

**UNIVERSIDAD DE EL SALVADOR  
FACULTAD DE INGENIERÍA Y ARQUITECTURA  
ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS**



**CURSO DE ESPECIALIZACIÓN DE INGENIERÍA DE DATOS  
ANÁLISIS DE INFORMACIÓN SOBRE ÓRDENES DE PRODUCCIÓN Y  
VENTAS DE LA EMPRESA "TOPAZIO" A TRAVÉS DEL SISTEMA ODOO  
Y MAGENTO**

PRESENTADO POR:

**HERNÁNDEZ CHÁVEZ, SALVADOR ENRIQUE  
ORDÓÑEZ VÁSQUEZ, ÁNGEL MAURICIO  
SANTAMARÍA MANCÍA, FRANCISCO BERNABÉ  
SIBRIAN ROSA, CRISTIAN ALEJANDRO**

PARA OPTAR AL TÍTULO DE:  
**INGENIERO DE SISTEMAS INFORMÁTICOS**

**CIUDAD UNIVERSITARIA, NOVIEMBRE DE 2024**

**UNIVERSIDAD DE EL SALVADOR**

RECTOR:

**M.SC. JUAN ROSA QUINTANILLA**

SECRETARIO GENERAL:

**LIC. PEDRO ROSALIO ESCOBAR**

**FACULTAD DE INGENIERÍA Y ARQUITECTURA**

DECANO:

**LUIS SALVADOR BARRERA**

SECRETARIO:

**ARQ. RAÚL FABIÁN**

**ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS**

DIRECTOR:

**ING. CÉSAR AUGUSTO GONZÁLES**

**UNIVERSIDAD DE EL SALVADOR**  
**FACULTAD DE INGENIERÍA Y ARQUITECTURA**  
**ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS**

Trabajo de Graduación previo a la opción al Grado de:

**INGENIERO DE SISTEMAS INFORMÁTICOS**

Título:

**ANÁLISIS DE INFORMACIÓN SOBRE ÓRDENES DE PRODUCCIÓN Y  
VENTAS DE LA EMPRESA “TOPAZIO” A TRAVÉS DEL SISTEMA ODOO  
Y MAGENTO**

Presentado por:

**HERNÁNDEZ CHÁVEZ, SALVADOR ENRIQUE**  
**ORDÓÑEZ VÁSQUEZ, ÁNGEL MAURICIO**  
**SANTAMARÍA MANCÍA, FRANCISCO BERNABÉ**  
**SIBRIAN ROSA, CRISTIAN ALEJANDRO**

Trabajo de Graduación Aprobado por:

Docente Asesor:

**ING. RENE FABRICIO QUINTANILLA GÓMEZ**

SAN SALVADOR, NOVIEMBRE 2024

Trabajo de Graduación Aprobado por:

Docente Asesor:

**Ing. Rene Fabricio Quintanilla Gómez**

# Índice

Introducción	11
Capítulo I: Especificación de proyecto	12
a) Situación actual	12
I. Antecedentes	12
II. Descripción del problema	13
III. Planteamiento del problema	14
b) Objetivos	16
Objetivo general	16
Objetivos Específicos	16
c) Alcances	17
d) Hipótesis	18
d) Justificación	19
e) Cronograma de actividades	20
f) Presupuesto	21
Capítulo II: Análisis y diseño de la propuesta de solución	22
a) Metodología de trabajo	22
Fase de diseño del Data Warehouse	23
Fase de la creación de tablero para visualización de datos	24
b) Descripción de la propuesta de solución	25
Descripción del dataset y diccionario de datos.	26
Resultado del Data Profiling	56
Especificación de las necesidades analíticas	61
Mapping de modelo dimensional y tablas del dataset	65
c) Descripción de la tecnología a utilizar	67
d) Diagrama arquitectónico de la solución	68
e) Descripción de cada componente de la solución	68
Capítulo III: Estrategia de implementación de propuesta de solución	69
a) Estrategia de implementación	69
b) Presupuesto de implementación (distribuido en rubros)	75
c) Análisis de resultados	76
Conexión a la base de datos del Data Warehouse	76

Paquetes de Transformación	77
Fact Sales:	81
Fact Lote:	82
Nuevos registros:	82
Fact Inventory Movements:	83
ETL Principal:	84
Reportes en Power BI	84
Dashboard de Ventas	85
Dashboard: análisis de producción	86
Conclusiones	90
Recomendaciones	91
Bibliografía	92
Anexos	93
Anexo 1. Imagen de entrada y salida de datos	93
Anexo 2. Script para dimensión de catalog product:	94
Anexo 3. Diagrama de modelo relacional del proyecto	95

## **Agradecimientos**

Agradezco a mis padres, Nancy Lissette Chávez y Ricardo Amílcar Hernández, por guiarme, apoyarme, animarme y velar por mí en cada etapa de mi vida sin descuido. A mis hermanos y familiares, Ricardo Chávez, Patricia Hernández entre más que no podría listar, por vivir como tuyas mis aflicciones como mis victorias. A mis amigos y compañeros presentes en mi vida y aquellos que ya no están, que han aportado a mi vida grandemente. Pero agradezco sobre todo a Dios, por enseñarme cuan grande es, por darme lo más valioso en mi vida que es mi familia, por demostrarme que de su mano no hay obstáculo imposible.

Por enseñarme a confiar en él.

Por enseñarme a confiar en mí.

Por recordarme qué la esperanza es algo bueno, tal vez lo mejor,  
y lo bueno nunca muere.

Salvador Enrique Hernández Chávez

Primeramente a Dios, por darme la salud y la resiliencia necesaria para llegar a este punto en mi vida, a mis padres, Guadalupe Vásquez de Ordoñez y Rigoberto Alfonso Ordoñez Velado, que sin duda han sido un pilar fundamental durante todo este camino, gracias por siempre creer y confiar en mí. A mi pareja Karen Fuentes, por creer y confiar en mí incluso en los días en los que sentía que no podía continuar más. A los compañeros y amigos que a lo largo de la carrera nos encontramos y compartimos no solo proyectos académicos si no también experiencias que nos han forjado como personas y como ingenieros en este camino que estamos a punto de culminar.

Ángel Mauricio Ordoñez Vásquez

Gracias a Dios por darme fuerzas y ánimos para luchar durante este camino, a mi madre que es un pilar fundamental en mi vida y que sin duda sin ella no hubiera podido avanzar hasta el día de hoy, a mis compañeros que me brindaron ayuda cuando yo lo necesité, a mis docentes que me enseñaron lecciones de vida y conocimiento como persona y sin duda esto marca un nuevo comienzo para crecer como persona y abrir nuevos horizontes, todavía hay camino que recorrer.

Francisco Bernabé Santamaría Mancía

Agradezco a Dios y al apoyo de mi familia, especialmente a Angelica Rosa Hernández, mi querida madre por todo el amor y ánimo brindado. A profesores, compañeros y amigos con los que he trabajado y que han hecho posible este proceso que estamos culminando satisfactoriamente.

Cristian Alejandro Sibrian Rosa

## Introducción

La evolución de las empresas y la necesidad de los negocios de gestionar los órdenes de producción a lo largo de los años es inmensa, actualmente el volumen de datos proveniente de distintas fuentes ha ido incrementando, por lo que al ser el principal recurso de la empresa es necesario la gestión correcta de la información. Es esencial la transformación adecuada de los datos históricos, para tener la facilidad del análisis de una forma más intuitiva para tomar mejores decisiones estratégicas para la mejora continua de los procesos operativos, la incidencia de la tecnología en los procesos de negocio, la aparición de nuevos productos y la evolución de los mercados, llevan a las empresas a permanecer en constante actualización, para poder reaccionar de una forma ágil y eficaz. La ingeniería de datos es la disciplina que se encarga de recolectar, trasladar y validar datos para su explicación, por tanto, se vuelve una necesidad para la empresa Topazio mejorar el proceso de órdenes de producción que se tiene dentro de la compañía, y para ello se plantea la creación de un Data Warehouse que ayudara al procesamiento de los datos y el análisis de la información.

Con la aparición de nuevas herramientas en la era digital que actualmente se vive, se vuelve necesario contar con metodologías que puedan analizar la información y sobre todo el manejo de grandes volúmenes de datos. Las bases de datos comunes relacionales se vuelven más limitadas sobre todo cuando se presentan datos no estructurados por lo que surgen una innumerable cantidad de nuevas tecnologías para completar esas necesidades. Por lo que considerando todos estos factores se propone implementar el modelo dimensional haciendo uso de herramientas tales como Odo, Magento, Power BI, SSS, Visual Studio, estas tecnologías facilitaran la extracción, transformación y la carga de datos a nuestro modelo dimensional.

La empresa Topazio actualmente no cuenta con algún método o prototipo de análisis de datos significativo para dar una mejor administración a futuro relacionado con nuevas decisiones a tomar como ventas, descuentos a clientes, productos más vendidos, tiempo de producción. Sin embargo, está al alcance de la organización proporcionar los recursos necesarios para que se elabore el análisis respectivo para las especificaciones que se han solicitado con el fin de mejorar la toma de decisiones en la empresa. Se utiliza Odo para registrar los movimientos de inventario, almacenando los productos que ingresan a la empresa y con el software Magento se realiza la venta de los productos en las diferentes tiendas. Con la creación de un Data Warehouse no es simplemente una infraestructura de almacenamiento de datos, sino una herramienta estratégica esencial para potenciar a las organizaciones en la era de la información, permitiéndoles convertir datos en decisiones y oportunidades en sus procesos empresariales.

# Capítulo I: Especificación de proyecto

## a) Situación actual

### I. Antecedentes

La empresa Topazio es una compañía que se dedica a la fabricación de camisas para caballero, esta empresa cuenta con más de 5 años en el mercado. Topazio a lo largo de los años ha crecido de manera sostenida gracias a su esfuerzo por crear productos de calidad y su enfoque de innovación en los procesos de fabricación por lo que distintos proveedores abastecen constantemente con materia prima de alta calidad, la empresa ha presenciado un alto crecimiento en sus ventas a lo largo del tiempo y la acumulación de clientes de diversos países. La organización se ha visto en la necesidad de conocer cuál es el comportamiento por el que atraviesan sus productos, así como los que se venden más, donde es que se vende más o incluso en que época hay más ventas para el almacén. La lógica de negocio implementada en la empresa Topazio se lleva a cabo cuando ellos inician el proceso de compra, se empieza por seleccionar el proveedor al cual solicitara los productos de la orden de compra y especifica la fecha requerida de la entrega, luego se eligen los artículos que necesita pedir al proveedor, estos productos son buscados por código del producto a su vez se establece la cantidad e impuestos que la orden de compra tendría.

Topazio utiliza Odoo como su sistema ERP para gestionar sus procesos de fabricación, este sistema permite tener un control sobre cada etapa de producción, así como la planificación de los lotes de producción, Odoo permite analizar el desempeño de los equipos de trabajo y productividad permitiendo conocer datos claves tales como: camisas producidas por hora, esto es esencial para la toma de decisiones operativas y optimizar los tiempos de producción y fabricación.

La organización Topazio ha decidido centrarse en puntos de venta físicos, donde utiliza la plataforma de Magento para gestionar las ventas en tiendas físicas, este sistema permite controlar el inventario de los puntos de venta, así como las transacciones y la gestión del stock. El uso de Magento permite optimizar la administración de inventarios en tiempo real, garantizando que siempre haya disponibilidad de productos en las tiendas. Además, esta plataforma facilita la gestión de múltiples tiendas físicas, ubicadas en diferentes puntos ya sea puestos locales o en centros comerciales, lo que permite a la empresa diversificar sus puntos de distribución y alcanzar clientes potenciales.

## II. Descripción del problema

En la actualidad, el alto incremento de datos suele generar una serie de inconvenientes para las empresas al momento de buscar información o un conjunto de datos que compartan alguna peculiaridad en común, al intentar detectar alguna anomalía o algo que este fuera de lo relativamente normal o realizar algún tipo de análisis en específico. Este tipo de inconvenientes tienen mayor presencia en las empresas que realizan a diario un considerablemente alto volumen de ventas y/o cuyos negocios llevan mucho tiempo en el mercado acumulando una gran cantidad de datos como consecuencia de las actividades del negocio y por lo que tomaría mucho tiempo consultar un dato en específico. Teniendo en cuenta que muchas empresas están empezando a evolucionar involucrando la actividad de sus negocios a la era digital y que los datos todas sus ventas han sido migradas de forma digital dejando atrás el papel; a pesar de esto, la búsqueda de información que se necesiten de forma casi inmediata o con alguna característica en común dentro de una inmensa cantidad de datos, representa una dificultad.

La empresa Topazio compra productos de distintos proveedores para realizar camisas, estos artículos se solicitan por medio de órdenes de compra. Sin embargo, la empresa no tiene la información precisa para poder realizar el respectivo análisis de producción y como se lleva la gestión de compras. Actualmente la cantidad de información y actualización de datos que se genera es inmensa, sumado a la variedad de formatos se crea una complejidad para la toma de decisiones. Este problema no solo afecta a los procesos operativos de la empresa, sino que también obstaculiza la capacidad de identificar oportunidades y disminuir riesgos. La empresa ha presenciado un alto incremento en sus ventas a lo largo del tiempo y la acumulación de datos de transacciones de diversos países, clientes y productos al momento de la realización de una venta. La organización se ha visto en la necesidad de conocer cuál es el comportamiento por el que atraviesan sus productos, así como los que se venden más, cuáles son los movimientos de inventario o incluso en que época hay más ventas para el almacén.

Los problemas se presentan a la hora de la toma de decisiones gerenciales en áreas como: compras, ventas y producción, es necesario conocer los valores y datos que entran en juego para poder realizar métricas y presentar reportes que apoyarán a estos departamentos, los cuales serán:

1. Que producto se vende más en un periodo de tiempo.
2. Que producto fue comprado menos en una cantidad de tiempo.
3. Cantidad de compras por periodo de tiempo.
4. Total, de productos por proveedor y en un periodo de tiempo.

La implementación de un Data Warehouse emerge como una solución estratégica para superar estos desafíos. La creación de un repositorio centralizado permitirá la

extracción, transformación y carga (ETL) de datos desde diversas fuentes hacia un entorno unificado. Esto facilitará análisis más rápidos y precisos, así como la generación de informes coherentes que respalden la toma de decisiones informada en todos los niveles de la organización.

### **III. Planteamiento del problema**

Cuando se trata de Ingeniería de Datos y se busca la manera de desarrollarla correctamente, hoy en día la tecnología brinda innumerables posibilidades de efectuar procesos de modelados de datos ya que no siempre tienden a estar en una estructura estándar como lo son las bases de datos relacionales, por ejemplo, y vienen representadas en formatos no legibles donde los datos no tienen una organización o estructura específica, simplemente están “existiendo”; así también para realizar procesos de extracción, transformación y carga de datos que para muchos nos podría resultar complicado y desconocido hacer una depuración de datos y posteriormente realizar una representación de ellos a partir del resultado de esa misma depuración para lograr identificar alguna característica en particular según los criterios y necesidades que se hayan tenido antes de realizar dichos procesos, aún más cuando la información es demasiada y variada, sin tomar en cuenta la falta de estructuración de los datos antes mencionada.

A pesar de la gran cantidad de herramientas disponibles en la era digital para el trabajo dentro de la Ingeniería de Datos se resulta desafiante realizar un manejo limpio y completo de los datos, lo cual da origen a la necesidad de identificar qué herramienta tecnológica se acopla más adecuadamente a la necesidad que se presente ya que algunas tienen mayor capacidad de procesamiento analítico que otras y ventajas convenientes para la organización. Sin embargo, no solo se debe de considerar la capacidad analítica y efectiva con la que estas herramientas trabajen, sino también el tiempo que estas requieran para realizar los procesos, ya que sería contraproducente hacer uso de una herramienta “poderosa” para el análisis pero que consuma mucho tiempo y recursos (hardware, software adicional como extensiones de paga, entre otros) para llevarlo a cabo. Otra dificultad que se suele manifestar, cuando se habla de trabajar con software desconocido, es lo fácil y práctica que sea la interacción con el usuario ya que muchas herramientas suelen cumplir su función, pero suele resultar complicado para muchos lograr comprender su correcto y óptimo funcionamiento.

La empresa Topazio, la cual se enfoca en la producción de camisas para caballeros, tiene dificultades en los procesos operativos debido a una gestión poco efectiva de sus fábricas (workcenters), con el propósito de abordar estas necesidades identificamos el problema a partir de entradas, procesos y salidas que nos llevará a la implementación del Data Warehouse. Por lo que se plantea una solución a cada una de las necesidades de la empresa Topazio.

Entradas: Son las que se obtienen a partir del sistema ERP de las ordenes de compras realizadas por la empresa Topazio.

Procesos: Los distintos recursos que se utilizaron para la extracción, transformación y carga de datos tales como SISS, asimismo como los que servirán para la presentación, visualización de datos y creación de informes con estos valores en este caso Power BI.

Salidas: El dashboard con los distintos reportes que se obtendrán de la transformación de datos tales como la cantidad de productos comprados en una fecha, el monto de compras, productos más comprados y productos más vendidos.

Una herramienta muy importante dentro de este proyecto será Power BI que se utiliza para convertir datos sin procesar en información significativa mediante el uso de visualizaciones y tablas intuitivas. Gracias a esta herramienta, se puede analizar los datos fácilmente y tomar decisiones comerciales importantes basadas en ellos; y que, al igual que la herramienta de ETL, Power BI también posee muchas ventajas para el cómodo manejo con el usuario. En la organización Topazio se logran identificar estos problemas antes mencionados por lo que se ha tomado a bien solventar dichas dificultades. El objetivo principal es adquirir e implementar herramientas digitales para realizar la Ingeniería de Datos con las ventas de la empresa. Esto permitirá enfocarse en la lógica de negocios y representar datos específicos según las necesidades para tomar decisiones que beneficien a la empresa.

## **b) Objetivos**

### **Objetivo general**

Diseñar y construir un modelo funcional para la correcta implementación de un Data Warehouse tomando de referencia principal la producción y venta de camisas para la empresa TOPAZIO que le permita optimizar la planificación de producción, para tener una gestión más ágil y a su vez mostrar información más concisa que apoye de forma positiva la toma de decisiones.

### **Objetivos Específicos**

1. Crear un modelo dimensional que refleje la estructura del prototipo de Data Waterhouse para la Empresa Topazio.
2. Utilizar enfoques y procedimientos especializados para extraer, transformar y cargar datos, con el propósito de definir claramente las especificaciones requeridas para realizar un análisis de datos exhaustivo, eficaz y preciso.
3. Identificar los datos más relevantes de la base transaccional que se integrarán en el modelo dimensional, para ser utilizados en la transformación y desarrollo del Data Warehouse de Topazio.
4. Usar la herramienta Microsoft Power BI para desarrollar informes basados en los datos procesados a través del ETL diseñado, con el propósito de facilitar a la alta gerencia el análisis y la toma de decisiones fundamentadas, ajustándose plenamente a los requerimientos y necesidades previamente establecidos.

### **c) Alcances**

1. Mejorar el proceso de toma de decisiones empresarial poniendo a disposición un almacén de datos confiable, de fácil acceso donde se puedan hacer consultas de manera ágil y segura.
2. Diseñar un modelo dimensional de Data Warehouse para los procesos de producción por lotes, movimientos de inventario y la venta en puntos físicos. Este modelo brindara información relevante para que pueda ser analizada y usada para facilitar la toma de decisiones.
3. Se identificarán y extraerán datos relacionados a los procesos de elaboración y venta de productos. Esto implica obtener información detallada del software ERP Odoo y del software para ventas Magento asegurando que estos datos sean relevantes y confiables.
4. Se crearán procesos estandarizados de extracción, transformación y carga de datos a los Data Warehouse con la herramienta SSIS en MS Visual Studio.
5. Diseñar y crear una visualización de datos con indicadores clave sobre la producción y venta de camisas que brinden información gráfica sobre el resultado de los procesos de producción por lotes, movimientos de inventario y ventas en tiendas físicas.

## d) Hipótesis

La realización de un análisis para conocer cómo se ejecutan los procesos operativos en una empresa es de suma importancia, La organización Topazio no cuenta con la información para analizar y mejorar sus procesos productivos tales como: tiempo de producción por producto, el movimiento de producto por ventas, productos más vendidos; por lo cual estas se ven como amenazas ante los desafíos que los negocios enfrentan, ante tal situación desarrollar un proceso ETL tiene como objetivo facilitar la obtención de información para la toma de decisiones para poder optimizar la gestión de órdenes de producción de la empresa Topazio ya que no cuentan con los recursos para conocer indicadores tales como tiempo de producción de cada workcenter, tiempo de producción por producto, con base a lo anterior el análisis está orientado al desarrollo de procesos ETL para el análisis de la información dentro de la empresa Topazio. El análisis se realizará teniendo como estudio la información de los procesos de producción, la fabricación de productos en las diferentes fábricas y el movimiento de los productos para su venta.

Se analizan las fases de los procesos de producción y ventas, también se analizan costos de producción como se pueden minimizar los gastos de operación, posteriormente se describen las ventajas y desventajas de crear un ETL, también se analizan los diferentes métodos aplicados a los costos cuanto se gasta en crear un producto, cuanto tiempo se tarda producir un lote y el movimiento de productos por ventas, también se describe las características, deficiencias y la forma en la que opera esta empresa, la propuesta de solución se plantea por medio del Data Warehouse y como este puede solventar la necesidad de saber que cuanto se produce a que costo y que productos son los que se venden más, esta solución se realiza mediante la aplicación de procesos ETL enfocado al desarrollo de los procesos de principales de la empresa Topazio, estos procesos tiene como objetivo principal comprender los movimientos de inventario, costos y ordenes de producción, esto facilita la mejora de los procesos productivos a través del análisis que se logra con la carga de información y resultados del Data Warehouse, Así como también permite establecer con certeza y confiabilidad los precios de venta y los márgenes de utilidad. Finalmente se plasmaron las conclusiones obtenidas sobre la base de los objetivos que fueron propuestos al inicio del trabajo.

