

**UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA
ESCUELA DE INGENIERIA DE SISTEMA INFORMATICOS**



CURSO DE ESPECIALIZACION DE INGENIERIA DE DATOS

**ANALISIS DE INFORMACION SOBRE MANEJO DE ACTIVO
FIJO DE LA EMPRESA INVERSIONES DIGITALES
GESTIONADA A TRAVES DEL SISTEMA ODOO.**

PRESENTADO POR:

**DURAN MARTINEZ, JONATHAN ALEXANDER
GARCÍA MALDONADO, CARLOS ALFREDO
MARTÍNEZ GUERRERO, CLAUDIA PATRICIA
NAVARRO ALFARO, MELANIE CAROLINA**

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

CIUDAD UNIVERSITARIA, ENERO 2024

UNIVERSIDAD DE EL SALVADOR

RECTOR:

MSC. JUAN ROSA QUINTANILLA

SECRETARIO GENERAL:

LIC. PEDRO ROSALIO ESCOBAR CASTANEDA

FACULTAD DE INGENIERÍA Y ARQUITECTURA

DECANO:

ING. LUIS SALVADOR BARRERA MANCIA

SECRETARIO:

ARQ. RAUL ALEXANDER FABIÁN ORELLANA

ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

DIRECTOR:

ING. CESAR AUGUSTO GONZALEZ

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE INGENIERIA Y ARQUITECTURA
ESCUELA DE INGENIERIA DE SISTEMA INFORMATICOS

Trabajo de Graduación previo a la opción al Grado de:
INGENIERO DE SISTEMAS INFORMÁTICOS

Título:

**ANALISIS DE INFORMACION SOBRE MANEJO DE ACTIVO
FIJO DE LA EMPRESA INVERSIONES DIGITALES
GESTIONADA A TRAVES DEL SISTEMA ODOO.**

Presentado por:

DURÁN MARTÍNEZ, JONATHAN ALEXANDER

GARCÍA MALDONADO, CARLOS ALFREDO

MARTÍNEZ GUERRERO, CLAUDIA PATRICIA

NAVARRO ALFARO, MELANIE CAROLINA

Trabajo de Graduación Aprobado por:

Docente Asesor:

Ing. RENE FABRICIO QUINTANILLA GOMEZ

SAN SALVADOR, ENERO 2024

Trabajo de Graduación Aprobado por:

Docente Asesor:

Ing. Rene Fabricio Quintanilla Gómez

Índice

Introducción	2
Capítulo I: Especificación de proyecto	3
a) Situación actual	3
i. Antecedentes	3
ii. Descripción del problema	5
iii. Planteamiento del problema	6
b) Objetivos	7
Objetivo general.....	7
Objetivos específicos.....	7
c) Alcances	8
d) Justificación	9
e) Cronograma de actividades	10
f) Presupuesto	11
Capítulo II: Análisis y diseño de la propuesta de la solución	13
a) Metodología de trabajo	13
b) Descripción de la propuesta de solución	21
c) Descripción de la tecnología a utilizar	23
d) Diagrama arquitectónico de la solución	29
e) Descripción de cada componente de la solución	29
Capítulo III Estrategia de Implementación de propuesta de solución.....	30
a) Estrategia de implementación	30
b) Presupuesto de implementación	60
c) Análisis de resultados.....	61
Conclusiones y recomendaciones	64
Bibliografía	65
Glosario	66
Anexos.....	68

Introducción

El propósito fundamental de este proyecto es realizar un análisis exhaustivo del manejo de activo fijo en la empresa "Inversiones Digitales", especializada en la fabricación y venta de sillas y escritorios de trabajo, distribuidos a través de cinco tiendas estratégicamente ubicadas en las principales ciudades del país. La importancia de este análisis radica en que los activos representan una parte sustancial del capital y los recursos de la empresa. Un manejo efectivo de los activos fijos impacta directamente en la eficiencia de los procesos de fabricación, la calidad de los productos y la satisfacción del cliente.

Con el fin de llevar a cabo esta evaluación, se diseñará una base de datos basada en Odoó que el sistema de gestión que la empresa utiliza para que organice y relacione la información de manera coherente. Las principales tablas incluirán datos sobre tiendas, empleados, productos, activos fijos, ventas, clientes y detalles de ventas. Estas relaciones facilitarán la gestión integral de la información y agilizarán el análisis de los datos. Este monitoreo permitirá a la empresa tomar decisiones fundamentadas sobre la asignación de recursos y la planificación de la producción y venta de sillas y escritorios.

Es esencial que la empresa cuente con la cantidad adecuada de activos disponibles en cada tienda para satisfacer eficientemente la demanda de los clientes. Además, este seguimiento contribuirá a identificar posibles problemas en el manejo de los activos fijos, facilitando la implementación de mejoras y la optimización de su rendimiento.

El enfoque principal del proyecto se centrará en establecer mecanismos y procesos eficientes para el seguimiento de la disponibilidad de los activos fijos. Esto proporcionará una visión clara y actualizada del inventario de activos en cada tienda. Además, se desarrollarán métricas y reportes que ofrecerán información relevante sobre la disponibilidad de los activos y su impacto en las operaciones del negocio.

En resumen, este proyecto abordará el seguimiento y control de la disponibilidad de los activos fijos en "Inversiones Digitales", identificando y resolviendo posibles problemas de disponibilidad para optimizar la eficiencia en la producción y venta de sillas y escritorios. Al garantizar que los activos estén disponibles cuando se necesiten, la empresa podrá ofrecer un servicio superior a sus clientes y maximizar su rentabilidad.

Capítulo I: Especificación de proyecto

a) Situación actual

i. Antecedentes

A lo largo de los años, las empresas se han dado cuenta de la importancia de la correcta manipulación de su información para el desarrollo de su negocio, por eso existe un concepto llamado Business Intelligence o su abreviatura en inglés BI.

El análisis de datos hace que se conviertan los datos operativos empresariales en información que respalde la toma de decisiones que los ayude a respaldar la toma de decisiones comerciales de manera más eficaz y eficiente para lograr tener un mejor lugar dentro del mercado en donde se encuentren y sobre su competencia.

Ahora Sabiendo esto, las organizaciones están buscando mejorar sus sistemas de toma de decisiones, ya que pueden verse abrumados por la cantidad y la complejidad de los datos disponibles en los sistemas de producción y transacción, con el análisis de datos se puede lograr que todos estos datos estén disponibles para todos los involucrados en el proceso y donde uno de los desafíos más importantes que enfrentan los profesionales de toda la industria en la actualidad es la participación y entendimiento en la tecnología de la información.

Inversiones Digitales es una empresa con 5 años en la industria de manufacturación de artículos de oficina, enfocados en la elaboración de sillas y escritorios, comercializando a lo largo del territorio salvadoreño.

Cuenta con 5 sucursales distribuidas en diferentes puntos del país, a continuación, se detallan las diferentes ubicaciones.

Zona central

San Salvador: En este departamento contamos con la casa matriz, donde se lleva la elaboración de todos los productos, para posteriormente distribuirlos a sus sucursales y asimismo realizar ventas al público.

Chalatenango: Ubicada en el centro de Chalatenango.

La Libertad: Ubicada en el centro de Santa Tecla.

Zona Occidental

Sonsonate: Ubicada en el centro de Juayúa.

Zona Oriental

San Miguel: Ubicada en el centro de San Miguel Centro.

En los últimos 2 años Inversiones Digitales ha presenciado un alto volumen en sus ventas, por consiguiente, el volumen de datos almacenados respecto a los activos que posee la empresa ha incrementado y se ha visto en la necesidad de una mejor organización y administración de datos para poder tomar mejores decisiones a nivel gerencial, de esta manera se llegó a la idea de crear un área de análisis de datos donde se pueda manejar los grandes volúmenes de datos de sus activos fijos con el fin de observar la información que pueda ser de utilidad para ellos.

Esta empresa se apoya en el sistema ERP ODOO, específicamente en el módulo de facturación/Contabilidad.

Entre los principales beneficios que ODOO ofrece en el manejo de activos fijos, se mencionan los siguientes:

- Automatización del ciclo de compra: ODOO facilita la gestión de activos fijos al crear facturas de compras desde el módulo de contabilidad donde se lleva el control de la compra de los activos fijos por tienda registrada.
- Gestión de múltiples empresas: Ahorra tiempo y esfuerzo con las reglas multiempresa de ODOO. Utiliza una sola instancia para sincronizar operaciones entre diferentes empresas. Crea órdenes de compra, comparte activos, proveedores y productos, y gestiona facturas para todas las empresas simultáneamente. Automatiza el proceso de facturación entre las empresas para ahorrar aún más tiempo.

Para comprender el proceso de adquisición de activos fijos, es fundamental entender la lógica del negocio y su relación con otras aplicaciones de ODOO.

- Relación Compra – Inventario: Sincroniza los niveles de existencias según las compras y realiza un seguimiento de los artículos adquiridos. Cada compra afecta al módulo de Inventario con un movimiento de ingreso.
- Relación Compra – Facturación: Convierte las órdenes de compra en facturas de proveedor para evitar la duplicidad de entradas donde se obtiene información significativa sobre las facturas de proveedores.
- Relación Compra – Venta: Convierte las órdenes de venta en facturas de clientes para evitar la duplicidad de entradas donde se obtiene información significativa sobre las ventas realizadas a los diferentes clientes.

ii. Descripción del problema

Inversiones Digitales es una línea de tiendas que se distribuyen en 5 sucursales donde cada tienda hace compras a diferentes proveedores para el abastecimiento de cada una de estas que se encuentran ubicadas en diferentes locaciones del país y según la demanda de los artículos se espera que las tiendas compren los activos a vender a los diferentes clientes.

Es por eso que tiene que poseer un sistema de organización de sus activos en donde se pueda lograr el cumplimiento de la demanda de adquisición de los clientes con los que cuenta la empresa por medio de una correcta gestión de compra de inventario en cada tienda.

Creando así la necesidad de comenzar a utilizar los servicios del análisis de datos para la empresa por medio de las diferentes herramientas disponibles como los son AWS S3, IAM, AWS Redshift, Talend Open Studio, Power Bi y PostgreSQL, para la correcta obtención y manejo de la información que la empresa ha recolectado durante estos años de gestión y poder proceder de esta manera a una correcta toma de decisiones para la empresa.

Los procesos principales con los que la empresa inversiones digitales cuenta son los siguientes.

- Compra de activos
- Venta de activos
- Reporte de compras
- Reporte de ventas
- Facturación

En el contexto de las adquisiciones de activos fijos para la empresa Inversiones Digitales por medio del uso del software ODOO, se ha establecido una serie de métricas cruciales que desempeñarán un papel fundamental en el análisis de datos y la toma de decisiones estratégicas. Estas métricas se detallan a continuación, proporcionando una visión más amplia del enfoque adoptado por la empresa:

Cantidad de Adquisiciones por Activo Fijo (conteo): Esta métrica se centra en cuantificar la cantidad de activos fijos adquiridos, permitiendo un seguimiento detallado de las unidades de activos en el inventario. Facilita la identificación de los activos más adquiridos y contribuye a una gestión más eficiente de los recursos.

Volumen de Adquisiciones por Activo Fijo (monto): Enfocándose en el aspecto financiero, esta métrica evalúa el valor monetario total de las adquisiciones de activos fijos. Proporciona una perspectiva clara sobre la inversión financiera realizada en cada tipo de activo, lo que ayuda a la empresa a entender la distribución de costos y priorizar inversiones estratégicas.

Diversidad de Activos Fijos por Proveedor (conteo): La relación con los proveedores es fundamental. Esta métrica se concentra en cuantificar la diversidad de activos fijos adquiridos a través de diferentes proveedores. Permite evaluar la dependencia de la empresa en relación con sus fuentes de suministro. Estas métricas proporcionan una base sólida para evaluar el rendimiento de las adquisiciones de activos fijos de Inversiones Digitales. Al analizar la cantidad y el valor de las adquisiciones, así como la diversidad de proveedores, la empresa puede tomar decisiones informadas para optimizar su cartera

de activos y mejorar la eficiencia operativa. Además, la implementación de estas métricas en el entorno ODOO facilita un seguimiento continuo y una actualización en tiempo real de los datos, respaldando así la toma de decisiones ágil y estratégica.

iii. Planteamiento del problema

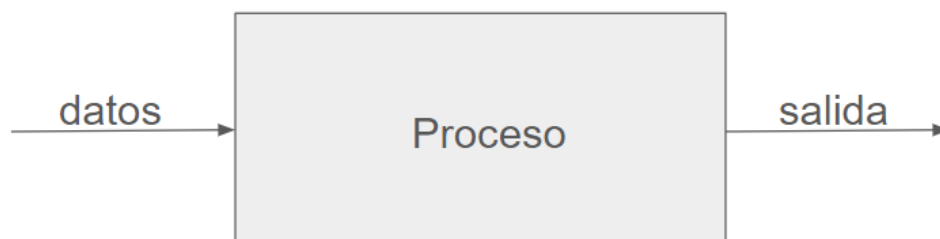
Con el constante progreso tecnológico, las oportunidades para modelar datos se amplían cada día. Este fenómeno resulta beneficioso, dado que, en algunas instancias, la información generada por las organizaciones no se presenta de manera estructurada para su almacenamiento en una base de datos relacional. En cambio, se manifiesta como lo conocemos hoy en día: un Data Lake. Estos conjuntos de datos carecen de estructura o presentan una semiestructura, pero pueden ser tratados mediante un proceso ETL. Este procedimiento implica la depuración y transformación de los datos, filtrando únicamente la información relevante para el análisis organizacional.

A continuación de manera más gráfica ejemplificaremos el planteamiento del problema.

Entradas: Como entrada tendremos la información generada mediante la plataforma OdoO, específicamente toda aquella información que tenga que ver con activos fijos de la empresa Inversiones Digitales.

Proceso: En esta parte se evidenciarán las diferentes herramientas que nos ayudarán para el análisis de los datos extraídos de las entradas, para dicha extracción necesitamos un Sistema Gestor Base de Datos, en este caso se ocupará PGAdmin, y para implementar el ETL necesitamos Talend Open Studio, AWS Services donde crearemos buckets, usuarios con permisos y el modelo dimensional y por último la herramienta que nos permitirá visualizar los datos de una forma muy estructurada y lógica como lo es Power BI.

Salida: En esta etapa tendremos la representación de la información de manera visual mediante gráficos para presentar informes a la alta gerencia y llegar al punto importante de la organización que es la toma de decisiones.



b) Objetivos

Objetivo general

Construir un prototipo para la implementación de un DataWareHouse sobre el análisis de datos de activo fijo de la empresa “Inversiones Digitales”, que será gestionado a través del sistema ODOO.

Objetivos específicos

- Comprender en detalle el proceso de manejo de activo fijo de la empresa mediante el sistema ODOO.
- Analizar y perfilar los datos disponibles en la empresa relacionados con el manejo de activo fijo.
- Describir de manera precisa el conjunto de datos involucrados en el proceso de manejo de activo fijo, incluyendo sus metadatos.
- Identificar las necesidades analíticas específicas que se pretenden abordar mediante la implementación de un modelo multidimensional.
- Diseñar y desarrollar un modelo multidimensional que satisfaga las necesidades analíticas relacionadas con el manejo de activo fijo.
- Establecer una correspondencia clara y precisa entre el modelo multidimensional propuesto y el conjunto de datos del sistema ODOO utilizado en el manejo de activo fijo.

c) Alcances

- Lograr que el personal de Inversiones Digitales obtenga una mejor estructura y diseño de los datos generados de sus activos fijos para alcanzar el éxito empresarial mediante la creación de procesos que extraigan, transformen y finalmente muestren la información procesada.
- Diseño de un modelo dimensional con la metodología de estrella que se alimentará de los datos procesados.
- Presentar la información procesada mediante gráficos para que la alta gerencia tome sus respectivas decisiones a partir del análisis que se muestra.

d) Justificación

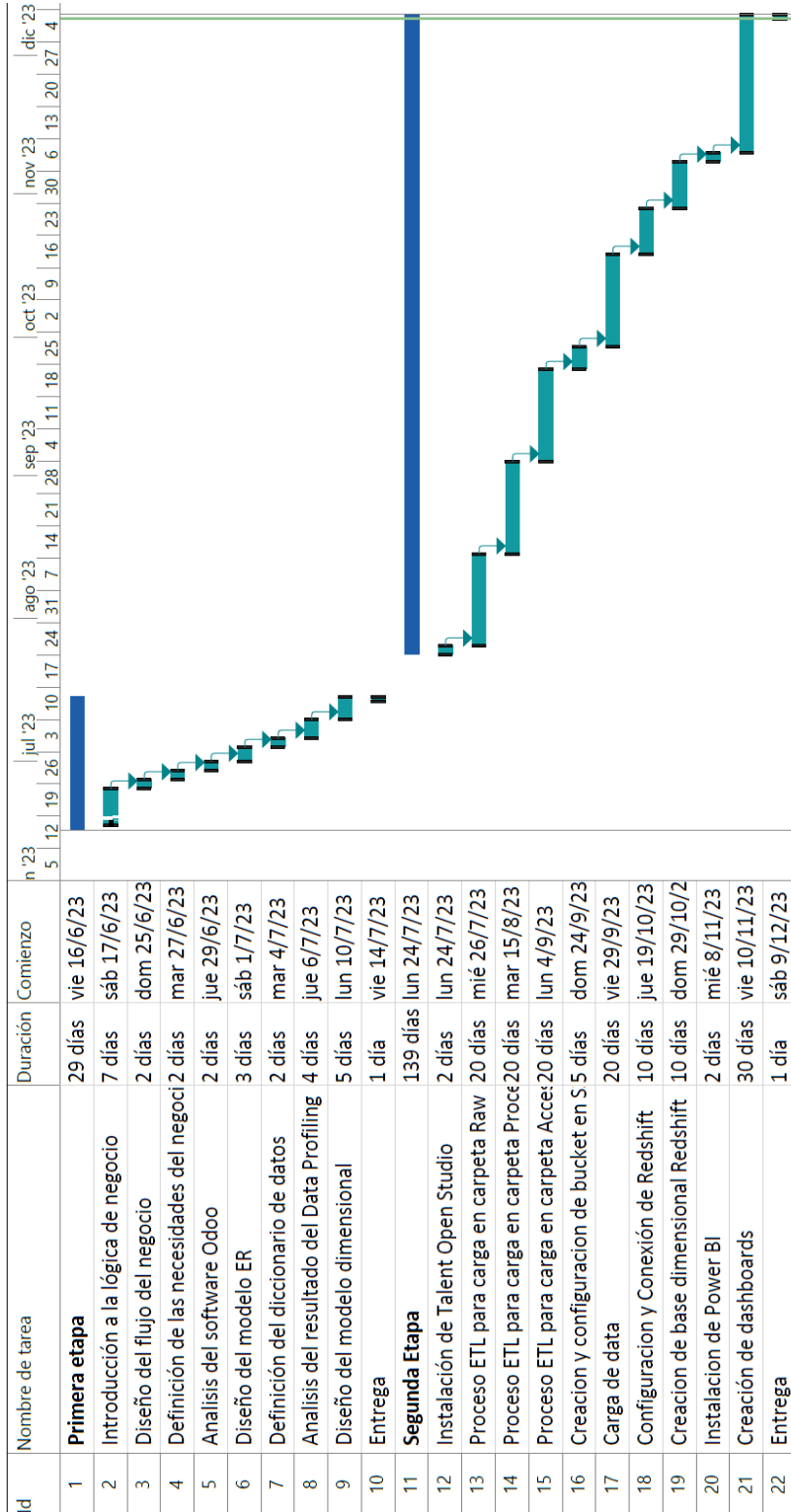
Dado al gran avance en el área de análisis de datos, para las empresas actuales es de suma importancia poseer procesos efectivos y eficientes para el manejo de datos en grandes volúmenes debido a la cantidad de transacciones que suceden dentro de estas. Es por eso que se necesita una mejor gestión de esta información, la cual es muy importante para la toma de decisiones en las empresas y poder tener mejores elecciones a futuro y mejorar los procesos en el presente que estas lleguen a poseer por medio del conocimiento y manejo correcto de sus datos.

