



Universidad  
Nacional  
de Loja

## Universidad Nacional de Loja

### Facultad de la Energía, las Industrias y los recursos Naturales no Renovables

#### Carrera de Ingeniería en Sistemas

**Implementación de Business Intelligence para la gestión de datos generados en el módulo de pedidos durante los años 2018-2021 en la empresa Marketing Digital Bravo & Beltrán**

**Implementation of Business Intelligence for the management of data generated in the order module during the years 2018-2021 in the Digital Marketing company Bravo & Beltrán**

**Trabajo de Titulación previo a la obtención del título de Ingeniero en Sistemas**

**AUTOR:**

Luis Miguel Quito Carrión

**DIRECTOR:**

Ing. Edwin René Guamán Quinche, Mg. Sc.

Loja – Ecuador

2023

## Certificación

Loja, 05 de mayo de 2023

Ing. Edwin René Guamán Quinche, Mg. Sc.  
**DIRECTOR DEL TRABAJO DE TITULACIÓN**

### **CERTIFICO:**

Que he revisado y orientado todo el proceso de elaboración del Trabajo de Titulación denominado: **Implementación de Business Intelligence para la gestión de datos generados en el módulo de pedidos durante los años 2018-2021 en la empresa Marketing Digital Bravo & Beltrán**, previo a la obtención del título de **Ingeniero en Sistemas**, de la autoría del estudiante **Luis Miguel Quito Carrión**, con **cedula de identidad Nro. 1150024527**, una vez que el trabajo cumple con todos los requisitos exigidos por la Universidad Nacional de Loja, para el efecto, autorizo la presentación del mismo para su respectiva sustentación y defensa.

Ing. Edwin René Guamán Quinche, Mg. Sc.  
**DIRECTOR DEL TRABAJO DE TITULACIÓN**

## **Autoría**

Yo, **Luis Miguel Quito Carrión**, declaro ser autor del presente Trabajo de Titulación y eximo expresamente a la Universidad Nacional de Loja y a sus representantes jurídicos, de posibles reclamos o acciones legales por el contenido del mismo. Adicionalmente acepto y autorizo a la Universidad Nacional de Loja, la publicación de mi Trabajo de Titulación en el Repositorio Institucional – Biblioteca Virtual.

**Firma:**

**Cédula de identidad:** 1150024527

**Fecha:** 25 de septiembre del 2023

**Correo electrónico:** luis.quito@unl.edu.ec

**Teléfono:** 0983183013

**Carta de Autorización por parte del autor, para la consulta, reproducción parcial o total y/o publicación electrónica del texto completo, del Trabajo de Titulación.**

Yo, **Luis Miguel Quito Carrión**, declaro ser autor del Trabajo de Titulación denominado: **Implementación de Business Intelligence para la gestión de datos generados en el módulo de pedidos durante los años 2018-2021 en la empresa Marketing Digital Bravo & Beltrán**, como requisito para optar por el título de **Ingeniero en Sistemas**, autorizo al Sistema Bibliotecario de la Universidad Nacional de Loja para que, con fines académicos muestre la producción intelectual de la Universidad, a través de la visibilidad de su contenido en el Repositorio Institucional.

Los usuarios pueden consultar el contenido de este trabajo en el Repositorio Institucional, en las redes de información del país y del exterior con las cuales tenga convenio la Universidad.

La Universidad Nacional de Loja, no se responsabiliza por el plagio o copia del Trabajo de Titulación que realice un tercero.

Para constancia de esta autorización, en la ciudad de Loja a los veinticinco días del mes de septiembre del dos mil veintitrés.

**Firma:**

**Autor:** Luis Miguel Quito Carrión

**Cédula de identidad:** 1150024527

**Dirección:** Loja, barrio Menfis Central

**Correo electrónico:** luis.quito@unl.edu.ec

**Teléfono:** +593 983 183 013

**DATOS COMPLEMENTARIOS:**

**Director del Trabajo de Titulación:** Ing. Edwin René Guamán Quinche, Mg. Sc.

## **Dedicatoria**

A mis queridos padres, Carmita y Víctor, por su constante apoyo, amor y sacrificio en mi proceso de formación profesional. Su incondicional respaldo ha sido la base fundamental para llegar hasta aquí y estoy agradecido eternamente por todo lo que han hecho por mí.

Dedico también este Trabajo de Titulación a todas las personas que me han apoyado en el camino hacia la culminación de mi carrera con éxito. A mi familia, compañeros, amigos y docentes, quienes de una u otra manera me han brindado su contribución para alcanzar este objetivo. Gracias a su valioso conocimiento, consejos y palabras de aliento día a día, he logrado culminar esta meta y alcanzar mi logro profesional.

***Luis Miguel Quito Carrión***

## **Agradecimiento**

Quiero expresar mi agradecimiento a la Carrera de Ingeniería en Sistemas de la Universidad Nacional de Loja por haberme brindado la oportunidad de formarme como profesional. También quiero agradecer a mis docentes por impartirme sus conocimientos día a día mediante su enseñanza.

De manera especial al Ing. Rene Guamán, mi Director del Trabajo de Titulación, por guiarme en todo momento, brindarme su valioso conocimiento, dedicar su tiempo y tener paciencia durante todo el desarrollo del Trabajo de investigación.

A la empresa Marketing Digital Bravo & Beltrán por haberme brindado la oportunidad de realizar mi Trabajo de Titulación en sus instalaciones.

A Dios por darme la sabiduría necesaria para lograr esta meta propuesta.

Por último, deseo expresar mi agradecimiento eterno a mis seres queridos por todo su esfuerzo, apoyo, paciencia y amor incondicional durante mi proceso de formación académica.

***Luis Miguel Quito Carrión***

## Índice de Contenidos

<b>Portada</b> .....	i
<b>Certificación</b> .....	ii
<b>Autoría</b> .....	iii
<b>Carta de autorización</b> .....	iv
<b>Dedicatoria</b> .....	v
<b>Agradecimiento</b> .....	vi
<b>Índice de contenidos</b> .....	vii
Índice de tablas.....	x
Índice de figuras.....	xii
Índice de anexos.....	xvii
<b>1. Título</b> .....	<b>1</b>
<b>2. Resumen</b> .....	<b>2</b>
2.1. Abstract.....	3
<b>3. Introducción</b> .....	<b>4</b>
<b>4. Marco teórico</b> .....	<b>6</b>
4.1 Business Intelligence.....	6
4.1.1 Antecedentes.....	6
4.1.2 Definición.....	7
4.1.3 Características.....	8
4.1.4 Beneficios de BI.....	8
4.1.6 Factores para la implementación de BI.....	9
4.2 Componentes de BI.....	10
4.2.1 Fuentes de Información.....	11
4.2.2 Data Warehouse.....	11
4.2.3 Data mart.....	14
4.2.4 Procesos ETL.....	15
4.2.5 Proceso analítico en línea OLAP.....	17
4.2.6 Herramientas de visualización.....	18
4.3 Modelado Dimensional.....	19
4.4 Herramientas para el desarrollo de BI.....	20

4.4.1 Pentaho.....	21
4.6 Metodología del ciclo de vida de Ralph Kimball.....	24
4.7 Trabajos Relacionados .....	27
<b>5. Metodología.....</b>	<b>29</b>
5.1 Tipo de Investigación.....	29
5.2 Contexto.....	29
5.3 Proceso.....	29
5.4 Recursos.....	31
5.5 Estándares.....	32
5.6 Metodología de desarrollo.....	32
5.7 Participantes.....	32
<b>6. Resultados.....</b>	<b>33</b>
6.1 OBJETIVO 1: Analizar los datos históricos generados en el módulo de pedidos para establecer el diseño del modelo.....	33
6.1.1 Análisis empresarial.....	33
6.1.2 Especificación de requerimientos .....	34
6.1.3 Diseño de la arquitectura.....	38
6.1.4 Selección e instalación de herramientas.....	40
6.2 OBJETIVO 2: Desarrollar el modelo de datos utilizando la plataforma de BI Pentaho.....	40
6.2.1 Modelado Dimensional .....	40
6.2.2 Modelo Lógico .....	43
6.2.3 Modelo Físico .....	44
6.2.4 Implementación de procesos ETL.....	45
6.3 OBJETIVO 3: Implementar una solución de BI a través de un Dashboard.....	55
6.3.2 Desarrollo de la aplicación BI .....	58
6.3.3 Despliegue de la aplicación BI.....	66
6.3.5 Pruebas de la solución BI .....	68
<b>7. Discusión.....</b>	<b>75</b>
7.1 Desarrollo de la propuesta alternativa.....	75
7.2 Valoración técnica, económica y social.....	78
<b>8. Conclusiones.....</b>	<b>82</b>
<b>9. Recomendaciones.....</b>	<b>83</b>

<b>10. Bibliografía.....</b>	<b>84</b>
<b>11. Anexos .....</b>	<b>88</b>

**Índice de tablas:**

<b>TABLA I. FACTORES PARA LA IMPLEMENTACIÓN DE BI.....</b>	<b>9</b>
<b>TABLA II. DESCRIPCION DE LOS COMPONENTES DEL PROCESO ETL.....</b>	<b>16</b>
<b>TABLA III. TIPOS DE MODELADO ETL.....</b>	<b>17</b>
<b>TABLAIV. COMPARACIÓN DE CARACTERÍSTICAS PENTAHO COMMUNITY &amp; ENTERPRISE EDITION.....</b>	<b>23</b>
<b>TABLA V. TRABAJOS RELACIONADOS.....</b>	<b>28</b>
<b>TABLA VI. PARTICIPANTES DEL TT.....</b>	<b>32</b>
<b>TABLA VII. REQUISITOS FUNCIONALES.....</b>	<b>35</b>
<b>TABLA VIII. REQUERIMIENTOS NO FUNCIONALES.....</b>	<b>35</b>
<b>TABLA IX. CASO DE USO LOGIN Y LOGOUT.....</b>	<b>36</b>
<b>TABLA X. CASO DE USO VISUALIZAR DASHBOARD.....</b>	<b>37</b>
<b>TABLA XI. CASO DE USO AGREGAR USUARIO.....</b>	<b>37</b>
<b>Tabla XII. FUENTE DE DATOS.....</b>	<b>38</b>
<b>Tabla XIII. MATRIZ BUS.....</b>	<b>41</b>
<b>TABLA XIV. DESCRIPCIÓN DE DIMENSIONES.....</b>	<b>42</b>
<b>TABLA XV. DESCRIPCIÓN DE MEDIDAS.....</b>	<b>43</b>
<b>TABLA XVI. DESCRIPCIÓN TABLA “dim_store”.....</b>	<b>44</b>
<b>TABLA XVII. RESULTADO DE LAS PRUEBAS DEL CUBO OLAP.....</b>	<b>68</b>
<b>TABLA XVIII. RESULTADOS DE LA PRUEBA DE CARGA Y ESTRÉS.....</b>	<b>70</b>
<b>TABLA XIX. VALORACIÓN ECONÓMICA: TALENTO HUMANO.....</b>	<b>79</b>
<b>TABLA XX. VALORACIÓN ECONÓMICA: HARDWARE Y SOFTWARE.....</b>	<b>80</b>
<b>TABLA XXI. VALORACIÓN ECONÓMICA: SERVICIOS.....</b>	<b>80</b>
<b>TABLA XXII. PRESUPUESTO FINAL.....</b>	<b>80</b>
<b>TABLA XXIII. ACRÓNIMOS Y ABREVIATURAS.....</b>	<b>93</b>
<b>TABLA XXIV. REQUERIMIENTO DE SEGURIDAD.....</b>	<b>98</b>
<b>TABLA XXV. REQUERIMIENTO DE ESCALABILIDAD.....</b>	<b>98</b>
<b>TABLA XXVI. REQUERIMIENTO DE FACILIDAD DE USO.....</b>	<b>98</b>

<b>TABLA XXVII. REQUERIMIENTO DE CONFIABILIDAD .....</b>	<b>99</b>
<b>TABLA XXVIII. REQUERIMIENTO DE DISPONIBILIDAD .....</b>	<b>99</b>
<b>TABLA XXIX. REQUERIMIENTO DE PORTABILIDAD .....</b>	<b>99</b>
<b>TABLA XXX. DESCRIPCIÓN DIMENSIÓN TIEMPO .....</b>	<b>102</b>
<b>TABLA XXXI. DESCRIPCIÓN DIMENSIÓN EMPRESA .....</b>	<b>102</b>
<b>TABLA XXXII. DESCRIPCIÓN DIMENSIÓN PRODUCTO .....</b>	<b>103</b>
<b>TABLA XXXIII. DESCRIPCIÓN DIMENSIÓN TIPO DE PAGO .....</b>	<b>103</b>
<b>TABLA XXXIV. DESCRIPCIÓN DIMENSIÓN USUARIO .....</b>	<b>103</b>
<b>TABLA XXXV. DESCRIPCIÓN HECHO PEDIDOS.....</b>	<b>103</b>
<b>TABLA XXXVI. RESULTADO DE PRUEBAS DE CARGA Y ESTRÉS.....</b>	<b>128</b>
<b>TABLA XXXVII. RESULTADOS DE PREGUNTAS RELACIONADAS CON LA SATISFACCIÓN DEL GERENTE CON RESPECTO A LA SOLUCIÓN BI DESARROLLADA. ....</b>	<b>129</b>

## Índice de figuras:

<b>Figura 1.</b> Línea de tiempo del BI .....	6
<b>Figura 2.</b> Componentes de Business Intelligence. ....	10
<b>Figura 3.</b> Características del Datawarehouse. ....	12
<b>Figura 4.</b> Estructura del Datawarehouse.....	13
<b>Figura 5.</b> Flujo de datos en el DW.....	14
<b>Figura 6.</b> Arquitectura Top-Down.....	15
<b>Figura 7.</b> Arquitectura Botton-Up. ....	15
<b>Figura 8.</b> Procesos ETL .....	15
<b>Figura 9.</b> Herramientas ETL.....	17
<b>Figura 10.</b> Esquema en estrella. ....	20
<b>Figura 11.</b> Esquema copo de nieve. ....	20
<b>Figura 12.</b> Herramientas de BI.....	21
<b>Figura 13.</b> Metodología ciclo de vida dimensional del negocio.....	25
<b>Figura 14.</b> Diagrama de procesos de pedidos.....	34
<b>Figura 15.</b> Caso de uso - Gerente.....	36
<b>Figura 16.</b> Reporte de Datos Modulo Pedidos .....	38
<b>Figura 17.</b> Back Room. ....	39
<b>Figura 18.</b> Front Room.....	39
<b>Figura 19.</b> Matriz Start Net de Tabla de Hechos Pedidos.....	42
<b>Figura 20.</b> Modelo lógico.....	43
<b>Figura 21.</b> Modelo Físico .....	44
<b>Figura 22.</b> Conexión de PDI con PostgreSQL .....	45
<b>Figura 23.</b> Job Principal .....	46
<b>Figura 24.</b> Job_Load_Stg.....	47
<b>Figura 25.</b> Transformación Order_Stg.....	47
<b>Figura 26.</b> Step de Lectura del archivo CSV en PDI .....	48
<b>Figura 27.</b> Step Replace to Delivery .....	48

<b>Figura 28.</b> Step Replace to Pick Up .....	48
<b>Figura 29.</b> Step Select Values .....	49
<b>Figura 30.</b> Step sort rows.....	49
<b>Figura 31.</b> Step Table output .....	49
<b>Figura 32.</b> Job Load_Dim.....	50
<b>Figura 33.</b> Transformación Order_Dim.....	50
<b>Figura 34.</b> Step table input Order_Dim.....	50
<b>Figura 35.</b> Step Dimensión lookup/update Dim_Order .....	51
<b>Figura 36.</b> Job Carga Tablas de hechos .....	51
<b>Figura 37.</b> Transformación Pre_Fact_Table.....	52
<b>Figura 38.</b> Database lookup Order.....	52
<b>Figura 39.</b> Conexión de Schema Workbench al Data Warehouse.....	53
<b>Figura 40.</b> Creación de Schema y Cubo .....	53
<b>Figura 41.</b> Dimensiones del cubo "ordersCube" .....	54
<b>Figura 42.</b> Jerarquía Dimensión Stores.....	54
<b>Figura 43.</b> Medidas del cubo "ordersCube" .....	55
<b>Figura 44.</b> Creación de medidas .....	55
<b>Figura 45.</b> Inicialización de Pentaho Server en Google Chrome .....	56
<b>Figura 46.</b> Publicación cubo OLAP .....	56
<b>Figura 47.</b> Campos para publicación de cubo .....	57
<b>Figura 48.</b> Publicación del cubo "ordersCube" exitosa.....	57
<b>Figura 49.</b> Administración de fuentes de datos en Pentaho Server.....	58
<b>Figura 50.</b> Conexión Pentaho-Data warehouse .....	58
<b>Figura 51.</b> Lectura de cubo con JPivot.....	59
<b>Figura 52.</b> Cubo ordersCube en Pentaho Server .....	59
<b>Figura 53.</b> Creación del Dashboard_donesitios .....	59
<b>Figura 54.</b> Componentes de un dashboard.....	60
<b>Figura 55.</b> Diseño del Layout Panel del RF01 .....	60

<b>Figura 56.</b> Configuración del Layout Panel del RF01 .....	61
<b>Figura 57.</b> Configuración del Datasource Panel del RF01.....	61
<b>Figura 58.</b> Creación de consulta MDX para el RF01 .....	62
<b>Figura 59.</b> Configuración del Component Panel del RF01 .....	62
<b>Figura 60.</b> Reporte del RF01 .....	63
<b>Figura 61.</b> Dashboard general de Pedidos.....	64
<b>Figura 62.</b> Dashboard por empresa .....	65
<b>Figura 63.</b> Diagrama de despliegue .....	66
<b>Figura 64.</b> Despliegue de solución BI al Servidor.....	66
<b>Figura 65.</b> Validación de la solución BI en el servidor .....	67
<b>Figura 66.</b> Login de Pentaho en el servidor .....	67
<b>Figura 67.</b> Pruebas de medidas por consulta MDX.....	68
<b>Figura 68.</b> Pruebas de medidas por consulta SQL.....	68
<b>Figura 69.</b> Resultado reporte usuarios con más pedidos .....	69
<b>Figura 70.</b> Pruebas de reportes por consulta SQL .....	70
<b>Figura 71.</b> Respuesta del gerente a pregunta sobre el cumplimiento de los requisitos solicitados en la solución BI.....	71
<b>Figura 72.</b> Respuesta del gerente a pregunta sobre el impacto de la solución BI en la toma de decisiones en la empresa. ....	72
<b>Figura 73.</b> Respuesta del gerente a pregunta sobre el diseño de la interfaz amigable y su usabilidad .....	72
<b>Figura 74.</b> Respuesta del gerente a pregunta sobre la utilidad de información presentada en el dashboard.....	73
<b>Figura 75.</b> Respuesta del gerente a pregunta sobre el tiempo de acceso al dashboard .....	73
<b>Figura 76.</b> Respuesta del gerente a pregunta sobre la mejora en la gestión de pedidos .....	74
<b>Figura 77.</b> Resultados consolidados de preguntas relacionadas con la satisfacción del gerente con respecto al dashboard desarrollado.....	75
<b>Figura 78.</b> Logo empresarial.....	88
<b>Figura 79.</b> Organigrama Empresarial.....	89

<b>Figura 80.</b> Diagrama de procesos pedidos. ....	90
<b>Figura 81.</b> Descarga de JRE.....	105
<b>Figura 82.</b> Versión de JRE.....	105
<b>Figura 83.</b> Acuerdo de licencia JRE.....	106
<b>Figura 84.</b> Instalación de JRE.....	106
<b>Figura 85.</b> Fin de instalación de JRE .....	107
<b>Figura 86.</b> Descarga de Pentaho Data Integration .....	107
<b>Figura 87.</b> Carpeta de Pentaho Data Integration.....	108
<b>Figura 88.</b> Ejecución de PDI .....	108
<b>Figura 89.</b> Entorno PDI .....	109
<b>Figura 90.</b> Repositorio de descarga de Schema Workbench. ....	109
<b>Figura 91.</b> Carpeta schema-workbench. ....	110
<b>Figura 92.</b> Entorno Mondrian .....	110
<b>Figura 93.</b> Repositorio de descarga Pentaho Server.....	111
<b>Figura 94.</b> Carpeta de Pentaho Server .....	111
<b>Figura 95.</b> Entorno Pentaho Server .....	112
<b>Figura 96.</b> Descarga de PostgreSQL .....	112
<b>Figura 97.</b> Ejecución del Instalador de PostgreSQL.....	112
<b>Figura 98.</b> Directorio de instalación PostgreSQL .....	113
<b>Figura 99.</b> Instalación de componentes de PostgreSQL .....	113
<b>Figura 100.</b> Asignación de contraseña para el super usuario.....	113
<b>Figura 101.</b> Puerto para levantar servidor de PostgreSQL.....	114
<b>Figura 102.</b> Configuración Regional.....	114
<b>Figura 103.</b> Configuraciones para la instalación de PostgreSQL .....	114
<b>Figura 104.</b> Instalación de PostgreSQL .....	115
<b>Figura 105.</b> Culminación de instalación de PostgreSQL .....	115
<b>Figura 106.</b> Dashboard de pgAdmin. ....	116
<b>Figura 107.</b> Job Load_Stg: steps de transformaciones comunes.....	117

<b>Figura 108.</b> Select values: transformación "Product_Stg" .....	117
<b>Figura 109.</b> Select values: transformación "Store_Stg" .....	117
<b>Figura 110.</b> Select values: transformación "TypePay_Stg". .....	118
<b>Figura 111.</b> Select values: transformación "User_Stg".....	118
<b>Figura 112.</b> Transformación Values_Stg.....	118
<b>Figura 113.</b> Select values: transformación "Values_Stg".....	118
<b>Figura 114.</b> Job Load_Dim: steps de transformaciones comunes.....	119
<b>Figura 115.</b> Table Input: transformación "Store_Dim". .....	119
<b>Figura 116.</b> Dimensión lookup/update: transformación "Store_Dim". .....	119
<b>Figura 117.</b> Table Input: transformación "TypePay_Dim".....	120
<b>Figura 118.</b> Dimensión lookup/update: transformación "TypePay_Dim".....	120
<b>Figura 119.</b> Transformación "Date_Dim".....	120
<b>Figura 120.</b> Jerarquía Dimensión Users.....	121
<b>Figura 121.</b> Jerarquía dimensión Orders.....	121
<b>Figura 122.</b> Jerarquía dimensión Products. ....	121
<b>Figura 123.</b> Jerarquía dimensión TypePays.....	122
<b>Figura 124.</b> Jerarquía dimensión Time.....	122
<b>Figura 125.</b> Servidor de DigitalOcean .....	123
<b>Figura 126.</b> Droplet pentaho-donesitios .....	123
<b>Figura 127.</b> Configuración de Putty.....	124
<b>Figura 128.</b> Agregar grupo de Hilos .....	126
<b>Figura 129.</b> Agregar petición HTTP .....	126
<b>Figura 130.</b> Componentes de visualización de pruebas.....	127
<b>Figura 131.</b> Componentes del plan de pruebas .....	127
<b>Figura 132.</b> Configuración de hilos .....	127
<b>Figura 133.</b> Configuración de petición HTTP .....	128

## Índice de anexos:

<b>Anexo 1.</b> Análisis Empresarial .....	88
<b>Anexo 2.</b> Especificación de requisitos.....	91
<b>Anexo 3.</b> Entrevista.....	100
<b>Anexo 4.</b> Descripción de tablas del Data mart .....	102
<b>Anexo 5.</b> Instalación de Herramientas .....	105
<b>Anexo 6.</b> Llenado de datos .....	117
<b>Anexo 7.</b> Definición de Jerarquías .....	121
<b>Anexo 8.</b> Configuración para el servidor del sistema. ....	123
<b>Anexo 9.</b> Ejecución de pruebas .....	126
<b>Anexo 10.</b> Validación y aceptación de la solución BI .....	130
<b>Anexo 11.</b> Certificado de traducción del resumen.....	131

## **1. Título**

**Implementación de Business Intelligence para la gestión de datos generados en el módulo de pedidos durante los años 2018-2021 en la empresa Marketing Digital Bravo & Beltrán**

## 2. Resumen

En la actualidad, las pequeñas y medianas empresas se enfrentan a desafíos significativos en cuanto a la gestión y el análisis de grandes volúmenes de datos. Estas compañías a menudo carecen de los recursos y la experiencia necesarios para convertir estos datos en información valiosa y útil para la toma de decisiones. BI se presenta como una herramienta que permite a estas organizaciones enfrentar estos desafíos. Al utilizar BI, las empresas pueden identificar patrones y tendencias que les permiten tomar decisiones más informadas y estratégicas. Además, la implementación de esta herramienta proporciona una mayor visibilidad y transparencia en los procesos empresariales, lo que mejora la confianza de los clientes y aumenta la rentabilidad de la organización. El no utilizar BI puede tener desventajas significativas para las empresas. La falta de información estratégica puede llevar a una toma de decisiones inadecuada y a la pérdida de oportunidades comerciales. Las empresas pueden encontrarse lidiando con una gran cantidad de datos no estructurados, lo que dificulta la identificación de patrones y tendencias, que retrasan la toma de decisiones.

En este contexto, el presente TT implementó una solución de Business Intelligence para el área de pedidos de la empresa Marketing Digital Bravo & Beltrán, que se encarga de gestionar adecuadamente los datos generados en la empresa y los convierte en información de valor, a la cual se puede acceder de manera más eficiente y oportuna a través de un dashboard, con el fin de dar soporte en el proceso de toma de decisiones al gerente de la organización.

Para alcanzar este objetivo el trabajo se desarrolló en tres fases. En la primera fase se analizó los datos de la organización con el fin de establecer los requerimientos y el diseño del modelo de datos, se obtuvieron mediante el uso del estándar IEEE 830. En la segunda etapa se definieron las dimensiones, medidas y tablas de hecho necesarias para la construcción del Datawarehouse, además de la implementación de procesos ETL para su población. Finalmente, en la tercera etapa se llevó a cabo la implementación y despliegue de la solución BI a través de un Dashboard. Es importante mencionar que para cada etapa del TT se utilizaron las fases de la metodología propuesta por Ralph Kimball, y como herramientas para el desarrollo de la solución BI la suite de Pentaho.

Los resultados obtenidos en el TT consisten en un dashboard que presenta información de los tres últimos años de los pedidos, empresas, productos y usuarios de la organización. A través de elementos visuales intuitivos, este dashboard hace que la información sea más fácil de comprender y acceder, lo que beneficia a la empresa a reducir tiempo de acceso a la información, mejorar la toma de decisiones y el desempeño general.

**Palabras Clave:** *Business Intelligence, Ralph Kimball, procesos ETL.*

## 2.1. Abstract

Currently, small and medium-sized enterprises face significant challenges regarding the management and analysis of large volumes of data. These companies often lack the necessary resources and expertise to turn this data into valuable and useful information for decision-making. BI is presented as a tool that allows these organizations to tackle these challenges. By using BI, companies can identify patterns and trends that enable them to make more informed and strategic decisions. In addition, the implementation of this tool provides greater visibility and transparency in business processes, improving customer confidence and increasing the profitability of the organization. Not using BI can have significant disadvantages for companies. The lack of strategic information can lead to inadequate decision-making and loss of business opportunities. Companies can find themselves dealing with a large amount of unstructured data, which makes it difficult to identify patterns and trends, delaying decision-making.

In this context, the present TT implemented a Business Intelligence solution for the orders area of the Marketing Digital Bravo & Beltrán company, which properly manages the data generated in the company and converts it into valuable information that can be accessed more efficiently and timely through a dashboard, in order to support the decision-making process of the organization's manager.

To achieve this objective, the work was developed in three phases. In the first phase, the organization's data was analyzed in order to establish the requirements and design of the data model, obtained through the use of the IEEE 830 standard. In the second stage, the necessary dimensions, measures, and fact tables were defined for the construction of the Data Warehouse, in addition to the implementation of ETL processes for its population. Finally, in the third stage, the implementation and deployment of the BI solution was carried out through a Dashboard. It is important to mention that for each stage of the TT, the phases of the methodology proposed by Ralph Kimball were used, and the Pentaho suite was used as a tool for the development of the BI solution.

The results obtained in the TT consist of a dashboard that presents information on the organization's orders, companies, products, and users for the last three years. Through intuitive visual elements, this dashboard makes the information easier to understand and access, which benefits the company in reducing access time to information, improving decision-making, and overall performance.

**Keywords:** *Business Intelligence, Ralph Kimball, ETL processes.*

### 3. Introducción

En el contexto de la sociedad de la información, la gestión de datos es un proceso necesario que deben llevar a cabo las empresas. Al contar con datos precisos y útiles, la gerencia de las empresas cuenta con una herramienta que permite tomar decisiones informadas y estratégicas que ayudan a mejorar el rendimiento general de la empresa. Muchas organizaciones optan por implementar nuevas tecnologías para obtener una ventaja competitiva sobre sus principales competidores [1]. BI se presenta como un soporte para la gestión de los datos de la empresa, convirtiéndolos en información de calidad [2]. El objetivo de BI es asistir a los encargados de la toma de decisiones en la adquisición de información adecuada y en la toma de decisiones precisas en la gestión de la empresa [3]. BI también tiene varias ventajas en su uso, incluyendo en primer lugar, la eliminación del trabajo basado en suposiciones [4]. Ayuda a identificar oportunidades de negocio [5]. Permite comprender el comportamiento de los consumidores [6]. Finalmente ayuda a crear objetivos realistas y aumenta la eficacia [7].

La implementación de Business Intelligence puede presentar una serie de desafíos para las empresas. En primer lugar, la integración de diferentes sistemas y bases de datos en una plataforma BI, puede requerir una compleja configuración técnica, lo que podría resultar en problemas de compatibilidad con ciertos sistemas que no son compatibles con las herramientas de desarrollo de BI [8]. Además, la falta de personal capacitado para manejar la implementación de BI podría retrasar o incluso impedir el éxito del proyecto [9]. Así mismo, los problemas de calidad de datos, como la falta de estandarización, estructuración o actualización adecuada de los datos, podrían dificultar la creación de informes precisos y significativos a través de BI [10].

En la actualidad las herramientas ofrecidas por BI, han hecho posible que las empresas se vean en la necesidad de implementar sistemas que manejen grandes volúmenes de datos para obtener información de valor, este es el caso de los datos generados en el módulo de pedidos de la empresa Marketing Digital Bravo & Beltrán, este conjunto de datos fue recolectado para determinar información útil, debido a que cada pedido contiene datos de interés sobre el cliente y productos, que no se puede aprovechar debido a la dificultad de entender los datos a simple vista. En este contexto se planteó la implementación de una solución de BI, tomando como base la metodología de Kimball, así como el uso de software libre, para la creación de un dashboard que se ajuste a las necesidades de la empresa, que permita visualizar información de valor y que brinde apoyo en el proceso de toma de decisiones a la gerencia de la empresa. Por ello, se planteó como objetivo principal “Implementar una solución BI para la gestión de datos generados en el módulo de pedidos en la empresa Marketing Digital Bravo & Beltrán”.

Para finalizar, el documento se compone en secciones. En el Marco Teórico, se incluyen varios capítulos de información bibliográfica relevante sobre Business Intelligence. En la sección de Metodología, se detallan los procesos, métodos, técnicas y estándares utilizados, así como los recursos y participantes involucrados en el desarrollo del TT. En la sección de Resultados, se presentan todos los artefactos que representan el diseño, construcción y despliegue de la solución BI, divididas en tres secciones, una para cada objetivo específico. La sección de Discusión analiza los resultados obtenidos y presenta las contribuciones del autor en función de los resultados y la literatura revisada. Las Conclusiones resumen los principales hallazgos del TT. Finalmente, en la sección de Recomendaciones se presentan sugerencias y aspectos a tener en cuenta para trabajos futuros.

## 4. Marco Teórico

El propósito de esta sección es presentar los conceptos relacionados con la temática, los cuales sustentan el desarrollo del presente TT. Esta información ha sido recopilada a través de un proceso de revisión bibliográfica. Se inicia con la conceptualización sobre Business Intelligence, a continuación, se definen los componentes y las tecnologías para el desarrollo de BI, seguido de la contextualización de la metodología de Ralph Kimball. Finalmente se presentan los diversos trabajos relacionados con el objeto de estudio.

### 4.1 Business Intelligence

Conforme a la sociedad de la información recalca la necesidad de tener mejores, rápidos y más eficientes métodos para la extracción y transformación de los datos en información de valor y esta a su vez convertirla en conocimiento que apoye en el proceso de toma de decisiones.

#### 4.1.1 Antecedentes

BI ha progresado con el tiempo y desde su aparición, ha ido evolucionando a un ritmo acelerado. En la Figura 1 realizada en base a [11], se ilustra el avance que ha tenido el BI a través del tiempo.

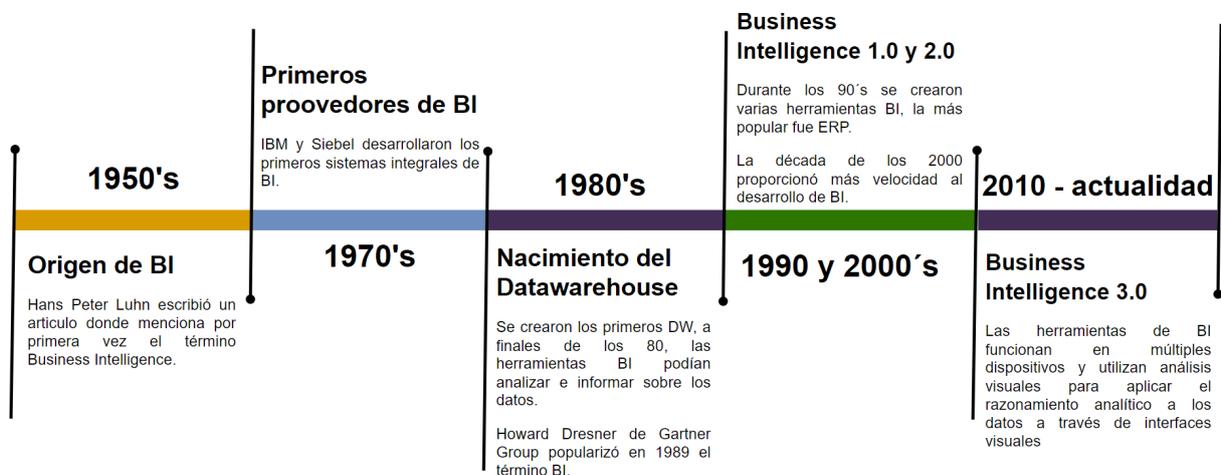


Figura 1. Línea de tiempo del BI

Lo que se conoce actualmente como BI nace en el año 1958, donde el investigador alemán de IBM Hans Peter Luhn en su artículo titulado “A Business Intelligence System” menciona por primera vez el término Business Intelligence, y lo define como “La capacidad de aprender las interrelaciones de los hechos presentados de manera que guíen la acción hacia un objetivo deseado” [7]; Con la publicación de su artículo, Luhn plantó la semilla del concepto de BI tal como se lo conoce en la actualidad.

En los años 70 aparecieron los primeros proveedores de BI (IBM y Siebel), las herramientas estuvieron disponibles para ayudar a acceder y organizar los datos de manera más efectiva [11]. Aunque el acceso a los datos había mejorado, la falta de infraestructura y la generación de los reportes era limitado a un número predefinido, cuando estos sistemas se sobrecargaban los usuarios esperaban por días o semanas para que los reportes sean completados con las diferentes pautas disponibles [12].

La década de 1980 vio el nacimiento del Datawarehouse, se utilizaron como repositorios centrales de datos integrados de una o más fuentes. El DW fue considerado como un componente central del BI [11]. Además, en 1989 Howard Dresner investigador de Gartner Group popularizo el termino BI como “Un término general que describe conceptos y métodos para mejorar la toma de decisiones comerciales mediante el uso de sistemas de soporte basado en hechos” [13].

El principal desarrollo de BI durante los años 90 fue la el desarrollo de herramientas de BI. Una de los más populares fue Enterprise Resource Planning (ERP). Las herramientas eran más fáciles de usar que las versiones anteriores, eran más eficientes y brindaban la funcionalidad que se necesitaba. Los empresarios ahora podían recopilar datos y obtener información al trabajar directamente con ellos.

La década de 2000, agregó más velocidad al desarrollo de BI y vio una concentración de BI en manos de IBM, Microsoft, SAP y Oracle. El análisis predictivo proporcionó un nuevo método de uso de datos, algoritmos y aprendizaje automático para pronosticar cambios futuros en el negocio. Las tecnologías en la nube pasan a primer plano a medida que las fuentes en tiempo real y las herramientas de visualización cambian la forma en que se ven los datos [11].

Actualmente se intenta que las herramientas y aplicaciones de BI sean lo más intuitivas posible, ya que genera conocimiento gracias a la gestión de datos, a través de reportes los gerentes pueden realizar una eficaz toma de decisiones.

#### **4.1.2 Definición**

Business Intelligence se define como el correcto proceso de todas las actividades relacionadas con la generación extracción, organización, análisis, computación y distribución del conocimiento en una organización; esto con el fin de mejorar su eficiencia, haciendo uso de las tecnologías más apropiadas [14].

Para [15], la inteligencia de negocios es el conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de una organización.

BI es un proceso interactivo para explorar y analizar información estructurada sobre un área, para descubrir tendencias o patrones, a partir de los cuales se puede 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 [16].

Partiendo de las definiciones extraídas de cada autor, se llega a la conclusión que BI es un conjunto de procesos que facilitan el tratamiento de datos en información, misma que sirve de apoyo a las organizaciones en la toma de decisiones, puesto que presenta a los usuarios información de manera ordenada, rápida y oportuna, permitiendo un mejor análisis gracias a indicadores claves que mejoran la gestión de la misma.

BI permite la correcta recolección de información, de manera que, combinando datos y análisis, se pueda obtener el conocimiento necesario para tomar estrategias de mejora continua, mismas que encaminen al cumplimiento de los objetivos empresariales.

#### **4.1.3 Características**

El Business Intelligence posee varios puntos de distinción que facilitan la toma de decisiones, de acuerdo con [17], algunos elementos que caracterizan al contexto de Business Intelligence:

- Integración y globalización de información: Todos los departamentos de la organización tienen la capacidad de brindar y acceder a cualquier tipo de información.
- Mejora el acceso a información: Brinda información de calidad, completa, correcta y coherente.
- Facilita la toma de decisiones: Gracias a las herramientas de visualización como gráficas, tablas y barras se logra tener una mejor visión general sobre la situación actual de la empresa.
- Orientación al usuario final: Debido al manejo de interfaces amigables e intuitivas, permite al usuario el fácil manejo de los sistemas BI, sin la necesidad de poseer conocimiento técnico para su uso.

#### **4.1.4 Beneficios de BI**

La implementación de sistemas BI trae consigo varios beneficios, según [18], [19] destacan:

- Transforma datos orientados a las aplicaciones en información orientada a la toma de decisiones.
- Integra y consolida diferentes fuentes de datos (internas y/o externas) y departamentos empresariales

- Aumenta la competitividad en el mercado.
- Brinda información completa, correcta, consistente, oportuna y accesible.
- Aumento de la eficiencia en el proceso de toma de decisiones.
- Apoya en la identificación de riesgos y oportunidades
- Optimización del tiempo

#### 4.1.6 Factores para la implementación de BI

En [20] afirma que para lograr el éxito en proyectos BI en una organización, los sistemas de información deben satisfacer las necesidades del negocio. Un modelo de negocio sólido debe incorporar todas las propuestas estratégicas que abordará el proyecto, los recursos necesarios, los posibles riesgos, los costes a asumir y los plazos de ejecución del proyecto. Así, el modelo proporcionará facilidad en la adopción de una nueva herramienta para la gestión de información. A continuación, en la TABLA I se presenta los criterios a tener en cuenta según [21], para asegurar el éxito en la construcción de proyectos BI.

TABLA I. FACTORES PARA LA IMPLEMENTACIÓN DE BI.

