



**Implementación de un sistema de gestión de datos para la toma de decisiones en
la empresa NUO tecnológica.**

Ing. Pozo Moreno, Diego Fabricio

Vicerrectorado de Investigación, Innovación y Transferencia de Tecnología

Centro de Posgrados

Maestría en gestión de sistemas de información e inteligencia de negocios

Trabajo de titulación, previo a la obtención del título de Magíster en Gestión de
sistemas de información e inteligencia de negocios

Msc. Raura Ruiz, Jorge Geovanny

24 de septiembre del 2020



Document Information

Analyzed document	Tesis Diego Pozo Version Final Uknd.docx (D82332392)
Submitted	10/21/2020 7:05:00 PM
Submitted by	
Submitter email	jgraure@espe.edu.ec
Similarity	3%
Analysis address	jgraure.espe@analysis.arkund.com

Sources included in the report

W	URL: https://docplayer.es/5066300-Pontificia-universidad-catolica-del-ecuador-facultad- ... Fetched: 8/2/2020 12:00:33 AM	5
W	URL: https://docplayer.es/689696-Desarrollo-de-una-aplicacion-de-business-intelligence- ... Fetched: 12/4/2019 7:39:12 AM	3
W	URL: https://www.cic.es/herramienta-de-inteligencia-de-negocio/ Fetched: 10/21/2020 7:07:00 PM	1
W	URL: https://docplayer.es/85100097-Universidad-nacional-del-altiplano-facultad-de-ingen ... Fetched: 3/25/2020 11:10:17 PM	4
SA	1598408437_3 PROYECTO GISEL ARREAGA C 140820 (LEMM).docx Document 1598408437_3 PROYECTO GISEL ARREAGA C 140820 (LEMM).docx (D78174783)	2
SA	1602033223_3 PROYECTO GISEL ARREAGA C COMP.pdf Document 1602033223_3 PROYECTO GISEL ARREAGA C COMP.pdf (D80895335)	2
W	URL: https://docplayer.es/93964332-Datawarehouse-municipalidad-de-cosquin.html Fetched: 6/6/2020 4:44:34 AM	1
SA	TESIS_JF_final BI.docx Document TESIS_JF_final BI.docx (D25038011)	1


Raura Ruiz, Jorge Geovanny

C.C.: 0501773063

DIRECTOR



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

CERTIFICACIÓN

Certifico que el trabajo de titulación, "**Implementación de un sistema de gestión de datos para la toma de decisiones en la empresa NUO tecnológica**" fue realizado por el Ing. **Pozo Moreno, Diego Fabricio** el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 24 de septiembre del 2020.

Firma:

Raura Ruiz, Jorge Geovanny

C.C.: 0501773063



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

RESPONSABILIDAD DE AUTORÍA

Yo, **Pozo Moreno, Diego Fabricio**, con cédula de ciudadanía n° 1721253019, declaro que el contenido, ideas y criterios del trabajo de titulación: **Implementación de un sistema de gestión de datos para la toma de decisiones en la empresa NUO Tecnológica** es de mi autoría y responsabilidad, cumpliendo con los requisitos legales, teóricos, científicos, técnicos y metodológicos establecidos por la Universidad de las Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Sangolquí, 24 de septiembre del 2020.

Firma:

Pozo Moreno, Diego Fabricio

C.C.: 1721253019



VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

AUTORIZACIÓN DE PUBLICACIÓN

Yo, **Pozo Moreno, Diego Fabricio** autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **Implementación de un sistema de gestión de datos para la toma de decisiones en la empresa NUO tecnológica** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 24 de septiembre del 2020.

Firma:

Pozo Moreno, Diego Fabricio

C.C.: 1721253019

DEDICATORIA

No es grande el que siempre triunfa, sino el que jamás se desalienta

José Luis Martín Descalzo

El presente trabajo investigativo lo dedico principalmente a Dios, por ser el inspirador y darme fuerza para continuar en este proceso de obtener uno de los anhelos más deseados en mi vida.

A mis padres Patricio y Fabiola por su amor, trabajo y sacrificio en todos estos años, gracias a ustedes he logrado llegar hasta aquí y convertirme en lo que soy. Ha sido un orgullo y privilegio ser su hijo, son los mejores padres.

A mis hermanos Johanna y Andrés por estar siempre presentes, acompañándome y por el apoyo moral, que me han brindado a lo largo de esta etapa de mi vida.

A mi novia Jenniffer por acompañarme durante todo este arduo camino y compartir conmigo alegrías y fracasos. Gracias por su amor incondicional y por su apoyo en este largo camino.

AGRADECIMIENTO

Agradezco a los docentes de la Maestría en Gestión de Sistemas de Información e Inteligencia de Negocios de la Universidad de las Fuerzas Armadas - ESPE, por haber compartido sus conocimientos a lo largo de la preparación de mi profesión, de manera especial, al Ing. Jorge Geovanny Raura Ruiz, Mgs, tutor de mi trabajo de investigación quien me ha guiado con su paciencia, y su rectitud como docente.

Agradezco a la empresa NUO Tecnológica por permitirme el acceso a la información confidencial de la empresa, al Ing. Victor Espinoza por todo en el apoyo en el transcurso de la carrera, especialmente al Ing. Patricio Avila Gerente General, quien depositó su confianza en mí para realizar mi trabajo de investigación.

Tabla de Contenido

Resumen	13
Abstract.....	14
Capítulo I.....	15
Introducción.....	15
Antecedentes	15
Planteamiento del problema	15
Objetivo general.....	17
Objetivos específicos	17
Justificación, importancia y alcance del proyecto	18
Preguntas de investigación	19
Hipótesis de investigación	20
Categorización de las variables de investigación	20
Capítulo II.....	21
Marco teórico	21
Marco teórico referencial	21
Inteligencia de negocios.....	21
Datos	21
Datos estructurados.....	21
Datos provocados.....	22
Datos tramitados.....	22
Datos compilados.....	22
Datos experimentales.....	22
Datos no estructurados.....	22
Datos no estructurados y semiestructurados	23
Datos generados por el usuario.	23
Almacenes de datos	23
Metodologías de diseño de almacenes de datos.	26
Principales aportaciones de un sistema de gestión de datos.	27
Características fundamentales de un Sistema de Gestión de Base de Datos.	28
Análisis de datos	29
Usos del análisis de datos.....	29
Proceso de análisis de datos.....	30
Herramientas de soporte para la toma de decisiones.....	31
Usabilidad de los sistemas de inteligencia de negocios	32

Capítulo III.....	34
Marco Metodológico.....	34
Enfoque de investigación.....	34
Método de investigación	34
Fases de la investigación.....	34
Situación actual.....	35
Organización de la entrevista.	36
Estado del Arte	37
Motivación.....	37
Objetivo de la búsqueda	37
Criterios de inclusión.....	37
Criterios de exclusión	38
Grupo de control.....	38
Cadena de búsqueda.....	39
Estudios candidatos	40
Cadena de búsqueda ideal	40
Selección de estudios.....	40
Filtros.....	42
Estudios primarios.....	42
Conclusiones o características del estado del arte	47
Capítulo IV	51
Construcción de la solución	51
Planificación del proyecto.....	53
Definición del proyecto	53
Justificación y objetivos.....	53
Alcance	53
Identificación de tareas.....	54
Secuencia de tareas	55
Estimación de recursos para el proyecto	55
Cronograma	57
Análisis de requerimientos	59
Línea tecnológica.....	60
Arquitectura tecnológica.....	60
Selección e instalación de productos	61
Modelado dimensional.....	61

Nivel de granularidad de dimensiones.....	62
Dimensiones	63
Tabla de hechos	64
Modelo físico.....	65
ETL diseño y desarrollo	66
Desarrollo de la aplicación.....	69
Modelo de diseño de BI.....	70
Despliegue	72
Capítulo V.....	75
Validación De La Solución	75
Síntesis de las preguntas de investigación	83
Capítulo VI	87
Conclusiones y Recomendaciones.....	87
Conclusiones	87
Recomendaciones	88
Referencias bibliográficas	90
Anexos	94

Índice de tablas

Tabla 1: Grupo de control	38
Tabla 2: Búsqueda de cadena ideal	39
Tabla 3: Selección de estudios.....	40
Tabla 4: Selección de estudios primarios	42
Tabla 5: Resumen de estudios primarios del estado del arte	47
Tabla 6: Identificación de tareas	54
Tabla 7: Secuencia de tareas	55
Tabla 8: Estimación de recursos para el proyecto	55
Tabla 9: Temas analíticos y procesos de negocio	59
Tabla 10: Matriz de procesos / dimensiones.....	59
Tabla 11: Temas analíticos y procesos de negocio	60
Tabla 12: Productos seleccionados	61
Tabla 13: Análisis dimensional de un proceso de ventas	62
Tabla 14: Nivel de granularidad de dimensiones	62
Tabla 15: DIM_FECHA.....	63
Tabla 16: DIM_PRODUCTO.....	63

Tabla 17: DIM_CLIENTE	64
Tabla 18: DIM_FACTURA	64
Tabla 19: FAC_VENTAS	64
Tabla 20: Medición de tiempos de tareas en el sistema de BI.....	76
Tabla 21: Número de asistencias a los usuarios del sistema BI	76
Tabla 22: Tiempos de ejecución de consulta del sistema de BI	77
Tabla 23: Nivel de participación de los usuarios con sistemas de bi	77

Índice de figuras

Figura 1: Diagrama de Ishikawa	17
Figura 2: Variable Dependiente e Independiente.....	20
Figura 3: Fases de la Metodología ad-hoc	34
Figura 4: Tareas de la metodología de Kimball	52
Figura 5: Esquema de desglose del trabajo	53
Figura 6: Cronograma del proyecto.....	58
Figura 7: Modelo dimensional.....	62
Figura 8: Modelo físico.....	65
Figura 9: Entrada de Datos vs Destino DSA	66
Figura 10: Diseño e Implementación del Data Warehouse	66
Figura 11: Repositorio de datos DSA	67
Figura 12: Proceso de ETL Dimensiones del Data Warehouse	67
Figura 13 Validación carga de datos a las dimensiones y a la tabla de hechos.....	68
Figura 14 Selección de la base de datos de origen	69
Figura 15 Configuración de la conexión con la base de datos	69
Figura 16 Selección de tablas de dimensiones y hechos	70
Figura 17 Panel de campos.....	71
Figura 18 Panel de visualizaciones	71
Figura 19 Diseño del reporte de ventas del área comercial	71
Figura 20 Total de ventas por año y por cliente	72
Figura 21 Cantidad de equipos vendidos por año y por mes.....	73
Figura 22 Proyección de venta por cantidad de equipos trimestral por cliente	73
Figura 23 Cantidad de equipos vendidos por año y por mes.....	73
Figura 24 Modelo conceptual propuesto de evaluación de usabilidad de BI	75
Figura 25 Resultados pregunta 1 de la encuesta usabilidad de sistema de BI.....	77
Figura 26 Resultados pregunta 2 de la encuesta usabilidad de sistema de BI.....	78

Figura 27 Resultados pregunta 3 de la encuesta usabilidad de sistema de BI.....	79
Figura 28 Resultados pregunta 4 de la encuesta usabilidad de sistema de BI.....	79
Figura 29 Resultados pregunta 5 de la encuesta usabilidad de sistema de BI.....	80
Figura 30 Resultados pregunta 6 de la encuesta usabilidad de sistema de BI.....	80
Figura 31 Resultados pregunta 7 de la encuesta usabilidad de sistema de BI.....	81
Figura 32 Resultados pregunta 8 de la encuesta usabilidad de sistema de BI.....	81
Figura 33 Resultados pregunta 9 de la encuesta usabilidad de sistema de BI.....	83

Resumen

A través del tiempo la tecnología ha reducido las barreras para realizar negocios, incrementar ingresos y mejorar procesos en las empresas. En el mundo de los negocios la toma de decisiones debe ser acertada y basada en el análisis de los datos de la empresa, actualmente muchas empresas no tienen su información en un sistema centralizado, dificultando así la toma de decisiones. El principal problema de la empresa en la cual se realizó el caso de estudio es la calidad y el tiempo de acceso a la información que se maneja para la toma de decisiones, debido a que cada empleado maneja distintas versiones de información y las almacena en distintas ubicaciones. El 37,50% del personal del área comercial, asegura que reciben información errónea constantemente, ocasionando pérdidas de negocios por la inconformidad causada a los clientes. El objetivo de esta investigación es implementar un sistema de gestión de datos que, mediante la integración y depuración de la información de la empresa, facilite al departamento comercial el acceso a la información optimizando los tiempos de respuesta para la toma de decisiones. Para esta investigación se propone una metodología personalizada ad-hoc que consta de las siguientes fases: observar el estado actual del fenómeno a estudiar, seguido del estudio del estado del arte, construcción de la solución y por último la validación de la solución. Para la implementación del sistema de gestión de datos se utilizó la metodología del modelo dimensional de Kimball. Los resultados indican un alto porcentaje de aceptación del personal del área comercial ya que el 88,9% de los encuestados están de acuerdo en que el sistema apoya a las actividades del área, debido a que este permitió reducir el tiempo de acceso a la información para la toma de decisiones en un 67.7%, anteriormente preparar y verificar la calidad de la información para la generación de reportes que apoyen a la toma de decisiones llevaba 3 días, en la actualidad se redujo a 1 día, ya que el sistema permite manejar información confiable y el acceso a la misma es en tiempo real, sin depender de terceros.

PALABRAS CLAVE:

- **TOMA DE DECISIONES**
- **SISTEMA CENTRALIZADO**
- **SISTEMA DE GESTIÓN DE DATOS**

Abstract

Over time, technology has reduced the barriers for doing business, increase incomes and improve business processes. In business world, the decision-making process must be accurated and based on the analysis of a company's data. Currently, many companies do not have their information in a centralized system, so it is hard to make decisions. The main problem of the company in which the case study was carried out is the quality and time of access to the information that is handled for decision-making, because each employee handles different versions of information and stores them in different locations. The 37.50% of the comercial department assure that they constantly receive wrong information, which causes lost business because of unsatisfied clients. The objective of this research is to implement a data management system that, through the integration and filtering of a company's information, provides the commercial department access to information, optimizing responding times for decision-making. For this research, a personalized ad-hoc methodology is proposed which consists of the following phases: observing the current state of the phenomenon to be studied, followed by the study of the state of the art, construction of the solution and finally the validation of the solution. For the implementation of a data management system, the Kimball dimensional model methodology was used. The results indicate a high percentage of acceptance of the personnel of the commercial area and that 88.9% of the respondents agree that the system supports the activities of the area, because it will reduce the time of access to information for decision-making by 67.7%, previously preparing and verifying the quality of the information for the generation of reports that support decision-making took 3 days, currently it has been reduced to 1 day, since the system allows handling information reliable and access to it is in real time, without relying on third parties.

KEYWORDS:

- **DECISION MAKING**
- **CENTRALIZED SYSTEM**
- **DATA MANAGEMENT SYSTEM**

Capítulo I

Introducción

Antecedentes

A través del tiempo la tecnología ha reducido las barreras para realizar negocios, incrementar ingresos, mejorar procesos e implementar nuevas herramientas dentro de las compañías.

Sin embargo, hoy por hoy, la implementación de esta ya no es un lujo, o una inversión sino una necesidad fundamental que permite a las grandes y pequeñas empresas estar a la vanguardia de los nuevos tiempos, con procesos competitivos tanto en el mercado nacional como internacional. En el mundo de los negocios la toma de decisiones debe ser acertada y basada en el análisis de los datos más relevantes que la empresa posee, en la actualidad muchas de las empresas no tienen su información organizada en un solo sistema centralizado, dificultando el acceso e integración de la información de interés para la toma de decisiones.

En toda empresa es relevante que todos sus colaboradores, sobre todo, los directamente involucrados, conozcan los procesos del negocio y que los conozcan de manera tal que todos manejen una misma versión. Es sólo así que se logra hacer un uso efectivo de las habilidades y aptitudes de los colaboradores y, a su vez, poder brindarles certeza en el desarrollo de sus funciones.

Planteamiento del problema

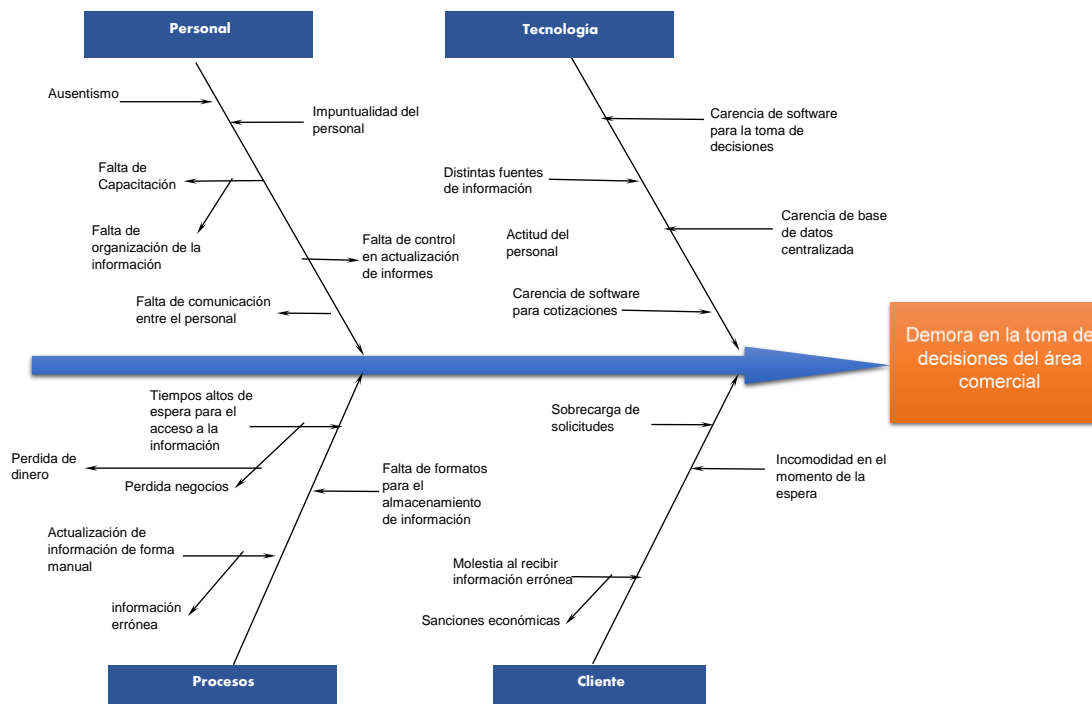
NUO tecnológica es una empresa especialista en crear soluciones tecnológicas de todo tipo en los segmentos como retail, manufactura y educación mediante la distribución de equipamiento con tecnología de punta que permiten automatizar los procesos de producción de cada uno de sus clientes. Mediante una entrevista semi estructurada realizada al gerente general y a dos gerentes regionales del área comercial, se identificó que el principal problema de la empresa es la calidad de la información que

se maneja para la toma de decisiones, debido a que cada uno de los empleados manejan distintas versiones de la información generada en la operación diaria de la empresa, información que es almacenada en sus computadores personales (Archivos de Excel) siendo pocos los empleados que la almacenan en sistemas transaccionales (ERP), causando así demoras en los procesos internos y respuestas a los requerimientos de los clientes.