El análisis de datos tiene muchas soluciones efectivas para lograr organizar y analizar los datos dentro de las empresas, de esta manera les brinda la oportunidad de tomar decisiones más inteligentes en donde antes todos los procesos eran manuales, llevaban más tiempo y esfuerzo, ahora con el avance de la tecnología todo esto se volvió más automatizado para poder de esta manera sacar un mejor provecho y obtener todas las ventajas que estos nuevos procesos pueden ofrecer para lograr ir un paso más adelante de las empresas que son competencia y poder así tener un mejor nivel sobre todas estas.

El módulo de contabilidad que la empresa de inversiones digitales tiene en uso proporciones grandes volúmenes de datos con los cuales, sin el debido trato del proceso de análisis de datos, no podrían ser explotados en su magnitud y no se lograría realizar las proyecciones necesarias para el manejo correcto de los activos fijos de la empresa.

Es por eso, que como empresa se decidió implantar el proceso de solución que les dé la oportunidad de obtener esta información, transformarla y cargarla dentro de las herramientas correcta para su correcto análisis y presentación en tableros gráficos con esta información y poder así apoyar la toma de decisiones.

e) Cronograma de actividades



f) Presupuesto

AWS

Para poder obtener los valores de los precios estimados mensuales y anuales basándose en las herramientas a utilizar como AWS, IAM, Redshift y S3. Se utilizó el AWS Pricing Calculator que es un servicio proporcionado por la propia plataforma, obteniendo los siguientes resultados de esa cotización:

Resumen de la estimación		
Costo inicial	Costo mensual	Costo total de 12 months
0,00 USD	185,07 USD	2220,84 USD
		Incluye el costo inicial

Estimación detallada

Nombre	Grupo	Región	Costo inicial	Costo mensual
Amazon Redshift	No se ha aplicado ningún grupo	US East (Ohio)	0,00 USD	182,50 USD
Estado: - Descripción: Resumen de la configuración: Nodos (1), Tipo de instancia (dc2.large), Utilización (solo bajo demanda) (100 %Utilized/Month), Modelo de precios (OnDemand)				
Amazon Simple Storage Service (S3)	No se ha aplicado ningún grupo	US East (Ohio)	0,00 USD	2,57 USD
Estado: - Descripción: Resumen de la configuración: Almacenamiento de S3 Estándar (100 GB por mes), Datos devueltos por S3 Select (100 GB por mes), Datos escaneados por S3 Select (100 GB por mes)				

Costo por desarrollador

En el proyecto se contará con 4 ingenieros que estarán a cargo de todas las actividades y tareas del mismo. Se ha acordado que se pagará \$10 por hora, por un tiempo máximo de 165 días.

Costo por desarrollo de proyecto = Cantidad de ingenieros X \$10 por hora X 165 días

Costo por desarrollo de proyecto = 4 ingenieros X \$10.00 X 8 horas X 165 días

Costo por desarrollo de proyecto = \$52,800

El costo total por el desarrollo de este proyecto es de:

Coste del proyecto = AWS + Costo por desarrollador

Coste del proyecto = \$2220,84 + \$52,800

Coste del proyecto = \$55.020,84

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

a) Metodología de trabajo

Para llevar a cabo la solución que espera Inversiones Digitales, se han definido cuatro etapas que ayudarán a conseguir la mejor solución posible.

En primera parte hacer un estudio meticuloso de los datos que posee la empresa, posteriormente con la información obtenida buscar construir un modelo dimensional que mejor se adapte a las necesidades de la empresa, cuando el modelo dimensional se tenga plasmado se procederá a definir los procesos ETL con los datos obtenidos, y estos datos será procesados mediante Talend Open Studio (OTS), para el tratamiento del Data Lake y almacenamiento de datos se utilizará Amazon Web Services (AWS) y finalmente se realizará el diseño de tableros para la representación de los datos ya procesados mediante la herramienta Microsoft Power BI.

Estudio del origen de los datos.

La empresa Inversiones Digitales se apoya de la plataforma Odoos 16 para llevar el control de todos los movimientos (ingresos, egresos, activos) de cada una de las sucursales que posee.

Este sistema se instala en el ordenador al mismo tiempo se instala pgAdmin 4 que es con el que se gestionará la base de datos de PostgreSQL, en el proceso de instalación pide crear la base de datos donde se almacenará toda la data generada.

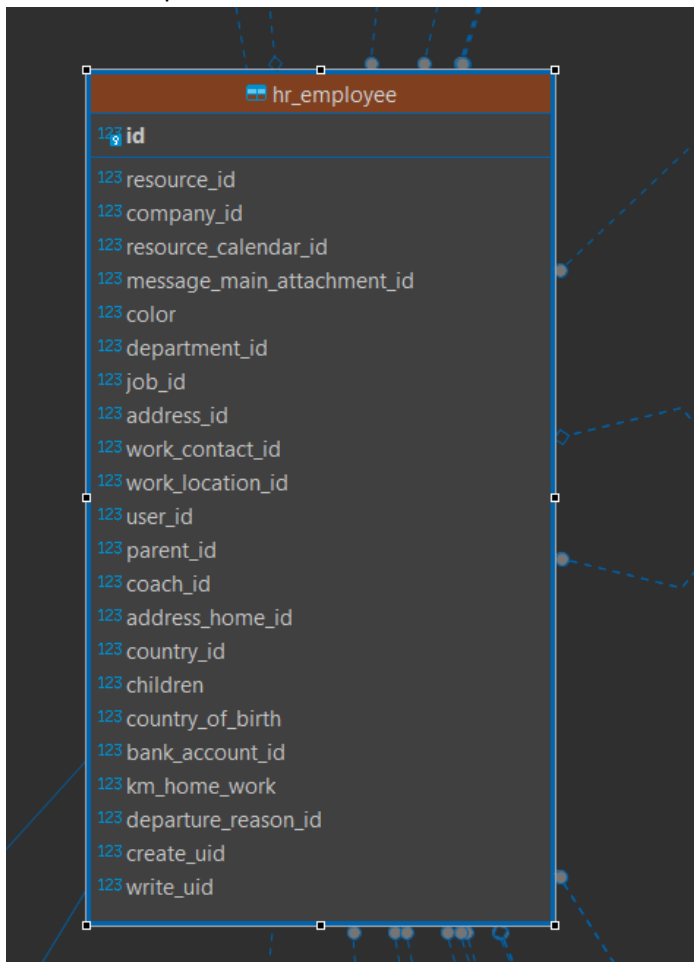
Posteriormente a la instalación y asegurarse que la instalación ha sido exitosa, se generan 771 tablas de las cuales sólo se ocuparán las relacionadas a la información de los activos fijos para generar el modelo dimensional.

Una parte importante en esta etapa es encontrar en qué tablas se almacena toda la data que interesa analizar.

A continuación, se muestran las tablas de la base de datos que corresponde a lo mencionado anteriormente.

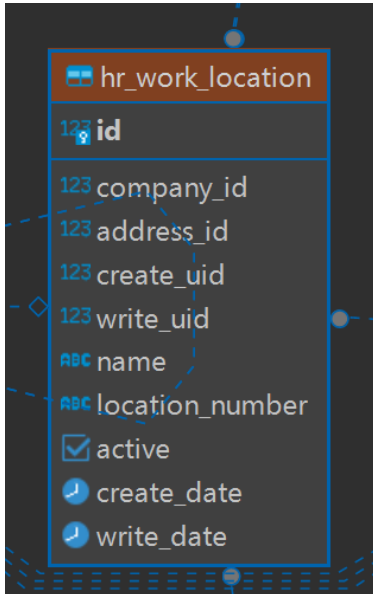
Hr_employee

Contiene parte de la información de los empleados en la empresa



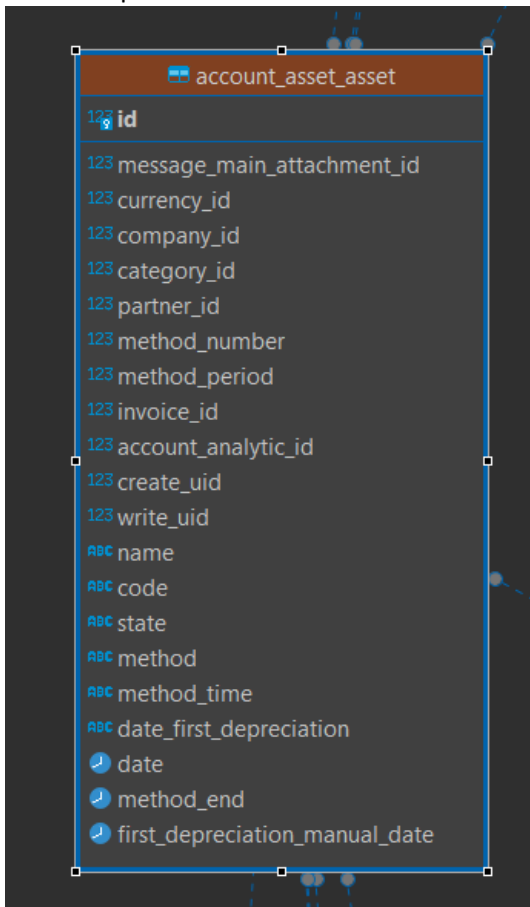
Hr_work_location

Contiene el restante de la información sobre la localización de los empleados de la empresa



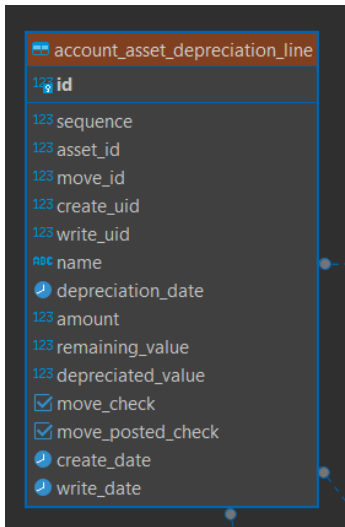
Account_asset_asset

Contiene parte de la información relevante sobre la depreciación de los activos



Account_asset_depreciation_line

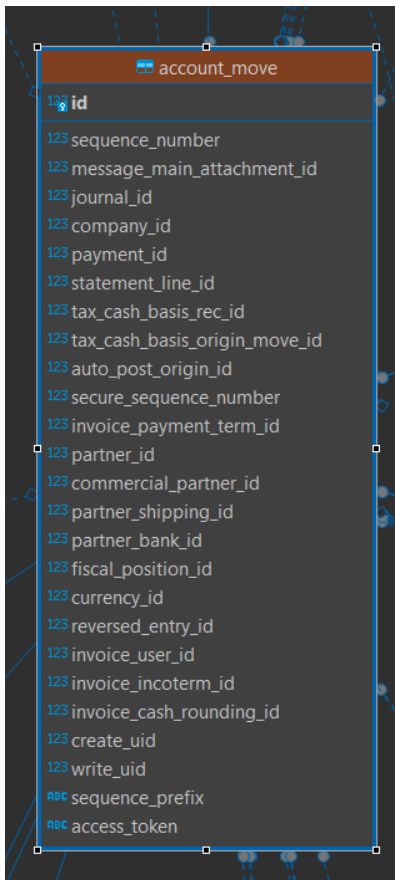
Contiene el resto de la información necesario sobre los valores de la depreciación de los activos



account_asset_depreciation_line	
PK	id
	sequence
	asset_id
	move_id
	create_uid
	write_uid
ASC	name
	depreciation_date
	amount
	remaining_value
	depreciated_value
	<input checked="" type="checkbox"/> move_check
	<input checked="" type="checkbox"/> move_posted_check
	create_date
	write_date

Account_move

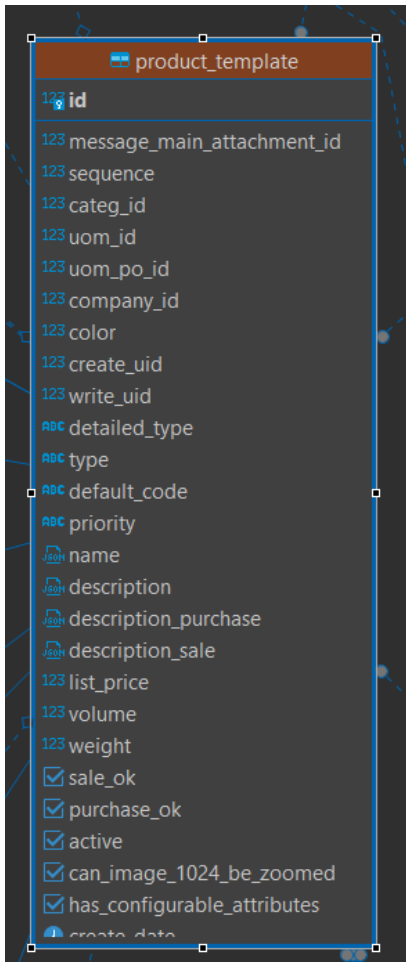
Contiene parte de la información necesaria para el análisis de compra de activos de proveedores y venta de activos con los clientes



account_move	
PK	id
	sequence_number
	message_main_attachment_id
	journal_id
	company_id
	payment_id
	statement_line_id
	tax_cash_basis_rec_id
	tax_cash_basis_origin_move_id
	auto_post_origin_id
	secure_sequence_number
	invoice_payment_term_id
	partner_id
	commercial_partner_id
	partner_shipping_id
	partner_bank_id
	fiscal_position_id
	currency_id
	reversed_entry_id
	invoice_user_id
	invoice_incoterm_id
	invoice_cash_rounding_id
	create_uid
	write_uid
ASC	sequence_prefix
ASC	access_token

Product_template

Contiene la información de los activos que maneja la empresa



Creación del modelo dimensional.

Para poder crear el modelo dimensional se escogió la propuesta que nos da a conocer Ralph Kimball, para lograr el objetivo de obtener información acerca de cuáles son las necesidades por las que está pasando Inversiones Digitales, y si la data generada en el sistema cumple o llena con las necesidades.

Este proceso para diseñar el modelo dimensional cuenta con los siguientes pasos:

1. Identificar los requisitos del proceso de negocio.

En esta etapa se identifica el proceso de negocio a llevar a cabo, para poder crear el modelo dimensional, se recogen los requisitos necesarios del proceso de negocio y se lleva a cabo la creación del modelado dimensional.

2. Identificar la granularidad.

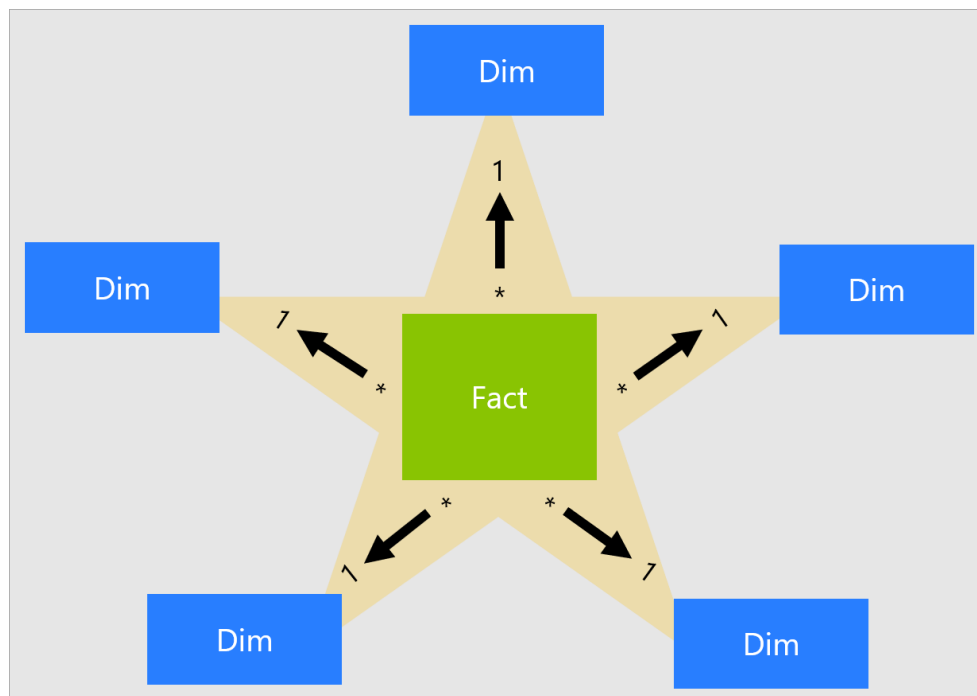
Se identificará la granularidad que tendrá la tabla de hechos y el proceso de negocio, en esta etapa se deberá identificar la tabla de hechos y posibles dimensiones y métricas.

3. Identificar las dimensiones.

Una vez determinado la granularidad del modelo, se identifican las dimensiones que en definitiva formarán parte del diseño para esta granularidad. Para poder identificar de una manera rápida las dimensiones, responden a las siguientes preguntas: ¿Qué?, ¿Cuándo?, ¿Dónde?, ¿Quién?, ¿Cómo? Y ¿Por qué? De los hechos.

4. Identificar las medidas o métricas.

Durante esta última etapa del modelo dimensiona, se identifican las métricas que irán incluidas en el modelo dimensional.



Procesos ETL

Para poder llevar toda la información necesaria desde el sistema transaccional con el que cuenta la empresa hasta el Data Warehouse, se deben diseñar los procesos encargados de la extracción, transformación y carga de datos (ETL).

Primordialmente se realizará la construcción de los procesos mencionados anteriormente, leyendo la data y comprendiendo el significado de cada dato extraído. En el caso de Inversiones Digitales, la data es almacenada en la base de datos PostgreSQL, se generan una infinidad de tablas de las cuales se seleccionan las apropiadas para el análisis de este proyecto. Luego de haber extraído de qué tablas y qué datos usaremos serán almacenados en el Data Lake, para ello se formarán 3 fases llamadas: Raw, Process y Access, donde en la carpeta Raw será almacenado el Data Lake.

Posterior a la extracción, tenemos toda la data lista para ser manipulada, pasa por el proceso de limpieza de datos, tratamiento de datos nulos y si se requiere de algún formato que se le necesite dar a los datos se realiza en esta parte.

En esta fase los datos se adaptan al modelo dimensional que se ha creado con anterioridad y estos se almacenan en la parte de Process, en esta fase se pueden aún más mejorar el estado de los datos para que se presenten con una mejor integridad.

Y por último se encuentra la fase de Access que es donde se almacena la data ya limpia y adaptada al modelo dimensional, y es de esta fase que se extraen los datos para ser utilizados en las herramientas de AWS.

Presentación de los datos.

