRICA	DESCRIPCIÓN	FÓRMULA
Volumen de ventas	Cantidad de boletos vendidos en un determinado período de tiempo.	Sumatoria de la cantidad de boletos vendidos.
Monto de ventas	Monto en nuevos soles de las ventas en un determinado periodo de tiempo.	Sumatoria de los montos vendidos.
Ocupabilidad de servicio	Cantidad de asientos ocupados en un determinado periodo de tiempo.	Sumatoria de los asientos ocupados.
Viajes	Cantidad de Viajes de un cliente en un determinado periodo de tiempo.	Sumatoria de los viajes.
Edades	Cantidad de Clientes de un cierto rango de edad en un determinado periodo de tiempo.	Sumatoria de clientes dentro del rango.

Nota: Elaboración Propia

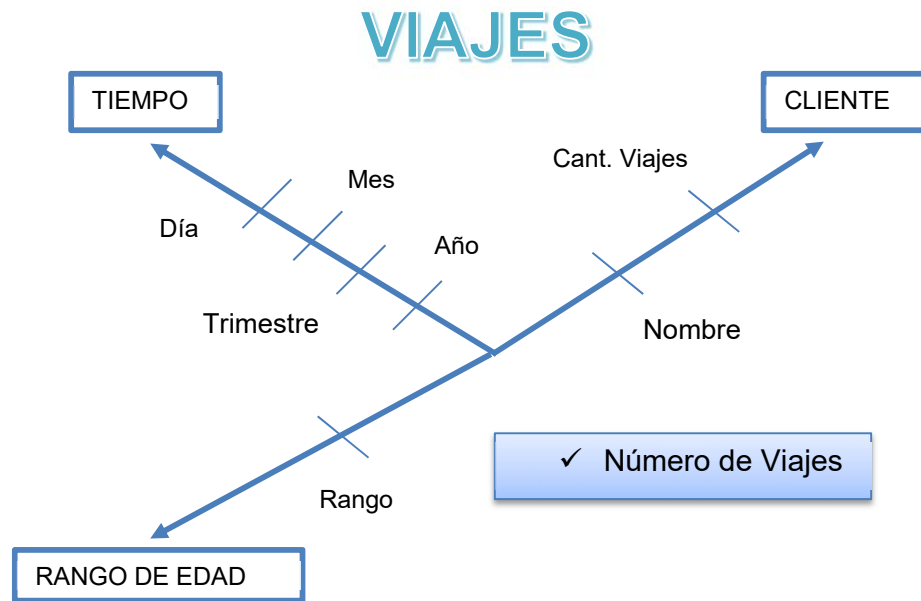


Fig. N° 26. Diagrama Star Net Viajes
(Elaboración Propia)

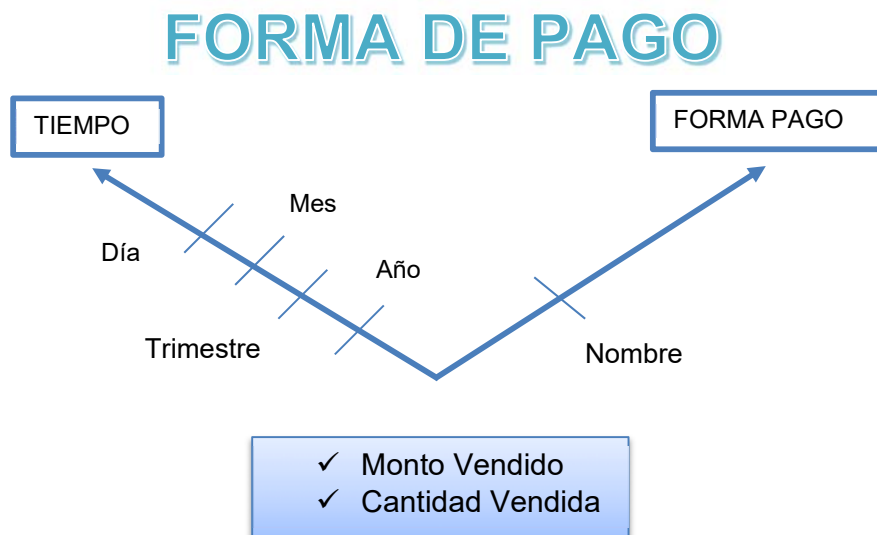


Fig. N° 27. Diagrama Star Net Forma de Pago
(Elaboración Propia)

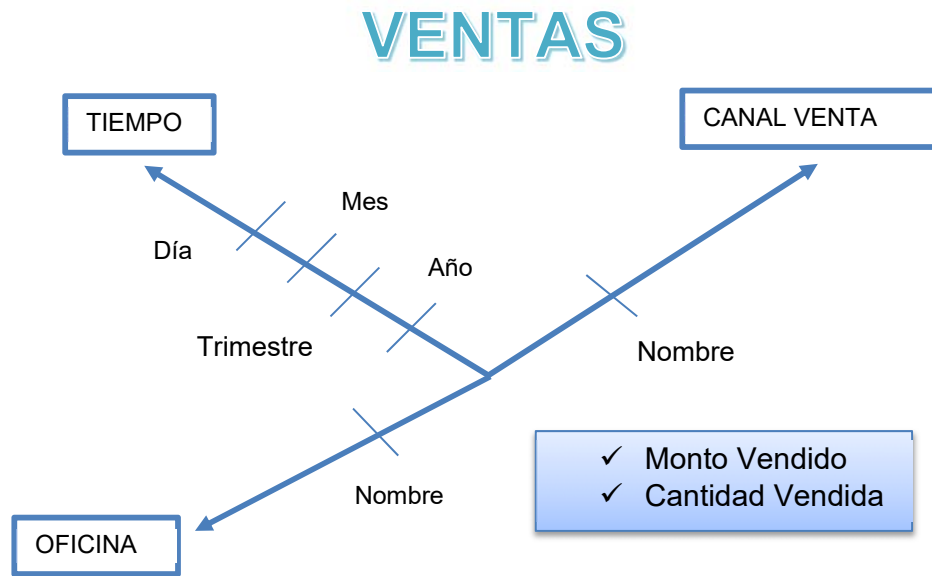


Fig. N° 28. Diagrama Star Net Ventas
(Elaboración Propia)

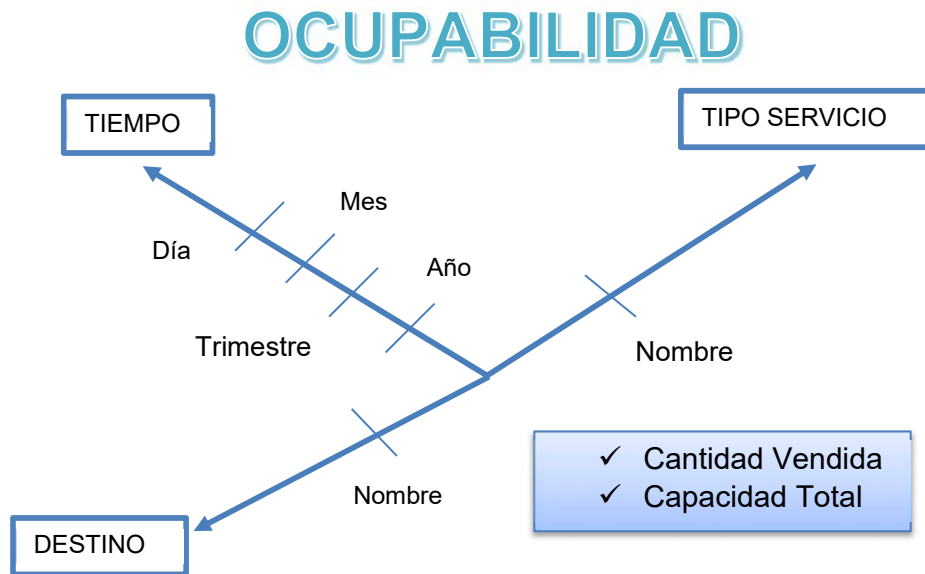


Fig. N° 29. Diagrama Star Net Ocupabilidad
(Elaboración Propia)

4.1.4. Arquitectura de la Solución

El diseño de la arquitectura para la solución nos servirá para ver el flujo de la información desde el inicio hasta el final y que componentes la constituyen.

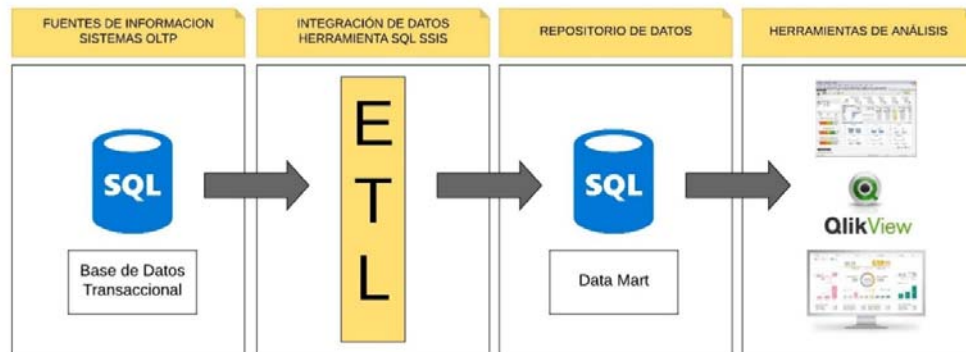


Fig. N° 30. Arquitectura de la Solución

(Elaboración propia)

Fuentes de Información

Las fuentes de información y/o fuente de datos que se usará será la base de datos transaccional que está en producción desde el año 2017, aquí se almacena la información del día a día de todas las agencias a nivel nacional de manera centralizada.

Integridad de Datos

Se extraerán y ordenarán los datos para el data mart así como la ejecución del proceso ETL. En la extracción se tomarán datos del sistema OLTP actual que está en una base de datos SQL Server 2014, para esto se mapearan las tablas que están involucradas en los requerimientos definidos en el punto 4.1.3.

Para la transformación se usarán diferentes criterios para que la información seleccionada sea entendible para los usuarios que van a utilizar está información para realizar el respectivo análisis.

Por ultimo una vez extraídos y transformados los datos, viene la carga de datos hacia el data mart en su respectivo modelo multidimensional consistente de sus

tablas de hechos y dimensiones, para que puedan ser utilizados por los usuarios respectivos.

Repositorio de Datos

En cuando al repositorio de datos, será almacenado en una base de datos SQL Server 2014 en el mismo servidor de base de datos del sistema OLTP.

Herramientas de Análisis

En el mercado hay muchas herramientas de análisis de datos, entre ellos QlikView, el cual fue elegido debido al conocimiento previo de la herramienta, reportería y análisis incluido, rápida implementación y facilidad de uso.

4.1.5. Modelo Dimensional

El modelo dimensional diseñado para el Data Mart del área comercial se centra en 4 puntos importantes:

- ✓ Ventas.
- ✓ Ocupabilidad.
- ✓ Forma de Pago.
- ✓ Viajes por Cliente.

En la Figura N° 31, podemos apreciar algunas de las tablas más importantes de la base de datos transaccional (punto de venta) de la empresa ITTSABUS, específicamente las tablas que guardan relación con el modelo multidimensional.

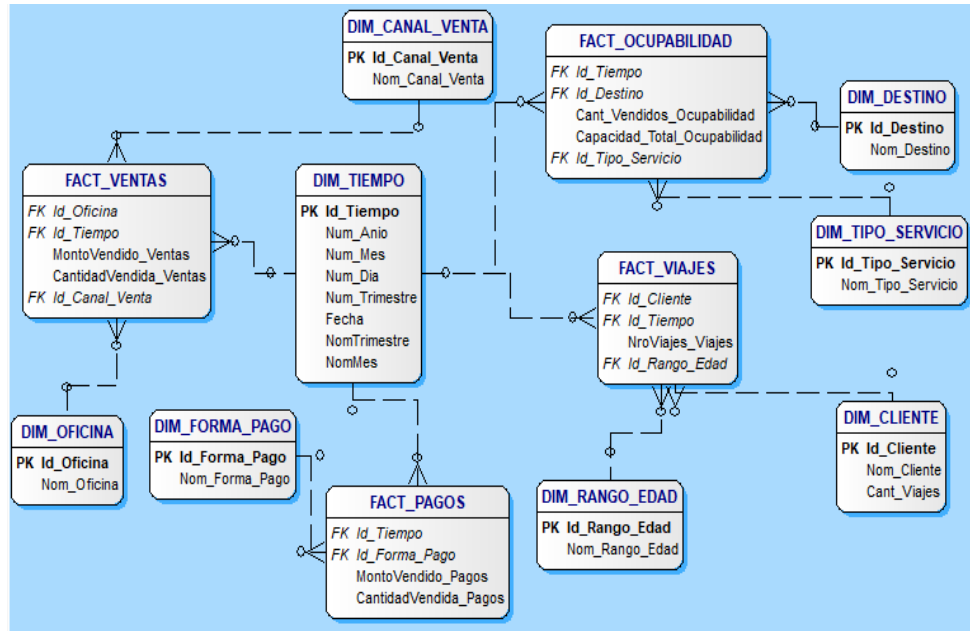


Fig. N° 32. Diagrama Entidad-Relación
(Elaboración propia)

Las dimensiones con sus respectivos atributos se almacenan físicamente en tablas de la base de datos relacional. A cada dimensión y hecho le corresponde una tabla en el diseño físico.

En la Tabla N° 16 están las dimensiones y su correspondiente tabla del modelo físico. Cabe mencionar que en esta tabla aparecen los atributos requeridos dentro de la funcionalidad solicitada.

Tabla N° 16. Dimensiones y jerarquías

DIMENSIÓN	NOMBRE DE TABLA	ATRIBUTO	JERARQUIA
Tipo de Servicio	Dim_Tipo_Servicio	Nom_Tipo_Servicio	Nom_Tipo_Servicio
Oficina	Dim_Oficina	Nom_Oficina	Nom_Oficina
Canal Venta	Dim_Canal_Venta	Nom_Canal_Venta	Nom_Canal_Venta
Destino	Dim_Destino	Nom_Destino	Nom_Destino
Tiempo	Dim_Tiempo	Tiempo	Año, Trimestre,

			Mes, Día
Rango de Edad	Dim_Rango_Edad	Nom_Rango_Edad	Nom_Rango_Edad
Cliente	Dim_Cliente	Nom_Cliente Cant_Viajes	Nom_Cliente, Cant_Viajes
Forma de Pago	Dim_Forma_Pago	Nom_Forma_Pago	Nom_Forma_Pago

Nota: Elaboración Propia

4.1.6. Diseño Físico

El diseño físico se centra en las tablas de hechos, la cual está relacionada con una o más tablas dimensionales. En la Figura N° 33 se puede apreciar el modelo físico del data mart para el área comercial.

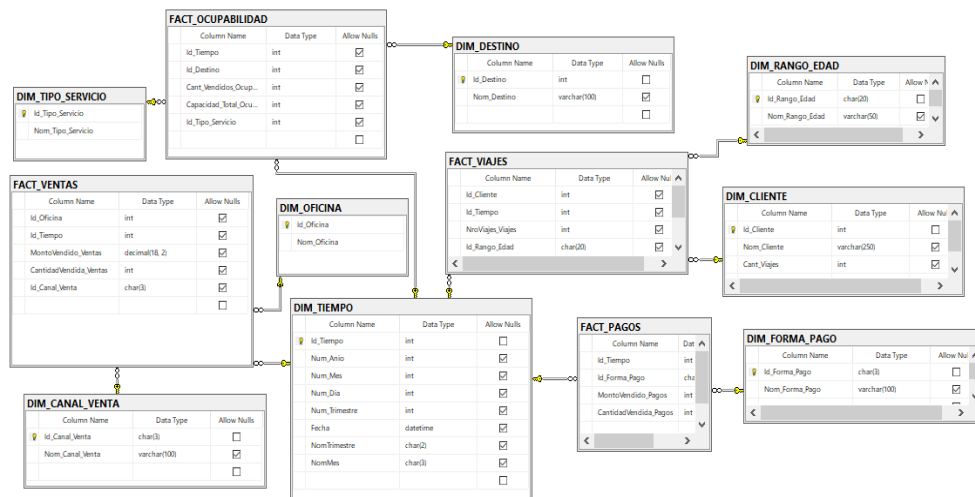


Fig. N° 33. Diagrama Físico

(Elaboración propia)

Cada tabla dimensional del modelo físico tiene una llave maestra que está relacionada a la tabla de hechos en su respectivo atributo mediante una clave foránea, lo cual le brinda al modelo una integridad referencial.

4.1.7. Procesos ETL

El proceso ETL que consiste en las etapas de Extracción, Transformación y Carga se ejecuta en el mismo servidor de producción del Core Punto de Venta dentro de

unas horas especificadas. Para la construcción del flujo de datos ETL se utilizó Microsoft Integration Services 2014 por ya contar con la licencia. Los orígenes de datos principalmente fueron obtenidos del Core Punto de Venta de ITTSA que tiene su base de datos en SQL Server 2014 la cual se encuentra en el mismo servidor. A continuación, se presenta la vista general de todo el proceso ETL. También se definió que el proceso ETL que alimenta el data mart será ejecutado a las 07:00am, 10:00am, 01:00pm y 06:00pm.

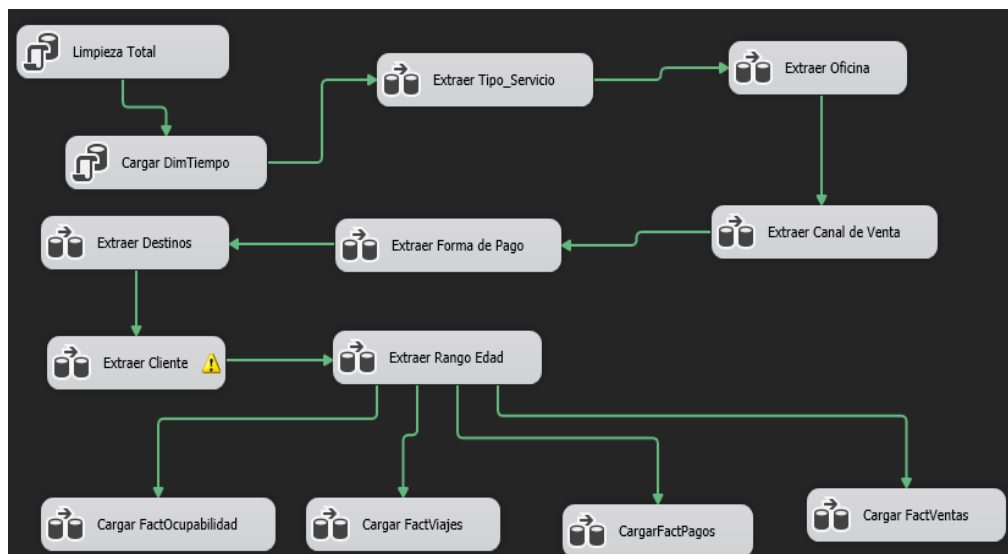


Fig. N° 34. Flujo de Datos ETL Completo

(Elaboración propia)

En la Figura N° 34 podemos apreciar todas las etapas del proceso ETL, empezaremos con la extracción y carga de las tablas dimensionales para terminar con la carga de las tablas de hechos.

Limpieza Total

Este flujo de datos realiza una limpieza de todas las tablas dimensionales y tablas de hechos, esto para tener siempre información actualizada cada vez que se ejecute la tarea en el motor de base datos Microsoft SQL Server 2014.