<b>Características</b>	
<b>Gerencia</b>	<ul style="list-style-type: none"> <li>- Apoyo y compromiso de gerencia</li> <li>- Participación activa durante todo el proceso</li> <li>- Equipo directivo bien cualificado</li> <li>- Proyecto que satisface las necesidades de gerencia</li> </ul>
<b>Vinculación empresarial</b>	<ul style="list-style-type: none"> <li>- Tener bien definido el modelo de negocio, procesos y problemas de la empresa, así como los requisitos de información</li> <li>- Visión estratégica del BI vinculada a las iniciativas de la empresa</li> <li>- Identificar los indicadores clave de rendimiento (KPI)</li> <li>- Establecer métricas y clasificaciones manejadas por la parte de negocio</li> <li>- Administrar la información manejada por el negocio</li> <li>- Formular una metodología y una gestión de proyectos manejada por la parte empresarial</li> <li>- Disponer de un marco teórico y actualizable gestionado por parte de la empresa</li> </ul>
<b>Encargado del proyecto</b>	<ul style="list-style-type: none"> <li>- Nivel avanzado en conocimientos empresariales</li> <li>- Gerente del proyecto</li> </ul>
<b>Estrategia</b>	<ul style="list-style-type: none"> <li>- Misión y visión bien definidas</li> <li>- Visión estratégica del proyecto de BI</li> <li>- Plan de negocio claro</li> </ul>
<b>Gestión del cambio</b>	<ul style="list-style-type: none"> <li>- Gestión del cambio adecuada y eficaz</li> <li>- Gestión del cambio orientada al usuario</li> </ul>
<b>Proyecto</b>	<ul style="list-style-type: none"> <li>- Planificación del proyecto</li> <li>- Definir y gestionar el alcance del proyecto</li> <li>- Gestión eficaz del proyecto</li> <li>- Diseño de soluciones basado en el usuario final</li> <li>- Clara relación entre los objetivos empresariales</li> <li>- Metodología del proyecto</li> <li>- Director de proyecto de BI competente</li> <li>- Desarrollo iterativo manejado por la parte empresarial</li> </ul>
<b>Recursos Humanos</b>	<ul style="list-style-type: none"> <li>- Apoyo de un experto en la fase inicial</li> <li>- Participación del usuario final durante el ciclo de vida del proyecto.</li> <li>- Equipo de competencias mixtas adecuado</li> <li>- Expectativas del usuario bien definidas</li> </ul>

<b>Conocimiento y habilidades</b>	<ul style="list-style-type: none"> <li>- Educación y aprendizaje adecuado y formal del usuario</li> <li>- Soluciones de aprendizaje sencillas</li> <li>- Conocimiento y habilidades del equipo</li> </ul>
<b>Información y tecnología</b>	<ul style="list-style-type: none"> <li>- Tecnología y herramientas adecuadas</li> <li>- Establecer un marco de trabajo, extensible y actualizable</li> <li>- Alta calidad de los datos y fuentes seguras</li> <li>- Interacción con otros sistemas</li> <li>- Seguridad de los datos</li> <li>- Gestión eficaz de los datos</li> <li>- Sistemas de datos de origen</li> <li>- Uso amigable del sistema de BI</li> <li>- Sostenibilidad del hardware y el software</li> <li>- Confianza, capacidad de actualización y flexibilidad del sistema</li> <li>- Tecnologías amigables orientadas al usuario</li> <li>- Modelo dimensional de datos y metadatos</li> <li>- Soporte de sistemas interactivos</li> <li>- Informes puntuales</li> </ul>
<b>Recursos</b>	<ul style="list-style-type: none"> <li>- Recursos tecnológicos adecuados</li> <li>- Presupuesto adecuado</li> <li>- Recursos humanos y financieros estratégicos</li> </ul>
<b>Métricas</b>	<ul style="list-style-type: none"> <li>- Uso actual del sistema</li> <li>- Utilidad percibida del sistema</li> <li>- Beneficios netos obtenidos</li> <li>- Satisfacción del usuario</li> <li>- Calidad del sistema</li> <li>- Calidad de la información</li> </ul>
<b>Entorno</b>	<ul style="list-style-type: none"> <li>- Cultura organizativa</li> <li>- Resolución de problemas no técnicos</li> </ul>

## 4.2 Componentes de BI

En la Figura 2 tomada de [16], se ilustra los componentes necesarios para la implementación de BI.

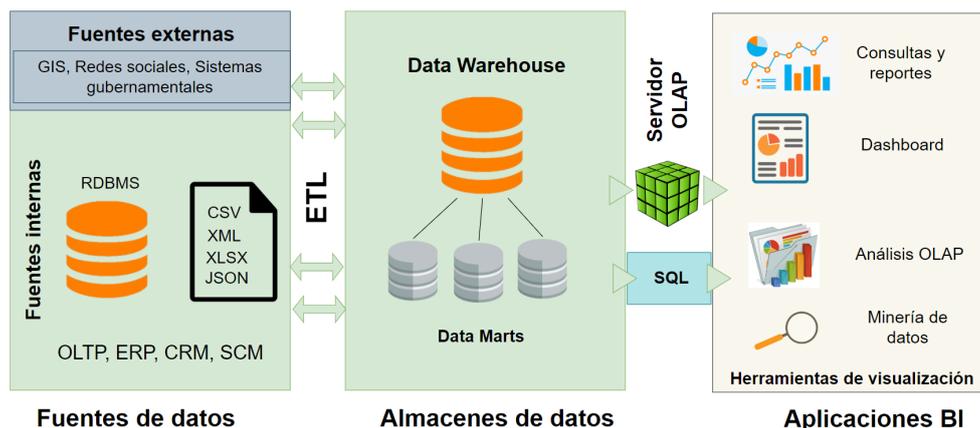


Figura 2. Componentes de Business Intelligence.

A continuación, se describen los componentes de BI:

- **Fuentes de datos:** Sirven de alimento al DW, se las obtiene a partir de distintas fuentes. (ver en el apartado Fuentes de Información)

- **Data Warehouse:** Se encarga de almacenar los datos de una forma que maximice su flexibilidad, facilidad de acceso y administración. (ver en el apartado 4.2.2 Data Warehouse)
- **Proceso ETL:** Para almacenar los datos en el DW primeramente deben pasar por procesos de transformación, limpieza y filtrado, con el fin de cargar los datos con una estructura estandarizada. (ver en el apartado 4.2.4 Procesos ETL)
- **Servidor OLAP:** Brinda capacidad de cálculo, consultas, funciones de planeamiento, pronóstico y análisis de grandes volúmenes de datos. (ver en el apartado 4.2.5 Proceso analítico en línea OLAP)
- **Herramientas de visualización:** Proporcionan facilidad para la visualización de información, a través de tablas, graficas, mapas, etc. (ver en el apartado 4.2.6 Herramientas de visualización)

#### 4.2.1 Fuentes de Información

La carga de información en un DW es un proceso complejo que se ve influenciado por múltiples factores. Entre ellos, destaca la cantidad de fuentes de información que se utilizan para la carga. El acceso a diferentes bases de datos requiere habilidades y conocimientos específicos en diversas sintaxis de SQL, lo que agrega un grado de dificultad adicional a este proceso.

Las fuentes de información a las que se puede acceder son:

- Sistemas operacionales o transaccionales, que incluyen aplicaciones desarrolladas a medida, ERP, CRM, SCM, etc.
- Sistemas de información departamentales: previsiones, presupuestos, hojas de cálculo, etc.
- Fuentes de información externa, en algunos casos comprada a terceros, como por ejemplo estudios de mercado. Las fuentes de información externas son fundamentales para enriquecer la información sobre clientes. En algunos casos es interesante incorporar información referente, por ejemplo, a población, número de habitantes, etc.

[16]

En general, la información que se carga en un DW suele ser de tipo estructurada, lo que significa que puede ser almacenada en forma de tablas. En su mayoría, se trata de datos numéricos que pueden ser analizados y procesados de manera eficiente.

#### 4.2.2 Data Warehouse

Un data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización con las propiedades siguientes: estable, coherente,

fiable y con información histórica. Al abarcar un ámbito global de la organización y con un amplio alcance histórico, el volumen de datos puede llegar a ser muy grande [15].

De acuerdo con Bill Inmon en [22], un Datawarehouse “es una colección de datos orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”. En la Figura 3, se ilustran las características del DW.



Figura 3. Características del Datawarehouse.

A continuación, se describen las características del DW [18]:

- **Orientado al negocio:** La información está clasificada de acuerdo a los aspectos que son de interés para la empresa, la clasificación que se realiza afecta al diseño y la implementación de los datos encontrados, puesto que la estructura difiere de los procesos operacionales orientados a las aplicaciones.
- **Integrado:** Las fuentes de datos producidas por distintos departamentos y aplicaciones ya sean internas o externas, deben ser estructuradas antes que se agreguen al DW.
- **Variante en el tiempo:** En una Data Warehouse se puede manejar grandes volúmenes de datos por esta razón los tiempos de ejecución de las consultas pueden aumentar, en cambio en un ambiente operacional la información se obtiene inmediatamente al momento del acceso.
- **No volátil:** En el ambiente operacional del Data Warehouse existe la carga y el acceso a los datos, por esta razón no cambian, se requiere mecanismos de control de recuperación o concurrencia.

### Estructura

El DW posee la capacidad de estructurar los datos en distintos niveles de esquematización y detalle. Ricardo Bernabéu en [18], presenta la estructura del DW, las cuales se pueden observar en la Figura 4.

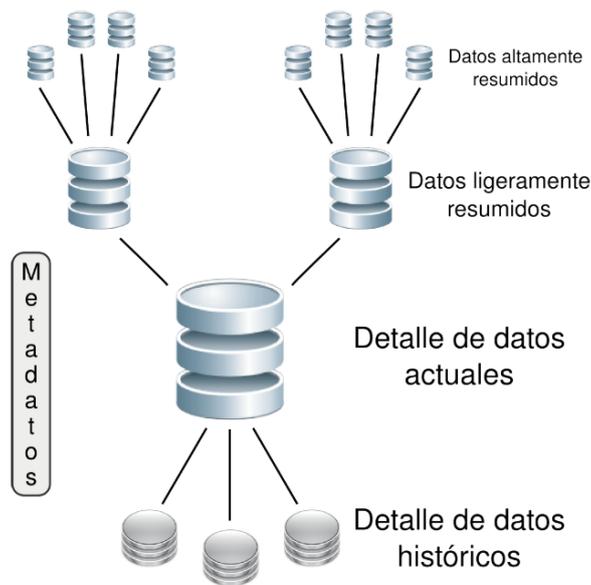


Figura 4. Estructura del Datawarehouse.

La estructura del DW se describe a continuación [17],[18]:

- **Detalle de datos actuales**, son datos de gran tamaño, se almacenan a nivel de detalle, su administración es compleja principalmente porque se necesita habilitar el acceso a la información de manera rápida y sencilla, además reflejan los eventos recientes.
- **Detalle de datos históricos**, representan datos antiguos que no son requeridos con mucha periodicidad, se almacenan a nivel de detalle normalmente en alguna forma de almacenamiento externo, la información que contienen es antigua para consultas esporádicas.
- **Datos ligeramente resumidos**, ofrecen fácil acceso, son los datos ya resumidos y tienen un bajo nivel de detalle, agrupan los datos bajo las condiciones establecidas en el análisis.
- **Datos altamente resumidos**, ofrecen fácil acceso, son los datos ya resumidos y los resumen aún más.
- **Metadatos**, representan la información acerca de los datos. De muchas maneras se sitúa en una dimensión diferente al de otros datos, ya que su contenido no es tomado directamente desde el ambiente operacional.

## Flujo de Datos

El flujo de datos en el DW posee un flujo de datos estándar y generalizado, se puede observar con mayor claridad en la Figura 5.

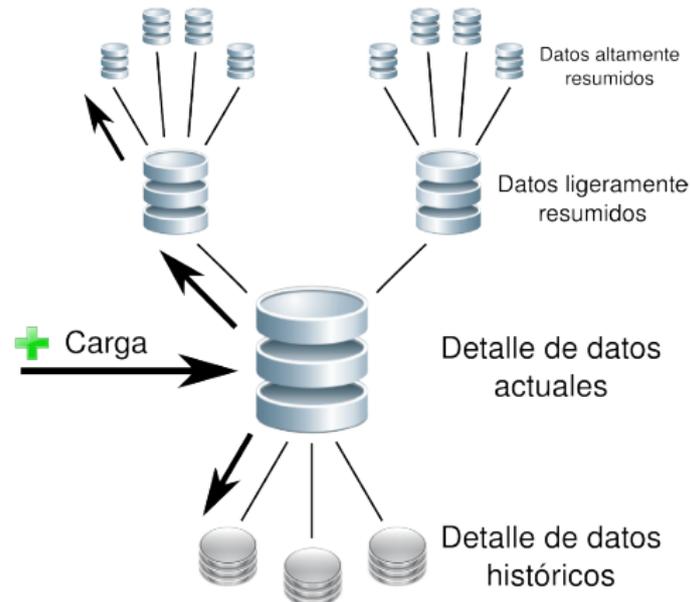


Figura 5. Flujo de datos en el DW

Cuando la información ingresa al DW se almacena a nivel de detalle de datos actuales. Los datos permanecerán allí hasta que ocurra alguno de los siguientes eventos [18]:

- Sean borrados del depósito de datos.
- Sean resumidos, ya sea a nivel de datos ligeramente resumidos o a nivel de datos altamente resumidos.
- Sean archivados a nivel de detalle de datos históricos.

### 4.2.3 Data mart

Un Data Mart es la implementación de un DW enfocado específicamente a un área funcional, problema en particular, departamento, tema o grupo de necesidades [18].

De acuerdo a las necesidades del proyecto, según [17] y [18], el data mart pueden adoptar las siguientes arquitecturas:

- **Top-Down:** Este tipo de arquitectura establece que primero se debe establecer el DW y luego se realiza el desarrollo, construcción y carga de los data mart. En la Figura 6 elaborada en base a [18], se ilustra la arquitectura especificada.

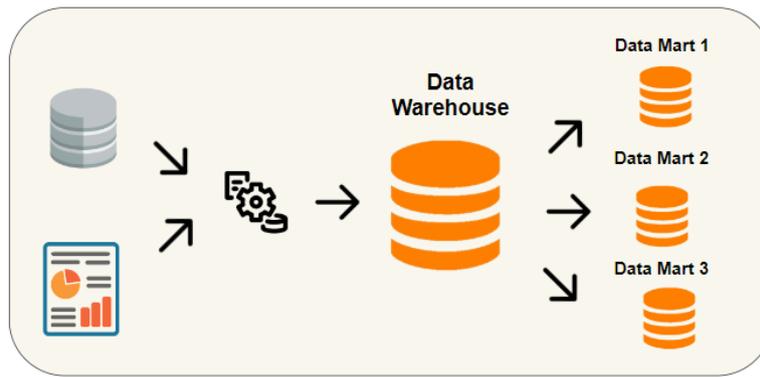


Figura 6. Arquitectura Top-Down.

- **Bottom-Up:** Este tipo de arquitectura establece que se deben definir previamente los data mart para luego ser integrados en un DW centralizado. En la Figura 7 elaborada en base a [18], se ilustra la descripción de la arquitectura.

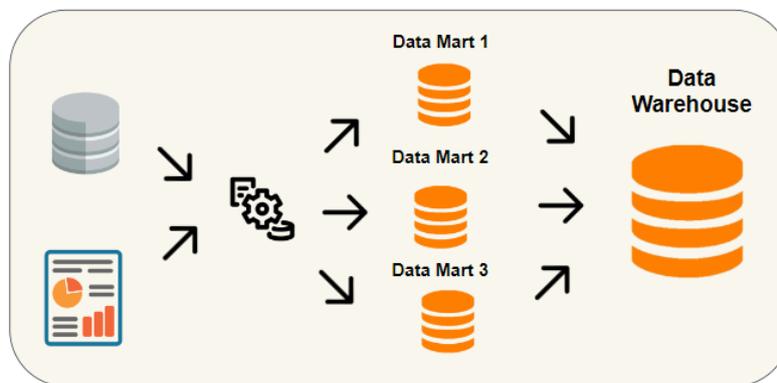


Figura 7. Arquitectura Botton-Up.

#### 4.2.4 Procesos ETL



Figura 8. Procesos ETL

Los procesos de extracción, transformación y carga (Extract-Transform-Load) consisten en actividades técnicas en el desarrollo de sistemas de Business Intelligence [23]. Su principal objetivo consiste en extraer, limpiar, transformar, resumir, formatear y guardar los datos que se almacenaran en el DW (ver Figura 8). Estos procesos forman parte del componente de

integración de datos, su correcta implementación se ve reflejada en la integridad, uniformidad, consistencia y disponibilidad que presentan los datos utilizados en la solución BI [24].

La construcción del ETL consta en tres subprocesos: componente de extracción, componente de transformación y componente de carga [25]. En la TABLA II, se presenta la descripción de cada uno de estos componentes identificando los elementos objetivo, las operaciones realizadas, y los resultados esperados.

TABLA II. DESCRIPCION DE LOS COMPONENTES DEL PROCESO ETL

Componente	Elementos Objetivos (entrada)	Operaciones realizadas (proceso)	Resultados de la tarea (salida)
<b>Extracción</b>	Fuente de datos, sistemas transaccionales, hojas de cálculo, archivos de texto	Selección	Datos crudos (en memoria)
<b>Transformación</b>	Datos crudos (en memoria)	Limpieza, transformación, personalización, cálculos y aplicación de funciones de agregación	Datos formateados, estructurados y resumidos de acuerdo a las necesidades. (en memoria)
<b>Carga</b>	Datos formateados, estructurados y resumidos de acuerdo a las necesidades (en memoria)	Inserción	Datos formateados, estructurados y resumidos con persistencia en el DW

A continuación, se presenta la descripción de cada uno de estos:

- **Componente de Extracción:** En esta etapa se lleva a cabo la captura de información desde los sistemas fuente de datos, bases de datos de sistemas transaccionales, hojas de cálculo, archivos de texto, servicios web, entre otros.
- **Componente de Transformación:** Se realizan actividades de limpieza, transformación, personalización, realización de cálculos y aplicación de funciones de agregación con el objetivo de obtener datos formateados, estructurados, y resumidos de acuerdo a los requerimientos.
- **Componente de Carga:** Finalmente se realiza la carga de los datos formateados, estructurados, y resumidos en el Data Warehouse, estos datos servirán de base para la elaboración de cubos OLAP y otras herramientas de consulta de información [25].

### Técnicas para modelar el proceso de ETL

El modelado del proceso de ETL, al igual que el de cualquier objeto computacional, puede representarse utilizando tres niveles de abstracción: conceptual, lógico y físico. En la TABLA

III [25], se puede apreciar una breve descripción y comparación de los mismos.

TABLA III. TIPOS DE MODELADO ETL

Tipos	Niveles de detalle	Objeto y conceptos
Conceptual	Bajo	Fuentes, atributos, transformaciones y estructura de destino.
Lógico	Medio	Tablas de origen, dimensiones, tablas de hechos, Atributos, operaciones (aritméticas, lógicas y etc.)
Físico	Alto	Tablas de origen, dimensiones, tablas de hechos Atributos, tipos de datos, precisión, restricciones, índices, entre muchas más.

### Herramientas ETL

Las herramientas ETL son esenciales en los proyectos de Business Intelligence debido a la importancia crítica en el proceso de extracción, transformación y carga [16].

Estas herramientas son de gran importancia para la gestión de datos en muchas organizaciones. Se utilizan para integrar, limpiar y transformar datos de diferentes fuentes en un formato común para su análisis y uso en aplicaciones empresariales. En la Figura 9, se muestran las herramientas ETL más utilizadas de código abierto y de paga. La herramienta utilizada para el desarrollo del proyecto es Pentaho Data Integration, que es una suite muy popular y que cuenta con herramientas ETL que permite la integración de datos de diferentes fuentes y su transformación para su uso en aplicaciones empresariales.



Figura 9. Herramientas ETL

#### 4.2.5 Proceso analítico en línea OLAP

*Online Analytical Processing*, se define como una tecnología que permite realizar el análisis de información en línea, la administración y ejecución de consultas de información en una empresa para la posterior toma de decisiones, estas herramientas brindan respuestas

inmediatas a preguntas complejas, utilizando la información histórica para entender la evolución de la información en el tiempo.

El proceso analítico en línea según [18], se encuentra caracterizado por:

- Disponer información organizada y resumida, en formato de tablas, gráficos y reportes, que ayuda a mantener la información organizada y disponible siempre para los usuarios.
- Facilitar la apreciación de un gran volumen de datos, a través de un simple análisis.
- Organizar la estructura de datos definiendo dimensiones, de esta forma se pueden analizar mediante jerarquías y combinaciones.
- Los cubos OLAP se generan mediante esquemas definidos en el DW, las dimensiones definen dominios, los miembros de una dimensión se agrupan en forma jerárquica, y cada celda del cubo contiene datos que relacionan elementos de las dimensiones.

#### **4.2.6 Herramientas de visualización**

Las herramientas de visualización de datos son aplicaciones y plataformas que permiten a los usuarios representar y explorar datos de manera visual para facilitar la comprensión de patrones, tendencias y relaciones en los datos [16]. Estas herramientas pueden incluir gráficos, tablas, diagramas, mapas y otras formas de visualización interactiva que permiten a los usuarios interactuar con los datos y explorarlos desde diferentes perspectivas.

Las herramientas de visualización son [26]:

- Generadores de informes: Los desarrolladores profesionales suelen emplearlos para elaborar informes estandarizados destinados a grupos, departamentos o a la organización.
- Herramientas de usuario final de consultas e informes: son empleadas por usuarios finales en la elaboración de informes destinados a ellos mismos o a terceros, no requieren programación.
- Herramientas OLAP: permiten manipular la información de forma multidimensional para explorarla desde distintas perspectivas y en diferentes periodos de tiempo.
- Herramientas de Dashboard: permiten a los usuarios finales tener una visión rápida y clara de la información crítica mediante el uso de iconos, tablas y gráficos intuitivos y fáciles de entender.
- Herramientas de planificación, modelización y consolidación: permite a la gerencia de las organizaciones crear planes de negocio y simulaciones con la información generada con BI.

- Herramientas datamining: Permiten crear modelos estadísticos que permiten descubrir e interpretar patrones desconocidos en la información mediante los cuales se pueden resolver problemas de negocio.

### **4.3 Modelado Dimensional**

El modelado dimensional es una técnica de diseño que presenta los datos en forma de modelo de relación entre entidades [22]. Cada tabla de dimensión tiene una clave primaria simple [27]. Todo modelo dimensional se compone de una tabla de hechos, y un conjunto de tablas más pequeñas llamadas tablas de dimensiones [18].

#### **Tablas de hechos**

Son tablas que contienen los hechos que serán utilizados por los analistas de negocio para apoyar el proceso de toma de decisiones. Los datos presentes en las tablas de hechos constituyen el volumen de la bodega, y pueden estar compuestos por millones de registros dependiendo de su granularidad y antigüedad de la organización. Los más importantes son los de tipo numérico. Están conformadas por dimensiones y medidas [18].

#### **Dimensiones**

Las tablas de dimensiones definen como están los datos organizados lógicamente y proveen el medio para analizar el contexto del negocio. Contienen datos cualitativos. Representan los aspectos de interés, mediante los cuales los usuarios podrán filtrar y manipular la información almacenada en la tabla de hechos [18].

Las tablas de dimensiones están asociadas a una tabla de hechos. Cada dimensión está definida por una clave primaria que sirve como base de referencia con la tabla de hechos. Las tablas de dimensión contienen atributos que se encuentran ordenados por jerarquías [23].

#### **Esquemas**

Existen dos técnicas que se utilizan en el modelado: esquema de estrella y copo de nieve: estrella y copo de nieve.

#### **Esquema en estrella**

Este esquema consiste en estructurar la información en procesos, vistas y métricas recordando así la forma de una estrella. El esquema en estrella, consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a esta, a través de sus respectivas claves. En la Figura 10, se puede apreciar un esquema en estrella.

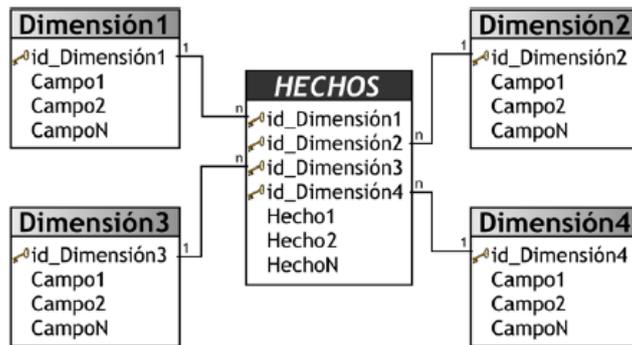


Figura 10. Esquema en estrella.

El esquema en estrella destaca por su velocidad y simplicidad para realizar análisis multidimensionales, además facilita las consultas al usuario final. Gracias a su gran rendimiento permite añadir dimensiones de manera individual sin alterar el rendimiento de la base de datos [18].

### Esquema Copo de Nieve

Este esquema permite que las tablas de dimensiones se encuentren normalizadas lo que permite simplificar las operaciones de selección de datos, resultando en información sin redundancia. Se representa como una tabla de hechos conectada con dimensiones anidadas. En la Figura 11, se puede apreciar un esquema de copo de nieve.

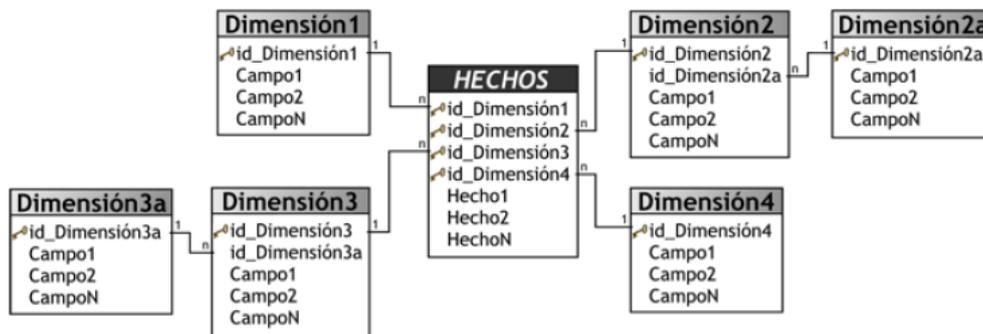


Figura 11. Esquema copo de nieve.

La utilización de este tipo de esquema permite una mejor segregación de datos en las dimensiones, pero otorga mayor complejidad a su estructura, esto hace que los tiempos de respuesta en los análisis multidimensionales sean mayores [18].

### 4.4 Herramientas para el desarrollo de BI

Las herramientas de BI son una serie de tecnologías y aplicaciones diseñadas para ayudar a las empresas a recopilar, analizar y presentar datos de manera efectiva. Estas herramientas incluyen software de análisis de datos, visualización de datos, informes y cuadros de mando, entre otros. Al utilizar estas herramientas, las empresas pueden analizar datos en tiempo real,

identificar tendencias y patrones, y tomar decisiones basadas en datos en lugar de en suposiciones. En la Figura 12Figura 9, se muestran las herramientas BI más utilizadas.



Figura 12. Herramientas de BI

Para el desarrollo del proyecto, se utilizará Pentaho.

#### 4.4.1 Pentaho

Pentaho Business Intelligence es una plataforma de tipo Open Source [14], que se encuentra orientada a la solución y centrada en procesos, se basa en cubrir amplias necesidades de análisis de datos y reportes empresariales, brindando apoyo al proceso de toma de decisiones [17]. La suite de Pentaho esta desarrollada en Java, lo que permite que sea flexible, y cubra una amplia gama de necesidades.

La plataforma fue desarrollada para mejorar la gestión y la toma de decisiones empresariales, está compuesta por diferentes programas que satisfacen los requisitos de BI, ofreciendo soluciones para la gestión y análisis de información, incluyendo procesos ETL, análisis OLAP, generación de informes, minería de datos, y creación de cuadro de mandos [28].

#### 4.6.1 Características

Las características que presenta la plataforma Pentaho BI son [17],[28]:

- Orientada a la solución y centrada en procesos
- Permite realizar análisis de datos para generar información rápida y oportuna a través de informes empresariales.
- La plataforma Pentaho esta desarrollada en Java, al igual que su ambiente de implementación.
- Maneja servicios críticos tales como programación, seguridad, integración,

automatización y flujo de trabajo.

- Es multiplataforma, tanto a nivel de servidor como de cliente.
- Esta construido bajo un servidor de aplicaciones compatibles con J2EE como JBOSS AS, WebSphere, Tomcat, WebLogic y Oracle AS, lo cual asegura la escalabilidad, integración y portabilidad.
- Se puede trabajar con diferentes bases de datos: JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird.
- Lenguajes de programación: Java, JavaScript, JSP, XSL.
- Interfaz de desarrollo: Java SWT, Eclipse, Web-based.
- Repositorio de datos basado en XML.
- Sus componentes están expuestos vía Web Services lo que apoya en la integración con Arquitecturas Orientadas a Servicios.

#### 4.6.2 Módulos de Pentaho

Pentaho BI dispone de módulos de diseño desarrollados para perfeccionar el flujo de modelado, transformación, almacenamiento y visualización de datos. A continuación, se detalla cada módulo [29]:

- **Data Integration:** este módulo a través de Kettle permite llevar a cabo procesos de extracción, transformación y carga.
- **Schema Workbench:** representa la interfaz visual para crear y probar cubos OLAP en Mondrian. Se pueden exportar los cubos creados directamente a Pentaho.
- **Business Analytics Platform:** es el servidor que proporciona varios servicios a los usuarios, como informes y herramientas de integración.
- **Report Designer:** representa la herramienta grafica para la creación de reportes, además permite exportarlos directamente a Pentaho
- **Aggregation Designer:** permite agregar y mantener tablas agregadas.
- **Metadata Editor:** dispone de un editor de metadatos útil para personalizar la capa de metadatos modelando la estructura de la base de datos para generar un modelo lógico de metadatos.

#### 4.6.3 Licencias de Pentaho

La plataforma Pentaho BI está disponible en dos versiones la Enterprise Edition, que es la versión comercial y la Community Edition, versión de código abierto. En la TABLA IV tomada de [30], se detallan las características que propias de cada versión.

TABLA IV. COMPARACIÓN DE CARACTERÍSTICAS PENTAHO  
COMMUNITY & ENTERPRISE EDITION.

<b>Software y servicios</b>	<b>Edición Community</b>	<b>Enterprise Edition</b>
Informes	Fuente abierta	Certificado
Análisis	Fuente abierta	Certificado
Cuadros de mando	Fuente abierta	Certificado
Integración de datos /ETL	Fuente abierta	Certificado
Plataforma Business Intelligence	Fuente abierta	Certificado
Minería de datos	Fuente abierta	Certificado
Interacción con foros de la comunidad	Fuente abierta	Certificado
Documentación Web de la comunidad	✓	✓
<b>Soporte profesional</b>		
Soporte telefónico		✓
Soporte por email		✓
Casos de soporte sin límites		✓
Acuerdo de nivel de servicio		✓
<b>Mantenimiento de software</b>		
Mantenimiento de software	Personal interno	Ingeniero Pentaho
Parches de versiones		✓
Arreglos incluidos en versiones posteriores.		✓
<b>Funcionalidad mejorada</b>		
Consola empresarial de Pentaho.		✓
Inicio de sesión única.		✓
Configuración de seguridad simplificada		✓
Diagnóstico de la aplicación.		✓
Utilidades del repositorio.		✓
Gestión del ciclo de vida.		✓
Reportes de auditoría.		✓
Vencimiento automatizado de contenido.		✓
Clustering.		✓
Supervisión del rendimiento.		✓
Gestión y monitoreo del ETL.		✓
<b>Software certificado</b>		
Software estabilizado.		✓
Ciclo de liberación administrado.		✓
Construcción optimizada.		✓
Documentación profesional.		✓
Base de conocimiento.		✓
Apoyo consultativo.		✓
Paquetes de asistencia remota.		✓
Paquetes de asistencia remota.		✓
Diseño e integración de paquetes.		✓
Paquetes de resolución de problemas y optimización.		✓
Foro en línea de la edición empresarial:		✓
Entrenamiento basado en la web.		✓
<b>Garantía de software</b>		
Indemnización por propiedad intelectual.		✓
Garantía de servicios.		✓

## 4.6 Metodología del ciclo de vida de Ralph Kimball

La metodología de Kimball fue concebida a mediados de la década de los 80's por miembros del Grupo Kimball y de Metaphor Computer Systems, una empresa pionera en apoyo a la toma de decisiones. Desde entonces, ha sido utilizado en proyectos de DW/ Bi en todas las industrias [31].

La metodología está basada en un Ciclo de Vida Dimensional del Negocio compuesto por 4 principios [32]:

- **Centrarse en el negocio:** se debe centrar en la identificación de los requerimientos del negocio, y usar estos para desarrollar soluciones sólidas que permitan cumplir con los objetivos de la empresa.
- **Construir una infraestructura de información adecuada:** diseñar una base de datos de alto rendimiento donde se reflejará los requerimientos de negocio identificados en la empresa.
- **Realizar entregas en incrementos significativos:** crear el DW en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos.
- **Ofrecer la solución completa:** proporcionar todos los elementos necesarios que brinden soluciones de valor a los usuarios del negocio. Desde un Datawarehouse bien diseñado, hasta la generación de aplicaciones para informes y análisis avanzado.

### 4.6.1 Arquitectura Kimball

Kimball propone una arquitectura para proyectos de BI en la cual se enfatiza que es necesario aprender el significado estratégico de los componentes de su arquitectura, que son: sistemas fuente de datos operacionales, el sistema de extracción transformación y carga (ETL), el área de presentación y las aplicaciones de Business Intelligence.

A continuación se presenta la descripción de cada componente [33]:

- **Sistemas operacionales fuente:** son los que capturan las transacciones de la organización, se tiene poco o nada de control sobre el contenido y el formato de los datos manejados aquí. Su prioridad es el rendimiento en procesamiento y disponibilidad.
- **Sistema ETL:** es un área de trabajo donde se inicializa la estructura de datos y una serie de procesos. El primer paso es la extracción, que consiste como lo describe la extracción de datos desde los sistemas de operacionales, a continuación, se realiza la transformación que incluye limpieza, combinación, estandarización y depuración de

datos, el paso final consiste en la estructuración y carga de la información en el área de presentación.

- Área de presentación: aquí la información es organizada y almacenada para el acceso y consulta directa de usuarios, configuradores de reportes y otras aplicaciones analíticas de BI.
- Aplicaciones BI: se refieren a la gran cantidad de capacidades proporcionadas a los usuarios de la organización para el uso de la información disponible en el área de presentación para la toma de decisiones analíticas.

#### 4.6.2 Metodología de Ralph Kimball

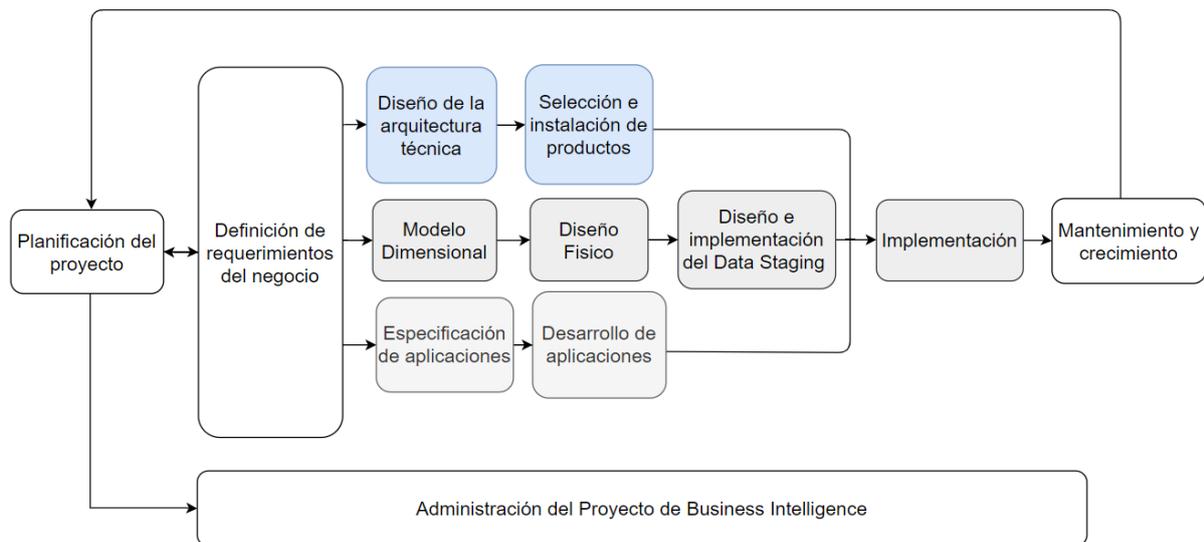


Figura 13. Metodología ciclo de vida dimensional del negocio.

Los pasos de la metodología Kimball se describe a continuación según [23], [32],[34], [35]:

- **Planificación del proyecto**

Se establece el propósito, los objetivos y el alcance del Data Warehouse - BI. Además, en esta etapa se identifica, planifica y designa las tareas y recursos necesarios para la ejecución del proyecto.

- **Definición de requisitos**

En esta fase se reúnen los requisitos de la empresa mediante técnicas de elicitación como entrevistas, encuestas y/o cuestionarios. Este procedimiento permitirá comprender el problema que se quiere resolver con la solución BI.

- **Diseño de la arquitectura**

El diseño arquitectónico sirve como marco organizativo para la integración de las herramientas a utilizarse. De acuerdo con lo planteado por Kimball, el diseño de la arquitectura se puede realizar mediante un proceso de ocho pasos:

- Establecer un grupo de trabajo de arquitectura
- Recopilar requisitos relacionados con la arquitectura
- Requisitos de Arquitectura de Documentos
- Desarrollar un modelo arquitectónico de alto nivel
- Diseñar y especificar los subsistemas
- Determinar la arquitectura- Fases de implementación
- Documentar la arquitectura técnica.
- Revisar y finalizar la Arquitectura técnica

- **Selección e Instalación de Productos**

De acuerdo al diseño de la arquitectura establecida, en esta fase se seleccionan e instalan las herramientas que mejor se adaptan para el desarrollo del proyecto.

- **Modelado Dimensional**

En esta etapa se desarrolla un proceso dinámico e iterativo, el modelado dimensional consta de varias fases desde la preparación de procesos hasta la iteración final de diseño. El proceso iterativo está definido en cuatro partes:

- Elegir el proceso de negocio
- Establecer el nivel de granularidad
- Definir las dimensiones.
- Identificar medidas y las tablas de hechos

- **Diseño Físico**

Los modelos dimensionales desarrollados y documentados necesitan ser traducidos a una base de datos física. Hay que cuidar que la base de datos no debe convertir el esquema dimensional en una estructura normalizada durante el proceso de diseño físico. Los detalles físicos de la implementación de la base de datos varían ampliamente según la plataforma y el proyecto.

Determinar el tamaño del DW y la estructura para soportar el diseño lógico. Se define los aspectos de configuración, hardware, software, el plan de indexación y las estrategias de partición en las tablas relacionales [36].

- **Diseño ETL**

Esta fase sirve como base en la implementación del Data Warehouse. El correcto diseño del sistema ETL permite extraer los datos de las diferentes fuentes origen de

datos, transformarlos para aumentar la calidad y consistencia de los mismos, y finalmente cargar la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis.

- **Especificación y desarrollo de aplicaciones BI**

Una parte fundamental del proyecto de DW/BI está en brindar a los usuarios una forma sencilla y estructurada, de acceder al almacén de datos. Este acceso estructurado se da a través de las aplicaciones de inteligencia de negocios, que proporcionan a los usuarios información de valor mediante informes y aplicaciones de análisis. Kimball divide a estas aplicaciones en dos categorías:

- Informes estándar: son informes simples, con parámetros de consulta fijo, que proporcionan al usuario información sobre un área específica de la empresa
- Aplicaciones analíticas: este tipo de aplicaciones son avanzadas, incluyen algoritmos y modelos de minería de datos, que ayudan a identificar oportunidades.

- **Implementación**

En la fase de implementación se busca llevar a cabo el enfoque realístico del proceso del sistema, ejecutándolo de modo que se tenga en cuenta cada detalle para el éxito de este. Se debe proporcionar los requisitos necesarios para que esto se lleve a cabo con eficacia, para tener mayor tiempo de enfoque en el almacenamiento de la data.

- **Mantenimiento**

Para administrar el entorno del Data warehouse es importante enfocarse en los usuarios de negocio, comunicarse con los mismos para establecer un flujo de retroalimentación. Es importante sentar las bases para el crecimiento y evolución de forma iterativa del DW.

#### **4.7 Trabajos Relacionados**

Según la bibliografía revisada, en la TABLA V, se presentan los trabajos relacionados con el objeto de estudio, estos trabajos fueron encontrados previo y durante la ejecución de la solución BI. Además, permiten sustentar el desarrollo del presente Trabajo de Titulación.