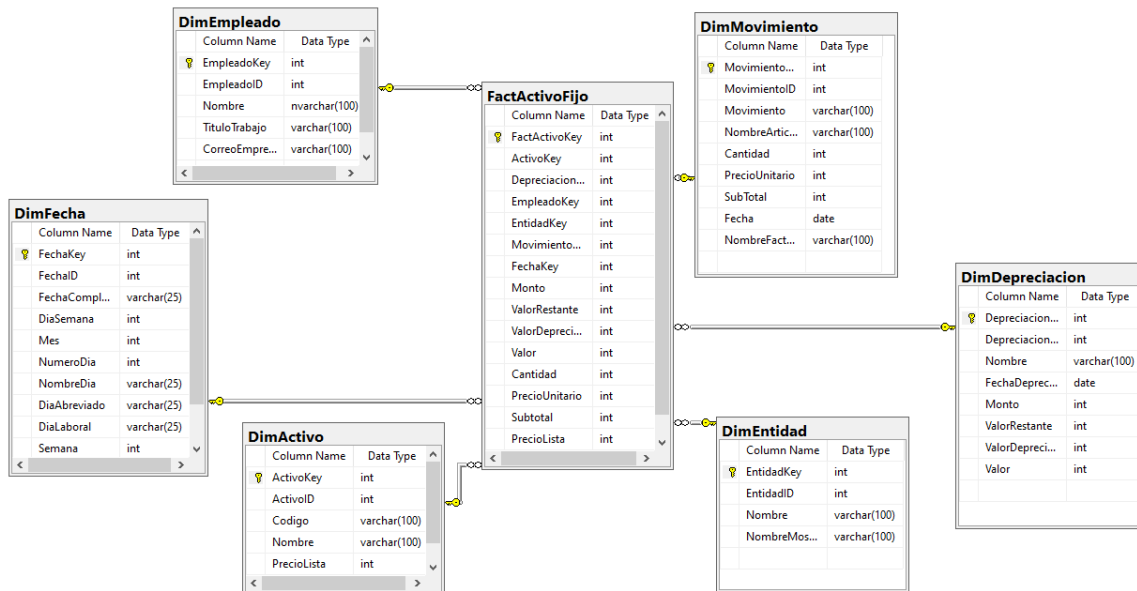
La representación de los datos ya procesados es realizada mediante tableros creados en Power BI, estos tableros contienen la información que se obtuvo del Data Warehouse, esta información es presentada de manera interactiva y entendible para el usuario final, con la intención de brindar una guía para la toma de decisiones de la alta gerencia.

b) Descripción de la propuesta de solución

Para la solución del proceso de activo fijo de la empresa Inversiones Digitales se ha identificado el siguiente conjunto de tablas del sistema transaccional, las cuales son conocidas como el Data Set las cuales se utilizarán como la fuente de datos para el Data Warehouse. Las tablas identificadas son las siguientes:



Con base en el diagrama ER, hemos propuesto el siguiente modelo dimensional, compuesto por 1 tabla de hechos y 6 dimensiones, siguiendo el Esquema de Estrella; como se muestra en la imagen



c) Descripción de la tecnología a utilizar

Para poder llevar a cabo esta solución, se han utilizado diversas herramientas que ayudan a lograr el objetivo de crear un Data Warehouse, en este apartado se describirán las herramientas utilizadas desde el momento de extraer la data desde el transaccional hasta la presentación de tableros con la extracción del análisis que se quiere llevar a cabo de acuerdo a la necesidad que ha presentado Inversiones Digitales.

Odoo

Es un sistema ERP (Enterprise Resource Planning) de código abierto que proporciona una suite de aplicaciones empresariales integradas. La versión "Community" es la edición gratuita y de código abierto de Odoo, que brinda a las empresas acceso a un conjunto básico de módulos para gestionar diversas funciones empresariales.

Características Principales:

- 1. Módulos Integrados:** Odoo Community 16 incluye una variedad de módulos que abarcan áreas como contabilidad, ventas, inventario, recursos humanos, CRM (Customer Relationship Management), fabricación, y más. Estos módulos están diseñados para trabajar de manera integrada, lo que permite una gestión eficiente de las operaciones comerciales.
- 2. Interfaz de Usuario Moderna:** Odoo se caracteriza por su interfaz de usuario intuitiva y moderna, lo que facilita la navegación y el uso para los usuarios finales.
- 3. Personalización:** La plataforma permite cierto grado de personalización para adaptarse a las necesidades específicas de cada empresa. Puedes agregar módulos adicionales según tus requisitos y personalizar flujos de trabajo.
- 4. Comunidad Activa:** Odoo cuenta con una comunidad activa de usuarios y desarrolladores que contribuyen con extensiones, mejoras y soporte. Esto crea un entorno colaborativo y favorece el intercambio de conocimientos.
- 5. Gestión de Proyectos y Tareas:** Incluye funcionalidades para la gestión de proyectos y tareas, facilitando la colaboración y el seguimiento del progreso de los proyectos.
- 6. eCommerce Integrado:** Odoo Community 16 proporciona una solución de comercio electrónico integrada que permite a las empresas gestionar sus operaciones de venta en línea.
- 7. Multiempresa:** La plataforma es capaz de gestionar múltiples empresas desde una sola instancia, lo que resulta útil para organizaciones con varias entidades comerciales.

Limitaciones:

- Aunque la versión Community es gratuita, algunas características avanzadas y módulos específicos pueden estar disponibles solo en la versión Enterprise de Odoo, que es de pago.
- La implementación y configuración inicial pueden ser desafiantes para usuarios sin experiencia en sistemas ERP, y puede ser beneficioso contar con la asistencia de expertos durante este proceso.

- La comunidad proporciona soporte principalmente a través de foros y documentación, lo que significa que puede haber limitaciones en comparación con el soporte premium disponible en la versión Enterprise.

Talend Open Studio

Esta herramienta de integración de datos de código abierto que proporciona un entorno gráfico para el diseño, la implementación y la gestión de procesos de integración de datos. Está desarrollado por Talend, empresa especializada en soluciones de integración de datos y gestión de datos maestros.

Las características principales son las siguientes:

1. Diseño gráfico de procesos:

TOS permite a los usuarios diseñar procesos de integración de datos utilizando una interfaz gráfica intuitiva. Puedes arrastrar y soltar componentes para construir flujos de trabajo de integración.

2. Conectividad y compatibilidad:

Ofrece conectividad con una amplia variedad de fuentes y destinos de datos, incluyendo bases de datos, servicios web, archivos planos, etc.

3. Transformación de datos:

Incluye conjunto completo de herramientas para transformar y manipular datos durante el proceso de integración.

4. Orquestación de procesos:

Permite orquestación de procesos complejos mediante la coordinación de tareas y flujos de trabajo.

5. Monitorización y gestión:

Proporciona capacidades para monitorear y gestionar los procesos de integración. Esto incluye la capacidad de detectar errores, realizar seguimientos y realizar ajustes según sea necesario.

6. Código abierto:

Al ser de código abierto, es gratuito para descargar y utilizar. Esto ha contribuido a su popularidad y ha llevado a una comunidad activa de usuarios y desarrolladores.

8. Ampliable y personalizable:

La arquitectura abierta permite a los usuarios extender y personalizar las funcionalidades según sus necesidades específicas.

Limitaciones:

- Aunque es una herramienta potente y versátil, la curva de aprendizaje puede ser empinada para aquellos que son nuevos en la integración de datos y ETL (Extract, Transform, Load).
- A medida que los proyectos de integración crecen en complejidad, algunos usuarios pueden encontrar limitaciones en la escalabilidad de la versión gratuita y optar por soluciones más avanzadas de Talend que ofrecen soporte y funcionalidades adicionales.

Amazon Web Services

Amazon Web Services (AWS) es una plataforma de servicios en la nube ofrecida por Amazon.com. Lanzada en 2006, AWS se ha convertido en uno de los principales proveedores de servicios en la nube a nivel mundial. Ofrece una amplia gama de servicios, desde cómputo en la nube hasta almacenamiento, bases de datos, machine learning, inteligencia artificial, análisis de datos y más.

Características Principales:

- 1. Escalabilidad y Flexibilidad:** AWS permite escalar recursos de manera flexible según las necesidades del usuario. Esto facilita a las empresas ajustar sus recursos de computación y almacenamiento de manera dinámica, lo que es esencial para gestionar cargas de trabajo variables.
- 2. Variedad de Servicios:** AWS proporciona una extensa variedad de servicios que cubren prácticamente todos los aspectos de la infraestructura de TI y el desarrollo de aplicaciones. Esto incluye servicios de cómputo, almacenamiento, bases de datos, redes, analítica, inteligencia artificial, aprendizaje automático y más.
- 3. Seguridad:** AWS pone un fuerte énfasis en la seguridad, ofreciendo características como cifrado de datos, autenticación multifactor, gestión de identidades y acceso, así como cumplimiento de normativas y estándares de seguridad.
- 4. Globalidad:** AWS opera en múltiples regiones alrededor del mundo, permitiendo a las empresas desplegar aplicaciones y datos en ubicaciones estratégicas para mejorar el rendimiento y la disponibilidad.
- 5. Pricing Flexible:** AWS utiliza un modelo de precios de pago por uso, lo que significa que los usuarios pagan solo por los recursos que consumen. Esto permite a las empresas controlar sus costos de manera más efectiva.
- 6. Integración y Ecosistema:** AWS ofrece integración con una amplia variedad de herramientas y servicios de terceros. Además, tiene un ecosistema robusto de partners y desarrolladores que ofrecen soluciones y servicios complementarios.
- 7. Soporte y Documentación:** AWS proporciona una amplia documentación, tutoriales y foros de soporte para ayudar a los usuarios a entender y utilizar eficientemente sus servicios. También ofrece planes de soporte técnico para empresas que requieren asistencia adicional.

Limitaciones:

- La complejidad de la plataforma puede resultar abrumadora para los usuarios nuevos, especialmente aquellos sin experiencia en la nube.
- El modelo de precios puede volverse complejo con la variedad de servicios ofrecidos, y los costos pueden aumentar si no se gestionan adecuadamente.
- La dependencia de la nube implica riesgos potenciales de disponibilidad y confidencialidad de datos.

Amazon Simple Storage Services (S3)

Amazon Simple Storage Service, comúnmente conocido como Amazon S3, es un servicio de almacenamiento en la nube ofrecido por Amazon Web Services (AWS). Lanzado en 2006, S3 es uno de los servicios más fundamentales y populares de AWS, diseñado para almacenar y recuperar cualquier cantidad de datos desde cualquier ubicación de la web.

Características Principales:

- 1. Almacenamiento Escalable:** Amazon S3 proporciona una solución de almacenamiento escalable y duradero que puede manejar grandes cantidades de datos. Los usuarios pueden almacenar desde unos pocos gigabytes hasta varios terabytes o más de información.
- 2. Durabilidad y Disponibilidad:** S3 está diseñado para ser altamente duradero y disponible. Los datos almacenados en S3 se replican automáticamente en varias ubicaciones dentro de una región de AWS, garantizando durabilidad y resistencia a fallas.
- 3. Facilidad de Uso:** La interfaz de usuario de Amazon S3 es intuitiva y fácil de usar. Los usuarios pueden gestionar sus recursos de almacenamiento a través de la consola web de AWS, la interfaz de línea de comandos (CLI) o mediante SDKs (kits de desarrollo de software) disponibles para varios lenguajes de programación.
- 4. Seguridad:** S3 ofrece opciones robustas de seguridad, incluyendo controles de acceso a nivel de objeto y políticas de bucket, cifrado de datos en reposo y en tránsito, y la integración con AWS Identity and Access Management (IAM) para la gestión de identidades y accesos.
- 5. Escala de Costos:** S3 utiliza un modelo de precios de pago por uso, lo que significa que los usuarios pagan solo por la cantidad de almacenamiento que consumen. Además, ofrece opciones de almacenamiento de bajo costo para datos menos frecuentemente accesibles.
- 6. Integración con Otros Servicios:** S3 se integra estrechamente con otros servicios de AWS, como AWS Lambda, AWS Glacier (para almacenamiento de archivos a largo plazo), Amazon CloudFront (para distribución de contenido) y más, lo que permite a los usuarios construir soluciones complejas y altamente escalables.
- 7. Versionado y Control de Acceso:** S3 ofrece funciones de versionado de objetos, lo que permite a los usuarios recuperar versiones anteriores de un objeto. Además, proporciona capacidades avanzadas de control de acceso, lo que permite definir políticas específicas para la gestión de permisos.

Limitaciones:

- Aunque el modelo de precios es flexible, los costos pueden aumentar si no se gestionan adecuadamente, especialmente con respecto a transferencias de datos y solicitudes.
- Algunas configuraciones avanzadas pueden requerir un conocimiento más profundo de los servicios de AWS.

Amazon Redshift

Amazon Redshift es un servicio de almacenamiento y procesamiento de datos basado en la nube ofrecido por Amazon Web Services (AWS). Está diseñado específicamente para análisis de datos y almacenamiento de grandes conjuntos de datos. Redshift utiliza la arquitectura de almacenamiento de datos en columna y está optimizado para consultas analíticas de alto rendimiento.

Características Principales:

1. **Almacenamiento de Datos en Columna:** Redshift almacena datos en formato de columna en lugar de formato de fila, lo que mejora significativamente el rendimiento de las consultas analíticas. Este enfoque es particularmente eficiente para operaciones de agregación y análisis en grandes conjuntos de datos.
2. **Escalabilidad:** Redshift es altamente escalable, permitiendo a los usuarios aumentar o disminuir la capacidad de almacenamiento y computación según sea necesario. Esto es crucial para empresas que manejan grandes cantidades de datos y experimentan picos de carga.
3. **Rendimiento de Consulta:** Ofrece un rendimiento de consulta rápido y eficiente, lo que facilita la ejecución de consultas analíticas complejas en grandes volúmenes de datos. La arquitectura masivamente paralela contribuye a un procesamiento rápido de consultas.
4. **Integración con Herramientas de BI:** Se integra fácilmente con diversas herramientas de Business Intelligence (BI) y visualización de datos, como Tableau, Looker y otras, lo que facilita el análisis y la presentación de datos.
5. **Seguridad:** Proporciona opciones avanzadas de seguridad, incluyendo cifrado de datos en tránsito y en reposo, control de acceso basado en políticas y la capacidad de integrarse con AWS Identity and Access Management (IAM).
6. **Copia y Carga de Datos:** Facilita la carga masiva de datos a través de herramientas como AWS Data Pipeline y AWS Glue. Además, es posible cargar datos de manera continua desde fuentes externas.
7. **Precio Basado en Consumo:** Al igual que muchos servicios de AWS, Redshift utiliza un modelo de precios de pago por uso, lo que significa que los usuarios pagan por la capacidad de almacenamiento y cómputo que consumen.

Limitaciones:

- Aunque es escalable, la adición de nodos adicionales puede aumentar los costos, y es importante optimizar y monitorear el rendimiento para garantizar la eficiencia en términos de costo.
- Las operaciones de carga y descarga de datos pueden experimentar tiempos más largos en comparación con otras bases de datos debido a la naturaleza columnar de Redshift.

Power BI

Es una plataforma de análisis de negocios desarrollada por Microsoft. Ofrece una variedad de herramientas para visualización de datos, análisis empresarial y creación de informes interactivos. Power BI permite a los usuarios conectarse a diversas fuentes de datos, transformar y modelar esos datos, y luego crear visualizaciones impactantes para obtener insights significativos.

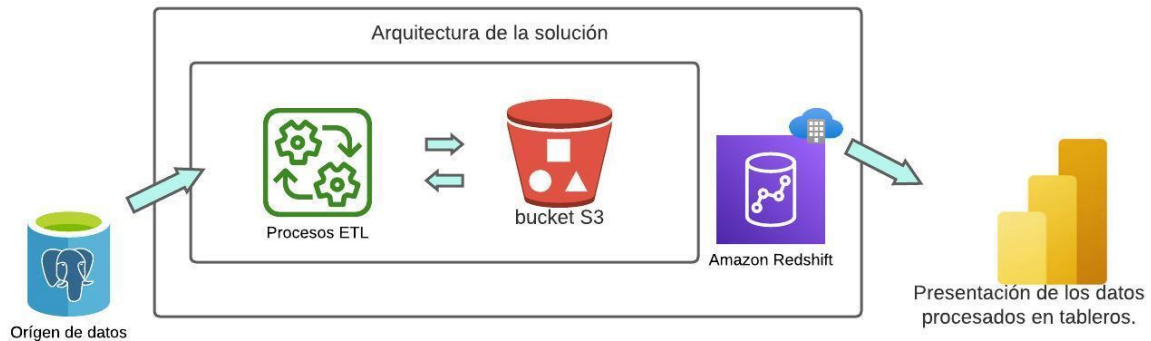
Características Principales:

- 1. Conectividad a Diversas Fuentes de Datos:** Power BI permite a los usuarios conectarse a una amplia variedad de fuentes de datos, incluyendo bases de datos locales, servicios en la nube, archivos de Excel, servicios web y más. Esto facilita la consolidación de datos dispersos.
- 2. Transformación y Modelado de Datos:** Los usuarios pueden transformar y modelar datos utilizando Power Query y Power Pivot, dos herramientas integradas en Power BI. Estas herramientas permiten la limpieza, la transformación y la combinación de datos para prepararlos para el análisis.
- 3. Visualizaciones Interactivas:** Power BI ofrece una amplia gama de visualizaciones interactivas, como gráficos, tablas dinámicas, mapas y más. Los usuarios pueden crear paneles interactivos y tableros de control para presentar datos de manera atractiva.
- 4. Integración con Microsoft Office y Azure:** Power BI se integra estrechamente con otras herramientas de Microsoft, como Excel, y con servicios en la nube de Azure. Esto facilita la colaboración y la conexión de datos desde diferentes plataformas.
- 5. Publicación y Compartición de Informes:** Los informes y paneles de Power BI pueden ser publicados en el servicio en la nube de Power BI, permitiendo a los usuarios compartir y colaborar en tiempo real. También es posible incorporar informes interactivos en otras aplicaciones y sitios web.
- 6. Capacidades de Análisis Avanzado:** Power BI incluye capacidades avanzadas de análisis, como la creación de medidas personalizadas utilizando el lenguaje de fórmulas DAX (Data Analysis Expressions) y la capacidad de crear modelos de datos complejos.
- 7. Seguridad y Cumplimiento:** Power BI ofrece características de seguridad y cumplimiento, incluyendo la capacidad de definir roles y permisos, así como la integración con Azure Active Directory para la gestión de identidades.

Limitaciones:

- Algunas funciones avanzadas pueden requerir conocimientos técnicos y experiencia en análisis de datos.
- Aunque hay una versión gratuita de Power BI, ciertas características avanzadas y capacidades de colaboración están disponibles en las versiones de pago.

d) Diagrama arquitectónico de la solución



e) Descripción de cada componente de la solución

Origen de datos:

La base de datos del sistema transaccional almacena de manera integral toda la información vinculada a las operaciones comerciales, incluyendo ventas, compras, clientes, productos, direcciones, empleados y activos de la empresa, llevadas a cabo a través de las diversas sucursales de Inversiones Digitales. Esta robusta base de datos en PostgreSQL no solo constituye una fuente completa de datos, sino también un punto de partida esencial para analizar cómo se almacenan los datos y para desarrollar la estrategia de diseño de los procesos ETL.

Procesos ETL:

Los procesos ETL se encargan de extraer la información esencial de la base de datos transaccional y trasladar el lago de datos mediante la suite Talend Open Studio (TOS). Una vez obtenidos los datos, los procesos de transformación se encargan de realizar modificaciones, sustituciones, reemplazos y limpieza de los datos. En la fase culminante, los datos son cargados a través de procesos específicos en la zona de Acceso del bucket S3.

Buckets S3:

Se utilizará el AWS S3 para el almacenamiento de los datos procedentes de la base de datos transaccional. En S3 se llevarán 3 almacenes de datos, los cuales cada uno almacenará la data desde la data cruda (almacenamiento RAW), la procesada (PROCESS) y la completamente transformada (ACCESS) mediante los procesos ETL que se implementen.

Amazon Redshift:

En Redshift, se creará el modelo dimensional, que consta de una tabla de hechos y dimensiones, para almacenar los datos previamente procesados en S3 mediante los procesos ELT. Redshift se desempeñará como nuestra base de datos en la nube, destacándose por su capacidad para almacenar y analizar extensas cantidades de registros y diversas fuentes de datos, es una excelente elección para el almacenaje de datos.

Power BI:

Esta etapa será el último paso del proyecto a realizar. Se realizará la conexión con Redshift para poder presentar los datos de manera gráfica ya que de esta forma las organizaciones analizan la información y llegar al punto importante que es la toma de decisiones.