El área comercial constituida por el gerente general, los gerentes de cuentas corporativas y los especialistas de precio y producto quienes son la cara de la organización frente al cliente, constantemente deben analizar la información de importancia de cada uno de sus clientes como: margen de ganancia, precios históricos de venta de determinado producto o servicio, estatus de las órdenes de compra nacionales e internaciones, stock de productos de alta rotación etc. Es por esto que ellos deben validar con todas las áreas involucradas que la información solicitada sea correcta y se encuentre actualizada para poder tomar decisiones acertadas, ocasionando así demoras y molestias en los clientes. El 37,50% del personal del área comercial informa que reciben información errónea constantemente, ocasionando reprocesos y perdidas de oportunidades de negocios. Esta problemática se la presenta en el siguiente diagrama de Ishikawa:

Figura 1

Diagrama de Ishikawa



Objetivo general

Implementar un sistema de gestión de datos que, mediante la integración y depuración de toda la información de la empresa, facilite al departamento comercial el acceso a la información más relevante optimizando los tiempos de respuesta para la toma de decisiones.

Objetivos específicos

OE1: Identificar el estado actual en el acceso a la información más relevante para la toma de decisiones en el departamento comercial para reconocer los principales inconvenientes y definir el problema principal que tiene el departamento comercial.

OE2: Identificar las técnicas para la implementación de sistemas de gestión de datos que son comúnmente utilizadas en las empresas para la centralización, integración y depuración de la información de una empresa a través de una revisión inicial de literatura.

OE3: Implementar un sistema de gestión de datos en la empresa NUO Tecnológica a través de la metodología de Kimball para la construcción de un almacén de datos.

OE4: Comprobar los resultados del sistema de gestión de datos implementado, mediante la definición de métricas que permitan medir la disminución del tiempo en la toma de decisiones por parte del área comercial.

Justificación, importancia y alcance del proyecto

En la actualidad uno de los elementos fundamentales de cualquier organización comercial se centra en la información que se maneja, debido al valor que posee dicha información, el manejo de esos datos se vuelve un factor trascendental en el desarrollo de las empresas, ya que muchas veces la correcta gestión y análisis de esos datos determinan el éxito del negocio, es por esto que muchas empresas deciden implementar un sistema de gestión de datos. Los aspectos principales que se buscan en estos sistemas son:

- Ayude a la organización a responder con precisión preguntas sobre el negocio.
- Mejoren la disponibilidad de datos ya recopilándolos de distintas ubicaciones y fuentes dispares (por ejemplo, marketing, ventas, finanzas, etc.) en un depósito central.
- Reduzcan el tiempo requerido para encontrar y analizar datos importantes (se logra eficiencia en los flujos de trabajo operativos dado que la actividad analítica se traslada a un sistema separado).
- Faciliten la toma de decisiones para aportar una ventaja competitiva de la organización.
- Un sistema que almacena información de manera estructurada apoya a la toma de decisiones ya que permite conocer el estado actual del negocio mediante la

consolidación de datos, análisis, generación de informes y revelación de tendencias.

- Este proyecto tiene como objetivo centralizar la información de la organización permitiendo al área comercial tener conocimiento del estado actual del negocio para la toma de decisiones, estandarizando y transformando los datos en un sistema de código único, permitiendo que cada uno de los departamentos que componen la empresa, puedan utilizar el mismo repositorio de datos para sus consultas y generación de informes individuales.

Preguntas de investigación

En función de cada uno de los objetivos específicos, se plantean las siguientes preguntas de investigación:

OE1-RQ1.- ¿El uso de técnicas cualitativas de levantamiento de información permitirá determinar el estado actual del acceso a la información para de la toma de decisiones en el departamento comercial?

OE1-RQ2.- ¿El estudio del estado actual, permite identificar el principal problema para la toma de decisiones en el área comercial?

OE2-RQ1.- ¿Los artículos encontrados en la revisión inicial de literatura permiten identificar técnicas para la implementación de sistemas de gestión de datos?

OE2-RQ2.- ¿Cuáles son los estudios relacionados con la integración y centralización de la información de una empresa proveniente de distintas fuentes?

OE3-RQ1.- ¿La metodología de diseño de sistemas de gestión de datos de Kimball permite identificar los beneficios del área comercial para el acceso a la información?

OE3-RQ2.- ¿Cómo beneficia la implementación de un sistema de gestión de datos al área comercial en la empresa NUO Tecnológica?

OE4-RQ1.- ¿La implementación de un sistema de gestión de datos permite al personal del área comercial el acceso a información confiable para la toma de decisiones?

OE4-RQ2.- ¿La implementación del sistema seleccionado permite que el personal del área comercial optimice tiempos en la toma de decisiones eliminando los reprocesos en la validación de información?

Hipótesis de investigación

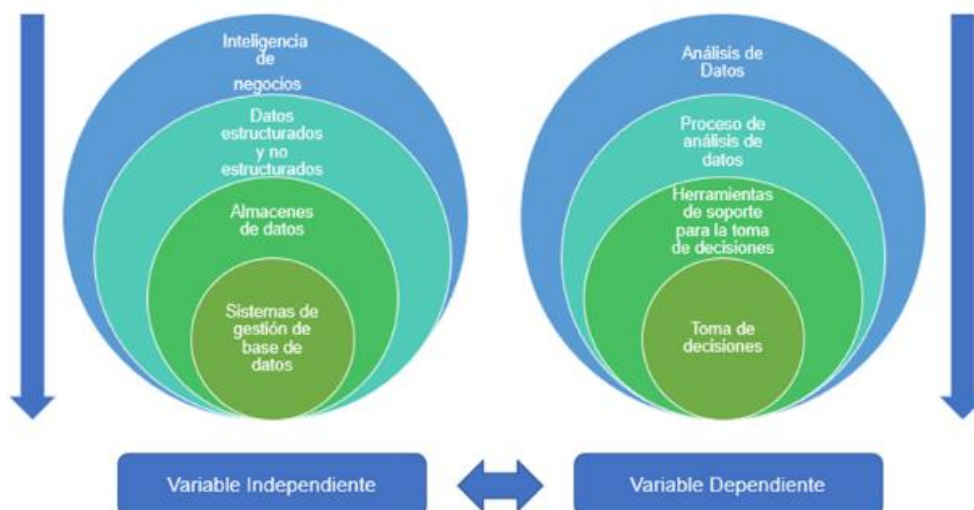
La implementación de un sistema de gestión de datos en la empresa NUO tecnológica permite al área comercial, el acceso a información de calidad para la toma de decisiones, reduciendo tiempos de respuesta y errores en el manejo de la información.

Categorización de las variables de investigación

- **Variable Independiente:** Sistema de Gestión de Datos.
- **Variable Dependiente:** Reducción de tiempos de respuesta y errores en el manejo de la información para la toma de decisiones.

Figura 2

Variable dependiente e independiente



Capítulo II

Marco teórico

Marco teórico referencial

Inteligencia de negocios

Es la técnica de convertir los datos en información, y la información en conocimiento, de forma que se pueda mejorar el proceso de toma de decisiones en los negocios.

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio (Matamala, Matamala, Cuello, & Leiva², 2011).

Datos

Los datos creados fuera de una organización se pueden agrupar en dos grandes categorías:

Datos estructurados.

Datos creados. Las empresas de datos generan información para la investigación de mercados. Esta acción puede consistir en elaborar encuesta a clientes o grupos de discusión. El proceso también incluye métodos más modernos como la creación de un programa de fidelización que recopile información de los consumidores o bien pidan a un usuario que creen una cuenta y entren cuando están comprando (Serrano-Cobos, 2014).

Datos provocados. Dar la oportunidad a las personas a expresar sus puntos de vista genera volumen de datos. Cada vez que el cliente valora un restaurante o la experiencia de compra, es un sumatorio de información (Rayo, 2016).

Datos tramitados. Las empresas recogen datos sobre todas las transacciones completadas, si la compra se ha completado a través de un carrito de la compra online o directamente en el punto de venta.

Las organizaciones también recopilan datos sobre el proceso de una compra online. Por ejemplo, un cliente puede hacer clic en un anuncio del home o un banner en google que les conduce a las páginas del producto que van a comprar.

Los bancos por ejemplo también son grandes consumidores de los datos transaccionales (Schonberger, 2013).

Datos compilados. Los censos de población, el número de coches matriculados cada año, el volumen de compra en un supermercado son datos compilados en grandes bases de datos de todos los hogares. Cada país tiene su propio sistema. Los individuos tenemos una serie de información básica registrada que incluye nuestra identificación, familia, estudios, antecedentes, datos médicos, etc. (Tascón, 2013).

Datos experimentales. Se crean datos experimentales cuando las empresas experimentan con diferentes acciones de marketing y mensajes para ver cuáles son los más efectivos hacia los consumidores.

También podemos interpretar los datos experimentales como una combinación de datos creados y transaccionales. (Schonberger, 2013).

Datos no estructurados. Son aquellos datos no almacenados en una base de datos tradicional (Joyanes, 2013). La información no estructurada no puede ser almacenada en estructuras de datos relacionales predefinidas. Se pueden establecer diferentes clasificaciones, se han considerado dos de ellas.

Datos no estructurados y semiestructurados. Los datos semiestructurados serían aquellos datos que no residen de bases de datos relacionales, pero presentan una organización interna que facilita su tratamiento, tales como documentos XML y datos almacenados en bases de datos NoSQL (Shenglei Pei, 2014).

Datos de tipo texto y no-texto. Datos no estructurados de tipo texto podrían ser datos generados en las redes sociales, foros, e-mails, presentaciones Power Point o documentos Word, mientras que datos no-texto podrían ser ficheros de imágenes jpeg, ficheros de audio mp3 o ficheros de video tipo flash.

Datos capturados. Los datos capturados se crean de forma pasiva debido a la conducta de una persona. Cada vez que alguien entra en un término de búsqueda en Google este motor, utiliza la información capturada para un futuro beneficio.

La información del GPS en los smartphones es otro ejemplo de recogida de datos pasiva que puede ser capturado con las plataformas tecnológicas del Big Data.

Datos generados por el usuario. Los datos generados por los usuarios se componen de todos los datos de las personas como son aquellos que están introduciendo en Internet diariamente. Desde tweets, mensajes de Facebook, a los comentarios en las noticias, videos subidos en YouTube.

Los individuos están creando una enorme cantidad de datos que las empresas pueden utilizar para mejorar su relación con los consumidores y obtener información sobre productos.

Almacenes de datos

Almacenes de datos se refiere a la base de datos corporativa que incorpora y depura la información que se encuentra en una o varias fuentes distintas, para procesar y analizar los datos desde diferentes perspectivas y agilizando su respuesta. La implementación de un data warehouse permite instaurar una solución eficaz y fiable de Business Intelligence.

La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información (modelos de tablas en estrella, en copo de nieve, cubos relacionales... etc.). Este tipo de persistencia de la información es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma (siempre en un entorno diferente a los sistemas operacionales) (Serrano-Cobos, 2012).

La ventaja principal de este tipo de sistemas se basa en su concepto fundamental, la estructura de la información. Este concepto significa el almacenamiento de información homogénea y fiable, en una estructura basada en la consulta y el tratamiento jerarquizado de la misma, y en un entorno diferenciado de los sistemas operacionales. Según definió Bill Inmon (2002) , el data warehouse se caracteriza por ser:

- **Integrado:** los datos almacenados en el data warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- **Temático:** sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del data warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.
- **Histórico:** el tiempo es parte implícita de la información contenida en un data warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el data warehouse sirve, entre otras cosas, para

realizar análisis de tendencias. Por lo tanto, el data warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.

- **No volátil:** el almacén de información de un data warehouse existe para ser leído, y no modificado. La información es por tanto permanente, significando la actualización del data warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía (p. 151).

Otra característica de un data warehouse es que contiene metadatos, es decir, datos sobre los datos. Los metadatos permiten saber la procedencia de la información, su periodicidad de refresco, su fiabilidad, forma de cálculo, etc.

Los metadatos son los que permiten reducir y mecanizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales (Inmon, 2002).

Los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, son:

Dar soporte al usuario final, ayudándole a acceder al data warehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Influye en levantar consultas, informes y análisis, mediante herramientas de Business Intelligence como DSS, EIS o CMI.

Dar soporte a los responsables técnicos del data warehouse en aspectos de auditoría, gestión de la información histórica, administración del data warehouse, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos, etc.

Por último, destacar que para comprender íntegramente el concepto de data warehouse, es importante entender cuál es el proceso de construcción del mismo,

denominado ETL (Extracción, Transformación y Carga), a partir de los sistemas operaciones de una compañía (Sretenović; Jovanović, 2016):

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

Metodologías de diseño de almacenes de datos. Existen muchas metodologías de diseño y construcción de data warehouse, entre todas estas se imponen dos metodologías, la de Kimball y la de Inmon, desde el punto de vista arquitectónico, la mayor diferencia entre estos dos autores es el sentido de la construcción del data warehouse, esto es comenzando por los data marts o ascendente (Bottom-up, Kimball) o comenzando con todo el data warehouse desde el principio, o descendente (Top-Down, Inmon).

Por otra parte, la metodología de Kimball, se basa en un modelo dimensional (no normalizado), al contrario de la de Inmon que se fundamenta en conceptos bien conocidos del diseño de bases de datos relacionales.

Metodología de Kimball. Esta metodología consiste en lo que Kimball define como el Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida del proyecto de data warehouse está basado en cuatro principios básicos:

- **Centrarse en el negocio:** Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis de este y la competencia consultiva de los implementadores.

- **Diseño de infraestructura:** consiste en el diseño de una base de datos única, centralizada, y de alto procesamiento donde se identifica los requerimientos de negocio más relevantes de la empresa.
- **Realizar entregas en incrementos significativos:** crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.
- **Ofrecer una solución:** consiste en suministrar las herramientas necesarias para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad-hoc, sistemas para el análisis y la elaboración de informes, capacitación, soporte, sitio web y documentación.

Metodología de Inmon. Según Inmon, un almacén de datos es una colección de datos orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad que la utiliza. Se trata, sobre todo, de un historial completo de la organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de los datos (Rivadera, 2010).

Principales aportaciones de un sistema de gestión de datos.

- Proporciona una herramienta para la toma de decisiones en cualquier área funcional, basándose en información integrada y global del negocio.

- Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- Proporciona la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- Simplifica dentro de la empresa la implantación de sistemas de gestión integral de la relación con el cliente.
- Supone una optimización tecnológica y económica en entornos de Centro de Información, estadística o de generación de informes con retornos de la inversión espectaculares (Serrano-Cobos, 2012)
- Sistema de gestión de base de datos

Un Sistema de Gestión de Bases de Datos (SGBD) consiste en una colección de datos interrelacionados y un conjunto de programas para acceder a los mismos. Normalmente en el núcleo de un sistema de información se sitúa un SGBD. El caso de los sistemas de información es un poco diferente ya que en principio las bases de datos espaciales no son adecuadas para su manejo con SGBD tradicionales (Rigaux & Voisard, 2002).

Características fundamentales de un Sistema de Gestión de Base de Datos.

Un SGBD permite el almacenamiento, manipulación y consulta de datos pertenecientes a una base de datos organizada en uno o varios ficheros (Guevara, 2019). En el modelo más extendido (base de datos relacional) la base de datos consiste, de cara al usuario, en un conjunto de tablas entre las que se establecen relaciones. A pesar de sus semejanzas (ambos manejan conjuntos de tablas) existen una serie de diferencias fundamentales entre un SGBD y un programa de hoja de cálculo, la principal es que un SGBD permite:

- En lugar de primarse la visualización de toda la información, el objetivo fundamental es permitir consultas complejas, cuya resolución está optimizada, expresadas mediante un lenguaje formal.
- El almacenamiento de los datos se hace de forma eficiente, aunque oculta para el usuario y normalmente tiene, al contrario de lo que ocurre con las hojas de cálculo, poco que ver con la estructura con la que los datos se presentan al usuario.
- El acceso concurrente de múltiples usuarios autorizados a los datos, realizando operaciones de actualización y consulta de los mismos garantizando la ausencia de problemas de seguridad (debidos a accesos no autorizados) o integridad (pérdida de datos por el intento de varios usuarios de acceder al mismo fichero al mismo tiempo (Rigaux & Voisard, 2002).