TABLA V. TRABAJOS RELACIONADOS

Titulo	Resumen	Referencias
Análisis, diseño e implementación de una solución Business Intelligence para la generación de indicadores y control de desempeño, utilizando la metodología Hefesto V2.0.	Realiza la implementación de Business Intelligence para la generación de indicadores y presentación de datos a través de reportes.	[17]
A Continuous Delivery Framework for Business Intelligence.	Presenta un marco de trabajo para la implementación de Business Intelligence.	[37]
Proposal of methodology for a Data Warehousing Process.	Propuesta de metodología para la implementación de procesos Data Warehousing, para generar indicadores de productividad.	[38]
The use of Business Intelligence tools for leadership and university administration.	Se utilizan consultas en cubos OLAP para el análisis de datos y mejorar la toma de decisiones.	[39]
Implementation of Business Intelligence for Sales Data Management Using Interactive Dashboard Visualization in XYZ Stores.	Implementación de un sistema de Business Intelligence para la administración de datos de ventas generados en las empresas, utilizando un Dashboard Interactivo.	[27]
Implementación de Business Intelligence como base para la gestión de información de la unidad educativa (SICS).	Se implementa un Dashboard con indicadores y reportes a partir de las consultas obtenidas en el cubo dimensional.	[34]
Open-Source Business Intelligence on a SME: a Case Study using Pentaho	Presenta el caso de estudio en una PYME, utilizando la herramienta de Open Source Pentaho, para la generación de un Dashboard que apoye en la toma de decisiones.	[29]
Analytical Data Mart for the Monitoring of University Accreditation Indicators	Se propone un Datamart analítico para facilitar la comprensión de los datos y apoyar la toma de decisiones a la administración.	[40]

## 5. Metodología

Para lograr cumplir los objetivos planteados en el Trabajo de Titulación, se aplicaron un conjunto de métodos y técnicas, las cuales permitieron llevar a cabo de forma adecuada el desarrollo del proyecto. La aplicación de la metodología de Ralph Kimball y técnicas como la entrevista fueron importantes para el éxito del proyecto.

La utilización de recursos de Hardware, Software, además de los recursos científicos, permitieron que el proyecto se desarrollara sin ningún contratiempo.

### 5.1 Tipo de Investigación

De acuerdo al enfoque del TT, se optó por una investigación mixta, es decir cualitativa-cuantitativa, debido a la necesidad de recopilar y analizar datos numéricos, no numéricos y descriptivos para implementar la solución BI y evaluar su impacto en la gestión de los datos en la empresa.

Además, fue fundamental contar con la perspectiva del gerente de la empresa para obtener información valiosa sobre los procesos internos y las necesidades de la organización. Para ello, se utilizó la técnica de la entrevista con el fin de obtener información detallada y relevante sobre la situación actual de la empresa y cómo la implementación de BI puede mejorar la gestión de datos de pedidos.

### 5.2 Contexto

El TT se realizó para “Marketing Digital Bravo & Beltrán”, empresa ubicada en las instalaciones del Centro Comercial Don Daniel de la ciudad de Loja (ver [mapa](#)<sup>1</sup>), enfocada en brindar soluciones eficientes en cuanto a tecnología y software. Donde se desarrolló una solución de Business Intelligence por medio de un Dashboard en el que se puede visualizar información sobre los pedidos generados durante los años 2018-2021 de manera más eficiente, y que brinda soporte a los encargados de la toma de decisiones.

En el desarrollo de la solución participo el Ing. Diego Bravo, gerente de la empresa quien colaboro con información relevante para el desarrollo del proyecto, además de la colaboración del Ing. Rene Guamán, como director del TT y del estudiante Luis Quito.

### 5.3 Proceso

Para el cumplimiento del objetivo general del proyecto propuesto, se realizó el siguiente proceso para cada uno de los objetivos específicos.

---

<sup>1</sup> <https://goo.gl/maps/LEotDKGqMZMypZK27>

### 1) Analizar los datos históricos generados en el módulo de pedidos para establecer el diseño del modelo.

En el primer objetivo se realizó un estudio de la empresa, además se ejecutaron las fases 2,3 y 4 de la metodología Kimball detallada en el marco teórico sección 4.6.2 . Por lo tanto, las etapas desarrolladas son:

- **Análisis Empresarial:** Se llevaron a cabo dos técnicas de investigación: la investigación en línea y una entrevista al gerente de la empresa. Estas técnicas se utilizaron con el objetivo de obtener una visión clara y precisa de la empresa.
- **Especificación de requerimientos:** Se definieron las generalidades del proyecto, además de los requisitos tanto funcionales como no funcionales, basado en las especificaciones del estándar IEEE 830.
- **Diseño de la arquitectura:** Se diseñó la arquitectura de la aplicación. El Back Room en la gestión de datos, y el Front Room para la entrega de información al usuario final.
- **Selección e instalación de herramientas:** Se realizó la instalación de las herramientas necesarias para la implementación de la solución BI.

### 2) Desarrollar el modelo de datos utilizando la plataforma de BI Pentaho.

Para cumplir este objetivo se enfocó en el diseño y desarrollo del Datawarehouse, además se implementaron procesos ETL. Es así que se ejecutaron las fases 5,6 y 7 de la metodología Kimball detallada en el marco teórico sección 4.6.2 . Las etapas desarrolladas son:

- **Modelado dimensional:** se definieron las dimensiones, medidas y tablas de hecho de la solución BI.
- **Creación del modelo físico:** se implementó una base de datos física en PostgreSQL, a partir del modelado dimensional.
- **Implementación de procesos ETL:** Se implementaron procesos ETL para poblar el DW en un formato acorde para la utilización de herramientas de análisis.

### 3) Implementar una solución de BI a través de un Dashboard.

En el último objetivo, se llevó a cabo el desarrollo y despliegue de la solución BI, así como también se validó y aceptó por parte de la empresa. Es así que se ejecutaron las fases 8 y 9 de la metodología Kimball detallada en el marco teórico sección 4.6.2 . Por lo tanto, las etapas desarrolladas son:

- **Desarrollo de la aplicación BI:** Se llevó a cabo la implementación del dashboard, en el que se puede visualizar información de valor para la empresa.

- **Despliegue de la aplicación BI:** El dashboard se implementó en un servidor de aplicaciones para que esté disponible en todo momento.
- **Pruebas y aceptación de la solución BI:** Se desarrollaron diversas pruebas para evaluar la solución de Business Intelligence desarrollada, y el gerente de la empresa la aprobó mediante la firma de un acta de acuerdo.

#### 5.4 Recursos

Con la finalidad de lograr el cumplimiento de los objetivos planteados, se hizo uso de los siguientes recursos:

- **Método Científico.** Este método fue la base para el desarrollo del Trabajo de Titulación realizado en sus diferentes etapas, desde el planteamiento del TT hasta su culminación. A partir de la problemática encontrada, se formuló una pregunta de investigación que explica la causa del problema en la gestión de datos y se sugiere una solución de BI para solventarla. Para probar la pregunta de investigación formulada se seleccionaron los datos de pedidos de la empresa, además se identificó la metodología, herramientas y técnicas más adecuadas para el tratamiento y presentación de estos datos. Además, con este método se buscó analizar y sintetizar los conceptos presentes en la revisión de literatura para fundamentar teóricamente el proceso de investigación. También permitió alcanzar los resultados de cada objetivo específico planteado, contrastar los datos con los trabajos relacionados y, finalmente, elaborar conclusiones y recomendaciones basadas en los resultados obtenidos.
- **Método analítico.** En el TT se utilizó este método para descomponer el objetivo general en varias fases que se establecieron como objetivos específicos, acompañadas de un conjunto de actividades. Durante la extracción de datos de la fuente de información, se descompuso la base de datos brindada por la empresa para obtener un reporte de los registros generados únicamente en el área de pedidos para evitar cargar datos innecesarios.
- **Investigación Bibliográfica:** Esta técnica dentro del TT sustentó la fundamentación teórica. Se aplicó la técnica de búsqueda por conveniencia para identificar los principales conceptos a través de literatura gris.
- **Estudio de casos:** Se empleó esta técnica para estudiar y analizar casos donde se implementó una solución de Business Intelligence para la gestión de datos empresariales, se identificaron 8 casos. Además, permitió recopilar información para fortalecer el marco

teórico y documentar detalladamente las actividades desarrolladas en cada uno de los objetivos establecidos.

- **Entrevista:** Se entrevistó al gerente de la empresa, con el objetivo de conocer la organización y evaluar el estado actual en la gestión de datos generados en la sección de pedidos. Esta entrevista fue fundamental durante la ingeniería de requerimientos de la solución desarrollada.

### 5.5 Estándares

- **IEEE 830:** Se aplicó durante la especificación de requisitos de la solución BI en el primer objetivo específico del TT.

### 5.6 Metodología de desarrollo

La metodología de desarrollo seleccionada para llevar a cabo la implementación de la solución BI es la definida por Ralph Kimball, descrita en el punto 4.6 del Marco Teórico.

### 5.7 Participantes

TABLA VI. PARTICIPANTES DEL TT

Nombre	Descripción
Luis Miguel Quito Carrión	Estudiante a cargo de la ejecución del TT.
Ing. Edwin Rene Guamán Quinche Mg. Sc.	Docente y director del TT.
Ing. Diego Bravo	Gerente de la Empresa Marketing Digital Bravo & Beltrán, colaborando a lo largo del TT desde el estudio de la empresa, recolección de requisitos y aceptación de la solución propuesta.

## **6. Resultados**

En esta sección, se presentan los resultados obtenidos durante la realización del Trabajo de Titulación. Para lograr el éxito del proyecto, se definieron tres objetivos específicos y se llevaron a cabo una serie de actividades detalladas a continuación.

### **6.1 OBJETIVO 1: Analizar los datos históricos generados en el módulo de pedidos para establecer el diseño del modelo.**

Para cumplir con el objetivo establecido, se realizó un análisis empresarial con el fin de identificar las actividades realizadas por la empresa y los datos disponibles. Luego, se especificaron los requisitos de software siguiendo las directrices del estándar IEEE 830. Se procedió a diseñar la arquitectura para los datos y el área de presentación, y finalmente se instaló las herramientas necesarias para la implementación exitosa de la solución BI.

#### **6.1.1 Análisis empresarial**

Para garantizar una implementación exitosa de la solución BI en la empresa, se llevó a cabo un análisis detallado de sus principales características. Este proceso fue esencial para comprender el funcionamiento y el comportamiento de la empresa, lo que permitió examinar y entender mejor sus necesidades de información. Además, el análisis ayudó a establecer un ámbito predefinido y comprender cada decisión tomada en cuanto al diseño e implementación del DW. Todo esto fue fundamental para construir y adaptar la solución BI de manera efectiva y asegurar su eficacia en la empresa. Los resultados detallados se pueden encontrar en el Anexo 1. Análisis Empresarial

Gracias al análisis llevado a cabo, se logró comprender en detalle el proceso de generación de información de los pedidos en la empresa. De esta manera, se pudieron identificar los diferentes pasos que se llevan a cabo desde que un cliente realiza un pedido hasta que se completa la entrega del mismo. Además, se analizaron los participantes y sistemas utilizados en cada fase del proceso para recopilar, procesar y transmitir la información necesaria de manera eficiente. En la Figura 14, se muestra el proceso para la generación de registros de pedidos en la empresa.

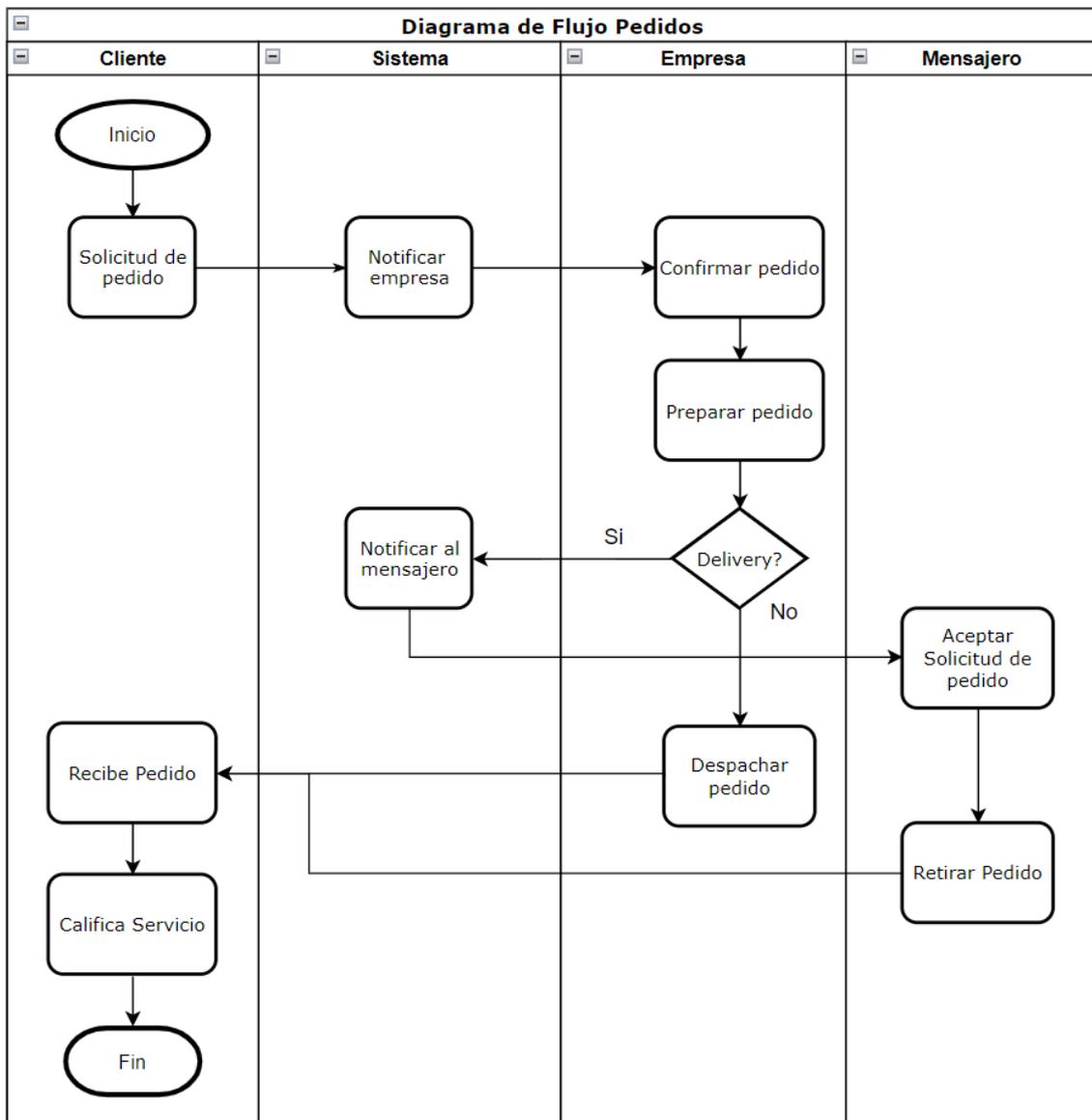


Figura 14. Diagrama de procesos de pedidos

### 6.1.2 Especificación de requerimientos

En esta sección se presentan los requerimientos funcionales y no funcionales, así como los casos de uso que se establecieron siguiendo el estándar IEEE 830 (ver Anexo 2), que fueron obtenidos por medio de la entrevista realizada al gerente de la empresa Marketing Digital Bravo & Beltrán (ver Anexo 3. Entrevista)

#### Requerimientos funcionales

En esta fase, se identificaron las necesidades de información requeridas por la gerencia, las cuales son esenciales para cumplir con los objetivos y estrategias de la empresa. Se definieron un total de 16 requisitos de BI, los cuales se detallan en la TABLA VII.

TABLA VII. REQUISITOS FUNCIONALES

<b>Código</b>	<b>Requisitos del proyecto</b>
RF02	Visualizar datos históricos de pedidos
RF03	Visualizar pedidos y monto facturado
RF04	Conocer número de empresas y usuarios
RF05	Mostrar empresa con más pedidos
RF06	Visualizar productos y montos por empresa
RF07	Mostrar usuarios con más pedidos
RF08	Conocer productos más pedidos
RF09	Mostrar usuarios y monto por empresa.
RF10	Visualizar productos por empresa.
RF11	Visualizar pedidos y monto por tipo de pago
RF12	Conocer pedidos y monto por tipo de entrega
RF13	Mostrar pedidos por empresa
RF14	Visualizar pedidos y monto por empresa según el tipo de pago
RF15	Visualizar pedidos y monto por empresa según el tipo de entrega
RF16	Calcular promedio de pedido por empresa
RF17	Visualizar pedidos y monto por categoría y subcategoría

### Requerimientos no funcionales

En esta fase se identificaron 6 requisitos no funcionales que debe cumplir la solución de Business Intelligence:

TABLA VIII. REQUERIMIENTOS NO FUNCIONALES

<b>Código</b>	<b>Requerimiento</b>	<b>Descripción</b>
RNF01	Seguridad	El sistema debe brindar seguridad tanto a nivel de transmisión de datos, seguridad de acceso a bases de datos, y al sistema en sí.
RNF02	Escalabilidad	El sistema debe tener la capacidad de incorporar cambios que puedan presentarse dentro de la empresa.
RNF03	Usabilidad	El sistema debe ser construido de tal manera que resulte sencillo de manejar para el usuario.
RNF04	Confiabilidad	La solución BI debe presentar información correcta, ya que servirá de base para llevar a cabo el proceso de toma de decisiones.
RNF05	Disponibilidad	La solución BI debe estar disponible el 100% de las veces que el usuario intente acceder.

## Casos de uso

En la Figura 15, se encuentra el caso de uso para el módulo de Login, donde el gerente puede ingresar y salir del sistema.

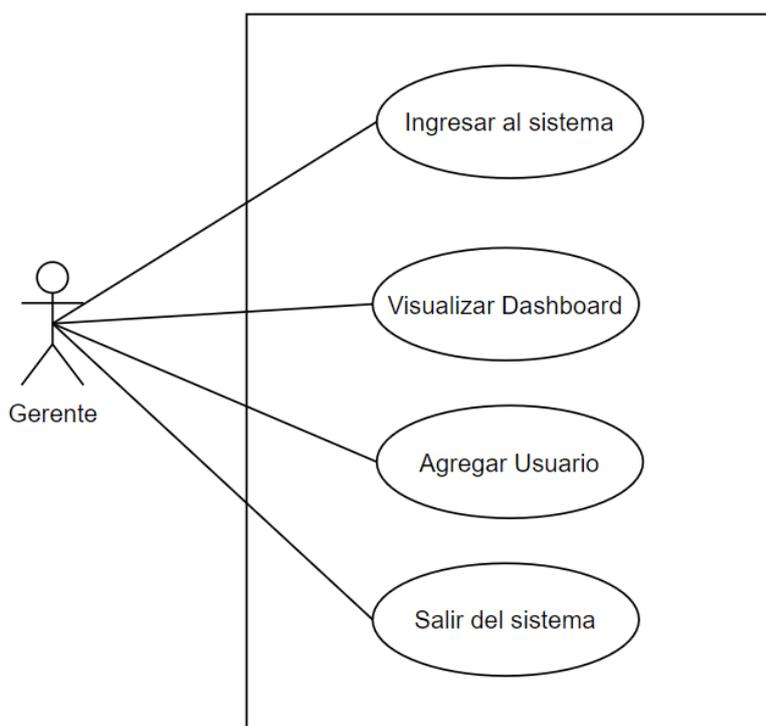


Figura 15. Caso de uso - Gerente

TABLA IX. CASO DE USO LOGIN Y LOGOUT

<b>ID</b>	CU001
<b>Nombre</b>	Login y Logout <b>Prioridad: Alta</b>
<b>Actor</b>	Gerente
<b>Descripción</b>	El sistema permitirá al usuario acceder ingresando su nombre de usuario y contraseña, además de brindar la opción de cerrar la sesión de manera segura al salir del sistema.
<b>Precondición</b>	El usuario tiene una cuenta válida en el sistema.
<b>Flujo principal Login</b>	<ol style="list-style-type: none"> <li>1. El usuario ingresa su nombre de usuario y contraseña.</li> <li>2. El sistema verifica que los datos ingresados son correctos y autentica al usuario.</li> <li>3. El sistema carga la página de inicio del usuario.</li> </ol>
<b>Flujo principal Logout</b>	<ol style="list-style-type: none"> <li>1. El usuario selecciona la opción de cerrar sesión.</li> <li>2. El sistema cierra la sesión del usuario y lo redirige a la página de inicio de sesión.</li> </ol>

TABLA X. CASO DE USO VISUALIZAR DASHBOARD

<b>ID</b>	CU002
<b>Nombre</b>	Visualizar Dashboard <b>Prioridad: Alta</b>
<b>Actor</b>	Gerente
<b>Descripción</b>	El módulo proporcionará al usuario una vista general de los datos de pedidos del sistema a través de un dashboard.
<b>Precondición</b>	El usuario tiene una cuenta válida y ha iniciado sesión en el sistema.
<b>Flujo principal</b>	<ol style="list-style-type: none"> <li>1. El usuario inicia sesión en el sistema.</li> <li>2. El sistema carga el dashboard.</li> <li>3. El usuario ve gráficos de tendencias de datos clave, como el número de pedidos totales, monto total, empresas y usuarios registrados.</li> <li>4. El usuario ve información actualizada sobre los productos más vendidos, los clientes que más compran, las empresas que más pedidos despachan y el promedio de venta.</li> <li>5. El usuario ve gráficos con información sobre los pedidos y el monto de acuerdo al tipo de pago, tipo de entrega</li> </ol>

TABLA XI. CASO DE USO AGREGAR USUARIO

<b>ID</b>	CU003
<b>Nombre</b>	Agregar Usuario <b>Prioridad: Alta</b>
<b>Actor</b>	Gerente
<b>Descripción</b>	El sistema permite al gerente agregar nuevos usuarios.
<b>Precondición</b>	El usuario tiene una cuenta válida y ha iniciado sesión en el sistema.
<b>Flujo principal</b>	<ol style="list-style-type: none"> <li>1. El usuario accede al panel de administración.</li> <li>2. El sistema muestra la opción para gestionar usuarios.</li> <li>3. El usuario selecciona la opción "Agregar usuario"</li> <li>4. El sistema muestra un formulario con varios campos informativos del nuevo usuario"</li> <li>5. El administrador ingresa los datos del nuevo usuario en el formulario</li> <li>6. Si los datos son válidos, el sistema agrega al nuevo usuario a la base de datos.</li> </ol>

### 6.1.3 Diseño de la arquitectura

La metodología Kimball proporciona dos capas para especificar la arquitectura del proyecto: Back Room y Front Room. La capa Back Room se enfoca en la obtención y preparación de los datos, mientras que la capa Front Room se encarga de entregar los datos al usuario final.

#### Origen de Datos

Para cumplir con esta fase, la empresa facilitó un archivo .sql que contiene la base de datos transaccional. Las características de la fuente de datos brindada por la institución se especifican en la Tabla XII.

Tabla XII. FUENTE DE DATOS

Sistema	Descripción	Tamaño	Estado
SISTEMA TRANSACCIONAL DE LA EMPRESA	Herramienta para almacenar y gestionar los datos de las transacciones que se efectúan en la empresa.	31 MB	En línea

Debido al alcance del proyecto, no se consideró la utilización de toda la base de datos proporcionada, sino que se generó un reporte en archivo .CSV únicamente de los registros generados en el área de pedidos para evitar cargar datos innecesarios. En la Figura 16, se muestra la sentencia SQL utilizada para obtener el archivo CSV que contiene los datos de pedidos generados, los cuales fueron utilizados para poblar el DW.

```
SELECT order_headers.id AS order_code, DATE_FORMAT(DATE(order_headers.created_at), '%d-%m-%Y') AS order_date, shipment_status AS ship_mode,
order_headers.total AS order_total, order_headers.iva_total AS order_iva, pay_mode_id AS pay_code, pay_modes.type_pay,
products.id AS product_code, products.name AS product_name, order_bodies.quantity AS product_quantity, order_bodies.price_unit AS product_price,
order_bodies.iva AS product_iva, order_bodies.price_total AS produc_price_total, stores.id AS store_code, stores.name AS store_name,
store_subsiadiaries.alias AS store_subsiadiary, categories.name AS store_category, subcategories.name AS store_subcategory,
users.id AS user_code, CONCAT_WS(' ', user_profiles.name, user_profiles.lastname) user_name
FROM pay_modes
INNER JOIN order_headers ON pay_modes.id = order_headers.pay_mode_id
INNER JOIN order_bodies ON order_headers.id = order_bodies.order_header_id
INNER JOIN products ON products.id = order_bodies.product_id
INNER JOIN stores ON stores.id = order_bodies.store_id
INNER JOIN store_subsiadiaries ON store_subsiadiaries.id = order_bodies.store_subsiadiary_id
INNER JOIN categories ON categories.id = stores.category_id
INNER JOIN subcategories ON subcategories.id = stores.subcategory_id
INNER JOIN users ON users.id = order_headers.user_id
INNER JOIN user_profiles ON user_profiles.user_id = users.id
ORDER BY order_headers.created_at DESC ;
```

Figura 16. Reporte de Datos Modulo Pedidos

#### Back Room

En esta capa se realiza la extracción de datos desde el archivo CSV generado (ver Figura 16). Estos datos son transformados y cargados mediante procesos ETL hacia el

repositorio de datos principal. En la Figura 17, se observa la arquitectura del Back Room de la solución BI.

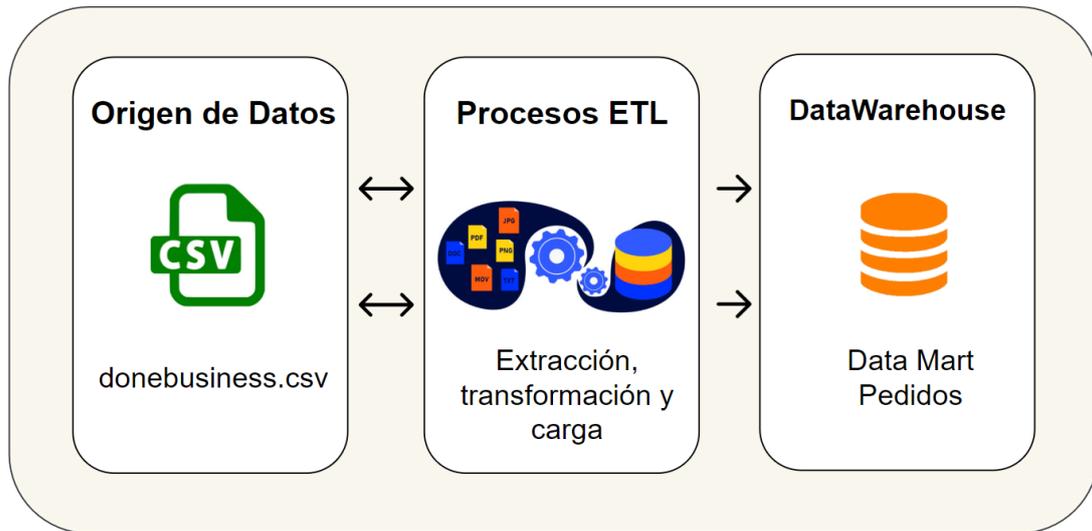


Figura 17. Back Room.

Además, en esta área ocurre el proceso de Data Staging (preparación de datos), donde se cargan, depuran y almacenan los datos.

### Front Room

En esta etapa se generó un cubo OLAP que proporciona los datos para la construcción del Dashboard. A continuación, en la Figura 18, se indica la arquitectura del Front Room que tiene la solución BI.

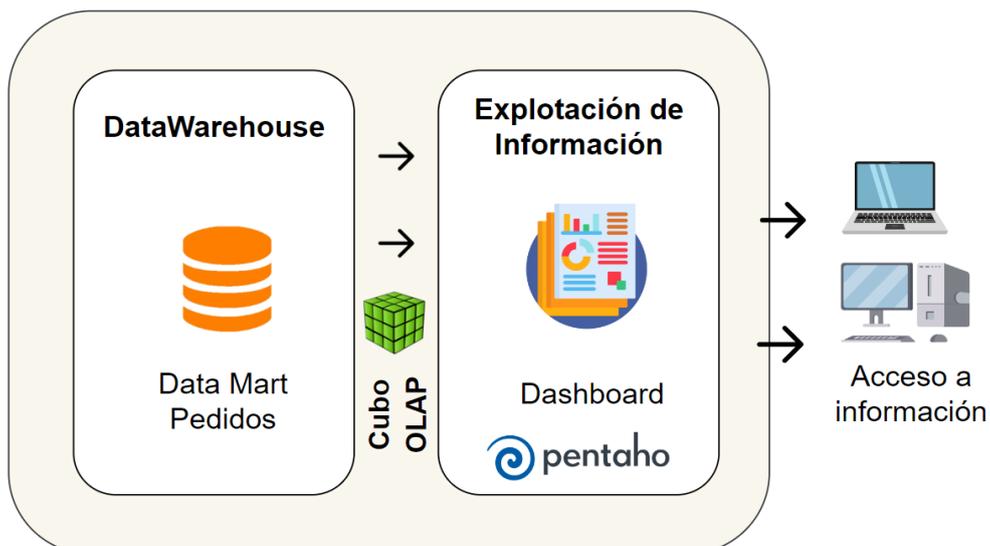


Figura 18. Front Room

#### 6.1.4 Selección e instalación de herramientas

Para mejorar el desarrollo de la solución de Business Intelligence, se ha dado prioridad a tres aspectos fundamentales: la existencia de documentación sobre el uso de la herramienta en trabajos relacionados, la facilidad de implementación y el costo. Es por eso que se eligió la suite de código abierto de Pentaho, la cual proporciona herramientas que simplifican la implementación de procesos ETL, análisis multidimensional y la creación de informes de manera sencilla y accesible.

Las herramientas instaladas para el cumplimiento del proyecto son:

- **Java Runtime Environment:** Necesario para la ejecución de programas Java.
- **Pentaho Data Integration:** Utilizado para llevar a cabo los procesos ETL.
- **Mondrian Schema Workbench:** Empleado para la creación del cubo OLAP.
- **Pentaho Server:** Utilizado para la creación de reportes y el dashboard.
- **PostgreSQL:** Necesario para la creación del Datawarehouse de la solución BI.
- **Jmeter:** Empleado para la ejecución de pruebas de carga y estrés de la solución BI.

Cada uno de los procedimientos realizados durante la instalación de las herramientas listadas se encuentran descritas en el Anexo 5. Instalación de Herramientas

#### 6.2 OBJETIVO 2: Desarrollar el modelo de datos utilizando la plataforma de BI Pentaho.

En el presente apartado se indica las actividades realizadas con el propósito de cumplir el objetivo 2.

##### 6.2.1 Modelado Dimensional

Después de obtener los requerimientos de la solución (ver en TABLA VII), se descompusieron los requerimientos planteados, con el objetivo de establecer las dimensiones y medidas que se utilizaron para elaborar el modelo lógico del Data mart.

##### Dimensiones

Las dimensiones hacen referencia a los objetos mediante los cuales se requiere examinar los hechos, por medio de ellas se puede responder a las preguntas planteadas. A continuación, se listan las variables de análisis identificadas en los requerimientos:

- Tiempo
- Empresa
- Producto
- Tipo de pago
- Tipo de entrega
- Cliente

## Medidas

Las medidas deben ser numéricas y representan lo que se desea analizar concretamente. Las medidas identificadas para el Datamart son los expresados a continuación:

- Número de pedidos
- Monto de pedidos
- Costo promedio de pedidos
- Número de productos

## Matriz Bus

En la Tabla XIII, se detalla el modelo de planificación del Datamart, que permitió definir las medidas para la tabla de hecho y la relación con las dimensiones.

Tabla XIII. MATRIZ BUS

<b>Medidas</b> \ <b>Dimensiones</b>	<b>Producto</b>	<b>Empresa</b>	<b>Cliente</b>	<b>Tipo Pago</b>	<b>Tiempo</b>
Número de pedidos	X	X	X	X	X
Monto de pedidos	X	X	X	X	X
Costo promedio pedidos		X			
Número de productos		X	X		

## Diagrama Start Net

Este diagrama muestra a través de líneas, círculos y rectángulos la representación del modelo dimensional, está compuesto por las medidas y la relación con las dimensiones y sus respectivos niveles.

En la Figura 19, se indica la Matriz Start Net correspondiente a la tabla de hechos ventas, que contiene métricas y a la derecha se encuentran las medidas a analizar (producto, empresa, cliente, tipo de pago, tipo de entrega tiempo).

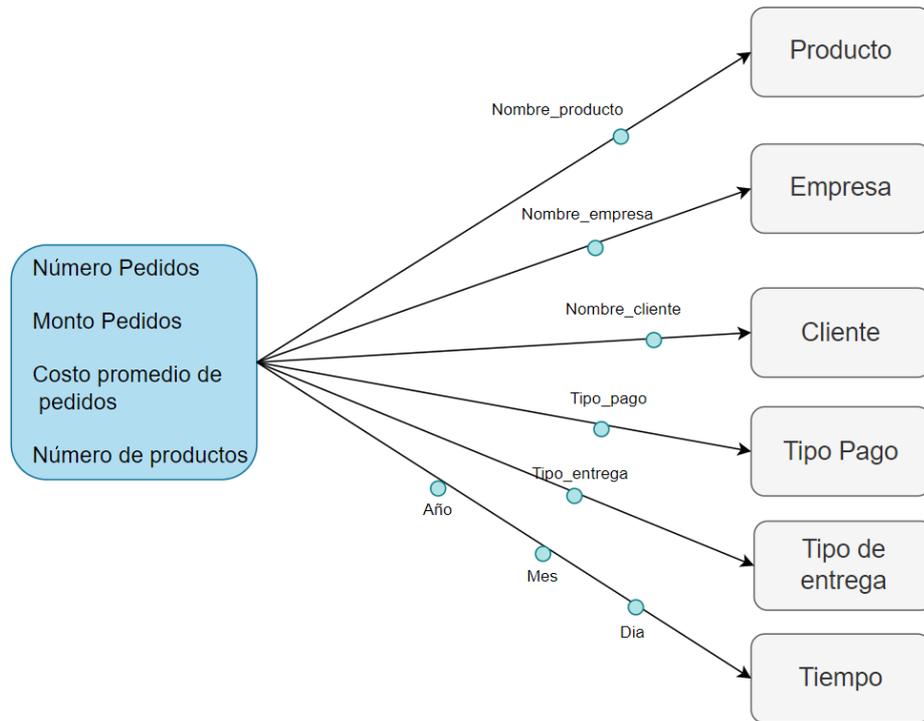


Figura 19. Matriz Start Net de Tabla de Hechos Pedidos

### Descripción de Dimensiones

A continuación, en la TABLA XIV se realiza una breve descripción de las dimensiones encontradas.

TABLA XIV. DESCRIPCIÓN DE DIMENSIONES

Dimensión	Descripción
<b>Tiempo</b>	Muestra los intervalos de tiempo de las transacciones realizadas en la empresa para un mejor análisis.
<b>Empresa</b>	Muestra información sobre las empresas que emiten productos.
<b>Producto</b>	Representa los productos despachados durante la solicitud de un pedido
<b>Tipo de Pago</b>	Indica los métodos de pago utilizados para cancelar el valor de un pedido
<b>Cliente</b>	Contiene información de los clientes que hacen uso del sistema de la empresa
<b>Tipo de entrega</b>	Representa información acerca del tipo de entrega (a domicilio o en el local)

### Descripción de Medidas

A continuación, en la TABLA XV se realiza una breve descripción de cada medida encontrada.

TABLA XV. DESCRIPCIÓN DE MEDIDAS

Medidas	Descripción
<b>Cantidad Pedidos</b>	Muestra la cantidad de pedidos realizados
<b>Monto Pedidos</b>	Muestra el monto de los pedidos generados
<b>Costo promedio de pedidos</b>	Representa el monto total de los pedidos con servicio de entrega a domicilio
<b>Número de productos</b>	Muestra la cantidad de productos

### Tabla de Hechos

El Data mart cuenta con una tabla de hechos denominada pedidos. Conforme a las medidas encontradas se definieron los atributos para la tabla de hechos:

- Cantidad de Ventas
- Monto de Ventas

### 6.2.2 Modelo Lógico

Como resultado del análisis dimensional, en la Figura 20 se muestra el diseño del modelo lógico del Data mart.

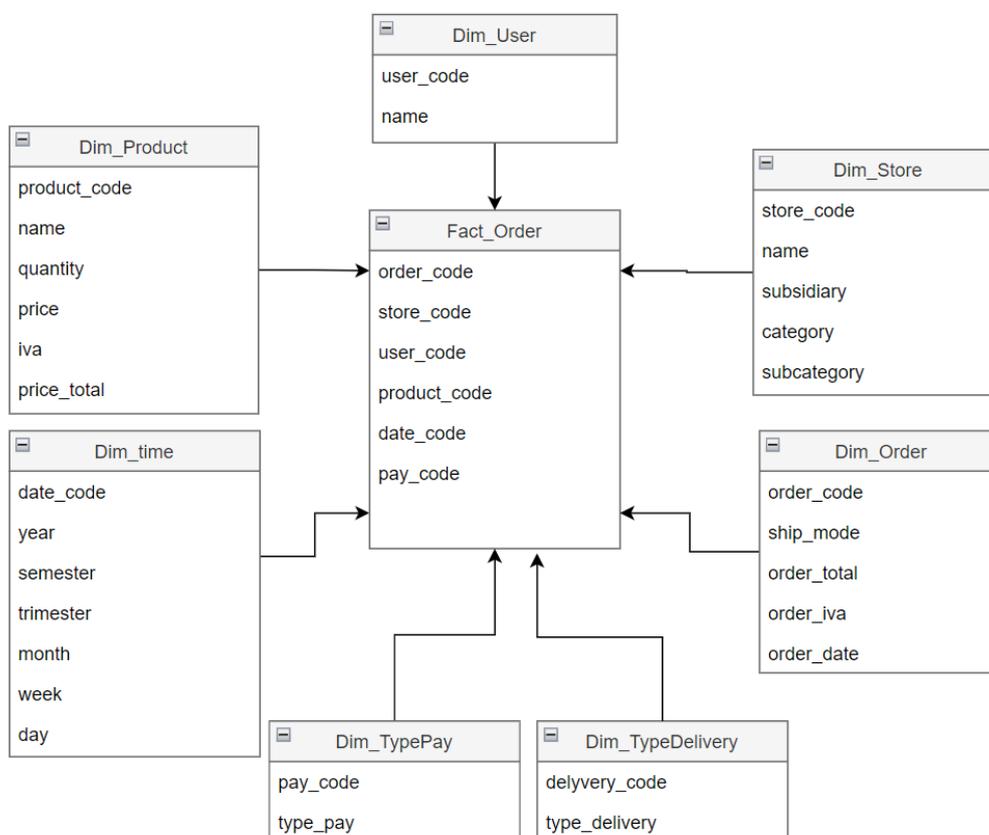


Figura 20. Modelo lógico del Data Mart

### 6.2.3 Modelo Físico

En la Figura 21, se ilustra el modelo físico dimensional de la base de datos. El modelo está basado en el esquema de estrella, presenta la tabla de hechos de pedidos relacionadas con sus respectivas dimensiones.

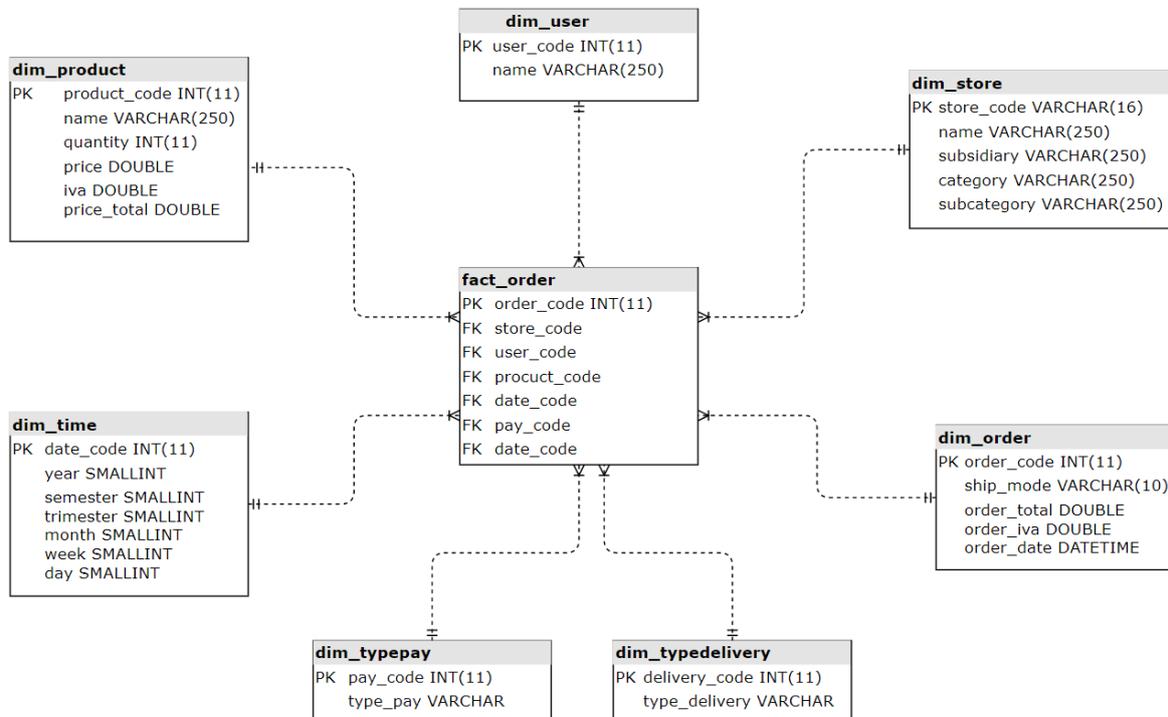


Figura 21. Modelo Físico en estrella

### Descripción de las Tablas

En esta sección se describe las tablas creadas en el modelo físico. A manera de ejemplo en la TABLA XVI, se muestra el nombre de la tabla “dim\_store”, una breve descripción de los registros que se almacenan, las columnas que contiene, descripción, tipo de dato, tamaño y tipo de llave que posee.