#### **d) Justificación**

La obtención de datos desde un sistema transaccional de manera específica era una forma común para la creación a futuro de reportes de los cuales las empresas se les facilitarían el análisis para futuras decisiones o acciones de las cuales dependerán la mejora de esta. La problemática de la empresa Topazio es que para realizar un análisis de datos se requiere de mucho tiempo y recursos, esto influye en la toma de decisiones específicas, ya que su lógica de negocio ha sido nada más que solo vender sin ningún otro enfoque favorable para el negocio, así de igual manera la demostración de la información almacenada de manera resumida y útil no ha sido la adecuada conforme a la necesidad que se presenta para mejorar el negocio de la organización, adicionalmente se necesita el software necesario para llevar a cabo toda esta metodología de análisis para futuras ocasiones, también existe la necesidad de la depuración de los datos ya que la cantidad de información es grande y suelen existir datos que no sean necesarios según lo que se necesite analizar y posteriormente presentar gráficamente de alguna manera.

Cada vez el procesamiento de grandes cantidades de datos se vuelve complejo para sistemas transaccionales, existen métodos que ayudan a trabajar con datos no estructurados a cierto punto se vuelven pocos eficientes. Una manera de optimizar y mejorar estas deficiencias es el uso de Data Warehouse, la cual nos permite gestionar diferentes orígenes de datos incluso si estos no son estructurados, facilitando el manejo de grandes volúmenes de datos conllevando a que nuestro usuario final tenga accesibilidad a ellos, favoreciendo las tareas analíticas lo cual es fundamental en cada negocio.

El proceso de compras en una empresa es una tarea fundamental del negocio por lo que facilitar el estudio de las variables que interesan a una compañía para futuras proyecciones o análisis de las conveniencias de ciertos productos al crecimiento de ellas es una de las principales facilidades que el Data Warehouse ayudara. La presentación de estos datos de una forma que el usuario será capaz de extraer y transformar para analizarlos a través de reportes de herramientas que ayudan a presentar datos de una forma amigable y entendible para las distintas áreas de interés de la empresa.

## e) Cronograma de actividades

Ues | Proyecto de especialización para empresa Topazio

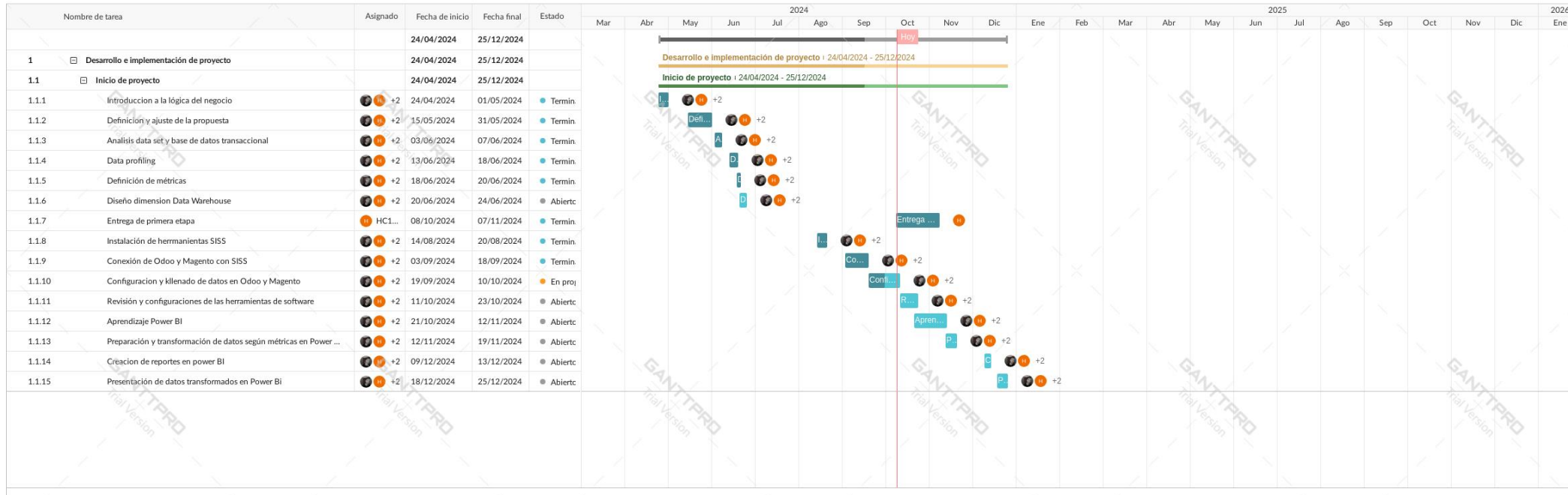


Figura 1. Cronograma de actividades para el desarrollo del Data Warehouse

## f) Presupuesto

A continuación, se presentan los costos incurridos durante el desarrollo del proyecto, se consideran recursos humanos, tecnológicos y servicios básicos que fueron necesarios para el desarrollo de la solución.

N°	Actividad o Recurso	Descripción	Cantidad	Precio Unitario (\$)	Costo Total (\$)	Costo Total Grupal (\$)
1	Energía Eléctrica	Representa un costo estimado mensual por cada integrante del equipo, para alimentar las computadoras y otros dispositivos informáticos relacionados con la realización del proyecto. Costo estimado por cada integrante.	8 meses	10	80	320
2	Connector Odoo-Magento	Conector desarrollado por Cybrosys, necesario para la comunicación entre Odoo y Magento y viceversa	1	54.86	54.86	54.86
3	Internet Residencial	Internet utilizado durante todo el periodo del proyecto, con una capacidad de 55Mbps. Costo estimado por cada integrante.	8 meses	26	208	832
4	Ingeniero de datos	Encargado de preparar, organizar y gestionar grandes cantidades de datos.	4	1200	4800	4800
5	Hardware	Memoria RAM 16GB, debido a requerimientos de software utilizado	2	80	160	160
6	Capacitaciones	Cursos relacionados a la utilización de las herramientas Odoo y Magento	2	35	70	70
					Total	6,236,86

**Tabla 1.** Presupuesto para el desarrollo del proyecto

## **Capítulo II: Análisis y diseño de la propuesta de solución**

### **a) Metodología de trabajo**

La metodología de trabajo que se utilizó es la metodología tradicional o también conocida como metodología en cascada, en la cual definimos cinco fases principales a partir de la definición del proyecto que incluye la fase de análisis de los procesos del negocio a través de los sistemas que usan para registrar sus flujos de trabajo, la exploración de datos desde datos transaccionales con las que cuentan que para Odoo es una base de datos en PostgreSQL y para Magento es una base de datos en MySQL. Con la información sobre los flujos de trabajo y la manera como se almacenan se procede a realizar un diseño de Data Warehouse en forma de dimensiones y tablas de hechos organizados en un diagrama de estrella. Luego en la siguiente fase se definen los detalles, procesos y herramientas para extraer, darle tratamiento y guardar los datos en el Data Warehouse. Por último, se procedió a diseñar y crear la presentación de datos.

### **Fase de análisis de los procesos del negocio**

Topazio es una empresa dedicada a la fabricación de prendas de vestir, se especializa en la producción de camisas. Para todos sus procesos internos usa el software Odoo. Por otro lado, usa el software Magento para la venta puntos de venta físicos ubicados en el país. Para poder analizar los procesos de interés se hizo una investigación exploratoria de los sistemas que almacenan la información. Se hizo uso de la herramienta Docker desktop para crear contenedores de Odoo en la versión 16 y Magento en su versión 2.3 ya que son las versiones que la empresa Topazio usa actualmente para sus procesos internos. Para ello se diseñaron archivos de configuración llamados docker-compose en donde se establecieron los contenedores, redes, volúmenes y la interacción entre estos que son necesarios para el funcionamiento de Odoo y Magento. (se puede agregar en anexo los archivos docker-compose para Magento y Odoo)

El siguiente paso fue definir como se desarrollan los procesos de producción y venta. Para ello se hizo uso de la interfaz gráfica de los sistemas. Se analizó el flujo de trabajo para definir productos, configurar sus categorías, precios, lista de materiales, centros de producción, estrategias de manejo de inventario, manejo de la producción y trazabilidad por lotes de los productos, así como la gestión de clientes y ventas.

### **Etapas exploración de datos**

Luego de tener conocimiento de los flujos de trabajo, es necesario conocer cómo se almacena la información de estos procesos, con qué estructura y en qué formatos. Para ello procedimos a la instalación del software DBeaver que es una herramienta de base de datos universal, libre y de código abierto para desarrolladores y administradores de bases de datos esto nos ayudaría a explorar las bases de datos que tenemos alojadas en los contenedores de Docker y contienen la estructura de cómo se almacena la información.

Como resultado pudimos identificar las tablas que se relacionan a los módulos de producción y venta usados por la empresa Topazio. Creamos un dataset con las tablas identificadas de Magento y Odo. Identificamos las columnas que contienen los datos clave como tiempo de producción, costo unitario, precio de venta, etc. Toda esta información es crucial para la creación de un modelo dimensional de Data Warehouse.

### **Fase de diseño del Data Warehouse**

Para la creación del modelo dimensional del Data Warehouse lo primero fue definir los procesos del negocio que son importantes para los objetivos del proyecto. Los procesos seleccionados son:

1. El proceso de producción, el cual definimos como uno de los más importantes de acuerdo con el objetivo general del proyecto. Este proceso incluye la administración de variantes de productos, los detalles de estos como tallas y colores. También define la lista de materiales que son requeridos para su elaboración, así como la trazabilidad por lotes una vez que ingresan al inventario. El procedimiento principal sería el de crear ordenes de producción para un producto en específico, crear las ordenes de trabajo necesarias a los centros de trabajo hasta obtener el producto final.
2. El proceso de administración de inventario. Este proceso se encarga del control de entradas y salidas de inventario, toda la información relevante de los productos que se mueven, el origen y destino y las razones de estos movimientos.
3. Proceso de venta. Este proceso es otro de los procesos con mayor importancia debido a que la información de obtenemos de sus operaciones es imprescindible para lograr los objetivos del proyecto. En el proceso de venta intervienen los productos con sus variantes, los locales que están distribuidos en centros comerciales, los precios de venta y las fechas y términos en que se realizan.

Con la información de los procesos, la estructura de los datos y su significado se diseñaron tres modelos dimensionales:

- a) Modelo dimensional de ventas
- b) Modelo dimensional de producción por lotes
- c) Modelo dimensional de movimientos de inventario

### **Fase extracción, transformación y carga**

Para la fase de diseño de los procesos de extracción, transformación y carga (ETL) se usa toda la información que poseemos de los procesos de producción y venta, principalmente la información sobre las fuentes de datos. Se procedió a instalar el software MS Visual Studio que es una herramienta de desarrollo potente que puede ser usar para completar todo el ciclo de desarrollo en un solo lugar, dentro de Visual Studio procedimos a instalar la herramienta SQL Server Integration Services que es una plataforma para crear soluciones de integración y transformación de datos a nivel empresarial.

Con esta herramienta se diseñaron e implementaron los siguientes procesos:

1. Extracción de datos. Se procedió a conectar al proyecto de SSS las bases de datos que sirven de origen de datos. Se crearon y configuraron los elementos de Integration Services que interactúan con las fuentes de datos para la extracción.
2. Transformación de datos. Se agregaron todas las transformaciones necesarias para que los datos estuvieran listos para ser cargados en el data Warehouse.
3. Carga de datos. Se crearon las tablas en una base de datos de SQL Server que se usó como destino para almacenar y actualizar los datos ya procesados.

### **Fase de la creación de tablero para visualización de datos**

Para esta fase se diseñaron tableros que presenten la información de una manera visual y sumamente fácil de entender por los usuarios. Para ello se instaló y configuro la herramienta MS Power BI. Con los datos ya cargados en nuestros Data Warehouse procedimos a conectarlos a la herramienta Power BI para la implementación de los tableros que representaran la información de los procesos de la empresa Topazio en sus procesos de producción y venta.

## **b) Descripción de la propuesta de solución**

Como parte de la propuesta de solución se propone el siguiente análisis para el modelado dimensional:

### 1. Proceso de negocios

La actividad seleccionada para su análisis es el proceso de producción y ventas.

### 2. Nivel de granularidad

#### a. ¿Qué detalle requiere el usuario del negocio?

El usuario de negocio necesita conocer indicadores ver la producción de camisas en función de horas y por equipo de producción.

#### a. ¿Qué detalle requiere el usuario del negocio?

El usuario de negocio necesita conocer indicadores ver la producción de camisas en función de horas y por equipo de producción.

El usuario de negocio necesita ver las ventas en función de productos, ventas por año/mes/día.

#### b. ¿Qué detalle es efectivamente posible con los datos?

Se pueden ver las ventas en función de todo lo que el usuario de negocio requiere de acuerdo con los datos del sistema transaccional que actualmente cuenta con los datos necesarios para poder presentar ese tipo de información.

### 3. Identificar las dimensiones

Las dimensiones que se proponen son: Product, Workcenter\_Producto, Time, Factory, Catalog\_Product, Customer, Store, Discount

### 4. Identificar las métricas

El usuario de negocio es quien decide qué es lo que quiere medir. Algunas de las métricas que se tomaran en cuenta son:

- a) Producción de camisas por equipo de producción
- b) Tiempo de producción por workcenter
- c) Cuál es el producto que más se vende
- d) Movimiento de producto por ventas

## Descripción del dataset y diccionario de datos.

Tras un análisis detallado de los flujos de ventas y las interacciones con ERP Odoó, y la plataforma magento se ha seleccionado una serie de tablas transaccionales que constituirán la base para la solución analítica del proceso producción y ventas. Estas tablas, que componen el "Data Set", provienen de un esquema unificado en la base de datos MYSQL y mariadb, serán fundamentales como fuente de datos para alimentar el Data Warehouse. A continuación, se expone un diagrama que ilustra el modelo relacional de este conjunto de datos seleccionado.

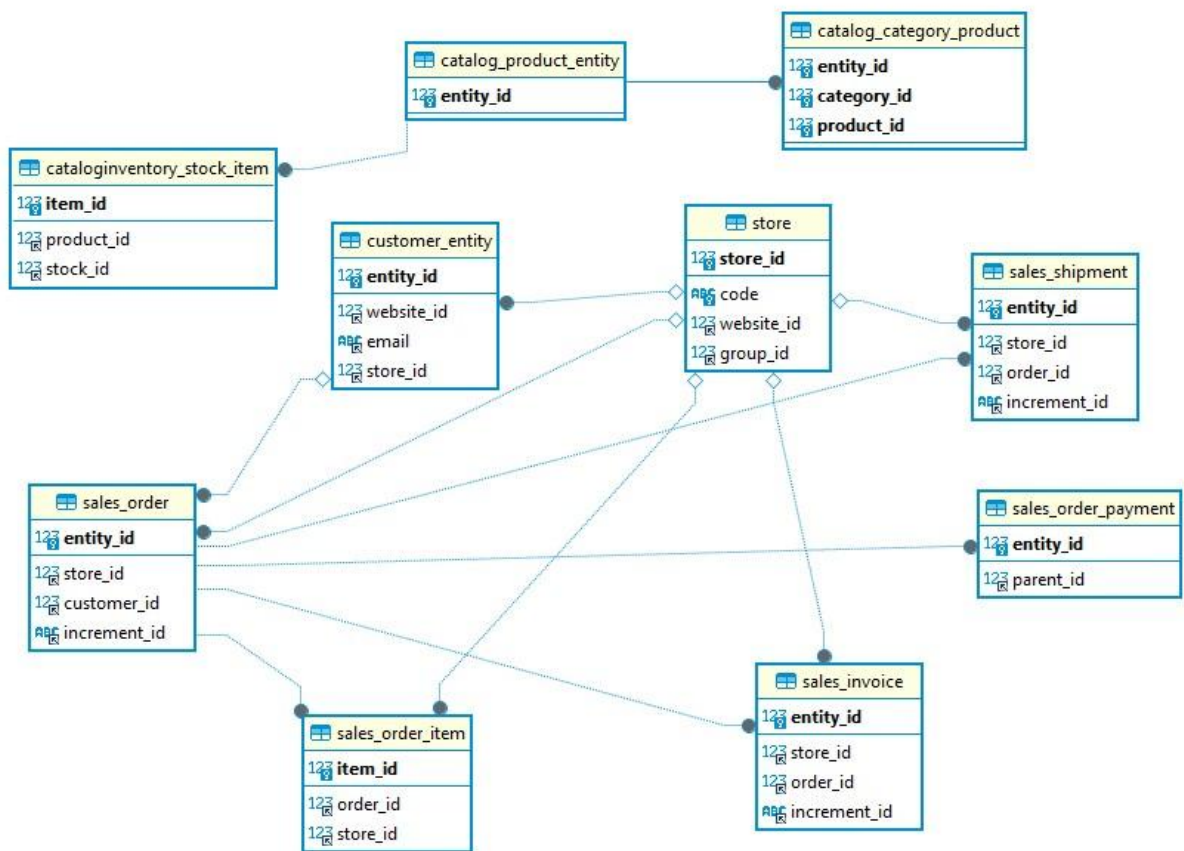


Figura 2. Diagrama de modelo relación para procesos de negocio

## Diccionario de datos

Tabla catalog\_product\_inventory

dbo. catalog_product_inventory			
Características de la tabla			
Nombre de la Tabla	catalog_product_inventory		
Descripción de la Tabla	Almacena información detallada sobre el inventario de los productos en el catálogo.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
inventory_id	INT (10) UNSIGNED	N/A	Identificador único del inventario. Es la clave primaria de la tabla.
product_id	INT(10) UNSIGNED	N/A	Identificador del producto asociado con el inventario.
stock	INT(10) UNSIGNED	N/A	Cantidad total del producto en el inventario.
reserved_stock	INT(10) UNSIGNED	N/A	Cantidad de producto reservada y no disponible para la venta inmediata.
available_stock	INT(10) UNSIGNED	N/A	Cantidad de producto disponible para la venta (stock - reserved_stock).
min_stock_threshold	INT(10) UNSIGNED	N/A	Umbral mínimo de stock para generar alertas de bajo inventario.
max_stock_threshold	INT(10)	N/A	Umbral máximo de

	UNSIGNED		stock permitido.
warehouse_location	VARCHAR(255)	255	Ubicación del producto en el almacén.
batch_number	VARCHAR(100)	100	Número de lote del producto, útil para la trazabilidad.
expiry_date	DATE	N/A	Fecha de caducidad del producto, si aplica.
last_stock_update	DATETIME	N/A	Fecha y hora de la última actualización de stock.
supplier_id	INT(10) UNSIGNED	N/A	Identificador del proveedor del producto.
reorder_level	INT(10) UNSIGNED	N/A	Nivel de stock al cual se debe reordenar el producto.
reorder_quantity	INT (10) UNSIGNED	N/A	Cantidad de producto que se debe reordenar.
safety_stock	INT(10) UNSIGNED	N/A	Cantidad de stock de seguridad para evitar rupturas de stock.
stock_status	ENUM('in_stock', 'out_of_stock', 'backordered')	N/A	Estado del stock del producto.
stock_on_order	INT(10) UNSIGNED	N/A	Cantidad de stock del producto que está en orden y pendiente de recepción.
stock_on_hold	INT(10) UNSIGNED	N/A	Cantidad de stock del producto que

			está en espera por alguna razón (ej. inspección de calidad).
damaged_stock	INT(10) UNSIGNED	N/A	Cantidad de stock del producto que está dañado y no disponible para la venta.
created_at	TIMESTAMP	N/A	Fecha y hora de creación del registro de inventario.
updated_at	TIMESTAMP	N/A	Fecha y hora de la última actualización del registro de inventario.

**Tabla 2.** Estructura del catálogo de productos de inventario

Tabla catalog\_product\_entity

dbo.catalog_product_entity			
Características de la tabla			
Nombre de la Tabla	catalog_product_entity		
Descripción de la Tabla	Almacena la información básica sobre los productos en el catálogo de Magento. Cada registro representa un producto único		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
entity_id	int	N/A	Identificador único del producto (PK).
attribute_set_id	int	10	Identificador del conjunto de atributos asignado al producto.

type_id	nvarchar	32	Tipo de producto (simple, configurable, virtual, etc.).
sku	nvarchar	64	Código SKU (Stock Keeping Unit) único del producto.
has_options	int	5	Indicador de si el producto tiene opciones personalizables
required_options	int	5	Indicador de si el producto tiene opciones obligatorias
created_at	timestamp	N/A	Fecha y hora en que se creó la categoría.
updated_at	timestamp	N/A	Fecha y hora en que se actualizó la categoría por última vez.

**Tabla 3.** Estructura de la entidad de catálogo de producto

Tabla catalog\_category\_product

dbo.catalog_category_product			
Características de la tabla			
Nombre de la Tabla	catalog_category_product		
Descripción de la Tabla	Establece una relación entre las categorías y los productos en el sistema Magento. Cada registro en esta tabla representa la asignación de un producto a una categoría específica.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo

category_id	int	10	Identificador único de la categoría (FK).
product_id	int	10	Identificador único del producto (FK).
position	int	11	Posición del producto dentro de la categoría.