Capítulo III Estrategia de Implementación de propuesta de solución

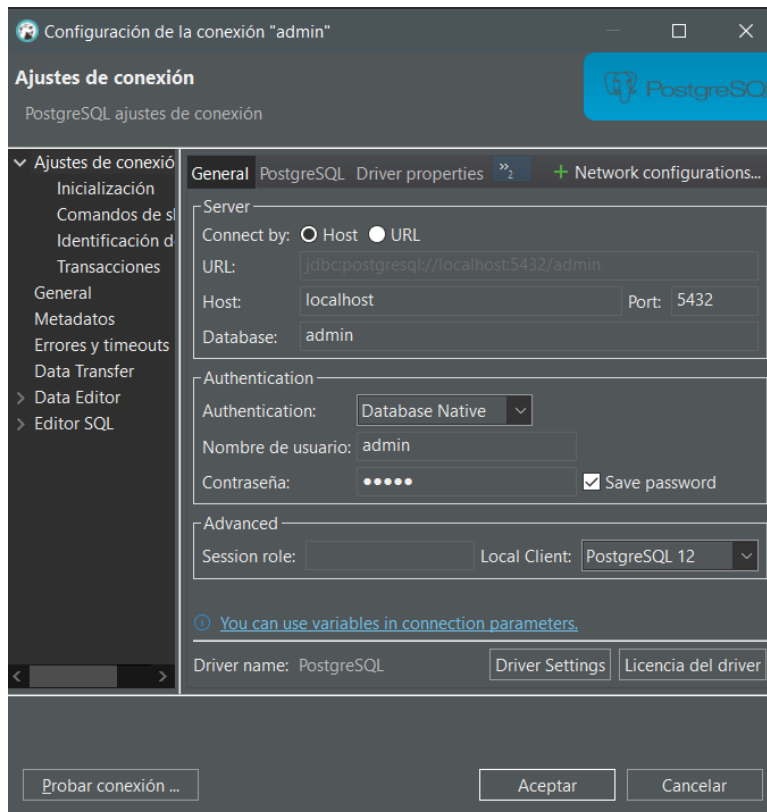
a) Estrategia de implementación

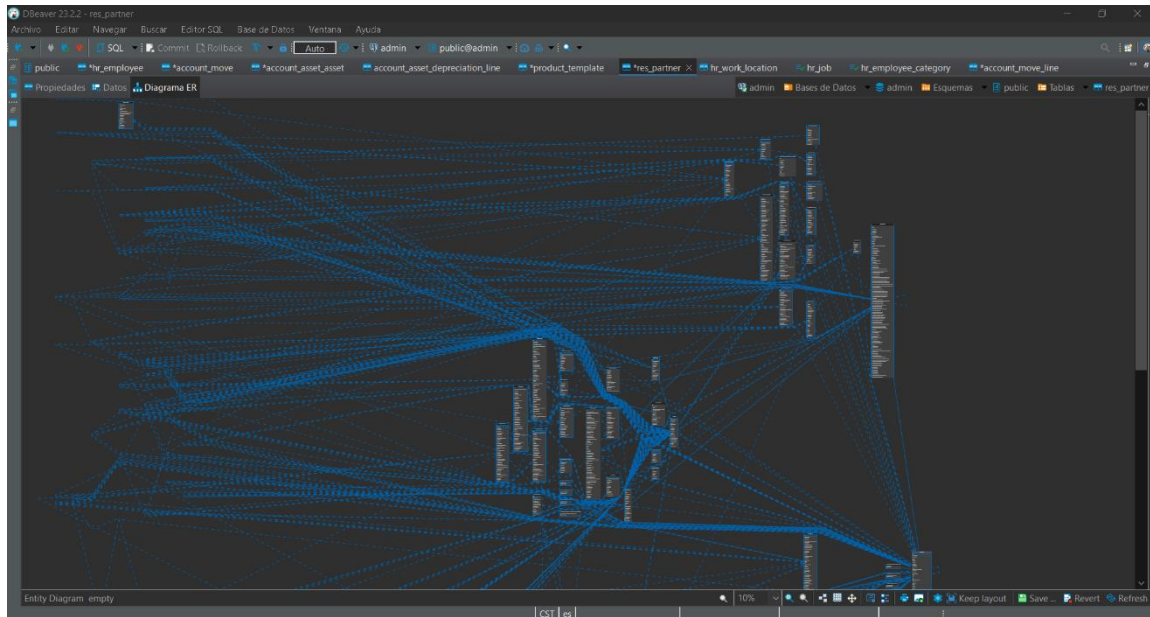
La estrategia por implementación será un producto por integración, en donde como se dijo la implementación se podrá integrar a otra solución que ya existe, que es el sistema transaccional que ya se tiene operando en la empresa de inversiones digitales que se integrada al módulo de contabilidad, que es el que ayuda con su información a la toma de decisiones por medio de la solución del dataWareHouse planteado.

La solución está compuesta por una secuencia de pasos en diferentes herramientas que nos llevarán a la solución final.

DBeaver

Esta herramienta se utilizo para el analisis y gestion de la base de datos por medio de la conexión de PostgreSQL donde se tiene la base de datos de inversiones digitales por medio del uso del módulo de contabilidad.





Conjunto de carpetas a utilizar para almacenamiento de los datos: Para almacenar toda la data recién extraída del transaccional hasta la data ya procesada se usarán tres carpetas donde se estructura la información:

Raw: En esta carpeta tendremos el data lake.

Process: En esta se almacenará la data procesada, donde eliminamos datos que no aportan utilidad, como transformación de valores nulos.

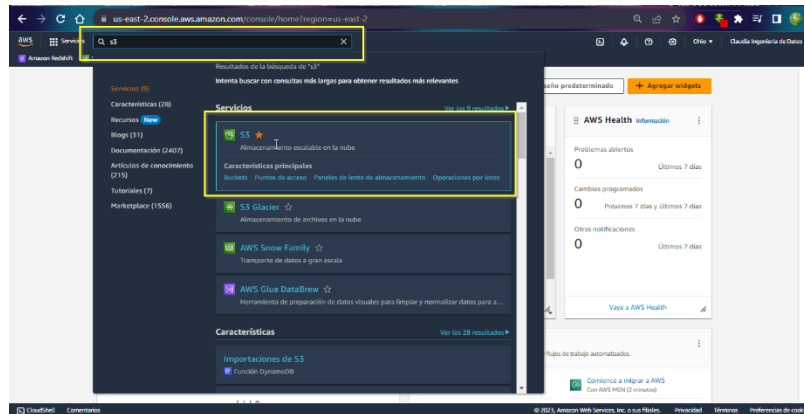
Access: Y en esta carpeta se almacena la data ya procesada que es la data final a mostrar mediante tableros.

Esta estructura de tablas se realizará en la herramienta S3 de AWS, para una mejor manipulación usaremos la herramienta S3 browser.

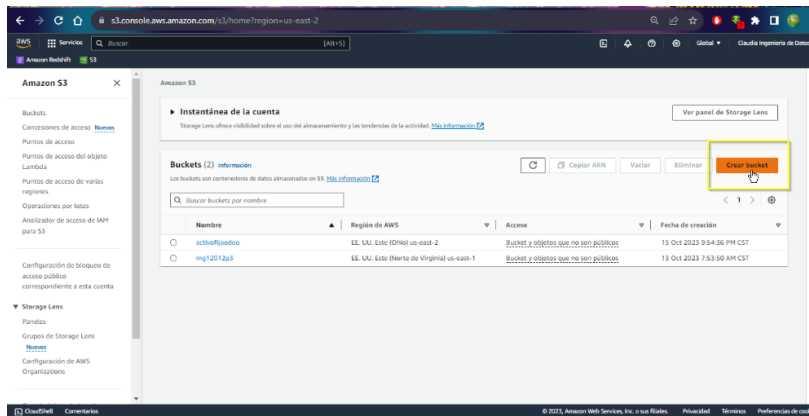
AWS S3

Para poder realizar la integración de servicios en la nube se utilizará Simple Storage Service mejor conocido como S3 de AWS. Para la creación de un bucket es necesario seguir los siguientes pasos:

1. Iniciar sesión en AWS y en la barra de búsqueda ingresamos S3 y damos click.



2. Damos clic en el botón "Crear bucket".



- Nos mostrará un panel de configuración general, ingresamos el nombre que poseerá el bucket, la región, en este caso seleccionamos la predeterminada que nos muestra, y damos clic en “Elegir el bucket”.

Amazon S3 > Buckets > Crear bucket

Crear bucket Información

Los buckets son contenedores de datos almacenados en S3. [Más información](#)

Configuración general

Nombre del bucket

El nombre del bucket debe ser único dentro del espacio de nombres global y seguir las reglas de nomenclatura del bucket. [Consulte las reglas para la asignación de nombres de buckets](#)

Región de AWS

EE. UU. Este (Ohio) us-east-2

Copiar la configuración del bucket existente: *opcional*
Solo se copia la configuración del bucket en los siguientes ajustes.

Elegir el bucket

En el apartado de “Control de versiones de buckets”, damos clic en el combo box de “Habilitar”.

Editar control de versiones de buckets Información

Control de versiones de buckets

El control de versiones es una forma de mantener múltiples variantes de un objeto dentro del mismo bucket. Puede utilizar el control de versiones para conservar, recuperar y restaurar todas las versiones de los objetos almacenados en su bucket de Amazon S3. Con el control de versiones, puede recuperarse con facilidad de las acciones involuntarias de los usuarios y de los errores en las aplicaciones. [Más información](#)

Control de versiones de buckets

Suspender
Esto suspende la creación de versiones de los objetos para todas las operaciones, pero conserva las versiones de los objetos existentes.

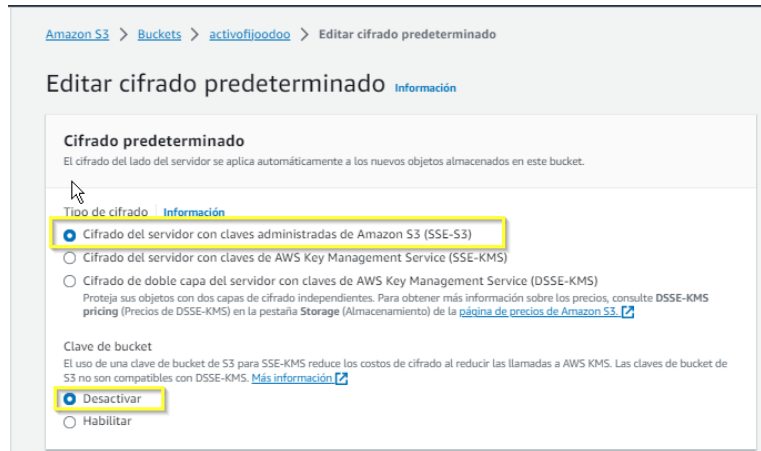
Habilitar

Eliminación de Multi-Factor Authentication (MFA)

Una capa adicional de seguridad que requiere la autenticación multifactor para cambiar la configuración del control de versiones del bucket y eliminar permanentemente las versiones de los objetos. Para modificar la configuración de las eliminaciones de MFA, utilice la CLI de AWS, el SDK de AWS o la API REST de Amazon S3. [Más información](#)

Deshabilitada

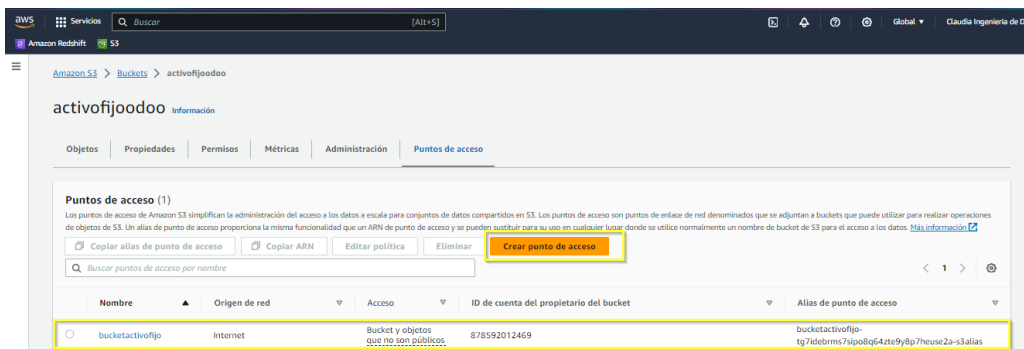
En el siguiente panel seleccionamos “cifrado del servidor con claves administradas” y “desactivar” en la clave del bucket.



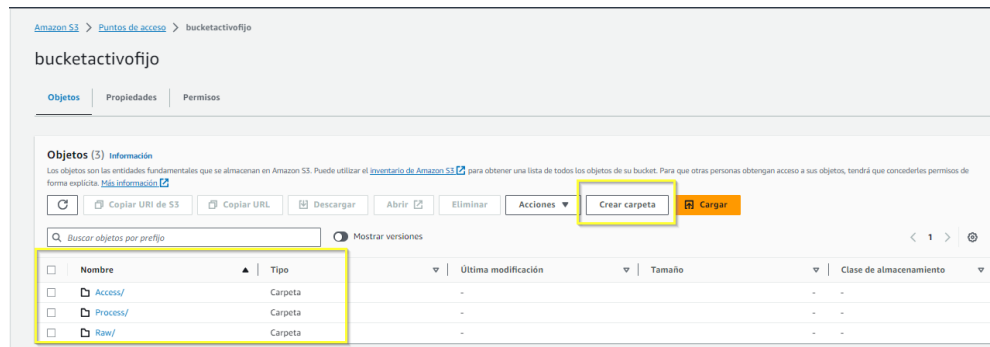
En el apartado de “Permisos”, configurar como activado el bloqueo de acceso al público.



Damos clic en el apartado de puntos de acceso y damos clic en “Crear punto de acceso” y nos creará el registro mostrado en la imagen.



Dentro del punto de acceso crearemos las carpetas con las que trabajaremos (Raw, Process, Access).



S3 Browser

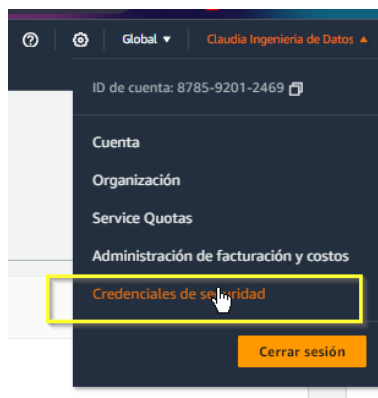
S3 Browser es un cliente de interfaz gráfica de usuario (GUI) diseñado para facilitar la gestión y el acceso a los servicios de almacenamiento en la nube que utilizan el protocolo de almacenamiento Simple Storage Service (S3) de Amazon Web Services (AWS). Este programa permite realizar diversas operaciones relacionadas con el almacenamiento en la nube, como cargar y descargar archivos, organizar carpetas, establecer permisos de acceso y más, utilizando una interfaz fácil de usar.

S3 Browser es especialmente útil para la interacción de los datos almacenados en la nube de manera más intuitiva. Proporciona funciones avanzadas para la gestión de objetos, la configuración de propiedades y la realización de tareas comunes de manera eficiente.

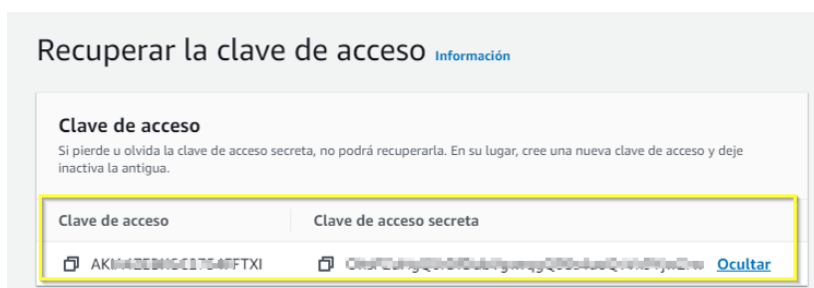
Para usar esta herramienta se deben seguir los siguientes pasos:

1. Primeramente, descargar la aplicación del sitio oficial: <https://s3browser.com/download.aspx>
2. Realizar la instalación, y posteriormente nos mostrará un formulario donde ingresaremos los datos de la cuenta a utilizar de AWS, previamente nos dirigimos a AWS para poder obtener las key de acceso.

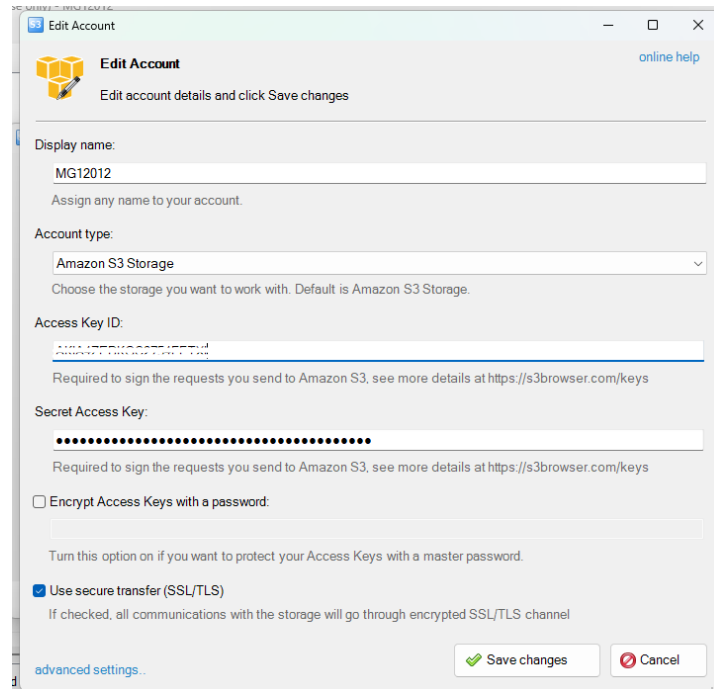
Damos clic en nuestro usuario y luego damos clic en “credenciales de seguridad”.



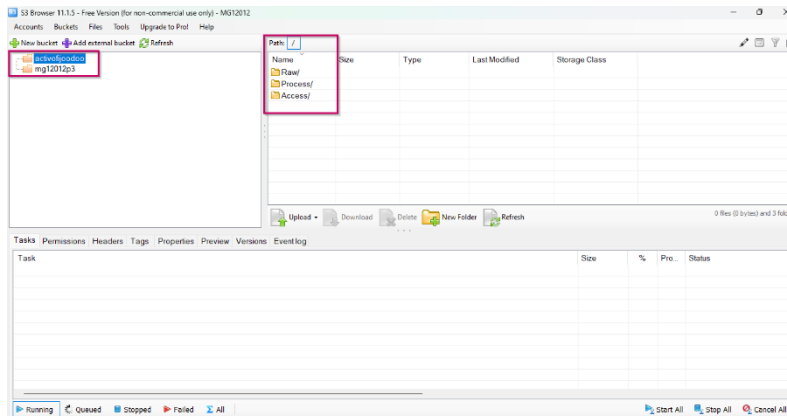
Y nos mostrará las siguientes claves de acceso, las copiamos y llevamos al formulario de S3 Browser.



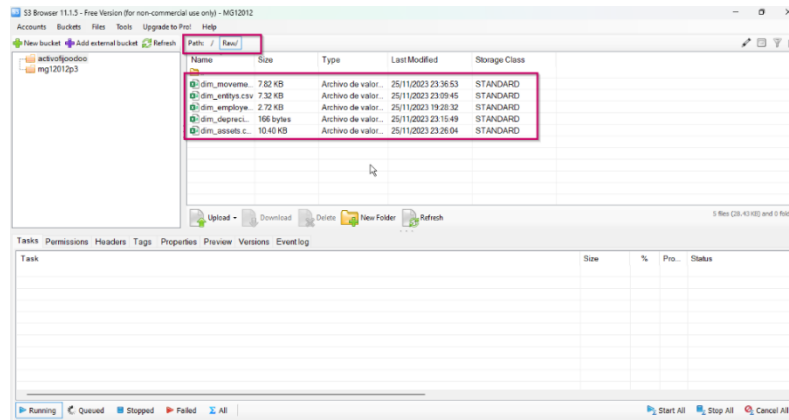
- Ingresamos el nombre, el tipo de cuenta, y las key generadas en la cuenta de AWS, guardamos los cambios, y con esto ya tendremos sincronizado AWS S3, cada vez que se realice un cambio desde esta herramienta se verá reflejado también en AWS S3.