TABLA XVI. DESCRIPCIÓN TABLA “dim\_store”.

Tabla: dim_store					
Descripción: Tabla de registros de empresas					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
store_code	Identificador de empresa	Varchar	16	✓	
name	Nombre de la empresa	Varchar	250		
subsidiary	Sucursal de la empresa	Varchar	250		
category	Categoría a la cual pertenece la empresa	Varchar	250		
subcategory	Subcategoría a la cual pertenece la empresa	Varchar	250		

La descripción de todas las tablas que componen el Datamart se encuentran descritas en el Anexo 4. Descripción de tablas del Data mart

## 6.2.4 Implementación de procesos ETL

Durante el desarrollo de esta fase se pobló el Datamart, además se tomó como base la sección 6.2.1 Modelado Dimensional, donde se encuentran especificadas las dimensiones y atributos que fueron cargados por los procesos ETL.

### Conexión a la base de datos

Para establecer el sistema BI, se requirió una base de datos en la cual se crearon todas las tablas definidas en el modelo físico (ver Figura 21). Para llevar a cabo esto, se estableció una conexión con la base de datos "dw\_donesitios" en PostgreSQL a través de PDI. Fue en esta base de datos donde se almacenaron todos los datos resultantes de la ejecución de los procesos ETL desarrollados.

En la Figura 22, se muestra el valor de los parámetros (Host, nombre de la base, número de puerto, usuario, y contraseña) necesarios para establecer la conexión. Para verificar que el enlace fue establecido correctamente se realizó un test donde se obtuvo el mensaje "conexión exitosa".

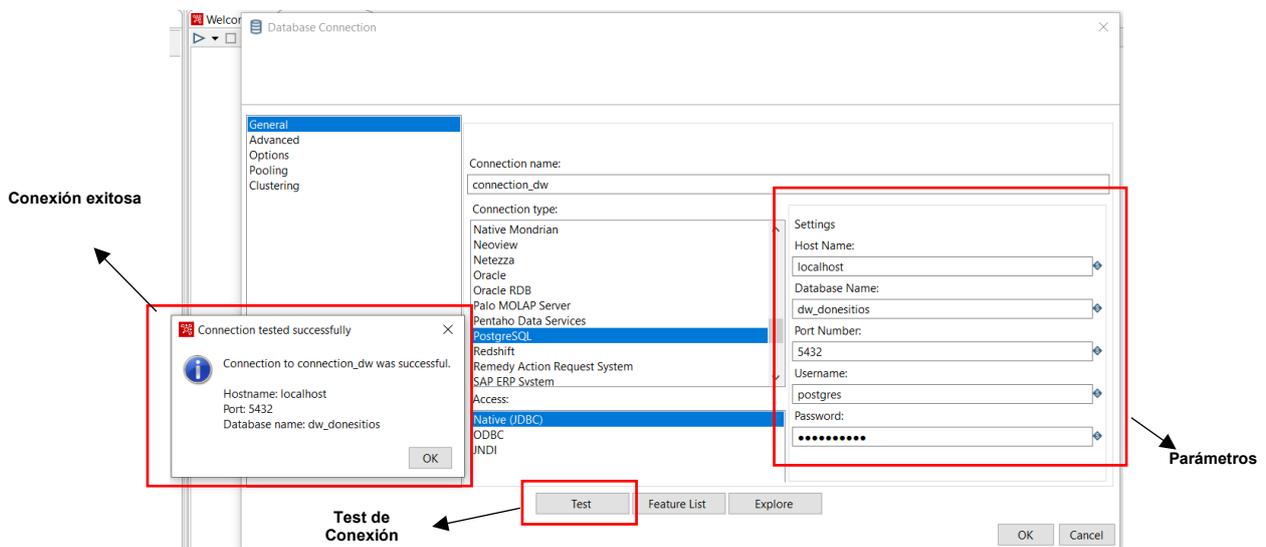


Figura 22. Conexión de PDI con PostgreSQL

### Procesos ETL

Para la implementación de los procesos ETL se utilizó como base la herramienta Pentaho Data Integration 9.2, que proporciona una abstracción grafica de los procesos que se está llevando a cabo con los datos a través de:



**Steps:** Realizan procesos de entrada, transformación o salida de registros.

 **Job:** Contienen la secuencia de tareas del proceso, pueden ser jobs, transformations y steps.

 **Transformation:** Contienen steps interconectados por hops que procesan registros desde el origen de datos hasta su salida.

**Hops:** Conectan los jobs, transformations, y steps entre sí, los datos siguen la dirección de la flecha.

 **Start:** Indica el inicio del proceso en cada job.

 **Success:** Indica la correcta finalización del job.

En la Figura 23, se ilustra el Job principal que contiene la secuencia de actividades desarrolladas durante la implementación de los procesos ETL.

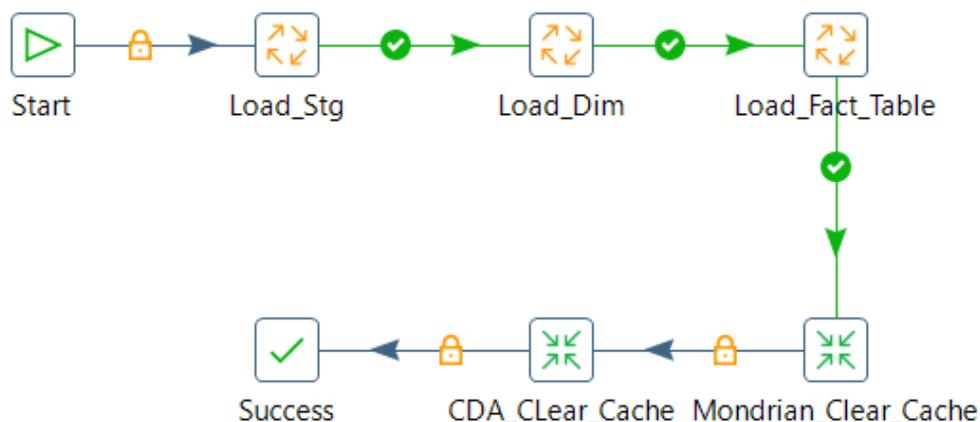


Figura 23. Job Principal

- **Start:** Indica el comienzo del job.
- **Load\_Stg:** Lleva a cabo la población de la base Staging, se recopilan todos los campos que se utilizaron de la fuente de datos. En la Figura 24, se ilustra las actividades desarrolladas en este job.
- **Load\_Dim:** Carga todas las tablas de las dimensiones del Datawarehouse, los datos son obtenidos de la base Staging. En la Figura 32, se detalla las actividades desarrolladas en este job.
- **Load\_Fact\_Table:** Carga los valores cuantitativos de cada dimensión. En la Figura

36, se indica las actividades desarrolladas en este job.

- **Mondrian\_Clear\_Cache:** Elimina la memoria Cache de Mondrian.
- **CDA\_Clear\_Cache:** Elimina la memoria Cache del Servidor de Pentaho.
- **Success:** Se ejecuta si todos los procesos se ejecutaron correctamente, caso contrario se interrumpirá la ejecución y mostrará una alerta en la consola.

A continuación, se realiza una descripción detallada sobre las actividades desarrolladas en el Job Principal.

#### ❖ Job Load\_Stg

Para que los datos puedan ser procesados y extraídos es necesario almacenarlos en el área de “staging”. En la Figura 24, se ilustra la secuencia en la ejecución de transformaciones que representan el almacenamiento en staging para cada dimensión y hecho.

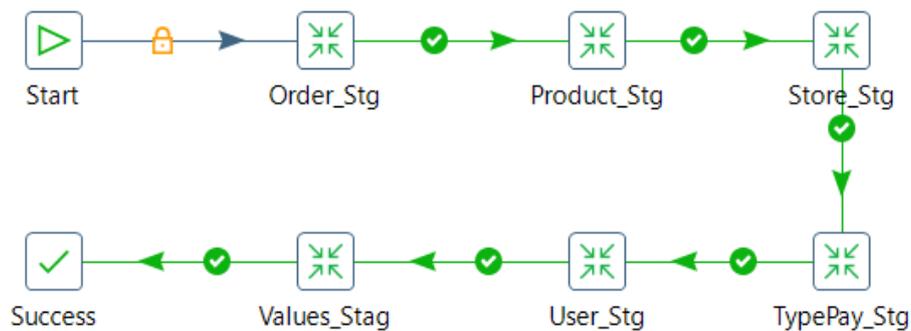


Figura 24. Job\_Load\_Stg

El Job Load\_Stg está compuesto por varias transformaciones, a manera de ejemplo, en la Figura 25 se detallan los steps que se llevaron a cabo en la transformación “Order\_Stg” para la carga de datos de ventas en el área de staging desde el archivo CSV.

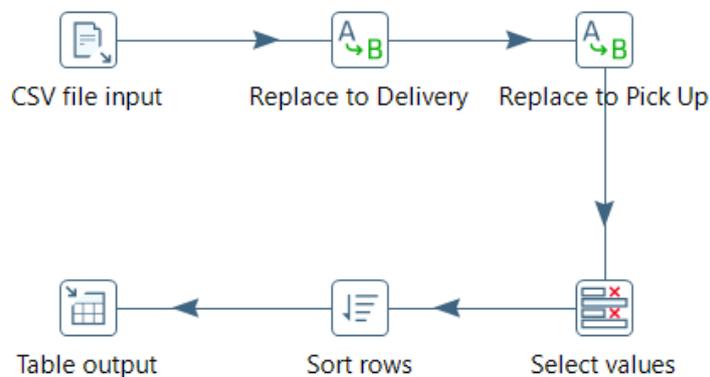


Figura 25. Transformación Order\_Stg

- *CSV file input*: este step permite leer los datos de archivos CSV en una transformación. En la Figura 26, se muestran los atributos y características cargados en PDI.

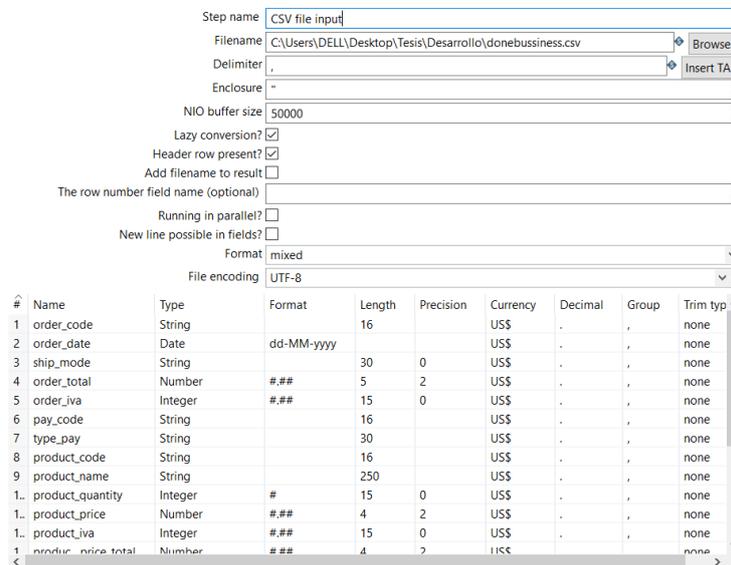


Figura 26. Step de Lectura del archivo CSV en PDI

- *Replace to Delivery*: se utilizó el step “replace in string” para que reemplace el tipo de venta a Delivery cuando la venta era realizada a domicilio en todo el dataset. En la Figura 27, se muestra el campo “ship\_mode”, en el cual se ejecutó el reemplazo.

#	In stream field	Out stream field	use RegEx	Search	Replace with
1	order_code		N		
2	ship_mode		N	0	Delivery

Figura 27. Step Replace to Delivery

- *Replace to Pick Up*: se utilizó el step “replace in string” para que reemplace el tipo de venta a Pick up in Store cuando la venta era retirada por el cliente en todo el dataset. En la Figura 28 se muestra el campo “ship\_mode”, en el cual se ejecutó el reemplazo.

#	In stream field	Out stream field	use RegEx	Search	Replace with
1	order_code		N		
2	ship_mode		N	1	Pick up in Store

Figura 28. Step Replace to Pick Up

- **Select Values:** Se obtiene los atributos propios que tendrá la tabla. En la Figura 29, se muestran los campos de las ventas.

#	Fieldname	Rename to	Length	Precision
1	order_code			
2	ship_mode			

Figura 29. Step Select Values

- **Sort rows:** En este step se ordenaron los datos en función del id de las ventas, además se puede controlar los datos repetidos, en la Figura 30 se muestra lo descrito anteriormente.

Step name: Sort rows

Sort directory: %%java.io.tmpdir%%

TMP-file prefix: out

Sort size (rows in memory): 1000000

Free memory threshold (in %):

Compress TMP Files?

Only pass unique rows? (verifies keys only)

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	order_code	Y	N	N	0	N

Figura 30. Step sort rows

- **Table output:** Finalmente se cargan los datos en una tabla de la base de datos. En la Figura 31, se muestran los campos que serán insertados en la base de datos.

Step name: Table output

Connection: connection\_dw

Target schema: Stg\_Order

Target table: Stg\_Order

Commit size: 1000

Truncate table

Ignore insert errors

Specify database fields

Main options | Database fields

Fields to insert:

#	Table field	Stream field
1	order_code	order_code
2	ship_mode	ship_mode

Figura 31. Step Table output

El proceso del resto de transformaciones especificadas en el Job Load\_Stg (ver en la Figura 24), se encuentra descritas en el Anexo 6. Llenado de datos.

## ❖ Job Load\_Dim

Este job ejecuta transformaciones que se encargan de extraer los datos cualitativos de la fuente de datos hacia las dimensiones. En la Figura 32, se detallan las diferentes transformaciones para cada una de las dimensiones.

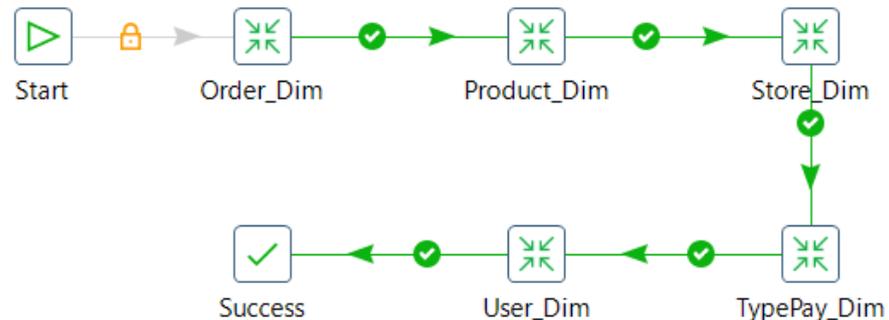


Figura 32. Job Load\_Dim

El Job Load\_Dim contiene varias transformaciones, a manera de ejemplo, en la Figura 33 se detallan los steps que se llevaron a cabo en la transformación "Order\_Dim" para la carga de datos de la dimensión ventas en la base de datos desde el área de staging.



Figura 33. Transformación Order\_Dim

- **Table input:** En este step se obtienen los datos almacenados en el área de staging, en la Figura 34 se muestra la configuración realizada

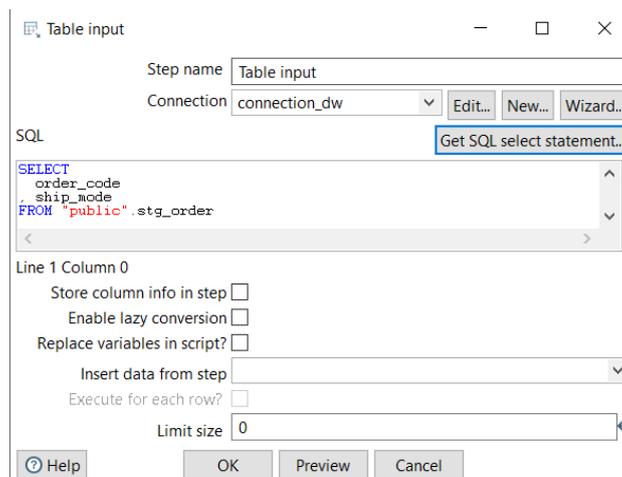


Figura 34. Step table input Order\_Dim.

- **Dimensión lookup/update:** permite implementar nuevos registros en la

dimensión, así como de actualizaciones. En la Figura 35, se muestra la configuración del step.

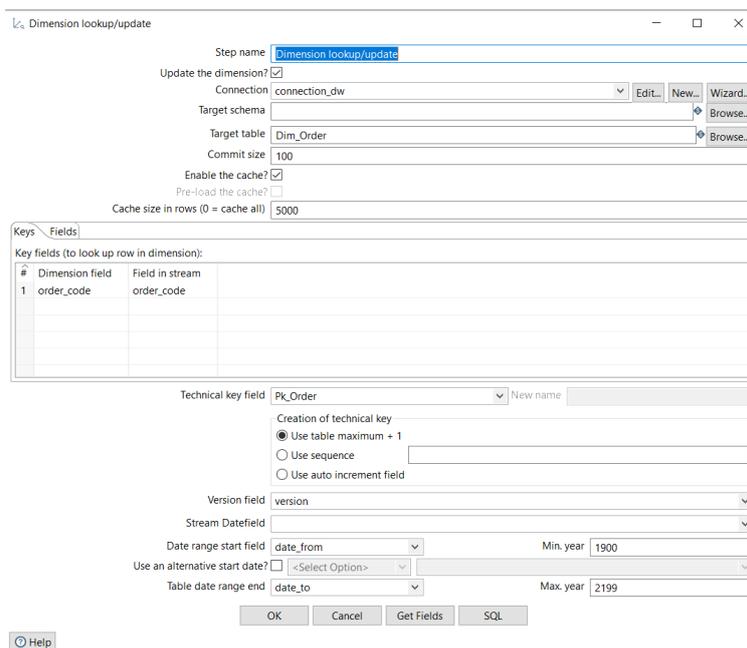


Figura 35. Step Dimensión lookup/update Dim\_Order

El proceso del resto de transformaciones especificadas en el Job Load\_Dim (ver Figura 32), se encuentra descritas en el Anexo 6. Llenado de datos.

### ❖ Job Load\_Fact\_Table

Este job ejecuta transformaciones que se encargan de extraer los datos cualitativos de la fuente de datos hacia las dimensiones. En la Figura 36 se detallan las diferentes transformaciones para cada una de las dimensiones.



Figura 36. Job Carga Tablas de hechos

El Job Load\_Fact\_Table contiene dos transformaciones, en la Figura 37 se detallan los steps que se llevaron a cabo en la transformación "Pre\_Fact".

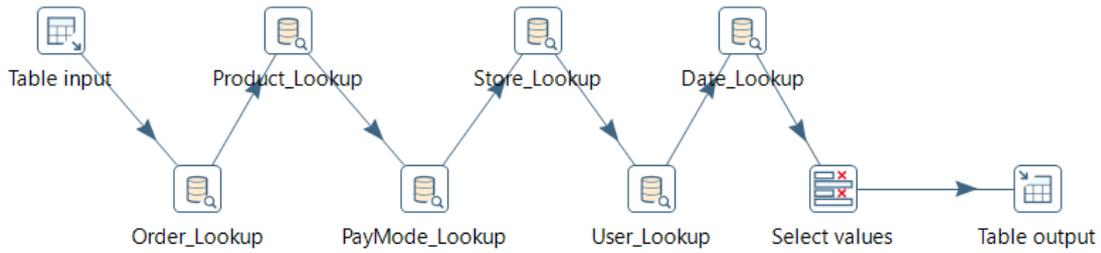


Figura 37. Transformación Pre\_Fact\_Table

- **Table input:** En este step se obtienen los valores cuantitativos almacenados en el área de staging
- **Database Lookup:** Se selecciona los campos correspondientes de cada una de las tablas de las dimensiones para obtener la llave generada asociada. En la Figura 38 se detalla la configuración realizada al step “Order\_Lookup”. Para el resto de dimensiones se debe seleccionar los valores correspondientes a cada una ellas.

La imagen muestra la configuración de un paso de 'Database lookup' con el nombre 'Order\_Lookup'. La conexión es 'connection\_dw' y el esquema de búsqueda es 'Dim\_Order'. Se han definido tres claves para la búsqueda:

#	Table field	Comparator	Field1	Field2
1	order_code	=	order_code	
2	date_from	<	order_date	
3	date_to	>=	order_date	

Los valores a devolver de la tabla de búsqueda son:

#	Field	New name	Default	Type
1	Pk_Order	Fk_Order		Integer

Se muestran también opciones como 'Enable cache?' (desactivado), 'Cache size in rows' (0) y 'Load all data from table' (desactivado). Los botones de acción incluyen 'Help', 'OK', 'Cancel', 'Get Fields' y 'Get lookup fields'.

Figura 38. Database lookup Order

- **Select Values:** Se obtiene los atributos propios que tendrá la tabla de hechos.
- **Table output:** Finalmente se cargan los datos en una tabla de la base de datos.

### Creación del cubo OLAP

Para diseñar el cubo multidimensional, se empleó la herramienta Mondrian Schema Workbench, la cual facilitó la creación y gestión de archivos XML que conforman el esquema desarrollado. Para ello, se llevaron a cabo una serie de actividades, detalladas a continuación:

Primero, se estableció una conexión entre el Data Warehouse alojado en PostgreSQL y Schema Workbench. En la Figura 39 se muestran los valores de los parámetros (host, nombre de la base de datos, número de puerto, usuario y contraseña) necesarios para establecer el enlace. Para verificar que la conexión se estableció correctamente, se realizó una prueba que arrojó el mensaje "conexión exitosa".

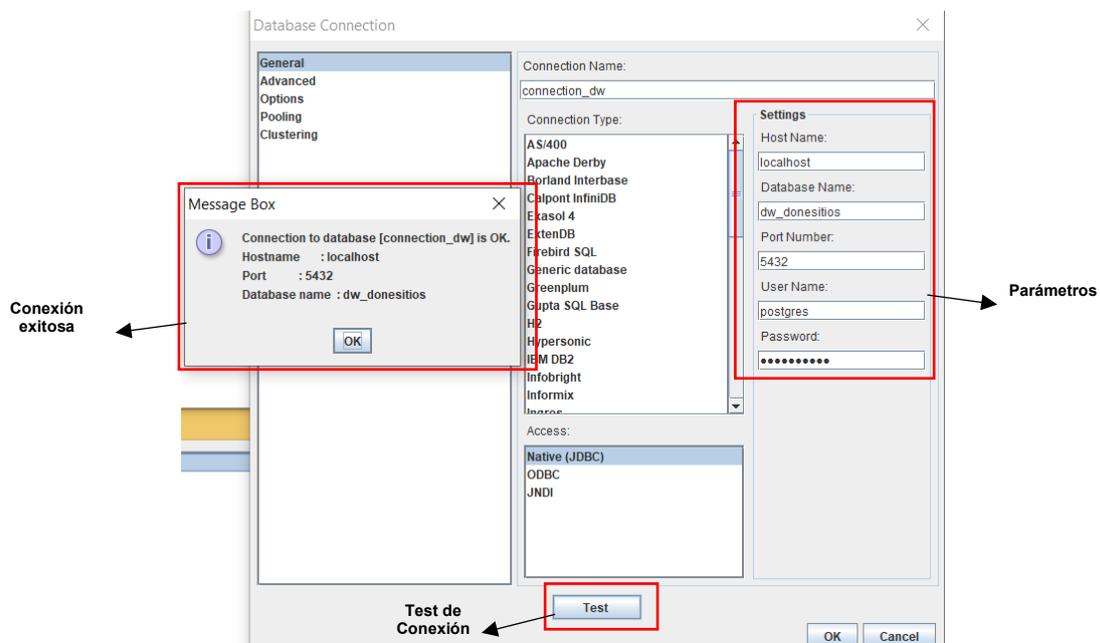


Figura 39. Conexión de Schema Workbench al Data Warehouse

En la Figura 40, se muestra el resultado de la creación del archivo XML (Schema), en el cual se define el cubo OLAP (ordersCube), las dimensiones y las medidas.

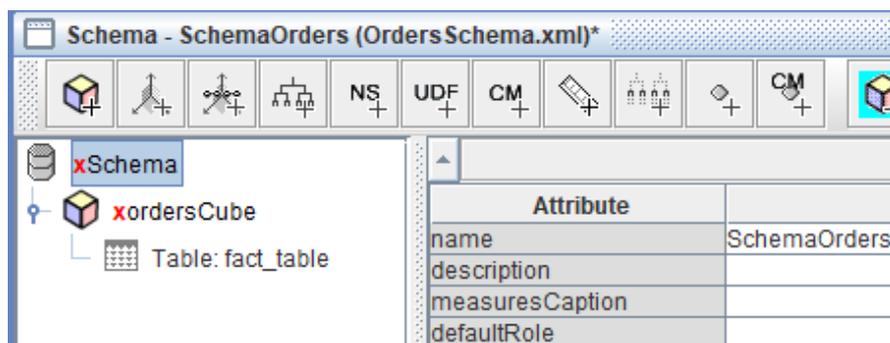


Figura 40. Creación de Schema y Cubo

Se definió en el cubo las dimensiones que fueron definidas en la sección 6.2.1 Modelado Dimensional. A continuación, en la Figura 41 se ilustra las dimensiones incluidas en el cubo "ordersCube".

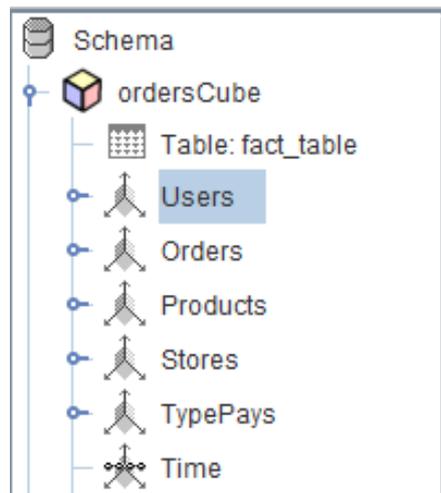


Figura 41. Dimensiones del cubo “ordersCube”

Es importante destacar que se pueden manejar varias jerarquías dentro de una misma dimensión, cada dimensión puede tener múltiples niveles.

A manera de ejemplo en la Figura 42, se muestra la jerarquía definida para la dimensión Stores (Category, Subsidiary, Stores). Cada jerarquía tiene varios niveles y una tabla que hace referencia a la dimensión que pertenece; por ejemplo, en la jerarquía “StoresIdentification” se presentan dos niveles (Code, Name) y la tabla de referencia (dim\_store) a la dimensión (Stores).

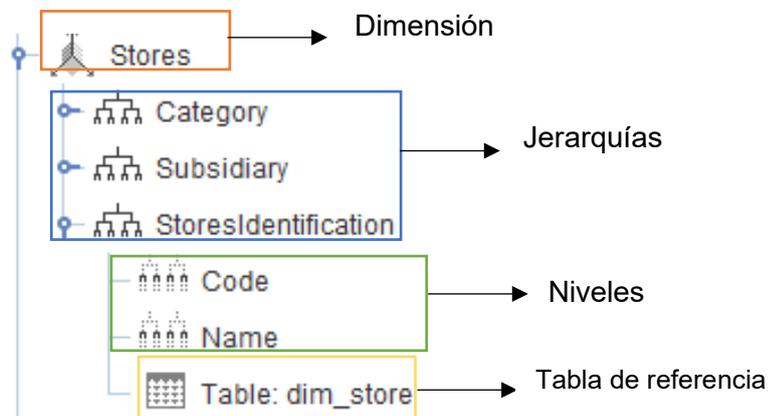


Figura 42. Jerarquía Dimensión Stores

Las jerarquías creadas para el resto de dimensiones especificadas en la Figura 41, se encuentran descritas en el Anexo 7. Definición de Jerarquías.

Finalmente, en la Figura 43 se ilustran todas las medidas creadas para la solución BI.



Figura 43. Medidas del cubo “ordersCube”

A manera de ejemplo, en la Figura 44 se muestra la configuración de los campos necesarios para la creación de la medida “Pedidos” que representa el número de pedidos.

Attribute	
name	Pedidos
description	
aggregator	distinct count
column	fk_order
formatString	
datatype	Numeric
formatter	
caption	
visible	<input checked="" type="checkbox"/>

Figura 44. Creación de medidas

- **Name:** nombre de la medida Pedidos.
- **Aggregator:** distinct count, cuenta valores únicos de la columna.
- **Column:** Hace referencia a la columna de la tabla.
- **Datatype:** Tipo de dato

### 6.3 OBJETIVO 3: Implementar una solución de BI a través de un Dashboard

Para el cumplimiento del objetivo planteado, se desarrolló un Dashboard generado a partir del cubo OLAP creado con anterioridad. Además, se evaluó el Dashboard a través de la información generada en los distintos reportes de la empresa con lo cual se pudo validar su correcta implementación.

Después de haberse creado el cubo es necesario la utilización de la herramienta Pentaho Server para leer el cubo y posteriormente crear el Dashboard. En primer lugar, se inicializó Pentaho Server mediante el archivo *start-pentaho.bat*, para acceder a su interfaz se utilizó el

navegador web Google Chrome. Por defecto Pentaho Server utiliza el puerto 8080, en caso que este puerto estuviese ocupado se lo debe cambiar por otro disponible.

En la Figura 45, se ilustra la inicialización de Pentaho Server en el navegador de Google Chrome.



Figura 45. Inicialización de Pentaho Server en Google Chrome

### Publicación de cubo OLAP

Una vez iniciado Pentaho Server se realizó la publicación del cubo, para ello en Schema Workbench se abre el proyecto que contiene el cubo creado, después en el menú *File* se eligió la opción "*Publish*" de la forma que se ilustra en la Figura 46.

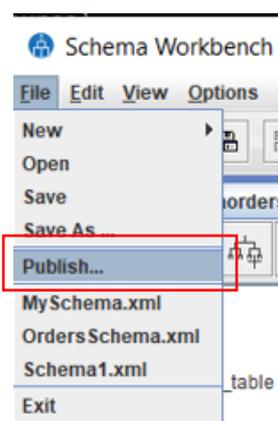


Figura 46. Publicación cubo OLAP

A continuación, en la Figura 47 se muestran los campos necesarios para la publicación del cubo (URL del servidor, usuario, contraseña y nombre). Es necesario tomar en cuenta que

para realizar la publicación del cubo debe estar ejecutándose el servidor web de Pentaho Server.

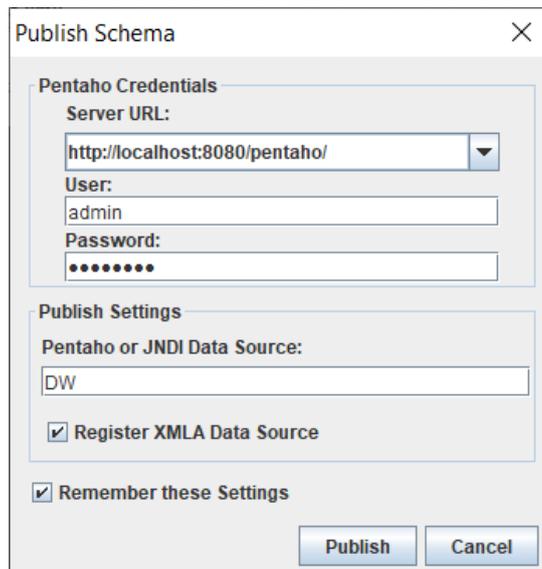


Figura 47. Campos para publicación de cubo

Si el cubo fue creado y publicado correctamente aparecerá un mensaje que indica que el cubo fue publicado con éxito, tal como se muestra en la Figura 48.

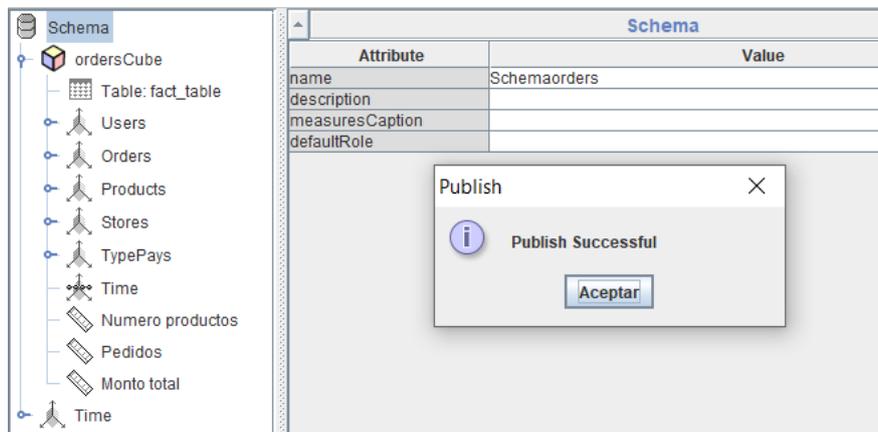


Figura 48. Publicación del cubo "ordersCube" exitosa

### Conexión entre Pentaho Server y el DW

Para establecer el enlace, se lo realizó mediante el panel de Pentaho Server en la opción "Manage Data Sources" tal como se ilustra en la Figura 49.

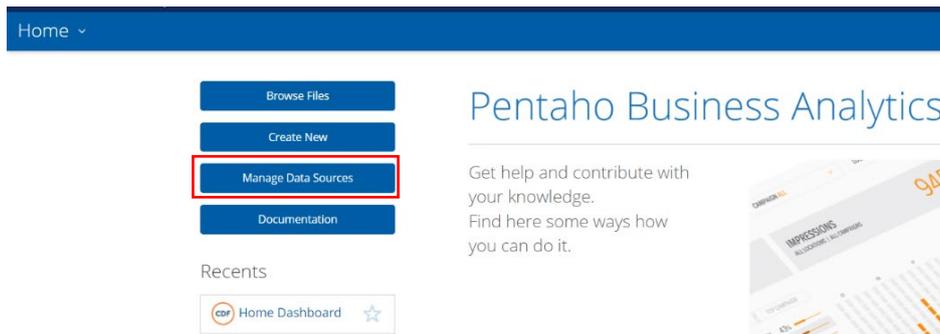


Figura 49. Administración de fuentes de datos en Pentaho Server

A continuación, en la Figura 50 se indica los parámetros (host, base de datos, puerto, usuario y contraseña) necesarios para establecer la conexión con el Data warehouse denominado “dw\_donesitios”. Es necesario realizar un *test* para comprobar que la conexión este bien establecida.

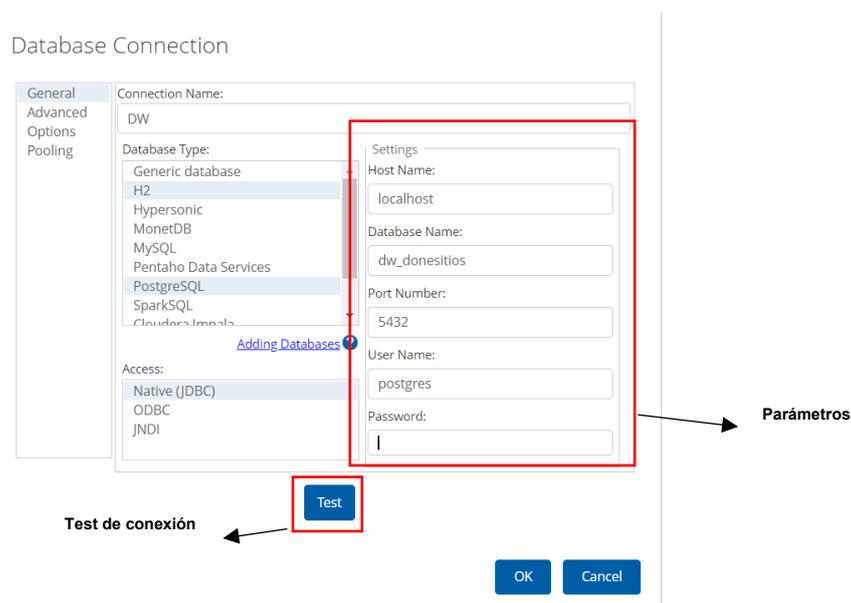


Figura 50. Conexión Pentaho-Data warehouse

### 6.3.2 Desarrollo de la aplicación BI

En esta etapa se utilizó la herramienta Pentaho Server para la creación del Dashboard, el cual contiene los reportes aprobados por el gerente de la empresa, que satisfacen los requisitos funcionales descritos en la TABLA I.

Para la lectura del cubo se instaló el plugin Pivot4J que permite el tratamiento de datos OLAP, se creó una nueva vista mediante la opción New JPivot View, donde se seleccionó el Schema y el cubo previamente publicado, en la Figura 51 se detalla lo mencionado.

## New JPivot View

Schema

Schemaorders ▾

Cube

ordersCube ▾

Ok

Figura 51. Lectura de cubo con JPivot

Después de haber seleccionado el cubo es posible realizar las consultas de acuerdo a la manera en cómo fueron definidas cada una de las dimensiones y medidas en Schema. En la Figura 52 se ilustra los campos del cubo “ordersCube” en Pentaho Server.

UsersIdentification	OrdersIdentification	ProductsIdentification	Category	PayName	Month	Medidas		
						Numero productos	Pedidos	Monto total
▣ All Users	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	35,274	14,106	124,967.53
Pablo David Rojas Apolo	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	99	42	360.45
Fernando Cevallos	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	12	7	45.48
Jina Stefania Arias Cordova	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	1	6.2
Mauricio Gregori Jiménez Bravo	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	1	12.69
1104868235 Gualpa Caraguay	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	62	7	28.72
1750496588 MARJURIE Sánchez	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	1	10
ANDREA Jaramillo	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	7	2	14
ANGIE MERCEDES GONZALEZ MOLINA	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	17	6	44.35
Aarón Conrado	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	34	8	38.55
Aarón Cabrera	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	3	2	19.49
Abigail Elizabeth Cruz Calle	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	13	5	33.9
Aby Quezada	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	7	2	13.75
Adi Merchan	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	1	9.75
AdriAna Torres	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	15	4	26.9
Adrian Castillo	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	3	1	5.35
Adrian Mauricio Toro Armijos	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	11	5	41.75
Adrian Ordoñez	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	3	3	12.2
Adriana Benites	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	2	0
Adriana Cuenca	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	6	2	18.7
Adriana Valarezo	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	2	1	20
Adriana Abarca	▣ All Orders	▣ All Products	▣ All Categories	▣ All Type Pays	▣ All Months	3	1	7.2

Figura 52. Cubo ordersCube en Pentaho Server

Para la creación del Dashboard se creó un nuevo Community Dashboard Editor (CDE) denominado *Dashboard\_donesitios*. En la Figura 53 se muestra la creación del cuadro de mando.

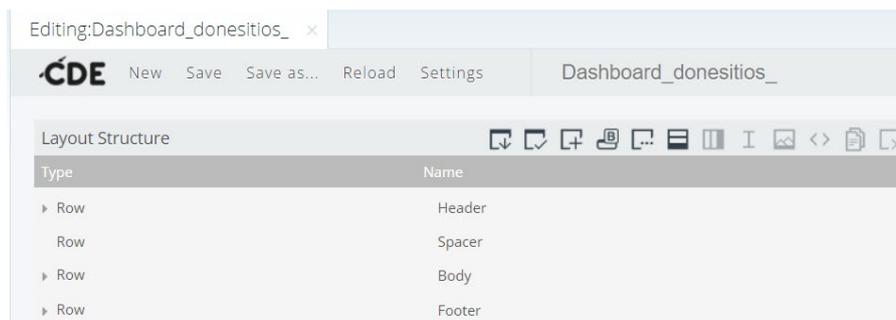


Figura 53. Creación del Dashboard\_donesitios

Para el desarrollo del dashboard en Pentaho Server se debe realizar un proceso que incluye configuraciones en tres componentes:



Figura 54. Componentes de un dashboard

- Layout: Es la representación gráfica de los datos. Se define el diseño y la estructura del dashboard, haciendo uso de estilos css, código javascript y html.
- Componentes: Se definen los componentes con los que se mostraran los datos obtenidos por la base de datos.
- Datasources: Se definen las consultas a realizar contra la base de datos.