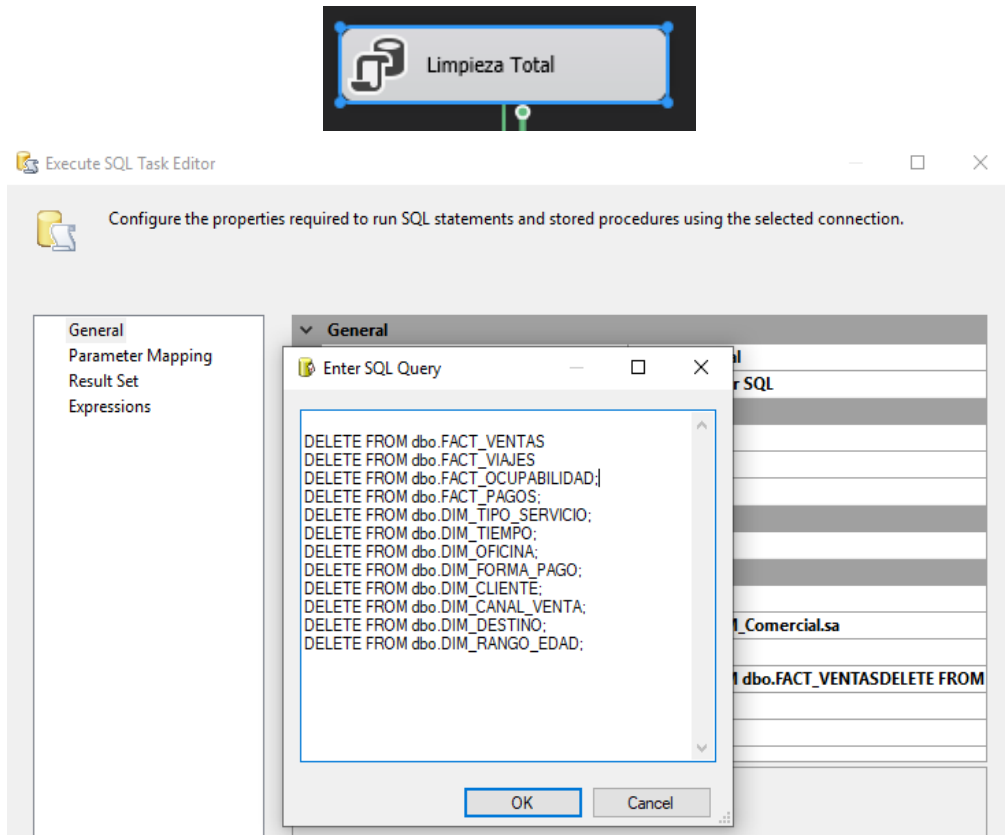
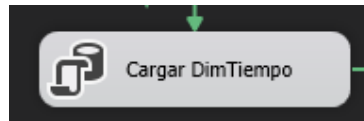


Fig. N° 35. Limpieza Total

(Elaboración propia)

Cargar Dimensión Tiempo

Este flujo de datos se encarga de ejecutar un script para generar datos desde el año 2015 al 2020.



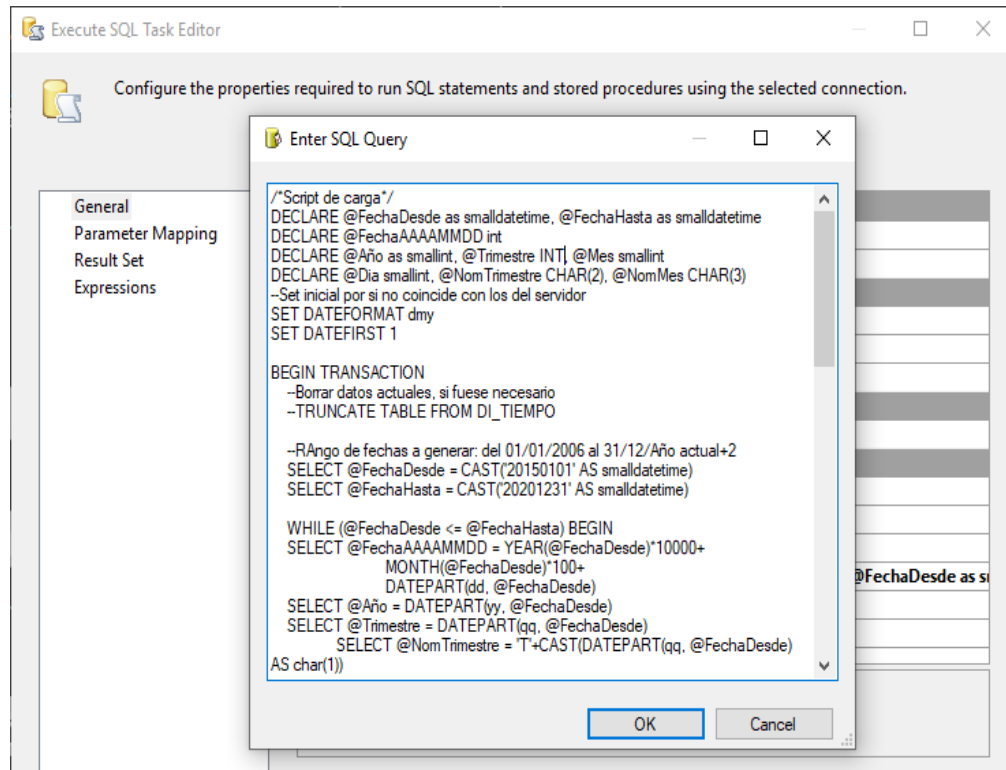
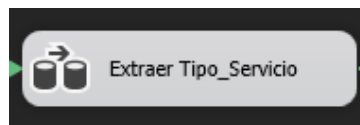
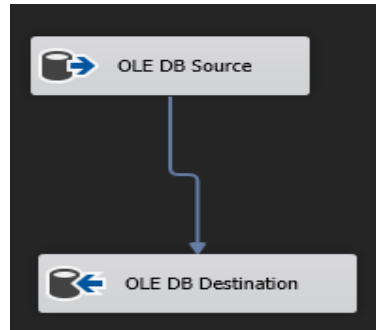


Fig. N° 36. Carga Dimensión Tiempo
 (Elaboración propia)

Extraer Tipo de Servicio

Este flujo de datos se encarga de leer los tipos de servicio de un origen de datos y cargarlo en la dimensión respectiva.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

Connection Manager
Columns
Error Output

OLE DB connection manager:
LocalHost.BDITTSABUSDEV.sa

Data access mode:
SQL command

SQL command text:

```
SELECT ID_CLASE_SERVICIO,NO_CORTO FROM  
ADTM_CLASE_SERVICIO  
WHERE ID_CATEGORIA_SERVICIO=1 and ES_CLASE_SERVICIO='A'
```

Destino:

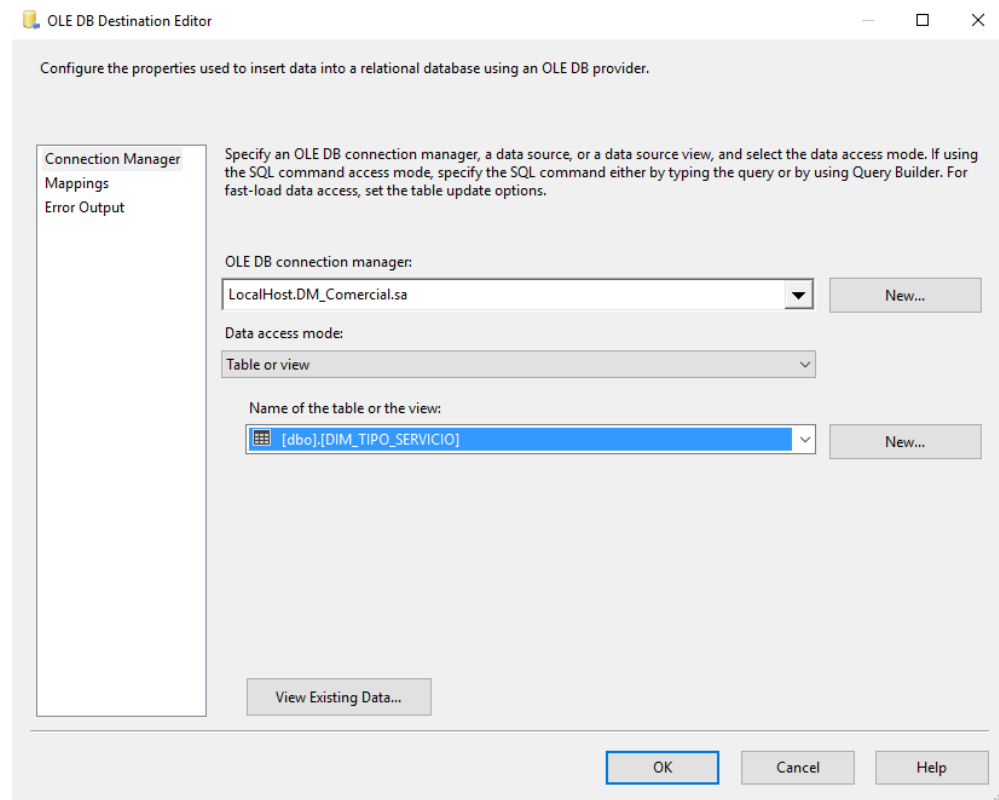
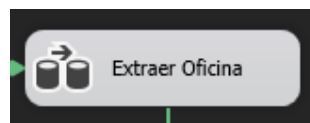


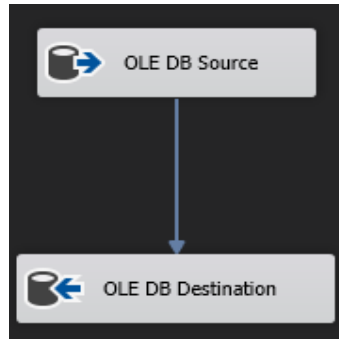
Fig. N° 37. Extraer Tipo de Servicio

(Elaboración propia)

Extraer Oficina

Este flujo de datos se encarga de leer las oficinas de una fuente origen y cargarla en la dimensión respectiva.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
 Columns
 Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

SELECT  ID_OFICINA, ID_AGENCIA, CO_TIPO_OFICINA, NO_OFICINA,
DE_DIRECCION, DE_TELEFONO1, DE_TELEFONO2, DE_CORREO, ES_OFICINA,
IN_TERMINAL, FE_CREA_REGISTRO, ID_CREA_REGISTRO, FE_MOD_REGISTRO,
        ID_MOD_REGISTRO,
CO_AUTORL_IMPR_MANIFIESTO_PASAJERO_SUNAT,
FE_IMPR_MANIFIESTO_PASAJERO_SUNAT,
CO_SERIE_IMPR_MANIFIESTO_PASAJERO,
NU_CORRELATIVO_IMPR_MANIFIESTO_PASAJERO,
        ID_OFICINA_EMBARCO_ENCOMIENDA, ID_OFICINA_ROYAL,
IN_TERMINAL_CARGO, in_jima, IN_HACE_CARGO,
IN_EMBARQUE_ENCOMIENDA, ID_CIUDADES
FROM    ADTM_OFICINA
WHERE   (ES_OFICINA = 'A') AND (IN_TERMINAL_CARGO = 0)
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

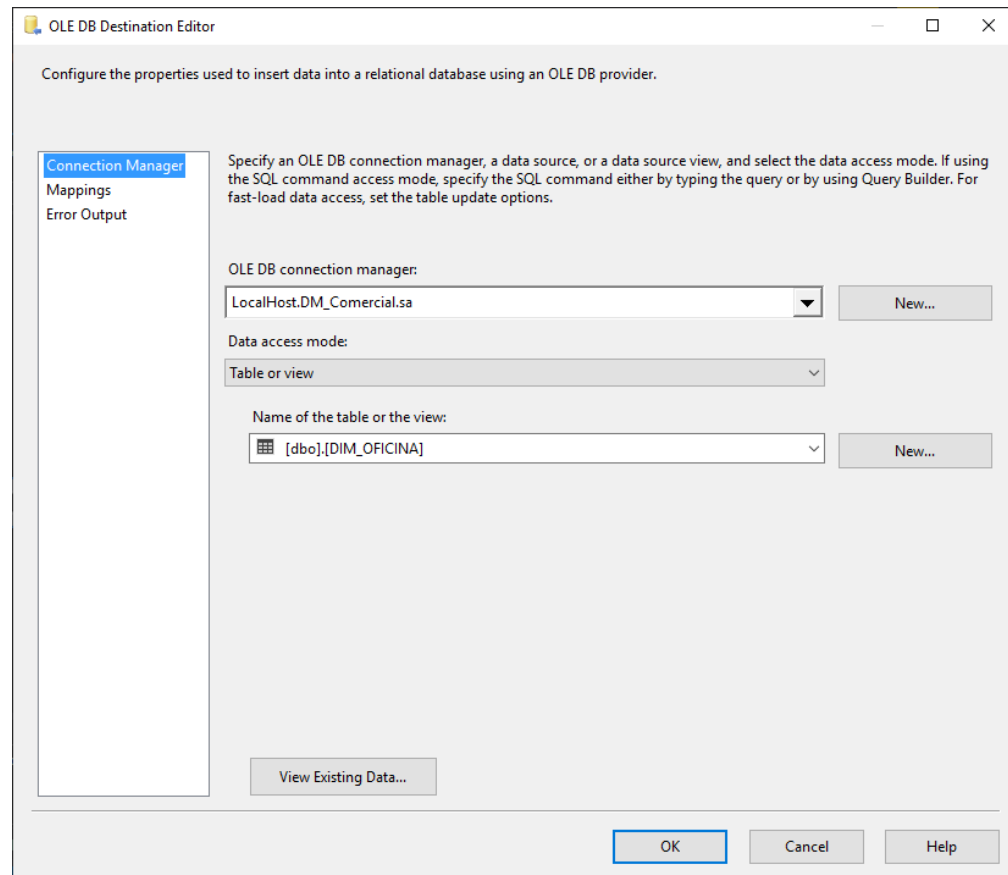
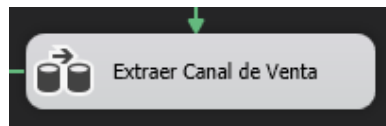


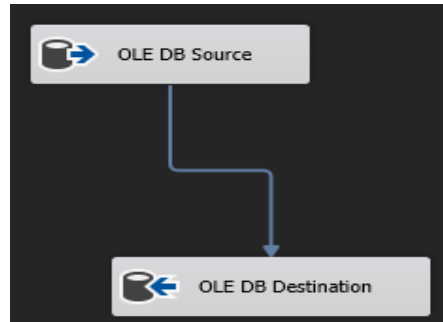
Fig. N° 38. Extraer Oficina

(Elaboración propia)

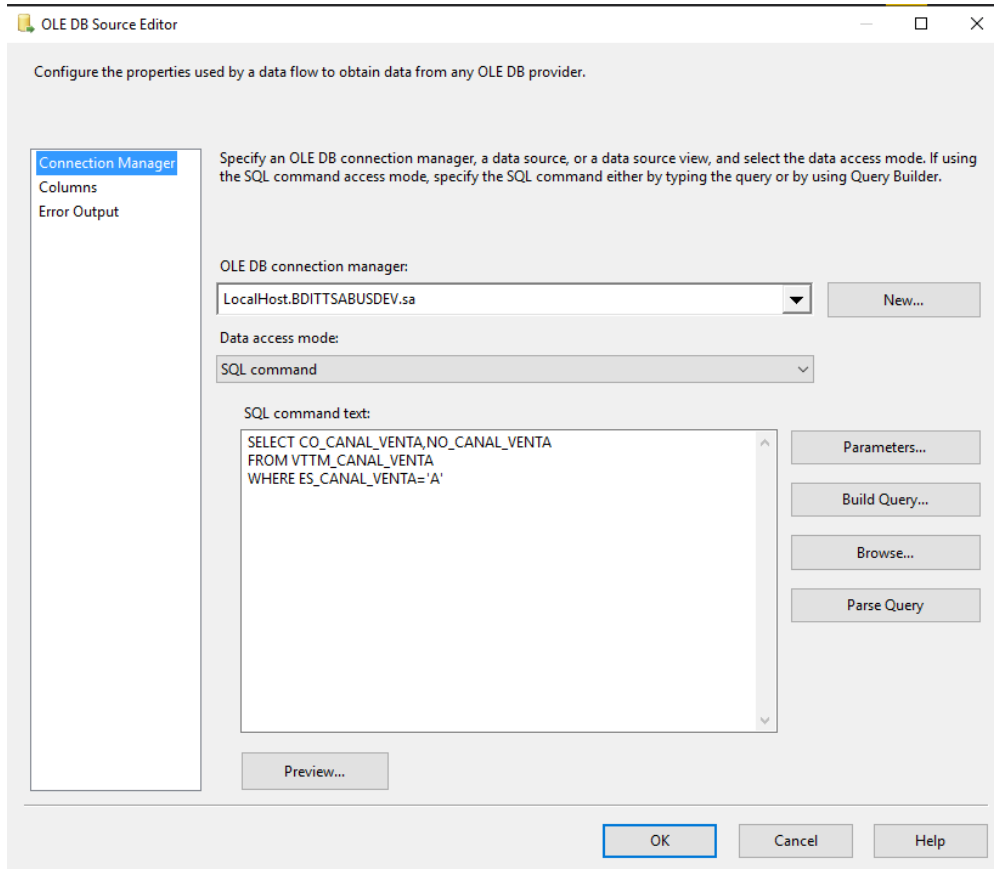
Extraer Canal de Venta

Este flujo de datos se encarga de leer los canales de venta de un origen de datos y cargarlo en la dimensión respectiva.