- Luego de iniciar sesión con nuestra cuenta podremos añadir los buckets necesarios para ejecutar nuestras tareas, y dentro del bucket generamos la estructura de tablas que se mencionaron previamente, que será donde se almacenará toda la data, que va desde el data lake hasta la data procesada lista para utilizar.



5. Ingresando a cada una de las carpetas podremos visualizar nuestros archivos generados, mostrando el tamaño, la fecha y hora de actualización y el tipo de almacenamiento.



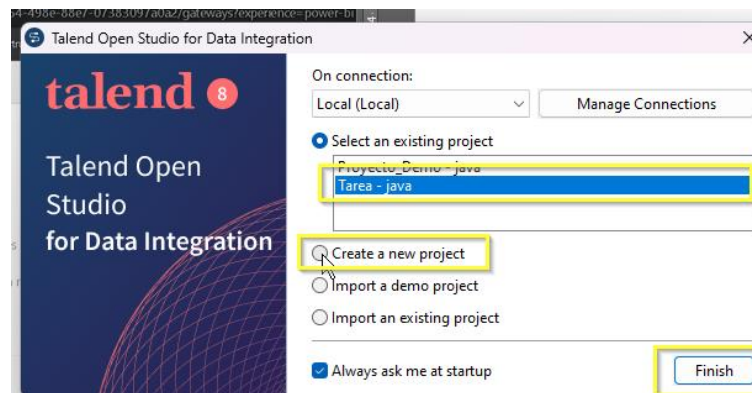
IAM

Después de haber creado el Bucket junto con su estructura de carpetas, el siguiente paso crucial implica la creación de un usuario IAM y la autorización de los correspondientes permisos de acceso al bucket. En caso de que el usuario IAM y el bucket de S3 pertenezcan a la misma cuenta de AWS, es factible otorgar acceso al usuario a una carpeta específica del bucket mediante la implementación de una política de IAM. A menos que la política de bucket niegue expresamente al usuario el acceso a la carpeta, no es necesario modificar la política de bucket siempre que la política de IAM garantice el acceso necesario. La política de IAM puede ser añadida a usuarios IAM individuales, o alternativamente, puede asociarse a un rol de IAM que varios usuarios tengan la capacidad de cambiar.

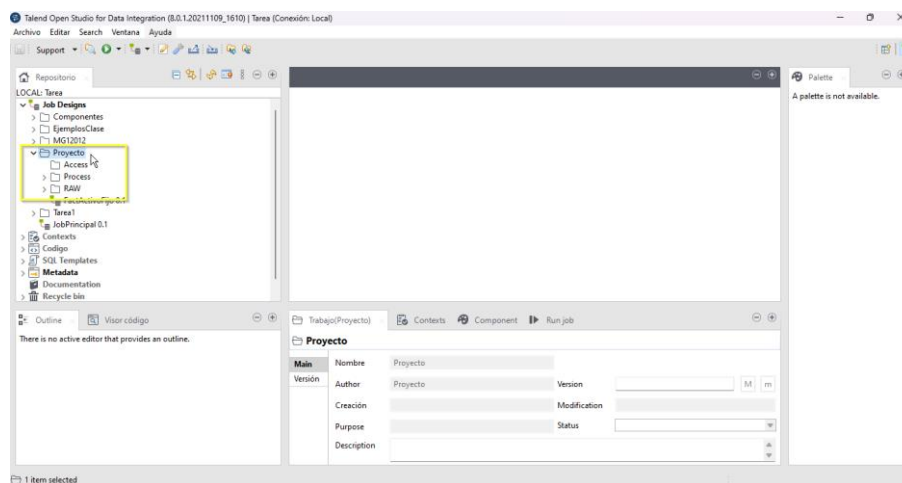
Talend Open Studio.

Luego de realizar la descarga de esta herramienta nos dirigimos a abrir la herramienta.

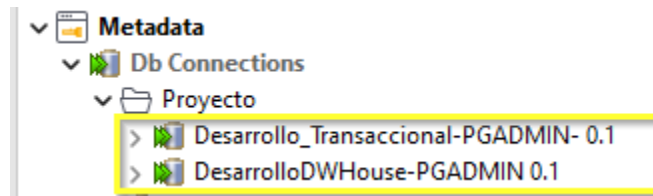
1. Damos clic en “Create a new project”, si se utiliza por primera vez, en este caso ya tenemos creado el proyecto, seleccionamos “Tarea” y damos clic en “Finish”.



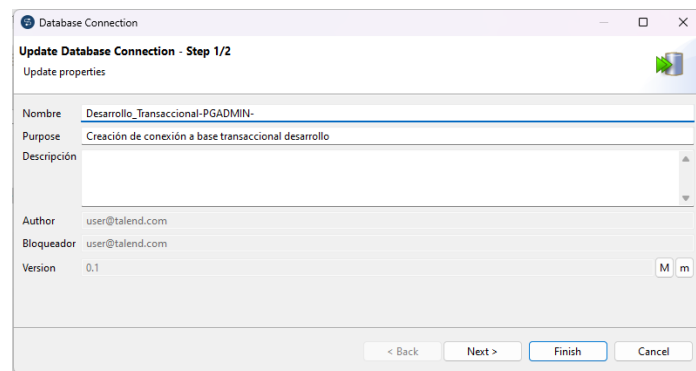
2. La ventana principal nos muestra el área de trabajo, y al lado izquierdo nos muestra los proyectos realizados con el desglose de carpetas con las que se estarán trabajando.



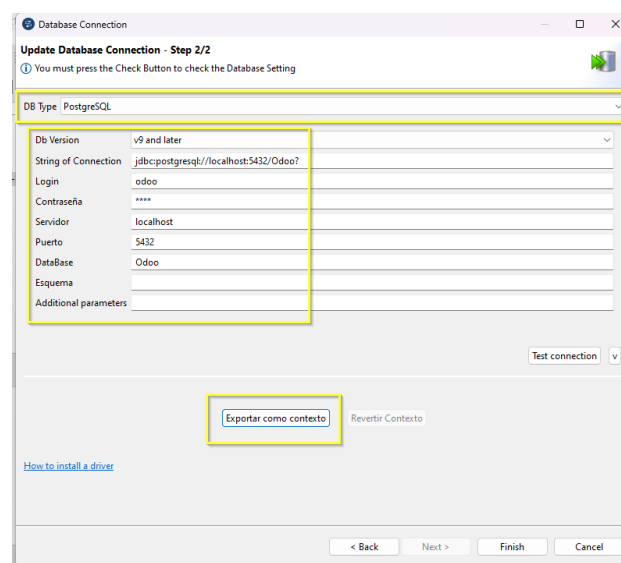
3. En el apartado de Metadata, creamos las conexiones, para el proyecto se creó una carpeta para llevar un mejor orden de las conexiones al transaccional y al Data Warehouse.



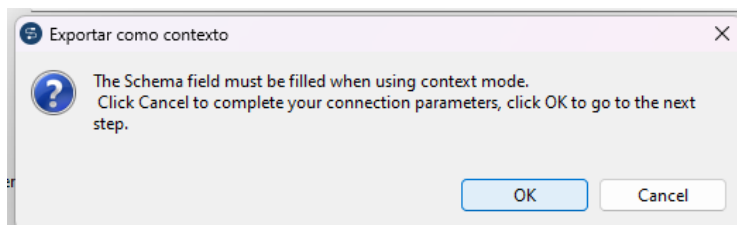
4. Para la conexión a la base transaccional asignamos un nombre y una breve descripción, damos clic en “Siguiente”.



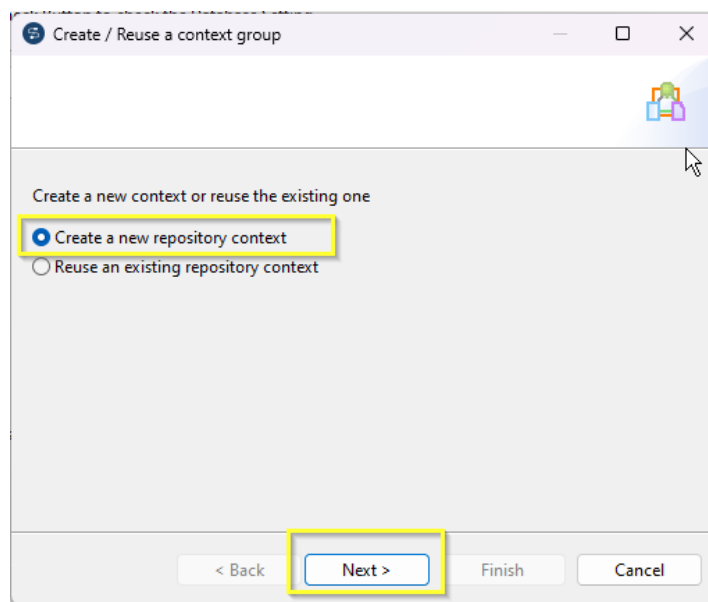
5. En el apartado de DB Type buscamos la base a la que queremos tener acceso, e ingresamos todos los datos necesarios en todos los campos solicitados, por último, damos clic en “Exportar como contexto”.



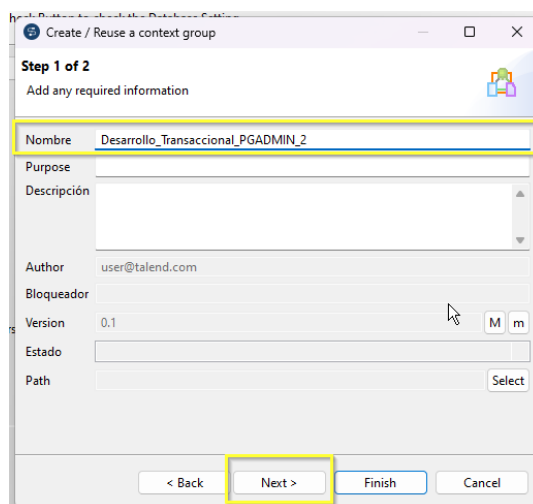
6. Damos clic en “Si”.



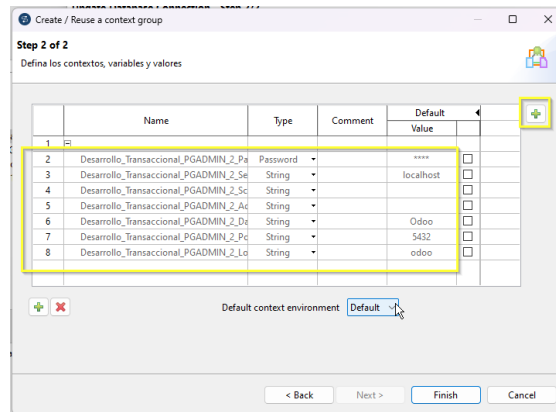
7. Damos clic en la primera opción y luego damos clic en “Next”.



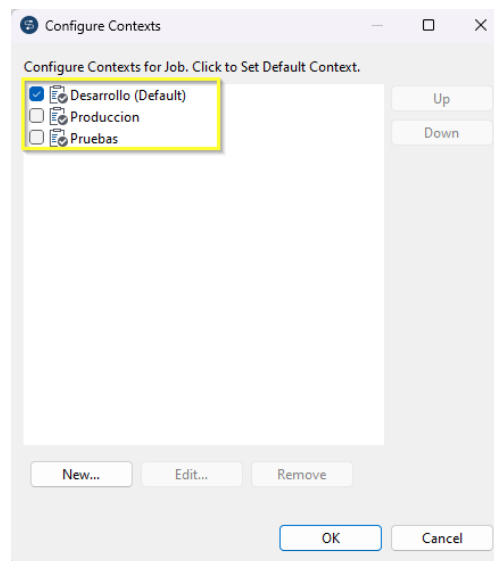
8. Ingresamos un nombre que identifique nuestro contexto y damos clic en “Next”.



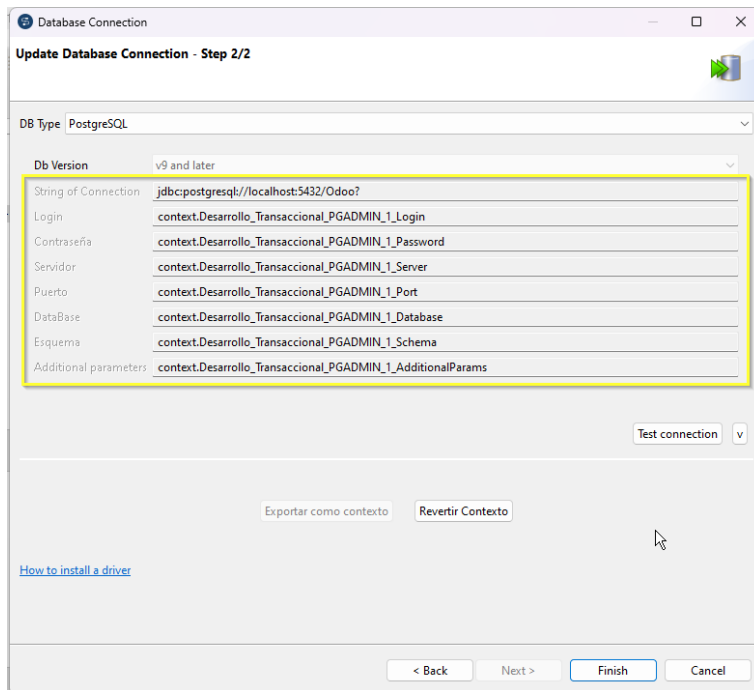
9. Nos aparecerá la siguiente pantalla, donde también se podrá configurar los contextos dando clic en la cruz verde.



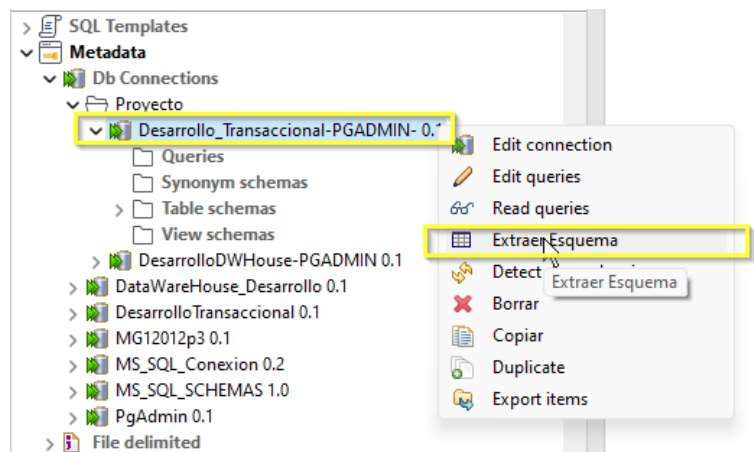
10. Configuramos los contextos y damos clic en “Ok” y por último en el botón “Finish” de la pantalla anterior.



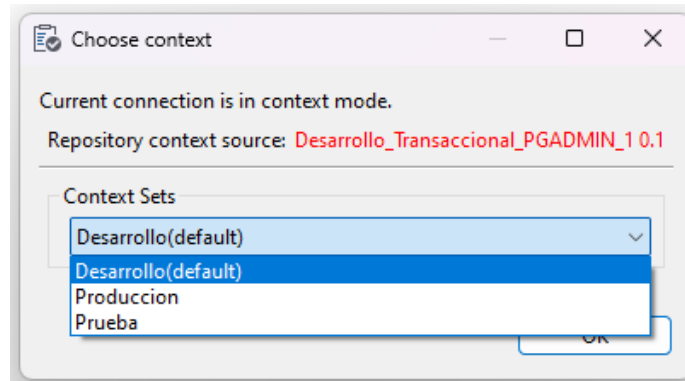
11. Nos quedará de la siguiente manera la configuración de los contextos.



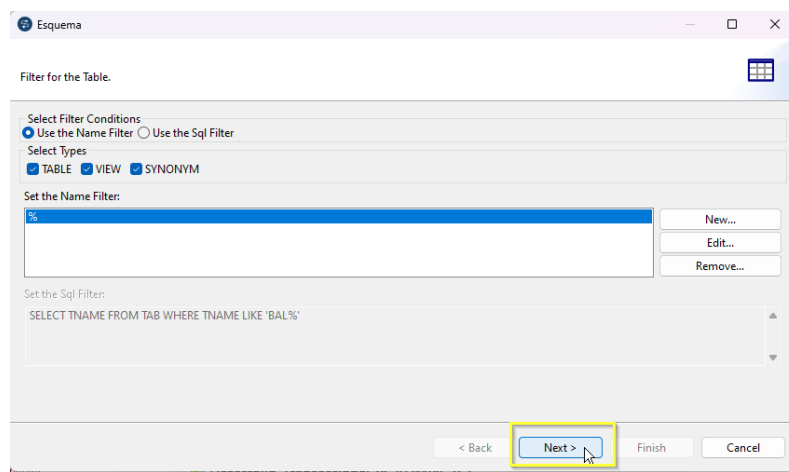
12. Nos dirigimos nuevamente al menú que nos aparece al lado izquierdo, damos clic en “extraer esquema”.



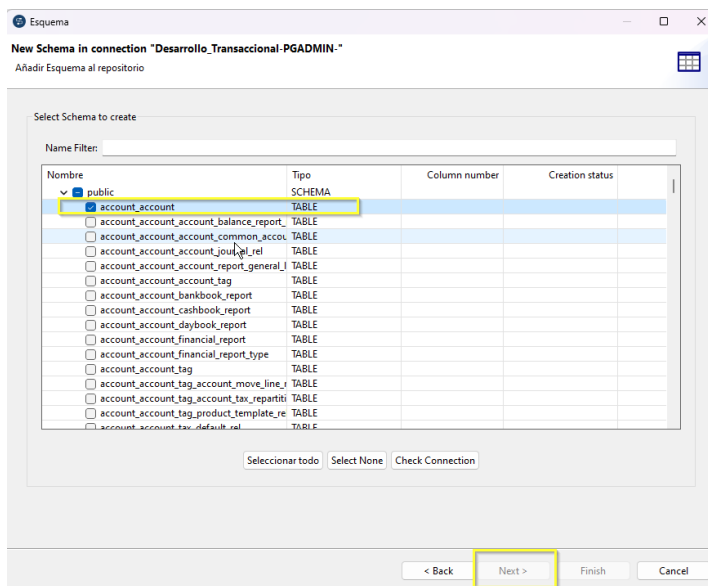
13. Nos muestra el set de contextos, seleccionamos el que necesitamos y damos clic en “Ok”.



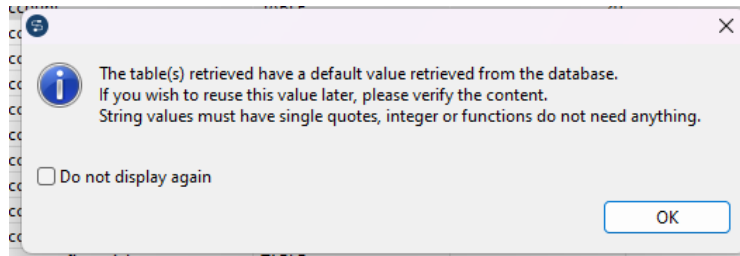
14. Se nos muestra la siguiente pantalla donde daremos clic en “Next”.



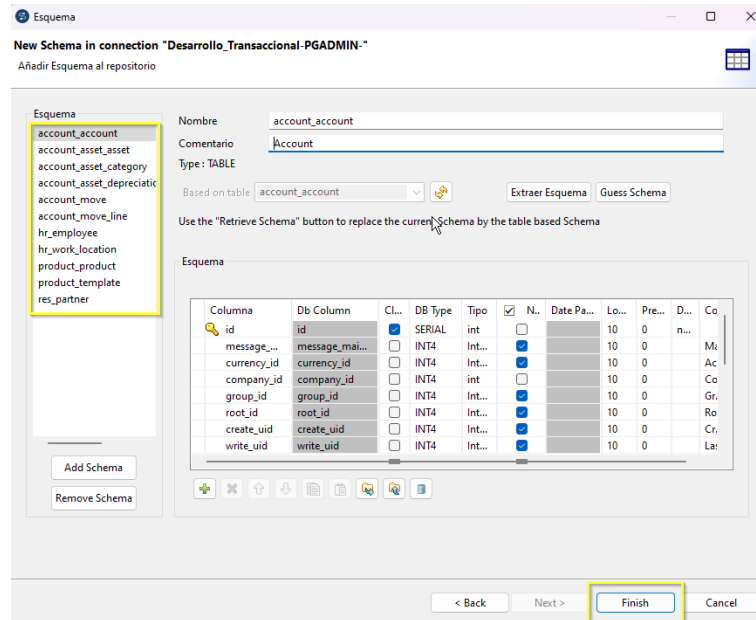
15. Nos aparecerá la base que consume Odoo, seleccionamos las tablas que utilizaremos para el modelo dimensional.