A manera de ejemplo, a continuación, se explica el proceso realizado en cada componente para la obtención del reporte del requisito RF01.

#### - Configuración en Layout Panel

En esta sección se definió el diseño y la estructura del dashboard. Al estar trabajando con HTML se definió la estructura en filas y columnas. En la Figura 55 se ilustra el diseño de la estructura en el Layout Panel para el RF01.

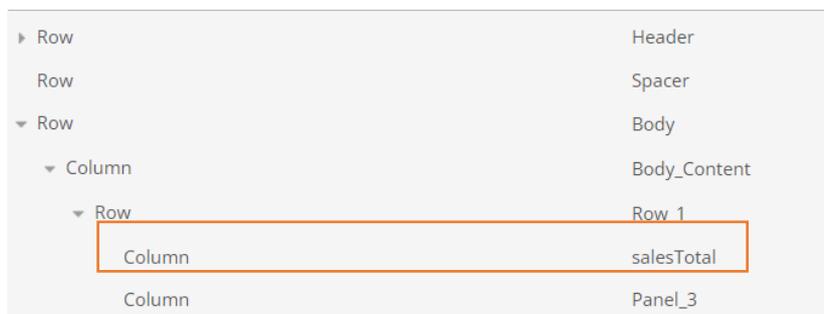


Figura 55. Diseño del Layout Panel del RF01

En la Figura 56, se muestra la configuración del Layout Panel del RF01.

Properties	
Property	Value
Name	salesTotal
Extra Small Devices	6
Small Devices	-
Medium Devices	-
Large Devices	-
Bootstrap Css Class	-
Height	400
BackgroundColor	<input type="text"/>
Corners	Round
Text Align	-
Css Class	-

Figura 56. Configuración del Layout Panel del RF01

**Name:** salesTotal, nombre de la columna donde se presenta el reporte.

#### - Configuración en Datasource Panel

En este panel se crearon las consultas a la base de datos, se crea una nueva *mdx over mondrianJndi* donde se configura las propiedades de la consulta. En la Figura 57, se muestra la configuración del Datasource del RF01.

Datasources		Properties	
Type	Name	Property	Value
Group	MDX Queries	Name	salesTotalQuery
mdx over mondrianJndi	salesTotalQuery	Access Level	Public
mdx over mondrianJndi	salesTotalTable	Jndi	DW
mdx over mondrianJndi	typePayQuery	Mondrian schema	Schemaorders
mdx over mondrianJndi	usersTopTable	Query	select NON EMPTY ((M (...)
mdx over mondrianJndi	salesStoreChartQuery	Parameters	<input type="text"/>
mdx over mondrianJndi	deliveryTypeQuery	Banded Mode	Compact
mdx over mondrianJndi	prueba	Calculated Columns	<input type="text"/>
mdx over mondrianJndi	filterStoresQuery	Columns	<input type="text"/>
mdx over mondrianJndi	tablaproducto	Output Columns	<input type="text"/>
mdx over mondrianJndi	categoryTotalQuery	Output Mode	Include
		Cache Keys	<input type="text"/>
		Cache Duration	3600
		Cache	True

Figura 57. Configuración del Datasource Panel del RF01

**Name:** salesTotalQuery, nombre que identifica a la fuente de datos para utilizarla en el resto de paneles.

**Jndi:** DW, hace referencia al Data Warehouse.

**Mondrian schema:** Schemaorders, esquema creado en schema-workbench.

**Query:** Sentencia MDX, para la obtención de esta consulta, a través de *JPivot View* se eligen los campos en las filas y columnas, para el RF01 se eligió la medida "Pedidos" y la fila "Fecha". En la Figura 58 se ilustra la consulta MDX

resultante.

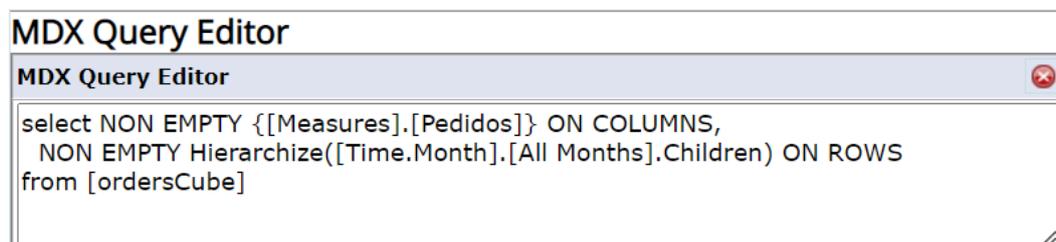


Figura 58. Creación de consulta MDX para el RF01

### - Configuración en Components Panel

En este panel se crearon los tipos de gráficos, tablas, selectores y filtros, se crea un nuevo *chart* que se configura de acuerdo a las propiedades del requisito. En la Figura 59, se muestra la configuración del panel component del RF01.

Name	salesTotalChart
Title	Historial de Pedidos
Listeners	[]
Parameters	[]
Datasource	salesTotalQuery
Height	400
Width	600
HtmlObject	salesTotal

Figura 59. Configuración del Component Panel del RF01

**Title:** Historial de Pedidos, nombre del reporte.

**Datasource:** salesTotalQuery, hace referencia al datasource creado anteriormente en la Figura 57.

**Height y Width:** Dimensiones del gráfico.

**HtmlObject:** salesTotal, hace referencia a la ubicación que tendrá el reporte dentro del maquetado HTML realizado anteriormente en la Figura 55.

Luego de configurar cada panel para presentar el informe del requisito RF01 el resultado es el indicado en la Figura 60.

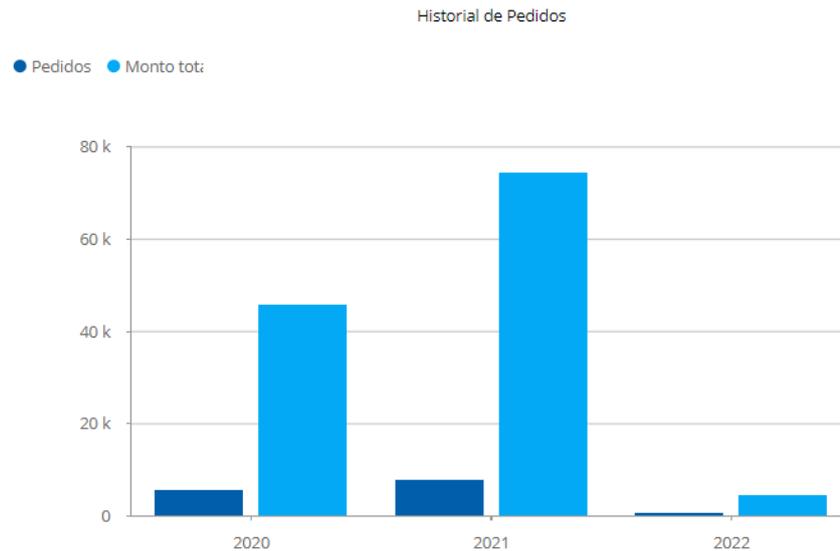


Figura 60. Reporte del RF01

Se construyeron dos dashboard, que el gerente de la empresa puede visualizar mediante la web ingresando su usuario y contraseña.

- **Estadísticas de pedidos en general:**

En la Figura 61, se ilustra el dashboard de pedidos general que cumple con los siguientes requisitos:

RF01: Historial de pedidos generados

RF02: Número de pedidos y monto facturado

RF03: Número de empresas y usuarios registrados

RF04: Empresas con más pedidos despachados

RF06: Usuarios con más pedidos generados y monto facturado

RF07: Productos más pedidos

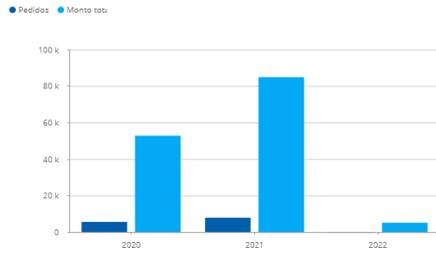
RF10: Número de pedidos y monto total por tipo de pago

RF11: Número de pedidos y monto total por tipo de entrega

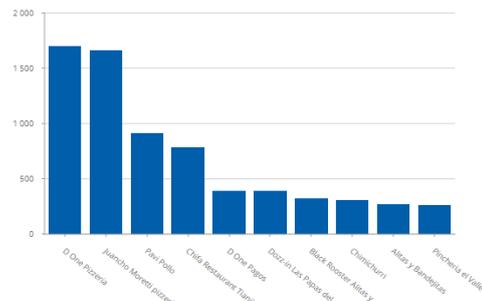
RF16: Pedidos y monto por categoría y Subcategoría

<b>Monto Total</b> \$	<b>Pedidos</b> ✓	<b>Empresas</b> ☰	<b>Usuarios</b> ☹
\$ 143,396.40	14,106	340	3,384

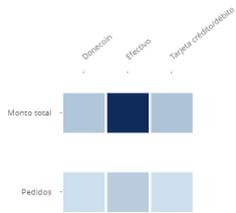
### Historial de Pedidos



### Pedidos por empresa



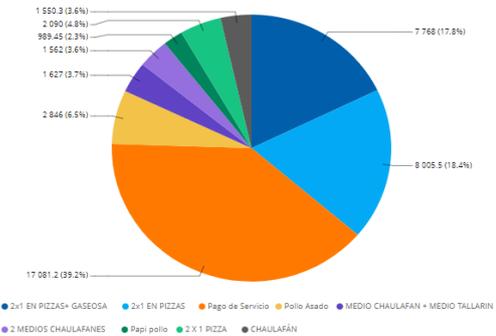
### Pedidos por pago



### Pedidos por entrega



### Productos más vendidos



### Usuarios con más pedidos

Usuarios	Pedidos	Monto Total
Vanessa del Rocio Beltrán Cabrera	168	1287.94
Diego Fernando Bravo Lozano	141	1096.43
John Jairo Villavicencio Sarango	136	1103.49
Favio Alexander Bravo Lozano	131	2341.13
Briggette Vásquez Sarmiento	122	590.45
Jackson Masache	119	737.46
John Villavicencio	72	506.14
Usuario Normal Donesitios	69	350.31
Paola Gabriela Bravo Lozano	68	534.39
Kevin Bryan Armijos Maza	67	3024.65

### Pedidos y monto por categoría

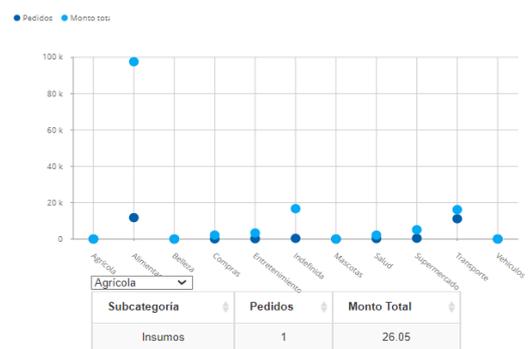


Figura 61. Dashboard general de Pedidos

- **Ventas por empresa**

En la Figura 59, se ilustra el dashboard de pedidos general que cumple con los siguientes requisitos:

RF05: Productos y monto más pedidos por empresa

RF08: Usuarios con más pedidos generados y monto facturado por empresa

RF09: Productos más pedidos por empresa

RF12: Historial de pedidos generados por empresa

RF13: Número de pedidos y monto total por empresa según el tipo de pago

RF14: Número de pedidos y monto total por empresa según el tipo de entrega

RF15: Promedio de venta por empresa

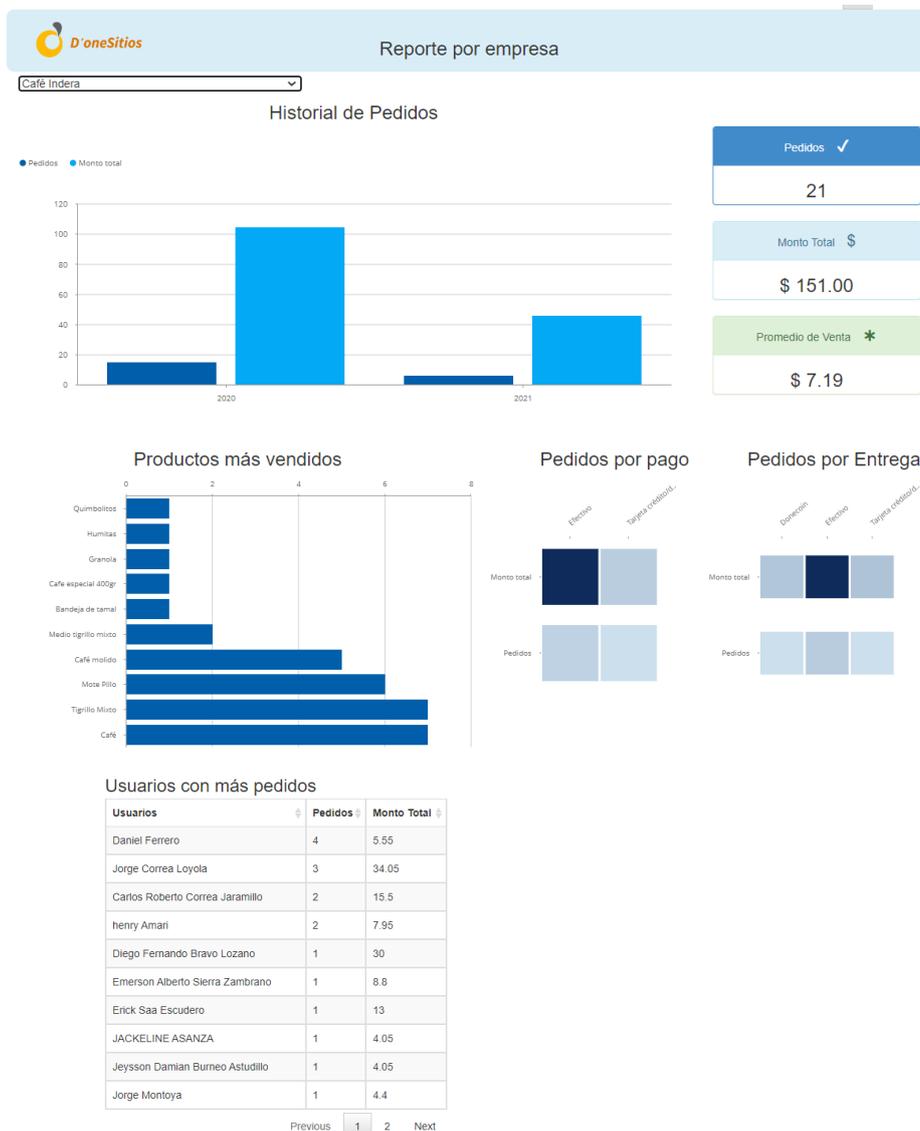


Figura 62. Dashboard por empresa

### 6.3.3 Despliegue de la aplicación BI

Una vez completado el desarrollo de todas las características de la solución BI, se procedió a subir la aplicación al servidor de aplicaciones. Para el despliegue de la aplicación en el servidor, se llevaron a cabo las configuraciones necesarias, las cuales se encuentran detalladas en el Anexo 8. Configuración para el servidor del sistema.

En la Figura 63, se muestra el diagrama de despliegue correspondiente para la implementación de la solución BI. Este consta de tres componentes: un equipo cliente que accede a través de un navegador web, un servidor de aplicaciones que aloja el servidor de Pentaho y PostgreSQL, y finalmente un servidor OLAP.

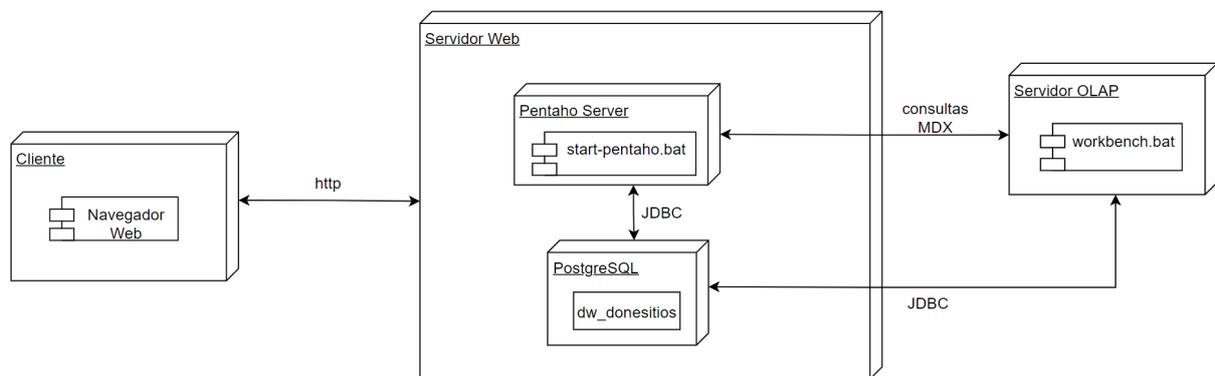


Figura 63. Diagrama de despliegue de la solución BI

Para implementar la solución BI en el servidor de DigitalOcean, se utilizó el comando SCP para cargar el contenido de la carpeta pentaho-server en el droplet pentaho-donesitios. En Windows, se usó el bash de Linux para ejecutar el comando. Desde la dirección correspondiente a la carpeta pentaho-server, se ingresó el siguiente comando en la terminal:

```
scp -r $PWD usuario@servidor : /rutaDestino
```

En la Figura 64, que muestra el Despliegue de la solución BI en el servidor, se puede observar la subida exitosa de los archivos al droplet en DigitalOcean.

```
PS C:\Users\DELL\Desktop\Tesis\Desarrollo\pentaho-server>
PS C:\Users\DELL\Desktop\Tesis\Desarrollo\pentaho-server> scp -r $PWD root@162.243.166.210:pentaho
root@162.243.166.210's password:
sales_data_sample.csv          100% 516KB 29.0KB/s 00:17
hibernate.lck                  100% 16    0.1KB/s 00:00
hibernate.log                  100% 0     0.0KB/s 00:00
hibernate.properties          100% 88    0.6KB/s 00:00
hibernate.script               100% 5632  20.0KB/s 00:00
migration.sql                  100% 568   3.4KB/s 00:00
quartz.lck                     100% 16    0.1KB/s 00:00
quartz.log                     100% 5120  11.4KB/s 00:00
quartz.properties             100% 89    0.7KB/s 00:00
quartz.script                  100% 7438  30.9KB/s 00:00
sampledata.backup              100% 165KB 19.3KB/s 00:08
sampledata.data                100% 1024KB 42.1KB/s 00:24
sampledata.lck                 100% 16    0.1KB/s 00:00
sampledata.log                 100% 0     0.0KB/s 00:00
sampledata.properties          100% 88    0.7KB/s 00:00
```

Figura 64. Despliegue de solución BI al Servidor

Al acceder mediante SSH al servidor, se pudo verificar que todos los archivos de la solución fueron subidos correctamente, tal como se muestra en la Figura 65.

```
root@pentaho-donesitios:~/server/pentaho-server#  
root@pentaho-donesitios:~/server/pentaho-server# ls  
data                promptuser.js      start-pentaho.sh  
import-export.bat  set-pentaho-env.bat stop-pentaho.bat  
import-export.sh   set-pentaho-env.sh stop-pentaho.sh  
licenses           start-pentaho-debug.bat third-party-tools  
pentaho-solutions start-pentaho-debug.sh tomcat  
pentaho.log        start-pentaho.bat
```

Figura 65. Validación de la solución BI en el servidor

Para iniciar el servidor, desde la carpeta pentaho-server se ejecutó el comando:

```
./start-pentaho.sh
```

Una vez que el servidor ha sido iniciado, es posible acceder a él desde un navegador web mediante la dirección <http://dashboard-donesitios/pentaho/Login>. Al hacerlo, se mostrará la pantalla de inicio de sesión de Pentaho tal como se ilustra en la Figura 66.

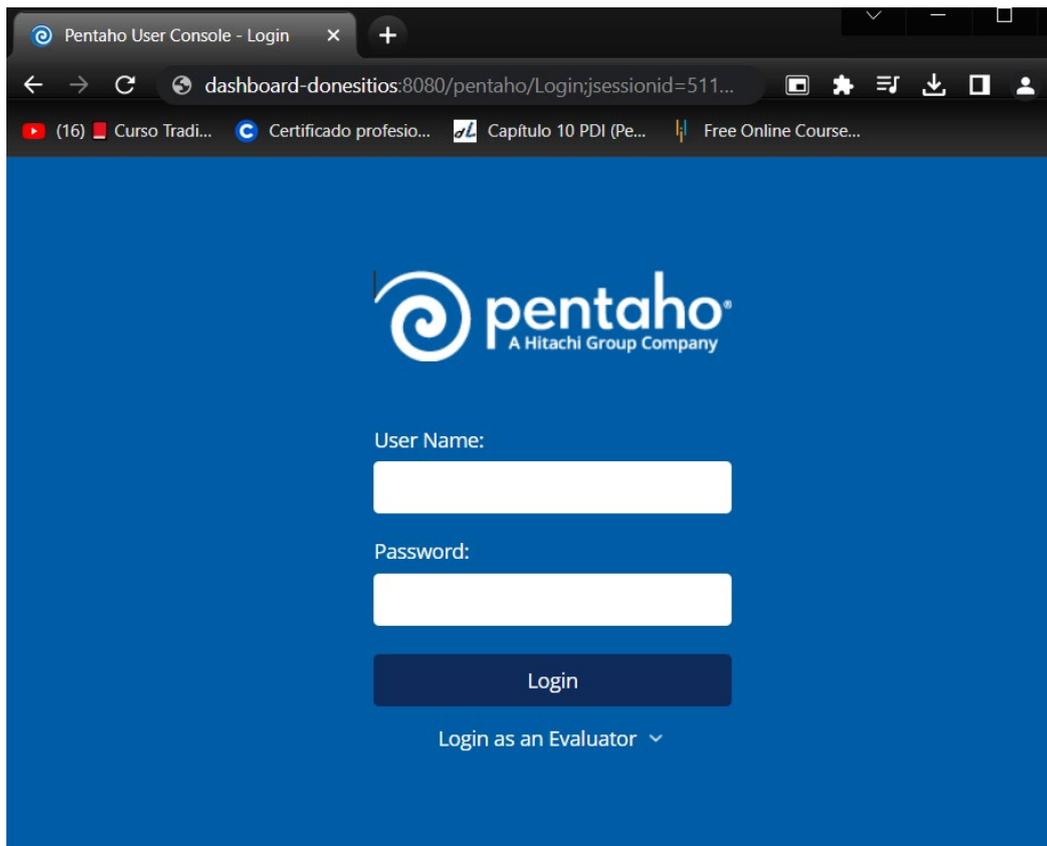


Figura 66. Login de Pentaho en el servidor

### 6.3.5 Pruebas de la solución BI

#### Pruebas de medidas de los cubos OLAP

A continuación, se comprobará el cálculo de las medidas del cubo OLAP. Para llevar a cabo las verificaciones, se calcularán manualmente los valores de las medidas de los cubos mediante consultas SQL desde el DW en PostgreSQL. Luego, se compararán estos valores con los obtenidos por las consultas MDX generadas en pentaho server.

Al realizar la consulta MDX mediante el cubo OLAP, se obtiene el resultado que se especifica en la.

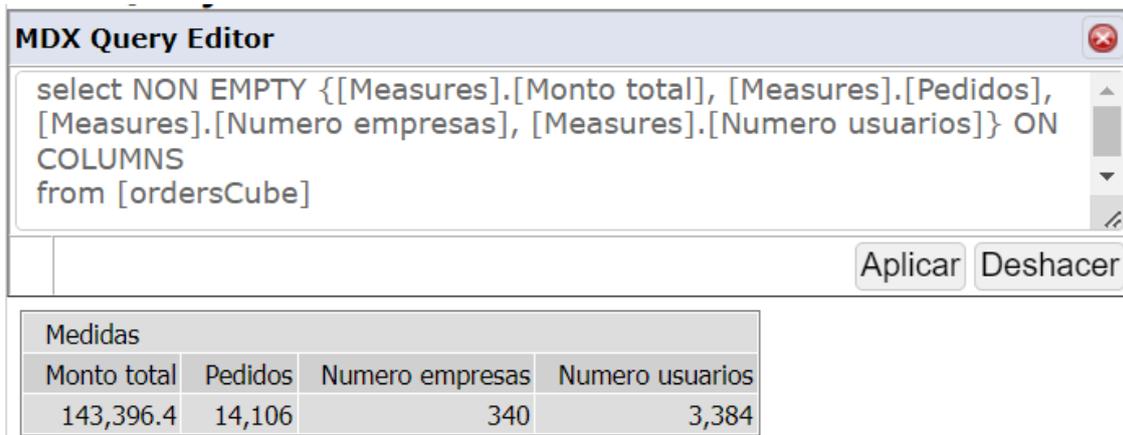


Figura 67. Pruebas de medidas por consulta MDX

La consulta SQL para obtener el cálculo de las medidas es la descrita en la Figura 68:

```
dw_donesitios=# select sum(produc__price_total) as Monto_total, count(distinct
order_code) as Pedidos, count(distinct store_code) as Empresas, count(distinct
user_code) as Usuarios  from stg_values;
monto_total | pedidos | empresas | usuarios
-----+-----+-----+-----
 143396.40 |   14106 |       340 |    3384
```

Figura 68. Pruebas de medidas por consulta SQL

A continuación, se presentan los resultados obtenidos para las consultas SQL y MDX realizadas

TABLA XVII. RESULTADO DE LAS PRUEBAS DEL CUBO OLAP

	Consulta SQL	Consulta MDX	Resultado
<b>Monto total</b>	143,396.4	143,396.4	Correcto
<b>Número de Pedidos</b>	14,106	14,106	Correcto
<b>Número de empresas</b>	340	340	Correcto
<b>Número de usuarios</b>	3,384	3,384	Correcto

Los resultados obtenidos muestran que ambas consultas arrojan los mismos valores para cada una de las medidas evaluadas, lo que indica que los datos del cubo OLAP están en concordancia con los datos del DW y que la implementación del cubo OLAP está funcionando correctamente.

### Evaluación de reportes

Para garantizar la precisión de los datos presentados en los informes, se definió una consulta que obtenga información de varias tablas. Se obtendrán los datos de dos formas: mediante una consulta SQL directa al Datawarehouse utilizando PostgreSQL y a través de los informes generados en Pentaho. Se compararán los resultados obtenidos por ambas consultas para garantizar su exactitud.

El reporte generado en Pentaho muestra a los usuarios que han realizado la mayor cantidad de pedidos y el monto total de sus compras. Los resultados obtenidos se presentan en la Figura 69.

Usuarios	Pedidos	Monto Total
Vanessa del Rocio Beltrán Cabrera	168	1287.94
Diego Fernando Bravo Lozano	141	1098.43
John Jairo Villavicencio Sarango	136	1103.49
Favio Alexander Bravo Lozano	131	2341.13
Briggette Vásquez Sarmiento	122	590.45
Jackson Masache	119	737.46
John Villavicencio	72	506.14
Usuario Normal Donesitios	69	350.31
Paola Gabriela Bravo Lozano	68	534.39
Kevin Bryan Armijos Maza	67	3024.65

Figura 69. Resultado reporte usuarios con más pedidos

La consulta SQL para obtener los usuarios con mayor cantidad de usuarios se presenta en la Figura 70, donde se muestra la consulta ejecutada y su resultado.

```
dw_donesitios=# select user_name, count(distinct order_code) as num_pedidos,
sum(produc_price_total) as total from dim_user inner join fact_table on dim_
user.pk_user = fact_table.fk_user inner join dim_order on fact_table.fk_order
= dim_order.pk_order group by user_name order by num_pedidos desc limit 10;
```

user_name	num_pedidos	total
Vanessa del Rocio Beltr�n Cabrera	168	1287.94
Diego Fernando Bravo Lozano	141	1098.43
John Jairo Villavicencio Sarango	136	1103.49
Favio Alexander Bravo Lozano	131	2341.13
Briggette Vsquez Sarmiento	122	590.45
Jackson Masache	119	737.46
John Villavicencio	72	506.14
Usuario Normal Donesitios	69	350.31
Paola Gabriela Bravo Lozano	68	534.39
Kevin Bryan Armijos Maza	67	3024.65

Figura 70. Pruebas de reportes por consulta SQL

Los datos presentados en la Figura 69 y Figura 70 coinciden en ambas consultas: la realizada manualmente a trav s de SQL y la realizada mediante Pentaho.

### Pruebas de Carga y estr s

Se llevaron a cabo pruebas de carga y estr s con el objetivo de evaluar el rendimiento de la visualizaci n de informaci n en el dashboard implementado en el servidor, utilizando diferentes cantidades de peticiones concurrentes: 50, 100, 500, 1000 y 2000. Los resultados obtenidos se detallan en la TABLA XVIII:

TABLA XVIII. RESULTADOS DE LA PRUEBA DE CARGA Y ESTR S

# Peticiones	Correctas	Err�neas	% Error	Tiempo m�nimo (ms)	Tiempo m�ximo (ms)
<b>50</b>	50	0	0%	714	2117
<b>100</b>	100	0	0%	848	5809
<b>500</b>	500	0	0%	342	7886
<b>1000</b>	996	4	0.40%	587	31257
<b>2000</b>	1949	51	2.55%	845	39231

En general, se pudo observar que el porcentaje de errores obtenidos fue muy bajo, siendo de 0% para las pruebas con 50, 100 y 500 peticiones concurrentes. No obstante, en las pruebas con 1000 y 2000 peticiones concurrentes se registraron algunos errores, alcanzando un porcentaje de 0.40% y 2.55% respectivamente.

En cuanto a los tiempos de respuesta, se observ  que el tiempo m nimo aument  ligeramente a medida que se increment  el n mero de peticiones. Los tiempos m ximos tambi n

aumentaron considerablemente en las pruebas con 1000 y 2000 peticiones, lo cual puede deberse a un mayor estrés en el sistema.

En resumen, los resultados de las pruebas indican que el acceso a la información en el dashboard puede soportar un número considerable de peticiones concurrentes sin mayores problemas. Los procesos realizados en el plan de pruebas se encuentran descritos en el Anexo 9. Ejecución de pruebas

### Pruebas de aceptación

Se realizó una encuesta al gerente de la empresa utilizando la técnica de medición propuesta por Likert para llevar a cabo una prueba de aceptación y evaluar si la solución BI cumple con sus expectativas. Los procesos realizados se encuentran descritos en el Anexo 9.

La Figura 71, muestra la respuesta del gerente en relación al cumplimiento de los requisitos solicitados, lo que indica que se han cumplido los requisitos de la solución BI.

1. ¿La solución BI para la gestión de datos de pedidos en la empresa cumple con los requerimientos solicitados?

1 respuesta

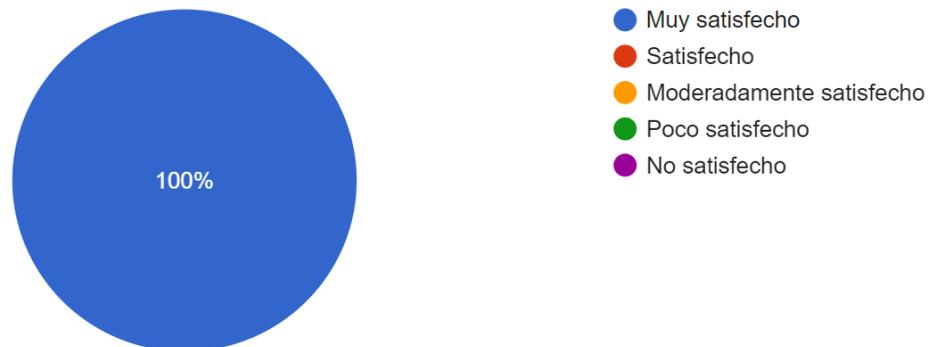


Figura 71. Respuesta del gerente a pregunta sobre el cumplimiento de los requisitos solicitados en la solución BI.

En la Figura 72, se evidencia el impacto del dashboard implementado en el proceso de toma de decisiones, lo que indica que la solución BI ha sido de gran ayuda para la empresa.

2. ¿Considera que el dashboard implementado ha tenido un impacto positivo en el proceso de toma de decisiones de la empresa?

1 respuesta

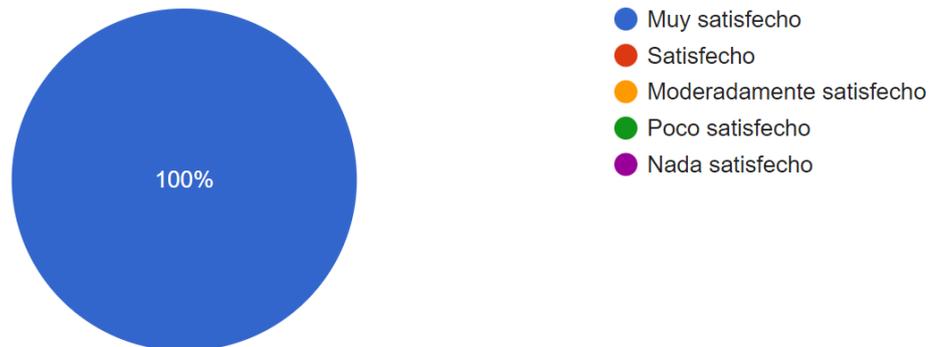


Figura 72. Respuesta del gerente a pregunta sobre el impacto de la solución BI en la toma de decisiones en la empresa.

La Figura 73 ilustra la respuesta del gerente de la empresa en relación a la interfaz amigable y la facilidad de uso de la solución BI. Este resultado indica que la solución BI es bien recibida en términos de interfaz y usabilidad del sistema.

3. ¿Considera usted que la solución BI presenta una interfaz de usuario amigable y de fácil usar?

1 respuesta

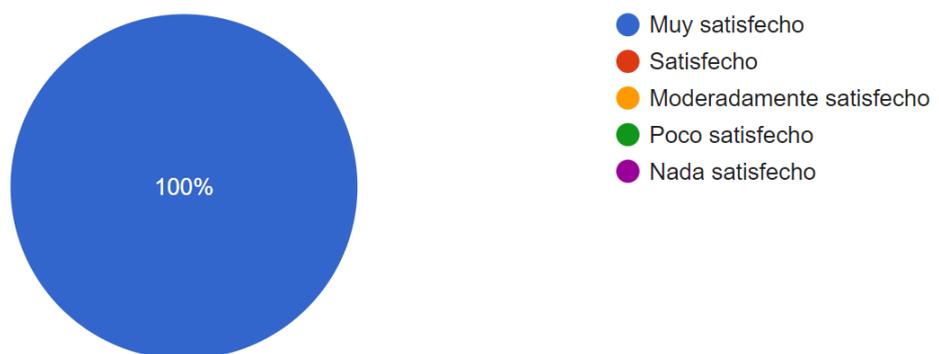


Figura 73. Respuesta del gerente a pregunta sobre el diseño de la interfaz amigable y su usabilidad

La Figura 74, presenta la respuesta del gerente sobre la información de valor mostrada en el dashboard. Por lo que se deduce que la información presentada en el dashboard es valiosa para la empresa

4. ¿La información presentada en el dashboard es útil para la empresa?

1 respuesta

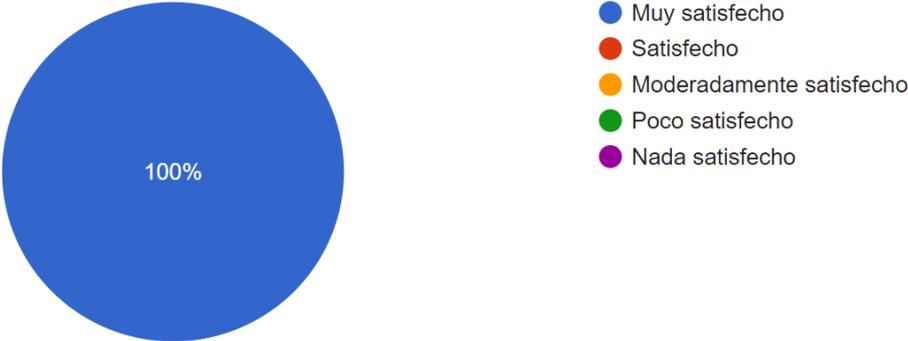


Figura 74. Respuesta del gerente a pregunta sobre la utilidad de información presentada en el dashboard

La Figura 75 representa la respuesta del gerente de la empresa acerca del tiempo de acceso a la información en el dashboard. El resultado indica que la solución BI es aceptada en términos de tiempo.

5. ¿Considera que los tiempos de respuesta del dashboard son aceptables?

1 respuesta

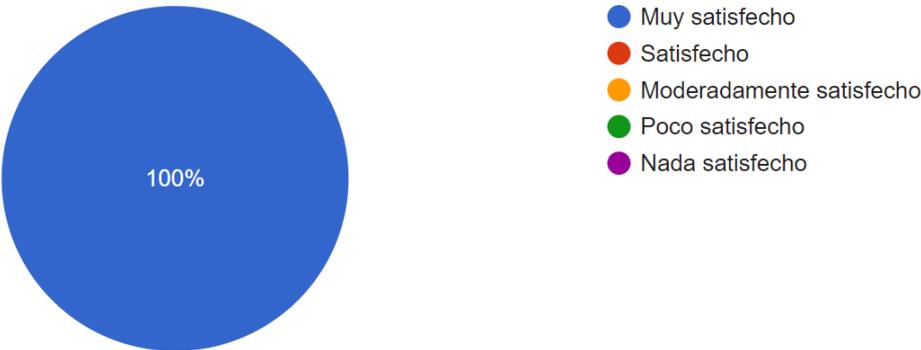


Figura 75. Respuesta del gerente a pregunta sobre el tiempo de acceso al dashboard

Según la respuesta del gerente en la Figura 76, la implementación de la solución BI ha permitido mejorar la gestión de los pedidos de la empresa.

6. ¿La solución BI ha permitido mejorar la gestión de pedidos en la empresa?

1 respuesta

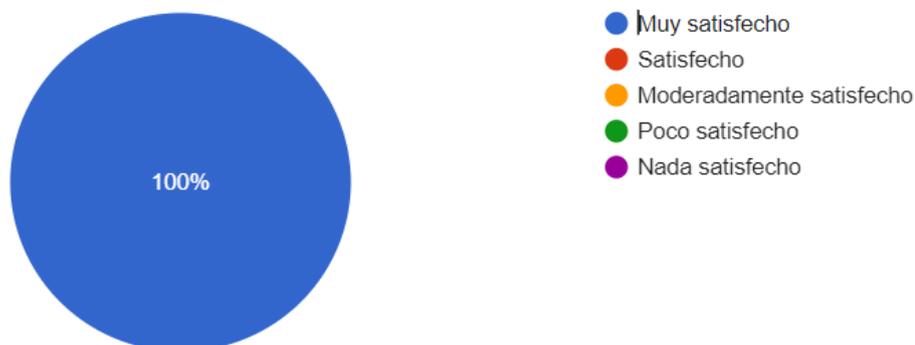


Figura 76. Respuesta del gerente a pregunta sobre la mejora en la gestión de pedidos

En base a los resultados obtenidos, la solución BI ha permitido al gerente:

- Involucrar a los trabajadores en los resultados de la empresa.
- Gestionar empresas que cuenten con productos de gran demanda en el sistema.
- Monitorear el cumplimiento y desempeño de las empresas
- Desarrollar nuevos servicios
- Identificar nuevas oportunidades de mercado.

Finalmente, Marketing Digital Bravo & Beltrán ha validado y aceptado la solución BI desarrollada en el presente TT. En el Anexo 10, se evidencia la aceptación de la solución BI implementada.

## 7. Discusión

### 7.1 Desarrollo de la propuesta alternativa

Para responder a la pregunta de investigación: ¿La implementación de una herramienta para la gestión de datos mejora la toma de decisiones en el módulo de pedidos en la empresa? Se llevó a cabo una encuesta al gerente de la empresa, en la que se consultó el nivel de satisfacción con el dashboard desarrollado.

En la Figura 77, se muestra el nivel de satisfacción del gerente de la empresa Marketing Digital Bravo & Beltrán con el dashboard desarrollado. Se obtuvieron estos resultados a partir de una encuesta en la que se consultó sobre el impacto, y la contribución de las distintas funcionalidades del dashboard en el proceso de toma de decisiones.