Análisis de datos

El análisis de datos es la ciencia que se encarga de examinar un conjunto de datos con el propósito de sacar conclusiones sobre la información para poder tomar decisiones, o simplemente ampliar los conocimientos sobre diversos temas. El análisis de datos es la ciencia que se encarga de examinar un conjunto de datos con el propósito de sacar conclusiones sobre la información para poder tomar decisiones, o simplemente ampliar los conocimientos sobre diversos temas (QuestionPro, 2018).

Usos del análisis de datos. El análisis de datos se utiliza en muchas industrias, independientemente del ramo, nos da las bases para tomar o no una decisión o cerciorarnos si una hipótesis es cierta o no.

- **Mercadotecnia:** el análisis de datos se ha usado principalmente para predecir el comportamiento de los consumidores, incluso para poder calificarlo. Conoce cómo hacer un análisis de datos para tu campaña de marketing.

- **Recursos Humanos:** el análisis de datos también es muy útil dentro de las empresas para mantener un buen clima laboral, y fuera de ella, calificando empleados potenciales.
- **Académicos:** Al igual que las empresas el análisis de datos también está presente en la educación, sirve para seleccionar a los alumnos de nuevo ingreso y para medir el rendimiento de los estudiantes (QuestionPro, 2018).

Proceso de análisis de datos. El proceso de Análisis de datos conlleva la recolección, transformación, limpieza y modelado de datos con el objetivo de descubrir información útil y trascendente para los intereses de la organización. Los resultados así obtenidos se comunican, se sugieren conclusiones y se usan para apoyar la toma de decisiones. La visualización gráfica de los datos se utiliza con frecuencia para retratar los datos y conseguir descubrir patrones útiles con mayor facilidad.

El proceso de análisis de datos consiste en las siguientes fases de naturaleza iterativa:

- **Especificación de los requisitos de datos:** los datos para el análisis se basan en una pregunta o un experimento basándose en los requisitos, se identifican los datos necesarios, desde la población o colección de datos a las variables o atributos específicos de los mismos.
- **Recopilación de datos:** la recopilación de datos es el proceso de recolección de información sobre las variables seleccionadas como requisitos de datos. El énfasis está en garantizar la recolección exacta y honesta de datos.
- **Procesamiento de datos:** los datos que se recogen deben ser procesados u organizados para su análisis, esto incluye estructurar los datos según sea necesario para las herramientas de análisis pertinentes.
- **Limpieza de datos:** los datos procesados y organizados pueden estar incompletos, contener duplicados o contener errores. La limpieza de datos es el

proceso de prevenir y corregir estos errores. Existen varios tipos de limpieza de datos que dependen del tipo de datos.

- **Análisis de los datos:** los datos que se procesan, organizan y limpian estarían listos para el análisis. Se disponen de varias técnicas de análisis de datos para comprender, interpretar y derivar conclusiones basadas en los requisitos.
- **Comunicación:** los resultados del análisis de los datos deben presentarse en un formato tal como lo requieran los usuarios para apoyar sus decisiones y acciones futuras. La retroalimentación de los usuarios podría resultar en un análisis adicional (Marketing-analítico, 2017).

Herramientas de soporte para la toma de decisiones

Hoy en día el entorno que rodea a las empresas es cada vez más cambiante, de modo que se necesita una mayor capacidad para la toma de decisiones rápidas que sean capaces de anteponerse y dar respuesta a los cambios. El Business Intelligence ofrece diversas herramientas que permiten facilitar la toma de decisiones mediante el análisis de los datos relativos a una organización, con el objetivo de comprender el funcionamiento de la empresa tanto en el pasado como en la actualidad para poder prever las estrategias más adecuadas para el futuro.

Las herramientas tradicionales de tratamiento de datos no permiten dar respuestas rápidas para la toma de decisiones, ya que tan solo se basan en la acumulación de datos.

Los principales inconvenientes de no utilizar herramientas BI son (Cruz, 2017):

- Poca flexibilidad para extraer datos.
- Necesidad de personal con competencias muy específicas para poder obtener métricas o informes.
- El tiempo de respuesta es significativamente mayor al que ofrecen las herramientas BI.

- Falta de integración entre los sistemas de obtención de datos y la generación de información.
- Datos inexactos, desfasados o insuficientes.
- Dificultades para adaptar la información al puesto de cada trabajador.

Usabilidad de los sistemas de inteligencia de negocios. Una de las fases más importantes en el desarrollo del software, hablese en este caso de los sistemas de inteligencia de negocios, es la fase de pruebas debido a que las mismas pueden consumir hasta el 50% del coste total del sistema y este porcentaje es mayor cuando se trata de sistemas críticos. Las pruebas tienen como objetivos eliminar todos los errores que puedan aparecer en la mínima cantidad de tiempo y esfuerzo, a la vez que buscan incrementar la confianza en los usuarios en que el sistema cumple con los requisitos deseados. Las pruebas en el desarrollo de software pueden ser unitarias, de integración, de sistema y de aceptación (Guamán., 2019).

Las pruebas tienen como objetivos eliminar todos los errores que puedan aparecer en la mínima cantidad de tiempo y esfuerzo, a la vez que buscan incrementar la confianza en los usuarios en que el sistema cumple con los requisitos deseados. Las pruebas en el desarrollo de software pueden ser unitarias, de integración, de sistema y de aceptación.

Por otro lado, la usabilidad se define como la medida de la calidad de la experiencia que tiene un usuario cuando interactúa con un sistema, para determinar que un sistema sea fácil de usar y aprender, sea eficiente en el uso de los elementos ofrecidos en pantalla y sea efectivo en el cumplimiento de las tareas que se pueden llevar a cabo a través de ella (Sanchez, 2011). Dentro de los factores que determinan la usabilidad podemos mencionar la accesibilidad, legibilidad, navegabilidad, facilidad de aprendizaje, velocidad de utilización, eficiencia del usuario y tasas de error (Sanchez, 2011). Según (Poropat, 2014), los factores que evalúan la usabilidad en los sistemas de inteligencia de

negocios son: facilidad de aprendizaje del sistema, la eficiencia del sistema, la flexibilidad del sistema, la participación de los usuarios del sistema y la satisfacción del usuario.

En relación, a los atributos, se tiene:

- **Flexibilidad del sistema:** Hace referencia proporcionar la capacidad de ampliar o cambiar las funcionalidades existentes de una manera adecuada y conveniente (Hasan, 2015).
- **Facilidad de aprendizaje del sistema:** es una parte de la eficacia. Es la capacidad del usuario para operar el sistema a un cierto nivel de competencia definido después de un periodo definido de entrenamiento (Poropat, 2014).
- **Eficiencia del sistema:** Evalúa en qué medida un usuario específico consume recursos con el fin de lograr sus objetivos con eficacia.
- **Participación de los usuarios del sistema:** es una cualidad de la experiencia del usuario con la tecnología que depende de la estética, la novedad y la facilidad de uso del sistema. Es la capacidad del usuario para atender y participar en la experiencia, y la evaluación general del usuario de la experiencia (Poropat, 2014).
- **Satisfacción de los usuarios:** Refleja el nivel de satisfacción de los usuarios a utilizar el sistema en un contexto particular (Hasan, 2015). Evalúa la forma en que el usuario se sienta sobre el sistema y cuáles son sus opiniones, en el caso de los sistemas de inteligencia de negocios se en que grado afecta la capacidad de uso del sistema de BI.

Capítulo III

Marco Metodológico

Enfoque de investigación

Los enfoques para una investigación pueden ser tres: cuantitativo, cualitativo y mixto en este caso particular se propone utilizar un enfoque cualitativo exploratorio.

Método de investigación

Para esta investigación se propone una metodología personalizada ad-hoc que consta de las siguientes fases: observar el estado actual del fenómeno a estudiar, seguido del estudio del estado del arte para verificar todas las soluciones propuestas en estudios relacionados, construcción de la solución y por último la validación de la solución propuesta. Este método permite combinar la reflexión racional o momento racional (la formación de hipótesis y la deducción) con la observación de la realidad o momento empírico.

Fases de la investigación

Figura 3:

Fases de la metodología AD-HOC



Fase 1: Situación actual.- Esta fase consta en observar el estado actual de la problemática propuesta, para esta fase se propone la técnica de observación científica la misma que permite observar con un objetivo claro, definido y preciso, para adquirir información sobre la problemática identificada.

Fase 2: Estado del arte.- Se debe realizar un proceso empírico, es decir un proceso a través de la experiencia, es decir se realiza el estado del arte de la problemática identificada para poder obtener información de estudios relacionados con la misma.

Fase 3: Construcción de la solución.- Se propone procesos y actividades que sirven para garantizar la verdad y confirmar que la hipótesis propuesta en esta investigación es verdadera.

Fase 4: Validación.- Se debe verificar que la solución propuesta funciona en el lugar que se plantea implementar.

En este capítulo presentamos los resultados de la primera y segunda fase de la investigación que corresponde a la situación actual y el estado del arte las mismas que se detallan a continuación:

Situación actual

Para el análisis del estado actual en el acceso a la información de la empresa NUO Tecnológica se utilizó la técnica de investigación cualitativa llamada entrevista semiestructurada, dicha entrevista se realizó a dos gerentes regionales del área comercial y al gerente general lo que permitió identificar las principales problemáticas que tiene el área comercial en el acceso a la información para la toma de decisiones como se muestra en el **Anexo 1**.

Para realización de la entrevista semiestructurada se siguieron los siguientes pasos:

Identificar a las personas adecuadas para la entrevista. Para este punto se identificó que se debe realizar la entrevista al gerente general y los gerentes regionales

de Quito y Guayaquil que son las sedes principales de la empresa NUO Tecnológica, quienes conocen muy bien el negocio y son los encargados de la toma de decisiones en la empresa, ya que tienen a su cargo dirigir a todo el equipo comercial.

Organización de la entrevista.

- **Presentación y propósito de la entrevista:** Se explicó a cada uno de los entrevistados que el motivo de la entrevista es identificar las problemáticas actuales en el manejo de la información dentro de la empresa para la toma de decisiones, con el objetivo de identificar la mejor solución para mitigar las problemáticas identificadas y generar un cambio en el manejo de la información dentro de la empresa.
- **Consentimiento y confidencialidad:** Se solicitó el consentimiento de cada uno de los entrevistados ya que la entrevista fue grabada y anónima con el objetivo de recopilar datos que apoyen al objetivo de esta investigación.
- **Lugar y horario:** La entrevista fue realizada en la sala de reuniones de la empresa NUO tecnológica, con una duración de 30 minutos cada entrevista.

Como resultado de las entrevistas realizadas se identificó que el problema principal para el acceso a la información dentro de la empresa es la demora y la calidad de la información que se maneja en el área comercial, puesto que para consultar los precios de venta históricos de cada uno de sus clientes en determinado producto o servicio, los gerentes de cuentas corporativas deben acudir a los especialistas en precio y producto que son los encargados de realizar las cotizaciones, las mismas que son almacenadas en cada uno de sus computadores en archivos de Excel y PDF. Si el especialista de precio y producto no está seguro de un precio histórico o no tiene conocimiento del mismo, debe solicitar al área financiera validar las facturas emitidas hasta la fecha en el sistema transaccional para poder conocer cuál fue el último precio de venta de un producto o servicio de determinado cliente, causando así demoras en el envío

de cotizaciones y errores en precios establecidos para cada uno de los clientes, adicional a esto al cerrar el año se conoce el total de facturación anual, pero no se puede discriminar con facilidad el monto de cada línea de producto por marca que oferta la empresa NUO Tecnológica. Esta situación dificulta el análisis del total de ventas por marca para verificar si se cumplió la meta propuesta y el compromiso adquirido con cada una de ellas. Este reporte actualmente se lo realiza en una hoja de Excel casi de manera manual y lleva al gerente comercial aproximadamente 5 días laborables, dejando varias de sus actividades de lado para poder realizar dicho reporte.

Estado del Arte

Para el análisis del estado del arte se utilizó la estrategia de revisión sistemática de literatura, utilizando las fases de criterios de inclusión y la estrategia de búsqueda, para ello como fuente de búsqueda de información de esta investigación se utilizó el repositorio académico IEEEExplore.

Motivación

La revisión de todos los estudios relacionados que permitan conocer las metodologías y soluciones propuestas con respecto a la implementación de un sistema de gestión de datos que apoye a la toma de decisiones en una empresa.

Objetivo de la búsqueda

Encontrar el estado del arte para la implementación de un sistema de gestión de datos que permita centralizar la información histórica de la empresa para la toma de decisiones.

Criterios de inclusión

- Se tomarán en cuenta los estudios que describan las metodologías para la implementación de un sistema de gestión de datos.
- El artículo debe contener información sobre el diseño de reportes (dashboard) personalizados.

- Los estudios deben tener información sobre metodologías y herramientas para la implementación de un sistema de gestión de datos.
- El artículo debe contener información sobre los beneficios de implementar un sistema de gestión de datos en las empresas.
- Estudios que contengan información sobre la toma de decisiones en las empresas.

Criterios de exclusión

- Los estudios que no contengan información sobre tipos de base de datos para el almacenamiento de la información.
- Los estudios que contengan información sobre el análisis de algoritmos de predicción.
- Los estudios que contengan información de toma de decisiones no relacionados con bases de datos empresariales.

Grupo de control

Para formar el grupo de control se obtuvieron varios artículos relacionados con el análisis, diseño e implementación de un sistema de gestión de datos y de herramientas de inteligencia de negocios para la toma de decisiones en las empresas, de los cuales se seleccionaron cuatro artículos los cuales confirman el grupo de control que se detalla en la Tabla 1:

Tabla 1:

Grupo de control

Grupo de Control	Título	Palabras Clave
EC01	El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos Autor: Leticia Calzada, José Luis Abreu Fecha: 2009-09-15	Business intelligence, information technology, decisions, information, computer science, data base, implementation, sales processes
EC02	Inteligencia de negocios: estrategia para el desarrollo de competitividad en empresas de base tecnológica Autor: Eduardo Ahumada Tello, Juan Manuel Perusquia Velasco Fecha: 2016-10-15	Business intelligence; Business competitiveness; Knowledge management; Business innovation; Decision making process
EC03	Análisis, Diseño, Construcción e Implementación de un Data Warehouse para Toma de Decisiones y Construcción de los kpi, para la Empresa Kronosconsulting Cia Ltda. Autores: Ximena Lozada Peñafiel, Holger Cruz Tamayo, Washington Pérez Argudo, Andrés de la Torre Díaz Fecha: 2017-06-17	Computer tools, data analysis, decision making, logistics processes Business Intelligence, Indicators, Data Warehouse, Scorecard
EC04	Implementación de un Datamart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso Autor: Julio YalanCastillo, Luis alomino Paniora Fecha: 2012-04-12	Datamart, systems, Datawarehouse, Bussines Intelligence, logistica shopping, Company

Cadena de búsqueda

Para armar las diferentes cadenas de búsqueda se clasifican las palabras claves, para esto se toman las palabras que más se repiten en cada grupo de control, a continuación, analizamos los resultados de búsqueda en el repositorio académico IEEEExplore.

Tabla 2:

Búsqueda de cadena ideal

N.-	Cadena de Búsqueda	# Estudios	EC01	EC02	EC03	EC04
C1	("Implementation" OR "application" OR "employment") AND ("data base" OR "data warehouse" OR "datamart") AND ("company" OR "Group" OR "club") AND ("sales processes") AND("Information" OR "data analysis")	19	X			X
C2	("data"OR"information") AND ("analysis" OR "investigation") AND ("Data warehouse") AND ("implementation") AND ("Business Intelligence")	36	X	X	X	
C3	("Data Warehouse" OR "datamart") AND ("logistics processes" OR "sales processes") AND ("decision making") AND ("implementation" OR "application")	31	X		X	
C4	("implementation" OR "application") AND ("Data Warehouse") AND ("company" OR "Group") AND ("information" OR "data") AND ("decision") AND ("reports" OR "picture" OR "summary")	8	X	X		

Estudios candidatos

La cadena de búsqueda ideal seleccionada es C2 ya que contiene la mayor cantidad de estudios relacionados con la investigación.

Cadena de búsqueda ideal

C2: ("data"OR"information") AND ("analysis" OR "investigation") AND ("Data warehouse") AND ("implementation") AND ("Business Intelligence")

Selección de estudios

Se selecciona los estudios que cumplan con los criterios de inclusión y se descartan los estudios que cumplen con los criterios de exclusión.