Origen:



OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

Connection Manager
Columns
Error Output

OLE DB connection manager:
LocalHost.BDITTSABUSDEV.sa

Data access mode:
SQL command

SQL command text:
SELECT CO_CANAL_VENTA,NO_CANAL_VENTA
FROM VTTM_CANAL_VENTA
WHERE ES_CANAL_VENTA='A'

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

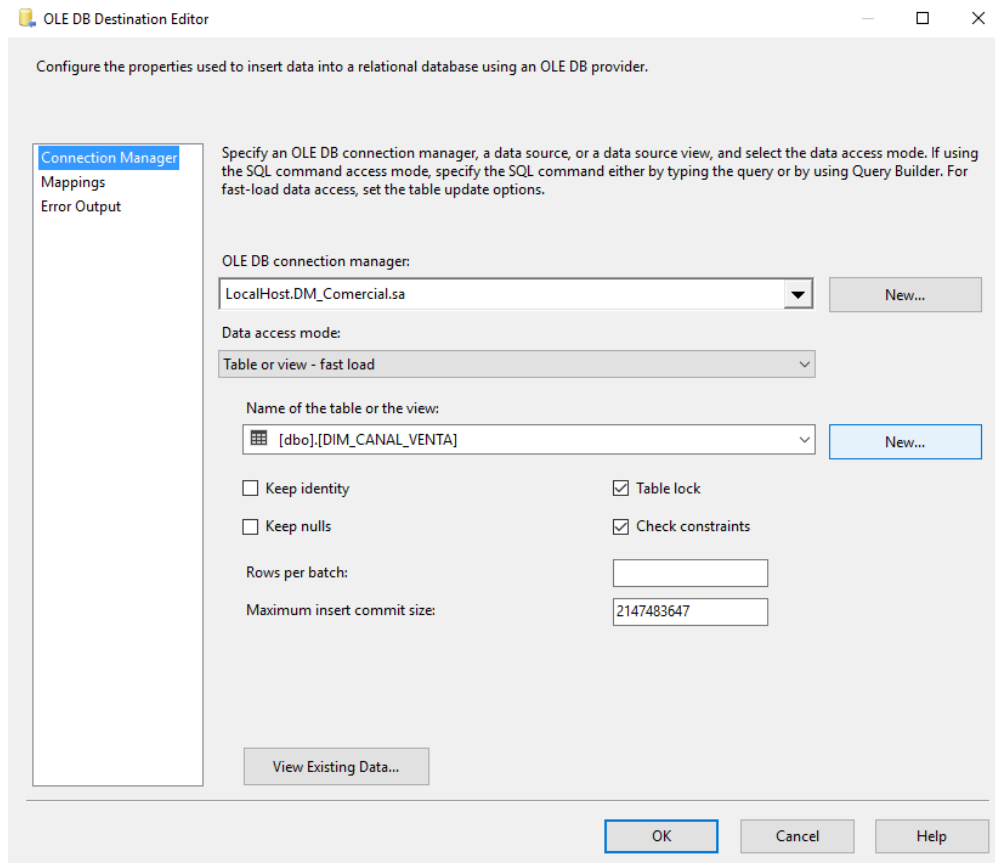
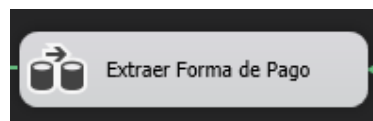


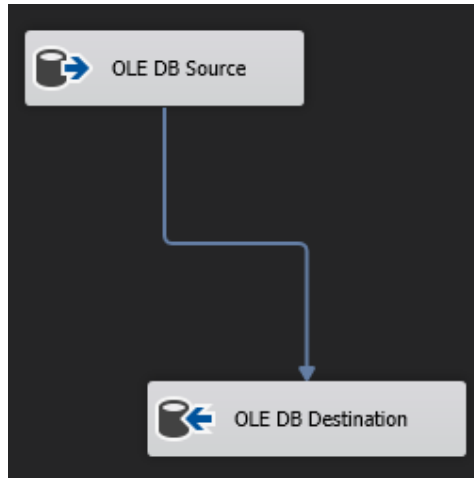
Fig. N° 39. Extraer Canal de Venta

(Elaboración propia)

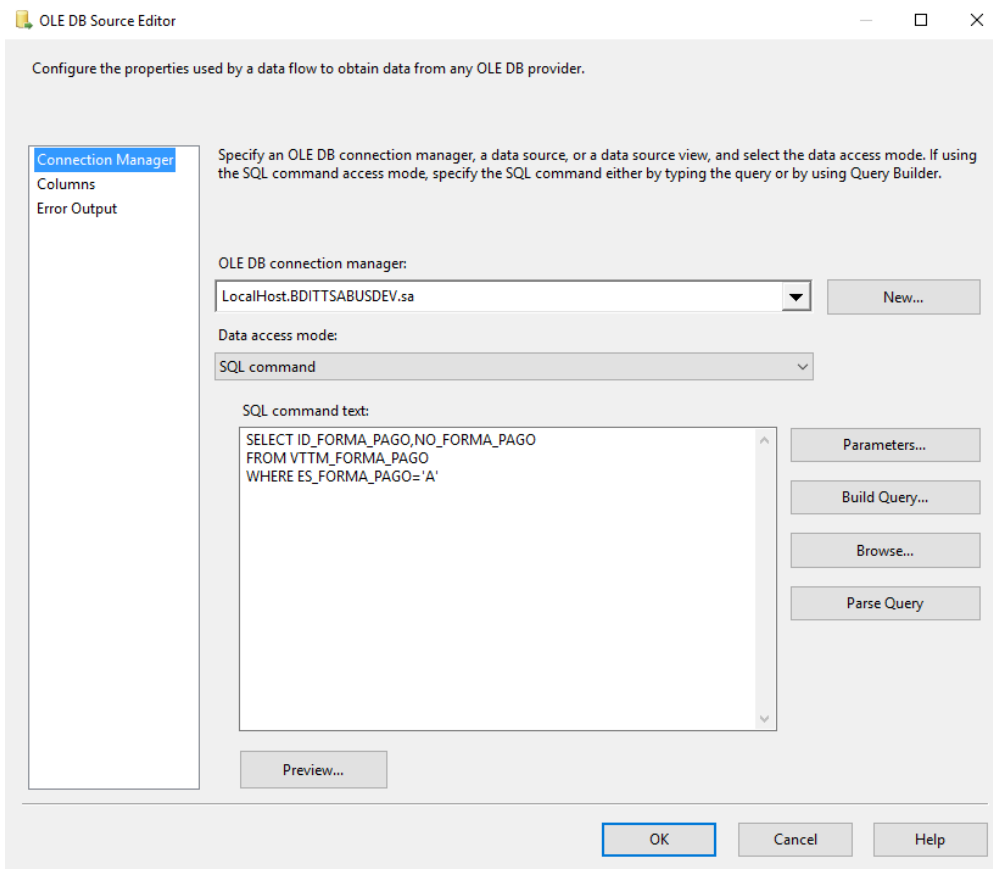
Extraer Forma de Pago

Este flujo de datos se encarga de leer las formas de pago desde un origen de datos y cargarlo en la dimensión respectiva.





Origen:



Destino:

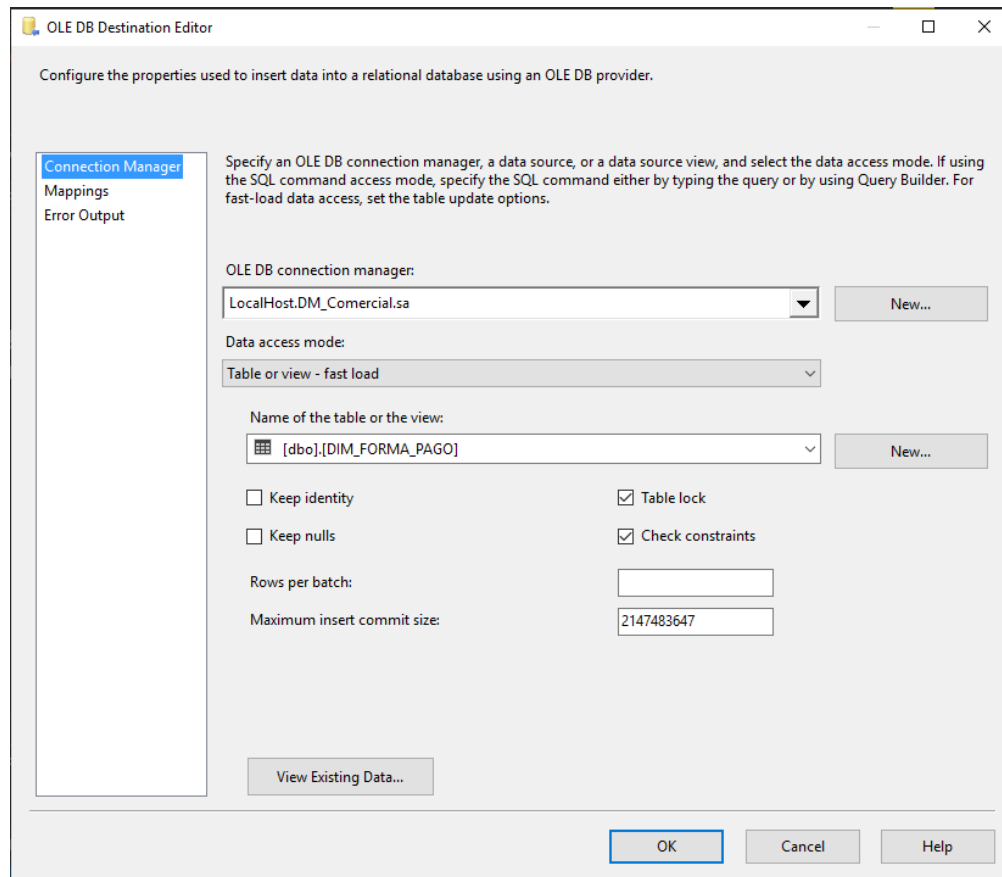
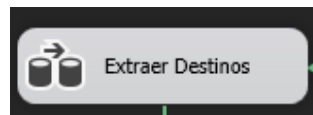


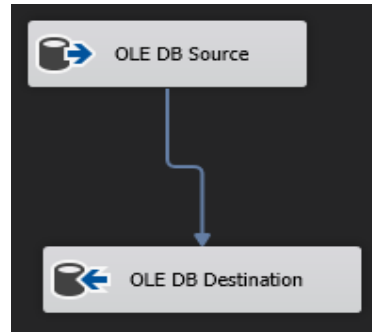
Fig. N° 40. Extraer Forma de Pago

(Elaboración propia)

Extraer Destino

Este flujo de datos se encarga de leer los destinos (origen-destino) de un origen de datos y los carga en la dimensión respectiva.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

Connection Manager
Columns
Error Output

OLE DB connection manager:
LocalHost.BDITTSABUSDEV.sa

Data access mode:
SQL command

SQL command text:

```
SELECT B.ID_DESTINO, (D.NO_OFICINA+'-'+E.NO_OFICINA) AS Ruta  
FROM PRM_DESTINO B  
INNER JOIN ADTM_OFICINA D ON D.ID_OFICINA=B.ID_OFICINA_ORIGEN  
INNER JOIN ADTM_OFICINA E ON E.ID_OFICINA=B.ID_OFICINA_DESTINO  
WHERE CO_TIPO_DESTINO='PAS' AND ES_DESTINO='A'
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

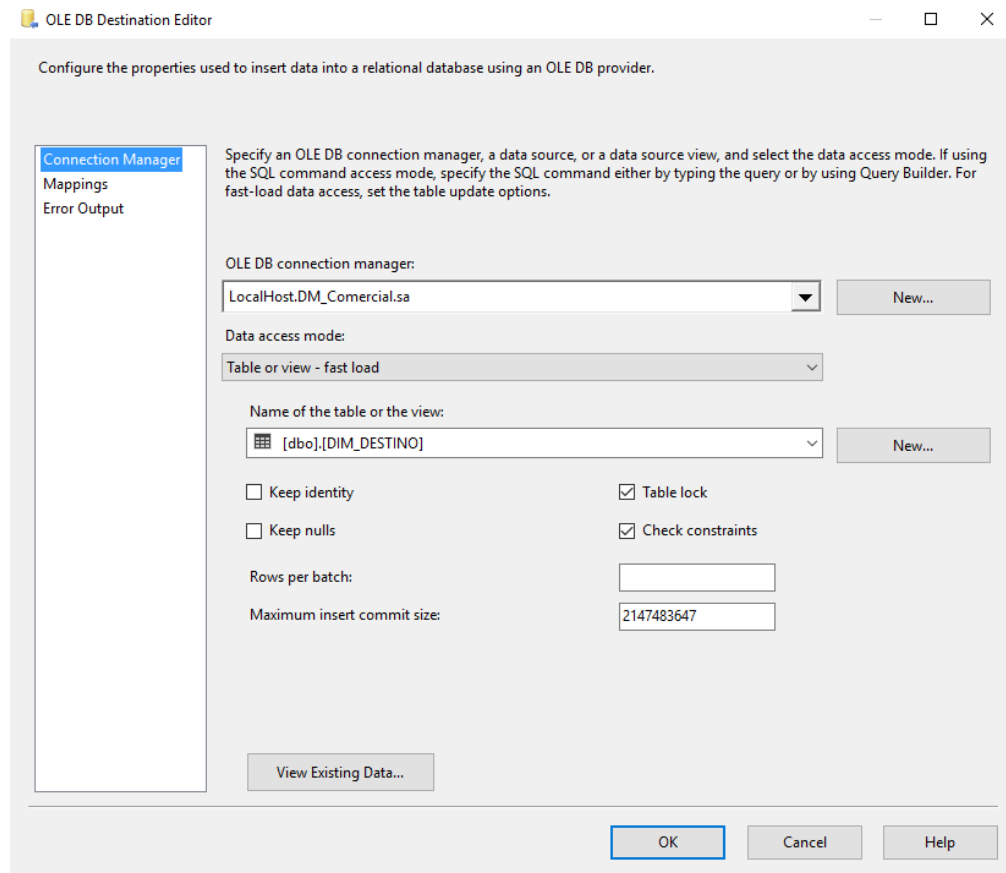
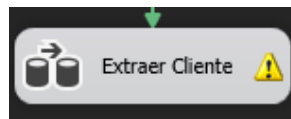


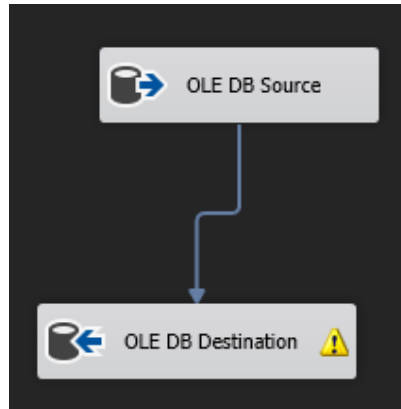
Fig. N° 41. Extraer Destino

(Elaboración propia)

Extraer Cliente

Este flujo de datos se encarga de leer todos los clientes de un origen de datos y los carga en la dimensión respectiva.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
 Columns
 Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

SELECT C.ID_PERSONA,CAST((C.NO_PERSONA + ' ' + C.AP_PERSONA +
'+ C.AM_PERSONA) AS varchar(250)) AS Nom_Cliente,COUNT(A.ID_BOLETO)
AS NroViaje
FROM VTTC_BOLETO A
INNER JOIN VTTM_PASAJERO B ON B.ID_PASAJERO=A.ID_PASAJERO
INNER JOIN ADTM_PERSONA C ON C.ID_PERSONA=B.ID_PERSONA
WHERE A.CO_ESTADO_BOLETO='EMB'
GROUP BY C.ID_PERSONA,C.NO_PERSONA,C.AP_PERSONA,C.AM_PERSONA
HAVING COUNT(A.ID_BOLETO) >= 80
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

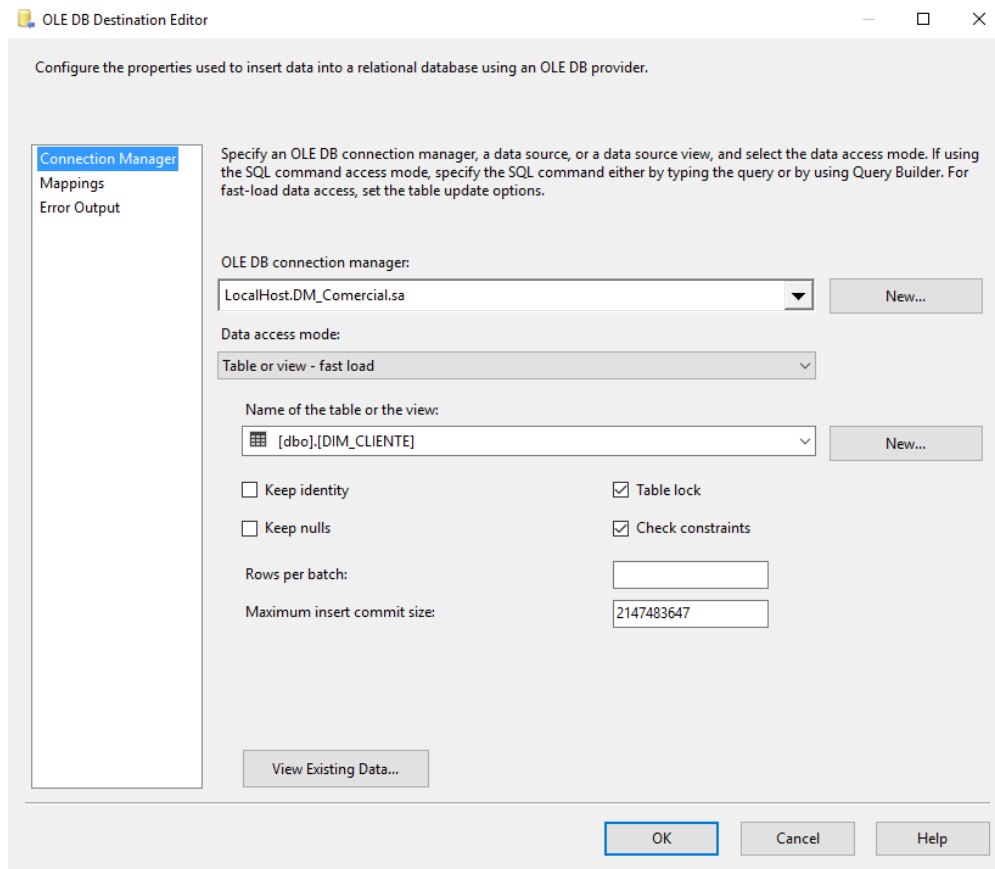
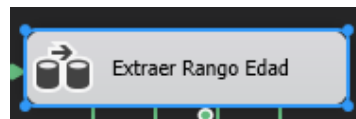


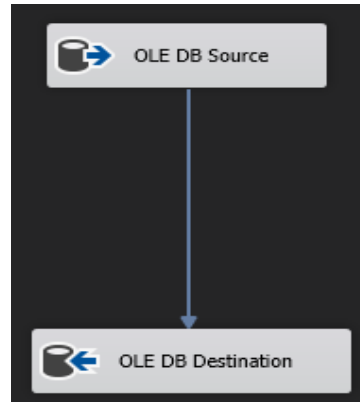
Fig. N° 42. Extraer Cliente

(Elaboración propia)

Extraer Rango de Edad

Este flujo de datos se encarga de leer las edades de los clientes, los transforma en un rango y luego los carga en la dimensión respectiva.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
 Columns
 Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

WITH AgeData as
(
  SELECT C.NU_DOCUMENTO
    ,DATEDIFF(YEAR, C.FE_NACIMIENTO, GETDATE()) AS [AGE]
    FROM VTTM_BOLETO A
      INNER JOIN VTTM_PASAJERO B ON
B.ID_PASAJERO=A.ID_PASAJERO
      INNER JOIN ADTM_PERSONA C ON C.ID_PERSONA=B.ID_PERSONA
      WHERE A.CO_ESTADO_BOLETO='EMB'
      GROUP BY C.NU_DOCUMENTO,C.FE_NACIMIENTO
),
GroupAge AS
(
  SELECT
    NU_DOCUMENTO
    ,CASE
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

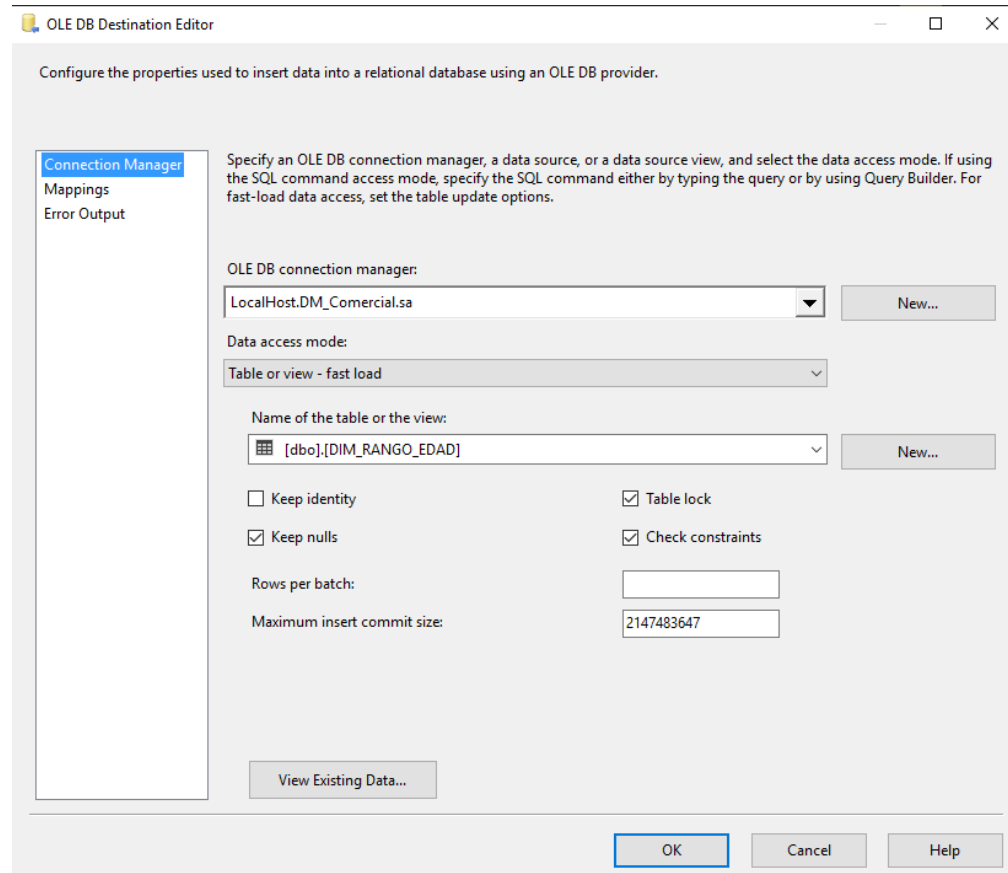
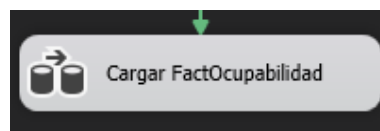


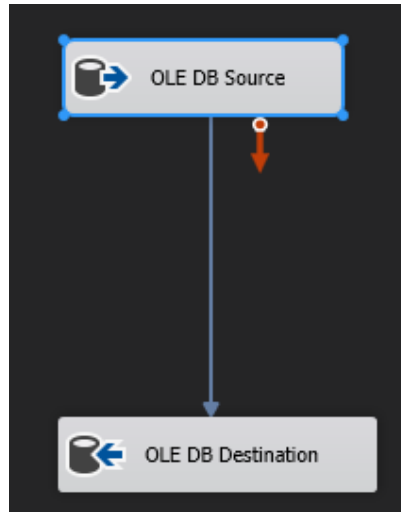
Fig. N° 43. Extraer Rango de Edad

(Elaboración propia)

Cargar Fact Ocupabilidad

Este flujo de datos se encarga de realizar la carga de la tabla de hechos Ocupabilidad.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

Connection Manager
 Columns
 Error Output

OLE DB connection manager:
 LocalHost.BDITTSABUDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

SELECT
Z.Id_Destino,Z.Id_Tiempo,Z.Id_Tipo_Servicio,Z.Vendidos,Z.CapacidadTotal
FROM
(
SELECT X.Id_Destino,X.Id_Tiempo,X.Id_Tipo_Servicio, SUM(CASE WHEN
X.CO_ESTADO_ASIENTO='VEN' THEN 1 ELSE 0 END) AS Vendidos,COUNT
(X.ID_PROGRAMACION) AS CapacidadTotal
FROM
(
SELECT E.Id_Destino, D.Id_Tiempo, H.Id_Tipo_Servicio,
C.CO_ESTADO_ASIENTO, C.ID_PROGRAMACION
FROM PRM_PROGRAMACION A
INNER JOIN PRM_DESTINO B ON B.ID_DESTINO=A.ID_DESTINO
INNER JOIN VTTR_ASIENTO_VENTA C ON
C.ID_PROGRAMACION=A.ID_PROGRAMACION
INNER JOIN VTTM_TRACKING_ASIENTO_VENTA F ON
  