**Tabla 4.** Estructura de la entidad de catálogo de producto

Tabla customer\_entity

dbo.customer_entity			
Características de la tabla			
Nombre de la Tabla	customer_entity		
Descripción de la Tabla	Almacena información principal sobre los clientes en el sistema Magento. Cada registro en esta tabla representa un cliente único, incluyendo detalles básicos como nombre, email, y fechas de creación y actualización.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
entity_id	int	10	Identificador único del cliente (PK).
entity_type_id	int	5	Identificador del tipo de entidad.
attribute_set_id	int	5	Identificador del conjunto de atributos asociado.
website_id	int	5	Identificador del sitio web asociado al cliente.
email	nvarchar	255	Dirección de correo electrónico del cliente.

group_id	int	3	Identificador del grupo de clientes al que pertenece el cliente.
increment_id	nvarchar	255	ID incremental del cliente.
store_id	int	5	Identificador de la tienda asociada al cliente.
created_at	timestamp	N/A	Fecha y hora de creación del registro del cliente.
updated_at	timestamp	N/A	Fecha y hora de la última actualización del registro del cliente.
is_active	int	5	Indicador de si el cliente está activo.
firstname	nvarchar	255	Primer nombre del cliente.
middlename	nvarchar	255	Segundo nombre del cliente.
lastname	nvarchar	255	Apellido del cliente.
prefix	nvarchar	40	Prefijo del nombre del cliente.
suffix	nvarchar	40	Sufijo del nombre del cliente.
dob	date	N/A	Fecha de nacimiento del cliente.
password_hash	nvarchar	255	Hash de la contraseña del cliente.
rp_token	nvarchar	255	Token de restablecimiento de contraseña.

rp_token_created_at	timestamp	N/A	Fecha y hora de creación del token de restablecimiento de contraseña.
default_billing	int	10	Identificador de la dirección de facturación predeterminada del cliente.
default_shipping	int	10	Identificador de la dirección de envío predeterminada del cliente.
taxvat	nvarchar	255	Número de identificación fiscal del cliente.
gender	int	11	Género del cliente.
confirmation	nvarchar	255	Código de confirmación del registro del cliente.
failures_num	int	11	Número de intentos fallidos de inicio de sesión.
first_failure	date	N/A	Fecha y hora del primer intento fallido de inicio de sesión.
lock_expires	date	N/A	N/A

**Tabla 5.** Estructura de la entidad cliente

Tabla inventory\_stock

dbo.inventory_stock	
Características de la tabla	
Nombre de la Tabla	inventory_stock
Descripción de la Tabla	Almacena información sobre los diferentes stocks de inventario en el sistema. Cada registro en esta tabla

	representa un stock único.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
stock_id	int	10	Identificador único del stock (PK).
website_id	int	5	Identificador del sitio web asociado al stock.
stock_name	nvarchar	255	Nombre del stock.

**Tabla 6.** Estructura del inventario

Tabla inventory\_stock\_item

dbo.inventory_stock_item			
Características de la tabla			
Nombre de la Tabla	inventory_stock_item		
Descripción de la Tabla	Se utiliza para almacenar información sobre el inventario de productos. Esta tabla contiene datos que se utilizan para realizar un seguimiento de la cantidad de stock disponible		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
item_id	int	10	Identificador único para cada registro (PK)
product_id	int	10	Identificador del producto al que pertenece el stock.
stock_id	int	10	Identificador del stock al que pertenece el ítem.
qty	int	10	Cantidad de stock disponible para el

			producto.
min_qty	int	10	Cantidad mínima de stock permitida para el producto.
use_config_min_qty	int	5	Indicador de si se debe utilizar la configuración global para la cantidad mínima.
is_qty_decimal	int	5	Indicador de si la cantidad de stock se maneja como un valor decimal.
backorders	int	5	Indicador de si se permiten pedidos en espera para el producto.
use_config_backorders	int	5	Indicador de si se debe utilizar la configuración global para los pedidos en espera.
min_sale_qty	int	13	Cantidad mínima de venta permitida para el producto.
use_config_min_sale_qty	int	5	Indicador de si se debe utilizar la configuración global para la cantidad mínima de venta.
max_sale_qty	int	10	Cantidad máxima de venta permitida para el producto.
use_config_max_sale_qty	int	10	Indicador de si se debe utilizar la configuración global para la cantidad máxima de venta.

is_in_stock	int	5	Indicador de si el producto está en stock.
low_stock_date	int	5	Fecha en la que se alcanzó el nivel bajo de stock.
notify_stock_qty	int	5	Cantidad de stock en la que se debe notificar el bajo inventario.
use_config_notify_stock_qty	int	13	Indicador de si se debe utilizar la configuración global para la notificación de bajo stock.
manage_stock	int	13	Indicador de si se debe gestionar el inventario para el producto.
use_config_manage_stock	int	13	Indicador de si se debe utilizar la configuración global para la gestión del inventario.
stock_status_changed_auto	int	10	Indicador de si el estado del stock cambió automáticamente.
use_config_qty_increments	int	10	Indicador de si se debe utilizar la configuración global para los incrementos de cantidad.
qty_increments	int	5	Incrementos de cantidad permitidos para la venta.
use_config_enable_qty_inc	int	5	Indicador de si se

			debe utilizar la configuración global para habilitar los incrementos de cantidad.
enable_qty_increments	int	5	Indicador de si se permiten incrementos de cantidad para la venta.
is_decimal_divided	int	13	Indicador de si la cantidad de stock se puede dividir en valores decimales.
website_id	int	10	Identificador del sitio web al que pertenece el stock del producto.
custom_attribute	int	5	Campo para almacenar atributos personalizados adicionales para el stock.

**Tabla 7.** Estructura de productos de inventario

Tabla sales\_invoice

dbo.sales_invoice			
Características de la tabla			
Nombre de la Tabla	sales_invoice		
Descripción de la Tabla	Almacena información sobre los diferentes stocks de inventario en el sistema. Cada registro en esta tabla representa un stock único.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo

entity_id	int	10	Identificador único de la factura (PK).
store_id	int	10	Identificador de la tienda donde se realizó la orden.
base_grand_total	float	21	Gran total de la factura en la moneda base.
shipping_tax_amount	float	21	Monto del impuesto de envío en la moneda de la orden.
tax_amount	float	21	Monto total de impuestos en la moneda de la orden.
base_tax_amount	float	21	Monto total de impuestos en la moneda base.
store_to_order_rate	float	21	Tasa de conversión de la moneda de la tienda a la moneda de la orden.
grand_total	float	21	Gran total de la factura en la moneda de la orden.
shipping_amount	float	21	Monto del envío en la moneda de la

			orden.
subtotal_incl_tax	float	21	Subtotal incluido impuestos en la moneda de la orden.
base_subtotal_incl_tax	float	21	Subtotal incluido impuestos en la moneda base.
store_to_base_rate	float	21	Tasa de conversión de la moneda de la tienda a la moneda base.
base_shipping_amount	float	21	Monto del envío en la moneda base.
total_qty	float	21	Cantidad total de productos en la factura.
base_discount_amount	float	21	Monto total del descuento en la moneda base.
base_to_global_rate	float	21	Tasa de conversión de la moneda base a la moneda global.
subtotal	float	21	Subtotal de la factura en la moneda de la orden.
base_shipping_tax_amount	float	21	Monto del impuesto de

			envío en la moneda base.
shipping_discount_amount	float	21	Monto del descuento de envío en la moneda de la orden.
shipping_incl_tax	float	21	Monto de envío incluido impuestos en la moneda de la orden.
base_shipping_incl_tax	float	21	Monto de envío incluido impuestos en la moneda base.
base_total_refunded	float	21	Monto total reembolsado en la moneda base.
discount_description	nvarchar	255	Descripción del descuento.
customer_note	nvarchar	255	Nota del cliente.
customer_note_notify	nvarchar	255	Indica si se notifica al cliente sobre la nota.
discount_amount	float	21	Monto total del descuento en la moneda de la orden.
subtotal_refunded	float	21	Monto subtotal reembolsado en la moneda

			de la orden.
adjustment_positive	float	21	Ajuste positivo de la factura.
base_adjustment_positive	float	21	Ajuste positivo de la factura en la moneda base.
adjustment_negative	float	21	Ajuste negativo de la factura.
base_adjustment_negative	float	21	Ajuste negativo de la factura.
shipping_refunded	float	21	Monto del envío reembolsado en la moneda de la orden.
base_shipping_refunded	float	21	Monto del envío reembolsado en la moneda base.
base_subtotal	float	21	Subtotal de la factura en la moneda base.
tax_refunded	float	21	Monto total de impuestos reembolsados en la moneda de la orden.
base_tax_refunded	float	21	Monto total de impuestos reembolsados en la moneda base.

discount_tax_compensation_amount	float	21	Monto de compensación de impuestos sobre el descuento en la moneda de la orden.
base_discount_tax_compensation_amount	float	21	Monto de compensación de impuestos sobre el descuento en la moneda base.
shipping_discount_tax_compensation_amount	float	21	Monto de compensación de impuestos sobre el descuento de envío en la moneda de la orden.
base_shipping_discount_tax_compensation_amount	float	21	Monto de compensación de impuestos sobre el descuento de envío en la moneda base.
shipping_invoiced	float	21	Monto de envío facturado en la moneda de la orden.
base_shipping_invoiced	float	21	Monto de envío

			facturado en la moneda base.
discount_invoice	float	21	Monto total del descuento facturado en la moneda de la orden.
base_discount_amount_refunded	float	21	Monto total del descuento reembolsado en la moneda base.
base_discount_invoiced	float	21	Monto total del descuento facturado en la moneda base.
qty	float	21	Cantidad de productos en la factura.
base_to_global_adjustment	float	21	Ajuste global en la mon

**Tabla 8.** Estructura ventas

Tabla sales\_order

dbo.sales_order			
Características de la tabla			
Nombre de la Tabla	sales_order		
Descripción de la Tabla	Almacena información sobre las órdenes de venta generadas en la tienda. Cada registro en esta tabla representa una orden de venta única.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo

entity_id	int	10	Identificador único de la orden (PK).
state	nvarchar	32	Estado de la orden (nuevo, procesando, completado).
status	nvarchar	255	Estado actual de la orden.
coupon_code	nvarchar	255	Código de cupón utilizado en la orden, si aplica.
protect_code	nvarchar	255	Código de protección para la orden.
shipping_description	nvarchar	255	Descripción del método de envío utilizado.
is_virtual	boolean	1	Indica si la orden es virtual.
store_id	int	5	Identificador de la tienda donde se realizó la orden.
customer_id	int	10	Identificador del cliente que realizó la orden.
base_discount_amount	float	13	Monto del descuento

			en la moneda base.
base_grand_total	float	13	Gran total en la moneda base.
base_shipping_amount	float	13	Monto del envío en la moneda base.
base_subtotal	float	13	Subtotal en la moneda base.
base_tax_amount	float	13	Monto de impuestos en la moneda base.
base_to_global_rate	float	13	Tasa de conversión de la moneda base a la moneda global.
base_to_order_rate	float	13	Tasa de conversión de la moneda base a la moneda de la orden.
discount_amount	float	13	Monto del descuento.
grand_total	float	13	Gran total en la moneda de la orden.
shipping_amount	float	13	Monto del

			envío en la moneda de la orden.
subtotal	float	13	Subtotal en la moneda de la orden.
tax_amount	float	13	Monto de impuestos en la moneda de la orden.
total_qty_ordered	float	13	Cantidad total de productos ordenados.
customer_email	nvarchar	255	Correo electrónico del cliente.
customer_firstname	nvarchar	255	Nombre del cliente.
customer_lastname	nvarchar	255	Apellido del cliente.
increment_id	nvarchar	255	ID incremental de la orden.
applied_rule_ids	nvarchar	3	IDs de las reglas de precios aplicadas a la orden.
base_currency_code	nvarchar	3	Código de la moneda base utilizada en la orden.
global_currency_code	nvarchar	3	Código de la moneda global

			utilizada en la orden.
order_currency_code	nvarchar	3	Código de la moneda de la orden.
store_currency_code	nvarchar	3	Código de la moneda de la tienda utilizada en la orden.
created_at	timestamp	N/A	Fecha y hora en que se creó la orden.
updated_at	timestamp	M/A	Fecha y hora de la última actualización de la orden.
total_item_count	int	10	Cantidad total de ítems en la orden.
customer_group_id	int	5	Identificador del grupo de clientes al que pertenece el cliente.
shipping_incl_tax	float	13	Monto del envío incluyendo impuestos en la moneda de la orden.
base_shipping_incl_tax	float	13	Monto del envío incluyendo impuestos

			en la moneda base.
discount_invoiced	float	13	Monto del descuento facturado en la moneda de la orden.
base_discount_invoiced	float	13	Monto del descuento facturado en la moneda base.
discount_refunded	float	13	Monto del descuento reembolsado en la moneda de la orden.
base_discount_refunded	float	13	Monto del descuento reembolsado en la moneda base.
subtotal_incl_tax	float	13	Subtotal incluyendo impuestos en la moneda de la orden.
base_subtotal_incl_tax	float	13	Subtotal incluyendo impuestos.
tax_invoiced	float	13	Monto de impuestos facturado en la moneda de la orden.

base_tax_invoiced	float	13	Monto de impuestos facturado en la moneda base.
tax_refunded	float	13	Monto de impuestos reembolsado en la moneda de la orden.
base_tax_refunded	float	13	Monto de impuestos reembolsado en la moneda base.
shipping_tax_amount	float	13	Monto del impuesto del envío en la moneda de la orden.
base_shipping_tax_amount	float	13	Monto del impuesto del envío en la moneda base.
weight	float	13	Peso total de los productos en la orden.

**Tabla 9.** Estructura de la orden de ventas

Tabla store

dbo.store	
Características de la tabla	
Nombre de la Tabla	store
Descripción de la Tabla	Almacena información sobre las tiendas dentro del

	sistema de comercio electrónico. Cada registro en esta tabla representa una tienda específica que puede estar asociada con un grupo de tiendas.		
Campos de la tabla			
Nombre del campo	Tipo de dato	Longitud del campo	Descripción del campo
store_id	int	N/A	Identificador único de la tienda (PK).
code	nvarchar	32	Código único de la tienda.
website_id	int	5	Identificador del sitio web al que pertenece la tienda.
group_id	int	5	Identificador del grupo de tiendas al que pertenece la tienda.
name	nvarchar	32	Nombre de la tienda.
sort_order	int	5	Orden de clasificación de la tienda.
is_active	boolean	1	Indicador de si la tienda está activa (1) o inactiva (0).

**Tabla 10.** Estructura de tienda

Tabla mrp\_production

dbo.mrp_production			
Características de la tabla			
Nombre de la Tabla	mrp_production		
descripción de la Tabla	Sirve para gestionar las ordenes de fabricación		
Campos de la tabla			
NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD DEL CAMPO	BREVE DESCRIPCION DEL CAMPO
Id	Int		Identificador único de la orden de

			fabricación
name	varchar	255	Nombre de la orden de fabricación
product_id	Int	4	Referencia al producto que se va a fabricar (product.product)
product_uom_id	Int	4	Unidad de medida del producto que se va a fabricar
lot_producing_id	Int	4	Identificador del lote que se está produciendo
picking_type_id	Int	4	Tipo de operación de almacén asociado a la orden de fabricación(stock.picking.type)
bom_id	Int	4	Identificador de la BOM
user_id	Int	4	Identificador del usuario que gestiona la orden de fabricación
company_id	Int	4	Identificador de la compañía a la cual pertenece la orden de fabricación
procurement_group_id	Int	4	Grupo de aprovisionamiento asociado a la orden de producción, gestiona la adquisición de componentes necesarios para la producción
date_planned_start	timestamp	N/A	Fecha planificada de inicio para la producción del producto.
date_planned_finished	timestamp	N/A	Fecha planificada de finalización para la producción del producto.
state	varchar	255	Estado de la orden (draft, confirmed, progress, done)
origin	varchar	255	Origen de la orden de fabricación, por ejemplo la orden de venta que generó esta orden de fabricación
priority	VarChar	255	Prioridad de la orden de fabricación

product_qty	Int	10	Cantidad de producto a fabricar
date_start	timestamp	N/A	Fecha que se inicia la fabricación del producto
date_finished	timestamp	N/A	Fecha que se termina la fabricación del producto

**Tabla 11.** Estructura de la tabla de producción

Tabla mrp\_workcenter

dbo.mrp_workcenter			
Características de la tabla			
Nombre de la Tabla	mrp_workcenter		
descripción de la Tabla	Esta tabla sirve para gestionar los centros de trabajo dentro del contexto del módulo de fabricación		
campos de la tabla			
NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD DEL CAMPO	BREVE DESCRIPCION DEL CAMPO
Id	Int		Identificador único del centro de trabajo
name	varchar	255	Nombre del centro de trabajo, campo obligatorio
time_efficiency	float	8	Eficiencia del centro de trabajo reflejada en porcentaje
cost_hour	float	8	Costo por hora de utilización de instalaciones del centro
company_id	Int	4	Hace referencia a la compañía que pertenece el centro de trabajo
active	bool	N/A	Campo de tipo booleano que indica si el centro está activo o no.

**Tabla 12.** Estructura de los centros de producción.

Tabla stock\_move

dbo. stock_move			
Características de la tabla			
Nombre de la Tabla	stock_move		
descripción de la Tabla	Registra las operaciones de inventario como recepciones, entregas y transferencias internas.		
Campos de la tabla			
NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD DEL CAMPO	BREVE DESCRIPCION DEL CAMPO
Id	Int		Identificador único del movimiento
name	string	255	Descripción del movimiento
product_id	Int	N/A	Producto relacionado con el movimiento. (Many2one, product.product)
product_uom_qty	Int	N/A	Cantidad de producto a mover
product_uom	Int	N/A	Unidad de medida del producto. (Many2one, product.uom)
location_id	Int	N/A	Ubicación de origen. (Many2one, stock.location)
location_dest_id	Int	N/A	Ubicación de destino. (Many2one, stock.location)
picking_id	Int	N/A	Referencia al picking asociado. (Many2one,

			stock.picking)
state	String	255	Estado del movimiento (draft, done, etc.). (Selection, opciones: 'draft', 'done', etc.)

**Tabla 13.** Estructura de movimientos de inventario

Tabla producto\_product

dbo. product_product			
Características de la tabla			
Nombre de la Tabla	product_product		
descripción de la Tabla	Nombre del almacén.		
campos de la tabla			
NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD DEL CAMPO	BREVE DESCRIPCION DEL CAMPO
Id	Int		Identificador único del producto
product_tmpl_id	Int	N/A	Referencia a la plantilla de producto. (Many2one, product.template)
default_code	string	64	Código interno del producto
barcode	string	128	Código de barras del producto
active	Boolean	N/A	Si la ubicación está activa o no

**Tabla 14.** Estructura de producto item

Tabla producto\_category

dbo. product_category			
Características de la tabla			

Nombre de la Tabla	product_category		
descripción de la Tabla	Categorías de productos, útiles para organizar y segmentar productos en análisis.		
Campos de la tabla			
NOMBRE DEL CAMPO	TIPO DE DATO	LONGITUD DEL CAMPO	BREVE DESCRIPCION DEL CAMPO
Id	Int		Identificador único de la categoría de producto
name	string	255	Nombre de la categoría
parent_id	Int	N/A	Categoría padre
complete_name	string	255	Nombre completo de la categoría (incluyendo jerarquía)
property_valuation	Strig	N/A	Método de valoración (manual, automated). (Selection, opciones: 'manual', 'automated')
property_cost_method	String	N/A	Método de coste (standard, real, average). (Selection, opciones: 'standard', 'real', 'average')

**Tabla 13.** Estructura de categoría de productos

## Resultado del Data Profiling

Se ha filtrado solo aquellas columnas de las tablas que juegan un papel importante tales como: ventas, inventario, producción.