Tabla 3:

Selección de estudios

N.-	Título	Criterios de Inclusión	Criterios de Exclusión
1	Design and implementation of data warehouse with data model using survey-based services data	X	
2	Study on Model about Competitive Intelligence System of Enterprise Based on Data Mining under Electronic Commerce Environment		X
3	Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System	X	
4	Research on Business Intelligence Application in Electric Grid Industry Based on Data Mining		X
5	The design and implementation of business intelligence in insurance industry	X	
6	A content-based data masking technique for a built-in framework in Business Intelligence platform		X
7	Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process análisis	X	
8	Banking Intelligence: Application of data warehouse in bank operations	X	
9	Innovation in business intelligence systems: Spatial component for advanced threat and risk análisis		X
10	Predictive analytics with aviation big data		X
11	Business intelligence in higher education: Enhancing the teaching-learning process with a SRM system		X
12	The determinants of integrated business intelligence and analytics in organisational performance process		X
13	Business intelligence systems: design and implementation strategies	X	
14	Generic key performance indicators	X	
15	Development of the data warehouse model for public authorities accounts in Croatia	X	
16	Datawarehouse design for educational data mining		X
17	Agile way of BI implementation	X	
18	Business intelligence system for operational decision making support: A case study on lube distribution	X	
19	MapReduce-based warehouse systems: A Survey	X	
20	X-RIME: Cloud-Based Large Scale Social Network Analysis		X
21	Reacquiring business information from web communities		X

N.-	Título	Criterios de Inclusión	Criterios de Exclusión
22	The current challenges and status of risk management in enterprise data warehouse projects in South Africa	X	
23	Zero Latency applied on a commercial Data Mart: Real-time information in support of decision making	X	
24	Data value in decision process: Survey on decision support system in small and medium enterprises	X	
25	Implementing and analyzing a data mart for the Arlington County initiative to manage Domestic Violence offenders		X
26	A Holistic View of Data Warehousing in Education		X
27	Knowledge and Process Based Decision Support in Business Intelligence System	X	
28	CorporateMeasures: A clinical analytics framework leading to clinical Intelligence		X
29	DB2 for OS/390 V5 vs. V6 outer join performance		X
30	Querying Workflow Logs		X
31	Cost analysis in construction projects using fuzzy OLAP cubes		X
32	Differential snapshot algorithms based on Hadoop MapReduce	X	
33	Cloud Based Solutions for Big Data		X
34	Recent Trends in Substation Automation and Enterprise Data Management	X	
35	Future Cloudscape		X
36	Strong Before Smart		X

Filtros

Para la selección de estudios se tomaron en cuenta todos aquellos artículos que tengan relación con la implementación de un sistema de gestión de datos en empresas, no en educación ni salud, adicional los artículos que se presentaron en conferencias internacionales.

Estudios primarios

Tabla 4:

Selección de estudios primarios

N.-	Título
1	Design and implementation of data warehouse with data model using survey-based services data
2	Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System
3	Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process análisis
4	Development of the data warehouse model for public authorities accounts in Croatia
5	MapReduce-based warehouse systems: A Survey
6	The current challenges and status of risk management in enterprise data warehouse projects in South Africa
7	Zero Latency applied on a commercial Data Mart: Real-time information in support of decision making
8	Data value in decision process: Survey on decision support system in small and medium enterprises
9	Knowledge and Process Based Decision Support in Business Intelligence System
10	Differential snapshot algorithms based on Hadoop MapReduce

(Boon Keong Seah, 2014) Design and implementation of data warehouse with data model using survey-based services data

Los autores proponen el diseño del almacén de datos, desarrollando un modelo de datos multidimensional para la creación de múltiples mercados de datos y el diseño de un proceso ETL para poblar los mercados de datos de la fuente de datos. Asegurando que el desarrollo de múltiples mercados de datos permitirá una generación de informes más sencilla al identificar una dimensión común entre los mercados de datos. Las capacidades de unión cruzada de los mercados de datos a través de dimensiones comunes demuestran la capacidad de profundizar fácilmente en los mercados de datos para el análisis e informe de datos cruzados.

(Shenglei Pei, 2014) Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System

Los autores proponen el uso de SQL Server para implementar la base de datos y un almacén de datos en una empresa de acero, para construir modelos de minería de datos y el diseño del sistema de integración de información. El sistema que proponen cubre la compra de acero, inventario, venta, distribución, contratos, finanzas, análisis

estadístico, soporte de decisiones y otras funciones de gestión. El sistema propuesto puede cumplir con la gestión científica y ordenada de las pequeñas y medianas empresas de acero para los recursos de información, y las necesidades de los gerentes de previsión y toma de decisiones.

(Khurram Shahzad, 2011) Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process analysis

Los autores proponen un enfoque para la integración que abarca dos niveles, conceptual y de implementación; el primero define los conceptos que se utilizan para relacionar los objetivos y procesar el almacén; este último se utiliza para implementar los conceptos y las relaciones, para vincular los datos del almacén y los registros de objetivos en las bases de datos relacionales. El enfoque propuesto mejora la forma en que los usuarios acceden a los datos del almacén e interpretan los datos. Para facilitar el uso del enfoque propuesto, los autores implementaron un prototipo. Los resultados de la evaluación, aunque preliminares, proporcionan evidencia de la utilidad de nuestro enfoque.

(Sretenović, Kovačić, & Jovanović, 2016) Development of the data warehouse model for public authorities accounts in Croatia

Los autores describen el proceso de creación y organización del almacén de datos y el proceso ETL de extracción, limpieza y transformación de datos en un modelo de almacenamiento de datos multidimensional, que se basa en el modelo de metadatos. Además, describe un concepto y métodos básicos utilizados en el desarrollo de sistemas de inteligencia empresarial para la construcción y el uso de un almacén de datos. Además, establece opciones para el análisis de los datos en el almacén, que se muestran en función del concepto de almacenamiento dimensional, y proporciona pautas para el desarrollo futuro.

(Suresh Rao & P., 2014) MapReduce-based warehouse systems: A survey

Los autores proponen el uso de MapReduce que es un paradigma informático que no solo ha obtenido mucha respuesta hoy en día a partir de la investigación y la industria, sino que también se ha reconocido como una herramienta más eficaz para el análisis de datos a gran escala. El marco de MapReduce y su implementación de código abierto Hadoop proporcionan una estructura accesible y tolerante a fallos para el análisis de datos a gran escala y se ha construido con éxito en sitios web de redes sociales y proveedores de servicios web importantes. Este documento hace un resumen de la estructura de ubicación de datos basada en MapReduce para el análisis de Big Data, a saber, el almacén de filas, el almacén de columnas, el almacén híbrido, RCFile, Mastiff y ORC File.

(Legodi & Barry, 2010) The current challenges and status of risk management in enterprise data warehouse projects in South Africa

Los autores proponen como objetivo principal de este estudio investigar el estado actual de la gestión de riesgos en proyectos de almacenamiento de datos y determinar dónde se encuentran las principales áreas de riesgo en estos proyectos. Debido a la naturaleza exploratoria de este estudio, se utilizó un proceso Delphi. En términos de gestión de riesgos, el impacto financiero de los riesgos es lo más importante a considerar durante el análisis de riesgos, la falta de administración de las partes interesadas es la principal causa de riesgo, la propagación del alcance tiene el mayor impacto en el riesgo y la gestión de riesgos es muy importante durante la factibilidad, Fases de diseño e implementación.

(Sassi, Arrivabene, & Romero, 2011) Zero Latency applied on a commercial Data Mart: Real-time information in support of decision making

El objetivo de los autores en este documento es examinar cómo se puede obtener información en tiempo real a través de la implementación de un ZLE en el DM y cómo, de

manera efectiva, respaldaría el proceso de toma de decisiones estratégicas en las organizaciones. Para lograrlo, se ha presentado una propuesta para crear un modelo de DM en ZLE con datos obtenidos del departamento comercial de una empresa de suministro de ingeniería civil. Este análisis revela resultados significativos, respaldando la idea original de desarrollar apoyo para la toma de decisiones.

(Pighin & Marzona, 2012) Data value in decision process: Survey on decision support system in small and medium enterprises

Este documento los autores describen una encuesta sobre el uso real de las metodologías de Business Intelligence (BI) en las pequeñas y medianas empresas (PYMES) del distrito de Udine, un área de media industrialización en el noreste de Italia. Se pidió a las empresas que respondieran un cuestionario preparado por nuestro personal. También obtuvieron información más profunda a través de entrevistas específicas. La muestra estaba compuesta por 45 empresas diferentes por área geográfica, dimensión y especialización de producto. En el documento describen las razones, el objetivo de conocimiento y la metodología de la encuesta. Finalmente, analizaron los resultados sobre el uso efectivo y deseado de BI como Sistema de Apoyo a la Decisión (DSS).

(Ou & Peng, 2006) Knowledge and Process Based Decision Support in Business Intelligence System

Con el fin de proporcionar al sistema de inteligencia empresarial la capacidad de tomar decisiones basadas en procesos, los autores proponen el concepto de gestión de procesos empresariales en el sistema actual de inteligencia empresarial. Agregan el componente del modelo de proceso en la base de modelo de inteligencia empresarial. Con la implementación del razonamiento basado en casos y la tecnología de razonamiento basada en reglas, los modelos de proceso se pueden construir y

administrar de manera eficiente. En este documento también ofrecemos una estrategia para la gestión del conocimiento en sistemas de inteligencia empresarial.

(Du & Zou, 2015) Differential snapshot algorithms based on Hadoop MapReduce

En este documento los autores proponen la instantánea diferencial de bajo costo y alta eficiencia que combina la base de datos de código abierto y Hadoop MapReduce.

El resumen de datos basados en instantáneas diferenciales generado por el algoritmo MD5 es muy efectivo, pero el costo de E / S es muy alto. Así que el documento propone la declaración SQL que consulta la base de datos mientras genera el resumen de tuplas solo una vez que I / O. Implementamos la declaración SQL en la base de datos de código abierto MySQL. Además, la programación paralela de MapReduce se utiliza para encontrar la diferencia de los archivos de la base de datos, lo que mejora la eficiencia y evita el error. El experimento verifica los diferentes rendimientos entre algoritmos diferenciales de algoritmos de instantáneas.

Conclusiones o características del estado del arte

Tabla 5:

Resumen de estudios primarios del estado del arte

Título del Estudio	Herramientas Utilizadas	Metodologías	Detalle
Design and implementation of data warehouse with data model using survey-based services data	Pentaho Data Integrator herramienta de integración que usa Spoon GUI	Metodología de Kimball, llamada Modelo Dimensional, Metodología de Bill Inmon Top-Down	El almacén de datos de una empresa consolida datos de fuentes heterogéneas para apoyar la toma de decisiones.
Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System	Uso de SQL Server	Arquitectura de tres niveles de sistemas de integración de información. - Implementar un data warehouse - Construcción de modelos de minería de datos	La implementación la base de datos y un data warehouse en una empresa de acero para el diseño del sistema de integración de información.

Título del Estudio	Herramientas Utilizadas	Metodologías	Detalle
Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process análisis	No Aplica	- Diseño del sistema de integración de información. Uso de data warehouse orientado a procesos con la definición de metamodelos, para la obtención de un modelo multidimensional temporal aplicando el enfoque de arquitectura basada en modelos.	La integración abarca dos niveles, conceptual y de implementación: el primero define los conceptos empleados en la integración, y la implementación describe las extensiones a las tablas de dimensiones y hechos de un process warehouse para representar los registros relacionados con las metas.
Development of the data warehouse model for public authorities in Croatia	No Aplica	Metodología de Kimball	En el enfoque dimensional de Kimball, los datos se dividen en tablas de hechos, las cuales se conectan a una o más tablas de dimensiones, que permiten la vista multidimensional al usuario, lo que se llama hipercubo.
MapReduce-based warehouse systems: A survey	MapReduce	No Aplica	El análisis de Big Data está creciendo rápidamente en la industria de la inteligencia empresarial o que hace que la solución de almacenamiento tradicional sea prohibitivamente cara.
The current challenges and status of risk management in enterprise data warehouse projects in South Africa	No Aplica	Método Delphi	La gestión de riesgos en proyectos de almacenamiento de datos que representan más del 10% de los presupuestos de tecnología de la información de la mayoría de las organizaciones y estos proyectos tienen una tasa de falla estimada del 50%.
Zero Latency applied on a commercial Data Mart: Real-time information	No Aplica	Zero Latency Enterprise (ZLE)	Obtener información en tiempo real a través de la implementación de una ZLE en los data mart de la empresa y cómo,

Título del Estudio	Herramientas Utilizadas	Metodologías	Detalle
in support of decision making			efectivamente, apoyaría el proceso de toma de decisiones estratégicas en las organizaciones
Data value in decision process: Survey on decision support system in small and medium enterprises	No Aplica	Encuesta	La encuesta realizada a 45 empresas sobre el uso real de las metodologías de BI en las pequeñas y medianas empresas, con un resultado del 24% de empresas utiliza sistemas Data Warehouse para el análisis de datos, entre las empresas que aún no cuentan con esta herramienta, el 26% adoptará una en el futuro y el 11% en el corto plazo.
Knowledge and Process Based Decision Support in Business Intelligence System	No Aplica	Gestión de Procesos Empresariales	Sistema de apoyo a la toma de decisiones basado en procesos para ayudar a las empresas a mejorar la velocidad y la eficacia de las operaciones comerciales.
Differential snapshot algorithms based on Hadoop MapReduce	Hadoop MapReduce / MySQL	Combinación de base de datos de código abierto y Hadoop MapReduce.	La captura de datos modificados desde el sistema de origen es el primer paso en el mantenimiento incremental de los almacenes de datos y la inteligencia empresarial, es un componente clave de la técnica ETL.

- El estudio de los autores (Boon Keong Seah, 2014) Design and implementation of data warehouse with data model using survey-based services data, permitió identificar dos metodologías para la implementación de un sistema de gestión de datos, la de Ralph de Kimball y la de Bill Inmon, el artículo también permitió identificar la importancia de la implementación de un modelo de datos multidimensional mediante la creación de múltiples data marts, para el apoyo en

la toma de decisiones en las empresas, ya que permite consolidar los datos de fuentes heterogéneas en el almacén de datos de la empresa.

- Varios de los estudios encontrados (Boon Keong Seah, 2014; Shenglei Pei, 2014; Khurram Shahzad, 2011; Sretenović, Kovačić, & Jovanović, 2016) coinciden en que la mejor solución para la centralización de la información dentro de una empresa es la implementación de un modelo de datos multidimensional y procesos ETL basado en la creación de múltiples data marts en base a las necesidades de cada empresa.

Capítulo IV

Construcción de la solución

El sistema de gestión de datos para la toma de decisiones se desarrollará utilizando la metodología de Ralph Kimball denominada Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle), que es una guía para la construcción de un data warehouse dentro de la organización, su filosofía se centra en que, en la mayoría de las organizaciones, la construcción de un data warehouse se origina por el interés y esfuerzo de un departamento. Es por este motivo que en una primera versión este data warehouse no es más que un data mart departamental. A medida que otros departamentos necesiten sus propios data marts, éstos se irán combinando con el primero manteniendo una metodología de estandarización mediante lo que Kimball denomina “dimensiones conformadas”, que serán las dimensiones comunes entre los diferentes departamentos.

Este ciclo de vida del proyecto de data warehouse, está basado en cuatro principios básicos (Marcelo, 2019):

- **Centrarse en el negocio:** Hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- **Diseño de infraestructura:** Diseñar una base de datos única, centralizada y de alto procesamiento donde se identifica los requerimientos de negocio más relevantes de la empresa.
- **Realizar entregas en incrementos significativos:** crear el almacén de datos (DW) en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación

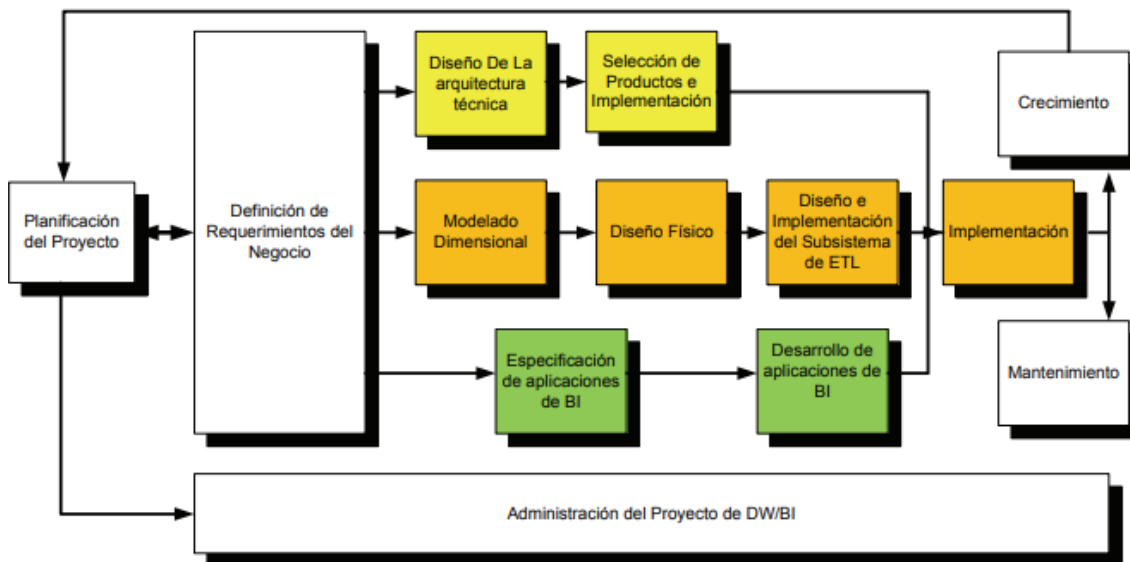
de los incrementos. En esto la metodología se parece a las metodologías ágiles de construcción de software.

- **Ofrecer una solución:** consiste en suministrar las herramientas necesarias para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

La construcción de una solución de DW/BI (Datawarehouse/Business Intelligence) es sumamente compleja, y Kimball nos propone una metodología que nos ayuda a simplificar esa complejidad. Las tareas de esta metodología (ciclo de vida) se muestran en la figura 4.

Figura 4:

Tareas de la metodología de Kimball (Rivadera, 2010)



Planificación del proyecto

Definición del proyecto

El desarrollo de este sistema de Business Intelligence nace de la necesidad de tener información centralizada, actualizada y resumida para el apoyo en la toma de decisiones en el departamento comercial de la empresa NUO Tecnológica, y así evitar la dependencia de otras áreas para tener acceso a la información de interés de manera inmediata.