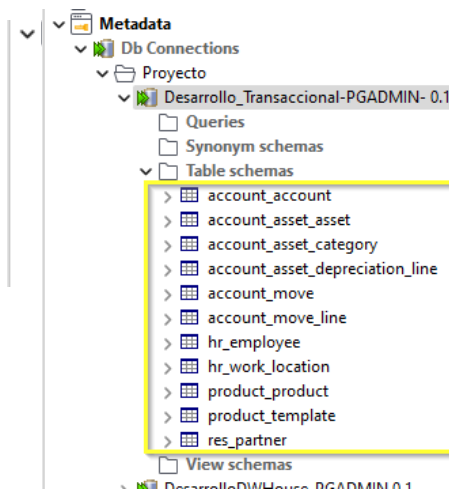
16. Damos clic en "Ok".



17. La siguiente pantalla nos muestra las tablas que vamos a utilizar, damos clic en "Finalizar".



18. Nos aparece en el apartado de la izquierda las tablas que se utilizarán.



Ingresamos un nombre que identifique nuestra conexión y una breve descripción, y damos clic en “Finish”.

Database Connection

Update Database Connection - Step 1/2

Update properties

Nombre: DesarrolloDWHouse-PGADMIN

Purpose: Conexión a DWH

Descripción:

Author: user@talend.com

Bloqueador: user@talend.com

Version: 0.1

Estado:

Path: Proyecto

< Back Next > Finish Cancel

19. Seleccionamos la base a la que se conectará, e ingresamos los datos necesarios para la conexión, damos clic en “Exportar como contexto”.

Database Connection

Update Database Connection - Step 2/2

You must press the Check Button to check the Database Setting

DB Type: PostgreSQL

Db Version: v9 and later

String of Connection: jdbc:postgresql://localhost:5432/DWHActivoFijo?

Login: odoo

Contraseña: ****

Servidor: localhost

Puerto: 5432

DataBase: DWHActivoFijo

Esquema:

Additional parameters:

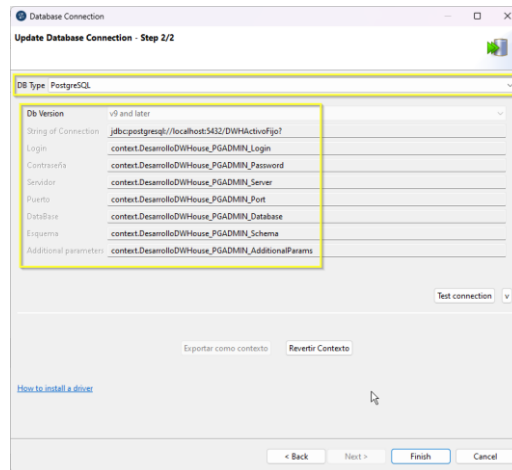
Test connection

Exportar como contexto Revertir Contexto

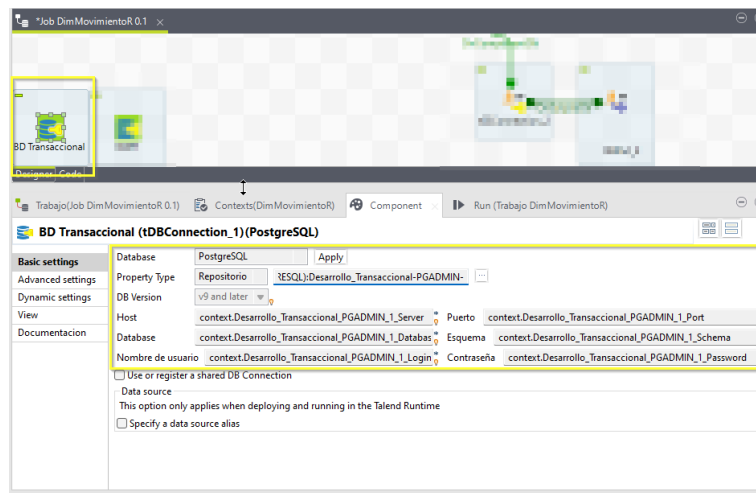
[How to install a driver](#)

< Back Next > Finish Cancel

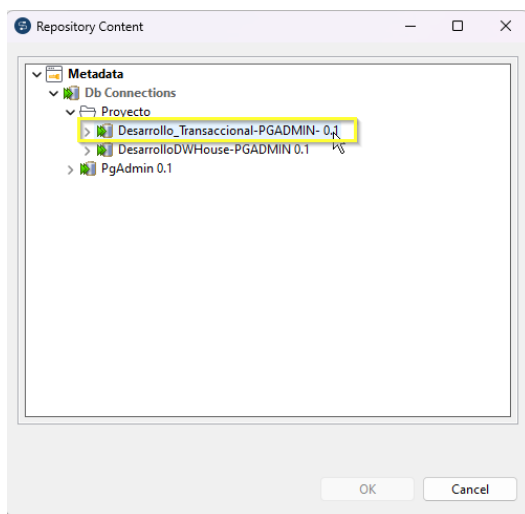
20. Nos quedará de la siguiente forma.



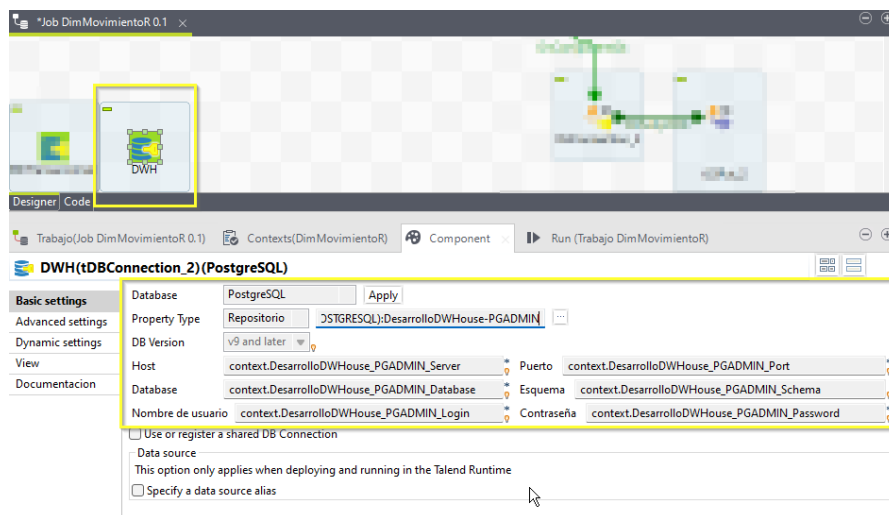
Ahora que tenemos todas las conexiones a la base y los contextos configurados se nos facilitará la extracción de la data mediante un objeto tDBCConnection, ya que en el apartado de “Property Type” seleccionamos “Repositorio”.



Luego seleccionamos la conexión creada previamente.

























Configuramos de la misma manera la conexión del Data Warehouse.



Con las siguientes configuraciones realizadas podemos comenzar a trabajar en los jobs los cuales serán repartidos en tres carpetas diferentes siguiendo el formato del S3 que son access, process y access

Los siguientes son los jobs de datos

- ▼  **Job Designs**
 - ▼  access
 -  DimActivoAccess 0.1
 -  DimDepreciacionAccess 0.1
 -  DimEmpleadoAccess 0.1
 -  DimEntidadAccess 0.1
 -  DimFechaAccess 0.1
 -  DimMovimientoAccess 0.1
 -  FactActivoFijo 0.1
 - ▼  process
 -  dim_activo_out 0.1
 -  dim_depreciacion_out 0.1
 -  dim_empleado_out 0.1
 -  dim_entidad_out 0.1
 -  dim_movimiento_out 0.1
 -  DimFecha 0.1
 - ▼  raw
 -  DimActivo 0.1
 -  DimDepreciacion 0.1
 -  DimEmpleado 0.1
 -  DimEntidad 0.1
 -  DimMovimiento 0.1

Aquí se muestra a continuación como quedan a continuación algunos de los jobs trabajados.

Raw

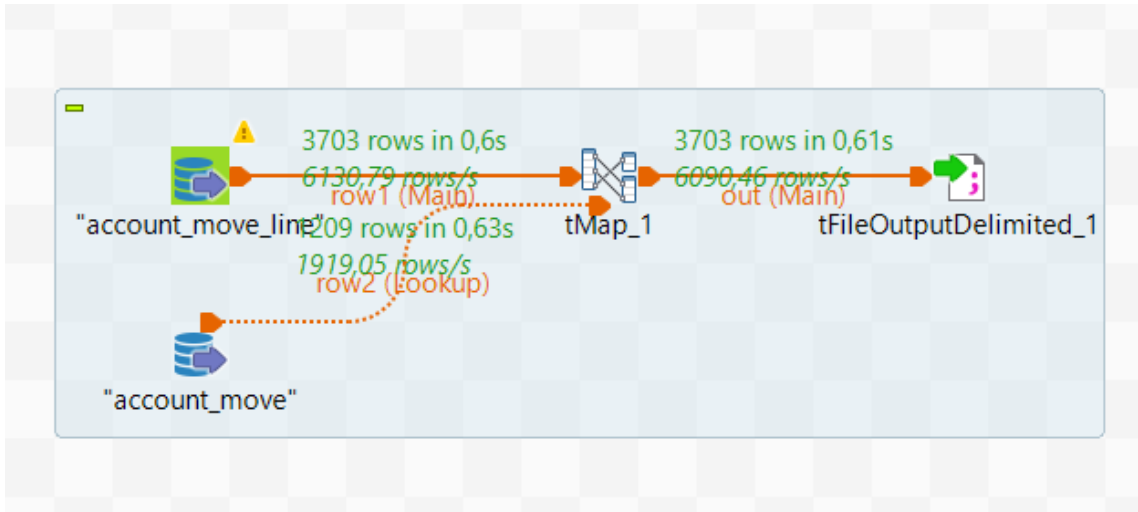


Diagrama de la interfaz de usuario de SAP Data Integration Studio para el job 'tmap_1'. Se muestran los detalles de los esquemas de entrada y salida.

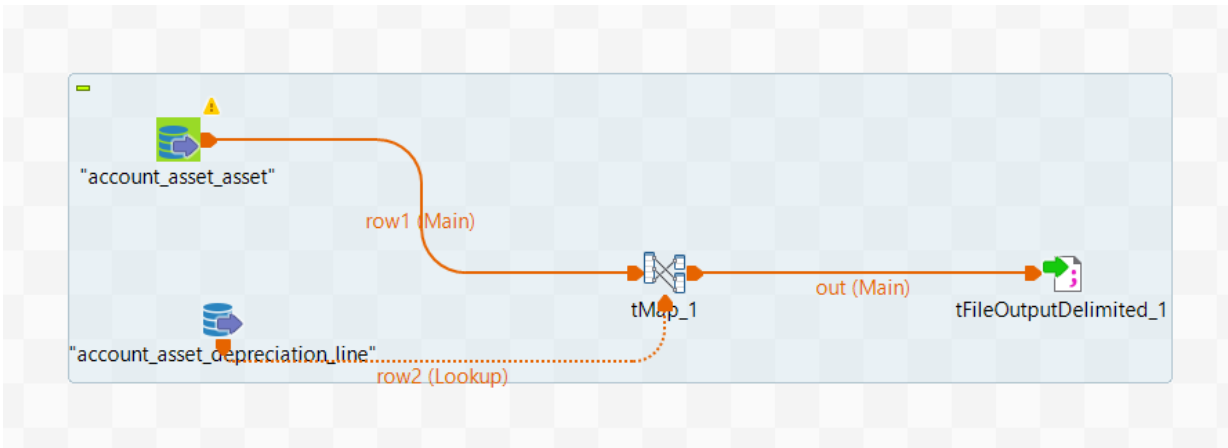
row1 (Entrada):

Column	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id		int			10	0	nextval...	
move_id		int			10	0		Journal En...
journal_id		Integer			10	0		Journal
company_id		Integer			10	0		Company
company_currency_id		Integer			10	0		Company ...
sequence		Integer			10	0		Sequence

out (Salida):

Column	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id		int			10	0	nextval...	
move_name		String			21474836...	0		Number
ref		String			21474836...	0		Reference
name		String			21474836...	0		Label
date		Date		"dd-MM-yyyy"	13	0		Date
quantity		BigDecimal			131089	0		Quantity

El editor de expresiones muestra la correspondencia entre las columnas de entrada y salida.



Talend Open Studio for Data Integration - tMap - tMap_1

Find:

Var

row1

Column
id
message_main_attachment_id
currency_id
company_id
category_id
partner_id
method_number
method_period
invoice_id
account_analytic_id
create_uid
write_uid
name
code
state
method
method_time
date_first_depreciation
date
method_end
first_depreciation_manual_date
analytic_distribution
note

out

Expresión	Column
row1.id	id
row1.message_main_attachment_id	message_main_attachment_...
row1.currency_id	currency_id
row1.company_id	company_id
row1.category_id	category_id
row1.partner_id	partner_id
row1.method_number	method_number
row1.method_period	method_period
row1.invoice_id	invoice_id
row1.account_analytic_id	account_analytic_id
row1.create_uid	create_uid
row1.write_uid	write_uid
row1.name	name
row1.code	code
row1.state	state
row1.method	method
row1.method_time	method_time
row1.date_first_depreciation	date_first_depreciation
row1.date	date
row1.method_end	method_end
row1.first_depreciation_manual_date	first_depreciation_manual_d...
row1.analytic_distribution	analytic_distribution
row1.note	note

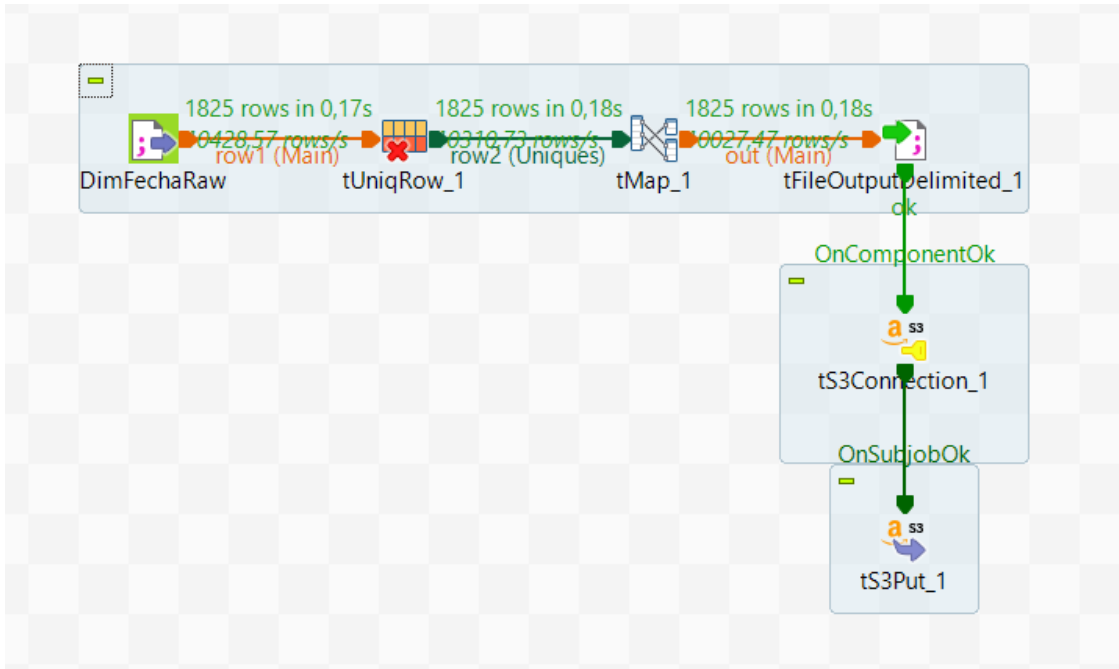
Editor de esquemas Editor de expresiones

row1	Columna	CL	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
<input checked="" type="checkbox"/>	id		int			10	0		nextval...
<input type="checkbox"/>	message_main_attachment_...	<input checked="" type="checkbox"/>	Integer		<input checked="" type="checkbox"/>	10	0		Main Atta...
<input type="checkbox"/>	currency_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Currency
<input type="checkbox"/>	company_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Company
<input type="checkbox"/>	category_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Category
<input type="checkbox"/>	partner_id	<input type="checkbox"/>	Integer		<input checked="" type="checkbox"/>	10	0		Partner

out	Columna	CL	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
<input checked="" type="checkbox"/>	id		int			10	0		nextval...
<input type="checkbox"/>	message_main_attachment_...	<input checked="" type="checkbox"/>	Integer		<input checked="" type="checkbox"/>	10	0		Main Atta...
<input type="checkbox"/>	currency_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Currency
<input type="checkbox"/>	company_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Company
<input type="checkbox"/>	category_id	<input type="checkbox"/>	int		<input type="checkbox"/>	10	0		Category
<input type="checkbox"/>	partner_id	<input type="checkbox"/>	Integer		<input checked="" type="checkbox"/>	10	0		Partner

Apply Ok Cancel

Process



Talend Open Studio for Data Integration - tMap - tMap_1

Find:

Auto map!