Sales\_order

	Row count	Null count	Blank count	Entirely unique count	Entropy lower case count	Total char count	Max chars	Min chars	Avg chars	Max whitespace	Min whitespace	Avg whitespace	Uppercase chars	Lowercase chars	Digit chars	Delimiter chars	Non-letter chars	Word count	Max words	Min words	
entity_id	100	0	0	0	0	53	3	1	34	5	2	7	0	0	0	91	0	91	50	1	1
state	100	0	0	0	0	134	23	35	11	4	6	4	50	0	400	91	0	141	100	2	2
status	100	0	0	10	0	187	62	4	11	13	9	8	250	200	0	91	0	141	50	1	1
coupon_code	100	0	0	24	0	45	24	7	2	3	16	11	20	65	0	0	0	0	50	1	1

protect_code	100	0	0	0	0	75	1	4	8	7	4	7	0	0	0	50	0	50	50	1	1
shipping_description	100	0	0	0	0	89	1	6	1	15	14	0	0	0	0	50	0	50	50	1	1
is_virtual	100	0	0	0	0	47	7	19	7	11	9	0	0	0	0	190	0	240	50	1	1
store_id	100	0	0	0	0	210	2	5	43	9	8	0	0	0	0	100	0	100	50	1	1

**Tabla 14.** Perfil de información de orden de venta

Sales\_shipment

	Row count	Number of links	Block count	Entirely separate count	Entirely separate count	Total character count	Maximum characters	Minimum characters	Average characters	Maximum white spaces	Minimum white spaces	Average white spaces	Upper case characters	Upper case characters (excluding first letters)	Lower case characters	Digit characters	Digit critical characters	Non-letter characters	Word count	Maximum words	Minimum words
entity_id	10	0	0	76	0	0	6	4	11	1	9	8	25	20	0	91	0	14	0	1	1

	0					2			3			0	0			1					
store_id	10000	0	0	0	0	74	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
total_weight	1000	0	0	21	0	12	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
total_qty	1000	0	0	0	0	0	2	7	2	3	1	1	20	15	0	0	0	0	0	1	1
email_sent	1000	0	0	46	0	0	1	4	8	7	4	7	0	0	0	50	0	50	0	1	1
order_id	1000	0	0	20	0	0	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
customer_id	1000	0	0	0	0	0	7	1	7	1	9	0	0	0	0	19	0	24	0	1	1
shipment_status	1000	0	0	76	0	0	6	4	11	1	9	8	25	20	0	91	0	14	0	1	1
increment_id	1000	0	0	0	0	74	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1

**Tabla 15.** Perfil de información de envío de la venta

sales\_invoice

Row count	Blank count	Entirely uppercase count	Entirely lowercase count	Total character count	Max characters	Min characters	Avg characters	Max whitespace	Min whitespace	Avg whitespace	Uppercase characters	Uppercase characters (excl .)	Lowercase characters	Digit characters	Digit characters	Non-letter characters	Word count	Max words	Min words
-----------	-------------	--------------------------	--------------------------	-----------------------	----------------	----------------	----------------	----------------	----------------	----------------	----------------------	-------------------------------	----------------------	------------------	------------------	-----------------------	------------	-----------	-----------

					e c o u n t									firs t lett er s)								
entity_id	10 0	0	0	0	0	50	1	4	8	7	4	7	0	0	0	50	0	50	50	1	1	
store_id	10 0	0	0	21	0	12	1	6	1	1 5	1 4	0	0	0	0	50	0	50	50	1	1	
base_grand_total	10 0	0	0	4	0	31	7	1 9	7	1 1	9 0	0	0	0	0	19 0	0	24 0	50	1	1	
shipping_tax_amount	10 0	0	0	1	0	0	6 2	4	11	1 3	9 8	21	87	0	91	0	14 1	50	1	1		
tax_amount	10 0	0	0	7	0	0	2 4	7	2	3	1 6	1 1	20 0	15 0	0	0	0	0	50	1	1	
base_tax_amount	10 0	0	0	0	0	0	1	4	8	7	4	7	0	0	0	50	0	50	50	1	1	
store_to_order_rate	10 0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	50	0	50	0	1	1	
grand_total	10 0	0	0	0	0	56	1	1	1	0	0	0	0	0	0	50	0	50	0	1	1	
shipping_amount	10 0	0	0	0	0	43	1	1	1	0	0	0	0	0	0	50	0	50	0	1	1	

**Tabla 16.** Perfil de información de venta

customer\_entity

	Row count	N	B	Entirely uppercase count	Entirely lowercase count	Total character count	Max characters	Min characters	Avg characters	Max whitespace	Min whitespace	Avg whitespace	Uppercase characters	Uppercase characters (excluding first letters)	Lowercase characters	Digit characters	Digit characters	Non-letter characters	Word count	Max words	Min words
entity_id	100	0	0	0	0	0	24	7	2	3	16	11	20	15	0	0	0	0	0	1	1
entity_type_id	100	0	0	5	0	0	4	4	4	0	0	0	20	15	0	0	0	0	50	1	1
attribute_set_id	100	0	0	10	0	0	1	1	1	0	0	0	0	0	0	50	0	50	50	1	1
website_id	100	0	0	1	0	0	1	1	1	0	0	0	0	0	0	50	0	50	0	1	1
email	100	0	0	76	0	0	62	4	11	13	9	8	25	20	0	91	0	14	0	1	1
group_id	100	0	0	0	0	74	1	6	1	15	14	0	0	0	0	50	0	50	0	1	1
increment_id	100	0	0	21	0	12	1	6	1	15	14	0	0	0	0	50	0	50	0	1	1
store_id	100	0	0	0	0	0	24	7	2	3	16	11	20	15	0	0	0	0	0	1	1
created_at	100	0	0	0	0	0	24	7	2	3	16	11	20	15	0	0	0	0	0	1	1

updated_at	100	0	0	21	0	12	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
										5	4										

**Tabla 17.** Perfil de información de la entidad usuario

Inventory\_stock\_item

	Row count	Number of links	Entity relationships count	Entity categories count	Total characters	Maximum characters	Minimum characters	Average characters	Maximum white spaces	Minimum white spaces	Average white spaces	Uppercase characters	Uppercase characters (excluding first letters)	Lowercase characters	Digits	Alphabetic characters	Non-letter characters	Word count	Maximum words	Minimum words
item_id	100	0	0	20	0	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
producto_id	100	0	0	0	0	7	1	7	1	9	0	0	0	0	19	0	24	0	1	1
stock_id	100	0	0	0	74	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1

qty	100	0	0	21	0	12	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1
min_qty	100	0	0	0	0	0	2	7	2	3	1	1	20	15	0	0	0	0	0	1	1
use_config_min_qty	100	0	0	0	0	0	2	7	2	3	1	1	20	15	0	0	0	0	0	1	1
is_qty_decimal	100	0	0	0	0	74	1	6	1	1	1	0	0	0	0	50	0	50	0	1	1

**Tabla 18.** Perfil de información de inventario disponible

## **Especificación de las necesidades analíticas**

Para el análisis de ventas de la empresa Topazio se realizará un estudio de mercado trimestral para visualizar los productos más vendidos, así como en donde se vende más un producto y en qué momento de la semana es que más se vende ese producto para poder tomar mejores decisiones y así aumentar las ventas y crecimiento de la empresa. Con el fin de analizar qué productos se venden más y mantiene el negocio en crecimiento y estabilidad financiera, se tomarán datos de los productos vendidos y cuales generan mayor ingreso para la empresa. Para evitar la acumulación de información en inventario de productos se utilizan prácticas de aplicación de descuentos estos con el fin de mejorar la salida y entrada de productos ya que algunos productos no se venden tanto como otros, es necesario tener estrategias para poder actuar ante estos escenarios, para poder realizar estos descuentos se necesita saber la información del volumen de ventas y salidas de los productos. Los productos no siempre se venden de la misma cantidad en todos los lugares, por ello, se ha detectado la necesidad de saber que producto se vende más según tienda ubicada, sabiendo en que tienda se vende más se pueden tomar mejores decisiones.

Para el análisis de producción se tomarán indicadores y procesos dentro de la empresa se necesitan saber cuántas camisas se producen por hora y por equipo de producción para ello se maneja la información por medio de lotes de producción para saber cuándo se ha producido en un determinado tiempo.

### **Modelo dimensional propuesto**

#### 1. Procesos de negocios

El proceso de negocio definido para la extracción y transformación de datos de este modelo es el proceso de "Gestión de Inventarios".

#### 2. Nivel de Granularidad

·¿Qué detalle requiere el usuario del negocio?

Quiero ver los movimientos por ubicación, el tipo de movimiento que fue y la razón del movimiento de ser posible.

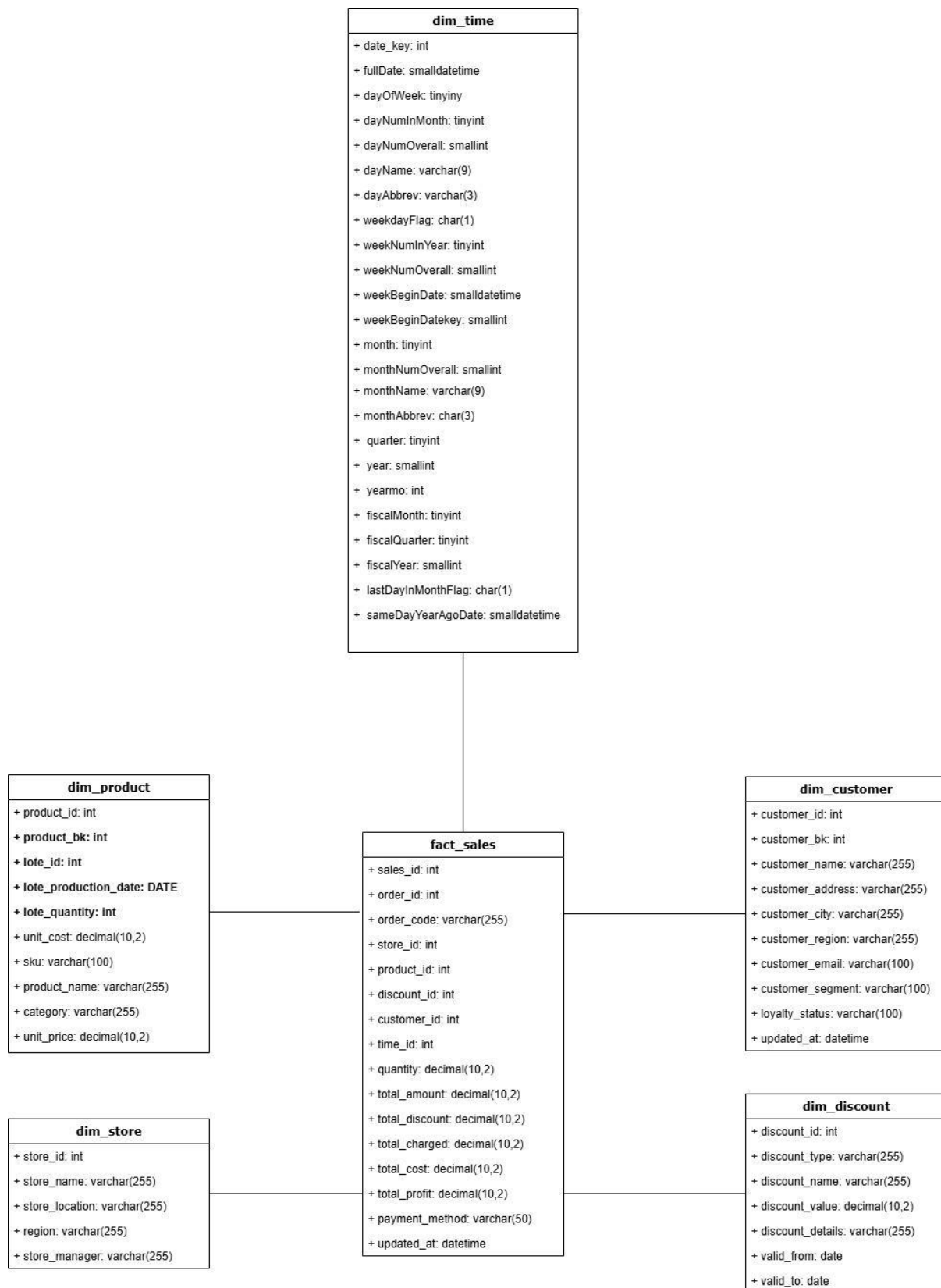
·¿Qué detalle es efectivamente posible con los datos?

Se pueden ver los movimientos ocurridos de inventario de acuerdo al tipo de inventario, la ubicación y detalles del movimiento.

#### 3. Identificación de dimensiones

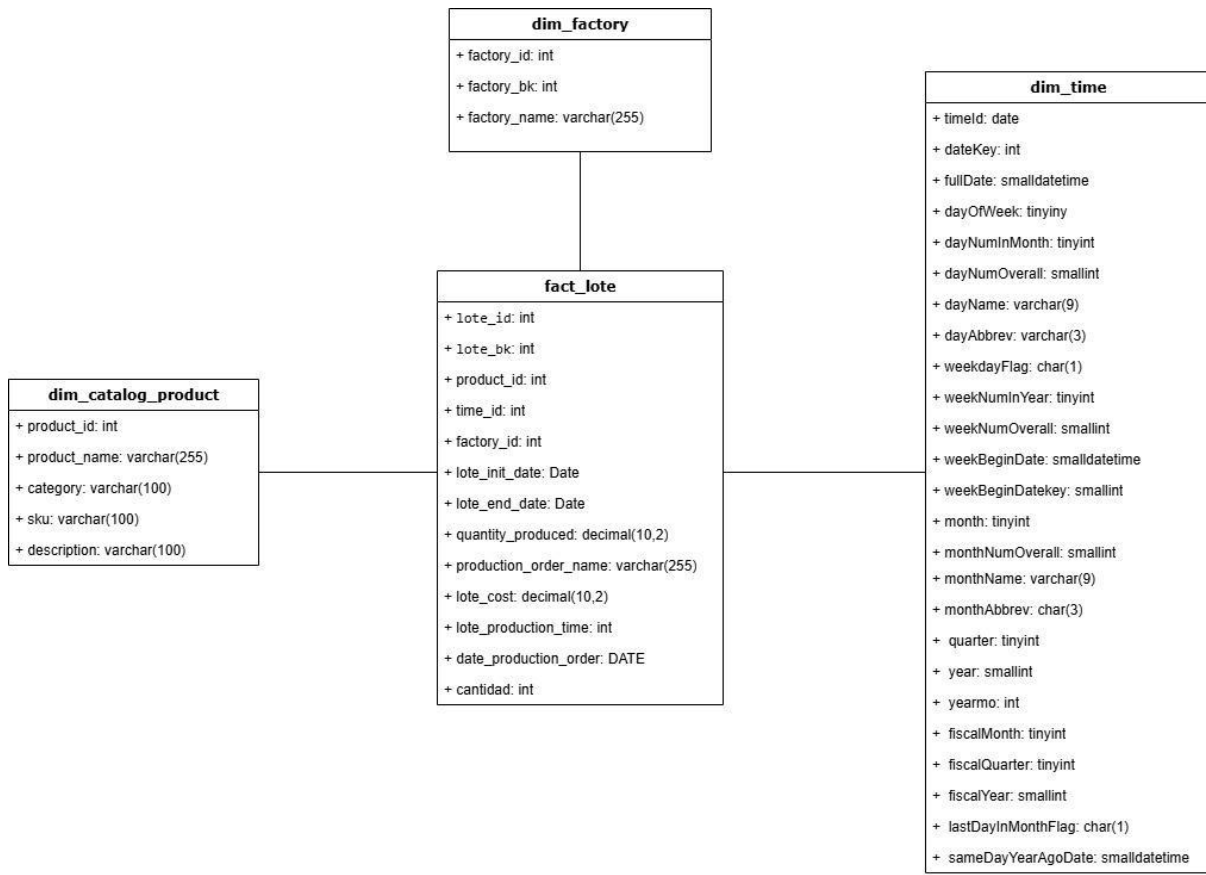
- a) Product
- b) Date
- c) Location

## Diagrama propuesto para el modelo dimensional de ventas



**Figura 3.** Diagrama dimensional de ventas

## Diagrama propuesto para el modelo dimensional de producción por lotes



**Figura 4.** Diagrama dimensional de producción por lotes

Diagrama propuesto para el modelo dimensional de movimientos de inventario de producto no terminado

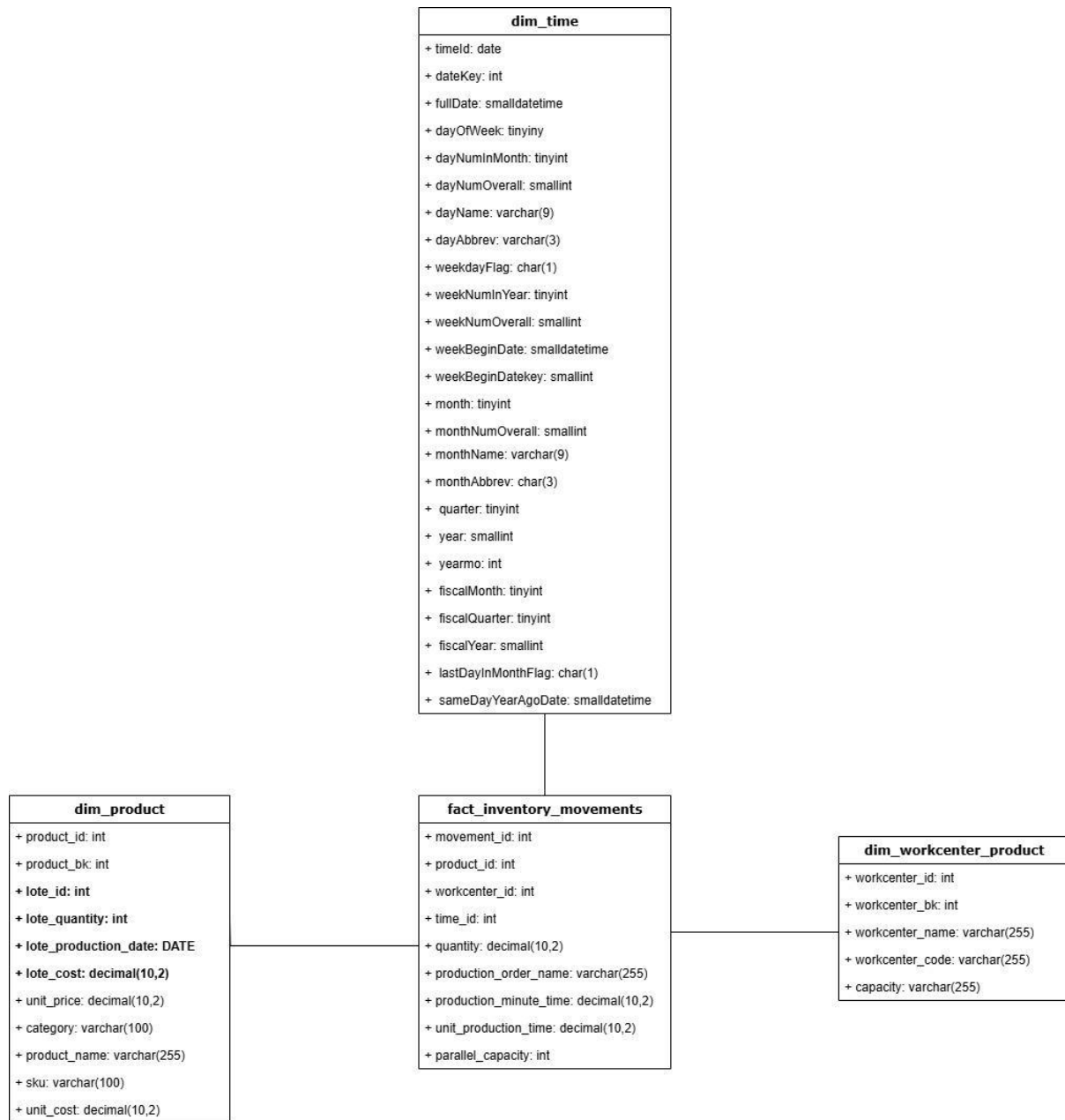


Figura 5. Diagrama dimensional de movimientos de inventario

## Mapping de modelo dimensional y tablas del dataset

### Dim Location

Table Name		Target											Source						
Column Name	Display Name	Description	Attribute Group	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Example Values	SCD Type	Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
location_id	Location Id	Llave primaria surrogada	Identifiers	int			PK		no		1, 2, 3...	Derived							
location_bk	Location Bk	Llave del negocio de store e workstore		int					no		1, 2, 3...	Production	odoo, magento					store.id    mrp.workcenter.id	Se obtiene el id en store o de la id.
location_name	Location Name	Nombre de la ubicación		varchar	255				no		Tienda 1	Production, Sales	odoo, magento					store.name    mrp.workcenter.name	
location_type	Location Type	Tipo de ubicación		varchar	50				no		Tienda							Tienda    Fabrica	
region	Region	Región de la ubicación		varchar	50				no		Centro	Production, Sales						store.region    'N/A'	

Figura 6. Mapping de dimensión Location

### Dim Product

Table Name		Target											Source						
Column Name	Display Name	Description	Attribute Group	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Example Values	SCD Type	Source System	Source Schema	Source Table	Source Field Name	Source Datatype	ETL Rules	Comments
product_id	Product Id	Llave primaria surrogada	Identifiers	int			PK		no		1, 2, 3...	Derived							
product_bk	Product Bk	Llave del negocio del catalogo de productos		int					no		1, 2, 3...	Production	odoo		product_template	id	int		
lote_id	Lote Id	Identificador del lote		int					no		1, 2, 3...	Production	odoo		mrp_production	id	int		
lote_production_date	Lote Production Date	Fecha de Producción de Lote de producto.		date					no		11/7/2024	Production	odoo		mrp_production	date_finished	timestamp		
lote_quantity	Lote Quantity	Cantidad producida de lote de producto.		int					no		1, 2, 3...	Production	odoo		mrp_production	product_qty	int		
unit_cost	Unit Cost	Costo unitario por producto.		decimal		10	2		no		5.57	Production	odoo					(stock_move_line.product_uom_qty * product_product.standard_price) / lote_quantity	
sku	SKU	codigo SKU del producto		varchar		100			no		SKU001	Sales	magento		catalog_product	sku	nvarchar		
product_name	Product Name	Nombre del producto.		varchar		255			no		Camisa Blanca	Production	odoo		product_template	name	varchar		
category	Category	Categoría del producto		varchar		100			no		Camisas de vestir	Production	odoo		product_category	name	varchar		
unit_price	Unit Price	Precio unitario del lote de producto.		decimal		10	2		no		12.99	Production	odoo		product_template	list_price	decimal		

Figura 7. Mapping de dimensión Product

## Dim Customer

TOPAZIO -- Datawarehouse																	DimCustomer	Example data model information wi
Table Name		dim_customer																
Table Type		Dimension																
View Name		dim_customer																
Display Name		Dim Customer																
Description		dimernsión de clientes registrados																
Used in schemas		dwh_topazio_g07																
Target											Source							
Column Name	Display Name	Description	Attribute Group	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Example Values	SCD Type	Source System	Source Scher	Source Table	Source Field Name	Source Datatype	ETL Rules
customer_id	Customer Id	llave primaria surrogada	identifiers	int			PK		no		1, 2, 3 ...						int	
customer_bk	Customer Bk	Llave del negocio.	identifiers	int					no		1, 2, 3 ...		Sales	magento	customer_entity	entity_id	int	
customer_name	Customer Name	Nombre del cliente		varchar	255				no		Salvador Hernandez		Sales	magento	customer_entity		varchar	first_name + ' ' + last_name
customer_address	Customer Address	Dirección del cliente		varchar	255				no		Soyapango, Sa...		Sales	magento	customer_address_entity			customer_address_entity.city + ' ' + customer_address_entity.region + ' ' + customer_address_entity.country
customer_city	Customer City	Ciudad de residencia del cliente		varchar	255				no		Soyapango		Sales	magento	customer_addr	city	varchar	
customer_region	Customer Region	Region de residencia del cliente		varchar	255				no		Centro		Sales	magento	customer_addr	region	varchar	
customer_email	Customer Email	Email del cliente		varchar	255				no		salvador@correo.c		Sales	magento	customer_entity	email	varchar	
customer_segment	Customer Segment	Segmento de clientes al que pertenece el cliente		varchar	255				no	Bajo Valor	Bajo Valor		Sales	magento				total_purchases_revenue: total de ventas en las tiendas. Rango de percentiles: 81 >= & <= 100: Alto valor 41 >= & <= 80: Medio valor < 80: Bajo Valor
loyalty_status	Loyalty Status	Evalua la actividad del cliente en el ultimo mes.		varchar	255				no	Activo	Activo		Sales	magento				Activo: cliente que ha realizado una compra en el ultimo mes.