Justificación y objetivos

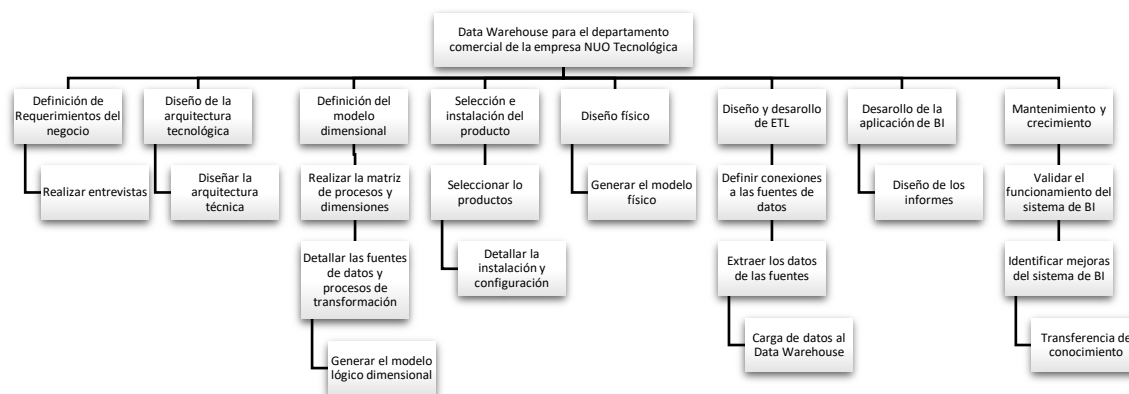
Este proyecto tiene como objetivo centralizar la información de la organización mediante la implementación de un data warehouse, permitiendo al área comercial de la empresa NUO Tecnológica tener conocimiento del estado actual del negocio para la toma de decisiones, estandarizando y transformando los datos en un sistema de código único, permitiendo que cada uno de los departamentos que componen la empresa, puedan utilizar el mismo repositorio de datos para sus consultas y generación de informes individuales.

Alcance

El desarrollo de este proyecto se delimita con el diseño e implementación de un data warehouse para el área comercial de la empresa NUO Tecnológica, el alcance se detalla a continuación en el esquema de desglose del trabajo, detallando los trabajos y entregables del proyecto.

Figura 5:

Esquema de desglose del trabajo



Identificación de tareas

Tabla 6:

Identificación de tareas

Tarea	Descripción
Realizar entrevistas	Definir las preguntas de negocio que deben ser resueltas con el proyecto y los requerimientos para los reportes
Diseñar la arquitectura técnica	Definir los campos necesarios de las fuentes de datos para el modelo de datos.
Realizar la matriz de procesos y dimensiones	Definir las tablas de dimensiones que se relacionaran directamente con la tabla de hechos según el proceso de negocio.
Detallar las fuentes de datos y procesos de transformación	Definir las fuentes de información y las transformaciones necesarias que permitirán extraer los datos necesarios para alimentar al nuevo sistema y cumplir con los requerimientos del negocio identificados.
Generar el modelo lógico dimensional	Definir y diseñar el modelo lógico que será utilizado.
Seleccionar los productos	Definir las distintas herramientas que se van a utilizar para implementar el proyecto.
Detallar la instalación y configuración	Realizar la instalación y configuración del software requerido para el sistema Data Warehouse.
Generar el modelo físico	Generar el modelo físico a implementar en la base de datos a partir del modelo lógico.
Definir conexiones a las fuentes de datos	Configurar las conexiones a las diferentes fuentes de datos para alimentar al nuevo sistema.
Extraer los datos de las fuentes	Construir y hacer pruebas unitarias de los procesos de extracción y transformación.
Carga de datos al Data Warehouse	Construir y hacer pruebas unitarias de los procesos de carga requeridos para poblar la bodega de datos.
Diseño de los informes	Desarrollo de los nuevos reportes para visualizar los datos.
Validar el funcionamiento del sistema de BI	Llevar a cabo el proceso de pruebas de todos los componentes del sistema para realizar la validación del ambiente integral de producción.
Identificar mejoras del sistema de BI	Identificar mejoras a realizar en el sistema en base a las recomendaciones de los usuarios.

Tarea	Descripción
Transferencia de conocimiento	Realizar la transferencia de conocimiento de la solución mediante la capacitación técnica y funcional.

Secuencia de tareas

Tabla 7:

Secuencia de tareas

Data Warehouse para el departamento comercial de la empresa NUO Tecnológica			
Actividad	Predecesora	Actividad	Predecesora
Definición de Requerimientos del negocio	-	Realizar entrevistas	-
Diseño de la arquitectura tecnológica	Definición de Requerimientos del negocio	Diseñar la arquitectura técnica	Realizar entrevistas
Definición del modelo dimensional	Diseño de la arquitectura tecnológica	Realizar la matriz de procesos y dimensiones	Diseñar la arquitectura técnica
Selección e instalación del producto	Definición del modelo dimensional	Seleccionar los productos	Generar el modelo lógico dimensional
Diseño físico	Selección e instalación del producto	Detallar la instalación y configuración	Seleccionar los productos
Diseño y desarrollo de ETL	Diseño físico	Generar el modelo físico	Detallar la instalación y configuración
Desarrollo de la aplicación de BI	Diseño y desarrollo de ETL	Definir conexiones a las fuentes de datos	Generar el modelo físico
Mantenimiento y crecimiento	Desarrollo de la aplicación de BI	Extraer los datos de las fuentes	Definir conexiones a las fuentes de datos
		Carga de datos al Data Warehouse	Extraer los datos de las fuentes
		Diseño de los informes	Carga de datos al Data Warehouse
		Validar el funcionamiento del sistema de BI	Diseño de los informes

Estimación de recursos para el proyecto

Tabla 8:

Estimación de recursos para el proyecto

Tarea	Descripción	Recurso (Cantidad)
Realizar entrevistas	Definir las preguntas de negocio que deben ser resueltas con el proyecto y los requerimientos para los reportes	Gerente del Proyecto (1) Consultor BI (1)
Diseñar la arquitectura técnica	Definir los campos necesarios de las fuentes de datos para el modelo de datos.	Gerente del Proyecto (1) Consultor BI (1)
Realizar la matriz de procesos y dimensiones	Definir las tablas de dimensiones que se relacionarán con la tabla de hechos.	Gerente del Proyecto (1) Consultor BI (1)
Detallar las fuentes de datos y procesos de transformación	Definir las fuentes de información y las transformaciones necesarias que permitirán extraer los datos necesarios para alimentar al nuevo sistema y cumplir con los requerimientos del negocio identificados.	Gerente del Proyecto (1) Consultor BI (1)
Generar el modelo lógico dimensional	Definir y diseñar el modelo lógico que será utilizado.	Gerente del Proyecto (1) Consultor BI (1)
Seleccionar los productos	Definir las distintas herramientas a utilizar para desarrollar el proyecto..	Gerente del Proyecto (1) Consultor BI (1)
Detallar la instalación y configuración	Realizar la instalación y configuración del software requerido para el sistema Data Warehouse.	Gerente del Proyecto (1) Consultor BI (1)
Generar el modelo físico	Generar el modelo físico a implementar en la base de datos a partir del modelo lógico.	Gerente del Proyecto (1) Consultor BI (1)
Definir conexiones a las fuentes de datos	Configurar las conexiones a las diferentes fuentes de datos para alimentar al nuevo sistema.	Gerente del Proyecto (1) Consultor BI (1)
Extraer los datos de las fuentes	Construir y hacer pruebas unitarias de los procesos de extracción y transformación.	Gerente del Proyecto (1) Consultor ETL (1)
Carga de datos al Data Warehouse	Construir y hacer pruebas unitarias de los procesos de carga requeridos para poblar la bodega de datos.	Gerente del Proyecto (1) Consultor ETL (1)
Diseño de los informes	Desarrollo de los nuevos reportes para visualizar los datos.	Gerente del Proyecto (1) Consultor BI (1)
Validar el funcionamiento del sistema de BI	Llevar a cabo el proceso de pruebas de todos los componentes del sistema para realizar la validación del ambiente integral de producción.	Gerente del Proyecto (1) Consultor BI (1)
Identificar mejoras del sistema de BI	Identificar mejoras a realizar en el sistema en base a las recomendaciones de los usuarios.	Gerente del Proyecto (1) Consultor BI (1)
Transferencia de conocimiento	Realizar la transferencia de conocimiento de la solución mediante la capacitación técnica y funcional.	Gerente del Proyecto (1) Consultor BI (1)

Cronograma

Para el desarrollo del proyecto de business intelligence es necesario realizar las tareas identificadas en los puntos anteriores en base a la metodología de kimball, actividades que tienen un tiempo estimado para ser culminadas en base al siguiente cronograma:

Análisis de requerimientos

Para el levantamiento de requerimientos se realizó una entrevista a los gerentes regionales del área comercial, mediante esta entrevista se identificaron varios temas analíticos y procesos de negocio, los mismos que permiten agrupar los siguientes requerimientos comunes:

Tabla 9:

Temas analíticos y procesos de negocio

Tema Analítico	Análisis o requerimiento inferido o pedido	Proceso de negocio de soporte	Comentarios
Gestión de Ventas	Obtener total de ventas	Ventas/ Maestro de Productos	Por cliente Por año Por SKU de producto Por marca de producto
	Obtener ventas por SKU del producto	Ventas/ Maestro de Productos	Por cliente Por año Por SKU de producto
Metas	Márgenes rentabilidad de	Ventas / Costos	Por cliente Por año Por SKU de producto
	Comparar ventas con meta anual	Ventas / Proyección anual	Por cliente Por año
Proyecciones	Proyección de ventas	Ventas/ Maestro de Productos	Por cliente Por año Por SKU de producto Por marca de producto

A partir del análisis de los requerimientos agrupados por temas analíticos y procesos de negocio, se construye la matriz de procesos / dimensiones (Bus Matrix):

Tabla 10:

Matriz de procesos / dimensiones

Proceso de negocio	Dimensiones			
	Tiempo	Cliente	Producto	Factura
Gestión de Ventas	X	X	X	X
Metas	X	X	X	X
Proyecciones	X	X	X	X

Línea tecnológica

A continuación, se muestra la arquitectura técnica del proyecto y las respectivas herramientas a ser utilizadas para el desarrollo del mismo, de acuerdo con las necesidades del negocio y tomando en cuenta las herramientas con las que cuenta la empresa NUO Tecnológica.

Arquitectura tecnológica

La arquitectura fue definida luego del análisis de la infraestructura de la empresa NUO Tecnológica, para este proyecto fue asignado un servidor HPE ProLiant DL20 Gen10 Server donde se realizó la configuración del ambiente necesario para el funcionamiento del data warehouse.

Tabla 11:

Temas analíticos y procesos de negocio

Marca	HP
Series	ProLiant DL20 Gen10 Server
Factor de forma	Rack (1U)
Procesador	Intel Xeon E-2124 (4-Core, 3.3 GHz, 71W)
Disco duro	2 x HPE 1TB SATA 7.2K LFF LP DS HDD
Sistema operativo	Microsoft Windows Server 2012
Fuente de alimentación	Voltaje de entrada de 100 a 240 Vca. 50/60Hz

Debido a que los archivos de las fuentes de datos se encuentran en diferentes formatos, para realizar la integración de dichos datos se propone utilizar la herramienta pentaho data integration.

Mediante los procesos de extracción, transformación y carga (ETL) todos los datos se integran en el ODS (Base de datos operacional), para luego ser cargados en el data warehouse alojado en el servidor asignado por la empresa.

Para el diseño de los reportes que son consumidos por el usuario se propone utilizar el software licenciado Power Bi, software que consume los datos desde el data warehouse.

Selección e instalación de productos

La selección de productos se llevó a cabo considerando las herramientas que actualmente posee la empresa, lo que contribuyó a disminuir los costos de este proyecto.

Tabla 12:

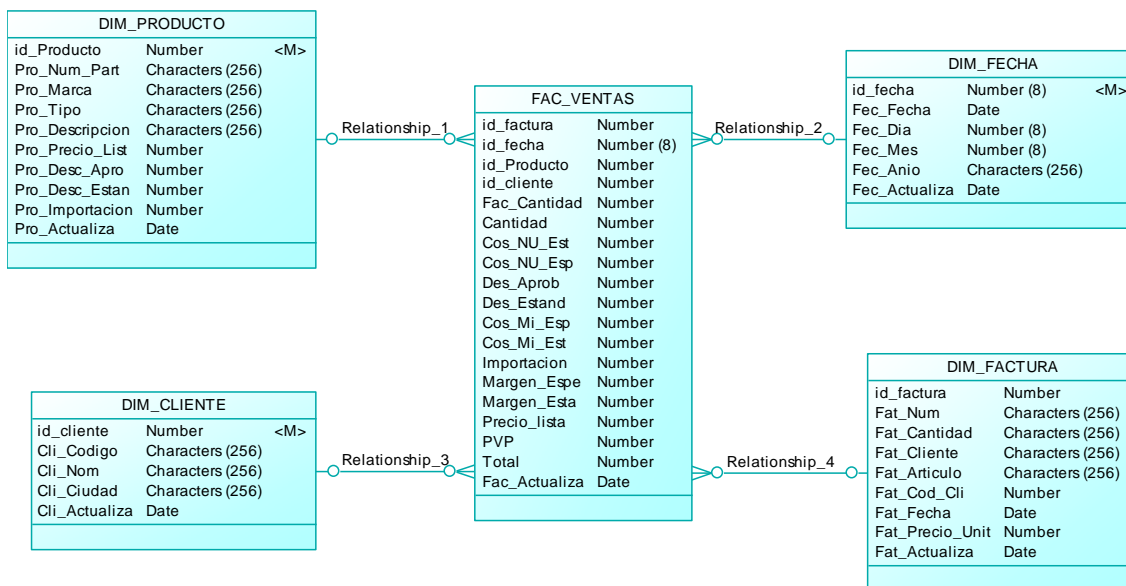
Productos seleccionados

Producto	Descripción	Software Licenciado	Observación
MySQL Server	Es un sistema de gestión de bases de datos relacional.	Si	La empresa cuenta con el licenciamiento
Pentaho Data Integration	Herramienta de la suite de Pentaho de extracción, transformación y carga.	Si	La empresa cuenta con el licenciamiento
Power BI	Herramienta que permite modelizar y analizar datos, presentarlos a través de paneles e informes; que puedan ser consultarlos de una manera muy fácil, atractiva e intuitiva por los usuarios.	Si	La empresa cuenta con el licenciamiento

Modelado dimensional

En base al análisis de requerimientos y a la matriz de procesos / dimensiones realizado en puntos anteriores, se identificaron tres procesos de negocio necesarios para crear el data mart del departamento comercial de la empresa NUO Tecnológica.

En la Figura 7 se observa el modelo dimensional propuesto como solución para la implementación de un data mart para el área comercial, conformado por una tabla de hechos (fac_ventas) la misma que contiene la clave primaria que se relaciona con las dimensiones dim_cliente, dim_producto, dim_factura y dim_fecha.

Figura 7:*Modelo dimensional*

De acuerdo a los requerimientos identificados en la fase anterior se construyó el esquema de análisis dimensional del proceso de negocio.

Tabla 13:*Análisis dimensional de un proceso de ventas*

PROCESO DE NEGOCIO: Ventas				
Tabla de hechos	Dimensiones			
	dim_fecha	dim_producto	dim_factura	dim_cliente
fac_ventas	X	X	X	X

Nivel de granularidad de dimensiones

Representa el nivel de detalle al que se desea guardar los datos del negocio que se desea analizar. Por ejemplo, los datos referentes a ventas, se pueden analizar a nivel de año, mes, día. Mientras mayor sea el nivel de detalle de los datos, se tendrán mayores posibilidades de análisis, ya que los mismos podrán ser resumidos” (Avella, 2014).

Tabla 14:*Nivel de granularidad de dimensiones*

Dimensión	Nivel de granularidad
dim_fecha	Id_fecha, año, mes, día
dim_producto	Id_producto, marca, tipo, producto
dim_factura	Id_factura, factura
dim_cliente	Id_cliente, cliente

Dimensiones

A continuación, se detallan las dimensiones con sus atributos correspondientes que conforman el modelo dimensional para la creación del data mart ventas.