```

Parameters...
 Build Query...
 Browse...
 Parse Query

Preview...

OK Cancel Help

Destino:

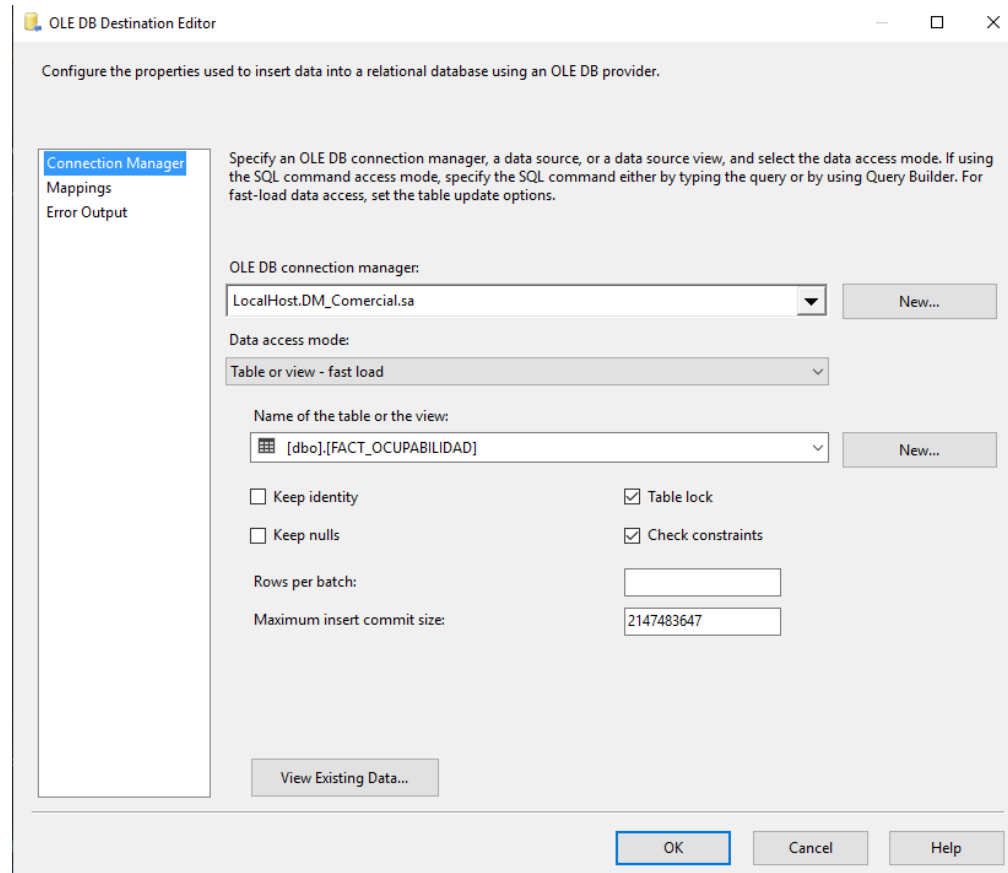
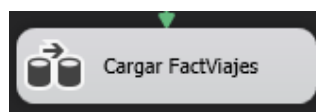


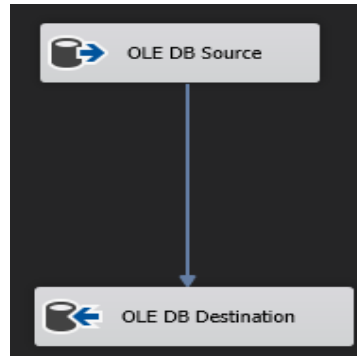
Fig. N° 44. Cargar Fact Ocupabilidad

(Elaboración propia)

Cargar Fact Viajes

Este flujo de datos se encarga de realizar la carga de la tabla de hechos Viajes.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
 Columns
 Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

SELECT D.Id_Cliente,H.Id_Tiempo,COUNT(A.ID_BOLETO) AS
NroViaje,I.Id_Rango_Edad
FROM VTTC_BOLETO A
INNER JOIN VTTM_PASAJERO B ON B.ID_PASAJERO=A.ID_PASAJERO
INNER JOIN ADTM_PERSONA C ON C.ID_PERSONA=B.ID_PERSONA
INNER JOIN DM_Comercial.dbo.DIM_CLIENTE D ON
D.Id_Cliente=C.ID_PERSONA
INNER JOIN DM_Comercial.dbo.DIM_TIEMPO H ON H.Id_Tiempo=CAST
(CONVERT(CHAR(10),A.FE_VIAJE,112) AS int)
INNER JOIN DM_Comercial.dbo.DIM_RANGO_EDAD I ON
I.Id_Rango_Edad=C.NU_DOCUMENTO
WHERE A.CO_ESTADO_BOLETO='EMB'
GROUP BY D.Id_Cliente,H.Id_Tiempo,I.Id_Rango_Edad
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

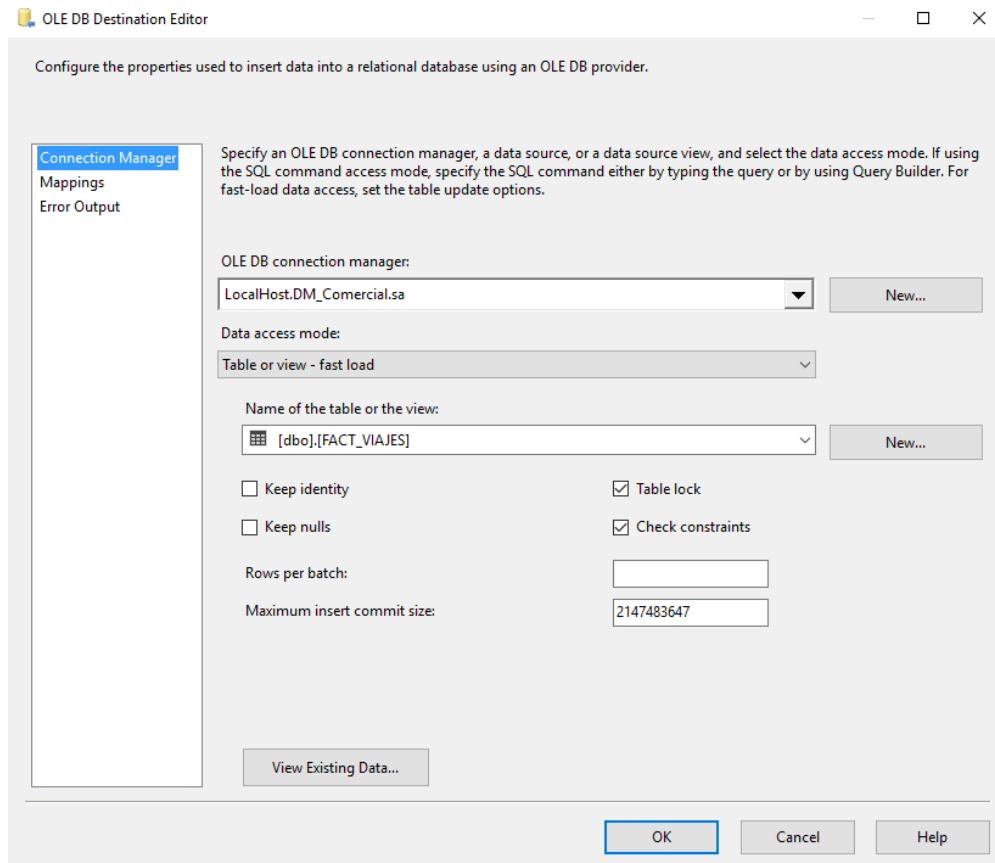
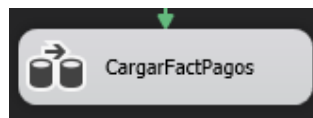


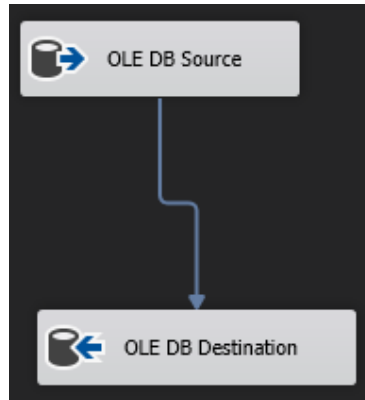
Fig. N° 45. Cargar Fact Viajes

(Elaboración propia)

Cargar Fact Pagos

Este flujo de datos se encarga de realizar la carga de la tabla de hechos Forma de Pago.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Connection Manager
 Columns
 Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

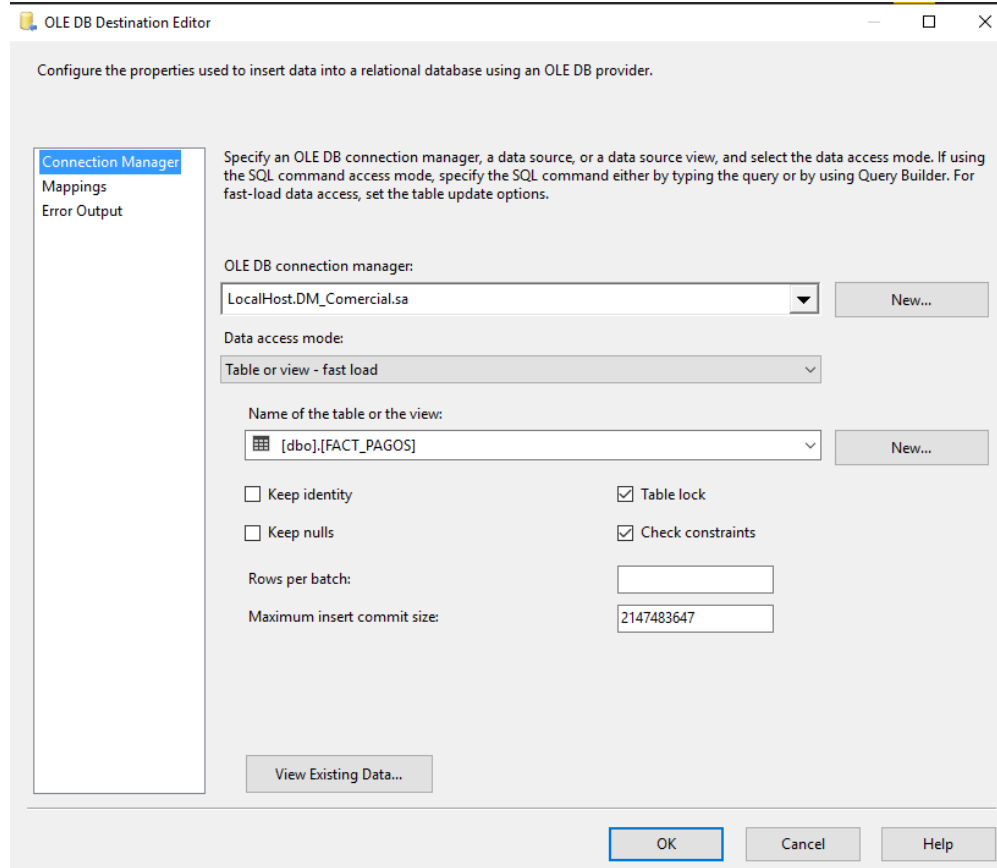
SELECT C.Id_Tiempo,D.Id_Forma_Pago,SUM(A.VA_MONTO_TOTAL) AS
MontoVendido,ISNULL(SUM(A.NU_BOLETOS),0) AS CantidadVendida
FROM VTTC_VENTA A
INNER JOIN VTTC_FORMA_PAGO B ON
B.ID_FORMA_PAGO=A.ID_FORMA_PAGO
INNER JOIN DM_Comercial.dbo.DIM_TIEMPO C ON C.Id_Tiempo=CAST
(CONVERT(CHAR(10),A.FE_CREA_REGISTRO,112) AS INT)
INNER JOIN DM_Comercial.dbo.DIM_FORMA_PAGO D ON
D.Id_Forma_Pago=A.ID_FORMA_PAGO
WHERE B.ES_FORMA_PAGO='A'
GROUP BY C.Id_Tiempo,D.Id_Forma_Pago
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:



OLE DB Destination Editor

Configure the properties used to insert data into a relational database using an OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

Connection Manager
Mappings
Error Output

OLE DB connection manager:
LocalHost.DM_Comercial.sa

Data access mode:
Table or view - fast load

Name of the table or the view:
[dbo].[FACT_PAGOS]

Keep identity Table lock
 Keep nulls Check constraints

Rows per batch:

Maximum insert commit size:

View Existing Data...

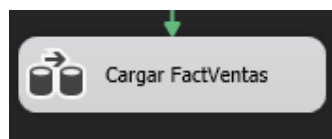
OK Cancel Help

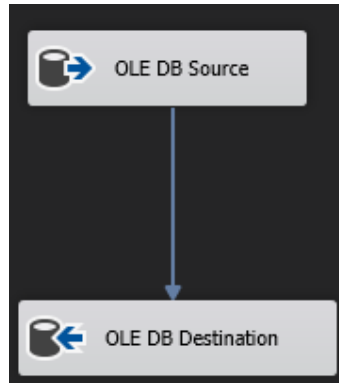
Fig. N° 46. Cargar Fact Pagos

(Elaboración propia)

Cargar Fact Ventas

Este flujo de datos se encarga de realizar la carga de la tabla de hechos Ventas.





Origen:

OLE DB Source Editor

Configure the properties used by a data flow to obtain data from any OLE DB provider.

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder.

Connection Manager
 Columns
 Error Output

OLE DB connection manager:
 LocalHost.BDITTSABUSDEV.sa New...

Data access mode:
 SQL command

SQL command text:

```

SELECT E.Id_Oficina,H.Id_Tiempo,SUM(A.PRECIO_ORIGINAL) AS
Monto_Vendido, COUNT(A.ID_BOLETO) AS Cantidad_Vendida,I.Id_Canal_Venta
FROM VTTC_BOLETO A
INNER JOIN VTTC_COMPROBANTE_PAGO B ON B.ID_DETALLE=A.ID_DETALLE
INNER JOIN VTTC_COMPROBANTE_PAGO C ON
C.ID_COMPROBANTE_PAGO=B.ID_COMPROBANTE_PAGO
INNER JOIN VTTM_VENTA D ON D.ID_VENTA=C.ID_VENTA
INNER JOIN ADTM_PUNTO_VENTA F ON
F.ID_PUNTO_VENTA=D.ID_PUNTO_VENTA
INNER JOIN ADTM_OFICINA G ON G.ID_OFICINA=F.ID_OFICINA
INNER JOIN DM_Comercial.dbo.DIM_OFICINA E ON
E.Id_Oficina=G.ID_OFICINA
INNER JOIN DM_Comercial.dbo.DIM_TIEMPO H ON H.Id_Tiempo=CAST
(CONVERT (CHAR(10),A.FE_CREA_REGISTRO,112) AS int)
INNER JOIN DM_Comercial.dbo.DIM_CANAL_VENTA I ON
I.Id_Canal_Venta=D.CO_CANAL_VENTA
  
```

Parameters...
Build Query...
Browse...
Parse Query

Preview...