Figura 8. Mapping de dimensión Customer

## Dim Store

TOPAZIO -- Datawarehouse																	DimStore	
Table Name		dim_store																
Table Type		Dimension																
View Name		dim_store																
Display Name		Dim Store																
Description		Dimension de tiendas de topazio.																
Used in schemas		dwh_topazio_g07																
Target											Source							
Column Name	Display Name	Description	Attribute Group	Datatype	Size	Precision	Key?	FK To	NULL?	Default Value	Example Values	SCD Type	Source System	Source Scher	Source Table	Source Field Name	Source Datatype	ETL Rules
store_id	Store Id	Llave primaria surrogada	Identifiers	int			PK		no		1, 2, 3...		Derived					
store_bk	Store Bk	Llave del negocio de tiendas.	Identifiers	int					no		1, 2, 3...		Sales	magento	store	id	int	
store_name	Store Name	Nombre de la tienda		varchar	255				no		Tienda Metropolitana		Sales	magento	store	name	varchar	
region	Region	Region de la tienda		varchar	255				no		Centro		Sales	magento	store	region	varchar	
store_manager	Store Manager	Manager de la tienda		varchar	100				no		Victor Manuel Vicente	1	Sales	magento	store	manager	varchar	

Figura 9. Mapping de dimensión Store

### **c) Descripción de la tecnología a utilizar**

Se han utilizado diferentes herramientas tecnológicas para el desarrollo de la solución. Aquí se describen las herramientas utilizadas desde el sistema transaccional hasta la presentación de los datos que es el producto final donde el usuario de negocio interactúa con la información solicitada.

#### **MariaDB**

Es un sistema de gestión de base de datos relacionales de código abierto con licencia GPLv2, derivado de MySQL. Es uno de los servidores de bases de datos más populares del mundo. Ver anexo

#### **Microsoft Visual Studio**

Es un entorno de desarrollo integrado (IDE) que permite a los desarrolladores crear aplicaciones, sitios web y servicios web. Es compatible con múltiples lenguajes de programación. Ver Anexo

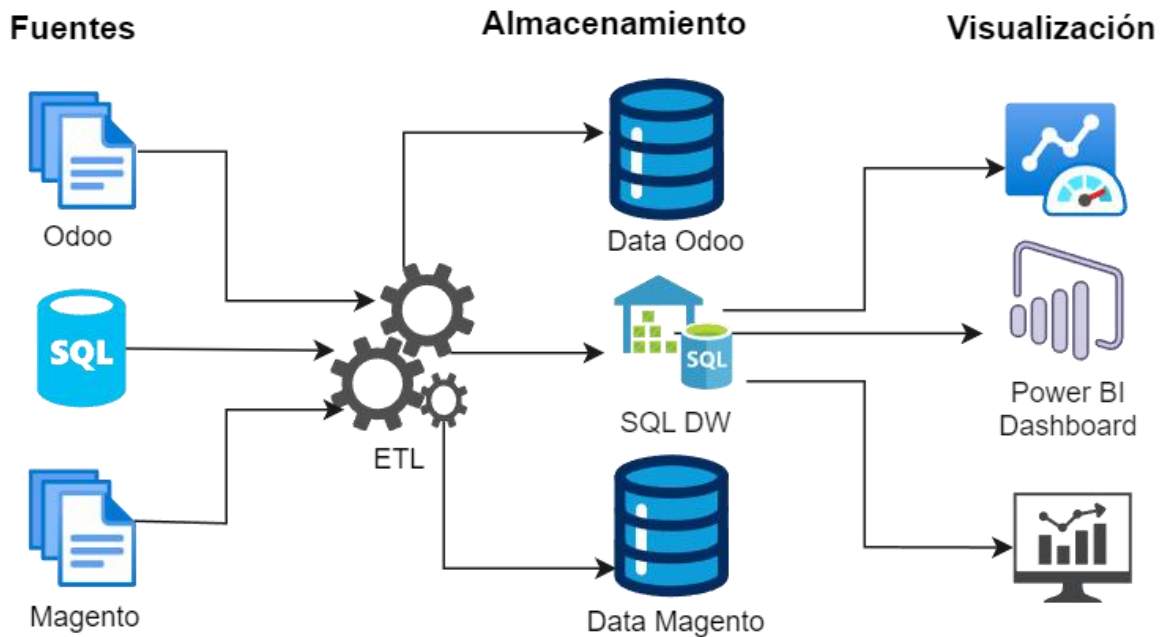
#### **SQL Server Integration Services (SSIS)**

Es una plataforma de Microsoft que permite crear soluciones para la integración y transformación de datos a nivel empresarial. SSIS es una herramienta ETL (Extract, Transform, Load) que permite crear, programar y administrar flujos de trabajo de integración de datos. Se utiliza para resolver problemas empresariales complejos, como copiar o descargar archivos, o cargar almacenes de datos.

#### **Power BI**

Es un servicio gratuito de análisis de negocio basado en la nube y visualización de datos, de negocio. Esta herramienta de Business Intelligence (BI), incorporada en la suite de productividad Microsoft Office 365, permite crear dashboard, informes o reportes interactivos con Power BI Desktop y acceder a los datos en cualquier lugar con las aplicaciones nativas en su versión móvil.

#### d) Diagrama arquitectónico de la solución



**Figura 10.** Diagrama arquitectónico de los procesos operativos

#### e) Descripción de cada componente de la solución

Fuentes de datos: las bases de datos de Odoo y Magento están pobladas con los datos que registran los flujos de tareas y resultados de los procesos de producción y venta. La base de datos de Odoo en PostgreSQL y la base de datos de Magento en MySQL son nuestros orígenes de datos. De la base de datos de Odoo extraeremos información de productos, variantes, precios, costos, ordenes de producción, ordenes de trabajo y todo lo relacionado a inventario. De la base de Magento podemos obtener la información de ventas, precios y lugar.

##### **Procesos de extracción, transformación y carga de datos.**

Como parte de estos procesos es encargarse de extraer todos los datos necesarios de las fuentes de datos a través de la herramienta que ofrece Visual Studio en un proyecto de Integration Services (SSIS). Conociendo de donde y como extraer los datos se procede a realizar las adecuaciones, modificaciones, conversiones y todos los tratamientos necesarios para asegurar legibilidad de los datos. Finalmente, luego de haber diseñado las estrategias de extracción y transformación de datos estos son cargados a la base de datos del Data Warehouse.

##### **Almacenamiento.**

El almacenamiento será en una base de datos de SQL Server, donde se crearán las dimensiones y tablas de hechos que almacenarán los datos resultantes de los procesos ETL. Para este proyecto se obtuvieron tres modelos dimensionales que almacenaran los datos los cuales son: modelo dimensional de ventas, modelo dimensional de producción por lotes y el modelo dimensional de movimientos de inventario.

## **Visualización.**

El diagrama arquitectónico de la solución cuenta con el componente de visualización para lo cual se utilizará la herramienta de Power BI. Se crearán tableros (dashboards) para presentar la información en formas mucho más amigables para los usuarios que necesitan consultar la información de manera rápida. Esto se hará colocando en lugares muy accesible los KPI, diferentes tipos de graficas de datos de ventas y producción, como el producto más vendido, las categorías más populares, cantidades vendidas y además tablas con datos informativos.

## **Capitulo III: Estrategia de implementación de propuesta de solución**

### **a) Estrategia de implementación**

La propuesta de solución diseñada para solventar la solicitud de la empresa Topazio sobre la necesidad del conocimiento de los indicadores de la producción de camisas, representa un producto compuesto, ya que esta solución trabaja en conjunto con Odo y Magento, como su fuente principal de recolección de datos, esto mediante el envío de data directamente de sus bases de datos. La solución presentada se compone de una serie de módulos que trabajan de la mano para lograr la extracción, transformación, carga y presentación de datos.

### **Sistema transaccional Odo y Magento**

Odo: Es un software de gestión empresarial básicamente todo en uno, es decir, ayuda a organizar y automatizar procesos clave, es por eso que Topazio lo utiliza principalmente para gestionar todo lo referente al proceso de producción, desde el abastecimiento de insumos hasta el movimiento de producto a las tiendas físicas.

Base de Datos PostgreSQL: Almacena toda la información referente a la fabricación de productos, inventarios, proveedores e insumos. El diseño de la base de datos de Odo está diseñada para almacenar un nivel de detalle alto, permitiendo no solo almacenar la información dicha anteriormente, sino también las fechas de inicio y fin de cada orden de fabricación. Por consiguiente, la gran mayoría de las tablas de esta base de datos son utilizadas para guardar información critica que refiere a todo el proceso de fabricación de camisas.

Magento: Topazio utiliza esta herramienta para la venta de productos, Topazio utiliza las herramientas avanzadas de magento para personalizar la experiencia de compra, gestionar productos, pagos, envíos y optimizar el rendimiento de ventas.

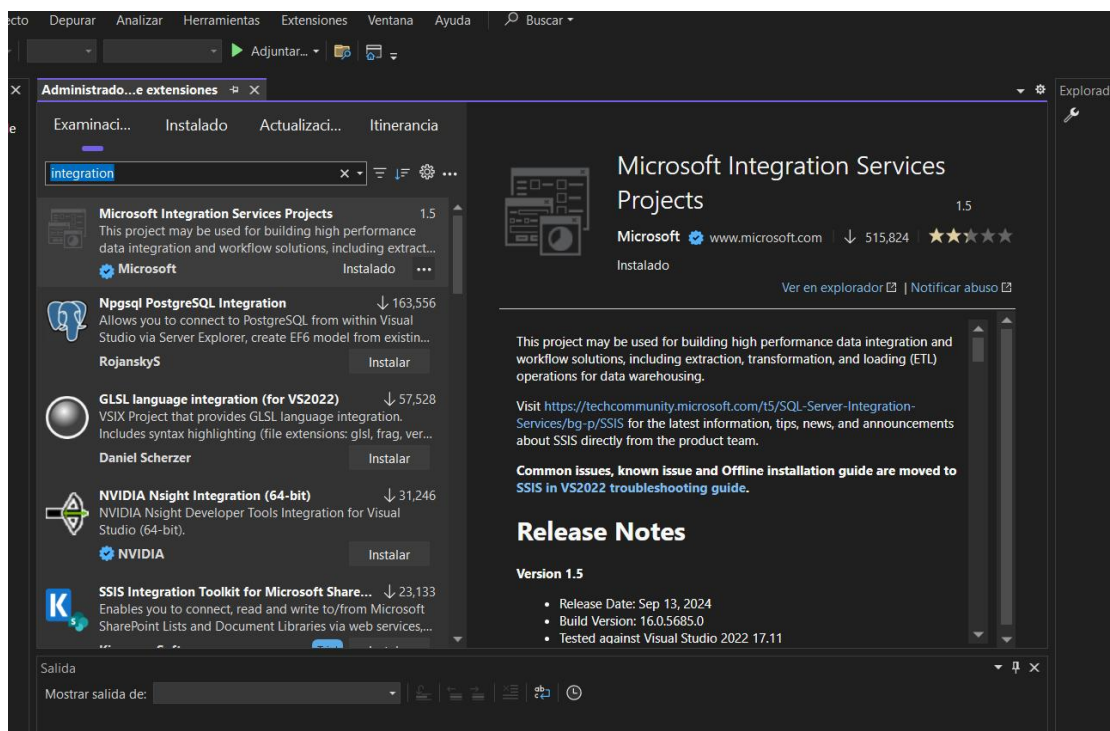
Base de datos MariaDB: En esta base de datos lo que nos interesa es almacenar todo el proceso de venta, desde la creación de la orden hasta el pago y la

facturación, esto principalmente mediante el esquema de comercio electrónico, que nos permite llevar un seguimiento detallado del ciclo de vida en cada venta registrada por magento.

Microsoft Visual Studio: VS mediante SQL Server Integration Services se convierte en una herramienta fundamental para la integración de datos, por lo que hoy en día la convierte en una de las principales herramientas para realizar procesos de extracción, transformación y carga (ETL) de volúmenes grandes de datos. Actualmente utilizaremos Visual Studio 2022 Community Edition por tanto no implica ningún costo de licencia y/o plan de pago.

Inicialmente, cargamos VS e instalamos Microsoft Integration Services Projects, herramienta para creación de procesos ETL.

Extensiones > Administrar extensiones > Buscamos Microsoft Integration Services



Projects e instalamos.

**Figura 11.** Creación de proyecto en visual studio

Como mencionamos anteriormente, Topazio utiliza Odoo y Magento para su gestión de negocio por lo que debemos realizar una extracción de los datos en cada base de datos de estos softwares.

Como primer punto debemos de definir las conexiones de datos hacia las fuentes de los datos, en el caso de Odoo extraemos los datos de su base de PostgreSQL y en Magento extraemos los datos de su base de MariaDB, ambas bases contenidas en un contenedor de Docker respectivamente:

Name	Image	Status	Port(s)	CPU (%)
✓ magento23		Running (3/3)		0.77%
✓ elasticsearch-1 64a641316afd	docker.io/bitnami/elasticsearch:6	Running		0.17%
✓ magento-1 b8d3509c7d46	bitnami/magento:2.3.4-debian-10-r62	Running	443:443 Show all ports (2)	0.33%
✓ mariadb-1 99ca9f6ec71a	docker.io/bitnami/mariadb:10.6	Running	3307:3306	0.27%
mysql 2dc373f3fa6d	mysql:latest	Exited (255)	3306:3306	0%
✓ odoo16		Running (2/2)		0.01%
✓ dbos-1 1411137a2424	postgres:14	Running	5434:5432	0%
✓ web-1 62d674295edf	odoo:16.0	Running	8069:8069	0.01%

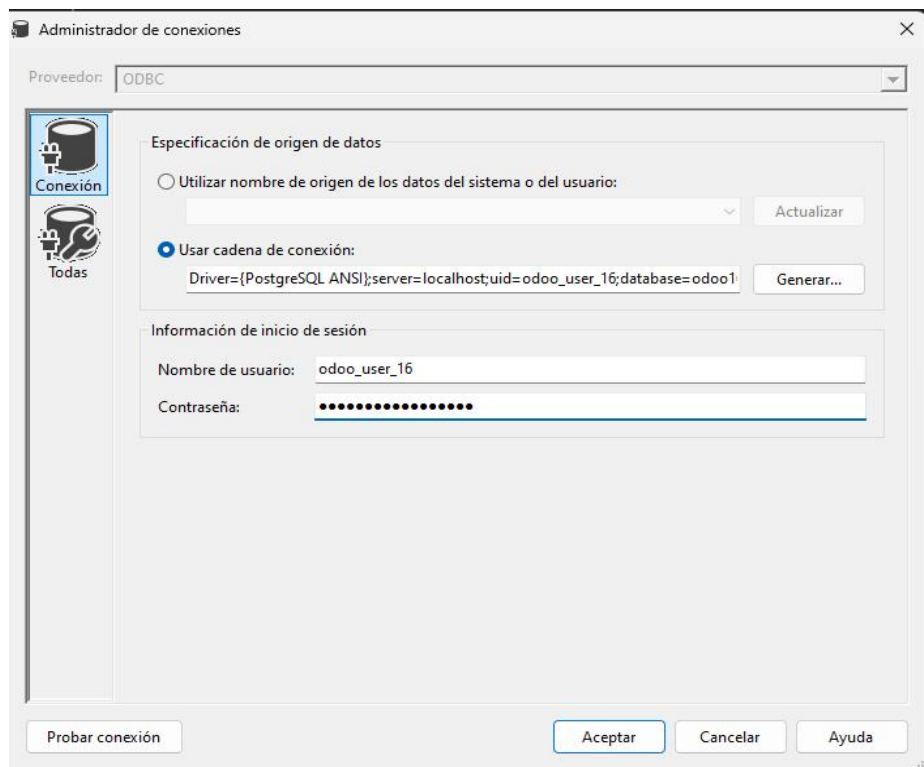
**Figura 12.** Servidor de Base de datos en Docker

Para cada una, hacemos una conexión que nos permita realizar consultas y/o updates a estas bases de datos con las credenciales definidas en los docker-compose de cada contenedor, a su vez hacemos una conexión con nuestra base de datos destino para inserción de la data:

### Conexión MARIADB (Magento)

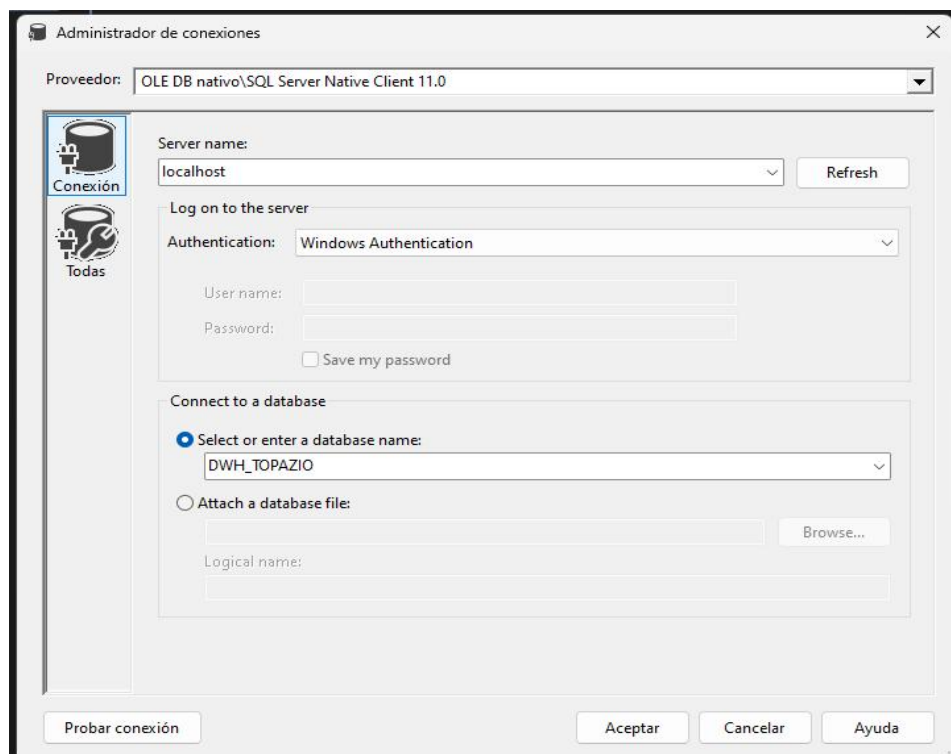
**Figura 13.** Conexión de mariadb para Magento

## Conexión PostgreSQL (Odoo)



**Figura 14.** Conexión de PostgreSQL para Odoo

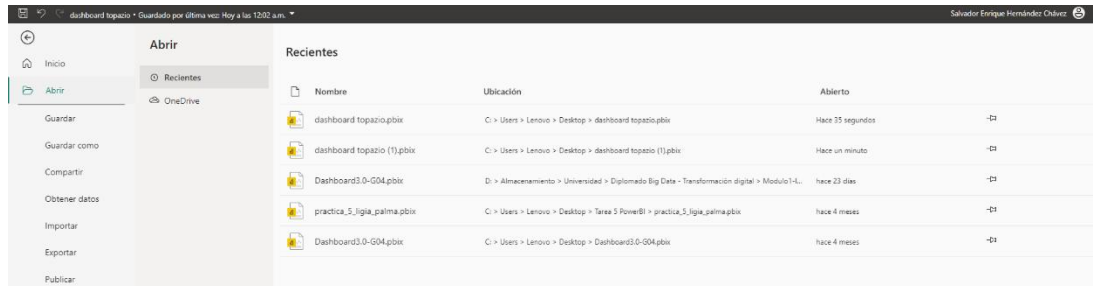
Posteriormente hacemos la conexión a la base que contendrá el DHW



**Figura 15.** Conexión de base de datos para DW

## Power BI

Para poder hacer uso de los reportes dinámicos hechos en Power BI, se debe seleccionar Archivo y dar clic en la opción “Abrir Informe”, luego con el explorador de archivos buscar el informe “dashboard topazio.pbix” o, si previamente fue abierto,

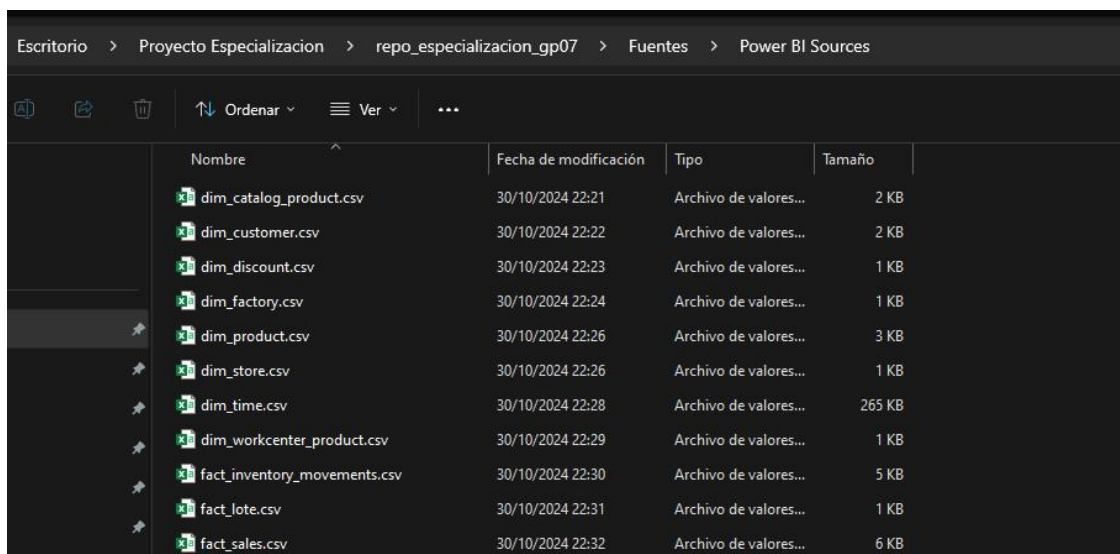


este aparecerá en informes abiertos recientemente:

**Figura 16.** Abrir archivo desde Power BI

## Actualizar Fuentes de datos

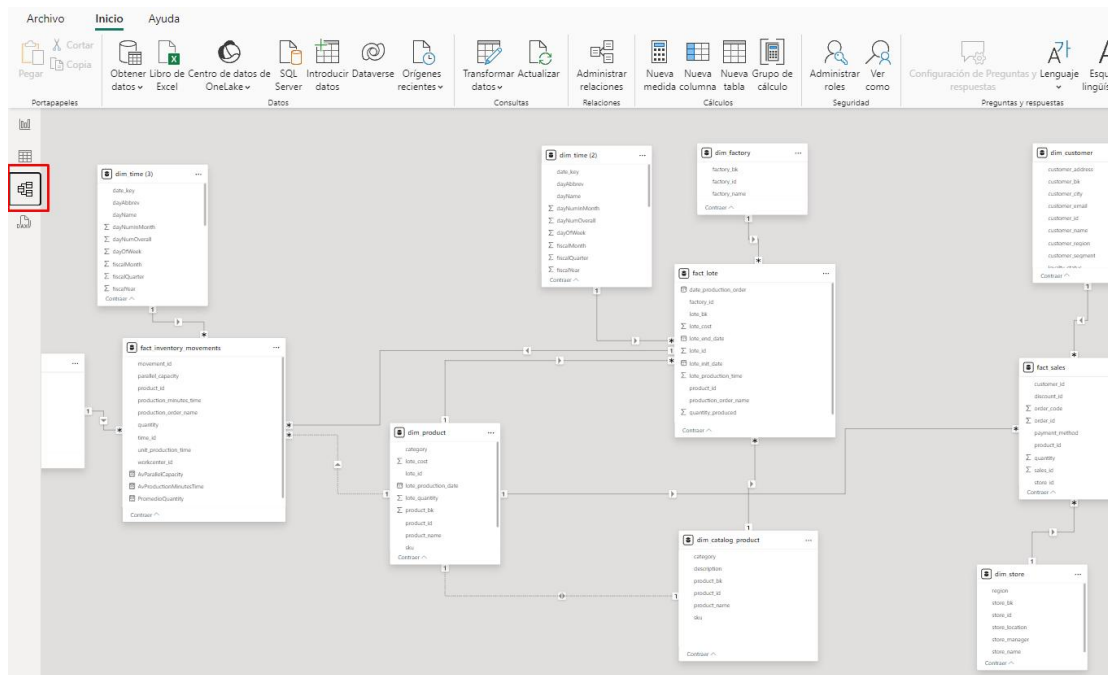
Para actualizar los datos que los informes presentan, estos son exportados en formato CSV, divididos por dimensión y por tabla de hechos respectivamente; estos han sido incluidos en un directorio llamado “Fuentes” -> “Power BI Sources” dentro



del directorio raíz de todo el proyecto.

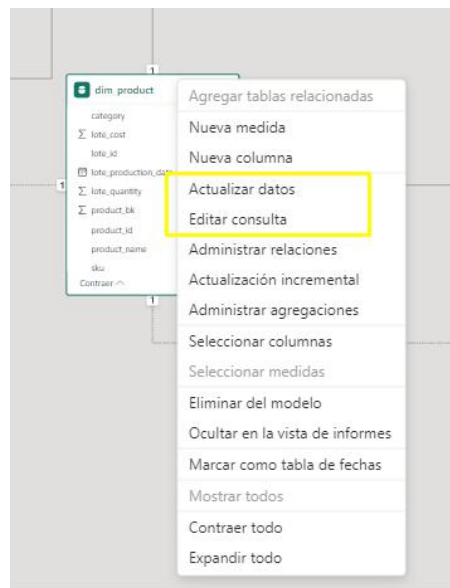
**Figura 17.** Visualización de archivos de dimensiones y tabla de hechos

Para comenzar el proceso de actualización de carga de datos, nos vamos a la opción de Vista del modelo, donde obtenemos una visualización de las entidades y dependencias entre estas cargadas de los modelos de Data Warehouse.



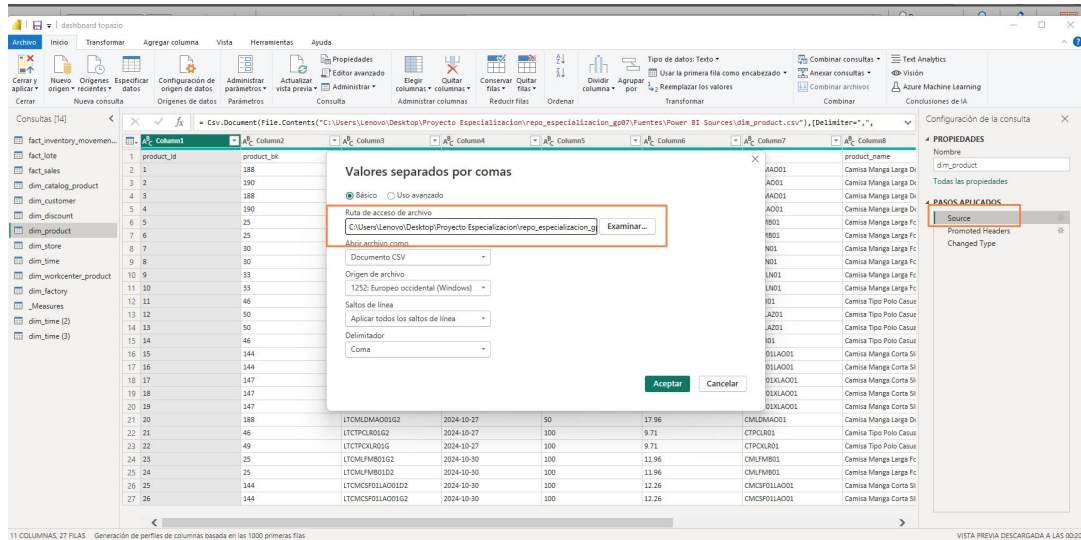
**Figura 18.** Visualización de entidades de los modelos

Ahora hacemos clic derecho sobre cualquier entidad, para este ejemplo se utilizará la dimensión “dim product”. Al hacer clic derecho obtenemos dos opciones que nos permiten actualizar los datos en la entidad, la primera como “Actualizar datos”, esta opción obtiene las actualizaciones en la fuente previamente establecida para esta entidad y actualiza la data en la entidad en cuestión. La segunda opción “Editar Consulta” nos permite establecer una fuente de datos diferente para la entidad, esto es para los casos donde la fuente de datos se encuentra en un directorio distinto al configurado previamente.



**Figura 19.** Visualización de una entidad independiente

Al hacer clic en “Editar Consulta” nos abre una ventana para la configuración de las fuentes de datos de la entidad, para nosotros cargar una nueva fuente hacemos clic derecho en Source en la sección de “Pasos Aplicados” y escogemos “Ir a



configuración”. Desde ahí escogemos el origen de datos por cada tabla.

**Figura 20.** Ventana de configuración de fuentes de datos

### b) Presupuesto de implementación (distribuido en rubros)

Para el presupuesto de implementación se han considerado recursos tecnológicos de licenciamiento, recurso humano y servicios básicos. Actualmente, la organización Topazio cuenta con una infraestructura mínima para almacenar información.

#### Recurso humano

Se calcula el sueldo de un analista de datos esto considerando el salario promedio.

Cantidad de recurso humano	Cantidad de horas requeridas	Precio por hora	Costo total
2	48	\$7.50	720

**Tabla 19.** Presupuesto por recurso humano

#### Recuso tecnológico

El diseño de la solución se ha realizado con software en su mayoría libre tales como: Odo, Magento, Docker. Para la implementación y puesta en marcha de esta solución se considera adquirir el licenciamiento de Power BI Pro, la cual permitirá crear, publicar, compartir informes con gráficos, y paneles modernos de los análisis

de la organización. Actualmente la licencia tiene un valor de \$10.22 por usuario al mes, si consideramos que la empresa requiere de dos usuarios analista de datos para la creación y el mantenimiento de dashboards y reportes, es necesario adquirir una licencia de Power BI Pro para cada usuario. Por tanto, el presupuesto de licenciamiento se presenta a continuación.

Licencia	Cantidad	Precio	Subtotal
Power BI	2	\$10.22	\$10.22
		Total mensual	\$20.44

**Tabla 20.** Presupuesto por recurso tecnológico

### Servicios básicos

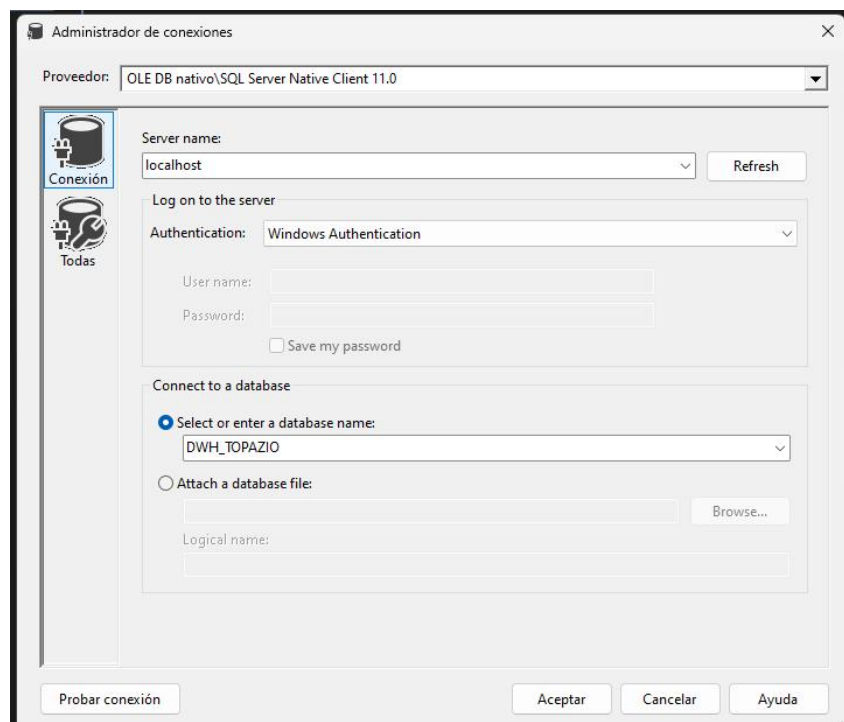
A continuación, se presentan los servicios básicos que la empresa Topazio necesita para su funcionamiento.

Nombre	Cantidad	Precio	Costo total
Energía Eléctrica	1 mes	\$10	\$10
Internet	1 mes	\$26	\$26

**Tabla 21.** Presupuesto por servicios básicos

### c) Análisis de resultados

#### Conexión a la base de datos del Data Warehouse



**Figura 21.** Conexión a la base de datos de DW

## Paquetes de Transformación

Con las conexiones establecidas, procedemos a crear un paquete individual para el proceso de extracción y transformación por dimensión y tabla de hechos correspondiente:

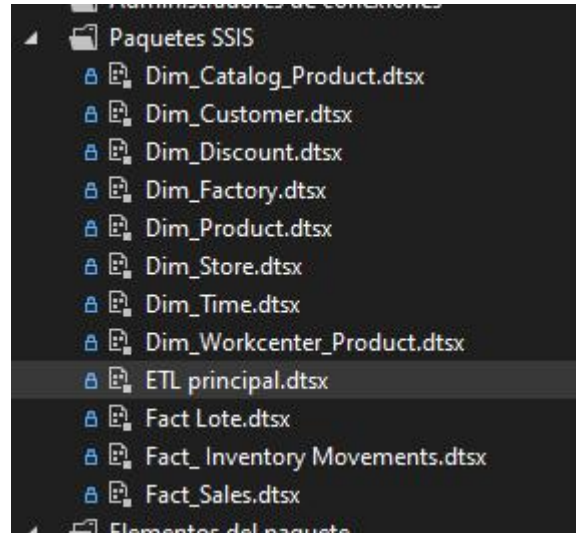
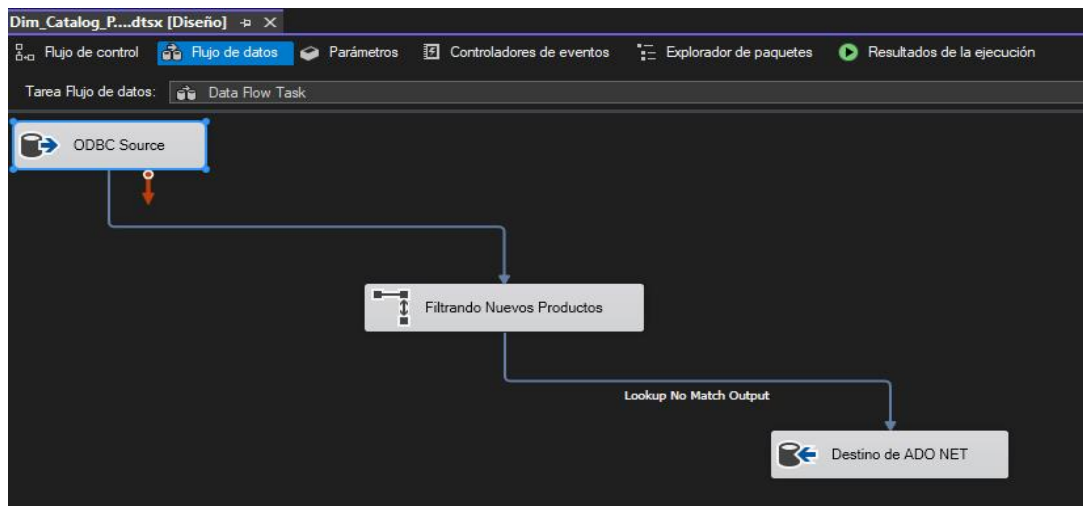


Figura 22. Paquetes del proyecto

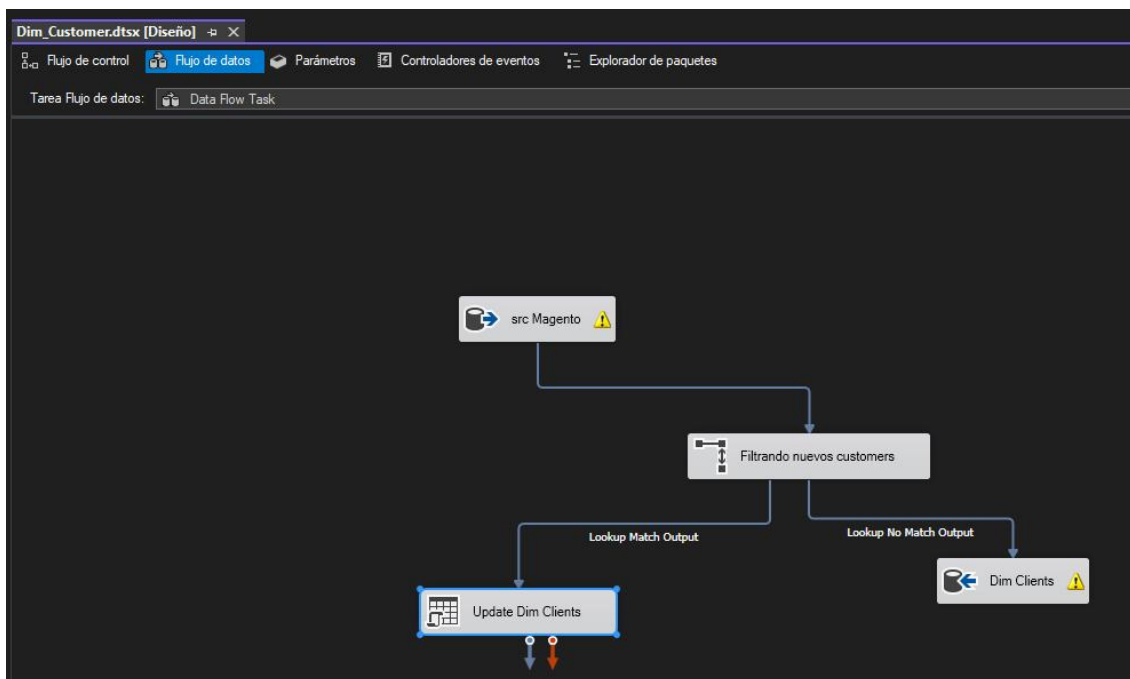
**Dim Catalog Product:** para este paquete se utiliza la conexión hacia magento con una



consulta SQL. El proceso de extracción y transformación sigue el siguiente flujo

**Figura 23.** ETL de dimensión de catálogo de producto

**Dim Customer:** en este paquete se realiza una consulta hacia la conexión de magento, a partir de esta se dividen los nuevos clientes a insertar en la dimensión como también se realiza un update del SCD Tipo 1 Loyalty Status y Customer



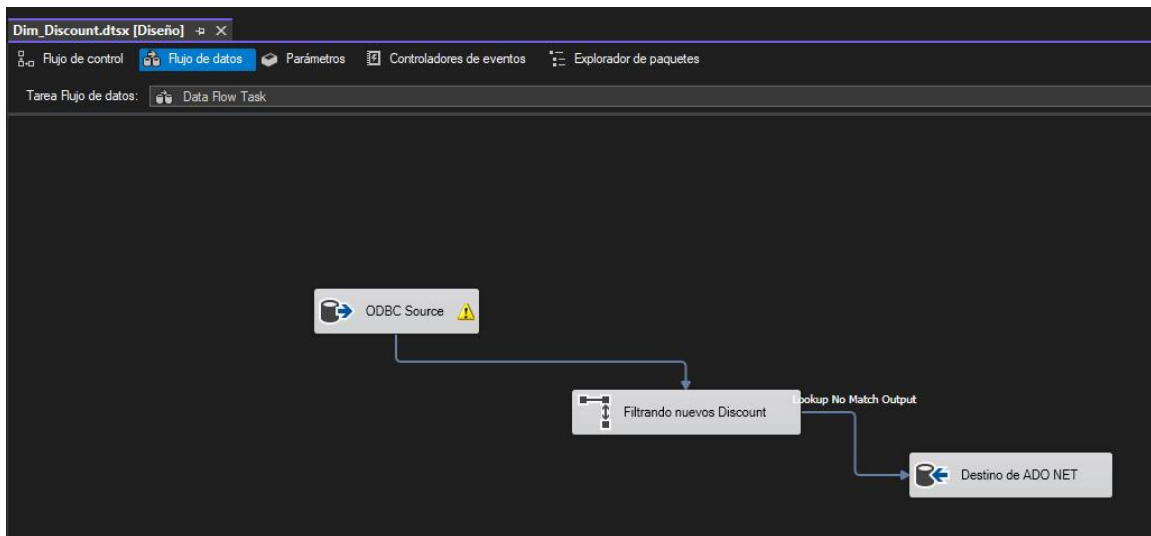
Segment.

**Figura 24.** ETL de dimensión customer

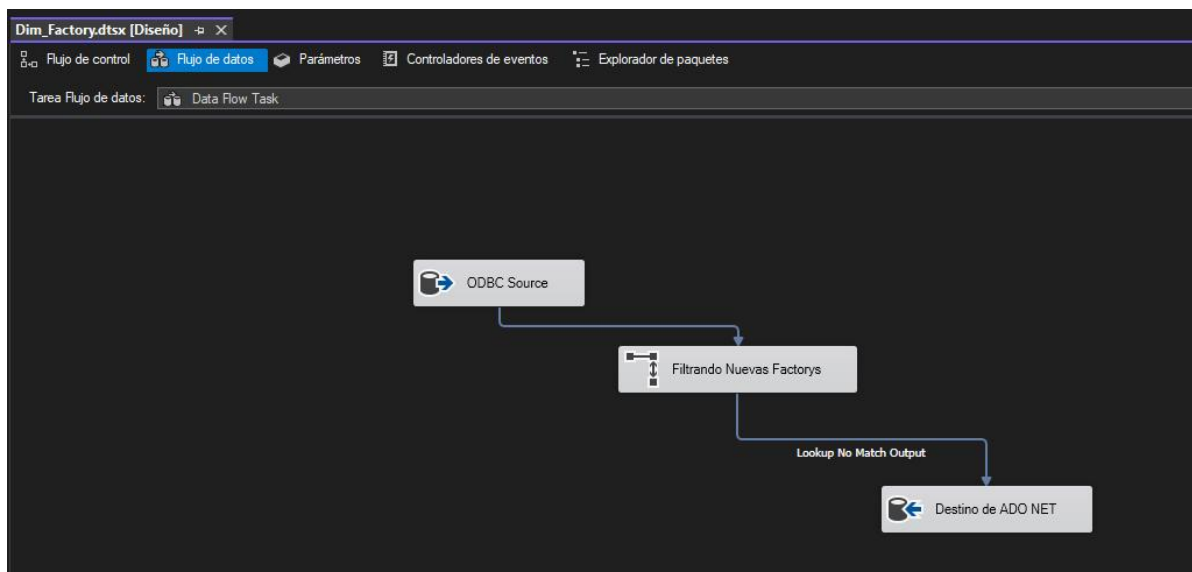
**Dim Discount:** paquete que se encarga de obtener los descuentos configurados en magento y filtrar aquellos que no se encuentran todavía en la dimensión para

insertarlos.