Tabla 15:

DIM_FECHA

Nombre de la Dimensión: Ventas		
Nombre de la Tabla: DIM_FECHA		
Tipo de Tabla: Tabla de dimensiones		
Descripción Campo	Campo	Tipo de dato
Código fecha auto incremental	id_fecha	Number
Fecha formato día/mes/año	Fec_fecha	Date
Día de la fecha	Fec_Dia	Number
Mes de la fecha	Fec_Mes	Number
Año de la fecha	Fec_Anio	Varchar2(4)
Actualización de registro de la dimensión	Fec_Actualiza	Date

Tabla 16:

DIM_PRODUCTO

Nombre de la Dimensión: Ventas		
Nombre de la Tabla: DIM_PRODUCTO		
Tipo de Tabla: Tabla de dimensiones		
Descripción Campo	Campo	Tipo de dato
Código producto auto incremental	id_producto	Number
Código del producto	Pro_Num_Part	Varchar2(10)
Marca del producto	Pro_Marca	Varchar2(10)
Tipo de producto	Pro_Tipo	Varchar2(10)
Descripción producto	Pro_Descripcion	Varchar2(50)
Precio de lista del producto	Pro_Precio_List	Number
Descuento especial aprobado del producto	Pro_Desc_Apro	Number
Descuento estándar producto	Pro_Desc_Estan	Number
Porcentaje de importación producto	Pro_Importacion	Number
Actualización de registro de la dimensión	Pro_Actualiza	Date

Tabla 17:*DIM_CLIENTE*

Nombre de la Dimensión: Ventas		
Nombre de la Tabla: DIM_CLIENTE		
Tipo de Tabla: Tabla de dimensiones		
Descripción Campo	Campo	Tipo de dato
Código cliente auto incremental	id_cliente	Number
Código del cliente	Cli_Codigo	Varchar2(10)
Nombre del cliente	Cli_Nom	Varchar2(10)
Ciudad del cliente	Cli_Cuidad	Varchar2(10)
Actualización de registro de la dimensión	Cli_Actualiza	Date

Tabla 18:*DIM_FACTURA*

Nombre de la Dimensión: Ventas		
Nombre de la Tabla: DIM_FACTURA		
Tipo de Tabla: Tabla de dimensiones		
Descripción Campo	Campo	Tipo de dato
Código factura auto incremental	Id_factura	Number
Número de factura	Fat_Num	Varchar2(20)
Cantidad de artículos de la factura	Fat_Cantidad	Varchar2(10)
Cliente de la factura	Fat_Cliente	Varchar2(10)
Artículo de la factura	Fat_Articulo	Varchar2(50)
Código del cliente de la factura	Fat_Cod_Cli	Number
Fecha de la factura	Fat_Fecha	Date
Precio unitario de artículo de la factura	Fat_Precio_Unit	Number
Actualización de registro de la dimensión	Fat_Actualiza	Date

Tabla de hechos

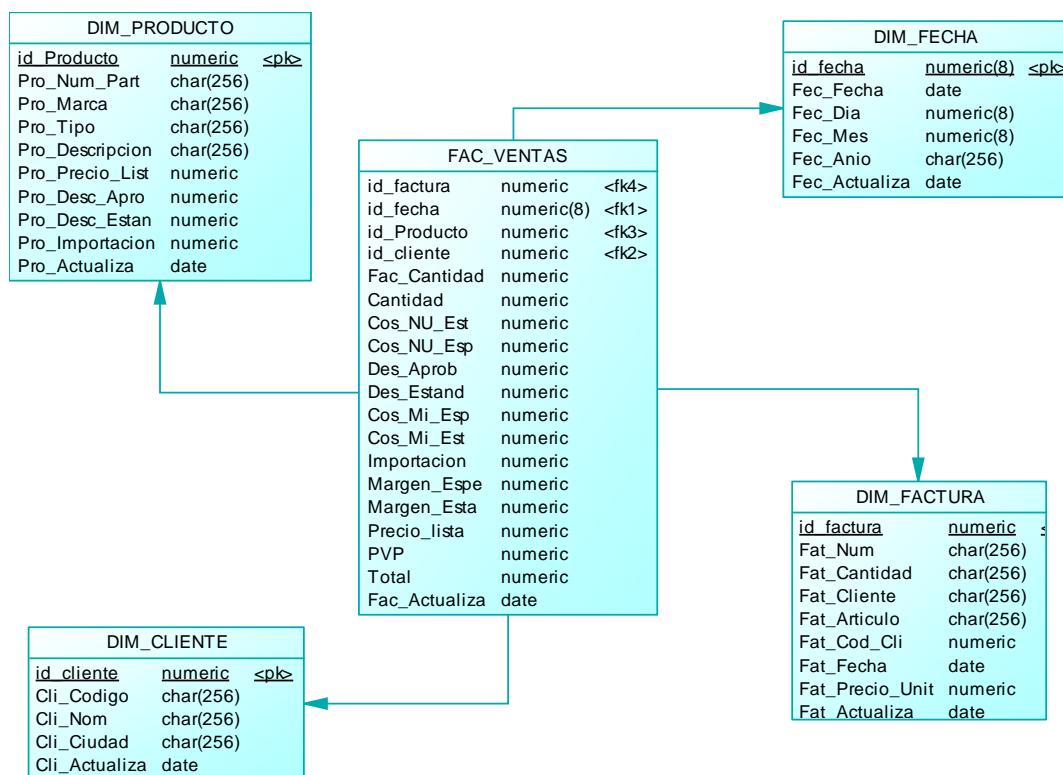
En la Tabla 18 se detalla la tabla de hechos Fac_Ventas con sus respectivos atributos en la cual se representa el proceso de negocio.

Tabla 19:*FAC_VENTAS*

Nombre de la Tabla de Hechos: Ventas		
Nombre de la Tabla: FAC_VENTAS		
Tipo de Tabla: Tabla de hechos		
Descripción Campo	Campo	Tipo de dato
Llave primaria de factura auto incremental	Id_factura	Number
Llave primaria de fecha auto incremental	id_fecha	Number

Nombre de la Tabla de Hechos: Ventas**Nombre de la Tabla: FAC_VENTAS****Tipo de Tabla: Tabla de hechos**

Llave primaria de producto auto incremental	id_producto	Number
Llave primaria de cliente auto incremental	id_cliente	Number
Descripción Campo	Campo	Tipo de dato
Cantidad de artículos de la factura	Cantidad	Number
Costo artículo EC	Cos_NU_Est	Number
Costo especial artículo EC	Cos_NU_Esp	Number
Descuento especial	Des_Aprob	Number
Descuento estándar	Des_Estand	Number
Costo Miami especial	Cos_Mi_Esp	Number
Costo Miami estándar	Cos_Mi_Est	Number
Porcentaje de importación	Importacion	Number
Margen con precio especial	Margen_Espe	Number
Margen con precio estándar	Margen_Esta	Number
Precio de lista del artículo	Precio_lista	Number
Precio de venta del artículo	PVP	Number
Total de la facturación	Total	Number

Modelo físico**Figura 8:***Modelo físico*

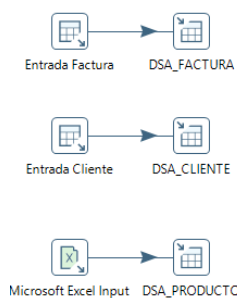
ETL diseño y desarrollo

Mediante 2 la herramienta pentaho data integrator se desarrolla la primera fase del proceso que corresponde al proceso ETL (Extracción, Transformación y Carga) de los datos provenientes de las distintas fuentes de datos de la empresa NUO Tecnológica de los años 2014, 2015, 2016, 2017, 2018, 2019 y 2020 que se encuentran en formato .mdf y .xls a la base de datos operacional.

Los orígenes de datos se estandarizaron de la siguiente forma Entrada Factura, para los procesos de destino en la base de datos operacional se denomina DSA_Factura, como se muestra en la Figura 11, este estándar fue aplicado a todas las tablas.

Figura 9:

Entrada de datos vs destino DSA

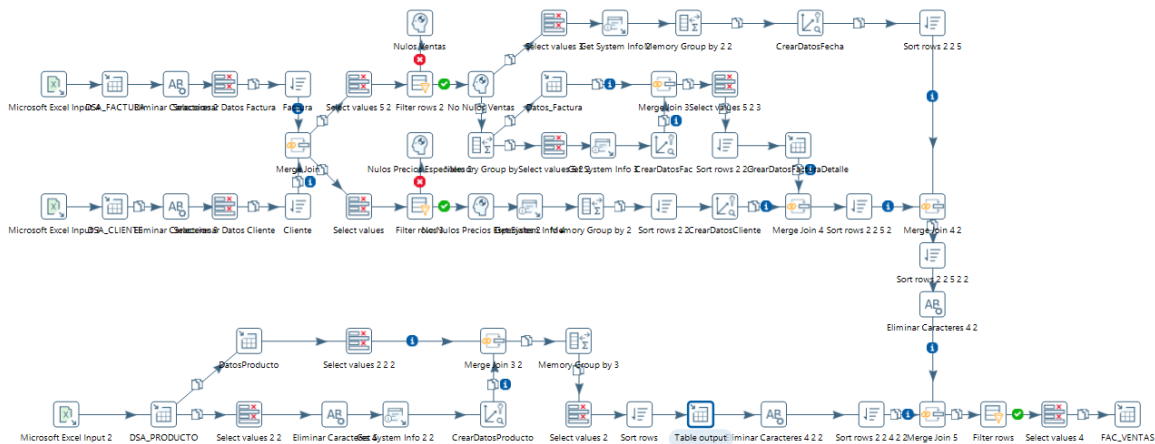


En la segunda fase se realiza la de depuración de los datos en el ODS, para continuar con el proceso de extracción y carga al data warehouse.

La plataforma que se utilizó para la implementación de los ETLs es Pentaho Data Integrator Versión 8.1. A continuación se detallan los diseños implementados para la creación del data warehouse.

Figura 10:

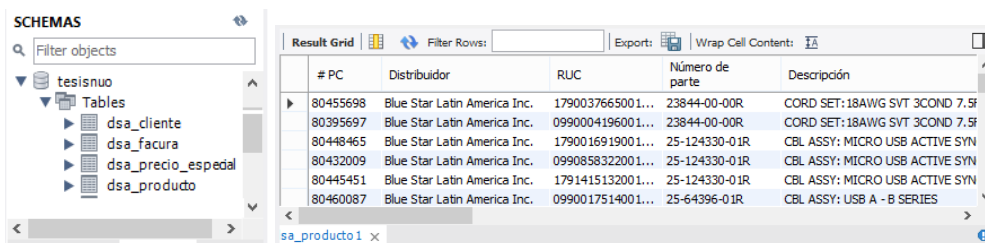
Diseño e implementación del data warehouse



Para la extracción de los datos de las fuentes de la empresa, se creó un repositorio de datos para cada una de las fuentes llamadas: dsa_factura, dsa_cliente y dsa_producto.

Figura 11:

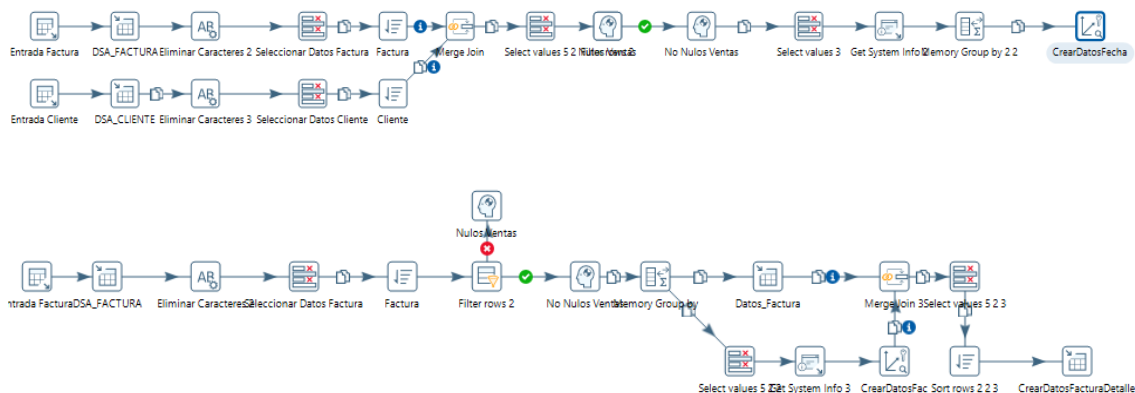
Repositorio de datos DSA

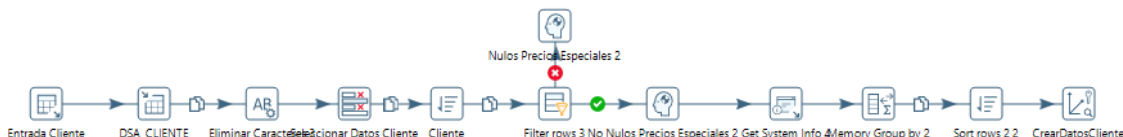


A continuación, se detalla el proceso ETL (Extracción, transformación y carga) de cada una de las dimensiones del Data Ware Hosue.

Figura 12:

Proceso de etl dimensiones del data warehouse





Pasos realizados para la transformación de datos:

- Se descartaron campos que no son necesarios en cada una de las dimensiones con la función “Select values”, de igual manera se realizó la validación que no existan datos nulos obligatorios dentro de las dimensiones con la función “Filter Rows”, esto permite tener un control y corrección a tiempo de dichos campos y así lograr tener información de calidad dentro del data warehouse.
- Se revisó que no exista duplicidad en los datos y que tenga congruencia entre las distintas fuentes de datos para evitar pérdidas de información.
- Mediante la herramienta pentaho se generaron nuevos índices autoincrementales para cada una de las dimensiones permitiendo mantener los índices originales de las fuentes de datos, para tener el registro relacionado perfectamente desde el origen hasta el destino.
- Se creó una estructura de base de datos dimensional adaptada a las necesidades de este proyecto.
- Por último, se realizó la carga de los datos al data warehouse.

Figura 13

Validación carga de datos a las dimensiones y a la tabla de hechos

id_factura	numfac	fecfac	coddi	codart	preuni
17505	001001-00002366	2017-01-13 00:00:00	0190072002001	CRD3000-1001RR	292.61
18199	002006-00000545	2020-02-04 00:00:00	0190072002001	\007	45
18171	002006-00000515	2020-01-20 00:00:00	0190072002001	\007	62
17066	002006-00000577	2020-02-20 00:00:00	0190110001001	ZD420-42C01M00EZ	415.65
18107	001001-00002062	2016-05-25 00:00:00	0190115798001	\006	50
16036	002006-00000570	2020-02-18 00:00:00	0190155722001	\007	1033.94
16036	002006-00000570	2020-02-18 00:00:00	0190155722001	\007	14161.91

id_factura	id_Cliente	id_fecha	id_Producto	Precio de Lista	Descuento Estandar	Descuento Aprobado	Importación %
16207	1463	4703	600	55	0.45	0.4287	0.27
15475	1463	4571	600	55	0.45	0.4287	0.27
16548	1463	4674	600	55	0.45	0.4287	0.27
16196	1463	4383	600	55	0.45	0.4287	0.27
15295	1463	4383	600	55	0.45	0.4287	0.27
15844	1463	4658	600	55	0.45	0.4287	0.27

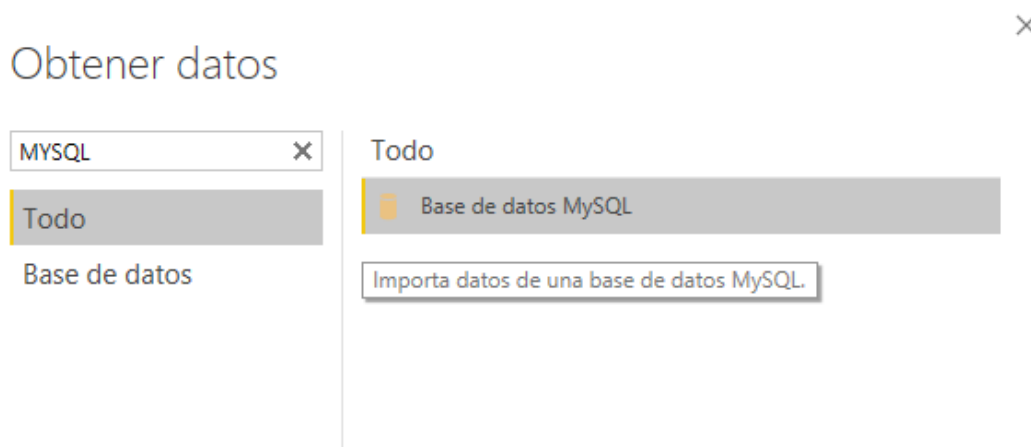
Desarrollo de la aplicación

Para la implementación de la aplicación de business intelligence primero se realizó la conexión de la herramienta Power BI con la base de datos, lo que permitió importar las tablas de dimensiones y la tabla de hecho a la herramienta Power BI.

- Seleccionamos la base de datos de origen donde se encuentra el data warehouse.

Figura 14

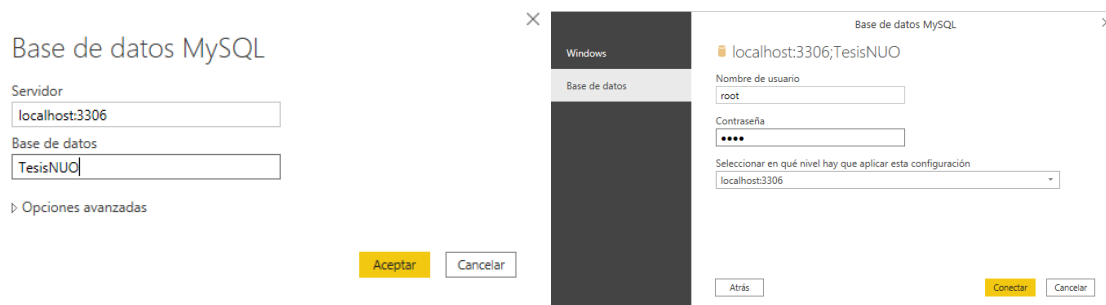
Selección de la base de datos de origen



- Se ingresa los datos del servidor, base de datos, nombre del usuario y contraseña para establecer la conexión con la base de datos de origen.

Figura 15

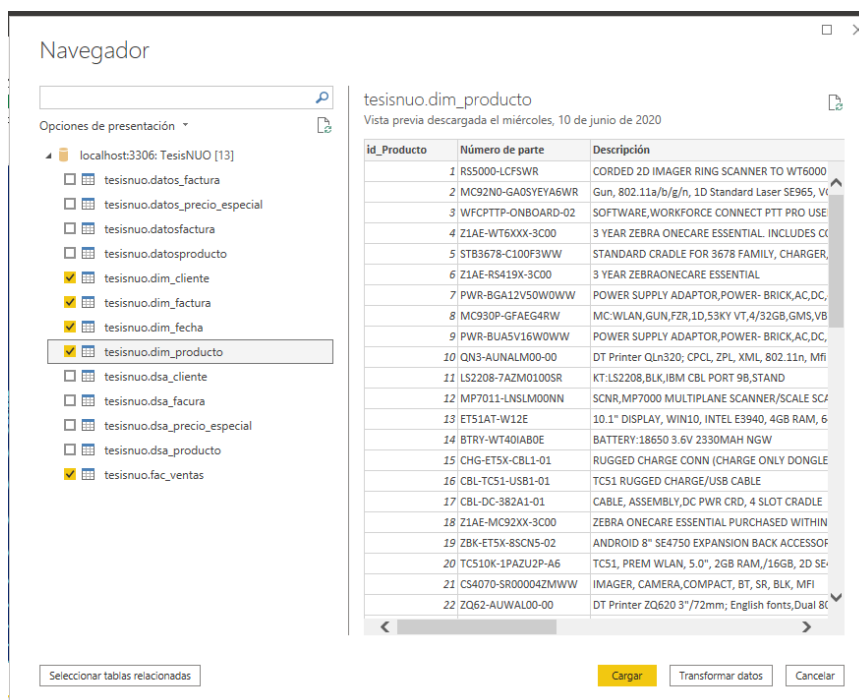
Configuración de la conexión con la base de datos



- Se selecciona las tablas de dimensiones: dim_producto, dim_fecha, dim_factura y dim_cliente, de igual forma la tabla de hechos fac_ventas.