OK Cancel Help

Destino:

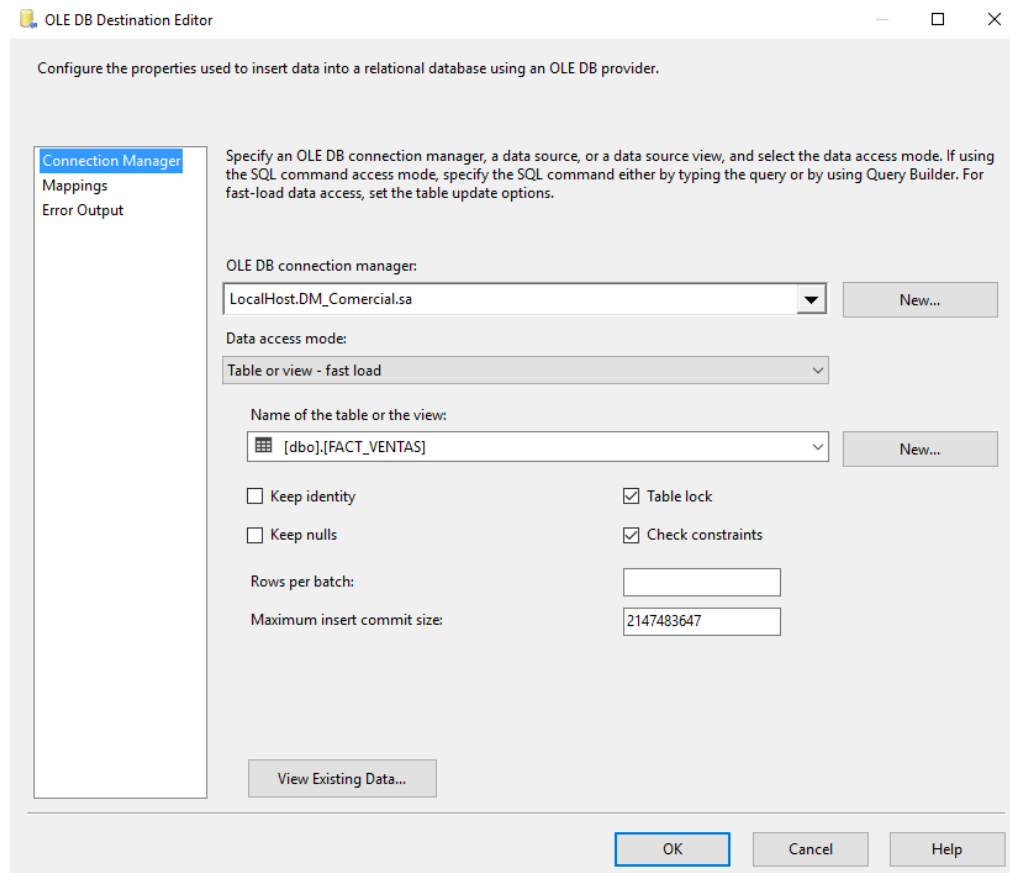


Fig. N° 47. Cargar Fact Ventas

(Elaboración propia)

4.1.8. Implementación del Data Mart

La elección de las herramientas de presentación y explotación de la información se consideró aspectos básicos como la usabilidad, diseño y conocimiento previo. Cabe resaltar que la implementación del data mart se llevó a cabo en la infraestructura actual de ITTSA.

Luego de ejecutar los procesos ETL, toda esa información es almacenada en una base de datos SQL Server 2014 en el mismo servidor del Core Punto de Ventas de ITTSA con un sistema operativo Windows Server 2012 de 64 bits.

La Figura N° 48 muestra la consola de gestión de SQL Server y se puede apreciar la base de datos del Data Mart.

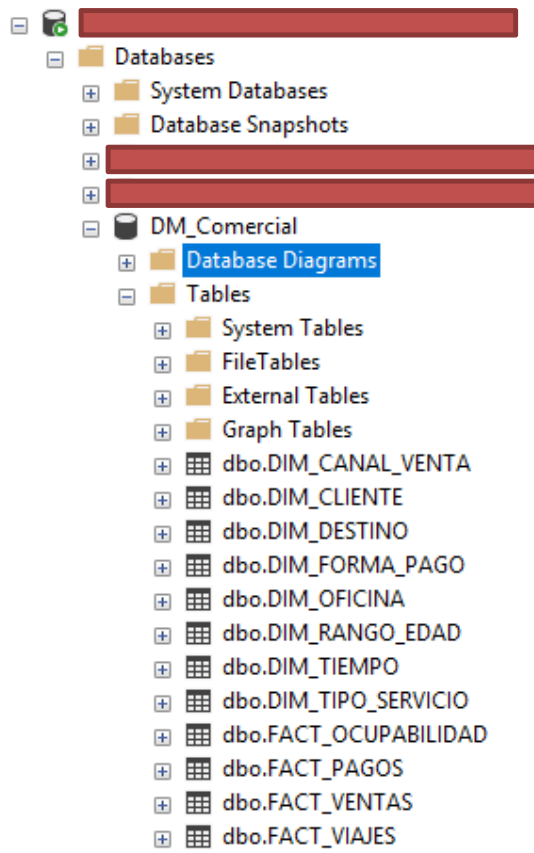


Fig. N° 48. Base de Datos Data Mart

(Elaboración propia)

La explotación de la información se hizo uso de la herramienta QlikView mediante dashboards que muestran la información necesaria para el área comercial, según los requerimientos solicitados.

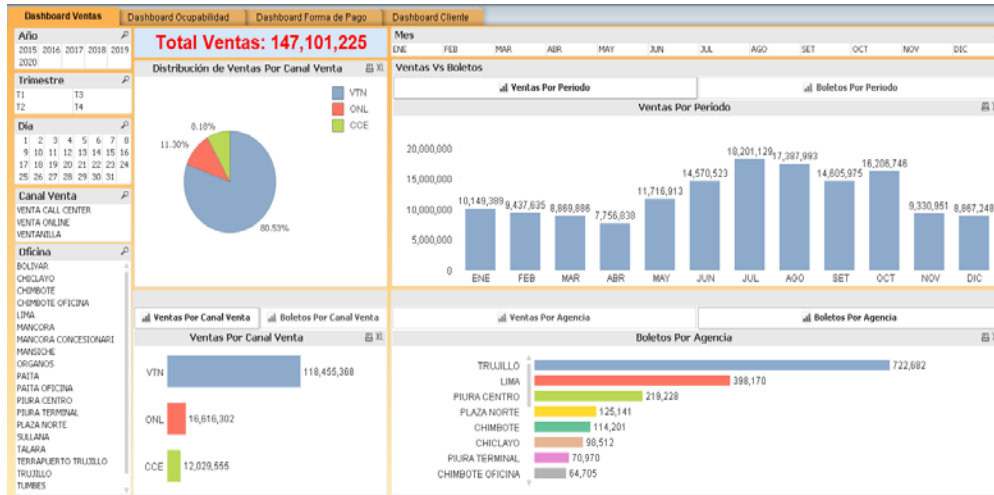


Fig. N° 49. Dashboard Ventas
(Elaboración propia)

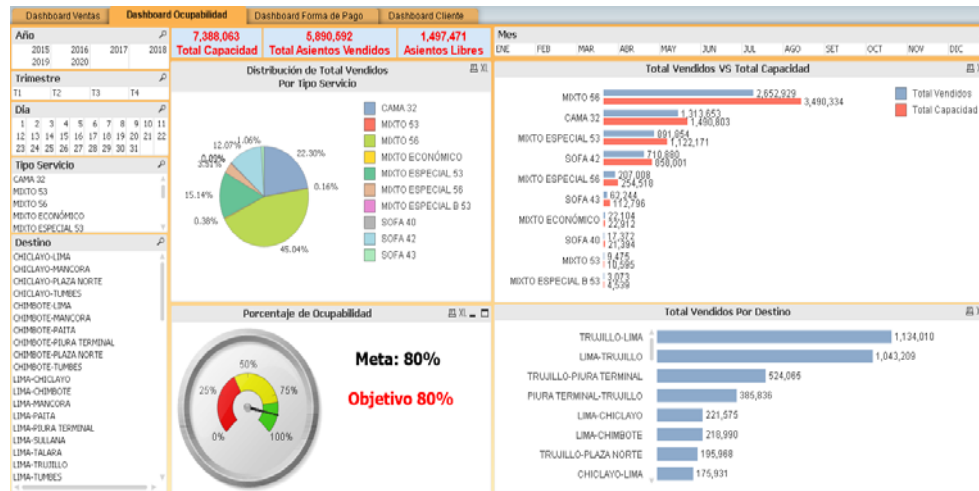


Fig. N° 50. Dashboard Ocupabilidad
(Elaboración propia)

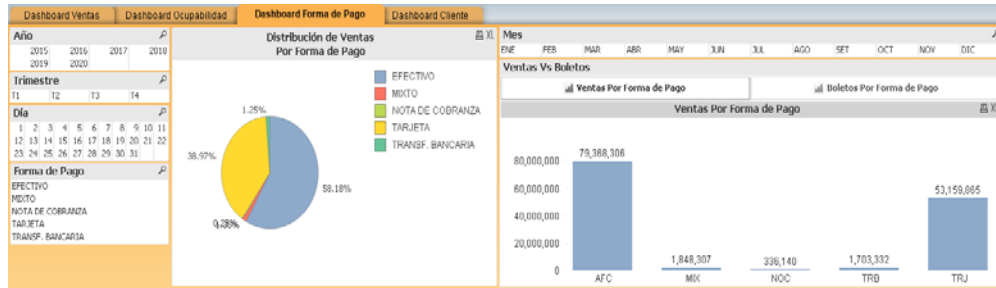


Fig. N° 51. Dashboard Forma de Pago

(Elaboración propia)

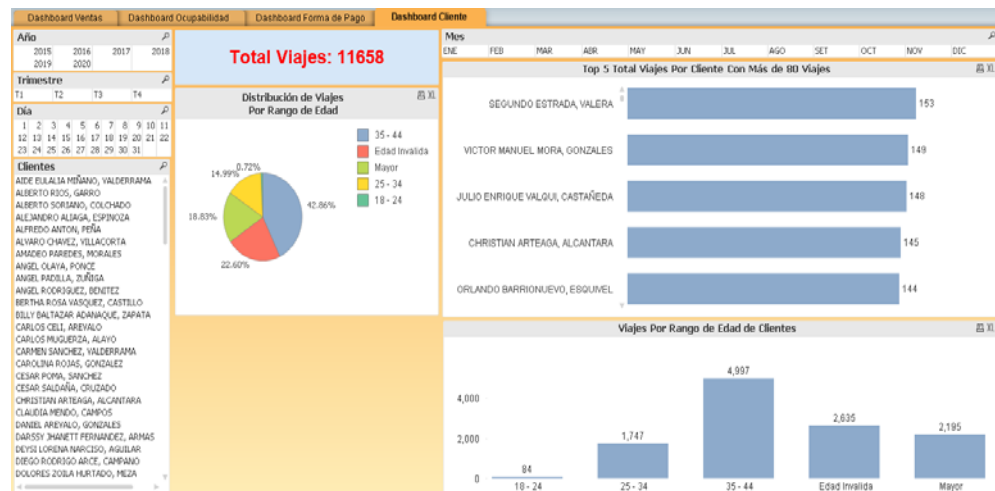


Fig. N° 52. Dashboard Cliente

(Elaboración propia)

CAPÍTULO 5. MATERIAL Y MÉTODOS

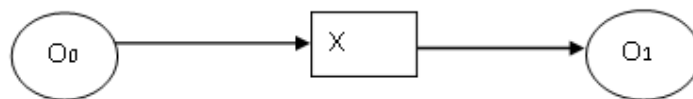
5.1. Tipo de diseño de investigación.

5.1.1. Tipo de Investigación

Descriptivo - Experimental.

5.1.2. Diseño de Investigación

Para comprobar la validez o refutación de la hipótesis, utilizaré el método de Diseño de Sucesión en Línea, llamada también Pre – Test, Post – Test, el que consiste.



En donde:

O₀: Proceso de toma de decisiones del área comercial **ANTES** de implementar el data mart.

X: Data Mart

O₁: Proceso de toma de decisiones del área comercial **DESPUES** de implementar el data mart.

5.2. Material.

5.2.1. Unidad de estudio.

La presente unidad de estudio es un ejecutivo del área comercial de la empresa ITTSABUS SRL.

5.2.2. Población.

La población es el área comercial de la empresa ITTSABUS SRL, la cual asciende a 4 ejecutivos.

5.2.3. Muestra.

Se tomará el 100.00% de la población de estudio del área comercial de la empresa ITTSABUS SRL.

5.3. Métodos.

5.3.1. Encuesta

Para poder recolectar y analizar datos se aplicará una encuesta al área comercial de la empresa ITTSABUS SRL.

5.3.2. Técnicas de Procesamiento y Análisis de Datos

En cuanto al procesamiento de datos éste consistirá en el cálculo, selección, clasificación y ordenación de tablas o cuadros, debidamente codificados y tabulados. La tabulación se realizará de forma electrónica y se someterá al tratamiento estadístico, para de esta manera determinar el significado de aquellas relaciones significativas y obtener como resultado la existencia de una coherencia entre Data Mart y Proceso de toma de decisiones. Para analizar los datos se utilizará el Software Excel o Spss versión 22, y se utilizará la prueba estadística t de student para corroborar la prueba de hipótesis planteada.

CAPÍTULO 6. RESULTADOS

6.1. Indicador Nivel de Satisfacción

Para el nivel de satisfacción se realizó una encuesta consistente en 15 preguntas en la escala de Likert (1 al 5). A continuación, se presentan los resultados del Pre-Test y Post-Test.

Tabla N° 17. Pre-Test Nivel de Satisfacción

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	1	2	2	1	1	3	1	2	1	3	1	1	3	2	2	26
2	2	1	2	1	2	3	1	3	1	3	1	1	3	3	2	29
3	1	2	2	2	3	3	1	2	1	3	2	1	3	3	1	30
4	2	2	1	1	2	4	1	4	1	4	1	1	3	3	1	31

Nota: Elaboración Propia

Tabla N° 18. Post-Test Nivel de Satisfacción

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	75
2	4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	74
3	5	4	5	5	5	5	5	5	5	5	5	5	4	5	5	73
4	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	75

Nota: Elaboración Propia

Luego se ha hecho una distribución de frecuencias en relación a los ejecutivos de acuerdo a cada pregunta y el puntaje que seleccionaron.

Tabla N° 19. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según siente que es productivo en cuanto a las actividades que realiza a diario.

Pregunta 01	Pre Test		Post Test	
Nunca	2	50.00	0	0.00
Casi nunca	2	50.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	1	25.00
Siempre	0	0.00	3	75.00

Total	4	100.00	4	100.00
--------------	----------	---------------	----------	---------------

Nota: Elaboración Propia

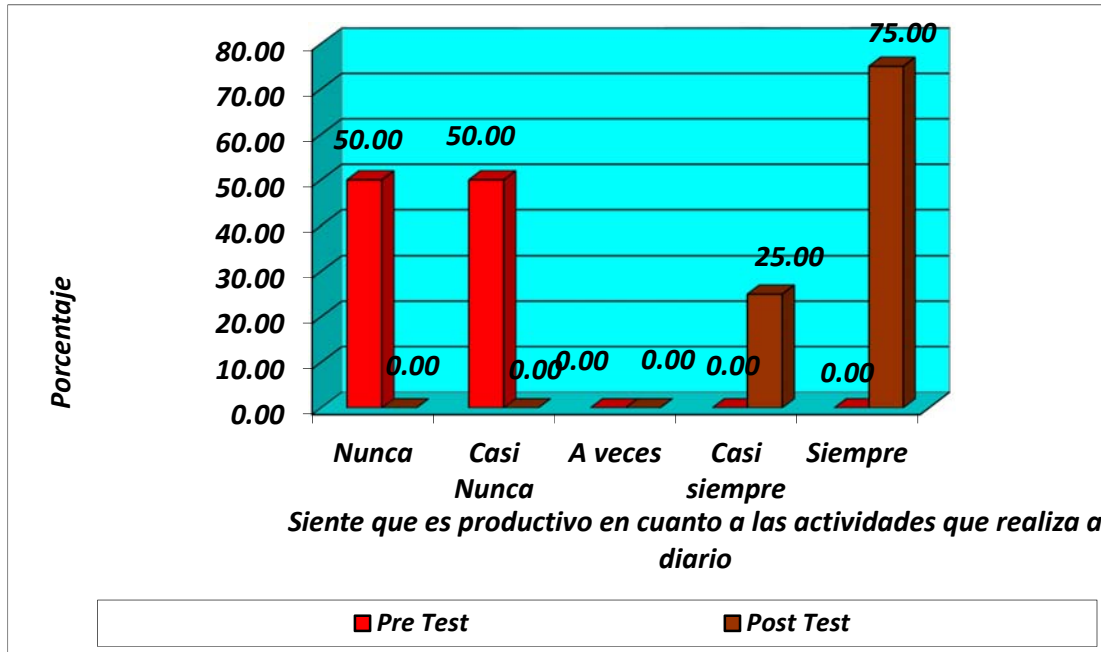


Fig. N° 53. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según siente que es productivo en cuanto a las actividades que realiza a diario

(Elaboración propia)

En la Tabla N° 19 y Figura N° 53 se observa que en la pregunta Siente que es productivo en cuanto a las actividades que realiza a diario en el Pre Test el 50.00% precisa nunca, el 50.00% casi nunca, mientras que en el Post Test el 25.00% precisa casi siempre y el 75.00% precisa Siempre.

Tabla N° 20. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según La información que les brindan los reportes del sistema actual le es suficiente

Pregunta 02	Pre Test		Post Test	
Nunca	1	25.00	0	0.00
Casi nunca	3	75.00	0	0.00
A veces	0	0.00	0	0.00

Casi siempre	0	0.00	1	25.00
Siempre	0	0.00	3	75.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

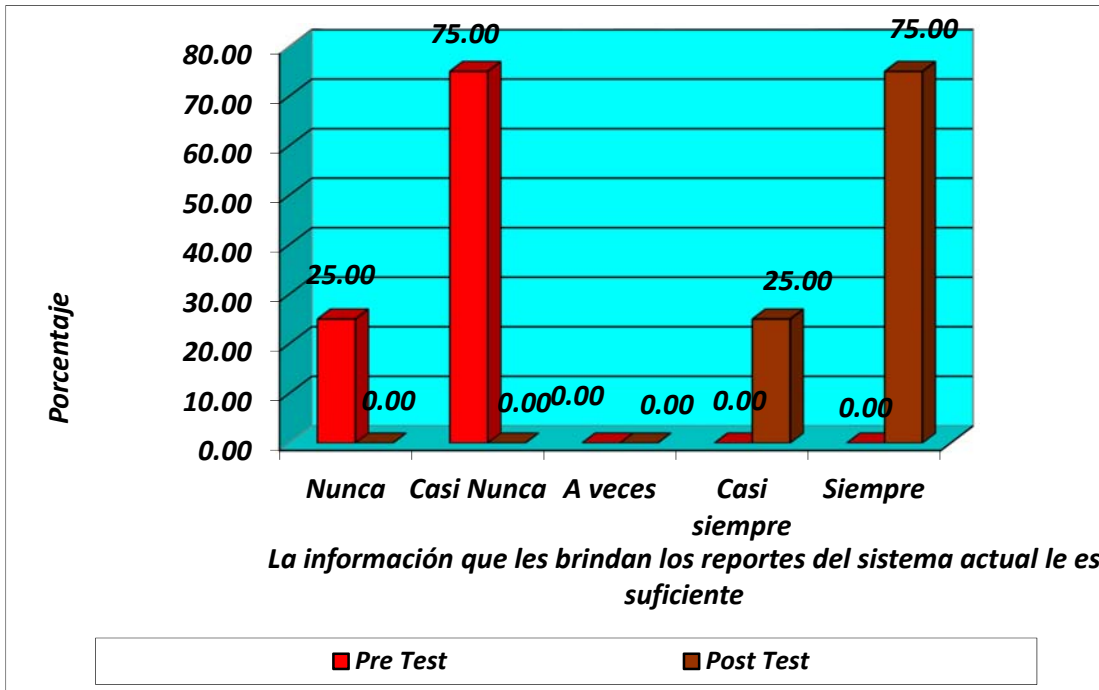


Fig. N° 54. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según la información que les brindan los reportes del sistema actual le es suficiente

(Elaboración propia)

En la Tabla N° 20 y Figura N° 54 se observa que en la pregunta la información que les brindan los reportes del sistema actual le es suficiente en el Pre Test el 25.00% precisa nunca, el 75.00% precisa casi nunca, mientras que en el Post Test el 25.00% precisa casi siempre y el 75.00% precisa siempre.