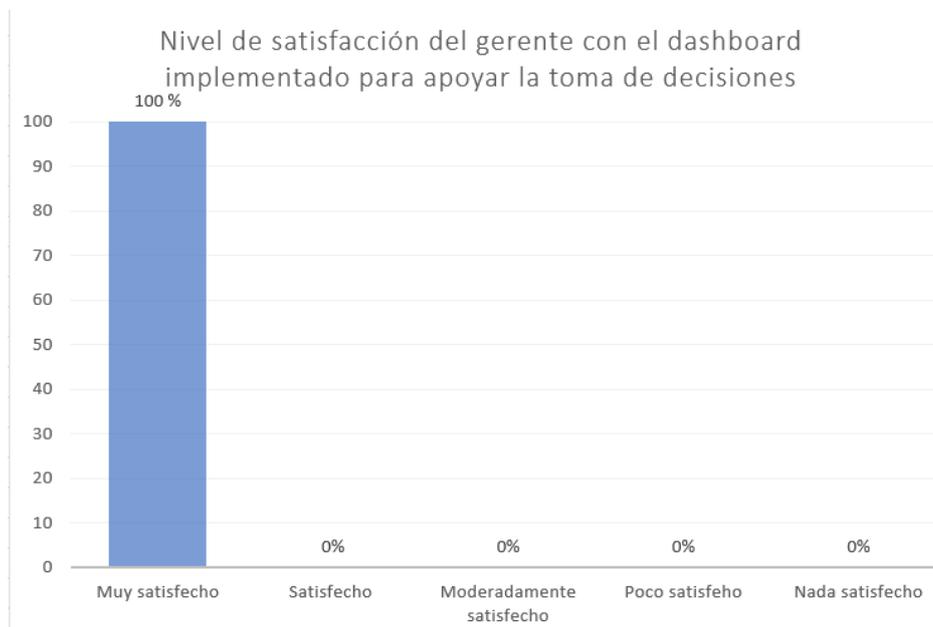


Figura 77. Resultados consolidados de preguntas relacionadas con la satisfacción del gerente con respecto al dashboard desarrollado.

El resultado obtenido fue una satisfacción del 100% por parte del gerente con la solución BI implementada. Esto indica que la solución generó un impacto positivo en el proceso de toma de decisiones de la empresa.

A diferencia de otras soluciones como las que se observa en [29], [17], [27] si bien desarrollan una solución BI, no presentan evidencia de como esta contribuye en el proceso de toma de decisiones. Por el contrario, en el presente TT, mediante una encuesta aplicada, se validó la satisfacción del gerente de la empresa con el dashboard implementado y la información que se presenta en este. De esta manera, se puede afirmar que el dashboard implementado sirve

como una herramienta para la gestión de información y que además brinda soporte en el proceso de toma decisiones informadas.

**OBJETIVO 1: Analizar los datos históricos generados en el módulo de pedidos para establecer el diseño del modelo.**

El estándar IEEE 830 para la recolección de requisitos en el presente TT proporcionó una estructura estandarizada para la documentación de requisitos, lo que permitió determinar 16 requisitos funcionales y 6 requisitos no funcionales de manera clara y completa (ver Anexo 2. Especificación de requisitos). Esta etapa se diferencia de otros proyectos, como se observa en [17], [34], donde la obtención de los requisitos se realiza sin una planificación y no presenta una documentación clara. El uso del estándar IEEE 830 ayudó a reducir la probabilidad de malentendidos o interpretaciones erróneas de los requisitos, que pudieron conducir a errores costosos en el proceso de desarrollo de software.

Durante el proceso 6.1.3 Diseño de la arquitectura, se estableció la arquitectura en Back Room y Front Room para la solución de Business Intelligence. Esta arquitectura permitió separar las actividades de procesamiento y transformación de datos (Back Room) de la presentación de información (Front Room), logrando una solución eficaz y eficiente que cumple con los requisitos de información propuestos. La implementación de esta arquitectura fue beneficiosa, ya que proporcionó una estructura clara y bien definida para el desarrollo del sistema de BI. Además, esta arquitectura es consistente con las propuestas presentada por [34]. En contraste, la arquitectura especificada en [38], podría presentar problemas de desempeño y escalabilidad en sistemas con grandes volúmenes de datos.

Una de las ventajas de la solución de Business Intelligence implementada en la empresa es la elección de herramientas de código abierto, como la suite de Pentaho. Al utilizar herramientas de código abierto, se evitó incurrir en costos de adquisición y licencias, lo que hizo que esta opción sea más viable para el proyecto. Este enfoque también está respaldado en [29], donde se menciona que trabajar con herramientas de código abierto es una alternativa rentable y efectiva para la implementación de soluciones de Business Intelligence en empresas medianas.

**OBJETIVO 2: Desarrollar el modelo de datos utilizando la plataforma de BI Pentaho.**

Durante la etapa 6.2.1 Modelado Dimensional, se realizó un análisis de los requerimientos con el que se logró determinar las dimensiones y medidas. A diferencia del enfoque utilizado en [17], [39], [38], en este TT se utilizó una matriz de bus para definir las medidas para la tabla de hecho y un Diagrama StarNet para representar el modelo dimensional. Estas técnicas permitieron desarrollar un modelo dimensional preciso y completo, que ayudó a analizar los datos de manera efectiva y tomar decisiones informadas.

La implementación de procesos ETL se llevó a cabo en la sección 6.2.4 Implementación de procesos ETL, utilizando PDI que es una de las opciones más populares en la industria para la construcción y procesamiento de datos, como se puede evidenciar en su amplio uso en diferentes implementaciones [17], [34], [41]. Durante el proceso de extracción, se cargaron los datos desde un archivo CSV, puesto que se necesitaba una carga de datos simple y rápida, además, los archivos CSV son fáciles de generar y editar y se pueden cargar en PDI fácilmente. En el proceso de transformación, se realizaron procesos para eliminar duplicados, reemplazar y corregir registros, lo que garantizó la calidad de los datos. Finalmente, el proceso de carga de los datos transformados se realizó de manera rápida y eficiente en el Datamart.

Para llevar a cabo el análisis de los datos en el Datamart se optó por utilizar cubos OLAP. Esta elección permitió llevar a cabo un análisis multidimensional que permitió explorar los datos desde diversas perspectivas y niveles de agregación, lo que proporcionó mayor flexibilidad en el análisis de datos. Es así que se utilizó la herramienta Schema Workbench para la creación del cubo multidimensional. Es importante destacar que el proceso de construcción del cubo multidimensional fue una etapa clave en el proceso de análisis de datos, ya que permitió visualizar la información de manera estructurada y eficiente. La definición de jerarquías y la utilización de funciones adecuadas para los hechos fueron fundamentales para obtener resultados precisos y significativos.

### **OBJETIVO 3: Implementar una solución de BI a través de un Dashboard.**

En el presente TT, se optó por desarrollar un dashboard como solución de presentación de información. La elección de este se justifica por presentar una visualización clara y concisa de grandes cantidades de datos en un solo lugar, la accesibilidad desde cualquier lugar y en cualquier momento, la interactividad y la capacidad de mostrar información en tiempo real, permitiendo una rápida toma de decisiones al gerente de la empresa.

Después de completar la construcción del dashboard, en contraste con [17], [34] se llevó a cabo el despliegue de la solución en un servidor de aplicaciones. Al desplegar la solución en el servidor de aplicaciones, se aseguró que el dashboard estuviera disponible en cualquier momento y en cualquier dispositivo. Esto permitió que la solución fuera accesible de manera remota y escalable, lo que significa que se puede adaptar a diferentes niveles de tráfico y usuarios. Además, se implementaron políticas de acceso, para proteger la información y garantizar que solo los usuarios autorizados tengan acceso al dashboard.

Otras soluciones como las que se observa en los trabajos relacionados [27], [39], [34] si bien realizan la implementación de un dashboard para la gestión de datos, no realizan pruebas para verificar la implementación. Por el contrario, en el presente TT se llevaron a cabo

diferentes tipos de pruebas. Se realizaron pruebas al cubo OLAP para verificar la correcta implementación de los mismos, también se realizaron pruebas a los reportes para verificar que la información presentada en el dashboard era la misma que la del DW. Los resultados obtenidos en ambas pruebas fueron positivos, lo que indica que el cubo OLAP y los reportes creados presentan la información correcta. Además, se llevaron a cabo pruebas de carga y estrés para evaluar el rendimiento de la solución implementada. Para ello, en Jmeter se realizaron pruebas con diferentes números de peticiones por segundo. Los resultados obtenidos demostraron que la solución puede soportar un número considerable de peticiones concurrentes sin mayores problemas. Sin embargo, se observó un margen de error del 0.4% y 2.55% para 1000 y 2000 peticiones por segundo, respectivamente, lo que podría deberse a un mayor estrés en el sistema. Finalmente, mediante pruebas de aceptación se validó la satisfacción del gerente con el dashboard implementado. De esta manera se confirmó que la implementación de una solución BI mejora el proceso de toma de decisiones en la empresa.

Por lo tanto, la solución propuesta es la mejor opción ya que existen evidencias que garantizan una mejora en el proceso de toma de decisiones. Además, esto fue validado por el gerente de la empresa. En contraste con otros trabajos relacionados, este TT cubre ciertas carencias al incluir un estándar definido para la especificación de requisitos, como el IEEE 830, así como la implementación de la solución BI en un entorno de producción. Además, se llevaron a cabo diferentes pruebas, incluyendo pruebas al cubo OLAP, pruebas de carga y estrés, y pruebas de aceptación, lo que permitió confirmar la correcta implementación del dashboard.

## **7.2 Valoración técnica, económica y social**

El desarrollo del presente TT se expresa detallando los beneficios prestados desde tres aspectos que se detallan a continuación:

### **7.2.1 Valoración Técnica**

Se utilizaron diferentes recursos hardware y software que fueron indispensables para dar cumplimiento con las actividades de cada objetivo planteado. Las herramientas utilizadas son las descritas a continuación:

- Plataforma de videoconferencia Zoom, utilizada para la comunicación con los actores involucrados.
- Herramienta de diagramado Draw.io, empleada para la creación de diagramas, modelos y procesos.
- Software Mendeley Desktop, para la administración de artículos científicos seleccionados

como referencias.

- Repositorio GitHub, para el almacenamiento de los archivos creados durante la implementación de la solución BI
- Pentaho Data Integration, permitió la implementación de procesos ETL.
- Schema Workbench, para la creación de cubos de información utilizados en el análisis multidimensional.
- Pentaho Server, empleado para la implementación de reportes y los Dashboard.
- Lenguaje de consulta MDX, para la obtención de datos de una base de datos multidimensional.
- Base de datos PostgreSQL, utilizado en un entorno de desarrollo para el almacenamiento del Datawarehouse.
- Digital Ocean, utilizado para alojar el dashboard desarrollado en un ambiente de producción.
- Jmeter, empleado para llevar a cabo las pruebas de carga y estrés al dashboard.

## 7.2.2 Valoración Económica

En esta sección se indica el presupuesto empleado durante del TT. Se requirió inversión de talento humano, hardware y software, y servicios.

### Talento Humano

En la TABLA XIX se detalla el tiempo y costo estimado de los actores involucrados en el TT. La asesoría del docente como tutor, es asumida por la Universidad Nacional de Loja.

TABLA XIX. VALORACIÓN ECONÓMICA: TALENTO HUMANO

Talento Humano	Número de Horas	Costo por Hora	Costo Total
estudiante	400	\$ 8	\$ 3.200,00
Director del Proyecto	50	-	-
Director TI de la empresa	5	\$ 12,50	\$ 62,50
<b>TOTAL</b>			<b>\$ 3262,50</b>

### Recursos Hardware y Software

Los recursos hardware y software utilizados en el desarrollo del TT son los especificados a continuación en la TABLA XX.

TABLA XX. VALORACIÓN ECONÓMICA: HARDWARE Y SOFTWARE

<b>Recursos Software</b>				
<b>Nombre</b>		<b>Costo Total</b>		
Microsoft Office		\$ 0,00		
Zoom		\$ 0,00		
GitHub		\$ 0,00		
Mendeley Desktop		\$ 0,00		
Google Chrome		\$ 0,00		
Draw.io		\$ 0,00		
Pentaho Data Integration		\$ 0,00		
Google Forms		\$ 0,00		
Schema Workbench		\$ 0,00		
Pentaho Server		\$ 0,00		
PostgreSQL		\$ 0,00		
DigitalOcean		\$ 36,00		
Apache Jmeter		\$ 0,00		
Firma Electrónica		\$ 34,00		
<b>Subtotal</b>		<b>\$ 70,00</b>		
<b>Recursos Hardware</b>				
<b>Cantidad</b>	<b>Nombre</b>	<b>Número de meses</b>	<b>Valor</b>	<b>Costo Total</b>
1	Laptop personal	5	\$ 30,00	\$ 150,00
<b>Subtotal</b>		<b>\$ 150,00</b>		
<b>TOTAL</b>				<b>\$ 220,00</b>

## Servicios

En la TABLA XXI se detallan los servicios y su valor económico que fueron necesarios durante la ejecución del TT.

TABLA XXI. VALORACIÓN ECONÓMICA: SERVICIOS

<b>Nombre</b>	<b>Meses</b>	<b>Costo Unitario</b>	<b>Costo Total</b>
Internet	5	\$ 25,00	\$ 125,00
Transporte	5	\$ 10,00	\$ 50,00
<b>TOTAL</b>			<b>\$ 175,00</b>

## Presupuesto Final

De acuerdo a los recursos económicos anteriormente presentados, se presenta la TABLA XXII, la cual contiene la sumatoria década uno de ellos. Además, se agregó un 10% del valor total para imprevistos.

TABLA XXII. PRESUPUESTO FINAL

<b>Recurso</b>	<b>Subtotal</b>
Recursos Humanos	\$ 3.262,50
Recursos Hardware y Software	\$ 220,00

Servicios	\$ 175,00
<b>Subtotal</b>	<b>\$ 3.657,50</b>
Imprevistos (10% del subtotal)	\$ 365.75
<b>TOTAL DEL PROYECTO</b>	<b>\$ 4.023,25</b>

### 7.2.2 Valoración Ambiental

Con respecto a la valoración ambiental, el TT se ejecutó en su totalidad con recursos digitales y tecnológicos, los cuales no tienen un mayor impacto al entorno ambiental, además se tuvo un bajo consumo de recursos materiales o de otros elementos que puedan llegar a dañar al planeta.

## 8. Conclusiones

Esta sección describe los eventos más relevantes encontrados en la ejecución del TT:

- La implementación de un Dashboard basado en Business Intelligence representa una inversión estratégica para la empresa Marketing Digital Bravo & Beltrán. Al proporcionar una visualización clara y accesible de los datos relevantes del módulo de pedidos, el Dashboard permite al gerente tomar decisiones más informadas, acelerar los procesos de negocio y mejorar la eficiencia operativa en general.
- La metodología de Ralph Kimball resultó altamente eficaz en la implementación de Business Intelligence. Siguiendo sus directrices, se obtuvo una comprensión precisa de todas las fases del proyecto, lo que condujo a una ejecución exitosa y puntual del proyecto. Esta metodología ofreció tanto una estructura robusta como un enfoque gradual para el diseño e implementación de la solución BI, lo que ayudó a minimizar errores y cambios significativos al concluir el proyecto.
- La ejecución de procesos ETL fue esencial para la integridad y la calidad de los datos en la empresa. La disponibilidad de información precisa y actualizada es crucial para la toma de decisiones estratégicas y la eficiencia operativa en general. Por lo tanto, es importante garantizar que la información en el Datawarehouse sea confiable y esté libre de errores antes de implementar soluciones BI.
- La correcta definición de la arquitectura de la solución de Business Intelligence fue fundamental para la construcción de un sistema independiente entre la lógica de negocios y la visualización de datos. Esto permitió una mayor flexibilidad en la gestión y la presentación de información.
- La selección de las herramientas adecuadas fue un aspecto fundamental para garantizar la productividad en el desarrollo de la solución y el éxito de los resultados esperados. Las herramientas y librerías OpenSource de la suite de Pentaho, permitieron la extracción, gestión y visualización de los datos de manera efectiva.
- La solución BI implementada ha logrado generar notables mejoras en cuanto al acceso a la información para la empresa Marketing Digital Bravo & Beltrán. Esta implementación ha resultado en ahorros de recursos significativos al mismo tiempo que proporciona una herramienta accesible en todo momento, brindando información y respaldo para el proceso de toma de decisiones.

## 9. Recomendaciones

Después de completar el TT, las recomendaciones son:

- Realizar un análisis exhaustivo de los datos disponibles en las fuentes de datos de la empresa para identificar y definir los indicadores clave que permitan medir y monitorear el desempeño de la empresa de manera efectiva.
- Establecer reuniones periódicas con los diferentes actores involucrados en el proceso de toma de decisiones, con el objetivo de identificar sus necesidades y requerimientos de información, y así brindar la información adecuada y oportuna para su toma de decisiones.
- Seleccionar herramientas adecuadas para la implementación de BI, evaluando su popularidad en el mercado, factibilidad de adquisición y facilidad de instalación y desarrollo.
- Utilizar una metodología de trabajo clara y definida, como la propuesta por Ralph Kimball, que permita seguir de manera secuencial las actividades necesarias durante la implementación de Business Intelligence, lo que garantiza una mejor planificación y evita cambios abruptos en el proyecto.
- Implementar Soluciones BI siempre y cuando la fuente de datos de origen contenga un gran volumen de registros, lo que permitirá identificar de manera efectiva patrones y tendencias clave para la toma de decisiones informada

## 10. Bibliografía

- [1] J. Espinoza and C. Palomino, "Desarrollo de un datamart para optimizar la generación de información estratégica de apoyo a la toma de decisiones en la vicepresidencia de banca comercial de interbank Perú (Tesis de pregrado)," 2016.
- [2] H. Wang and S. Wang, "A knowledge management approach to data mining process for business intelligence," *Ind. Manag. Data Syst.*, vol. 108, no. 5, pp. 622–634, 2008, doi: 10.1108/02635570810876750.
- [3] A. Pourshahid, I. Johari, G. Richards, D. Amyot, and O. S. Akhigbe, "A goal-oriented, business intelligence-supported decision-making methodology," pp. 1–36, 2014, doi: doi:10.1186/s40165-014-0009-8.
- [4] J. Ranjan, "Business Intelligence: Concepts, components, techniques and benefits," *J. Theor. Appl. Inf. Technol.*, vol. 9, no. 1, pp. 60–70, 2009.
- [5] Y. Levy and M. M. Ramim, "A Study of Online Exams Procrastination Using Data Analytics Techniques," vol. 8, 2012, [Online]. Available: [http://www.openu.ac.il/research\\_center/chais2011/papers.html](http://www.openu.ac.il/research_center/chais2011/papers.html).
- [6] H. Chen, R. H.L.Chiang, and V. C. Storey, "Business Intelligence and Analytics: From Big Data To Big Impact," *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2018, [Online]. Available: <http://www.jstor.org/stable/41703503>.
- [7] M. Kasem and E. E. Hassanein, "Cloud Business Intelligence Survey," *Int. J. Comput. Appl.*, vol. 90, no. 1, pp. 23–28, 2014, doi: 10.5120/15540-4266.
- [8] A. Saxena, "A Survey on Big Data Analytics: Challenges, Open Research Issues and Tools," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 10, no. 12, pp. 1899–1908, 2022, doi: 10.22214/ijraset.2022.48294.
- [9] D. Suša Vugec, V. Bosilj Vukšić, M. Pejić Bach, J. Jaklič, and M. Indihar Štemberger, "Business intelligence and organizational performance: The role of alignment with business process management," *Bus. Process Manag. J.*, vol. 26, no. 6, pp. 1709–1730, 2020, doi: 10.1108/BPMJ-08-2019-0342.
- [10] K. Javier, "Higher Education Meets Business Intelligence," *West. Oregon Univ.*, p. 30, 2017, [Online]. Available: <https://digitalcommons.wou.edu/cgi/viewcontent.cgi?article=1143&context=theses>.
- [11] C. Lago, "150 years of business intelligence: A brief history," 2018. <https://www.cio.com/article/221963/history-of-business-intelligence.html> (accessed Aug. 06, 2022).

- [12] H. N. Castro Zúñiga and E. M. Vera Baren, "Elaboración de una herramienta Business Intelligence para el análisis de los datos generados por la autogestión de una cooperativa de transporte público del cantón Milagro," *Sustain.*, vol. 4, no. 1, pp. 1–56, 2020, [Online]. Available: <https://pesquisa.bvsalud.org/portal/resource/en/mdl-20203177951%0Ahttp://dx.doi.org/10.1038/s41562-020-0887-9%0Ahttp://dx.doi.org/10.1038/s41562-020-0884-z%0Ahttps://doi.org/10.1080/13669877.2020.1758193%0Ahttp://sersc.org/journals/index.php/IJAST/article>.
- [13] M. J. Murillo Junco and G. Cáceres Castellanos, "Business intelligence y la toma de decisiones financieras: una aproximación teórica\*," *Logos Cienc. Tecnol.*, vol. 5, no. 1, pp. 119–138, 2013, [Online]. Available: <https://www.redalyc.org/pdf/5177/517751547010.pdf>.
- [14] A. M. Salazar Montalvan, "Implementación de una solución de Business Intelligence como apoyo a la toma de decisiones en el proceso de mantenimiento de servicios de clientes de la empresa Claro en el area de instalación & mantenimiento HFC Chiclayo," p. 178, 2019, [Online]. Available: <http://tesis.usat.edu.pe/handle/20.500.12423/2068>.
- [15] J. Conesa Caralt and J. Curto Díaz, *Introducción al Business Intelligence*, El Ciervo., vol. 163. Barcelona, 2012.
- [16] L. Cano, *Business Intelligence: Competir Con Información*. Madrid: Banesto, Fundación Cultur, 2007.
- [17] V. N. Bustos Barrera, Sofía Anabel; Mosquera Artieda, "Análisis, Diseño E Implementación De Una Solución Business Intelligence Para La Generación De Indicadores Y Control De Desempeño, En La Empresa Otecel S.a, Utilizando La Metodología Hefesto V2.0," *Esc. Politécnica del Ejército..*, p. 176, 2013, [Online]. Available: <http://repositorio.espe.edu.ec/handle/21000/6305>.
- [18] R. D. Bernabeu, "Hefesto. Data Warehousing: Investigación y Sistematización de Conceptos. Hefesto: Metodología para la construcción de un datawarehouse," p. 146, 2010, [Online]. Available: <https://www.businessintelligence.info/resources/assets/hefesto-v2.pdf>.
- [19] V. Karina, "Elementos De Inteligencia De Negocio ( Bi ) Aplicada al analisis de datos en la cooperativa de taxis ' Benjamín Carrión ' De La Ciudad De Loja ." Loja, 2015.
- [20] K. Salmasi, "Identification and classification of organizational level competencies for BI success," *J. Intell. Stud. Bus.*, vol. 6, no. 1, pp. 48–58, 2016, [Online]. Available: <https://ojs.hh.se/index.php/JISIB/article/view/142>.

- [21] J. M. V. García and B. H. D. Pinzón, "Key success factors to business intelligence solution implementation," *J. Intell. Stud. Bus.*, vol. 7, no. 1, pp. 48–69, 2017, doi: 10.37380/jisib.v7i1.215.
- [22] W. H. Inmon, *Building the Data Warehouse: Timely. Practical. Reliable.*, vol. 13, no. 401. 2005.
- [23] R. Kimball, L. Reeves, M. Ross, and W. Thornthwaite, "The Data Warehouse Lifecycle Toolkit Table of Contents," *Architecture*, pp. 1–405, 2008.
- [24] A. Simitsis and P. Vassiliadis, "A method for the mapping of conceptual designs to logical blueprints for ETL processes," *Decis. Support Syst.*, vol. 45, no. 1, pp. 22–40, 2008, doi: 10.1016/j.dss.2006.12.002.
- [25] A. Bustamante Martínez, E. Amaru, G. Lista, and L. C. Gómez Flórez, "ETL Processes modeling techniques: an alternatives review and its application in a BI solution development project," *Sci. Tech. Año XVIII*, vol. 18, no. 1, pp. 185–191, 2013, [Online]. Available: <https://www.redalyc.org/pdf/849/84927487027.pdf>.
- [26] L. González, "Inteligencia de Negocios - Business Intelligence (BI)," p. 27, 2012, [Online]. Available: <https://basesdatoscms.files.wordpress.com/2012/10/resumen-businessintelligence.pdf>.
- [27] R. Akbar, M. Silvana, M. H. Hersyah, and M. Jannah, "Implementation of Business Intelligence for Sales Data Management Using Interactive Dashboard Visualization in XYZ Stores," *2020 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2020 - Proc.*, pp. 242–249, 2020, doi: 10.1109/ICITSI50517.2020.9264984.
- [28] Gravatar, "Pentaho: todo en una sola plataforma," 2020. <https://gravitar.biz/pentaho/> (accessed Dec. 05, 2021).
- [29] N. Leite, I. Pedrosa, and J. Bernardino, "Open Source Business Intelligence on a SME: A Case Study using Pentaho," *Iber. Conf. Inf. Syst. Technol. Cist.*, vol. 2019-June, no. June, pp. 1–7, 2019, doi: 10.23919/CISTI.2019.8760740.
- [30] Pentaho Corporation, "Pentaho Enterprise and Community Editions Feature Comparison." p. 6, 2008.
- [31] Kimball Group, "Kimball DW/BI Lifecycle Methodology," 2018. <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dw-bi-lifecycle-method/>.
- [32] G. R. Rivadera, "La metodología de Kimball para el diseño de almacenes de datos (Data warehouses)," *Cuad. la Fac.*, vol. 5, pp. 56–71, 2018, [Online]. Available:

<http://www1.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadere-formateado.pdf>.

- [33] R. Kimball and M. Ross, *The Data Warehouse Toolkit*, Third. United States of America, 2013.
- [34] W. Quillupangui, "Implementación de Business Intelligence como base para la gestión de información de la unidad educativa (SICS)," p. 105, 2013.
- [35] J. L. Lozano Pinchi, "Implementación de una solución Business Intelligence para apoyar en la toma de decisiones en la Empresa Agro Micro Biotech SAC," *Univ. Priv. Antenor Orrego*, 2019.
- [36] A. Pinto, "Universidad técnica del norte instituto de postgrado maestría en tecnología e innovación educativa," pp. 1–3, 2019.
- [37] M. Puonti, J. Järvi, and T. Mikkonen, "A continuous delivery framework for business intelligence," *Front. Artif. Intell. Appl.*, vol. 301, no. February, pp. 248–262, 2018, doi: 10.3233/978-1-61499-834-1-248.
- [38] W. Castillo-Rojas, Q. F. Medina, and M. F. Farina, "Proposal of methodology for a data Warehousing process: Use case: Generation of indicators of academic productivity of a university)," pp. 1–10, 2017, doi: 10.1109/cimps.2017.8169943.
- [39] J. Gubalova, "The use of Business Intelligence tools for leadership and university administration," *ICETA 2015 - 13th IEEE Int. Conf. Emerg. eLearning Technol. Appl. Proc.*, no. November 2015, 2016, doi: 10.1109/ICETA.2015.7558464.
- [40] L. Ortiz and M. Hallo, "Analytical Data Mart for the Monitoring of University Accreditation Indicators," *EDUNINE 2019 - 3rd IEEE World Eng. Educ. Conf. Mod. Educ. Paradig. Comput. Eng. Career, Proc.*, pp. 1–6, 2019, doi: 10.1109/EDUNINE.2019.8875826.
- [41] E. V. ARROBO LAPO and E. F. LOZADA TORRES, "BUSINESS INTELLIGENCE PARA LA GENERACIÓN DE INDICADORES Y CONTROL DE DESEMPEÑO, EN EL ÁREA DE COMPRAS AGRÍCOLAS DE LA EMPRESA 'PRONACA'.," *J. Chem. Inf. Model.*, vol. 21, no. 2, pp. 1689–1699, 2017, [Online]. Available: [https://www.oecd.org/dac/accountable-effective-institutions/Governance Notebook 2.6 Smoke.pdf](https://www.oecd.org/dac/accountable-effective-institutions/Governance%20Notebook%202.6%20Smoke.pdf).
- [42] Marketing Digital Bravo & Beltrán, "Donesitios," 2021. <https://www.donesitios.com/>.
- [43] S. Engineering and S. Committee, *IEEE recommended practice for software requirements specifications*, vol. 1998, no. October. New York, 2011.

## 11. Anexos

### Anexo 1. Análisis Empresarial

Sobre la empresa

La empresa Marketing Digital Bravo & Beltrán es una empresa fundada en la ciudad de Loja-Ecuador, se enfoca especialmente en el desarrollo de software, brindando soluciones tecnológicas de calidad a los problemas de las personas. Su principal producto a destacar es D'Onesitios, una aplicación móvil multiplataforma capaz de brindar solución a múltiples necesidades de los usuarios, a través de ella los usuarios pueden hacer uso de diferentes servicios como entregas a domicilio, solicitar taxis, y realizar pagos [42]. En la Figura 78 se muestra el logo de la empresa.



Figura 78. Logo empresarial.

#### **Misión:**

Marketing Digital Bravo & Beltrán, es una empresa especializada en el diseño, desarrollo e implementación de estrategias efectivas de Software. Su misión es ofrecer soluciones avanzadas que permitan mejorar la competitividad y productividad, logrando el crecimiento organizacional de sus clientes.

Los servicios profesionales que ofrecen se distinguen por la atención personalizada a los clientes, aplicando las mejores prácticas, procesos y tecnologías disponibles en el mercado.

#### **Visión:**

Consolidarse como la empresa líder en el mercado de servicios de Software, manteniendo un nivel de exigencia con procesos de mejora continua, basados en la innovación, calidad y servicio al cliente. Lo anterior marca un diferencial único que identifica y a la vez hace más competitiva a la empresa en el mercado.

#### **Organigrama empresarial:**

En la Figura 79, se ilustra la estructura organizacional de la empresa en la que se desarrolló la solución BI.

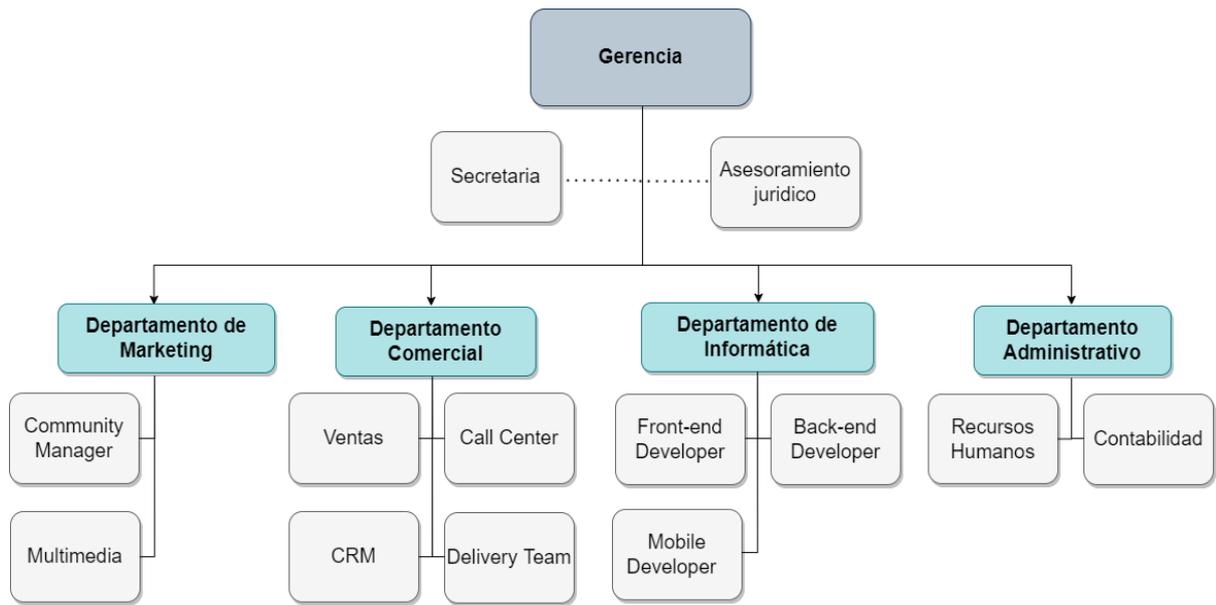


Figura 79. Organigrama Empresarial

**Procesos:**

El principal proceso que se lleva a cabo es ventas a través de pedidos.

- **Solicitud de pedido**

Lo realizan los usuarios a través de la aplicación móvil para adquirir los productos que requieren. En la Figura 80, se muestra las actividades necesarias para realizar un pedido desde la solicitud hasta la entrega al usuario.

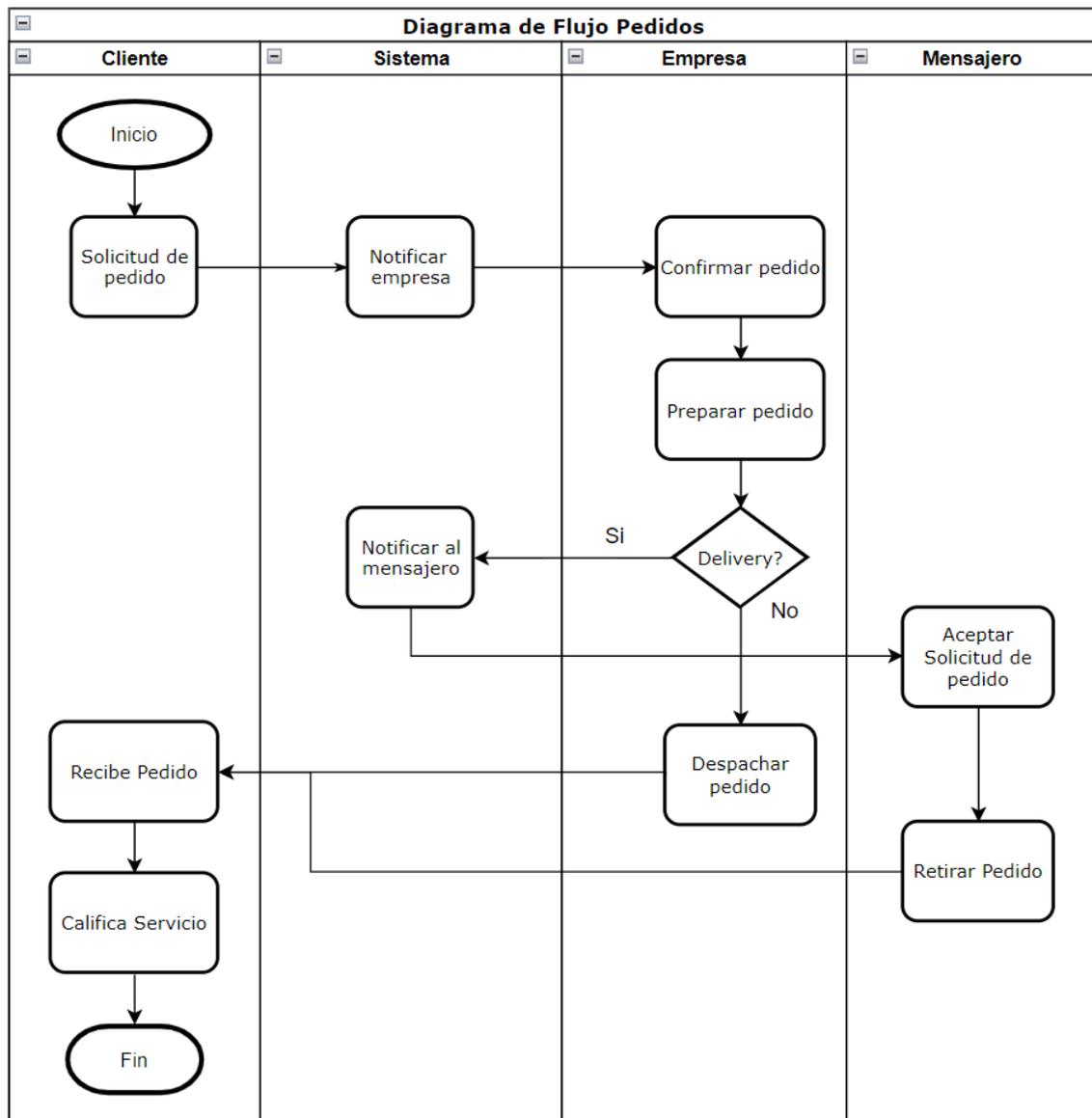


Figura 80. Diagrama de procesos pedidos.

## **Anexo 2. Especificación de requisitos**

# ESPECIFICACION DE REQUISITOS DE SOFTWARE

**Proyecto: "Implementación de BI para la gestión de datos generados en el módulo de pedidos en la empresa Marketing Digital Bravo & Beltrán"**

### **Historial de revisiones**

<b>Versión</b>	<b>Fecha</b>	<b>Descripción</b>	<b>Autor</b>
1.0	27-01-22	Emisión inicial	Luis Miguel Quito Carrión
2.0	03-02-22	Corrección de requisitos	Luis Miguel Quito Carrión
3.0	15-02-22	Revisión final	Luis Miguel Quito Carrión

## 1. Introducción

El proyecto "**Implementación de BI para la gestión de datos generados en el módulo de pedidos en la empresa Marketing Digital Bravo & Beltrán**" tiene como objetivo principal mejorar el proceso de toma de decisiones a través de la implementación de un dashboard que presente mediante gráficos estadísticos, tablas y elementos visuales la información de valor de la empresa. En este documento se definirán los requerimientos necesarios para el desarrollo de la solución BI y se establecerán las bases para su implementación exitosa.

### 1.1 Propósito

El propósito de este documento es especificar y definir de manera clara y precisa todas las funcionalidades, restricciones y requisitos de la solución BI que permitirá gestionar los datos generados en el módulo de pedidos de la empresa Marketing Digital Bravo & Beltrán.

### 1.2 Ámbito del Sistema

El sistema que se desarrollará consiste en un dashboard empresarial basado en software Open Source, en el que se puede visualizar información sobre el rendimiento de los pedidos en la empresa, y apoyar en el proceso de toma de decisiones al gerente de la institución.

### 1.3 Personal Involucrado

<b>Nombre</b>	Luis Miguel Quito Carrión
<b>Cargo</b>	Investigador
<b>Responsabilidad</b>	Recolección y análisis de información, implementación de BI, documentación.
<b>Información de contacto</b>	luis.quito@unl.edu.ec

<b>Nombre</b>	Edwin René Guamán Quinche
<b>Cargo</b>	Director del Proyecto
<b>Responsabilidad</b>	Revisión y control
<b>Información de contacto</b>	rguaman@unl.edu.ec

<b>Nombre</b>	Diego Fernando Bravo
<b>Cargo</b>	Gerente de la empresa
<b>Responsabilidad</b>	Brindar información de interés para el proyecto, revisión del sistema BI.
<b>Información de contacto</b>	gerencia@donesitios.com

## 1.4 Definiciones, acrónimos y abreviaturas

En la siguiente tabla, se detallan algunos términos utilizados con sus respectivas definiciones.

TABLA XXIII. ACRÓNIMOS Y ABREVIATURAS

<b>Acrónimo/Abreviatura</b>	<b>Definición</b>
<b>Data Warehouse (DW)</b>	Sistema de almacenamiento y administración de datos de diferentes fuentes, que proporcionan información significativa
<b>Data mart (DM)</b>	Subconjunto del DW que se extrae para satisfacer las necesidades de un área específica dentro de la empresa
<b>ETL</b>	Extraction, Transformation, Loading, procesos de extracción, transformación y carga de los datos en un DW.
<b>OLAP</b>	On-line Analytical Processing, permite a los usuarios extraer y consultar datos de manera fácil y selectiva para su análisis
<b>BI</b>	Business Intelligence
<b>IEEE</b>	Institute of Electrical and Electronics Engineers (Instituto de Ingeniería Eléctrica y Electrónica).
<b>RF</b>	Requisito Funcional
<b>RNF</b>	Requisito no funcional

## 1.5 Referencias

<b>Título del documento</b>	<b>Referencia</b>
IEEE Recommended Practice for Software Requirements Specifications	"IEEE Recommended Practice for Software Requirements Specifications," in IEEE Std 830-1998 , vol., no., pp.1-40, 20 Oct. 1998, doi: 10.1109/IEEESTD.1998.88286 [43].

## 1.6 Resumen

Este documento fue elaborado con el fin de detallar todas las consideraciones técnicas en el desarrollo de la solución BI, para la gestión de los datos de pedidos de la empresa Marketing Digital Bravo & Beltrán. Esta organizado en tres secciones. En la primera sección se presenta una descripción general de las especificaciones de recursos del sistema. En la segunda sección del documento, se presentan las principales funciones que debe realizar el sistema, los datos involucrados y los factores, restricciones, suposiciones y dependencias que influyen en él. Finalmente, la tercera sección del documento define requisitos funcionales (RF) y requisitos no funcionales (RNF) que se van a implementar en el desarrollo de la solución BI.

## **2. Descripción General**

### **2.2 Perspectiva del producto**

La solución de BI para la gestión de datos en el área de pedidos que se va a desarrollar, depende de los datos que se generen dentro de la empresa y tiene como objetivo ser una herramienta que apoye en la toma de decisiones eficiente. Es una herramienta diseñada para entornos web, que se podrá integrar al resto de módulos existentes en el sistema de la empresa.