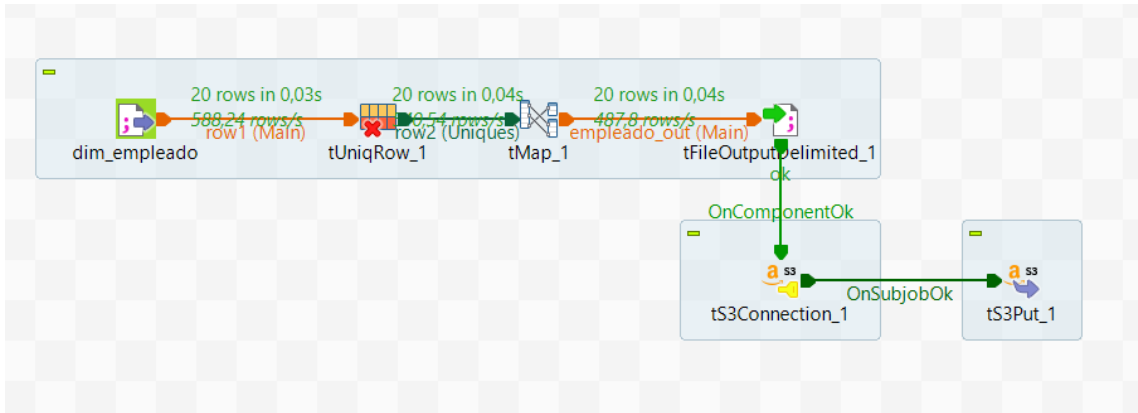
row2	Column
	date_key
	full_date
	day_of_week
	day_num_in_month
	day_num_overall
	day_name
	day_abbrev
	weekday_flag
	week_num_in_year
	week_begin_date
	week_begin_date_key
	month
	mont_num_overall
	mont_name
	quarter
	year
	yearmo
	fiscal_month
	fiscal_quarter
	fiscal_year
	last_day_in_month_flag
	same_day_year_ago_date

Editor de esquemas / Editor de expresiones

row2	Column	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
	date_key	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		8	0		
	full_date	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		10	0		
	day_of_week	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		1	0		
	day_num_in_month	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
	day_num_overall	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
	day_name	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		11	0		

out	Expression	Column
	Numeric sequence("S1",1,1)	idFecha
	row2.date_key	date_key
	row2.full_date	full_date
	row2.day_of_week	day_of_week
	row2.day_num_in_month	day_num_in_month
	row2.day_num_overall	day_num_overall
	row2.day_name	day_name
	row2.day_abbrev	day_abbrev
	row2.weekday_flag	weekday_flag
	row2.week_num_in_year	week_num_in_year

Apply Ok Cancel



talend Open Studio for Data Integration - tMap - tMap_1

row2

Column	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id		Integer		<input checked="" type="checkbox"/>	2	0		
resource_id		Integer		<input checked="" type="checkbox"/>	2	0		
company_id		Integer		<input checked="" type="checkbox"/>	1	0		
resource_calendar_id		Integer		<input checked="" type="checkbox"/>	1	0		
message_main_attachment_id		String		<input checked="" type="checkbox"/>	0	0		
color		Integer		<input checked="" type="checkbox"/>	1	0		

empleado_out

Columna	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id_empleado		Integer		<input checked="" type="checkbox"/>	2	0		
row2.id		Integer		<input checked="" type="checkbox"/>	2	0		
name		String		<input checked="" type="checkbox"/>	16	0		
job_title		String		<input checked="" type="checkbox"/>	31	0		
work_email		String		<input checked="" type="checkbox"/>	30	0		

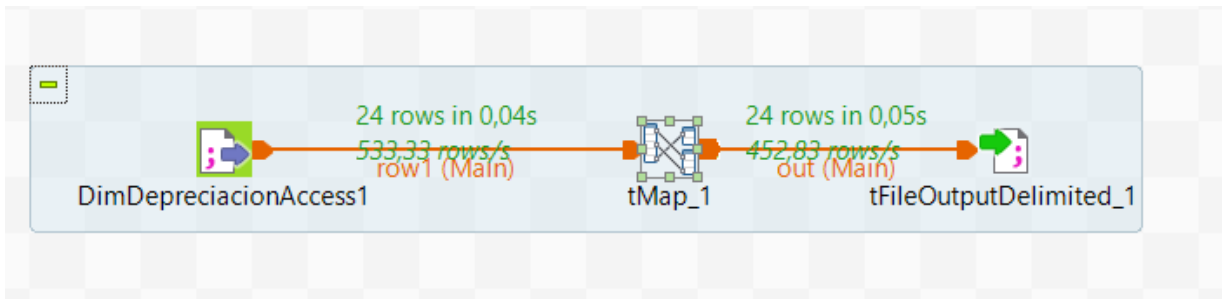
Editor de esquemas Editor de expresiones

Empleado_out

Columna	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id_empleado		Integer		<input checked="" type="checkbox"/>	2	0		
row2.id		Integer		<input checked="" type="checkbox"/>	2	0		
name		String		<input checked="" type="checkbox"/>	16	0		
job_title		String		<input checked="" type="checkbox"/>	31	0		
work_email		String		<input checked="" type="checkbox"/>	30	0		

Apply OK Cancel

Access



Talend Open Studio for Data Integration - tMap - tMap_1

Find:

Var

Auto map!

row1

Column
id_empleado
id
name
job_title
work_email

out

Expresión	Column
row1.id_empleado	id_empleado
row1.id	id
row1.name	empleado
row1.job_title	título_de_trabajo
row1.work_email	correo_empresaial

Editor de esquemas Editor de expresiones

row1

Columna	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id_empleado	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
id	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
name	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		16	0		
job_title	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		31	0		
work_email	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		30	0		

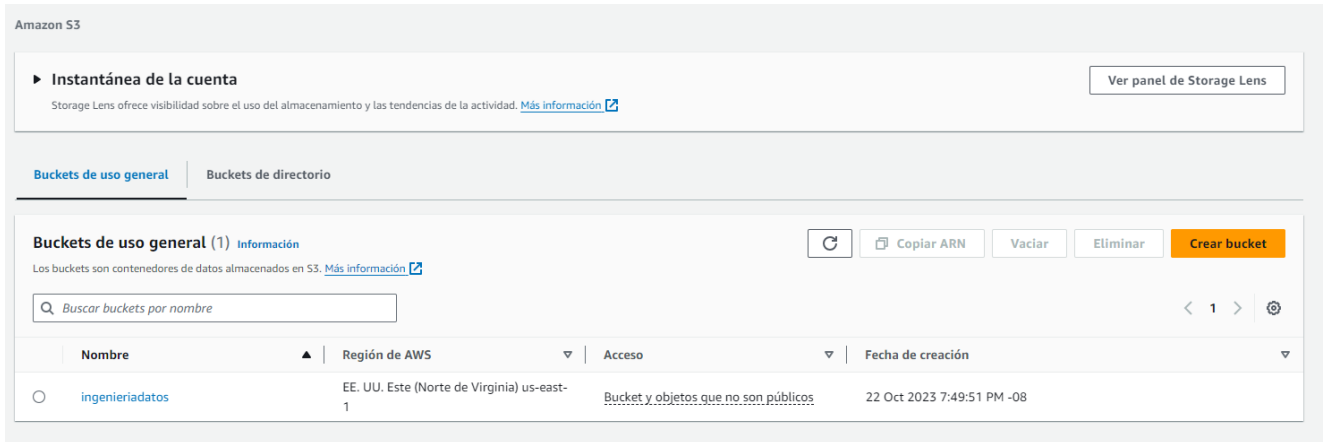
out

Columna	Cl.	Tipo	N.	Date Pattern (Ctrl+...)	Longitud	Precision	Default	Comentario
id_empleado	<input checked="" type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
id	<input type="checkbox"/>	Integer	<input checked="" type="checkbox"/>		2	0		
empleado	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		16	0		
título_de_trabajo	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		31	0		
correo_empresaial	<input type="checkbox"/>	String	<input checked="" type="checkbox"/>		30	0		

Apply Ok Cancel

Amazon S3

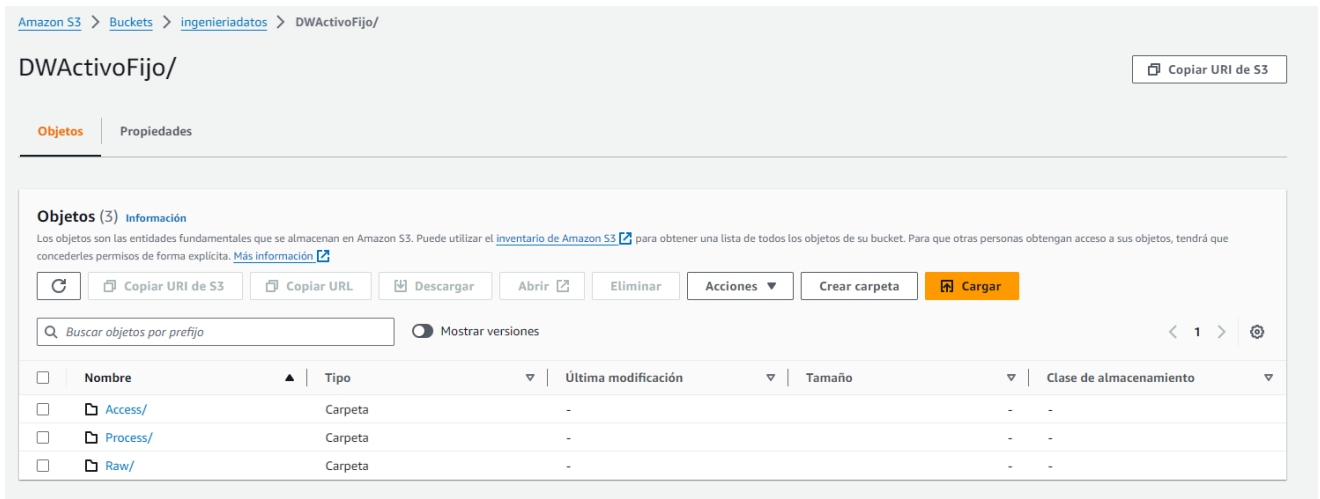
Se creó un bucket llamado “ingenieriadatos” en Amazon S3 para la implementación de la solución. La creación del Bucket utilizado en el proyecto lo podemos visualizar en la imagen



Creación de la estructura de almacenamiento en S3 con las siguientes carpetas dentro del bucket:

1. RAW
2. PROCESS
3. ACCESS

La estructura de las carpetas creadas en Amazon S3 para el almacenamiento en la nube del proyecto, la podemos visualizar en la imagen



Para poder acceder al Bucket de Amazon S3 es necesario realizar una configuración al componente tS3Connection en Talend Open Studio, agregando el Access Key y la Secret Key para lograr el acceso a nuestro bucket “ingenieriadatos”.

Amazon Redshift

Se llevó a cabo la creación de un clúster en AWS Redshift con el nombre “ingenieriadatoscluster”, lo podemos visualizar en la siguiente imagen, en el cual vamos a poder almacenar los datos de la capa de Access obtenidos del proceso ETL de Talend Open Studio y su carga de datos en Amazon S3 para su posterior visualización gráfica en Power BI y ayudar a la empresa en sus futuras decisiones estratégicas.

The screenshot shows the AWS Redshift console interface. At the top, there's a navigation bar with 'Vaya a Serverless', 'Adquirir nodos reservados', and 'Crear clúster' buttons. Below this is a section titled 'Información general sobre los recursos' for the 'ingenieriadatoscluster' in the US East (N. Virginia) region. It displays a summary of resources: 1 total node, 1 on-demand node, 0 reserved nodes, 0 available reserved nodes (0 of 0 used), 3 automated snapshots, and 2 manual snapshots. Below this is a table for 'Información general acerca del clúster (1)', showing the cluster name 'ingenieriadatoscluster' and its state as 'Available'. To the right, there's a 'Recursos compartidos de datos' section with a warning to authorize access to other AWS accounts. At the bottom, there are controls for 'Métricas del clúster', including a refresh button, a dropdown for 'Cualquier clúster', a dropdown for 'Última hora', and a 'Ver en CloudWatch' button.

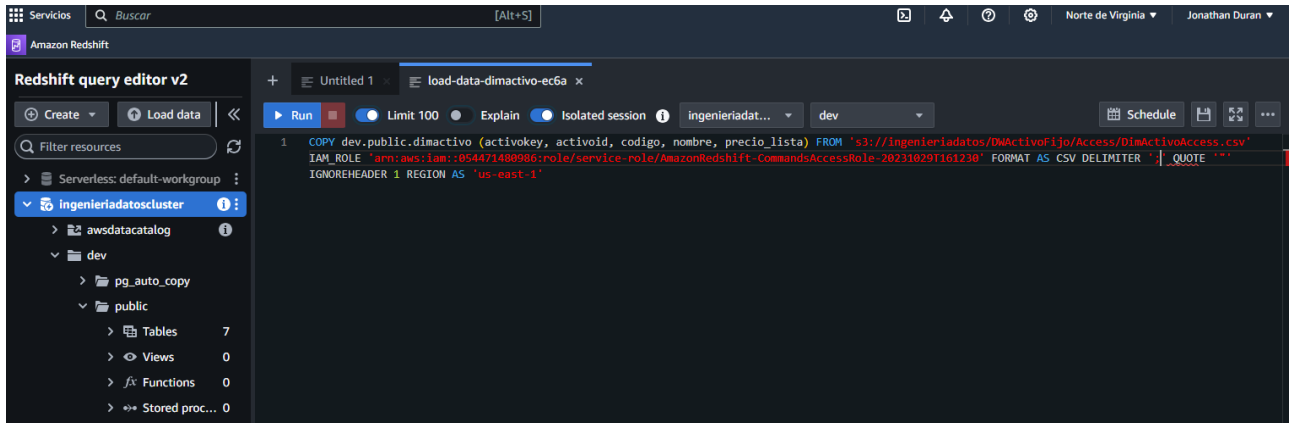
Para lograr la conexión entre Redshift y la capa de Access de S3 se ejecutó el siguiente procedimiento: Ejecutamos un script con el cual creamos la estructura de la tabla de hecho y sus dimensiones. Como ejemplo: Dimensión Activo.

The screenshot shows the AWS Redshift query editor interface. The top bar includes the AWS logo, 'Servicios', a search bar, and a keyboard shortcut '[Alt+S]'. The main editor area is titled 'Redshift query editor v2' and shows a SQL script for creating a table named 'DimActivo'. The script is as follows:

```
1 CREATE TABLE DimActivo(  
2 ActivoKey INT NOT NULL,  
3 ActivoID INT NOT NULL,  
4 Código VARCHAR(100) NULL,  
5 Nombre VARCHAR(100) NULL,  
6 PrecioLista INT NULL,  
7 CONSTRAINT ActivoSK PRIMARY KEY (ActivoKey)  
8 )
```

The interface also shows a left sidebar with navigation options like 'Editor', 'Queries', 'Notebooks', 'Charts', 'History', and 'Scheduled queries'. The 'Queries' section is expanded to show the 'ingenieriadatoscluster' and its sub-folders: 'awsdatacatalog', 'dev', 'dev1', 'dev2', 'dev3', and 'sample_data_dev'. The 'Run' button is visible, along with options for 'Limit 100', 'Explain', and 'Isolated session'. The bottom of the editor shows a 'Result 1' section and a 'Summary' tab.

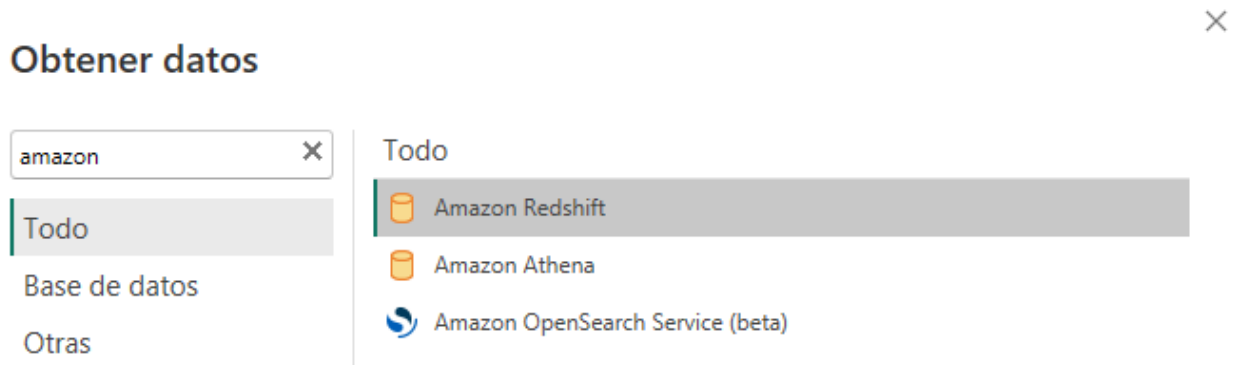
Para cargar los datos, se debe de importar desde la capa de Access del Bucket en S3 hacia la dimensión correspondiente como en la siguiente imagen



POWER BI

Finalmente se utilizará la herramienta Power BI ya que es una solución de análisis empresarial basado en la nube, y permite unir diferentes fuentes de datos, analizarlos y presentar un análisis de estos a través de informes y paneles. Realizando una conexión con la base de datos de AWS Redshift podemos acceder a los datos almacenados en la base de datos con el modelo del DataWarehouse definido en la implementación de la solución, obtenido del ETL generado con Talend Open Studio y almacenado en la capa de presentación en AWS S3.

Primero debemos de seleccionar el origen de datos, en nuestro caso será Amazon Redshift



Luego colocamos el end point y nombre de la base de datos de Redshift

Amazon Redshift

Server

ingenieriadatoscluster.cxkakybbrx4z.us-east-1.redshift.amazonaws.com:5439

Database

dev

▶ Opciones avanzadas

Aceptar

Cancelar

Colocamos nuestro usuario y contraseña de base de datos de Redshift y luego se nos mostraran las tablas de la base de datos para seleccionarlas a nuestro dashboard

Navegador

🔍

Opciones de presentación ▾

Amazon Redshift [5]

▶ catalog_history

▶ pg_auto_copy

▶ pg_automv

▶ pg_s3

public [7]

dimactivo

dimdepreciacion

dimempleado

dimentidad

dimfecha

dimmovimiento

factactivofigo

dimactivo

activokey	actvoid	codigo	nombre	precio
1	5	FURN_7777	Silla con rodos	
2	18	FURN_1118	Escritorio con espejo	
3	17	FURN_0269	Silla de oficina negra	
4	21	FURN_7888	Escritorio con vidrio templado	
5	15	E-COM11	Escritorio gris	
6	10	E-COM06	Escritorio de esquina	
7	3	EST-8274	Escritorio Gerencial	
8	19	FURN_8900	Escritorio negro	
9	20	FURN_9001	Escritorio ejecutivo	
10	22	FURN_0789	Escritorio privado de trabajo	
11	24	FURN_8855	Mini escritorio	
12	7	FURN_9999	Silla sencilla de espera	
13	23	FURN_6666	Silla de oficina azul	
14	6	FURN_8888	Silla con colchon extra afelpado	
15	14	E-COM10	Silla elevada	
16	12	E-COM08	Silla de plastico azul	
17	13	E-COM09	Silla ejecutiva de cuero	
18	4	ECOM-8273	Silla de reuniones formal	
19	11	E-COM07	Silla gerencial	
20	27	FURN_8999	Silla Gamer Pro	
21	9		Escritorio moderno	
22	29	ECOM-99823	Silla de madera	
23	28	E-COM007	Escritorio doble nivel	

Seleccionar tablas relacionadas

Cargar

Transformar datos

Cancelar

b) Presupuesto de implementación

Para la empresa de inversiones digitales es importante contratar servicios profesionales para la implementación de la solución, y después de eso comenzar con las capacitaciones a los recursos de la empresa para el uso óptimo de las herramientas necesarias para la visualización de los informes estratégicos. Este costo se realizará tomando en cuenta la implementación y capacitación que se llevará a cabo en un lapso de 15 días. Se realizará el cálculo tomando en cuenta el salario otorgado por la empresa que es de \$800 mensuales, se realizará por ambos integrantes del proyecto. Realizando los cálculos correctos determinamos que el salario diario de cada persona es de \$26.66 por día y por hora trabajada, el desglose de este presupuesto es el siguiente:

Recurso humano: 2

Precio por hora: \$3.33

Cantidad de horas para la implementación: 120

Costo total: \$799.20

Hardware

Es importante para la empresa tener un equipo que posea los recursos mínimos necesarios y que sea capaz de soportar la ejecución simultanea de las diferentes herramientas que se utilizaran en el desarrollo de la solución. Tomando en cuenta esto se decidió por la compra de un equipo con las siguientes características:

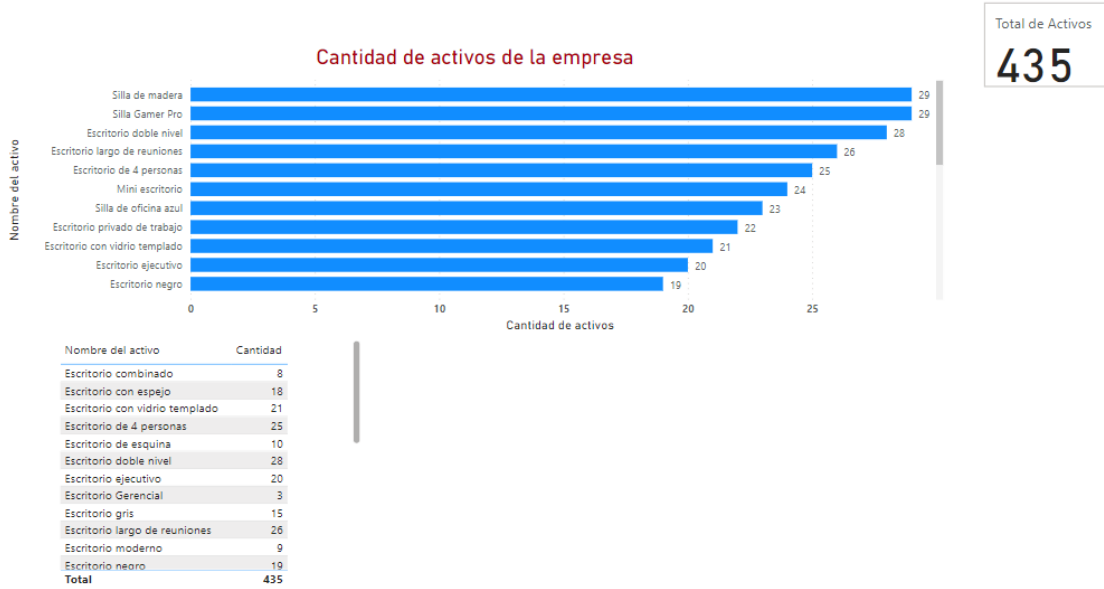
- INTEL CORE i5-4570.
- RAM 8GB.
- SSD 240GB.