**Figura 25.** ETL de dimensión discount



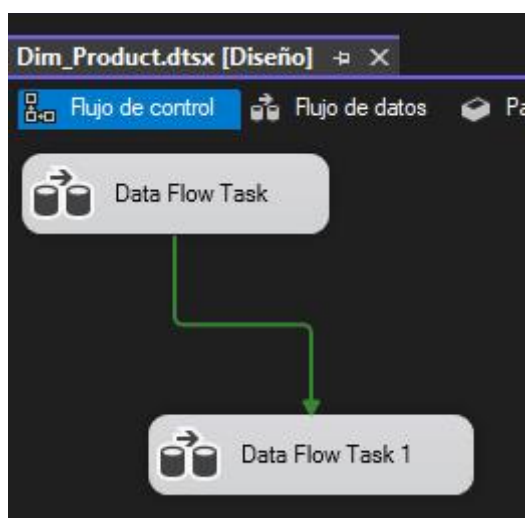
**Dim Factory:** paquete que se encarga de extraer de la conexión de Odoos las fábricas configuradas dentro e insertarlas dentro de la dimensión:



**Figura 26.** ETL de dimensión factory

**Dim Product:** este paquete cuenta con dos secuencias de

flujo de ejecución:

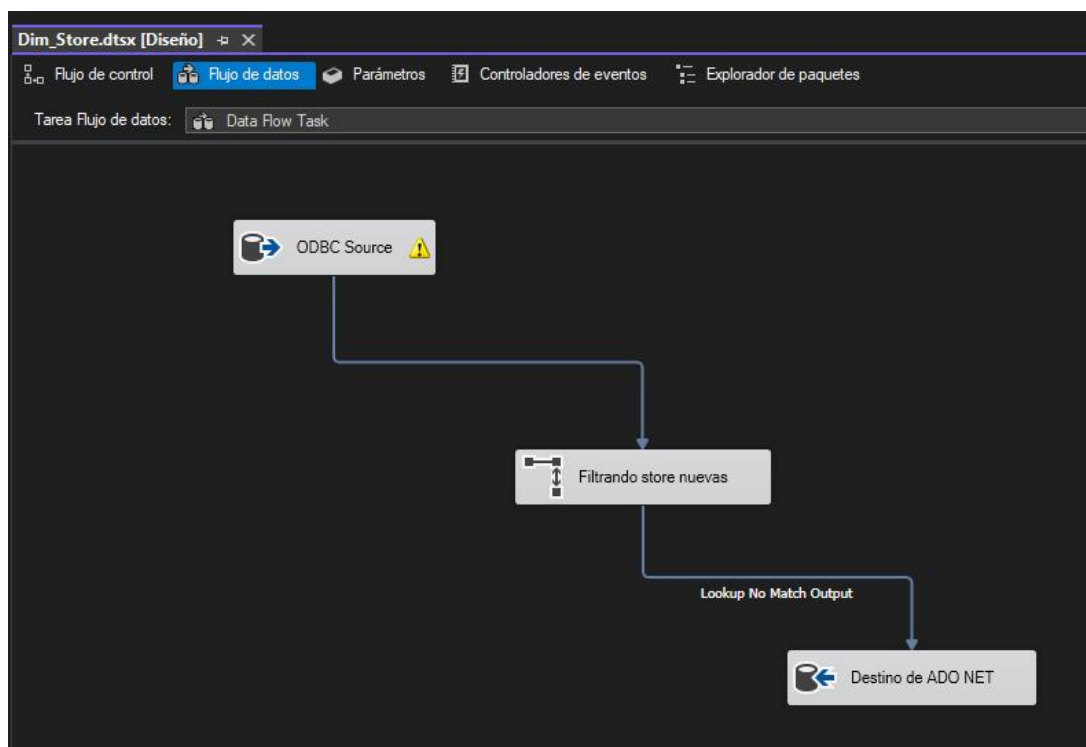


**Figura 27.** ETL de dimensión product

El primer flujo se encarga de obtener el precio del producto configurado actualmente y este es almacenado en caché.

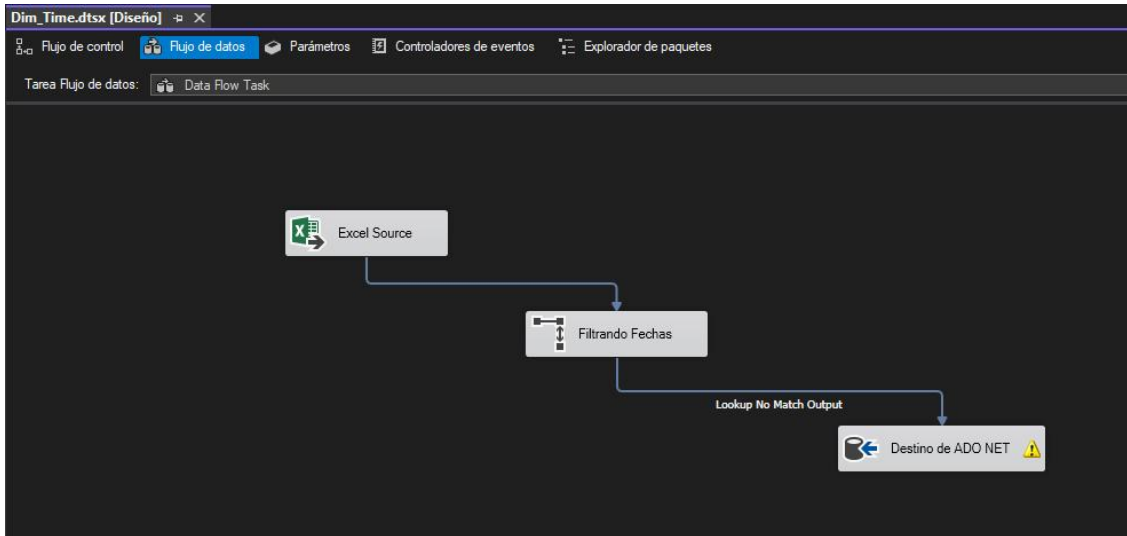
El segundo flujo se encarga de obtener los lotes de productos realizados y asignar precio y categoría de acuerdo con los resultados almacenados en caché.

**Dim Store:** este paquete hace una consulta a la conexión de magento y obtiene las tiendas filtrando aquellas ya registradas en la dimensión.



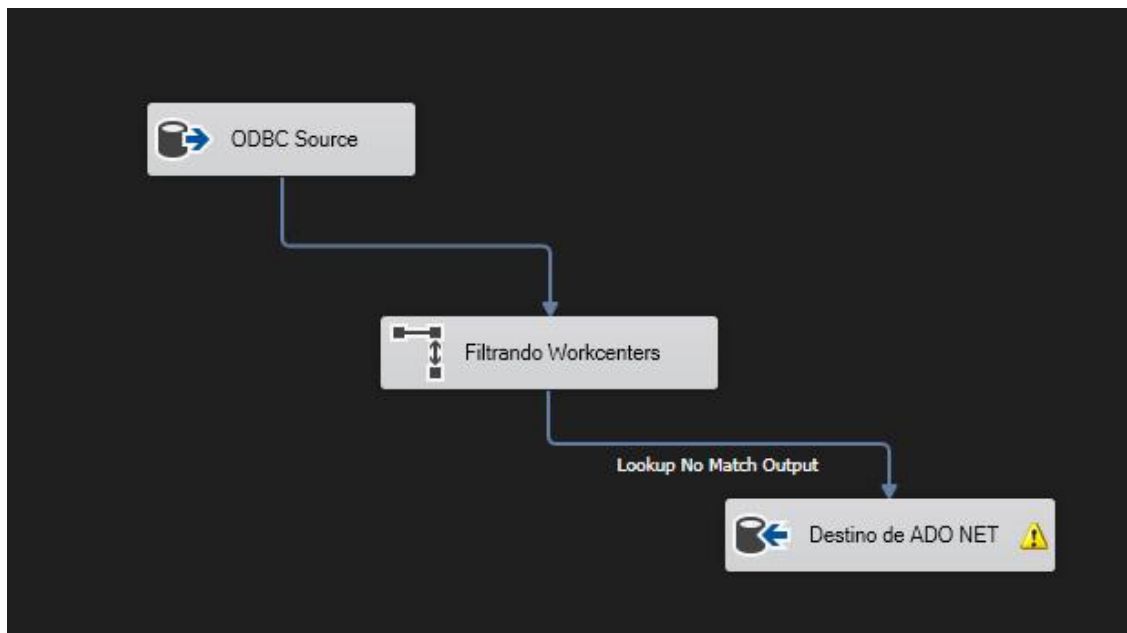
**Figura 28.** ETL de dimensión store

**Dim Time:** este paquete obtiene a partir de un archivo plano de Excel una conexión hacia el subdirectorio de Fuentes dentro de la carpeta de proyecto para realizar una inserción dentro de la dim\_date:



**Figura 29.** ETL de dimensión time

**Dim WorkCenter Product:** obtiene las workcenter configuradas en Odoo y filtra los



registros no insertados previamente en la dimensión:

**Figura 30.** ETL de dimensión workcenter

**Fact Sales:** este paquete hace una consulta para obtener las ventas de producto registradas, a partir de la herramienta LookUp y sus parámetros de venta, obtiene el identificador de cada dimensión las cuales describen el registro específico que se guardará en la fact table; a su vez filtrando por código de orden de venta para

aquellas ventas que no han sido registradas todavía:

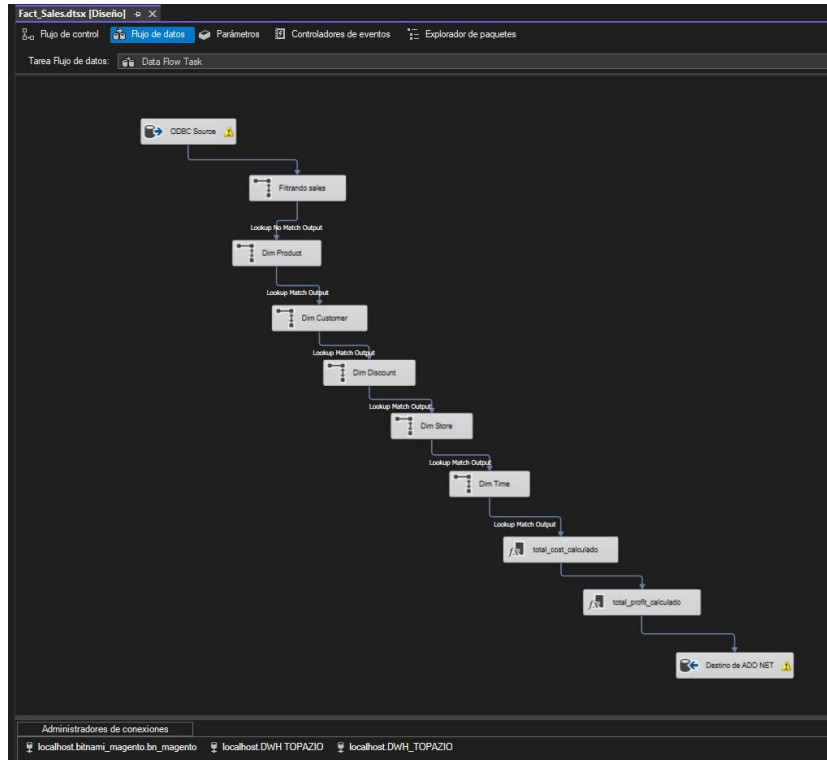
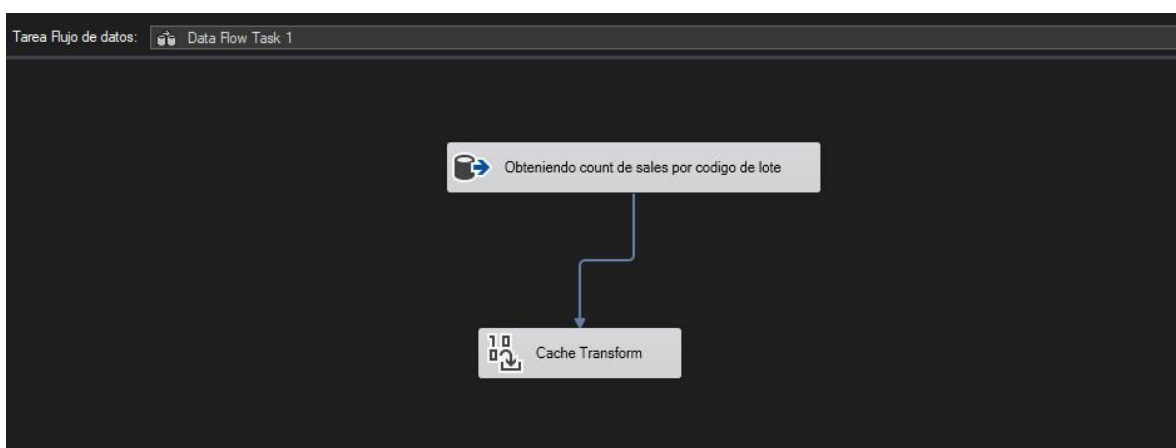


Figura 31. tabla de sales

ETL de la hechos

**Fact Lote:** este paquete contiene dos flujos de ejecución, el primer flujo se encarga de contar la cantidad de producto registrada que fue vendida, dividida por Lote de

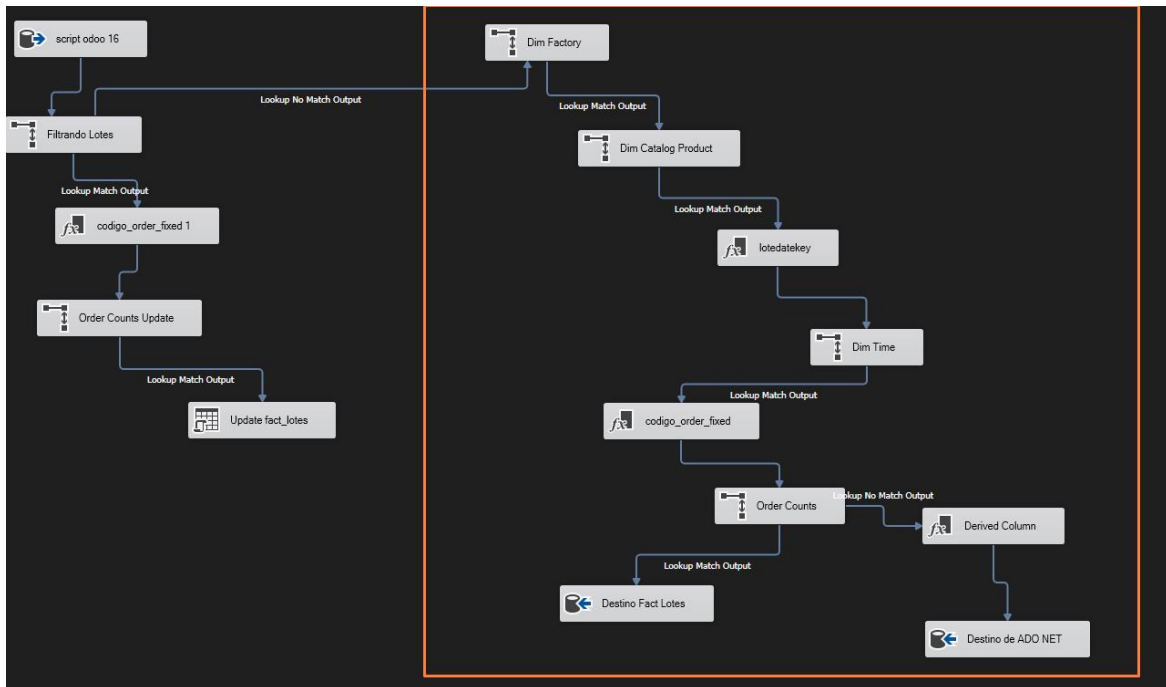


Producto.

Figura 32. ETL de la tabla de hechos lote

El segundo flujo de ejecución hace una consulta hacia Odoon donde obtiene los lotes registrados de producción, se filtran los registros ya insertados de los nuevos y siguen dos flujos diferentes:

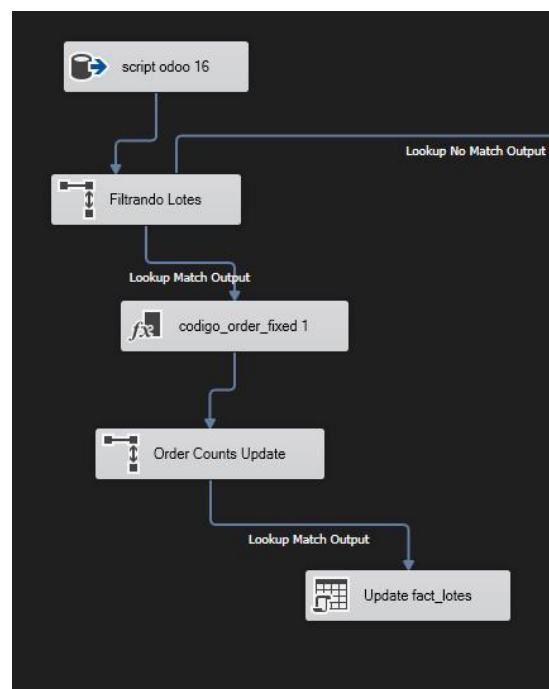
**Nuevos registros:** estos usan la herramienta LookUp para obtener los identificadores de las dimensiones descriptivas del registro del lote, al final obtiene los resultados del conteo de producto vendido por lote y es asignado a la columna



cantidad en esta.

**Figura 33.** ETL de productos vendidos

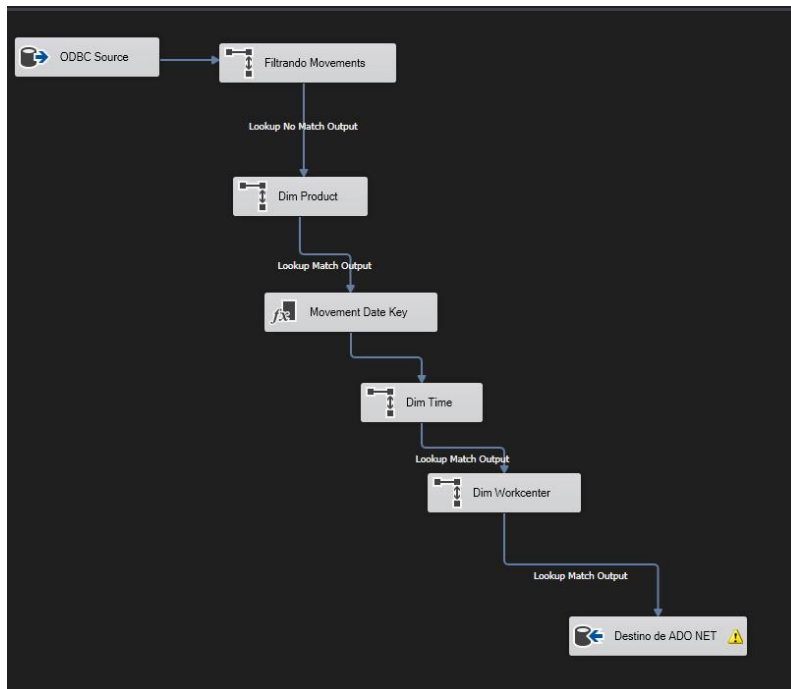
**Registros ya ingresados:** en este flujo se obtiene la cantidad actualizada de productos vendidos del lote y es actualizado el registro correspondiente.



**Figura 34.** ETL de productos registrados

**Fact Inventory Movements:**

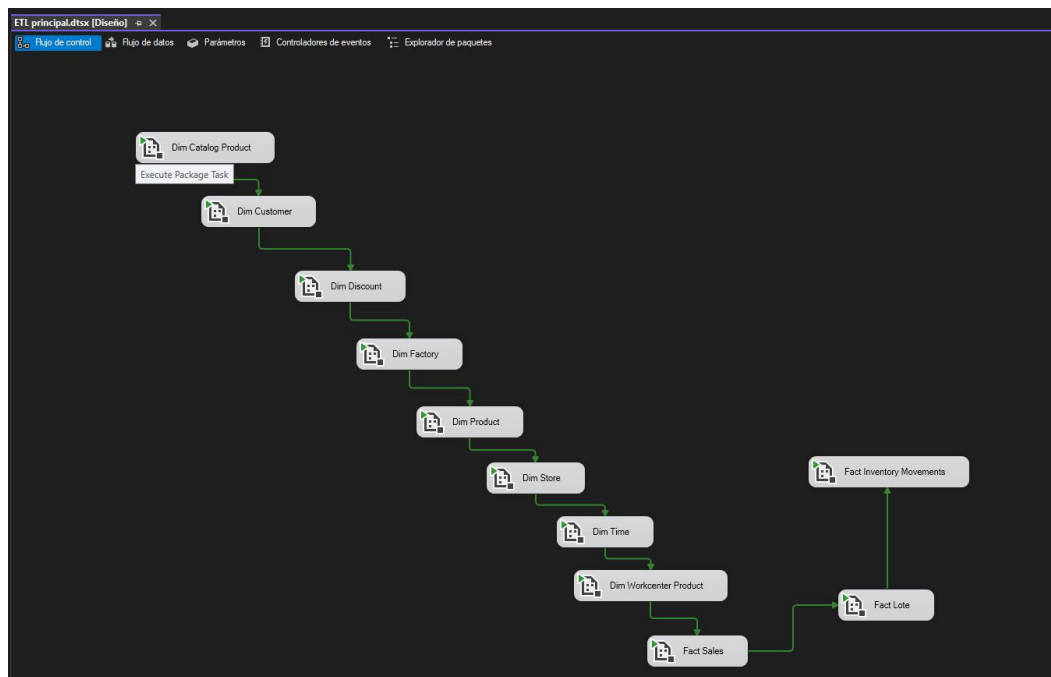
Este paquete hace un filtro de los resultados de los movimientos obtenidos desde Odoo y obtiene los identificadores de las dimensiones que describen cada registro.