### **2.3 Funcionalidad del producto**

La funcionalidad del sistema BI se puede visualizar en el dashboard, mediante gráficos estadísticos (diagramas de barras, pasteles, tablas, graficas, etc.) los datos históricos de los pedidos generados en la empresa. La solución BI, pretende ser un apoyo para la gerencia de la empresa Marketing Digital Bravo & Beltrán, en el proceso de toma de decisiones.

### **2.4 Características de los usuarios**

La persona que hará uso de la solución BI es el gerente de la empresa, quien tiene conocimiento de informática intermedio, por lo que no presenta inconvenientes para el uso de la solución BI.

### **2.5 Restricciones**

- La solución BI debe ser representada por un dashboard.
- La interfaz del dashboard debe ser sencilla e intuitiva.
- El acceso sera permitido solamente a usuarios autenticados como administradores.
- Se debe construir el sistema con herramientas de software libre.
- La aplicación BI debe ser accedida mediante interfaz web.
- El sistema debe ser multiplataforma.

### **2.6 Suposiciones y dependencias**

- Los requisitos listados son estables.
- Los reportes generados estarán basados con los datos de pedidos brindados por la empresa.
- La solución sera desarrollada en un ambiente de desarrollo.

## **3. Requisitos específicos**

En esta sección se presentan los requisitos específicos que deberán ser satisfechos por la solución. Todos los requisitos aquí expuestos son esenciales, es decir, no sería aceptable un sistema que no satisfaga alguno de los requisitos aquí presentados.

### 3.2 Requisitos funcionales

En esta fase se identificaron las necesidades de información requeridas por la gerencia, que son esenciales para lograr cumplir objetivos y estrategias de la empresa. Se definieron 16 requisitos de BI y 1 requisito de Software.

#### Requerimiento funcional 1

<b>Código</b>	RF01
<b>Nombre</b>	Autenticación de usuarios
<b>Descripción</b>	El sistema permite el acceso a usuarios debidamente autenticados a través de la validación de sus credenciales
<b>Prioridad</b>	Alta

#### Requerimiento funcional 2

<b>Código</b>	RF02
<b>Nombre</b>	Visualizar datos históricos de pedidos
<b>Descripción</b>	El sistema permite visualizar los datos históricos de pedidos generados
<b>Prioridad</b>	Alta

#### Requerimiento funcional 3

<b>Código</b>	RF03
<b>Nombre</b>	Visualizar pedidos y monto facturado
<b>Descripción</b>	El sistema permite visualizar el número total de pedidos y el monto total facturado
<b>Prioridad</b>	Alta

#### Requerimiento funcional 4

<b>Código</b>	RF04
<b>Nombre</b>	Visualizar número de empresas y usuarios
<b>Descripción</b>	El sistema permite conocer el número total de empresas y usuarios registrados en la base de datos
<b>Prioridad</b>	Media

#### Requerimiento funcional 5

<b>Código</b>	RF05
<b>Nombre</b>	Mostrar empresa con más pedidos
<b>Descripción</b>	El sistema debe mostrar las empresas con mayor número de pedidos despachados.

<b>Prioridad</b>	Alta
------------------	------

---

#### Requerimiento funcional 6

<b>Código</b>	RF06
<b>Nombre</b>	Visualizar productos y montos por empresa
<b>Descripción</b>	El sistema permite visualizar los productos vendidos y el monto generado por empresa
<b>Prioridad</b>	Media

---

#### Requerimiento funcional 7

<b>Código</b>	RF07
<b>Nombre</b>	Mostrar usuarios con más pedidos
<b>Descripción</b>	El sistema deberá mostrar los usuarios con mayor número de pedidos realizados y el monto total
<b>Prioridad</b>	Media

---

#### Requerimiento funcional 8

<b>Código</b>	RF08
<b>Nombre</b>	Listar productos más pedidos
<b>Descripción</b>	El sistema permite conocer los productos más vendidos
<b>Prioridad</b>	Media

---

#### Requerimiento funcional 9

<b>Código</b>	RF09
<b>Nombre</b>	Mostrar usuarios y monto por empresa.
<b>Descripción</b>	El sistema debe mostrar los usuarios con más pedidos y el monto vendido por empresa
<b>Prioridad</b>	Media

---

#### Requerimiento funcional 10

<b>Código</b>	RF10
<b>Nombre</b>	Visualizar productos por empresa.
<b>Descripción</b>	El sistema debe permitir visualizar los productos más pedidos por empresa
<b>Prioridad</b>	Media

---

### Requerimiento funcional 11

<b>Código</b>	RF11
<b>Nombre</b>	Visualizar pedidos y monto por tipo de pago
<b>Descripción</b>	El sistema permite visualizar el número de pedidos y monto total según el tipo de pago
<b>Prioridad</b>	Media

### Requerimiento funcional 12

<b>Código</b>	RF12
<b>Nombre</b>	Listar pedidos y monto por tipo de entrega
<b>Descripción</b>	El sistema permite conocer el número de pedidos y monto total por tipo de entrega
<b>Prioridad</b>	Media

### Requerimiento funcional 13

<b>Código</b>	RF13
<b>Nombre</b>	Mostrar pedidos por empresa
<b>Descripción</b>	El sistema debe mostrar el número de pedidos y el monto total por empresa
<b>Prioridad</b>	Alta

### Requerimiento funcional 14

<b>Código</b>	RF14
<b>Nombre</b>	Visualizar pedidos y monto por empresa según el tipo de pago
<b>Descripción</b>	El sistema permite visualizar el número de pedidos y monto total por empresa según el tipo de pago
<b>Prioridad</b>	Media

### Requerimiento funcional 15

<b>Código</b>	RF15
<b>Nombre</b>	Visualizar pedidos y monto por empresa según el tipo de entrega
<b>Descripción</b>	El sistema permite visualizar el número de pedidos y monto total por empresa según el tipo de empresa
<b>Prioridad</b>	Media

### Requerimiento funcional 16

<b>Código</b>	RF16
<b>Nombre</b>	Calcular promedio de pedido por empresa
<b>Descripción</b>	El sistema permite calcular el monto promedio de pedidos por empresa
<b>Prioridad</b>	Alta

### Requerimiento funcional 17

<b>Código</b>	RF17
<b>Nombre</b>	Visualizar pedidos y monto por categoría y subcategoría
<b>Descripción</b>	El sistema permite visualizar los pedidos y el monto por categoría y subcategoría
<b>Prioridad</b>	Media

### 3.3 Requisitos no funcionales

En esta fase se identificaron 6 requisitos no funcionales que debe cumplir la solución de Business Intelligence:

TABLA XXIV. REQUERIMIENTO DE SEGURIDAD

<b>Nombre:</b> Seguridad	<b>Código:</b> RNF01
<b>Descripción:</b>	El sistema debe brindar seguridad tanto a nivel de transmisión de datos, seguridad de acceso a bases de datos, y al sistema en sí.
<b>Prioridad:</b> Alta	

TABLA XXV. REQUERIMIENTO DE ESCALABILIDAD

<b>Nombre:</b> Escalabilidad	<b>Código:</b> RNF02
<b>Descripción</b>	El sistema debe tener la capacidad de incorporar cambios que puedan presentarse dentro de la empresa. Para cumplir con este requerimiento, se hace uso de la metodología de Ralph Kimball, la cual brinda escalabilidad a la solución BI.
<b>Prioridad:</b> Alta	

TABLA XXVI. REQUERIMIENTO DE FACILIDAD DE USO

<b>Nombre:</b> Facilidad de uso	<b>Código:</b> RNF03
<b>Descripción</b>	El sistema debe ser construido de tal manera que resulte sencillo de manejar para el usuario, este requerimiento se cumple mediante la implementación de interfaces de usuario intuitivas.
<b>Prioridad:</b> Alta	

TABLA XXVII. REQUERIMIENTO DE CONFIABILIDAD

<b>Nombre:</b> Confiabilidad	<b>Código:</b> RNF04
<b>Descripción</b>	La solución BI debe presentar información correcta, ya que servirá de base para llevar a cabo el proceso de toma de decisiones.
<b>Prioridad:</b> Alta	

TABLA XXVIII. REQUERIMIENTO DE DISPONIBILIDAD

<b>Nombre:</b> Disponibilidad	<b>Código:</b> RNF05
<b>Descripción</b>	La solución BI debe estar disponible el 100% de las veces que el usuario intente acceder.
<b>Prioridad:</b> Alta	

TABLA XXIX. REQUERIMIENTO DE PORTABILIDAD

<b>Nombre:</b> Portabilidad	<b>Código:</b> RNF06
<b>Descripción</b>	El dashboard resultante debe ser visible con cualquier navegador
<b>Prioridad:</b> Alta	

### Anexo 3. Entrevista

A continuación, se presenta la entrevista realizada para la obtención de información general de la empresa y los requerimientos de información necesarios para la implementación de la solución propuesta en el TT.

---

<b>Entrevistado:</b>	Ing. Diego Bravo <b>Gerente en Marketing Digital Bravo &amp; Beltrán</b>
<b>Fecha:</b> 13/12/2021 15H:0-16H:00	<b>Duración:</b> 1 hora
<b>Lugar:</b> Oficinas de Marketing Digital Bravo & Beltrán	
<b>Responsable:</b>	Luis Miguel Quito Carrión <b>Autor del Trabajo de Titulación</b>

---

#### ENTREVISTA

---

<b>Pregunta 1</b>	<b>¿En qué consiste la empresa y cuál es su giro de negocio?</b> La compañía está enfocada especialmente al desarrollo del software, brindar soluciones eficientes en cuanto a tecnología y a software. La empresa se especializa en el desarrollo de software y soluciones tecnológicas eficientes, con el objetivo de satisfacer las necesidades y demandas de sus clientes. Además, cuenta con un equipo altamente capacitado en tecnologías de vanguardia para garantizar la calidad y la eficiencia de sus productos y servicios.
<b>R:</b>	
<b>Pregunta 2</b>	<b>¿Cuántos clientes se maneja en el sistema y cuántos usuarios están activos?</b> Actualmente hemos llegado a tener 22300 clientes, tenemos una media de casi 5 mil usuarios, y un poquito más de 500 que están activos en la aplicación, pero dentro del sistema la aplicación D'onesitios tiene alrededor de casi 35mil descargas este momento, solo en la Play Store, en la App Store que tiene cerca de 4mil y en Huawei tenemos unos cuantos más. Estamos más o menos entre casi llegando a 40 mil descargas de la aplicación, pero usuarios netos debemos tener cerca de unos 10 mil. Que generan información digamos tiene tráfico con el sistema. Sí, que están activos este momento. Hay gente que se desactiva después de un tiempo y vuelve a activarse, pero los que están activos es un promedio más o menos de 10mil.
<b>R:</b>	
<b>Pregunta 3</b>	<b>¿Cómo se almacenan y segmentan los datos generados por los sistemas de la empresa?</b> Estamos trabajando con 2 servidores, uno que es para guardar todo lo que es data y otro para almacenar imágenes. La segmentación es de acuerdo a las tablas que se ingresan al inicio con datos generales del cliente en ese momento, nombres, direcciones, teléfonos, correos electrónicos y luego como se va alimentando la aplicación/plataforma, por ejemplo: direcciones del cliente o tipos de pedido que hace o formas de pago que ingresan, toda la información se va complementando.
<b>R:</b>	
<b>Pregunta 4</b>	<b>Actualmente con esta información, ¿De qué manera se realiza el estudio de este tipo de datos generados por los clientes?</b>

---

R: Realmente no estamos trabajando a un nivel de todavía darles un tratamiento, darle una finalidad a esta pero nuestro trabajo como empresa este momento es tomar decisiones en cuanto a esta información, no tanto para tratar los datos del cliente si no para darles un mejor servicio. Por ejemplo, lo usamos para el tema de call center, lo usamos para sugerir en algunos casos productos para asignar premios, para verificar direcciones, pero, todavía no hemos pasado nivel donde ya podamos darle un valor más importante a esta información

---

**Pregunta 5**

**¿Qué necesidades presenta la empresa con sus datos actualmente?**

R: La implementación de nuevos sistemas para el tratamiento de esta información a través de sistema de inteligentes que generen reportes sencillos, para poder brindarles mejores funciones y recomendaciones a los clientes y también se podría segmentar de mejor manera para poder llegar a ellos con productos y servicios que realmente necesitan. Se necesitan por ejemplo reportes específicamente en el área de pedidos, ya que es de nuestro interés como empresa conocer datos referentes a ellos, por ejemplo, es necesario conocer el número de pedidos generados por fechas, productos y empresas que más pedidos tiene, usuarios que más compras realizan y cuanto gastan en promedio por pedido, además de información de tipo de pago, promedio de ventas por empresa y tipo de entrega del pedido.

---

**Pregunta 6**

**¿Qué impacto crees que pueda tener la implementación de un sistema de datos para la gestión de los mismos?**

R: Realmente uno de nuestros objetivos como empresa este momento no es solo brindarles opciones a la gente a través de los productos o servicios que brindamos, si no también es ir recopilando esta información para poder través de esta información, darles un sentido más importante a los productos o al trabajo que hacemos. El impacto va a ser fuertísimo porque el rato que implementemos nuevos sistemas y nuevos procesos para esta data, vamos a tener resultados increíbles porque estamos arrasando mucho con el cliente. Claro que, con el consentimiento de cada uno de estos clientes para tratarlo dentro de la empresa, y poder luego llegar de una manera, pero increíble con nuestros clientes, dándoles soluciones efectivas. Lo que trabajamos es para poder entender los clientes a través de esta data y para eso de ley tenemos que implementar un montón de cosas detrás que es lo que la data, como dicen este "nuevo petróleo" del siglo 21.

---

**Responsables**



Luis Miguel Quito Carrión  
estudiante



Ing. Diego Bravo  
Gerente de la empresa Marketing Digital  
Bravo & Beltrán

#### Anexo 4. Descripción de tablas del Data mart

A continuación, se describe las tablas que componen el data mart, en cada una de ellas se muestra el nombre de la tabla, una breve descripción de los registros que se almacenan, las columnas que contiene, descripción, tipo de dato, tamaño y tipo de llave (primary key o foreign key) que posee.

En la TABLA XXXI, se describe la tabla “dim\_time”.

TABLA XXX. DESCRIPCIÓN DIMENSIÓN TIEMPO					
Tabla: dim_time					
<b>Descripción:</b> Contiene valores de fechas de las transacciones que se han realizado en la organización para obtener un reporte más específico.					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
date_code	Identificador en el tiempo en que se realizó el pedido	Varchar	16	✓	
year	Año en el que se encuentra la fecha	Date			
semester	Identificador del trimestre en el que se encuentra la fecha	Date			
trimestre	Identificador del trimestre en el que se encuentra la fecha	Date			
month	Número del mes en el que se encuentra la fecha	Date			
week	Número de semana en el que se encuentra la fecha	Date			
day	Día de la semana en el que se encuentra la fecha	Date			

En la TABLA XXXI, se describe la tabla “dim\_store”.

TABLA XXXI. DESCRIPCIÓN DIMENSIÓN EMPRESA					
Tabla: Dim_Store					
<b>Descripción:</b> Tabla de registros de empresas					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
store_code	Identificador de empresa	Varchar	16	✓	
name	Nombre de la empresa	Varchar	250		
subsidiary	Sucursal de la empresa	Varchar	100		
category	Categoría a la cual pertenece la empresa	Varchar	100		
subcategory	Subcategoría a la cual pertenece la empresa	Varchar	100		

En la TABLA XXXII, se describe la tabla “dim\_product”.

TABLA XXXII. DESCRIPCIÓN DIMENSIÓN PRODUCTO					
Tabla: Dim_Producto					
Descripción: Tabla de registros de productos y sus características					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
product_code	Identificador del producto	Varchar	16	✓	
name	Nombre del producto	Varchar	250		
quantity	Cantidad del producto	Int	10		
price	Precio del producto	Double	10		
iva	Iva del producto	Double	10		
price_total	Precio total del producto	Double	10		

En la TABLA XXXIII TABLA XXXI, se describe la tabla “dim\_store”.

TABLA XXXIII. DESCRIPCIÓN DIMENSIÓN TIPO DE PAGO					
Tabla: Dim_TypePay					
Descripción: Contiene registros sobre los tipos de pago					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
pay_code	Identificador del tipo de pago	Varchar	16	✓	
type_pay	descripción sobre el tipo de pago	Varchar	50		

En la TABLA XXXIV, se describe la tabla “dim\_user”.

TABLA XXXIV. DESCRIPCIÓN DIMENSIÓN USUARIO					
Tabla: Dim_User					
Descripción: Se registran los datos del cliente					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK
user_code	Identificador del cliente	Varchar	16	✓	
name	Nombre completo del cliente	Varchar	250		

En la TABLA XXXV, se describe la tabla “fact\_order”.

TABLA XXXV. DESCRIPCIÓN HECHO PEDIDOS					
Tabla: Fact_order					
Descripción: Contiene valores de fechas de las transacciones que se han realizado en la organización para obtener un reporte más específico					
Columna	Descripción	Tipo de dato	Tamaño	PK	FK

order_code	Clave foránea de dimensión pedidos	la	Varchar	16	✓
store_code	Clave foránea de dimensión empresa	la	Varchar		✓
user_code	Clave foránea de dimensión usuarios	la	Varchar		✓
product_code	Clave foránea de dimensión productos	la	Varchar		✓
date_code	Clave foránea de dimensión fecha	la	Varchar		✓
pay_code	Clave foránea de dimensión tipo de pago	la	Varchar		✓

## Anexo 5. Instalación de Herramientas

El principal objetivo de esta sección es proporcionar al usuario una guía para la instalación de las herramientas que se utilizaron en el desarrollo del TT.

### Instalación de Java Runtime Environment (JRE)

El entorno de Pentaho está basado en Java, por lo que es necesario contar previamente con JRE instalado. Para su descarga se ingresó al enlace: <https://www.java.com/es/download/>, se selecciona el botón de “Aceptar e iniciar descarga gratuita” (véase en la Figura 81); automáticamente se inicia la descarga de la última versión disponible de Java para el sistema operativo y arquitectura que se esté usando, en este caso se utilizó un equipo con Windows 10 de 64 bits.

#### Java de 64 bits para Windows

Recomendado Version 8 Update 321 (Tamaño de archivo: 81.99 MB)

Fecha de publicación: 18 de enero de 2022

 **Actualización importante de la licencia de Oracle Java**

**La licencia de Oracle Java ha cambiado para las versiones publicadas a partir del 16 de abril de 2019.**

El nuevo [acuerdo de licencia de Oracle Technology Network para Oracle Java SE](#) es sustancialmente diferente a las licencias de Oracle Java anteriores. La nueva licencia permite ciertos usos, como el uso personal y de desarrollo, sin coste alguno (aunque podría haber otros usos autorizados en licencias de Oracle Java anteriores que ya no estén disponibles). Revise las condiciones con atención antes de descargar y utilizar este producto. Puede consultar las preguntas frecuentes [aquí](#).

La licencia comercial y el soporte están disponibles con una [suscripción de Java SE](#) de bajo coste.

Oracle también ofrece la última versión de OpenJDK con la [licencia pública general](#) de código abierto en [jdk.java.net](http://jdk.java.net).

 Hemos detectado que está utilizando Google Chrome y quizá no pueda utilizar el complemento Java desde este explorador. A partir de la versión 42 (abril 2015), Chrome ha desactivado la el método estándar por el que los exploradores soportan plugins. [Más información](#)



Figura 81. Descarga de JRE

Para el presente proyecto se instaló la versión de JRE-8u321, ya que cumple con los requisitos necesarios para la instalación de PDI.

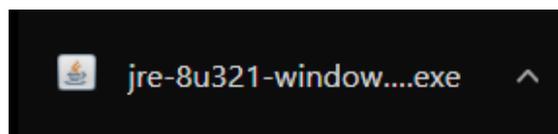


Figura 82. Versión de JRE

Se ejecutó el instalador, el cual presenta la pantalla de bienvenida, así como el acuerdo de licencia para su instalación, a la cual se debe seleccionar el botón “Instalar” para continuar con el proceso, de la forma que se indica en la Figura 83.

En caso que se requiera cambiar la dirección de instalación se debe marcar la opción de “Cambiar la Carpeta de Destino”, caso contrario se instalará por defecto en C:\Program Files (x86).

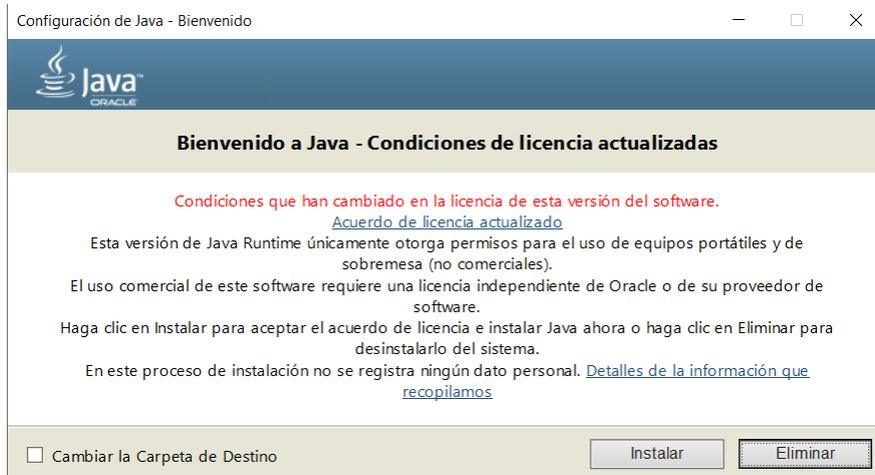


Figura 83. Acuerdo de licencia JRE

El proceso de instalación comienza, durante el proceso se van a extraer todos los elementos necesarios para la correcta instalación de JRE. En la Figura 84, se muestra el progreso en la instalación de Java.



Figura 84. Instalación de JRE

Una vez terminada la instalación de JRE, se muestra una pantalla donde se indica el estado final en la instalación de Java, finalmente se selecciona el botón “Cerrar” para dar por finalizado el proceso de instalación. En la Figura 85 se indica la correcta instalación de Java.

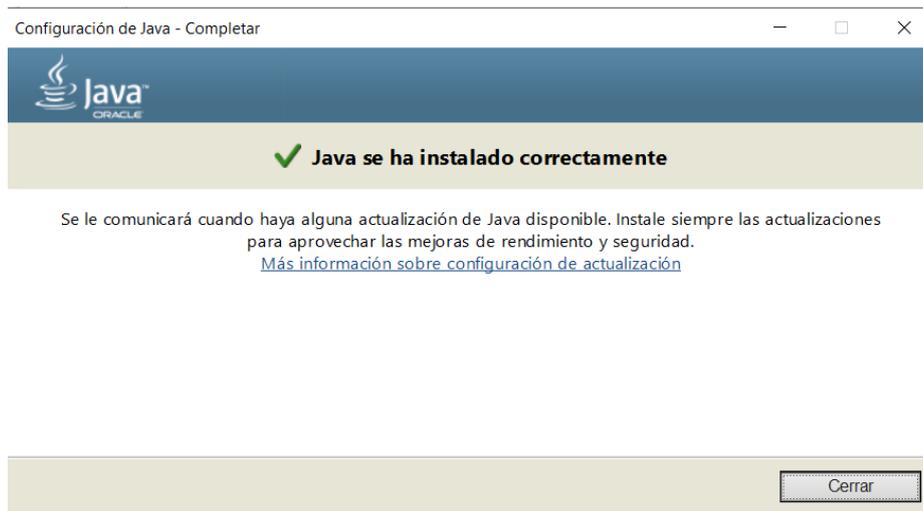


Figura 85. Fin de instalación de JRE

## Instalación de PDI

Para la instalación de Pentaho Data Integration, en primer lugar, se descargó el instalador desde el enlace: <https://sourceforge.net/projects/pentaho/>, de la forma que se indica en la Figura 86.

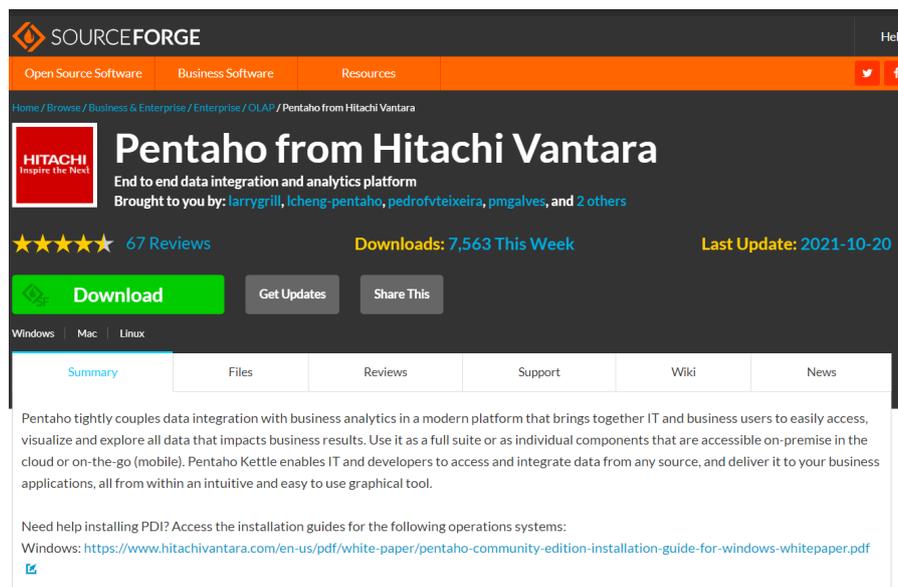


Figura 86. Descarga de Pentaho Data Integration

Una vez finalizada la descarga se obtiene la carpeta con todos los archivos necesarios para el funcionamiento de Pentaho. En la Figura 87 se ilustra la carpeta de PDI.

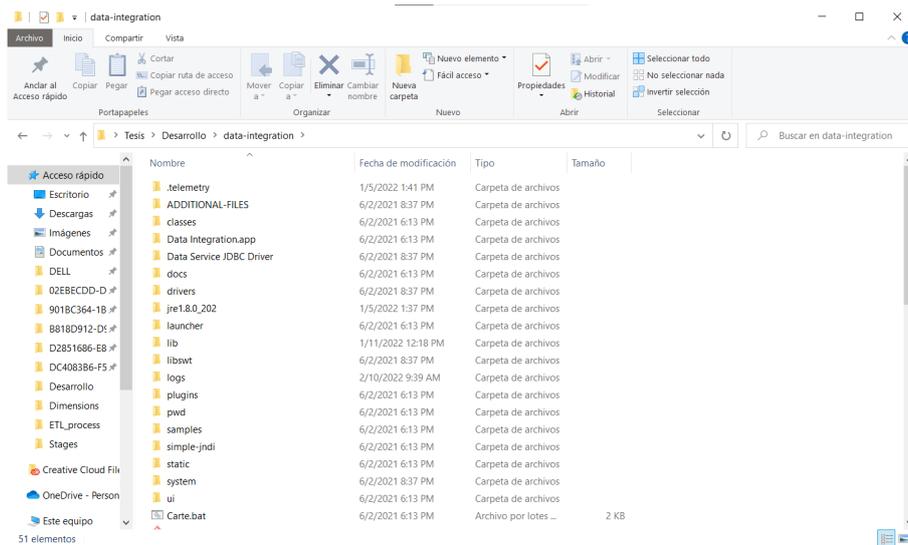


Figura 87. Carpeta de Pentaho Data Integration

Posteriormente, para poner en marcha Data Integration se ejecutó el archivo “Spoon.bat” que se encarga de abrir el entorno de PDI, tal como se muestra en la Figura 88.

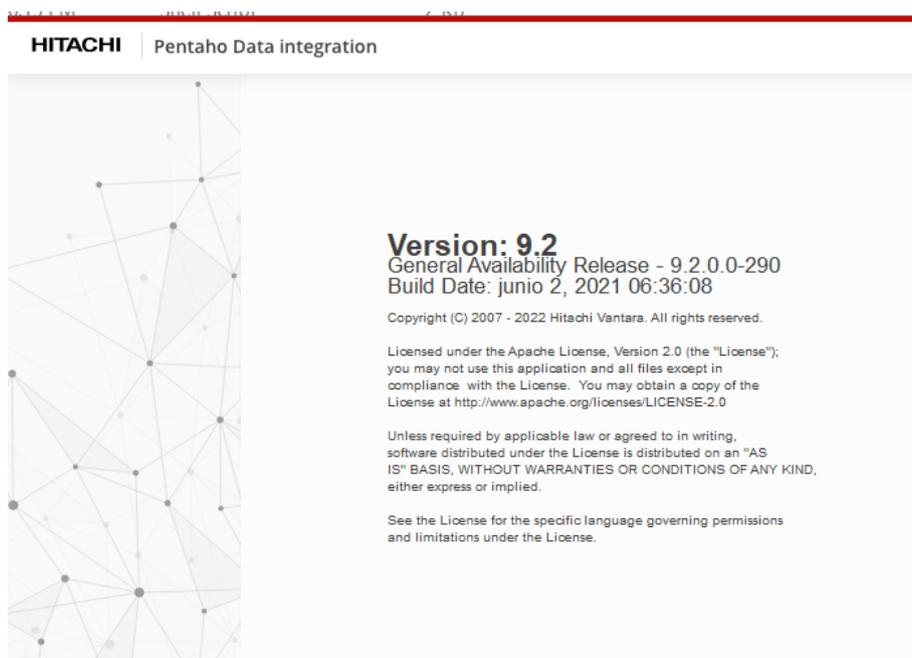


Figura 88. Ejecución de PDI

Finalmente, en la Figura 89, se ilustra el entorno de trabajo de Pentaho Data Integration instalado correctamente y listo para su uso.

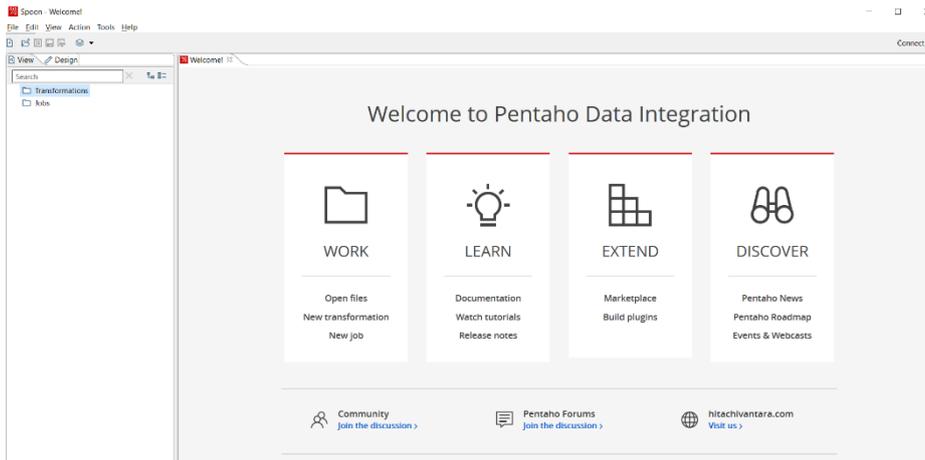


Figura 89. Entorno PDI

## Instalación de Mondrian

Para la instalación de Mondrian, se debe descargar el instalador desde la dirección: <https://sourceforge.net/projects/mondrian/files/>, en la sección “Files” se selecciona la opción “*schema workbench*”, tal como se ejemplifica en la **Figura 90**.

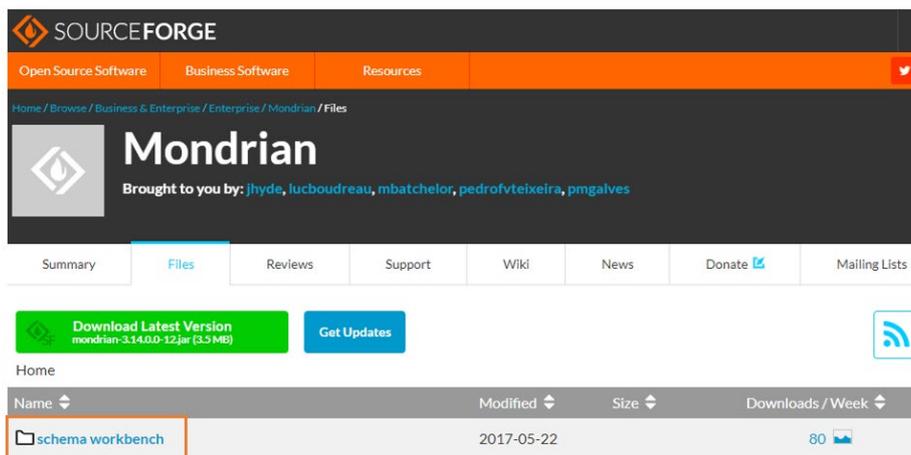


Figura 90. Repositorio de descarga de Schema Workbench.

Una vez finalizada la descarga se obtiene la carpeta “*schema-workbench*” que contiene todos los archivos necesarios para el funcionamiento de Mondrian. En la Figura 91 *Figura 87* se ilustra la carpeta de Schema Workbench.

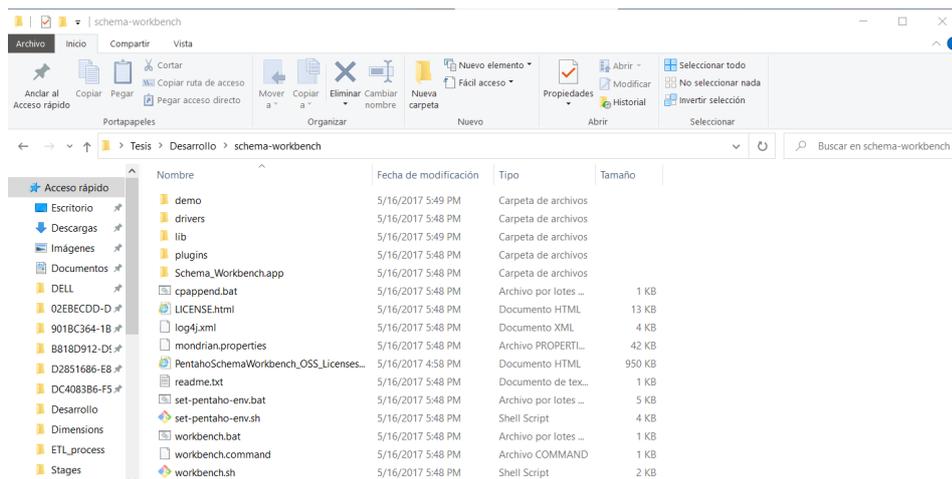


Figura 91. Carpeta schema-workbench.

Para poner en marcha el programa se ejecutó el archivo “*workbench.bat*” que se encarga de abrir Schema Workbench. Finalmente, en la Figura 92, se ilustra el entorno de trabajo de Mondrian instalado correctamente y listo para su uso.

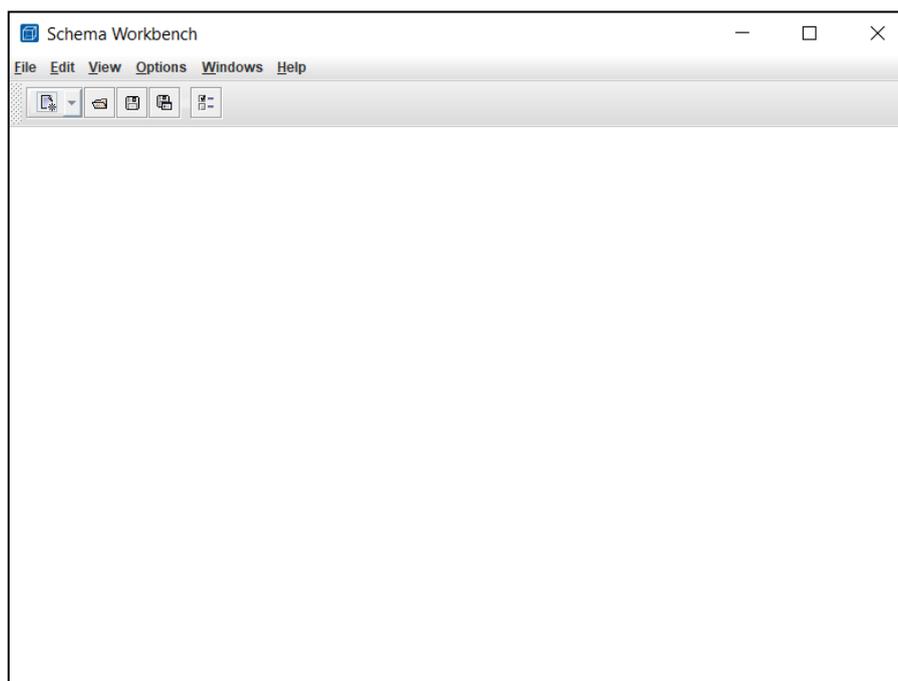


Figura 92. Entorno Mondrian

### Instalación de Pentaho Server

Para la instalación de Pentaho Server, se debe ingresar al sitio web: y descargar la carpeta “*Business Intelligence Server*”, de la forma que se indica en la *Figura 93*.

	Open Source Software	Business Software	Resources
Pentaho 8.1	2018-05-15		96
Pentaho 8.0	2017-11-15		96
Pentaho Metadata	2017-11-15		8
Big Data Shims	2017-11-15		4
Data Integration	2017-05-22		625
Business Intelligence Server	2017-05-22		90

Figura 93. Repositorio de descarga Pentaho Server

Una vez finalizada la descarga se obtiene la carpeta “pentaho-server-ce-7.1.0.0-12” que contiene todos los archivos necesarios para el funcionamiento de Pentaho Server. En la Figura 94 se ilustra la carpeta resultante.

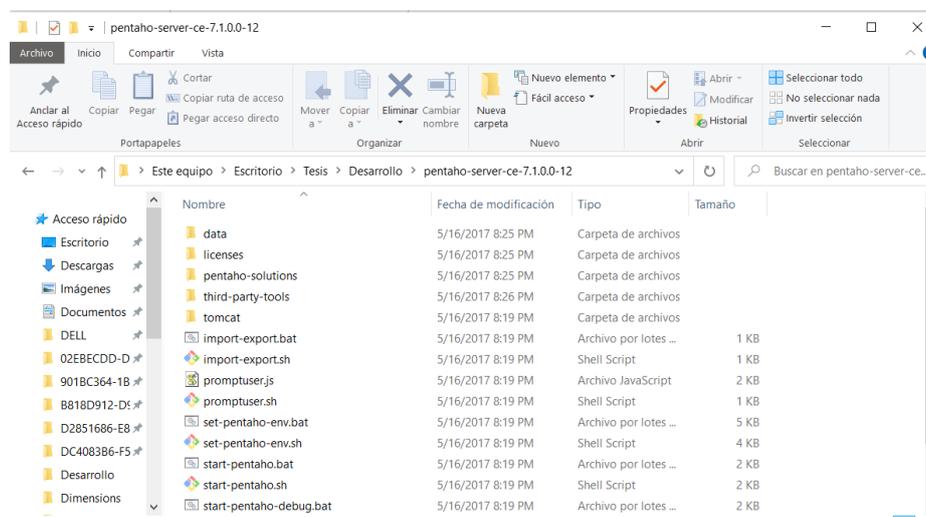


Figura 94. Carpeta de Pentaho Server

Para poner en marcha el programa se ejecutó el archivo “start-pentaho.bat”, posteriormente se abre Google Chrome con la ruta: localhost:8080 que contiene el dashboard general de Pentaho Server. En la Figura 95, se ilustra el entorno de trabajo listo para su uso.