Tabla N° 21. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según tiene disponible la información necesaria para realizar su trabajo

Pregunta 03	Pre Test		Post Test	
Nunca	1	25.00	0	0.00

Casi nunca	3	75.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

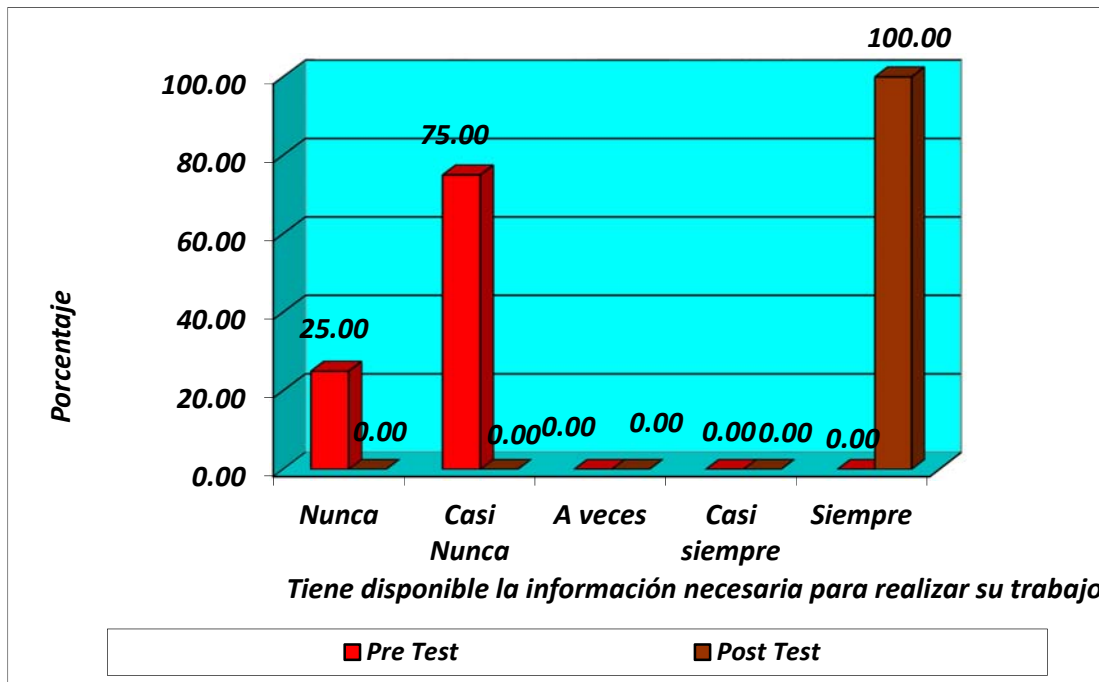


Fig. N° 55. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según tiene disponible la información necesaria para realizar su trabajo

(Elaboración propia)

En la Tabla N° 21 y Figura N° 55 se observa que en la pregunta tiene disponible la información necesaria para realizar su trabajo en el Pre Test el 25.00% precisa nunca, el 75.00% precisa casi nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 22. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Está a gusto con el tiempo que toma obtener la información comercial

Pregunta 04	Pre Test	Post Test
-------------	----------	-----------

Nunca	3	75.00	0	0.00
Casi nunca	1	25.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

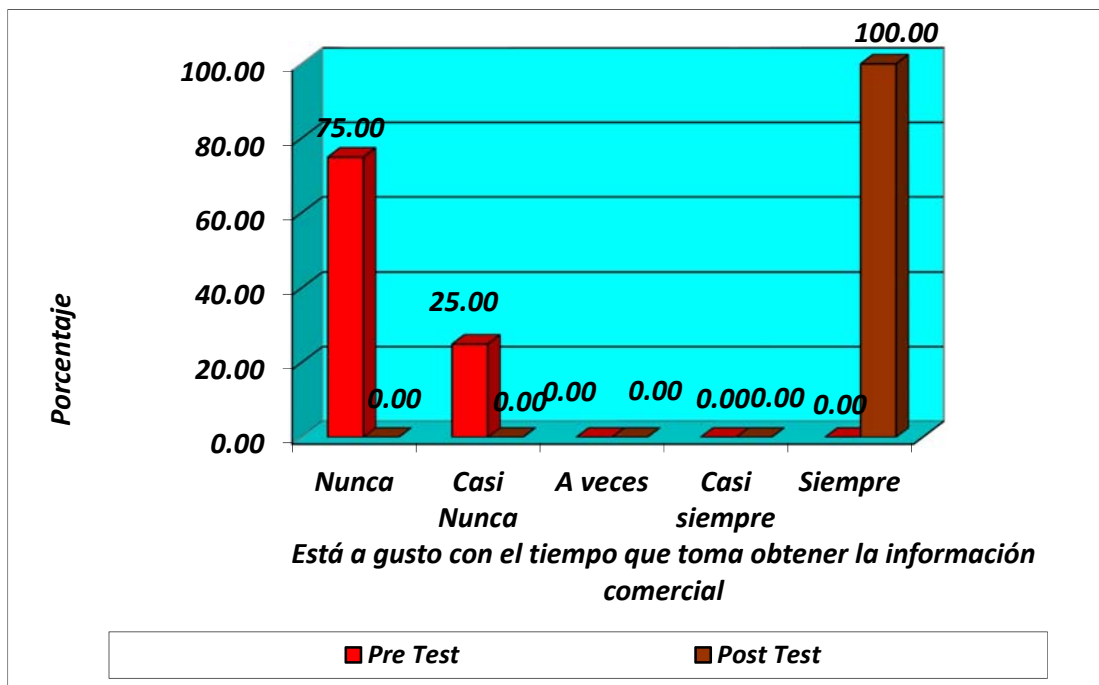


Fig. N° 56. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Está a gusto con el tiempo que toma obtener la información comercial

(Elaboración propia)

En la Tabla N° 22 y Figura N° 56 se observa que en la pregunta Está a gusto con el tiempo que toma obtener la información comercial en el Pre Test el 75.00% precisa nunca, el 25.00% precisa casi nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 23. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información de ocupabilidad de los buses

Pregunta 05	Pre Test		Post Test	
Nunca	1	25.00	0	0.00
Casi nunca	2	50.00	0	0.00
A veces	1	25.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

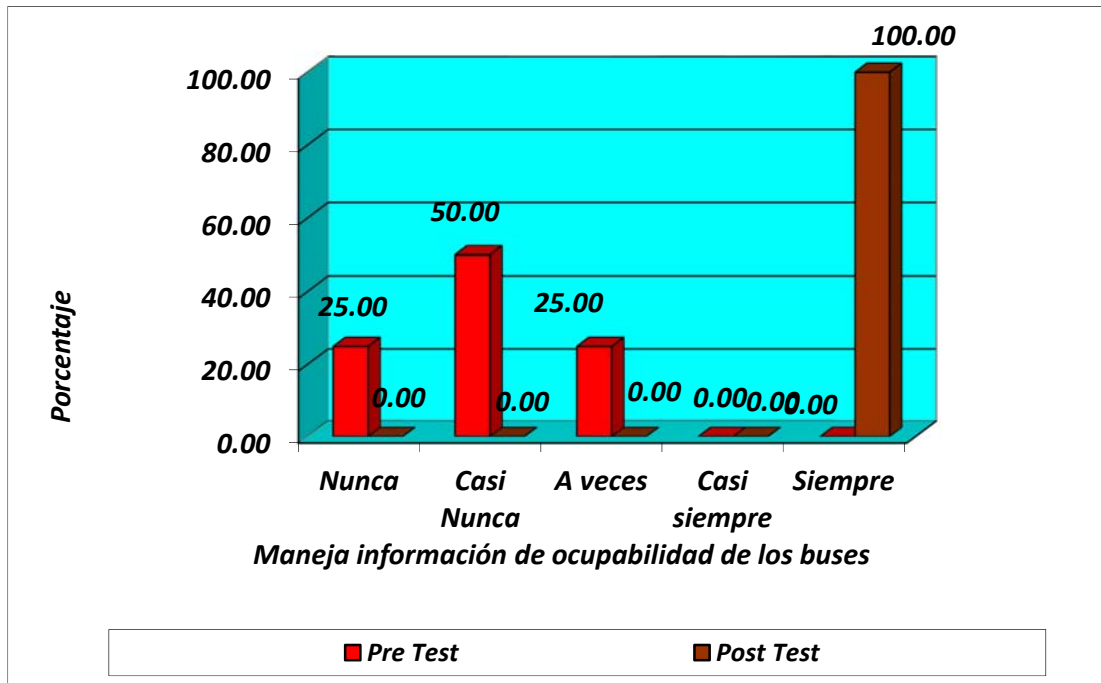


Fig. N° 57. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información de ocupabilidad de los buses

(Elaboración propia)

En la Tabla N° 23 y Figura N° 57 se observa que en la pregunta Maneja información de ocupabilidad de los buses en el Pre Test el 25.00% precisa nunca, el 50.00% precisa casi nunca, el 25.00% precisa a veces mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 24. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información detallada de los canales de venta.

Pregunta 06	Pre Test		Post Test	
	Count	Percentage	Count	Percentage
Nunca	0	0.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	3	75.00	0	0.00
Casi siempre	1	25.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

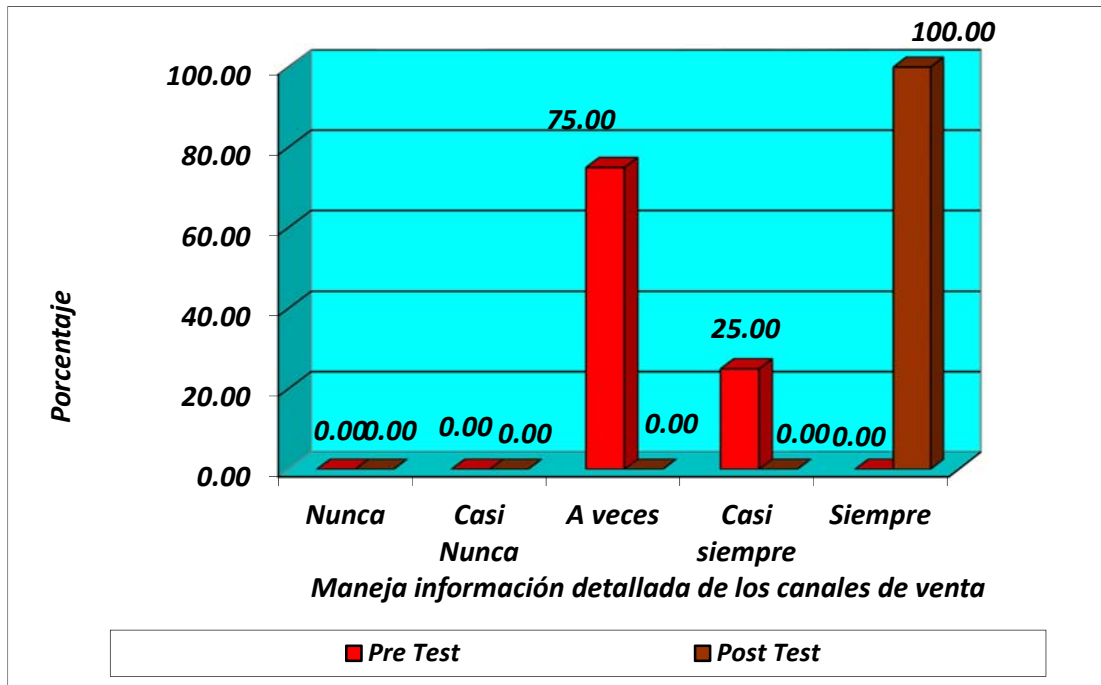


Fig. N° 58. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información detallada de los canales de venta

(Elaboración propia)

En la Tabla N° 24 y Figura N° 58 se observa que en la pregunta El sistema actual le brinda información detallada de los canales de venta en el Pre Test el 75.00% precisa a veces, el 25.00% precisa casi siempre mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 25. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Puede realizar su trabajo con rapidez.

Pregunta 07	Pre Test		Post Test	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Nunca	4	100.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

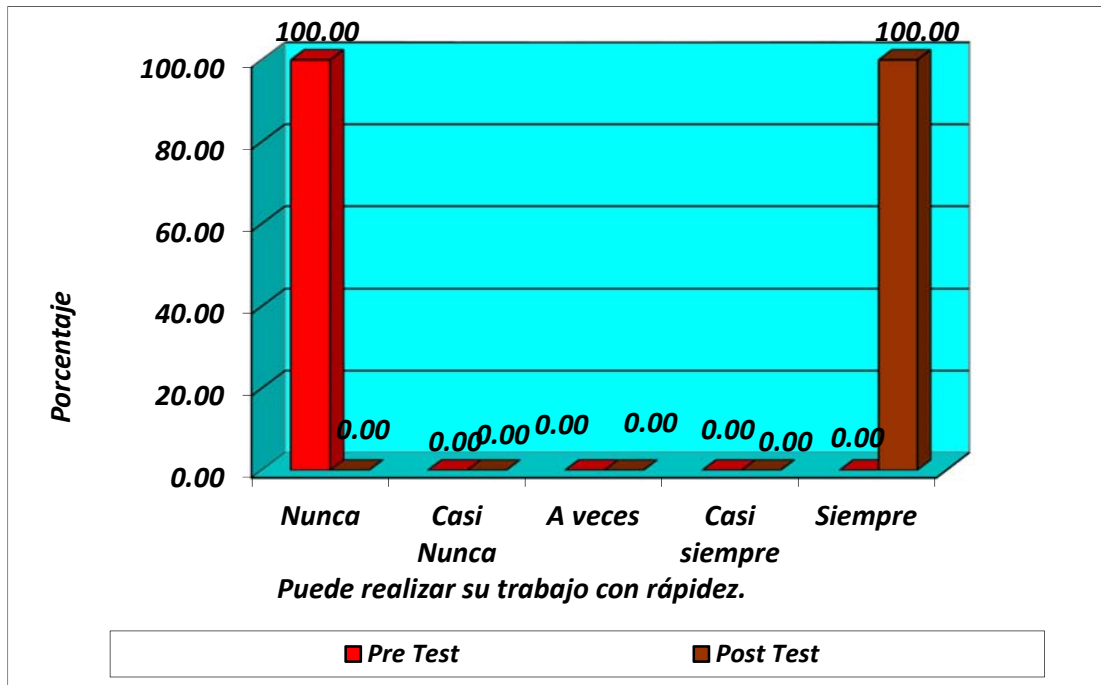


Fig. N° 59. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Puede realizar su trabajo con rapidez

(Elaboración propia)

En la Tabla N° 25 y Figura N° 59 se observa que en la pregunta Puede realizar su trabajo con rapidez en el Pre Test el 100.00% precisa nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 26. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Tiene disponibilidad de la información comercial las 24hrs del día

Pregunta 08	Pre Test		Post Test	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Nunca	0	0.00	0	0.00
Casi nunca	2	50.00	0	0.00
A veces	1	25.00	0	0.00
Casi siempre	1	25.00	2	50.00
Siempre	0	50.00	2	50.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

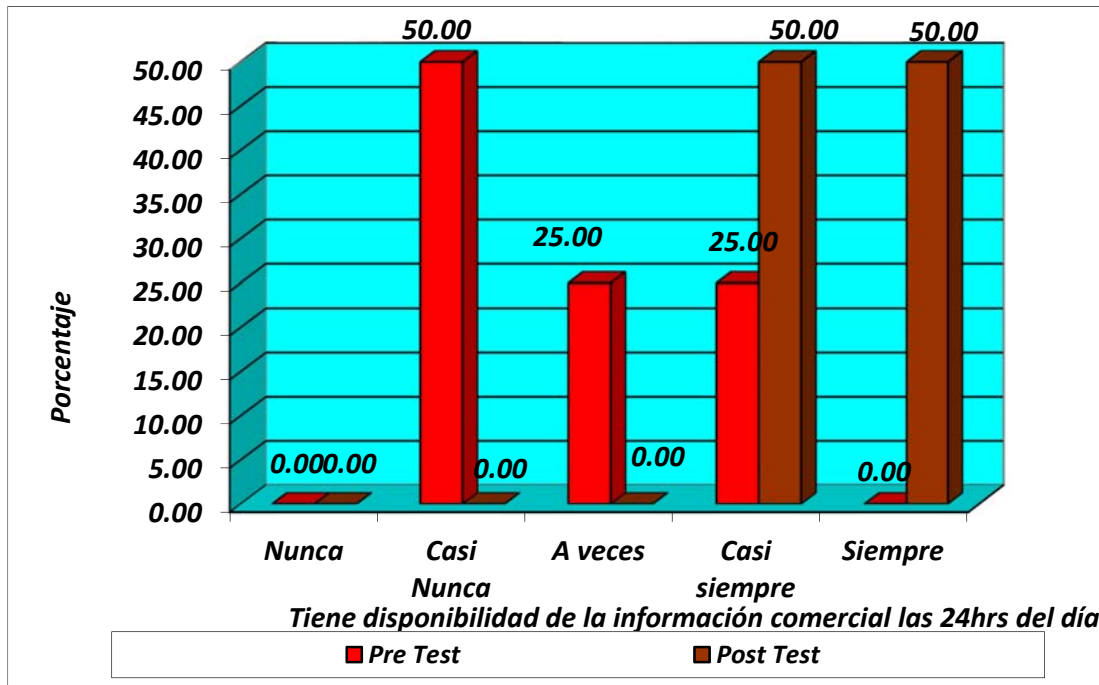


Fig. N° 60. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Tiene disponibilidad de la información comercial las 24hrs del día

(Elaboración propia)

En la Tabla N° 26 y Figura N° 60 se observa que en la pregunta Tiene disponibilidad de la información comercial las 24hrs del día en el Pre Test el 50.00% precisa casi nunca, el 25.00% a veces, el 25.00% casi siempre mientras que en el Post Test el 50.00% precisa casi siempre, el 50.00% siempre.

Tabla N° 27. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Cree usted que actualmente puede analizar la información de manera adecuada.