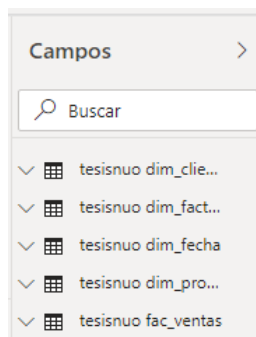
Figura 16

Selección de tablas de dimensiones y hechos



Modelo de diseño de BI

Una vez establecida la conexión con la base de datos se puede observar las tablas dimensiones y tabla de hecho en el panel de campos de Power BI para empezar a diseñar los reportes.

Figura 17*Panel de campos*

En el panel de visualizaciones se observa las diferentes opciones para realizar los reportes que brinda la herramienta Power Bi como matrices, pasteles, mapas, histogramas entre otros, para empezar con el diseño se selecciona el diseño requerido y se lo arrastra al área de trabajo.

Figura 18*Panel de visualizaciones*

Para finalizar se prepara el área de trabajo donde se encuentran todos los objetos del reporte, mediante el panel de configuración podemos cambiar el color de cada uno de los objetos, tamaño de fuentes.

Figura 19*Diseño del reporte de ventas del área comercial*



Despliegue

El producto final de este trabajo de titulación es el sistema de business intelligence (BI) que mediante la generación de reportes brinda soporte a la toma de decisiones al área comercial de la empresa NUO Tecnológica de acuerdo a sus requerimientos.

A continuación, se muestran algunos indicadores con sus respectivos informes:

- ¿Cuál es el total de ventas por año y por cliente de la empresa?

Figura 20

Total de ventas por año y por cliente

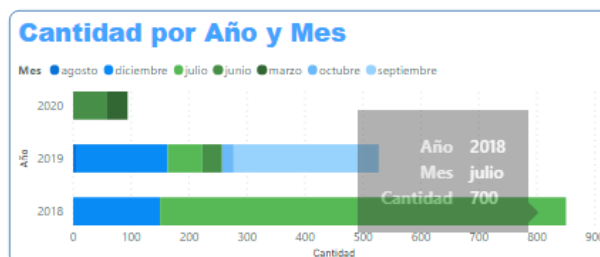


En la Figura 20 se observa el total de ventas por año vs el total de ventas por cliente desde el año 2014 hasta el 2020, lo que permite verificar el crecimiento anual en ventas de la empresa NUO tecnológica, llegando a su pico máximo en el año 2019 con un total de \$7,52 millones y seguido por el año 2018 en el que se alcanzó un total de \$7,39 millones.

- ¿Cuál es la cantidad de equipos de determinado modelo vendidos por año de determinado cliente?

Figura 21

Cantidad de equipos vendidos por año y por mes



En la Figura 21 se observa la cantidad de equipos vendidos en determinado modelo por año y por mes, lo que permite identificar los meses con mayor demanda de dicho modelo de equipo para el manejo de stock.

- ¿Cuál es la proyección de venta por cantidad de equipos de determinado modelo por cliente?

Figura 22

Proyección de venta por cantidad de equipos trimestral por cliente



En la Figura 22 se observa una proyección trimestral de venta por cantidad de equipos de determinado modelo, lo que permite al departamento comercial planificar las compras de dicho modelo de equipos en base al comportamiento de las ventas históricas.

- ¿Cuál es el último precio histórico de venta por equipo y por cliente?

Figura 23

Cantidad de equipos vendidos por año y por mes

Cliente	Marca	Número de Parte	PVP Unitario	Año	Mes	Día
CAMACHO MORAN NELSON RODOLFO	ZEBRA	CBA-U51-S16ZAR	\$ 26,96	2019	febrero	4
CAMACHO MORAN NELSON RODOLFO	ZEBRA	CBA-U51-S16ZAR	\$ 26,96	2019	marzo	8
CAMACHO MORAN NELSON RODOLFO	ZEBRA	CBA-U51-S16ZAR	\$ 26,96	2019	octubre	8
CAMACHO MORAN NELSON RODOLFO	ZEBRA	CBA-U51-S16ZAR	\$ 26,96	2020	marzo	2
CAMACHO MORAN NELSON RODOLFO	ZEBRA	CBA-U51-S16ZAR	\$ 26,96	2020	marzo	5

En la Figura 23 se observa el último precio de venta de determinado producto con un detalle por cliente, marca, número de parte y fecha de venta, lo que permite a los integrantes del departamento comercial identificar el último precio de venta para futuras cotizaciones o negociaciones de precios, de manera inmediata sin tener que depender de la persona encargada de la facturación que anteriormente debía buscar en el sistema transaccional la última factura del cliente en la que se facturó el producto de interés, reduciendo drásticamente el tiempo de respuesta a los clientes.

En este capítulo se describe el diseño e implementación de un data warehouse en la empresa NUO Tecnológica utilizando la metodología de Kimball denominada ciclo de vida dimensional del negocio, como resultado se obtuvo un almacén de datos centralizado, para apoyar la toma de decisiones, mediante el diseño de un proceso de extracción, transformación y carga (ETL) con la herramienta pentaho data Integrator que permitió consolidar, reformatear y limpiar los datos de múltiples fuentes de datos y cargarlos al almacén de datos. Mediante el uso de la herramienta Power BI se implementó un sistema de inteligencia de negocios, para el diseño de informes y análisis de la información del departamento comercial de la empresa.

Capítulo V

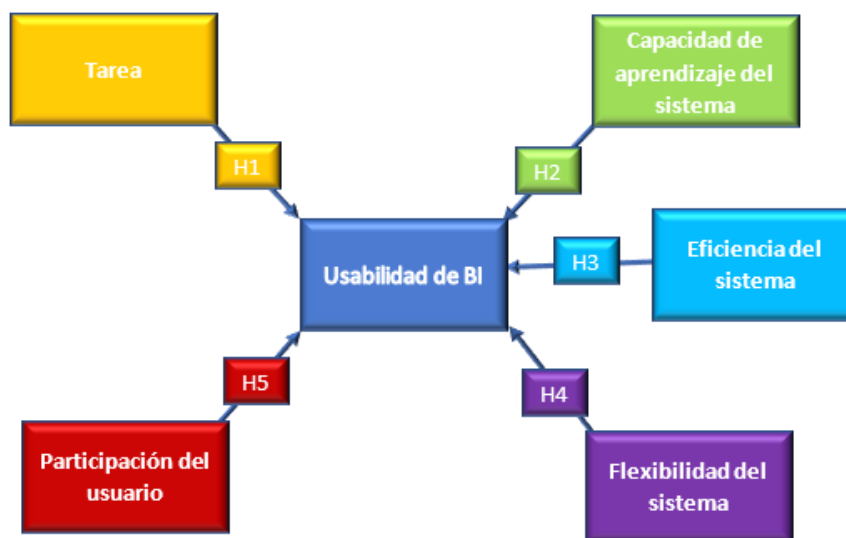
Validación De La Solución

En este capítulo se busca verificar si la implementación del sistema de gestión de datos en la empresa NUO tecnología permitió reducir los tiempos de acceso a la información para el apoyo a la toma de decisiones.

Una vez realizada la capacitación al gerente general, gerentes de cuentas corporativas y especialistas en precio y producto quienes forman parte del departamento comercial de la empresa NUO Tecnológica, se aplicó una encuesta a 9 integrantes del departamento comercial basado en el modelo de evaluación de la usabilidad en los sistemas de inteligencia de negocios propuesta por Elvis Proropat (Poropat, 2014) con el objetivo de evaluar la usabilidad del sistema de Business Intelligence (BI). Este modelo demuestra los factores que influyen en la usabilidad del sistema de BI:

Figura 24

Modelo conceptual propuesto de evaluación de usabilidad de BI



Tarea: Cuando la tarea en la que se utiliza el sistema de BI es más sofisticada, más usuarios del sistema de BI perciben que el sistema de BI es útil (H1).

Se plantearon las siguientes tareas con el objetivo de medir el tiempo que tarda cada usuario en realizarlas. La medición se realizó mediante el uso de un cronómetro digital para tres usuarios:

1. Total de ventas en el año 2019 de la empresa.
2. Cantidad vendida de equipos TC20 Zebra en el año 2018.
3. Ultimo descuento aprobado para el equipo Extreme Networks AP7532 y la cantidad de equipos venidos en el transcurso del 2020.

Tabla 20:

Medición de tiempos de tareas en el sistema de BI

Número de tarea	Tiempo Usuario 1	Tiempo usuario 2	Tiempo usuario 3	Promedio
1	19 seg	17 seg	26 seg	21 seg
2	40 seg	33 seg	38 seg	37 seg
3	43 seg	34 seg	54 seg	44 seg

Capacidad de aprendizaje del sistema: Cuando el sistema de BI es más fácil de aprender, esto mejora su percepción de usabilidad (H2).

Se realizó la medición del número de asistencias que recibieron cada uno de los usuarios en cada una de las tareas definidas en el punto anterior.

Tabla 21:

Número de asistencias a los usuarios del sistema BI

Número de tarea	Número de asistencias usuario 1	Número de asistencias usuario 2	Número de asistencias usuario 3
1	0	0	0
2	1	0	1
3	1	1	2

Eficiencia del sistema: Si el sistema de BI es más eficiente, los efectos positivos de eso se pueden ver a través de la usabilidad del sistema de BI (H3).

Para medir el tiempo que tarda el sistema en ejecutar una consulta y mostrarla al usuario se utilizó la herramienta de power BI analizador de rendimiento.

Tabla 22:

Tiempos de ejecución de consulta del sistema de BI

Número de tarea	Tiempo de ejecución
1	537ms
2	788ms
3	450ms

Flexibilidad del sistema: Cuando el sistema de BI sea más flexible, esto afectará positivamente su usabilidad (H4).

Participación del usuario: Tiene un efecto positivo en la usabilidad del sistema de BI, cuando el usuario del sistema de BI está más comprometido con el sistema de BI y sus procesos (H5).

Tabla 23:

Nivel de participación de los usuarios con sistemas de bi

Número Usuario	Nivel de experiencia sistemas de BI (1-10)	Nivel de experiencia computadores (1-10)
1	8	10
2	4	10
3	5	10
4	2	10
5	2	10
6	3	10
7	4	10
8	6	10
9	4	10

A continuación, se presenta los resultados de la encuesta.

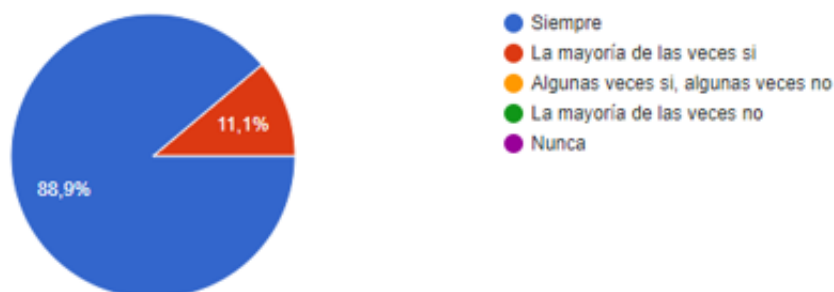
Tarea

Figura 25

Resultados pregunta 1 de la encuesta usabilidad de sistema de BI

1.- ¿El sistema de Business Intelligence (BI) le permite tener acceso a información histórica de ventas de la empresa, por ejemplo: Cantidad de equipos y precios históricos vendidos por cliente?

9 respuestas



En relación con la Pregunta 1, se tiene como resultado que el 88,9% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre les permite tener acceso a la información histórica de ventas por cantidad de equipos y por cliente, mientras que el 11,1% respondió la mayoría de las veces sí.

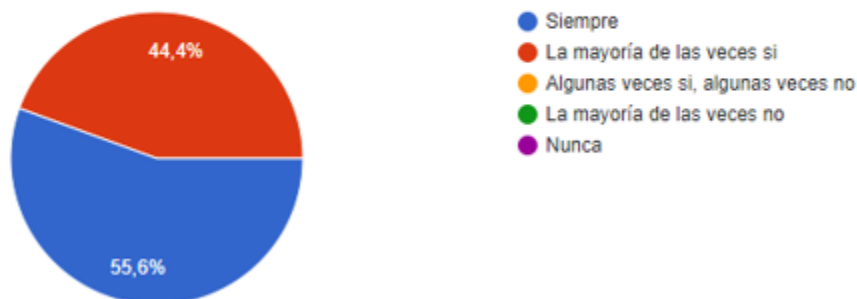
Capacidad de aprendizaje del sistema

Figura 26

Resultados pregunta 2 de la encuesta usabilidad de sistema de BI

2.- ¿El uso del sistema de Business Intelligence (BI) es intuitivo y de fácil aprendizaje?

9 respuestas



En relación con la Pregunta 2, se tiene como resultado que el 55,6% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre es intuitivo y de fácil aprendizaje, mientras que el 44,4% respondió la mayoría de las veces sí.

Eficiencia del sistema

Figura 27

Resultados pregunta 3 de la encuesta usabilidad de sistema de BI

3.- ¿El sistema de Business Intelligence (BI) apoya al desarrollo de las actividades de su área?

9 respuestas



En relación con la Pregunta 3, se tiene como resultado que el 88,9% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre apoya al desarrollo de actividades en el área comercial, mientras que el 11,1% respondió la mayoría de las veces sí.

Figura 28

Resultados pregunta 4 de la encuesta usabilidad de sistema de BI

4.- ¿El sistema de Business Intelligence (BI) le permite tener el acceso a la información de intereses en cualquier lugar donde la requiera?

9 respuestas



En relación a la Pregunta 4, se tiene como resultado que el 66,7% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre les permite tener

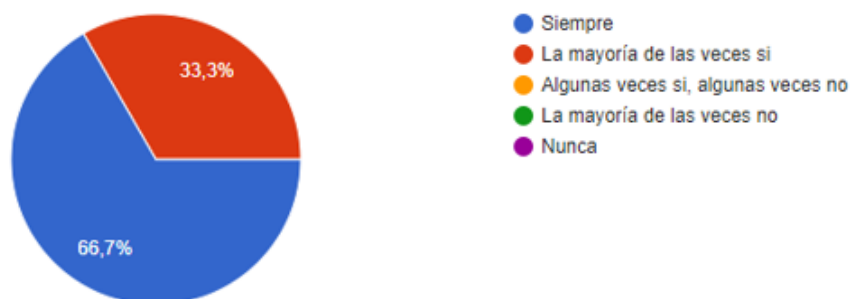
acceso a la información de interés en cualquier lugar donde la requiera, mientras que el 33,3% respondió la mayoría de las veces sí.

Figura 29

Resultados pregunta 5 de la encuesta usabilidad de sistema de BI

5.- ¿El sistema de Business Intelligence (BI) brinda facilidad para obtener los datos de interés?

9 respuestas



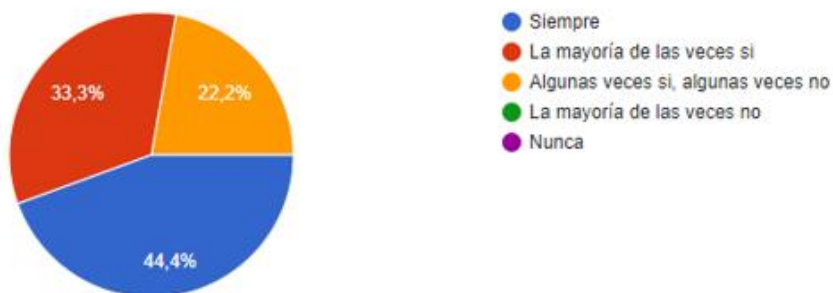
En relación a la Pregunta 5, se tiene como resultado que el 66,7% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre brinda facilidad para obtener los datos de interés, mientras que el 33,3% respondió la mayoría de las veces sí.

Figura 30

Resultados pregunta 6 de la encuesta usabilidad de sistema de BI

6.- ¿El sistema de Business Intelligence (BI) permite realizar reportes con facilidad de acuerdo con sus necesidades?

9 respuestas



En relación con la Pregunta 6, se tiene como resultado que el 44,4% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre permite realizar reportes con facilidad de acuerdo a sus necesidades, un 33,3% respondió la mayoría de las veces sí, mientras que un 22,2% respondió Algunas veces, algunas veces no.

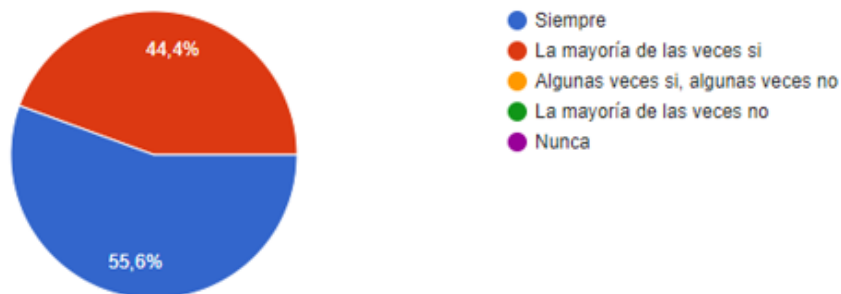
Flexibilidad del sistema

Figura 31

Resultados pregunta 7 de la encuesta usabilidad de sistema de BI

7.- ¿El sistema de Business Intelligence (BI) brinda la información necesaria para la toma de decisiones en el negocio?