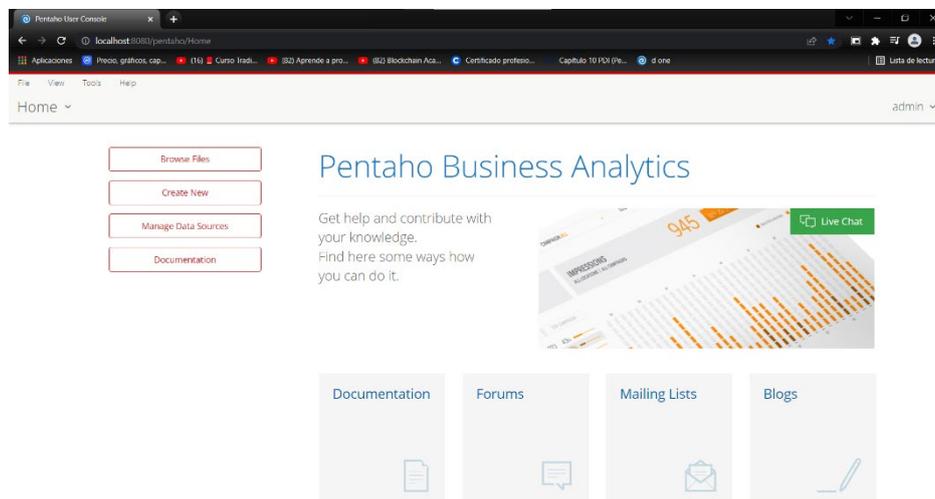
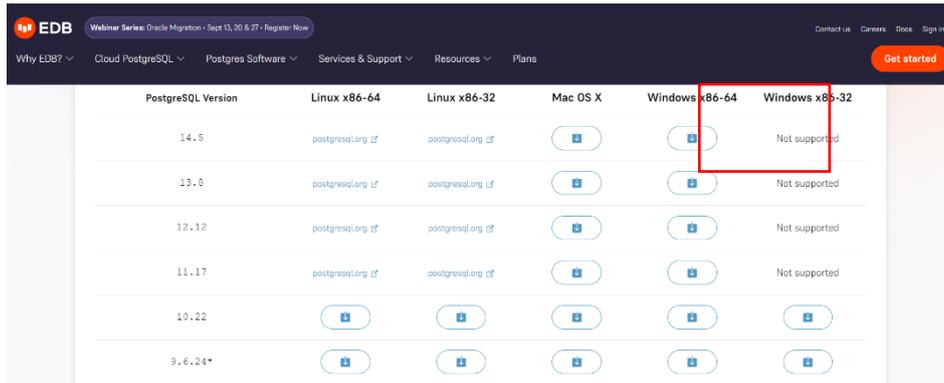


Figura 95. Entorno Pentaho Server

## Instalación de PostgreSQL

Para la instalación de PostgreSQL, se debe ingresar al sitio web: <https://www.enterprisedb.com/downloads/postgres-postgresql-downloads> y descargar la última versión del programa de acuerdo al sistema operativo en el que se esté trabajando, para el desarrollo del TT se utilizó la versión 14.5 para el entorno Windows 10. En la Figura 96, se ilustra el instalador descargado para la instalación de PostgreSQL.



PostgreSQL Version	Linux x86-64	Linux x86-32	Mac OS X	Windows x86-64	Windows x86-32
14.5	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	Not supported
13.9	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	Not supported
12.12	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	Not supported
11.17	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	<a href="#">postgres.org</a>	Not supported
10.22	<a href="#">postgres.org</a>				
9.6.24*	<a href="#">postgres.org</a>				

Figura 96. Descarga de PostgreSQL

Una vez finalizada la descarga se ejecutó el archivo, mismo que inicia el asistente de instalación de PostgreSQL para Windows 10, que se indicada en la Figura 97.

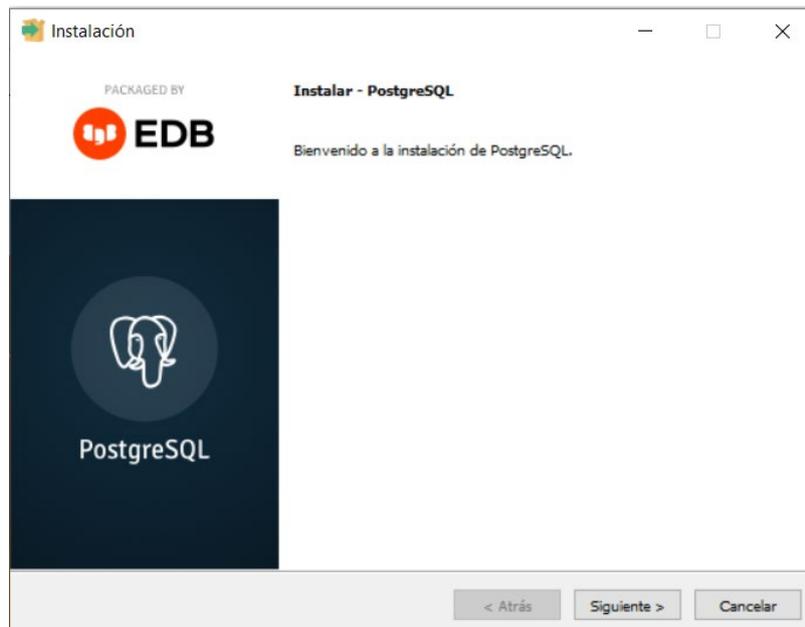


Figura 97. Ejecución del Instalador de PostgreSQL

A continuación, se definió el directorio donde se realizará la instalación, en la Figura 98 se indica la dirección utilizada para la instalación de PostgreSQL.



Figura 98. Directorio de instalación PostgreSQL

En la Figura 99 se muestra la ventana donde se definió los componentes que se instalaron.

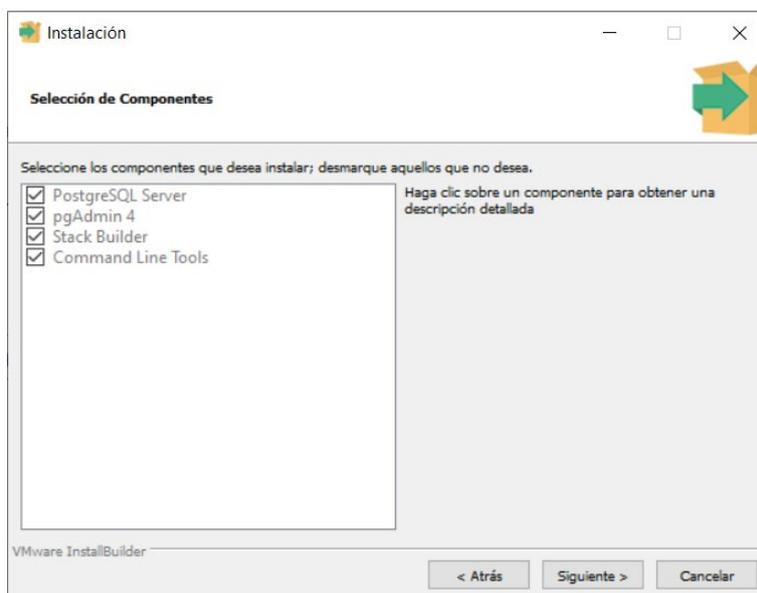


Figura 99. Instalación de componentes de PostgreSQL

A continuación, se asignó la contraseña para el super usuario de PostgreSQL. En la Figura 100 se indica el llenado de la contraseña.

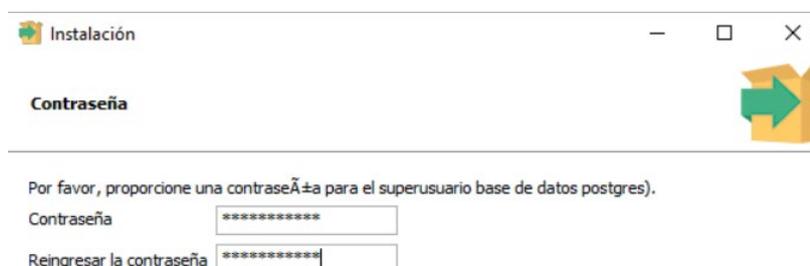


Figura 100. Asignación de contraseña para el super usuario

En la Figura 101, se ilustra la ventana donde se especificó el puerto en el cual se ejecuta el servidor de PostgreSQL.

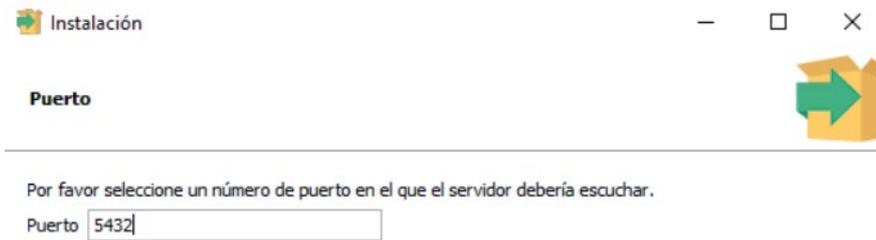


Figura 101. Puerto para levantar servidor de PostgreSQL

A continuación, en la Figura 102, se indica la configuración regional seleccionada.

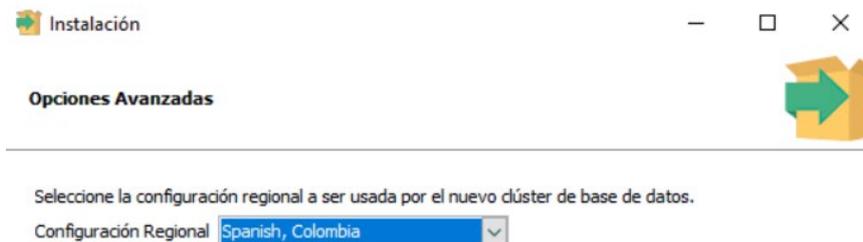


Figura 102. Configuración Regional

Seguidamente se indica una ventana con un resumen de las configuraciones que se usaran para la instalación, de la forma que se indica en la Figura 103.

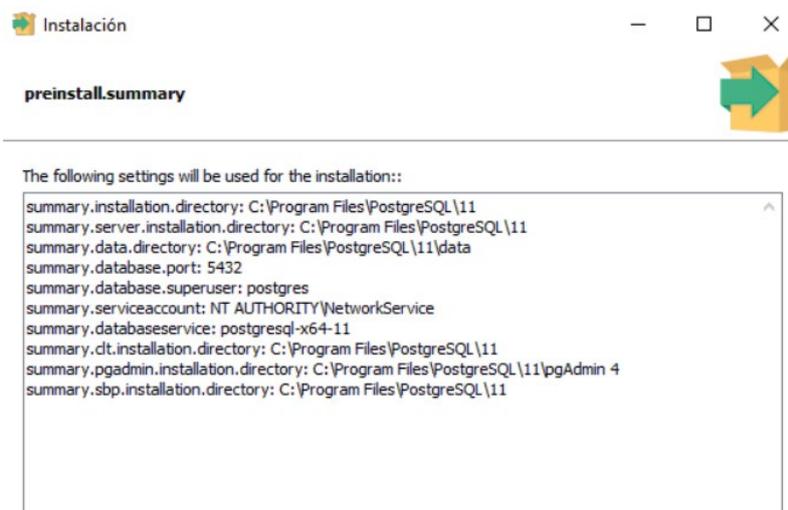


Figura 103. Configuraciones para la instalación de PostgreSQL

En la Figura 104, se indica el proceso de instalación del gestor de base de datos.

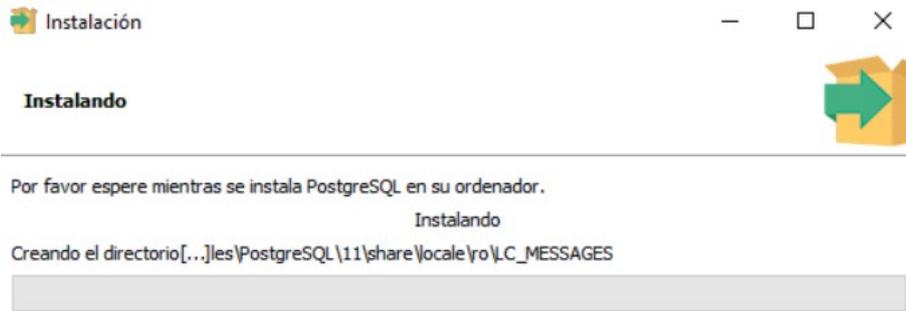


Figura 104. Instalación de PostgreSQL

Cuando el proceso de instalación haya culminado se mostrará una ventana como la detallada en la Figura 105, la cual indica la correcta instalación de PostgreSQL en el computador.



Figura 105. Culminación de instalación de PostgreSQL

Una vez finalizada la instalación se inició PostgreSQL, se ejecutó el archivo instalado “pgAdmin 4” el cual inicializa el programa, durante la ejecución se debe ingresar la contraseña de super usuario que se asignó previamente (véase en la Figura 100). A continuación, en la Figura 106 se ilustra el entorno de trabajo de PostgreSQL listo para su uso.

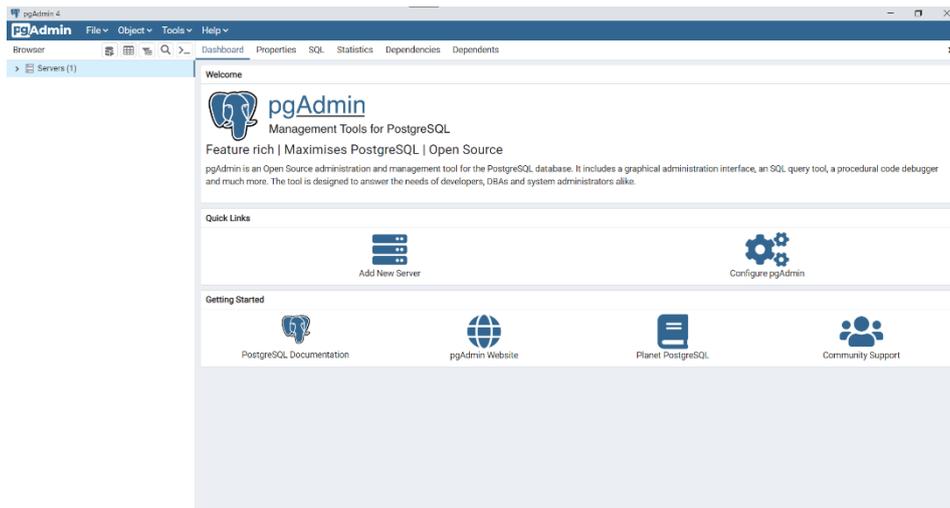


Figura 106. Dashboard de pgAdmin.

## Anexo 6. Llenado de datos

### ❖ Job Load\_Stg

En esta sección se detallan las actividades realizadas durante la extracción, transformación y carga de datos del área staging para cada transformación del Job “Load\_Stg” (véase en la Figura 24).

Los steps desarrollados son comunes para las transformaciones. En la Figura 107, se ilustra los steps que se llevaron a cabo en cada transformación.



Figura 107. Job Load\_Stg: steps de transformaciones comunes.

Los valores seleccionados en el step “*Select values*”, fueron elegidos de acuerdo a la transformación, de manera que:

- En la Figura 108, se muestran los valores seleccionados en el step “Select values” para la transformación “Product\_Stg”.

#	Fieldname	Rename to	Length	Precision
1	product_code			
2	product_name			

Figura 108. Select values: transformación “Product\_Stg”

- En la Figura 109, se muestran los valores seleccionados en el step “Select values” para la transformación “Store\_Stg”.

#	Fieldname	Rename to	Length	Precision
1	store_code			
2	store_name			
3	store_subsiary			
4	store_category			
5	store_subcategory			

Figura 109. Select values: transformación "Store\_Stg"

- En la Figura 110, se muestran los valores seleccionados en el step “Select values” para la transformación “TypePay\_Stg”.

^#	Fieldname	Rename to	Length	Precision
1	pay_code			
2	type_pay			

Figura 110. Select values: transformación "TypePay\_Stg".

- En la Figura 111, se muestran los valores seleccionados en el step "Select values" para la transformación "User\_Stg".

^#	Fieldname	Rename to	Length	Precision
1	user_code			
2	user_name			

Figura 111. Select values: transformación "User\_Stg".

Para la transformación "Values\_Stg" los steps desarrollados son los especificados en la Figura 112.



Figura 112. Transformación Values\_Stg.

- En la Figura 113, se muestran los valores seleccionados en el step "Select values" para la transformación "Values\_Stg".

^#	Fieldname	Rename to	Length	Precision
1	order_code			
2	order_date			
3	order_total			
4	order_iva			
5	pay_code			
6	product_code			
7	product_quantity			
8	product_price			
9	product_iva			
1..	produc_price_total			
1..	store_code			
1..	user_code			

Figura 113. Select values: transformación "Values\_Stg".

- **Job Load\_Dim**

En esta sección se detallan las actividades realizadas en cada transformación durante el proceso ETL para cada dimensión especificada en el Job Carga de dimensiones (véase en la Figura 32).

Los steps desarrollados son comunes para todas las transformaciones. En la Figura 114, se ilustra los steps que se llevaron a cabo en cada transformación.



Figura 114. Job Load\_Dim: steps de transformaciones comunes.

Los valores y configuraciones realizadas en el step “Table input” y “Dimension lookup/update”, fueron realizados de acuerdo a cada dimensión, de manera que:

- En la Figura 115, se muestran los valores seleccionados en el step “Table input” para la transformación “Store\_Dim”.

```

SELECT
  store_code
, store_name
, store_subsidiary
, store_category
, store_subcategory
FROM "public".stg_store
  
```

Figura 115. Table Input: transformación "Store\_Dim".

En la Figura 116, se muestran las configuraciones realizadas en el step “Dimensión lookup/update” para la transformación “Store\_Dim”.

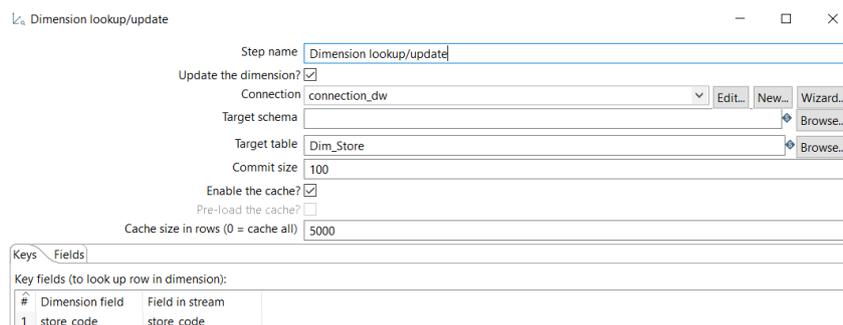


Figura 116. Dimensión lookup/update: transformación "Store\_Dim".

- En la Figura 117, se muestran los valores seleccionados en el step “Table input” para la transformación “TypePay\_Dim”.

```

SELECT
    pay_code
    , type_pay
FROM "public".stg_typepay

```

Figura 117. Table Input: transformación "TypePay\_Dim".

En la Figura 118, se muestran las configuraciones realizadas en el step "Dimensión lookup/update" para la transformación "TypePay\_Dim".

Figura 118. Dimensión lookup/update: transformación "TypePay\_Dim".

Para el almacenamiento de registros de fechas, se creó la transformación "Date\_Dim", en la Figura 119 se ilustra la configuración de steps realizada para la dimensión tiempo.

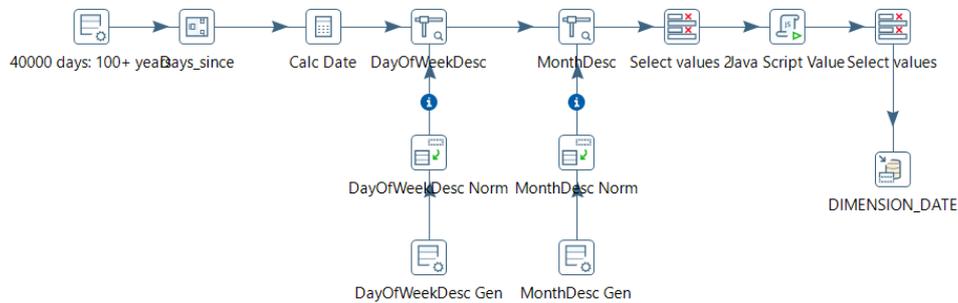


Figura 119. Transformación "Date\_Dim"

## Anexo 7. Definición de Jerarquías

En el presente apartado se detalla la jerarquía de cada dimensión especificada en la Figura 41, de manera que:

- En la Figura 120, se ilustra la jerarquía, niveles y tabla de referencia de la dimensión Users.

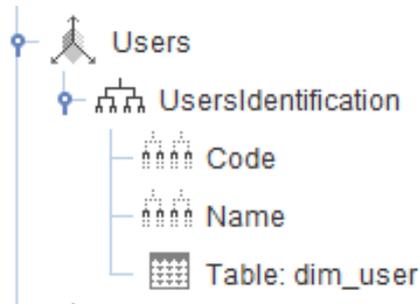


Figura 120. Jerarquía Dimensión Users.

- En la Figura 121, se ilustra la jerarquía, niveles y tabla de referencia de la dimensión Orders.

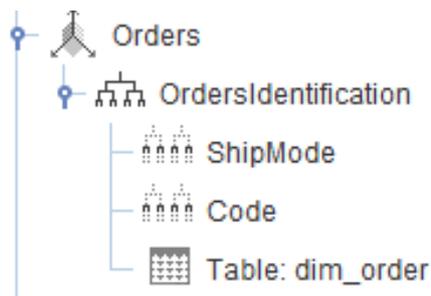


Figura 121. Jerarquía dimensión Orders.

- En la Figura 122, se ilustra la jerarquía, niveles y tabla de referencia de la dimensión Products.

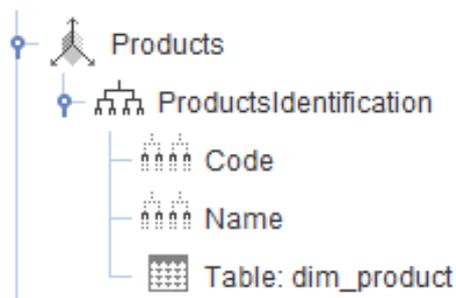


Figura 122. Jerarquía dimensión Products.

- En la Figura 123, se ilustra la jerarquía, niveles y tabla de referencia de la dimensión TypePays.

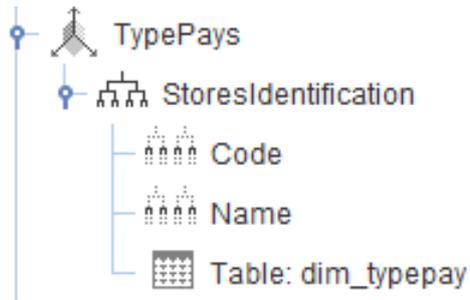


Figura 123. Jerarquía dimensión TypePays.

- En la Figura 124, se ilustra las jerarquías, niveles y tabla de referencia de la dimensión Time.

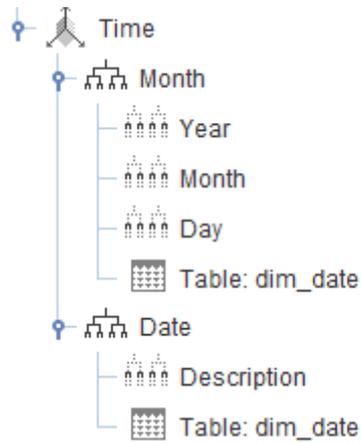


Figura 124. Jerarquía dimensión Time.

## Anexo 8. Configuración para el servidor del sistema.

Para comenzar, es necesario contar con un servidor. En este caso, se creó una cuenta en DigitalOcean a través del siguiente enlace: <https://www.digitalocean.com/>.

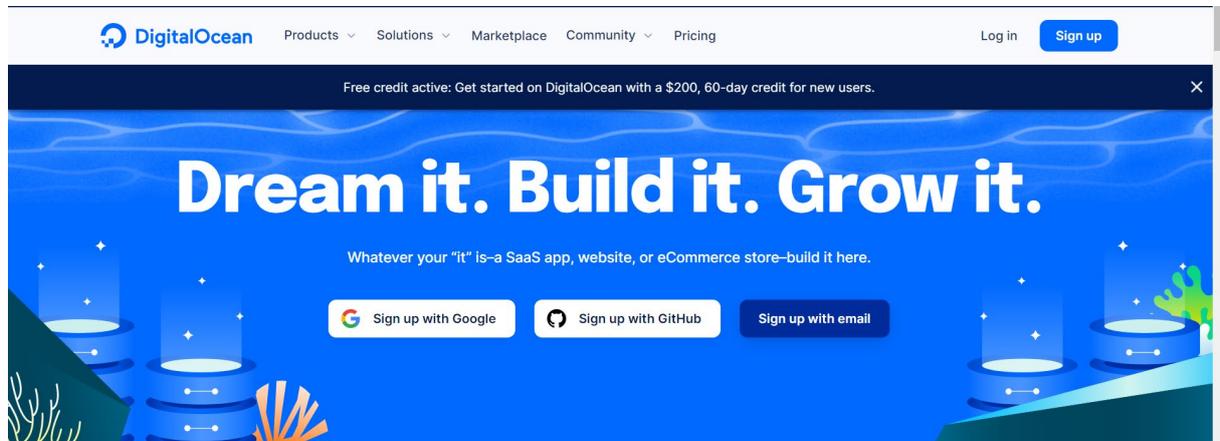


Figura 125. Servidor de DigitalOcean

A continuación, se procede a crear el droplet en el que se instalará el sistema operativo y las herramientas necesarias para el despliegue de la solución BI. Una vez finalizada la instalación, se proporcionará la dirección IP correspondiente para acceder a la aplicación. En la Figura 126, se muestra el droplet pentaho-donesitios y su dirección IP: 162.243.166.210.

### Droplets

<input type="text" value="Search by Droplet name"/>				<a href="#">Create Droplet</a>
Name	IP Address	Created	Tags	
 <b>pentaho-donesitios</b> 2 GB / 25 GB Disk / NYC1 - Ubuntu 22.04 (LTS) x64	162.243.166.210	5 days ago		 <a href="#">More</a>

Figura 126. Droplet pentaho-donesitios

Se puede acceder al servidor desde cualquier consola que permita acceso mediante SSH. En este caso, se hizo uso de la herramienta PUTTY. En la Figura 127, se muestra la configuración para acceder al servidor mediante la herramienta Putty.

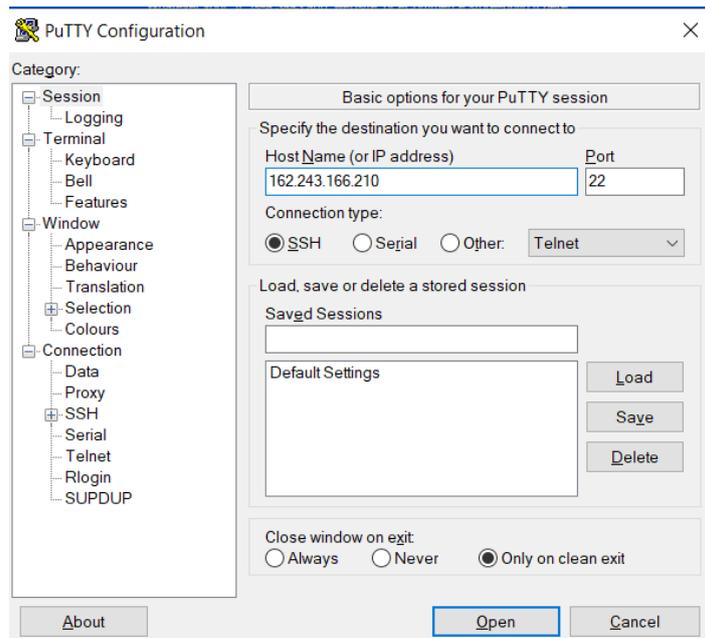


Figura 127. Configuración de Putty.

A continuación, se instalan las herramientas necesarias para la ejecución de la solución BI:

### Instalación de PosgreSQL

- Agregar el repositorio de Oracle Java al sistema operativo:

```
sudo add-apt-repository ppa:webupd8team/java
```

- Instalar Oracle Java 8 utilizando el siguiente comando:

```
sudo apt install oracle-java8-installer
```

- Una vez instalado, se puede verificar que Java esté instalado y funcionando correctamente con el siguiente comando:

```
java -version
```

### Instalación de PosgreSQL

- Instalación PostgreSQL utilizando el siguiente comando:

```
sudo apt install postgresql postgresql-contrib
```

- Iniciar el servicio de PostgreSQL:

```
sudo systemctl start postgresql
```

- Verifica que el servicio está en ejecución utilizando el siguiente comando:

```
sudo systemctl status postgresql
```

- Configuración de usuario y la base de datos de PostgreSQL, creando un usuario y asignándole una contraseña.

```
CREATE USER nombre_usuario WITH PASSWORD 'contraseña'
```

## Anexo 9. Ejecución de pruebas

### Pruebas de carga y estrés

Para la ejecución de las pruebas de carga y estrés se utilizó la herramienta JMeter que permitió diseñar, ejecutar y ver el resultado de la ejecución del plan de pruebas. Las pruebas fueron realizadas a la solución BI que se encuentra alojada en la dirección <http://dashboard-donesitios:8080/pentaho>. A continuación, se describen los pasos realizados durante la realización de las pruebas:

En primer lugar, se creó un grupo de hilos (Thread Group) en el plan de pruebas, tal como se muestra en la Figura 128.

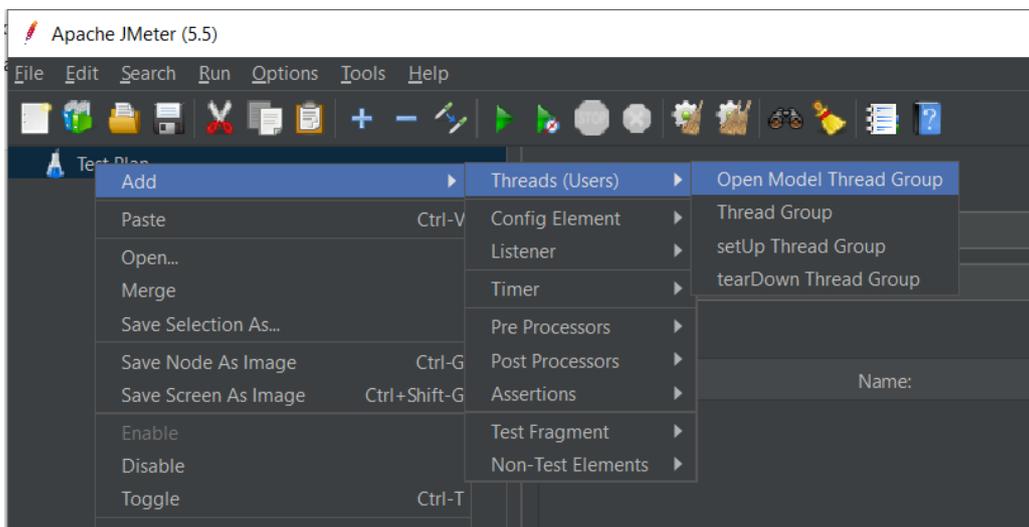


Figura 128. Agregar grupo de Hilos

A continuación, se agregó una petición HTTP de la forma que se indica en la Figura 129.

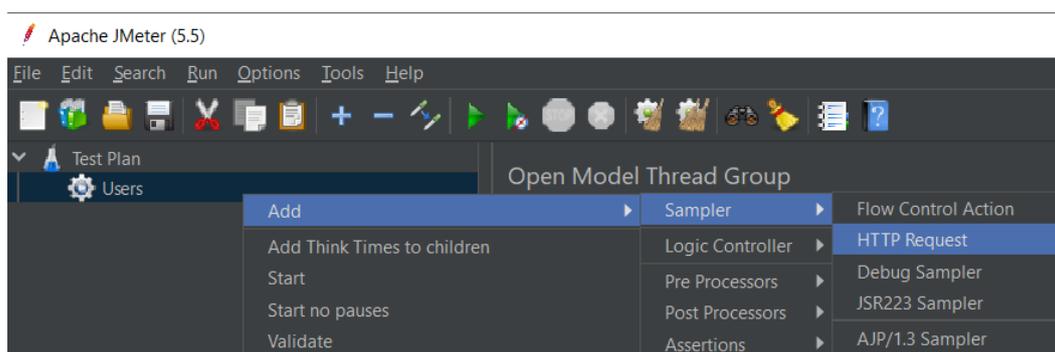


Figura 129. Agregar petición HTTP

Posteriormente, se incluyeron los componentes View results Tree y Summary Report para visualizar los resultados obtenidos durante la ejecución de las pruebas. En la Figura 130, se muestra cómo se añadieron los componentes.

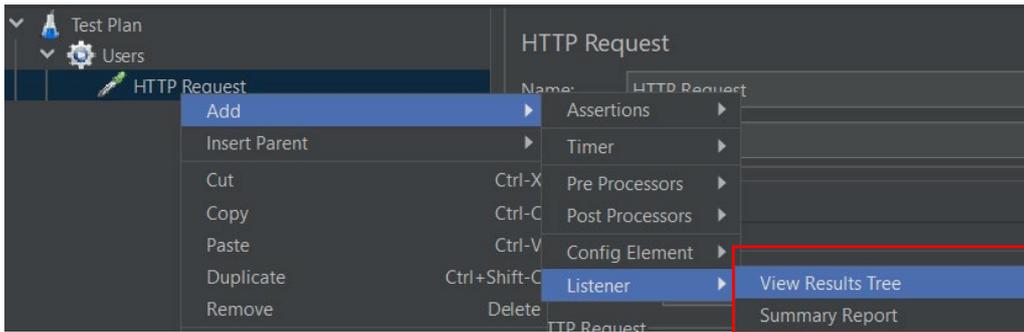


Figura 130. Componentes de visualización de pruebas

Finalmente, en la Figura 131 se muestra los componentes que se utilizaron para llevar a cabo las pruebas de carga y estrés de la solución BI

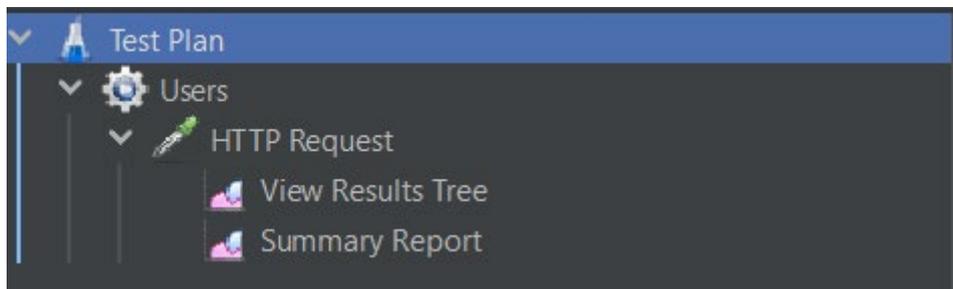


Figura 131. Componentes del plan de pruebas

### Configuración de hilos para el plan de pruebas.

En este punto se configuro el grupo de hilos asignando el nombre, número de hilos (número de peticiones), periodo de subida en segundos (tiempo que toma en aumentar gradualmente el número de peticiones) y el bucle, tal y como se muestra en la Figura 132.

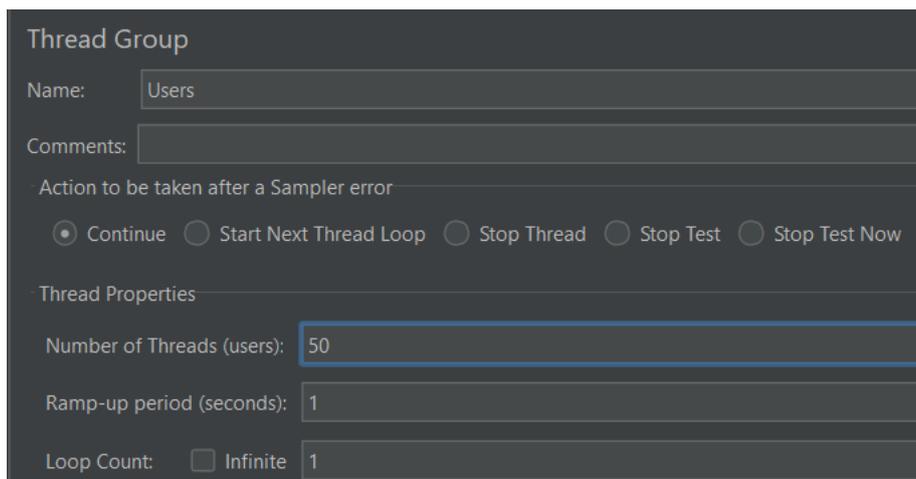


Figura 132. Configuración de hilos

## Configuración de petición HTTP

Durante la configuración de la petición HTTP, se estableció el nombre a la petición, protocolo Http, dirección IP del servidor, tipo de petición y la ruta a probarse. En la Figura 133, se ilustra la configuración realizada.



Figura 133. Configuración de petición HTTP

Para presentar los resultados obtenidos, a continuación, en la TABLA XXXVI, se muestran los tiempos mínimos y máximos de respuesta para cada grupo de peticiones, junto con la cantidad de solicitudes correctas e incorrectas en cada grupo. Todo esto se realizó tras la simulación de varias peticiones concurrentes y la recopilación de datos a partir de 50, 100, 500, 1000 y 2000 solicitudes.

TABLA XXXVI. RESULTADO DE PRUEBAS DE CARGA Y ESTRÉS

# Peticiones	Correctas	Erróneas	% Error	Tiempo mínimo (ms)	Tiempo máximo (ms)
<b>50</b>	50	0	0%	714	2117
<b>100</b>	100	0	0%	848	5809
<b>500</b>	500	0	0%	342	7886
<b>1000</b>	996	4	0.40%	587	31257
<b>2000</b>	1949	51	2.55%	845	39231

## Pruebas de aceptación

Durante las pruebas de aceptación, el gerente de la empresa inició sesión en el dashboard utilizando sus credenciales personales y evaluó la forma en que se accede a la información de pedidos. Para medir el nivel de satisfacción de la solución de Business Intelligence (BI), se aplicó una encuesta en Google Forms utilizando la escala de Likert.

A continuación, en la TABLA XXXVII se muestran los resultados obtenidos en cada una de las preguntas realizadas al gerente de la empresa.

TABLA XXXVII. RESULTADOS DE PREGUNTAS RELACIONADAS CON LA SATISFACCIÓN DEL GERENTE CON RESPECTO A LA SOLUCIÓN BI DESARROLLADA.

	Muy satisfecho	Satisfecho	Moderadamente Satisfecho	Poco satisfecho	No satisfecho
¿La solución BI para la gestión de datos de pedidos en la empresa cumple con los requerimientos solicitados?	X				
¿Considera que el dashboard implementado ha tenido un impacto positivo en el proceso de toma de decisiones de la empresa?	X				
¿Considera usted que la solución BI presenta una interfaz de usuario amigable y de fácil usar?	X				
¿La información presentada en el dashboard es útil para la empresa?	X				
¿Considera que los tiempos de respuesta del dashboard son aceptables?	X				
¿La solución BI ha permitido mejorar la gestión de pedidos en la empresa?	X				

Los resultados de la encuesta muestran que el gerente de la empresa está muy satisfecho en todos los aspectos evaluados en relación al dashboard desarrollados para la gestión de datos de pedidos. Esto sugiere que la solución cumple con los requerimientos solicitados y que el dashboard implementado ha tenido un impacto positivo en el proceso de toma de decisiones de la empresa. Además, la solución presenta una interfaz de usuario amigable y de fácil uso, y la información presentada en el dashboard es útil para la empresa. Los tiempos de respuesta del dashboard también son aceptables, y la solución BI ha permitido mejorar la gestión de pedidos en la empresa.

## **Anexo 10. Validación y aceptación de la solución BI**

Marketing Digital Bravo & Beltrán ha realizado la revisión y uso de la solución desarrollada, en el que se pudo acceder a información de manera más sencilla.

Al acceder al dashboard construido para Marketing Digital Bravo & Beltrán, se pudo constatar la facilidad de acceso a información sobre los pedidos generados, esto permite a la empresa ahorrar tiempo y recursos en la generación de esta información, además facilita la toma de decisiones a la gerencia de la empresa con información valiosa para su gestión.

Por lo tanto, se ha validado y aceptado los reportes presentados en el dashboard construidos para Marketing Digital Bravo & Beltrán.

Atentamente,



---

Ing. Diego Bravo

**GERENTE DE LA EMPRESA MARKETING DIGITAL BRAVO & BELTRÁN**

**Anexo 11.** Certificado de traducción del resumen

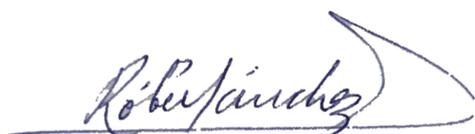
**CERTIFICADO DE TRADUCCIÓN**

Yo, **ROBER DE JESUS SANCHEZ ARMIJOS**, con número de cédula **1102523048** y con título de Licenciado en Ciencias de de la Educación, Mención Idioma Inglés, registrado con número **1031-02-275026**.

**CERTIFICO:**

Que la traducción del resumen del presente Trabajo de Titulación denominado: **“Implementación de Business Intelligence para la gestión de datos generados en el módulo de pedidos durante los años 2018-2021 en la empresa Marketing Digital Bravo & Beltrán”** de autoría de **Luis Miguel Quito Carrión**, portador de la cédula de identidad número **1150024527**, estudiante de la carrera de Ingeniería en Sistemas de la Facultad de la energía, las industrias y los recursos naturales no renovables de la Universidad Nacional de Loja, siendo el mismo una versión correcta de traducción literal del español al inglés. Además, se certifica la fidelidad de la traducción mas no se asume responsabilidad por la autenticidad o el contenido del documento en la lengua de origen.

Es todo en cuanto puedo certificar en honor a la verdad, facultando al interesado hacer uso del presente en lo que se creyera conveniente.



---

**Lic. ROBER DE JESUS SANCHEZ ARMIJOS**

**Nro. de registro SENESCYT: 1031-02-275026**

**C.I: 1102523048**

**CEL: 0960256707**