Pregunta 09	Pre Test		Post Test	
	Count	Percentage	Count	Percentage
Nunca	4	100.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

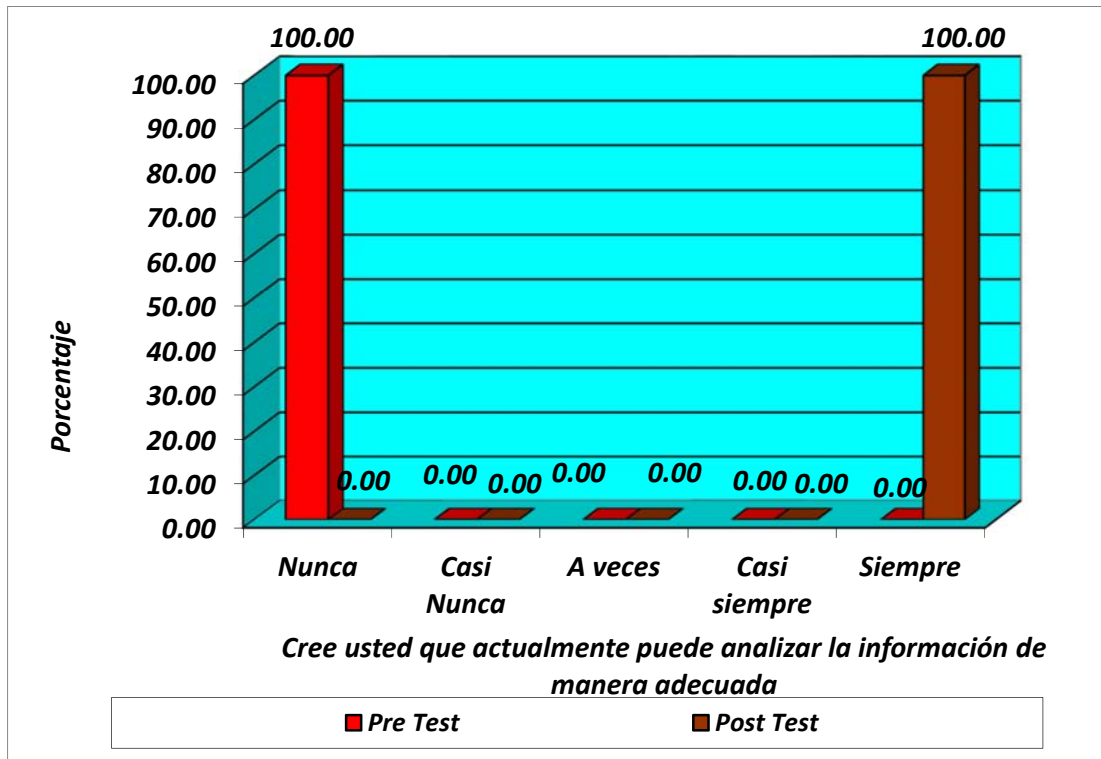


Fig. N° 61. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Cree usted que actualmente puede analizar la información de manera adecuada.

(Elaboración propia)

En la Tabla N° 27 y Figura N° 61 se observa que en la pregunta Cree usted que actualmente puede analizar la información de manera adecuada en el Pre Test el 100.00% precisa nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 28. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Cree tener información de ventas exacta

Pregunta 10	Pre Test		Post Test	
Nunca	0	0.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	3	75.00	0	0.00
Casi siempre	1	25.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

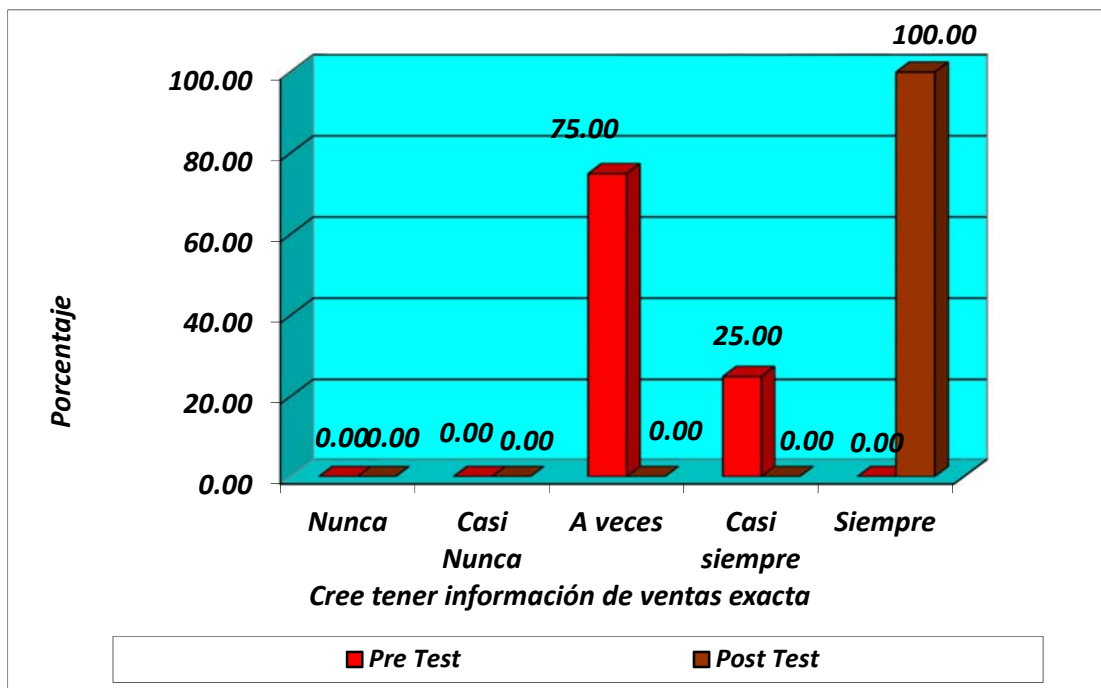


Fig. N° 62. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Cree tener información de ventas exacta

(Elaboración propia)

En la Tabla N° 28 y Figura N° 62 se observa que en la pregunta Cree tener información de ventas exacta en el Pre Test el 75.00% precisa a veces el 25.00% precisa casi siempre, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 29. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Está conforme con el tiempo que involucra la toma de decisiones

Pregunta 11	Pre Test		Post Test	
Nunca	3	75.00	0	0.00
Casi nunca	1	25.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

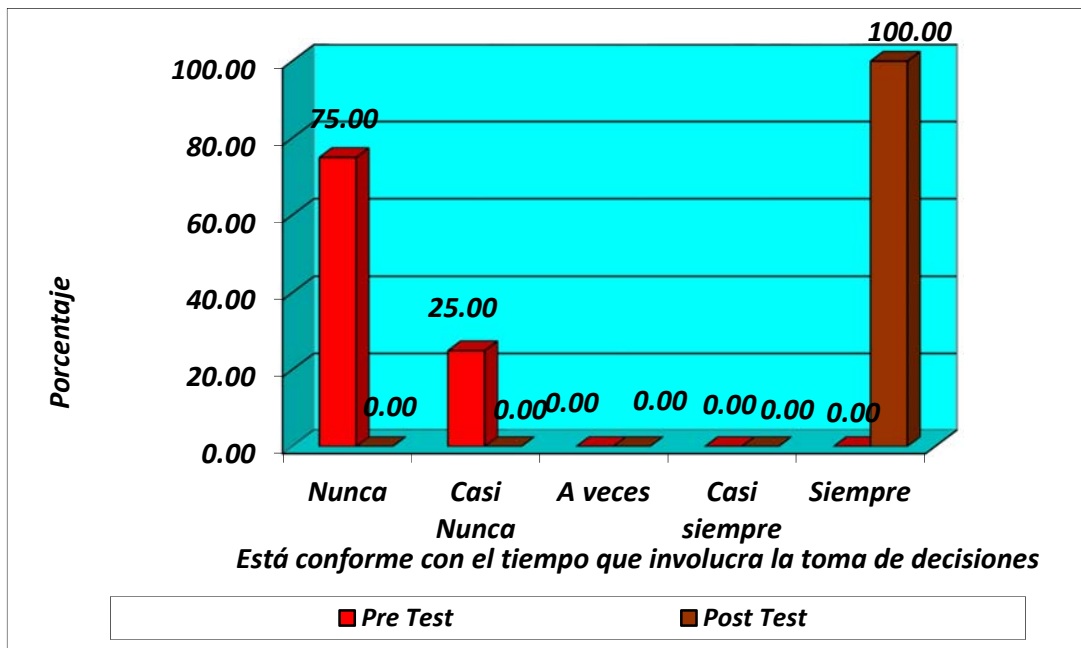


Fig. N° 63. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Está conforme con el tiempo que involucra la toma de decisiones

(Elaboración propia)

En la Tabla N° 29 y Figura N° 63 se observa que en la pregunta Está conforme con el tiempo que involucra la toma de decisiones en el Pre Test el 75.00% precisa nunca, el 25.00% precisa casi nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 30. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Maneja información necesaria de los clientes

Pregunta 12	Pre Test		Post Test	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Nunca	4	100.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

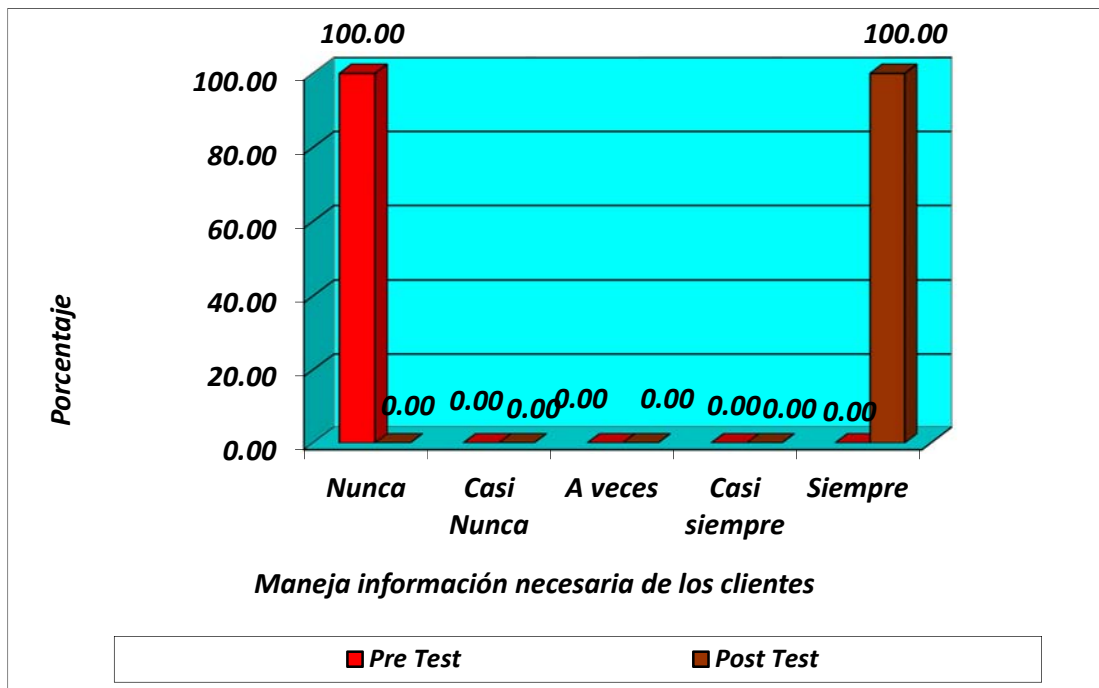


Fig. N° 64. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Maneja información necesaria de los clientes

(Elaboración propia)

En la Tabla N° 30 y Figura N° 64 se observa que en la pregunta Maneja información necesaria de los clientes en el Pre Test el 100.00% precisa nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 31. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según La información actual que manejan les apoya en predecir las ventas

Pregunta 13	Pre Test		Post Test	
Nunca	0	0.00	0	0.00
Casi nunca	0	0.00	0	0.00
A veces	4	100.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

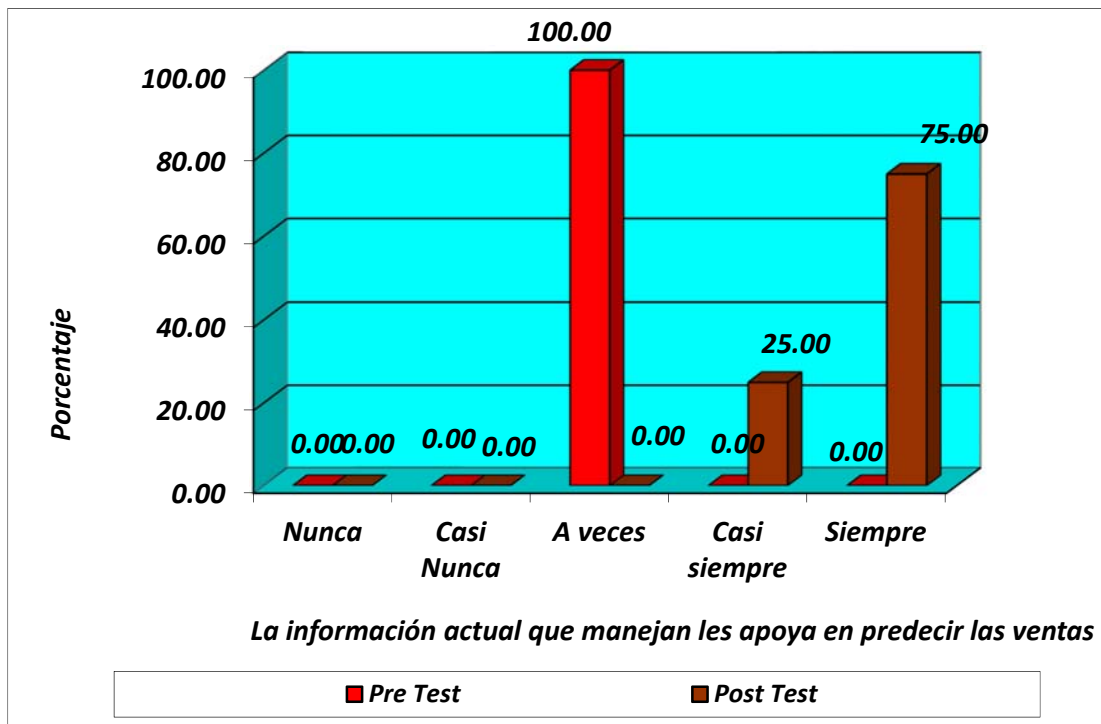


Fig. N° 65. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según La información actual que manejan les apoya en predecir las ventas

(Elaboración propia)

En la Tabla N° 31 y Figura N° 65 se observa que en la pregunta La información actual que manejan les apoya en predecir las ventas en el Pre Test el 100.00% precisa a veces, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 32. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Manejan información de indicadores para la toma de decisiones

Pregunta 14	Pre Test		Post Test	
Nunca	0	0.00	0	0.00
Casi nunca	1	25.00	0	0.00
A veces	3	75.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

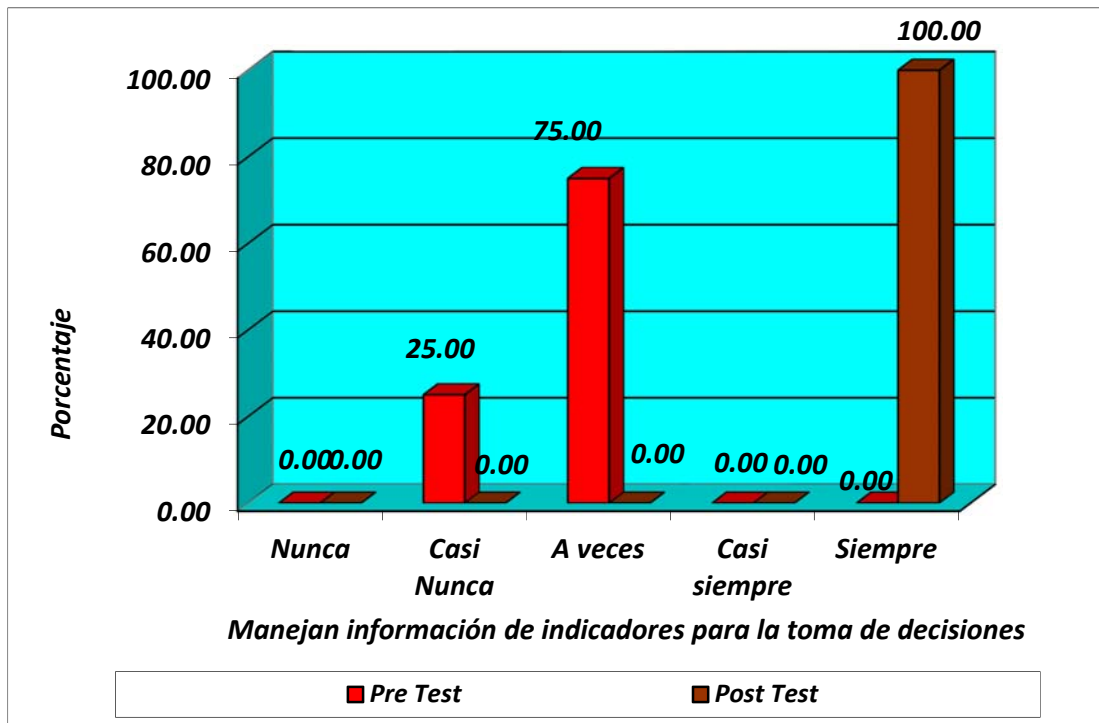


Fig. N° 66. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Manejan información de indicadores para la toma de decisiones

(Elaboración propia)

En la Tabla N° 32 y Figura N° 66 se observa que en la pregunta Manejan información de indicadores para la toma de decisiones en el Pre Test el 25.00% precisa casi nunca, el 75.00% precisa a veces, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 33. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Considera los reportes que maneja fiables

Pregunta 14	Pre Test		Post Test	
Nunca	2	50.00	0	0.00
Casi nunca	2	50.00	0	0.00
A veces	0	0.00	0	0.00
Casi siempre	0	0.00	0	0.00
Siempre	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

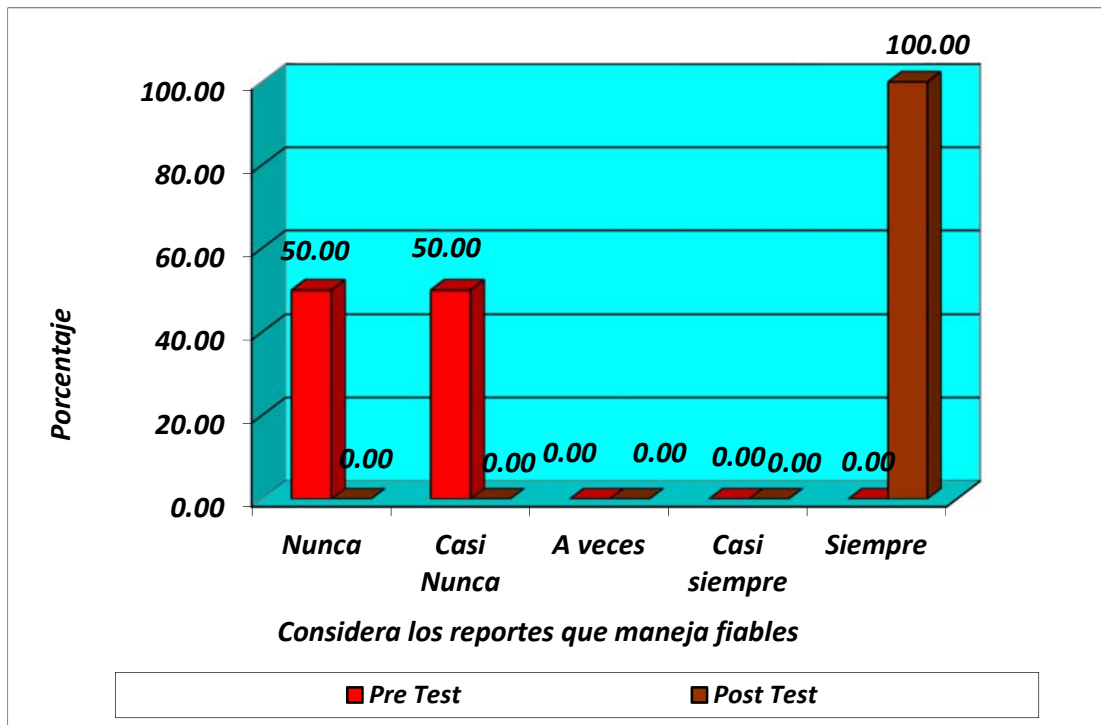


Fig. N° 67. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Considera los reportes que maneja fiables

(Elaboración propia)

En la Tabla N° 33 y Figura N° 67 se observa que en la pregunta Considera los reportes que maneja fiables en el Pre Test el 50.00% precisa nunca, el 50.00% precisa casi nunca, mientras que en el Post Test el 100.00% precisa siempre.