**Figura 35.** ETL de tabla de hechos inventory

**ETL Principal:** este paquete se encarga de Orquestar el orden de ejecución de los

paquetes hacia el llenado completo de la base de datos del Data Warehouse.



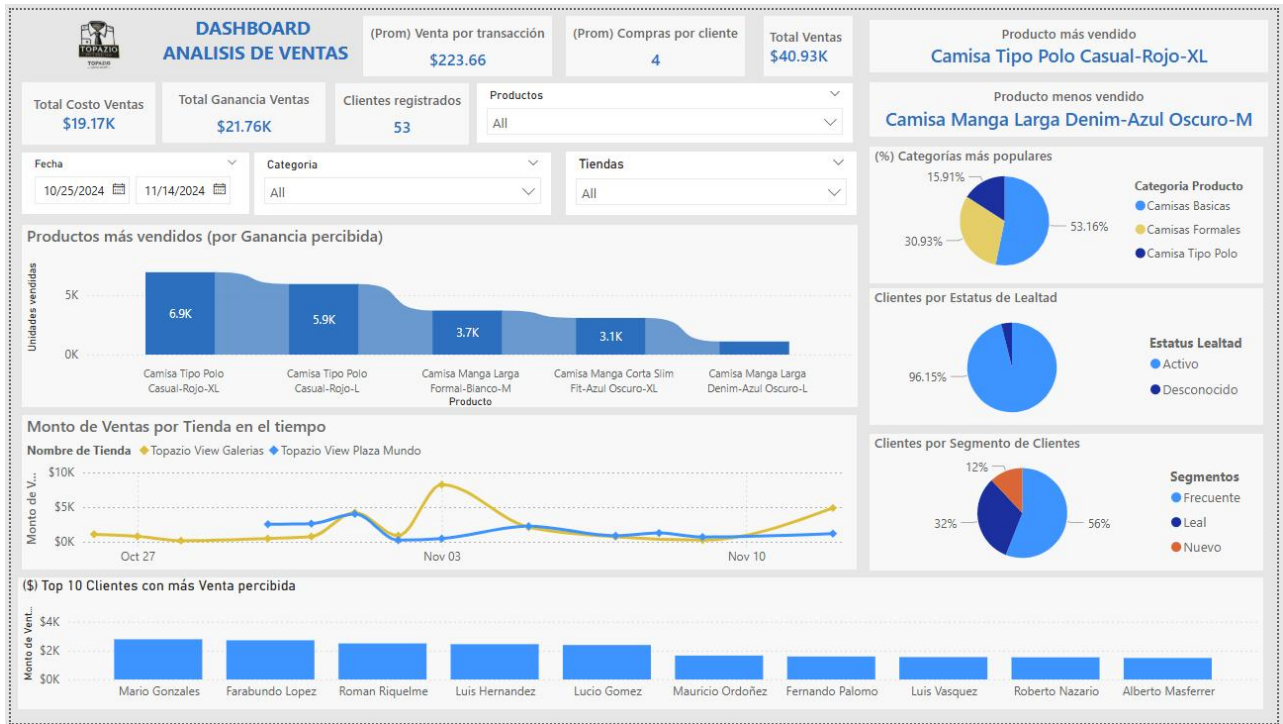
**Figura 36. ETL principal**

### Reportes en Power BI

Dentro de los reportes configurados se presentan:

1. Dashboard de Ventas
2. Dashboard de Lotes de Productos.
3. Dashboard de Movimientos de Inventario.

### Dashboard de Ventas



**Figura 36.** Dashboard de ventas para la empresa Topazio

El dashboard de ventas nos presenta un conjunto de gráficas y KPIs obtenidos de acuerdo con la data recolectada.

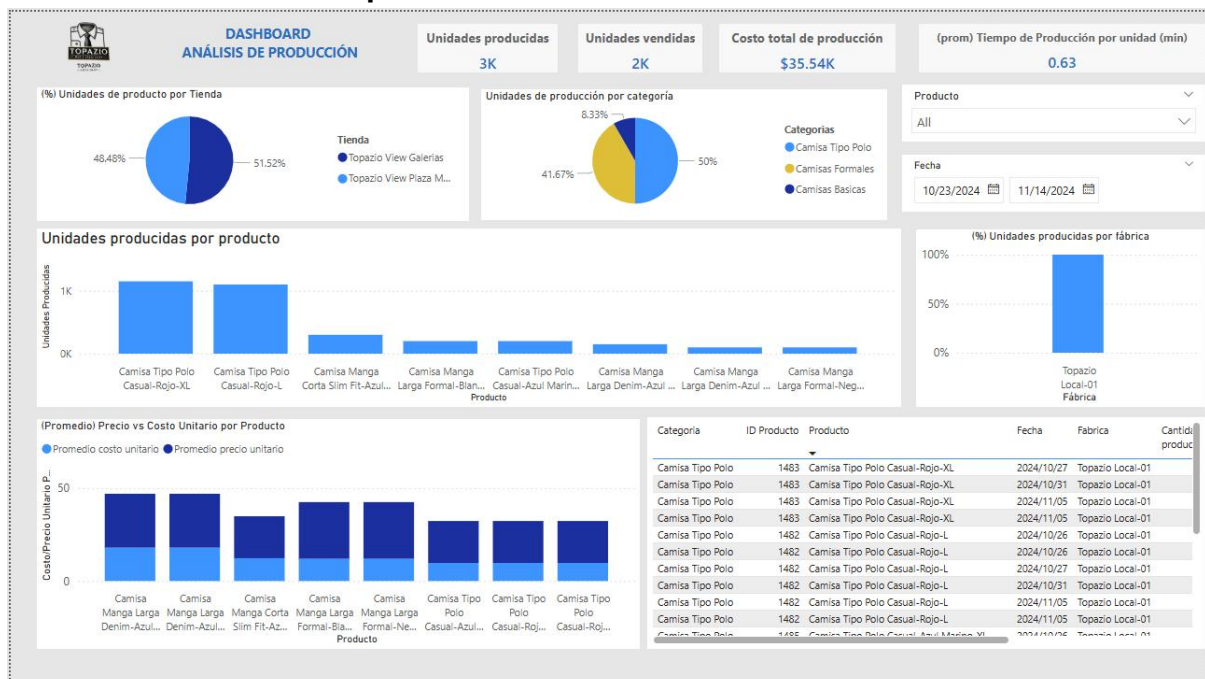
**Entre los KPI tenemos:**

1. Total Ventas: sumatoria de dinero obtenido por todas las ventas independientemente de la tienda.
2. Total Costo Ventas: sumatoria del costo de producto vendido en todas las ventas registradas Independiente de la tienda.
3. Total Ganancia Ventas: el margen de ganancia entre el Total de Ventas menos los costos totales de ventas.
4. Promedio de Venta por Transacción: promedio de dinero que ingresa a la empresa por venta promedio.
5. Producto más vendido: presenta el SKU del producto más vendido en general.
6. No de Clientes registrados: Cantidad de clientes de los que se tiene un registro.
7. Promedio de compras por cliente: Promedio de compras realizadas por un mismo cliente.

**Gráficas:**

- a) Productos más vendidos por unidad: gráfica que permite ver el top de los productos más vendidos a nivel de unidad, es decir aquellos con unidades más vendidas en todas las ventas.
- b) Monto total de ventas por cliente: Gráfica de barras que nos permite ver que clientes son aquellos que generan un mayor impacto en las ventas de la empresa.
- c) Cliente por estatus de lealtad: esta gráfica nos permite ver un estatus de los clientes dentro de la cliente, clasificándolos como Activo o Inactivo, dependiendo del tiempo entre su última compra y la fecha actual.

### Dashboard: análisis de producción



**Figura 37.** Dashboard de producción de la empresa Topazio

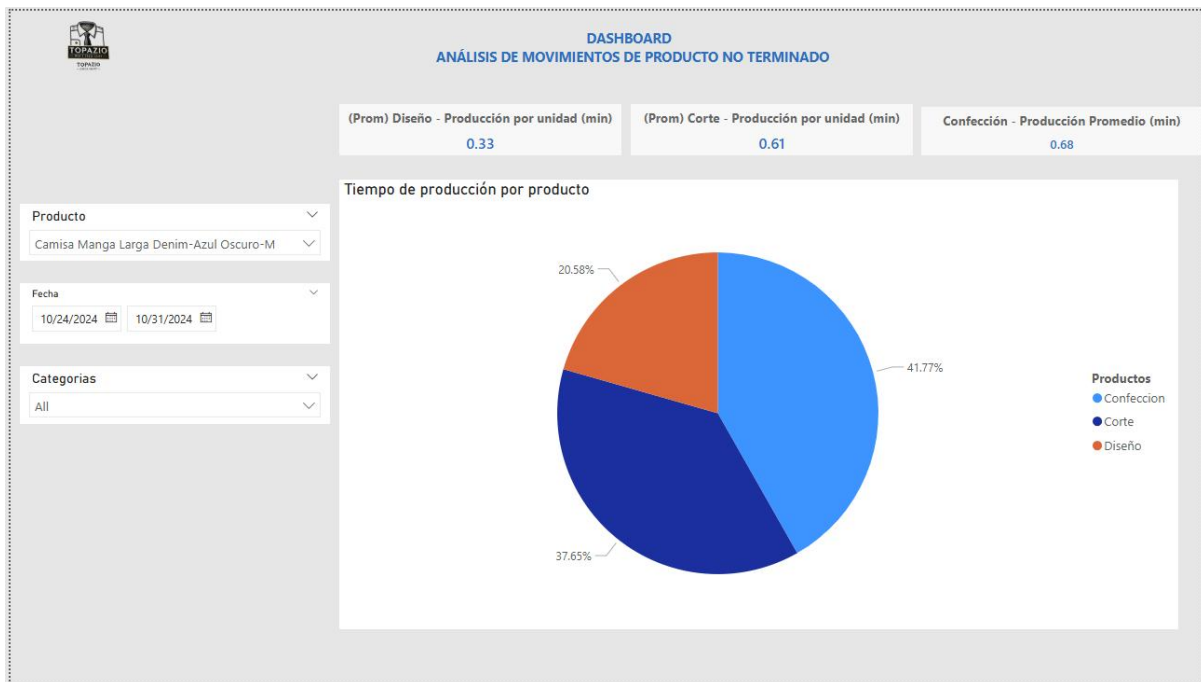
Este tiene KPIs tales como:

1. No. Total de Lotes fabricados.
2. Costo total de Producción: La suma de costos totales de todos los lotes fabricados.
3. Tiempo promedio de producción (min): Este indicador nos muestra el tiempo promedio que un lote toma para ser fabricado, esto expresado en minutos.

Entre sus gráficas tenemos:

- a) Lotes por tienda: este gráfico nos permite ver cuanto porcentaje de los lotes es enviado por tienda.
- b) No de Lotes producidos por fecha: este gráfico nos permite visualizar la cantidad de lotes que son realizados en una fecha especificada y por producto mediante su SKU.
- c) Lotes generados por gráfica: nos permite ver cuántos lotes del total general son hechos por fábrica.
- d) Precio promedio vs Costo unitario por producto: este gráfico nos permite ver en promedio la diferencia de costos vs precio unitario de producto fabricado.
- e) Cantidad producida por costos de productos: este gráfico nos permite ver una progresión de como el producto se comporta en cuanto a la cantidad de productos fabricados y su relación directa con los costos de producción.
- f) Tabla resumen de lotes fabricados: Una tabla que nos permite visualizar datos generales del producto e indicadores como el tiempo de producción y la cantidad producida.
- g) Producción de lotes x Categoría: Nos permite visualizar que categorías de producto son más fabricadas en general por lote.

## **Dashboard movimientos de Inventario**



**Figura 38.** Dashboard de movimientos de inventario de la empresa Topazio

Entre sus KPI tenemos:

1. Tiempo promedio en Estación (min): el promedio de tiempo expresado en minutos que se tarda un lote de producto dentro de una estación de trabajo.

Gráficos:

- a) Tiempo promedio de Producción por estación por capacidad paralela: nos permite ver el tiempo promedio de producción en una estación vs su capacidad paralela de producción.

## Conclusiones

Para la realización de este proyecto se logró analizar, diseñar y construir tres modelos funcionales para la implementación de Data Warehouse que ayudaran a facilitar el análisis de la producción y venta de camisas para la empresa Topazio. Los tres modelos se realizaron teniendo en cuenta procesos de producción terminando con la venta en puntos físicos ubicados en centros comerciales. Estos modelos están diseñados para ofrecer a los usuarios una gestión ágil que permite obtener información concisa que apoye el proceso de toma de decisiones.

Con base en el análisis de los flujos de trabajo y los diseños de base de datos de Odo y Magento y como estos trabajan en conjunto para la gestión de la producción y venta de productos de la empresa Topazio se lograron diseñar y construir procesos para la extracción, transformación y carga de datos para poblar los diseños dimensionales. Durante el proceso de diseño se lograron implementar reglas esenciales para el manejo de valores nulos, conversión de datos, campos calculados, limpieza de bases de datos y manejo del crecimiento de datos con SCD.

Se realizaron tableros de visualización de datos donde se presentan reportes de resultados de los flujos trabajo de producción y venta. Estos reportes se diseñaron teniendo en cuenta que serán utilizados por usuarios que buscan obtener información de manera rápida y concisa para tomar decisiones por lo que se utilizó lenguaje no técnico con detalles que sirven a un nivel alto de la organización. Para ello se hizo uso de graficas de pastel, graficas de barras, grafica de dispersión, títulos informativos, filtros interactivos donde cada elemento del reporte reaccionad de manera dinámicamente ante algún cambio. Estos reportes aportan información muy valiosa de los procesos clave dentro de la empresa Topazio.

Tomando en cuenta la hipótesis planteada al comienzo de este proyecto y el desarrollo y resultados obtenidos concluimos que la hipótesis es verdadera debido a que: se logró hacer un análisis de los procesos de producción y venta desde Odo y Magento dos potentes herramientas utilizadas en conjunto para administrar los procesos de producción y venta de camisas. Con base en estos datos se lograron diseñar modelos dimensionales para los procesos clave. Se tuvo éxito en la extracción, transformación y carga de datos desde las fuentes de datos de los dos sistemas que se utilizan unificando y almacenando la información en tres modelos de estrella. Por último, se logró la creación de informes de datos dirigidos a las personas encargadas de tomar decisiones. Esta información es de alto nivel lo que facilita su comprensión lo cual tiene un gran impacto en la toma de decisiones ágiles para la producción y venta de camisas.

## Recomendaciones

Las etapas de diseño del modelo dimensional y la consecuente etapa de desarrollo de la solución ETL deberían de seguir una metodología de trabajo ágil, esto, pues como cualquier proceso de desarrollo de soluciones de información, se ven afectadas por lo rápido que el ambiente de la empresa y las necesidades de los clientes cambian. Por esto se recomienda la implementación de metodologías útiles que permitan hacer sincronización entre las necesidades del cliente y la solución que se le prepara, para así orientar ésta última con el propósito de cumplir estas necesidades en base a criterios alcanzables.

Se recomienda la utilización de prototipados en el desarrollo de dashboards de visualización de datos, ya sean funcionales o no funcionales, en el proceso de refinamiento de estos a los requerimientos del cliente.

Las herramientas utilizadas son muy efectivas al momento de hacer un uso óptimo de ellas. Sin embargo, esto puede significar limitantes por lo que se recomienda adquirir una versión de paga si se desea otro tipo de complemento.

Poseer hardware óptimo para soportar el procesamiento masivo de datos, ya que la cantidad de información que se maneja es en volúmenes muy grandes.

## Bibliografía

- Ralph Kimball, M. R. (2013). *The Data Warehouse Toolkit (3rd ed.)*. Wiley.
- Abad, D. (2022). *Modelo EAV en Magento*. Obtenido de <https://dabad.es/magento/magento-modelo-eav/>
- Anónimo. (2023). *Anonimo*. Obtenido de PostgreSQL Documentation: <https://www.postgresql.org/docs/14/index.html>
- dotCloud. (2019). *Docker documentation*. Obtenido de <https://docs.docker.com/desktop/install/windows-install/>
- Microsoft. (2018). *Power BI documentation*. Obtenido de <https://learn.microsoft.com/en-us/power-bi/>
- Microsoft. (2021). *Crear un paquete ETL sencillo en SSIS*. Obtenido de <https://learn.microsoft.com/es-es/sql/integration-services/ssis-how-to-create-an-etl-package?view=sql-server-ver16>
- Microsoft. (2022). *SQL Server documentation*. Obtenido de <https://www.microsoft.com/es/sql-server/>
- Microsoft. (2022). *Windows Subsystem for Linux Documentation*. Obtenido de <https://learn.microsoft.com/en-us/windows/wsl/>
- Microsoft. (2024). *Integration Services (SSIS), proyectos y soluciones*. Obtenido de <https://learn.microsoft.com/es-es/sql/integration-services/integration-services-ssis-projects-and-solutions?view=sql-server-ver16>
- Mundy, J. (2011). *Design a Dimension*. Obtenido de <https://www.kimballgroup.com/2011/11/design-tip-140-is-it-a-dimension-a-fact-or-both/>
- Odoo. (2020). *Documentación de Odoo*. Obtenido de <https://www.odoo.com/documentation/16.0/es/>
- Rider, S. (s.f.). *DBeaver Documentation*. Obtenido de 2020: <https://dbeaver.com/docs/dbeaver/>

# Anexos

## Anexo 1. Imagen de entrada y salida de datos

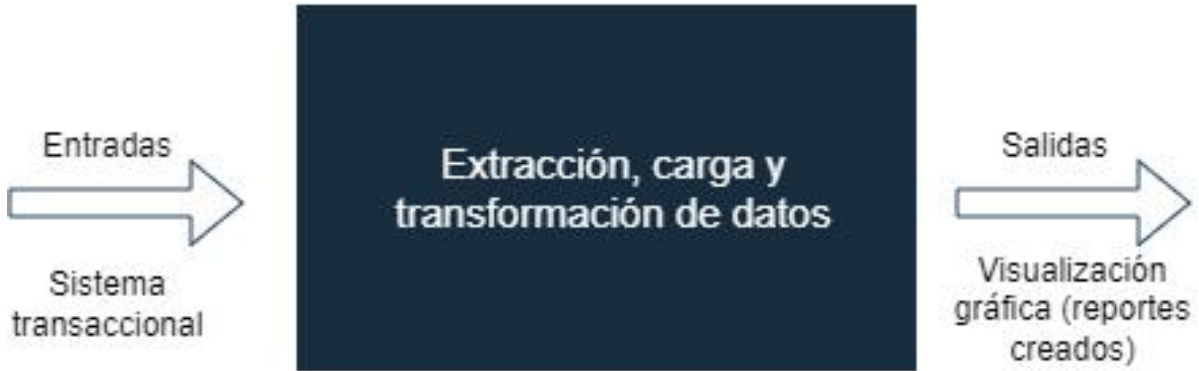


Figura 39. Entrada, procesamiento y salida de datos

## Cronograma de actividades para el desarrollo del proyecto

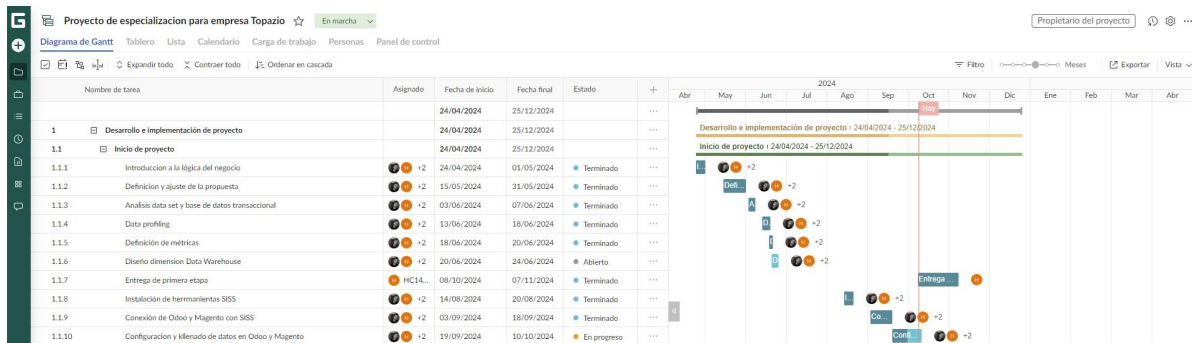


Figura 40. Cronograma de actividades

## Anexo 2. Script para dimensión de catalog product:

```
SELECT
    p.entity_id AS product_id,
    pn.value AS product_name,
    catp.category_id,
    cat.value AS category,
    p.sku,
    CAST(REPLACE(REPLACE(pd.value, '<p>', ''), '</p>', '')) AS VARCHAR(255)) AS
description
FROM
    catalog_product_entity AS p
LEFT JOIN
    catalog_product_entity_varchar AS pn ON p.entity_id = pn.entity_id
    AND pn.attribute_id = (SELECT attribute_id FROM eav_attribute WHERE
attribute_code = 'name' AND entity_type_id = 4)
LEFT JOIN
    catalog_category_product AS catp ON p.entity_id = catp.product_id
LEFT JOIN
    catalog_category_entity_varchar AS cat ON catp.category_id = cat.entity_id
    AND cat.attribute_id = (SELECT attribute_id FROM eav_attribute WHERE
attribute_code = 'name' AND entity_type_id = 3)
LEFT JOIN
    catalog_category_entity_text AS pd ON p.entity_id = pd.entity_id
    AND pd.attribute_id = (SELECT attribute_id FROM eav_attribute WHERE
attribute_code = 'description' AND entity_type_id = 3)
```