9 respuestas



En relación con la Pregunta 7, se tiene como resultado que el 55,6% de los encuestados se encuentran de acuerdo en que el sistema de BI siempre brinda la información necesaria para la toma de decisiones, mientras que el 44,4% respondió la mayoría de las veces sí.

Figura 32

Resultados pregunta 8 de la encuesta usabilidad de sistema de BI

8.- ¿Cree usted que es posible replicar este sistema de Business Intelligence (BI) en otras áreas de la empresa para el apoyo en la toma de decisiones? Si la respuesta es "SI" mencione el área en que se podría replicar el sistema de BI?

9 respuestas

Si, área de distribución

Si, BI puede ser aplica a cualquier área de una empresa para cumplir con las estrategia de departamento. Ejm: área financiera, área de servicio al cliente, área de operaciones.

Si, en áreas como bodega en la que se puede tener el inventario en tiempo real de forma inmediata sin dependencia de terceros

Si.

Si, mercadeo, financiero

Si, es muy importante tenerlo en soporte técnico.

Si se podría aplicar en los departamentos de Soporte técnico, Industria, retail, financiero.

Absolutamente

El área de inventario

En relación con la Pregunta 8, se puede decir que el 100% de los encuestados están de acuerdo en que el sistema de BI se puede replicar en otras áreas de la empresa como:

- Área de distribución.
- Área financiera.
- Área servicio al cliente.
- Área de operaciones.
- Área de bodega.
- Área de mercadeo.
- Área de soporte técnico.
- Área de industria.
- Área de retail.

Participación del usuario

Figura 33

Resultados pregunta 9 de la encuesta usabilidad de sistema de BI

9.- ¿Cree usted que el sistema de Business Intelligence (BI) permite reducir tiempo para la toma de decisiones en el departamento comercial?

9 respuestas

Si
Si
Sí, BI permite agilizar proceso de toma de decisiones, gracias a su analítica predictiva.
Si. El contar con una herramienta comercial tan potente es importante.
Si, ya que tengo toda la data disponible lista para crear reportes necesarios.
Si debido a que permite tener acceso a la información sin depender de ninguna otra persona, facilitando la generación de reportes y apoyando a la toma de decisiones.
Totalmente
Si, ya que permite tener acceso a la información de interés en tiempo real

En relación con la Pregunta 9, se puede decir que el 100% de los encuestados están de acuerdo en que el sistema de BI permite reducir tiempo para la toma de decisiones por distintas razones:

- Permitiendo agilizar procesos debido a su analítica descriptiva.
- Permitiendo tener acceso a la información en tiempo real.
- La importancia de contar con una herramienta tan potente en el departamento comercial.
- Se tiene toda la data disponible para la generación de reportes.

Síntesis de las preguntas de investigación

Con el objetivo de resumir los principales resultados alcanzados en esta investigación, se presenta una síntesis de las preguntas de investigación que se plantearon al inicio de este trabajo de titulación.

OE1-RQ1.- ¿El uso de técnicas cualitativas de levantamiento de información permitirá determinar el estado actual del acceso a la información para de la toma de decisiones en el departamento comercial?

Mediante las entrevistas realizadas al gerente general y a la gerente de ventas de la región 1 de la empresa NUO Tecnológica, se pudo determinar que el acceso a la información para la toma de decisiones era muy complicado, principalmente por que la información de la empresa se encontraba almacenada en diferentes fuentes de información. Por ejemplo, en el sistema transaccional y las computadoras personales de cada uno de los integrantes del área comercial, por lo que era necesario reunir a todos los involucrados para poder generar los reportes de proyecciones de ventas, análisis de los equipos más vendidos, crecimiento en ventas por cliente y marca. Todos estos reportes se realizaban de manera manual unificando la información de interés en un archivo Excel lo que llevaba aproximadamente de 3 a 5 días laborables.

OE1-RQ2.- ¿El estudio del estado actual, permite identificar el principal problema para la toma de decisiones en el área comercial?

El análisis del estado actual permitió identificar todas problemáticas que existían en el área comercial de la empresa para la toma de decisiones. De todas estas problemáticas se identificó que la principal es la calidad de la información que se maneja para la toma de decisiones, debido a que cada uno de los empleados manejan distintas versiones de la información generada en la operación diaria de la empresa, información que es almacenada en sus computadores personales, siendo pocos los empleados que la almacenan en sistemas transaccionales, causando así demoras en los procesos internos y respuestas a los requerimientos de los clientes.

OE2-RQ1.- ¿Los artículos de investigación encontrados en la revisión inicial de literatura permiten identificar técnicas para la implementación de sistemas de gestión de datos?

Como se ha puesto en evidencia en el capítulo III, varios artículos encontrados en la revisión inicial de literatura detallan las distintas técnicas para la implementación de los sistemas de gestión de datos en las empresas, lo que permitió conocer las técnicas existentes y definir la que más se adapta a la empresa NUO Tecnológica.

OE2-RQ2.- ¿La revisión del estado del arte permitirá identificar estudios relacionados con la integración y centralización de la información proveniente de distintas fuentes de datos en una empresa?

En la revisión del estado del arte se identificó los estudios relacionados con la integración y centralización de la información en las empresas que manejan varias fuentes de datos, facilitando la selección de la herramienta para integración de datos pentaho data integrator (PDI) para aplicar las técnicas ETL (Extracción, transformación y carga) permitiendo integrar la data de las distintas fuentes que posee la empresa, adicional permitió identificar la metodología de Kimball, llamada modelo dimensional para la construcción de un data warehouse que permite la centralización de la información con el propósito de optimizar la toma de decisiones, con base en las consultas hechas en una base de datos relacional que están ligadas con la medición o un conjunto de mediciones de los resultados de los procesos de negocio.

OE3-RQ1.- ¿La metodología de diseño de sistemas de gestión de datos de Kimball permite identificar los beneficios del área comercial para el acceso a la información?

La metodología del modelo dimensional de Kimball permitió crear un data mart en base a las necesidades identificadas en el presente trabajo de titulación del área comercial de la empresa NUO Tecnológica para la toma de decisiones, fijando así una base para la creación de nuevos data mart de acuerdo con las necesidades de otras áreas de la empresa en el futuro.

OE3-RQ2.- ¿Cómo beneficia la implementación de un sistema de gestión de datos al área comercial en la empresa NUO Tecnológica?

La implementación de un sistema de gestión de datos beneficia al área comercial agilizando el proceso en la toma de decisiones, ya que permite el acceso a la información de interés en tiempo real, debido a que la data se encuentra lista para la creación de reportes el sistema de Business Intelligence lo que permite verificar el estado actual de la empresa de manera fácil e intuitiva.

OE4-RQ1.- ¿La implementación de un sistema de gestión de datos permite al personal del área comercial el acceso a información confiable para la toma de decisiones?

La implementación de un sistema de gestión de datos permite al personal del área comercial el acceso a información confiable para la toma de decisiones, debido a que la información se encuentra integrada desde las distintas fuentes de datos dentro del data warehouse, mediante el uso de técnicas ETL (Extracción, transformación y carga) que proporciona la herramienta pentaho data integrator, que fue la herramienta seleccionada en el presente trabajo de titulación.

OE4-RQ2.- ¿La implementación del sistema seleccionado permite que el personal del área comercial optimice tiempos en la toma de decisiones eliminando los reprocesos en la validación de información?

Sobre la optimización de tiempos para la toma de decisiones mediante la implementación del sistema de Business Intelligence se tiene que el 100% de los usuarios están de acuerdo con que el sistema permite optimizar tiempo en la toma de decisiones en el área comercial, debido a tienen acceso a información confiable en tiempo real desde sus computadores personales y dispositivos móviles cuando la requieran, utilizando la herramienta Power BI para la visualización de reportes que apoyan a la toma de decisiones.

Capítulo VI

Conclusiones y Recomendaciones

Conclusiones

- Se identificó el estado actual en el acceso a la información de la empresa NUO Tecnológica mediante entrevistas semiestructuradas realizadas a los gerentes regionales del área comercial, definiendo así los requisitos del negocio en base a la matriz de procesos y dimensiones de la metodología de Kimball, lo que permitió identificar cual es la información más relevante para la toma de decisiones en el departamento comercial.
- La revisión de literatura permitió determinar que existen varios artículos de investigación que detallan las técnicas más utilizadas para la implementación de un sistema de gestión de datos, lo que permitió identificar que la implementación de un data warehouse es la técnica más adecuada para la integración y centralización de la información de las distintas fuentes de datos como archivos de Excel y bases de datos transaccionales que posee la empresa NUO Tecnológica.
- El sistema desarrollado a partir de la metodología del modelo dimensional de Kimball fue valido debido a que durante el análisis de los datos almacenados en la base de datos transaccional se identificó que los datos son muy extensos y desordenados, adicional a esto no se contaba con la información a detalle por producto, información que se encontraba almacenada en cada uno de los computadores de los especialistas en precio y producto del área comercial, razón por la cual se creó una nueva fuente de datos en un archivo de Excel con toda la información de los productos ofrece la empresa, ya que sin esta fuente de datos no se podría obtener los análisis requeridos por la gerencia comercial. La

herramienta pentaho data integrator fue de gran utilidad para lograr la integración y centralización de la información de calidad y depurada.

- La publicación del sistema de Business Intelligence mediante el uso de la herramienta Power BI a partir de la implementación del data warehouse ha constituido una herramienta muy importante dentro del departamento comercial, ya que en base a la encuesta de usabilidad del sistema realizada a los integrantes del área comercial. Los resultados en general indican que el sistema cumple con los objetivos propuestos: un 100% están de acuerdo en que el sistema de BI permite reducir el tiempo para la toma de decisiones mediante la generación de informes que anteriormente demoraban de 3 a 5 días laborables ya que era necesario consolidar toda la información y que a partir de la implementación del sistema toma alrededor de 1 día el análisis de los datos para la toma de decisiones, ya que se tiene la data lista en el data warehouse para ser analizada de acuerdo a las necesidades de cada usuarios, un 66,7% indicó que el sistema permite tener acceso a la información en cualquier lugar y cuando la requiera, un 55,6% indicó que el sistema brinda la información necesaria para la toma de decisiones, un 55,6% indicó que el sistema es intuitivo y de fácil aprendizaje, un 44,4% indicó que el sistema permite realizar reportes con facilidad de acuerdo a las necesidades de los usuarios, un 88,9%.

Recomendaciones

- Se recomienda el uso de la metodología dimensional de Kimball cuando se quiere construir un data mart como primer elemento del sistema de análisis en este caso el de ventas, y luego ir añadiendo otros que compartan las mismas dimensiones ya definidas o incluyendo otras nuevas.
- Se recomienda el uso de pentaho data integrator, como herramienta de integración de datos ya que una de sus características principales es la de no

entrar en detalle de la implementación del cómo se hace cada una de las tareas, simplemente permite especificar qué es lo que queremos hacer.

- Se recomienda la implementación de sistemas de Business Intelligence en todas las áreas de la empresa ya que permite transformar los datos en información útil y con ello generar conocimiento para poder tomar las mejores decisiones para cada área.

Referencias bibliográficas

- Abreu, L. C. (2009). El impacto de las herramientas de inteligencia de negocios en la toma de decisiones de los ejecutivos. *International Journal of Good Conscience*, 16-52.
- Aucancela, G. M. (2019). La usabilidad en los sistemas de inteligencia de negocios, un caso práctico. *Ciencia Digital*, 321-327.
- Avella, A. (17 de Junio de 2014). *Calameo*. Obtenido de Calameo: <https://es.calameo.com/books/002299301667571c7ab05>
- Boon Keong Seah, N. E. (2014). Design and implementation of data warehouse with data model using survey-based services data. *Fourth edition of the International Conference on the Innovative Computing Technology*, 58-64.
- Cruz, C. P. (12 de Junio de 2017). *Gestiopolis*. Obtenido de Gestiopolis: <https://www.gestiopolis.com/herramientas-sistema-soporte-las-decisiones-dss/>
- Du, W., & Zou, X. (2015). Differential snapshot algorithms based on Hadoop MapReduce. *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 1203-1208.
- Eduardo Ahumada Tello, J. M. (2016). Businessintelligence:Strategy for competitiveness development in technology-based firms. *Contaduría y Administración*, 127-158.
- Guamán., M. A. (2019). La usabilidad en los sistemas de inteligencia de negocios. *Ciencia Digital*, 319-330.
- Guamán., M. A. (2019). La usabilidad en los sistemas de inteligencia de negocios, un caso práctico. *Ciencia Digital*, 321-327.
- Guevara, L. V. (21 de Octubre de 2019). *Readthedocs*. Obtenido de Readthedocs: https://gestionbasesdatos.readthedocs.io/_/downloads/es/latest/pdf/
- Hasan, L. A. (2015). An Integrated Measurement Model for Evaluating Usability Attributes. *Security and Advanced Communication*, 1-6.

- Inmon, W. (2002). *Building the Data Warehouse (Third Edition)*. New York: John Wiley & Sons, Inc.
- Josep Curto Díaz, J. C. (2010). *Introducción al Business Intelligence*. Barcelona: UOC.
- Joyanes, L. (27 de Mayo de 2013). *Big Data: Análisis de grandes volúmenes de datos en organizaciones*. Editorial Alfaomega. Obtenido de <http://sasybi.blogspot.com/2014/05/big-data-gestion-de-datos-no.html>
- Khurram Shahzad, J. Z. (2011). Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process analysis. *2011 FIFTH INTERNATIONAL CONFERENCE ON RESEARCH CHALLENGES IN INFORMATION SCIENCE*, 1-11.
- Legodi, I., & Barry, M.-L. (2010). The current challenges and status of risk management in enterprise data warehouse projects in South Africa. *PICMET 2010 TECHNOLOGY MANAGEMENT FOR GLOBAL ECONOMIC GROWTH*, 1-5.
- Marcelo, G. E. (2019). Análisis de metodologías para desarrollar Data Warehouse aplicado a la toma de decisiones. *Ciencia Digital*, 404-405.
- Marketing-analítico*. (1 de Marzo de 2017). Obtenido de Marketing-analítico: <https://www.marketing-analitico.com/analitica-web/proceso-analisis-datos/>
- Matamala, C. Z., Matamala, C. Z., Cuello, K. C., & Leiva2, G. A. (2011). Análisis de rendimiento académico estudiantil. *Ingeniare. Revista chilena de ingeniería*, 370.
- Ou, L., & Peng, H. (2006). Knowledge and Process Based Decision Support in Business Intelligence System. *First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*, 780-786.
- Pighin, M., & Marzona, A. (2012). Data value in decision process: Survey on decision support system in small and medium enterprises. *2012 Proceedings of the 35th International Convention MIPRO*, 1647-1654.

- Poropat, E. (2014). EVALUATION OF BUSINESS INTELLIGENCE SYSTEM USABILITY. *UNIVERSITY OF LJUBLJANA*, 30-40.
- QuestionPro. (19 de Julio de 2018). Obtenido de QuestionPro: <https://www.questionpro.com/es/analisis-de-datos.html>
- Ralph Kimball, J. C. (2004). *The Data Warehouse*. Indianapolis: Wiley Publishing, Inc.
- Ralph Kimball, M. R. (2010). *Relentlessly Practical Tools for Data Warehousing and Business Intelligence*. Indianapolis: Wiley Publishing, Inc.
- Rayo, Á. M. (2016). BIG DATA FOUNDATIONS. *Computer Training*.
- Rigaux, P., & Voisard, M. S. (2002). *Spatial Databases With Application To GIS*. San Francisco: Morgan Kaufmann.
- Rivadera*, G. R. (2010). La metodología de Kimball para el diseño de almacenes de. *Instituto Tecnológico de Buenos*, 57-60.
- Rivadera, G. R. (2019). La metodología de Kimball para el diseño de almacenes de datos (Data warehouses). *Cuadernos De Ingeniería*, (5), 56-71.
- Sanchez, W. (2011). La usabilidad en Ingeniería de Software: definición y características. *Revista de Ingeniería e Innovación de la Facultad de Ingeniería Universidad Don Bosco*, 7-21.
- Sassi, R. J., Arrivabene, A., & Romero, M. (2011). Zero Latency applied on a commercial Data Mart: Real-time information in support of decision making. *2011 IEEE 3rd International Conference on Communication Software and Networks*, 219-223.
- Schonberger, V. M. (2013). *Big data la revolución de los datos masivos*. Madrid: Turner Publicaciones.
- Serrano-Cobos, J. (2012). *Fundamentos de Bases de Datos*. Madrid: Jorge Serrano-Cobos.
- Serrano-Cobos, J. (2014). Big data y analítica web. Estudiar las corrientes y pescar en un océano de la data. *MASmedios*, 562.

- Serrano-Cobos, J., Serrano-Cobos, J., & Serrano-Cobos, J. (2012). *Fundamentos de Bases de Datos*. Madrid: Jorge Serrano-Cobos.
- Shenglei Pei, G. J. (2014). Application of Data Warehouse and Data Mining in the Steel Enterprise Information Integration System. *2014 Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics*, 181-184.
- Sretenović, M., Kovačić, B., & Jovanović, V. (2016). Development of the data warehouse model for public authorities accounts in Croatia. *2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 1509-1514.
- Sureshrao, G. S., & P., A. H. (2014). MapReduce-based warehouse systems: A survey. *2014 International Conference on Advances in Engineering & Technology Research (ICAETR - 2014)*, 1-8.
- Tascón, M. (12 de Junio de 2013). *Dialnet*. Obtenido de Dialnet: <https://dialnet.unirioja.es/servlet/articulo?codigo=4423775>

Anexos