Tabla N° 34. Resumen Pre y Post Nivel de Satisfacción

Ejecutivos	Pre Test		Post Test	
	Total	Promedio	Total	Promedio
1	26	1.73	75	5.00
2	29	1.93	74	4.93
3	30	2.00	73	4.87
4	31	2.07	75	5.00
TOTAL	116	7.73	297	19.80

Nota: Elaboración Propia

En la tabla N° 34, se puede apreciar que ha habido un incremento en el nivel de satisfacción como se puede apreciar en el Pre Test fue de 7.73 en promedio para los 4 ejecutivos y en el Post Test fue de 19.80 en promedio para los 4 ejecutivos, por lo cual tuvimos un incremento de nivel de satisfacción en promedio de los 4 ejecutivos de 12.07.

Mínimo Puntaje: 26

Máximo Puntaje 75

Rango: $75-26=49$

Amplitud de intervalos No Satisfecho y Satisfecho: $49/2=24.5\sim 25$

Intervalos: No Satisfecho – [25-50], Satisfecho – [50-75].

Tabla N° 35. Distribución de las personas entrevistadas de la empresa de Transportes ITTSABUS SRL según Nivel de Satisfacción

Nivel de Satisfacción	Pre Test		Post Test	
No satisfecho [25-50]	4	100.00	0	0.00
Satisfecho [50-75]	0	0.00	4	100.00
Total	4	100.00	4	100.00

Nota: Elaboración Propia

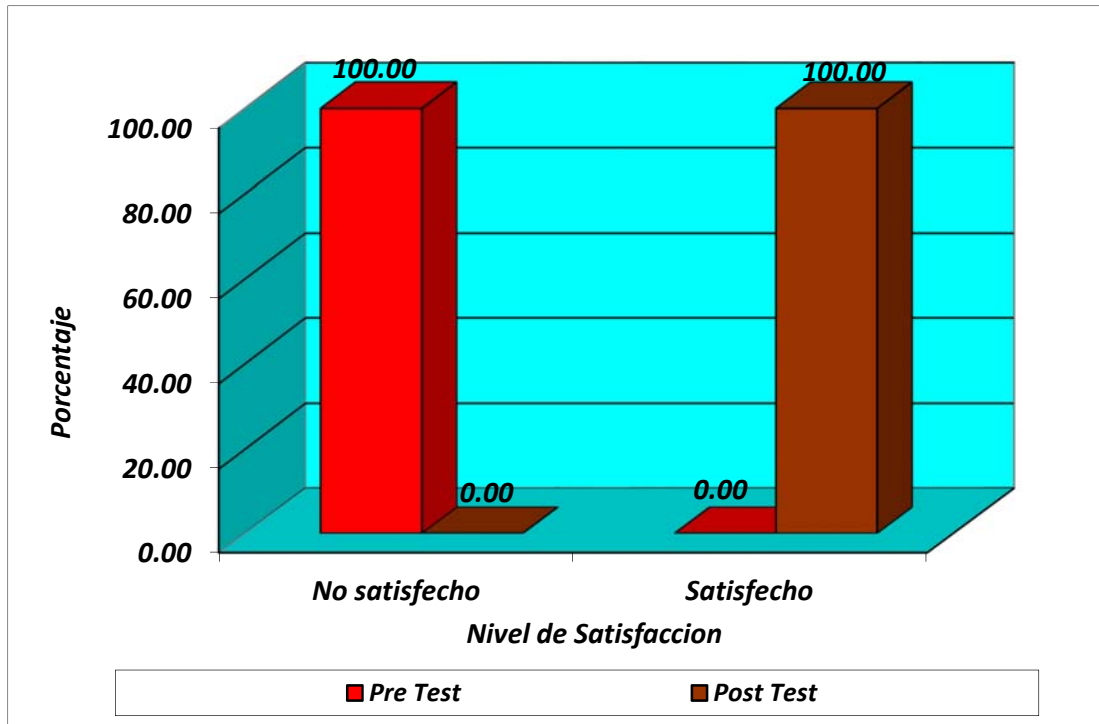


Fig. N° 68. Porcentaje de las personas entrevistadas de la empresa de transportes ITTSABUS SRL según Nivel de Satisfacción

(Elaboración propia)

En la Tabla N° 35 y Figura N° 68 se observa que el Nivel de Satisfacción en el Pre Test fue de 100.00% No satisfecho mientras que en el Post Test fue de 100.00% Satisfecho.

6.2. Indicador Tiempo del Proceso de Toma de Decisiones

Se tomó como muestra 30 días en un periodo de 2 meses, para los tiempos en promedio de los 4 ejecutivos.

Hipótesis Nula (Ho): La implementación de un data mart (Tiempo del proceso de toma de decisiones) no mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

Hipótesis Alternativa (Ha): La implementación de un data mart (Tiempo del proceso de toma de decisiones) mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

Tabla N° 36. Datos Tabulados Tiempo Proceso de Toma de Decisiones

Número	Pre Prueba (min)	Post Prueba (min)	Diferencia (min)
1	194	19.40	174.60
2	197	34.40	162.60
3	195	44.40	150.60
4	196	14.40	181.60
5	196	19.40	176.60
6	198	14.40	183.60
7	197	16.40	180.60
8	196	17.40	178.60
9	195	22.40	172.60
10	195	31.40	163.60
11	196	33.40	162.60
12	194	39.40	154.60
13	198	14.40	183.60
14	197	22.40	174.60
15	197	20.40	176.60
16	198	23.40	174.60
17	195	44.40	150.60
18	196	46.40	149.60
19	196	34.40	161.60
20	197	19.40	177.60
21	194	21.40	172.60
22	197	22.40	174.60
23	197	23.40	173.60
24	198	20.40	177.60
25	197	25.40	171.60
26	198	26.40	171.60
27	195	30.40	164.60
28	197	33.40	163.60
29	198	24.40	173.60
30	198	15.40	182.60
TOTAL	5892	775.00	5117.00
PROMEDIO	196.40	25.83	170.57

VARIANZA	1.70	87.01	98.93
DESVIACION	1.30	9.33	9.95
COEF. VAR.	0.66	36.11	5.83

Nota: Elaboración Propia

NIVEL DE SIGNIFICANCIA: $\alpha = 0.05$

ESTADÍSTICA DE PRUEBA: T de student (muestras dependientes)

$$T = \frac{D - 0}{\frac{S}{\sqrt{n}}} = \frac{170.57}{\frac{9.95}{\sqrt{30}}} = 93.89$$

Grado de libertad: $n - 2 = 30 - 2 = 28$ $T_{\text{tabla}} = 1.701$ con un nivel de significancia del 5%.

REGIONES

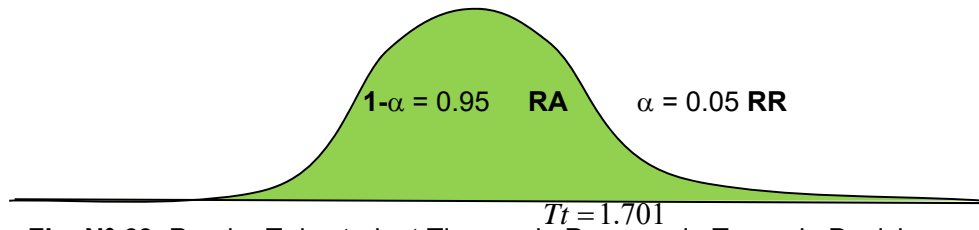


Fig. N° 69. Prueba T de student Tiempo de Proceso de Toma de Decisiones

(Elaboración Propia)

DECISIÓN: Ho se Rechaza, por lo tanto, la implementación de un data mart (Tiempo del proceso de toma de decisiones) mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L., mediante la prueba estadística T de Student para muestra dependientes a un nivel de significancia del 5%.

6.3. Indicador Cantidad de Reportes con Errores

Se tomó como muestra 30 días en un periodo de 2 meses, para la cantidad de reportes con errores en promedio de los 4 ejecutivos.

Tabla N° 37. Datos Tabulados Cantidad de Reportes Con Errores

Número	Pre Prueba	Post Prueba	Diferencia
1	2	0	2

2	1	0	1
3	0	0	0
4	0	0	0
5	0	0	0
6	2	0	2
7	1	0	1
8	0	0	0
9	0	0	0
10	1	0	1
11	1	0	1
12	1	0	1
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	1	0	1
18	2	0	2
19	0	0	0
20	2	0	2
21	0	0	0
22	0	0	0
23	1	0	1
24	0	0	0
25	0	0	0
26	1	0	1
27	0	0	0
28	1	0	1
29	0	0	0
30	1	0	1
TOTAL	18	0	18
PROMEDIO	0.60	0	0.60
VARIANZA	0.52	0	0.52
DESVIACION	0.72	0	0.72
COEF. VAR.	120.66	0	120.66

Hipótesis Nula (Ho): La implementación de un data mart (Cantidad de Reportes con errores) no mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

Hipótesis Alternativa (Ha): La implementación de un data mart (Cantidad de Reportes con errores) mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L.

NIVEL DE SIGNIFICANCIA: $\alpha = 0.05$

ESTADÍSTICA DE PRUEBA: T de student (muestras dependientes)

$$T = \frac{D-0}{\frac{S}{\sqrt{n}}} = \frac{0.60}{\frac{0.72}{\sqrt{30}}} = 4.56$$

Grado de libertad $n-2=30-2=28$ $T_{\text{tabla}}=1.701$ con un nivel de significancia del 5%.

REGIONES

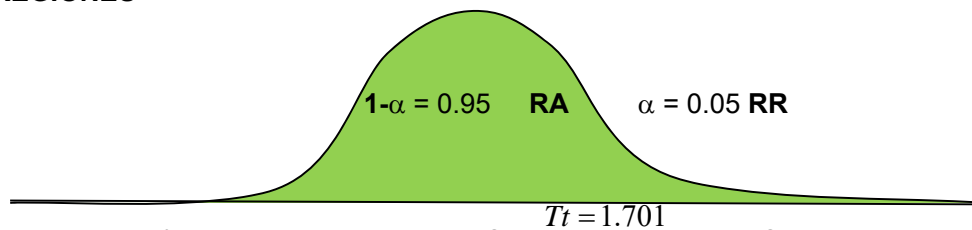


Fig. N° 70. Prueba T de student Cantidad de Reportes Con Errores

(Elaboración Propia)

DECISIÓN: H_0 se Rechaza, por lo tanto, la implementación de un data mart (Cantidad de Reportes con errores) mejora el proceso de toma de decisiones en el área comercial de la empresa ITTSABUS S.R.L., mediante la prueba estadística T de Student para muestra dependientes a un nivel de significancia del 5%.

CAPÍTULO 7. DISCUSIÓN

Después de analizar los resultados obtenidos se puede obtener lo siguiente:

- Que el Nivel de cumplimiento en el Pre Test el 100.00% es de No Satisfecho mientras que en el Post Test el 100.00% está Satisfecho, comparado con Alex Jesús Durand Mendoza (2014), quien en sus resultados obtuvo un nivel de satisfacción alto, le presente tesis tiene niveles de satisfacción alto (100%). Siendo la razón la calidad de interfaz e información brindada en la misma mediante la herramienta de explotación Qlik View.
- Martha Luz Tuñoque Julcas, Oswaldo Vilchez Zapata (2016), quien en sus resultados obtuvo que la implementación del data mart reduce de 4.1 horas a 10 segundos el proceso de generación de información de los centros de costos de la empresa constructora Beaver, en la presente tesis también se logra una reducción de tiempos en un porcentaje de 86.85% que representan 5117 minutos en promedio de los 4 ejecutivos del área comercial.
- Jubitza Lisbeth Salazar Tataje (2017), quien en sus resultados obtuvo que mediante la implementación de un data mart como herramienta de inteligencia de negocios se logra incrementar las ventas de la empresa Azaleia Perú y reducir los errores de la generación de reportes en un 100%, en la presente tesis también se logra una reducción de cantidad de reportes con errores en un 100% que representan 0 reportes con errores en las mediciones realizadas.

CONCLUSIONES

Se logró implementar el data mart mejorando el proceso de toma de decisiones del área comercial, lo cual queda evidenciado en lo siguiente:

- Se logró mejorar el proceso de toma de decisiones en relación al nivel de satisfacción del personal en un 100.00% en los 4 ejecutivos del área que toman decisiones en el día a día.
- Se logró mejorar el proceso de toma de decisiones en relación a los tiempos de proceso logrando una reducción del 86.85%.
- Se logró mejorar el proceso de toma de decisiones en relación a la cantidad de reportes con errores logrando una reducción del 100.00%.

RECOMENDACIONES

- Realizar encuestas a personal del área comercial que van a campo para saber si ahora sus actividades han mejorado en cuando a si ahora están mejor enfocadas y tienen metas establecidas por parte de sus Jefes Directos.
- Establecer una retroalimentación constante con los usuarios de la herramienta para ver en que otros aspectos se puede mejorar en la dimensión tiempo del proceso de toma de decisiones.
- Monitorear que el sistema siga funcionando de una manera óptima y establecer controles relacionados a la seguridad a la información para que el data mart no se vea afectado por algún usuario malintencionado.

REFERENCIAS

- ✓ Lluís Cano J. (2007). *Business Intelligence: Competir Con Información*. 1st ed. Barcelona: Banesto;
- ✓ McLeod, R. (2001). *Sistemas de información Gerencial - 658.011*. Pearson Education.
- ✓ Kenneth, L. C. (2004). *Sistemas de Información Gerencial – Administración de la empresa digital, 658.04/L29*. Pearson Education.
- ✓ J.S.Hammond, P. (2003). *La Toma de Decisiones*. Ediciones Deusto S.A.
- ✓ Brien, J. A. (2001). *Sistemas de Información Gerencial “Manejo de la Tecnología de la información en la empresa interconectada en red”, 658.004,I*. Editorial Mc Graw Hil.
- ✓ Navita Kumari. (2013). *Business Intelligence in a Nutshell*. Vol 1, Issue 4.
- ✓ Yusnier Reyes Dixson, Lissette Nuñez Maturel (2015). *La inteligencia de negocio como apoyo a la toma de decisiones en el ámbito académico*.
- ✓ Peña A., A. (2006). *Inteligencia de Negocios: Una Propuesta para su Desarrollo en las Organizaciones*. Instituto Politécnico Nacional, Dirección de Publicaciones, México.
- ✓ Josep Curto Díaz (2011). *Introducción al Business Intelligence*.
- ✓ Marinela Mircea, Bogdan Ghilic, Marian Stoica. (2011). *Combining Business Intelligence with Cloud Computing to delivery agility in actual economy*. Department of Economic Informatics. The Bucharest Academy of Economic Studies.
- ✓ Chuck Ballard, Daniel M. Farrell, Amit Gupta, Carlos Mazuela, Stanislav Vohnik. (2006). *Dimensional Modeling: In a Business Intelligence Environment*.
- ✓ Gregory S. Nelson. (2010). *Business Intelligence 2.0: Are we there yet?*. ThotWawe Technologies.
- ✓ EMA. (2014). *Unlocking Your IT Data with User-Driven BI*.
- ✓ *Consideraciones procesos ETL en entornos Big Data: Caso Hadoop*. (2014). Recuperado de: <http://www.dataprix.com/blog-it/big-data/consideraciones-procesos-etl-entornos-big-data-caso-hadoop>.
- ✓ *¿Cual es la diferencia entre una base de datos OLTP y OLAP?*. (2013). Recuperado de: <https://jorsodom.wordpress.com/2013/09/19/cual-es-la-diferencia-entre-una-base-de-datos-oltp-y-olap>.
- ✓ *¿Qué es Business Intelligence?*. (2018). Recuperado de: https://www.sinnexus.com/business_intelligence.

- ✓ Fundamentos BI. (2013). Recuperado de: <https://www.latino-bi.com/espanol/fundamentos-bi/introduccion-al-bi.php>.
- ✓ Gestión multidimensional en los Proyectos de BI.- Requerimientos del Director de Proyecto. (2018). Recuperado de: <https://decisionesytecnologia.wordpress.com/category/metodologia-business-intelligence>.
- ✓ Inteligencia de Negocios. (2014). Recuperado de: <http://inteligenciadenegociosval.blogspot.com/2014/01/metodologia-de-kimball.html>.
- ✓ Bernabeu Ricardo Dario. (2010). HEFESTO: Metodología para la Construcción de un Data Warehouse.
- ✓ Enfoques de Desarrollo DW. (2012). Recuperado de: <https://es.scribd.com/document/329929539/4-Metodologia-disegno-DW1-pdf>.
- ✓ Mendoza De los Santos, A. C. (2015). Curso de Inteligencia de Negocios. Trujillo: UPN.

ANEXOS

ANEXO N° 1. Cuestionario Para Medir la Variable Dependiente

Nivel de Satisfacción

I. Instrucciones

Estimado(a), el presente instrumento tiene la finalidad de recoger información sobre el Nivel de Satisfacción. Le pedimos que sea sincero en sus respuestas.

II. Información específica

Estimado, marque solo una de las opciones:

Nunca	Casi Nunca	A veces	Casi Siempre	Siempre
1	2	3	4	5

N°	Items	1	2	3	4	5
1	Siente que es productivo en cuanto a las actividades que realiza a diario					
2	La información que les brindan los reportes le es suficiente					
3	Tiene disponible la información necesaria para realizar su trabajo					
4	Está a gusto con el tiempo que toma obtener la información comercial					
5	Maneja información de ocupabilidad de los buses					
6	Maneja información detallada de los canales de venta					
7	Puede realizar su trabajo con rapidez					
8	Tiene disponibilidad de la información comercial las 24hrs del día					
9	Cree usted que actualmente puede analizar la información de manera adecuada					
10	Cree tener información de ventas exacta					
11	Está conforme con el tiempo que involucra la toma de decisiones					
12	Maneja información necesaria de los clientes					
13	La información actual que manejan les apoya en predecir las ventas					
14	Manejan información de indicadores para la toma de decisiones					
15	Considera los reportes que maneja fiables					

ANEXO N° 2. Ficha de Observación Para Medir la Variable Dependiente Porcentaje de Reducción del Tiempo del Proceso de Toma de Decisiones

FICHA DE OBSERVACION					
Se hará una medición del tiempo en promedio de los 4 ejecutivos en 30 días en un intervalo de tiempo de 2 meses desde que se solicita el reporte hasta que es entregado y se concluye la toma de decisión.					
Item	Pre Exp. (Min)	Fecha Pre Exp.	Post Exp. (Min)	Fecha Post Exp.	Observaciones
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
.					
.					
30					

ANEXO N° 3. Ficha de Observación Para Medir la Variable Dependiente Cantidad de Reportes Con Errores

FICHA DE OBSERVACION					
Se hará una medición de la cantidad de reportes con errores en promedio de los 4 ejecutivos en 30 días en un intervalo de tiempo de 2 meses desde que se solicita el reporte hasta que es entregado y se concluye la toma de decisión.					
Item	Pre Exp. (Unidad)	Fecha Pre Exp.	Post Exp. (Unidad)	Fecha Post Exp.	Observaciones
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
.					
.					
30					