Precio valorado de \$329.0

c) **Análisis de resultados** (breve manual de resultados que evidencie las capturas relevantes de la propuesta de la solución).

Especificación de Necesidades que el Modelo Propuesto resolverá.

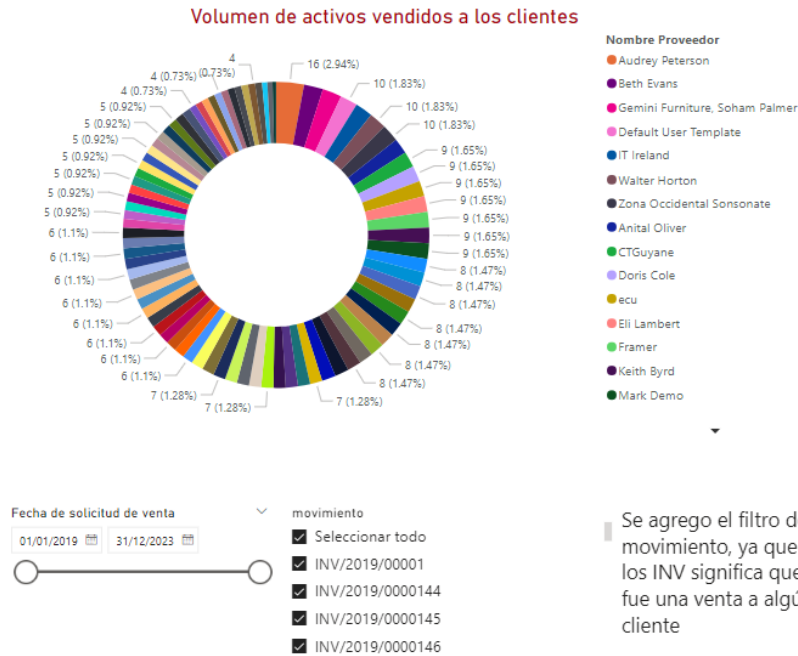
Reporte de cantidad de activos de la empresa(conteo): Genera un informe que permite visualizar la cantidad de activos con los cuales cuenta la empresa.



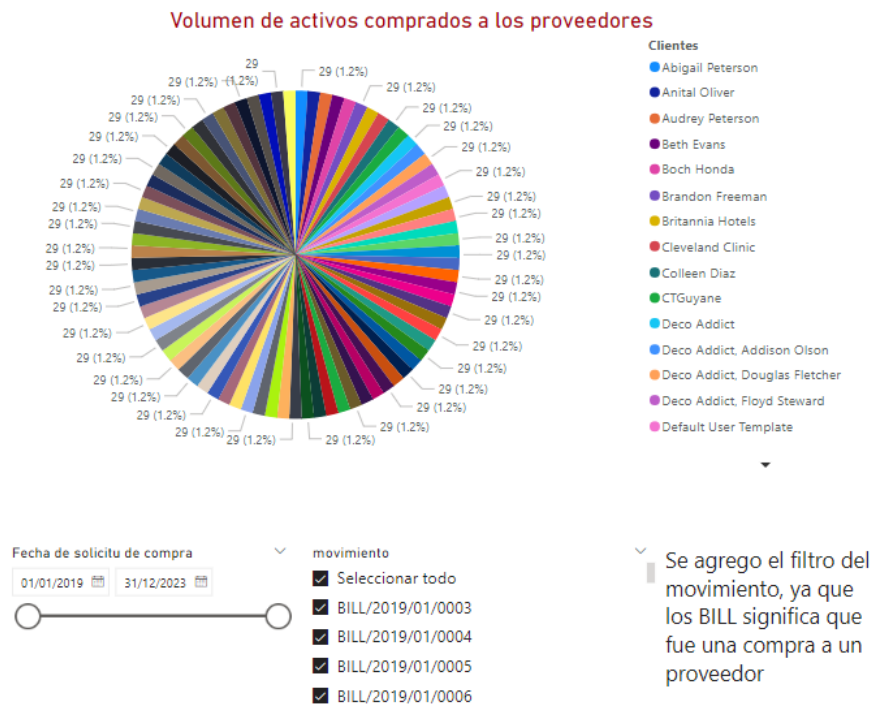
Reporte de cantidad de activos asignados por empleado(conteo): Genera un informe que permite visualizar la cantidad de activos que han sido asignados a los empleados en un rango seleccionado o fecha específica.



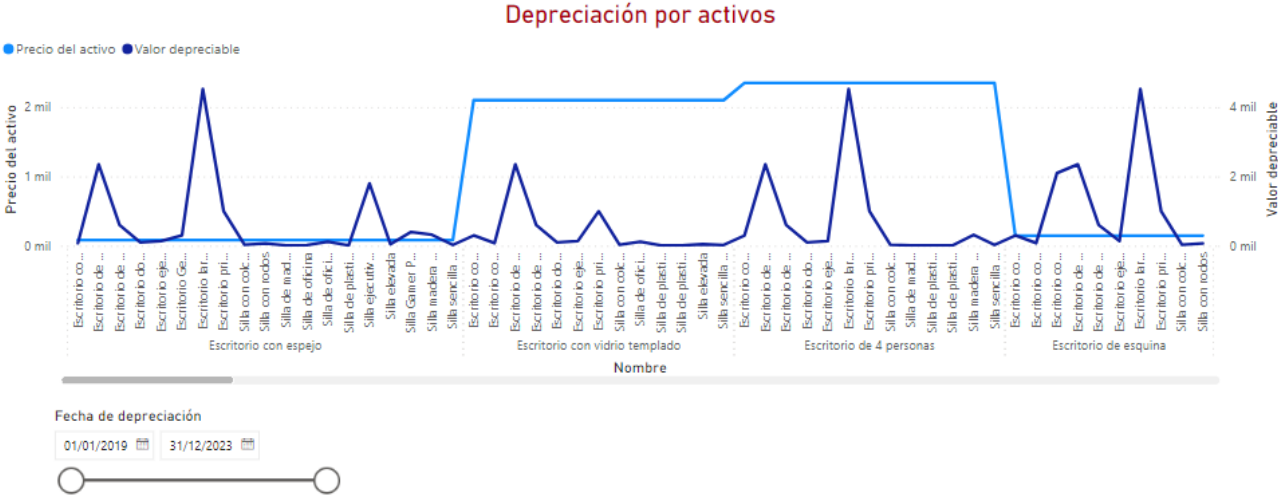
Reporte de volumen de activos vendidos a los clientes(conteo): Genera un informe que permite visualizar el total de todos los activos vendidos a cada uno de los clientes en un rango seleccionado o fecha específica de venta.



Reporte de volumen de activos comprados a los proveedores(conteo): Genera un informe que permite visualizar el total de cada compra a cada proveedor en un rango seleccionado o fecha específica.



Reporte de depreciación por activos (valor de depreciación): Genera un informe que permite visualizar el valor de la depreciación de cada activo contra el valor del precio del mismo activo en un rango seleccionado o fecha específica.



Conclusiones y recomendaciones

Conclusiones

- Durante la creación de la solución de Data Warehouse para la empresa Inversiones Digitales, se emplearon los conocimientos obtenidos durante nuestra especialización en ingeniería de datos.
- Se desarrolló un modelo dimensional utilizando la metodología estándar para la construcción de un Data Warehouse, adaptándolo a las necesidades específicas de Inversiones Digitales, una empresa dedicada a la fabricación y venta de sillas y escritorios de trabajo.
- Implementamos una capa de presentación en la solución de Data Warehouse que permite el consumo de datos mediante una aplicación de Business Intelligence (BI). Esta aplicación posibilita la presentación de visualizaciones de datos mediante informes interactivos, facilitando la interpretación para los usuarios.

Recomendaciones

- Se sugiere que la implementación de la solución de Data Warehouse sea llevada a cabo preferiblemente por un ingeniero de datos con conocimientos en servicios de AWS, Talend Open Studio, PostgreSQL y Power BI.
- Se recomienda seguir utilizando la metodología utilizada durante el desarrollo de la solución de Data Warehouse, asegurándose de que esté alineada con las necesidades planteadas por la organización. Esto debería hacerse antes de realizar cualquier cambio en la solución, garantizando así su adaptación continua a los requisitos empresariales.

Bibliografía

- Kimball, R., & Ross, M. (2013). *The data warehouse toolkit (3rd ed.)*. Wiley
- Bowen, J. (2013). "Getting Started with Talend Open Studio for Data Integration." Packt Publishing.
- Barton, R. (2013). "Talend Open Studio Cookbook." Packt Publishing.
- Baron, J., et al. (2016). "AWS Certified Solutions Architect Official Study Guide." Sybex.
- HOW TO GET ACCESS KEY ID AND SECRET ACCESS KEY OF AMAZON S3 ACCOUNT? (s.f.). Obtenido de Supsysitic: <https://supsysitic.com/documentation/id-secret-access-keyamazon-s3/>
- What is Amazon S3? - Amazon Simple Storage Service. (s.f.). Obtenido de Amazon AWS: <https://docs.aws.amazon.com/AmazonS3/latest/userguide/Welcome.html>
- What is IAM? - AWS Identity and Access Management. (s.f.). Obtenido de Amazon AWS: <https://docs.aws.amazon.com/IAM/latest/UserGuide/introduction.html>
- Kalin, M. (2019). "Learning Amazon S3." O'Reilly Media.
- Nickel, J. (2019). "Mastering Identity and Access Management with Microsoft Azure." Packt Publishing.
- Brunn, H. (2021). "Odoos 14 Development Cookbook." Packt Publishing.
- ¿Qué es IAM? - AWS Identity and Access Management. (s. f.). https://docs.aws.amazon.com/es_es/IAM/latest/UserGuide/introduction.html
- Ferrari, A., & Russo, M. (2018). "Analyzing Data with Power BI and Power Pivot for Excel." Microsoft Press.
- Powell, B. (2016). "Mastering Power BI." Packt Publishing.
- Provost, F., & Fawcett, T. (2013). "Data Science for Business." O'Reilly Media.
- Bragg, S. M. (2005). "Fixed Asset Accounting: A Comprehensive Guide." Wiley.
- AWS Pricing Calculator. (s.f.). Obtenido de AWS Pricing Calculator: <https://calculator.aws/#/>

Glosario

A

Access: Zona del bucket destinada a guardar los datos definitivos a tratar para la presentación de la información cargada por medio de redshift y mostrada en power bi.

C

Clúster de Amazon Redshift: un cluster es un grupo de nodos que tienen un nodo líder encargado de crear planes de acción que son ejecutados en paralelo por los nodos de cómputo cuando llega una query.

D

Dashboard: en el contexto de Power BI son cuadros de mando o visualizaciones de un conjunto de datos lo que apoya a la toma de decisiones en la empresa.

Data Warehouse: es un almacén de datos electrónico de gran capacidad capaz de almacenar datos de forma segura, fiable, fácil de recuperar y administrar.

Dimensiones: en el modelado dimensional, las dimensiones denotan el contexto del hecho que se está queriendo analizar, se enlazan a la tabla de hechos por medio de un campo clave.

E

Esquema estrella: en el modelado dimensional, un esquema estrella está conformado por una tabla de hechos al centro y varias tablas de dimensiones enlazadas a la tabla de hechos

ETL: siglas provenientes del inglés “Extract TransformLoad” utilizadas para los procesos de extracción, transformación y carga de datos,13GGranularidad: en el modelado dimensional, la granularidad define el nivel mínimo de detalle de los datos con que cuenta el sistema fuente.

I

IAM: siglas provenientes del inglés para “Identity and Access Management”, servicio de AWS para la administración de identidades y acceso a los recursos de AWS de forma segura.

Ingeniería de Datos: se encarga de construir y mantener la estructura de datos y las arquitecturas necesarias que alimentan aplicaciones que utilizan grandes cantidades de datos.

M

Modelo Dimensional: resultado final del modelado dimensional que consiste en un diagrama compuesto por dimensiones y tablas de hechos.

P

Power BI: es un conjunto de herramientas de Microsoft, utilizada para el análisis de datos capaz de crear visualizaciones interactivas para el usuario de negocio.

Process: zona del bucket destinada para el almacenamiento de los datos ya disponibles para su presentación luego de haber pasado por los procesos ETL.

R

Raw: zona del bucket destinada para el almacenamiento de los datos crudos tal y como se obtienen desde la fuente de origen.

Redshift: Es un servicio de AWS utilizado para el almacenamiento masivo de datos llegando hasta el orden de los petabytes. Amazon Redshift está orientado al almacenamiento de datos en la nube tanto para datos estructurados como semi estructurados.

S

S3: de las siglas en inglés "Simple Storage Service", S3 es uno de los servicios web ofrecidos por Amazon desde 2006 para el almacenamiento de objetos, entre sus principales puntos destacables S3 ofrece una solución para el almacenamiento de objetos segura, duradera y escalable a un bajo costo.

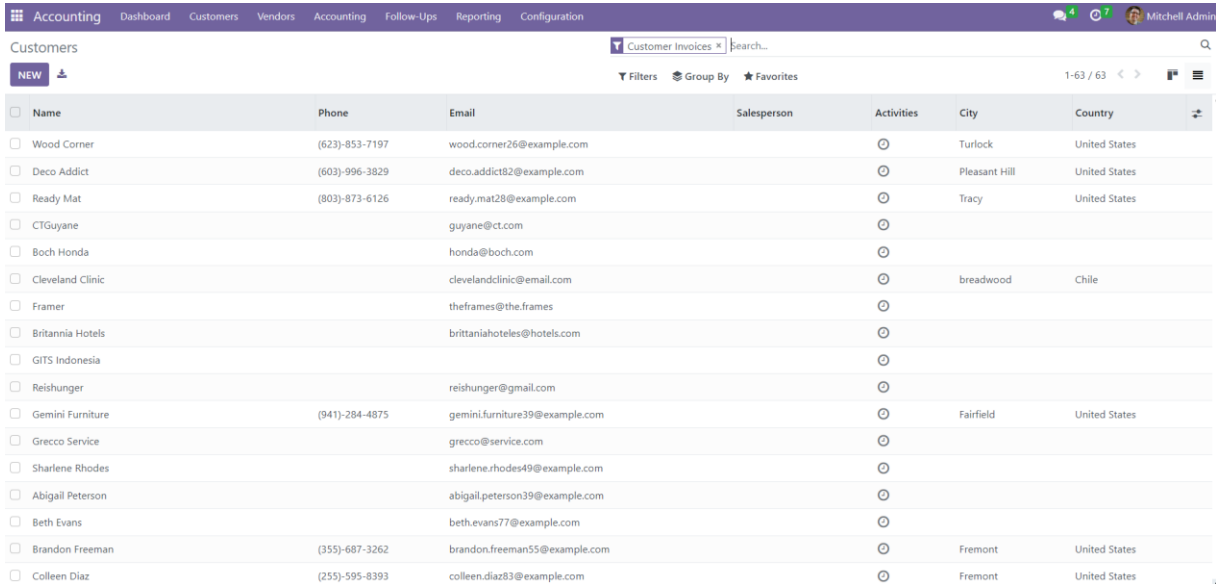
T

Talend open studio: Es una suite que aporta un conjunto muy completo de herramientas de integración de datos que se ofrece en una versión de código libre Community Edition CE totalmente funcional.

Anexos

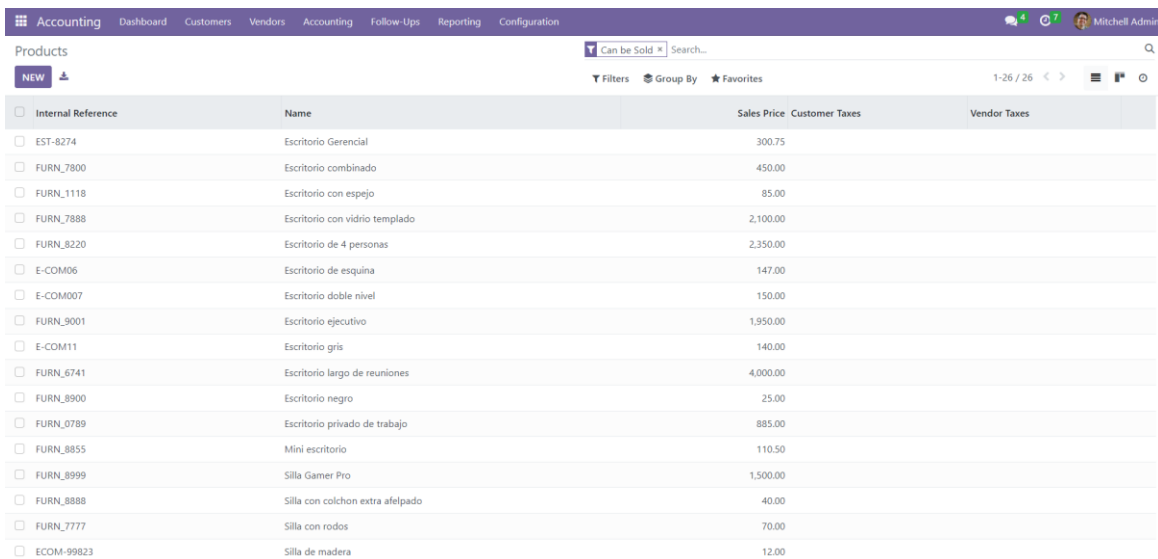
Visualización del proceso de compra y venta de activos en el sistema ODOO.

En el apartado de clientes es donde se tienen toda la información de los clientes de la empresa con su información de contacto y otros detalles de estos



<input type="checkbox"/>	Name	Phone	Email	Salesperson	Activities	City	Country
<input type="checkbox"/>	Wood Corner	(623)-853-7197	wood.corner26@example.com		○	Turlock	United States
<input type="checkbox"/>	Deco Addict	(603)-996-3829	deco.addict82@example.com		○	Pleasant Hill	United States
<input type="checkbox"/>	Ready Mat	(803)-873-6126	ready.mat28@example.com		○	Tracy	United States
<input type="checkbox"/>	CTGuyane		guyane@ct.com		○		
<input type="checkbox"/>	Boch Honda		honda@boch.com		○		
<input type="checkbox"/>	Cleveland Clinic		clevelandclinic@email.com		○	breadwood	Chile
<input type="checkbox"/>	Framer		theframes@the.frames		○		
<input type="checkbox"/>	Britannia Hotels		britanniahoteles@hotels.com		○		
<input type="checkbox"/>	GITS Indonesia				○		
<input type="checkbox"/>	Reishunger		reishunger@gmail.com		○		
<input type="checkbox"/>	Gemini Furniture	(941)-284-4875	gemini.furniture39@example.com		○	Fairfield	United States
<input type="checkbox"/>	Grecco Service		grecco@service.com		○		
<input type="checkbox"/>	Sharlene Rhodes		sharlene.rhodes49@example.com		○		
<input type="checkbox"/>	Abigail Peterson		abigail.peterson39@example.com		○		
<input type="checkbox"/>	Beth Evans		beth.evans77@example.com		○		
<input type="checkbox"/>	Brandon Freeman	(355)-687-3262	brandon.freeman55@example.com		○	Fremont	United States
<input type="checkbox"/>	Colleen Diaz	(255)-595-8393	colleen.diaz83@example.com		○	Fremont	United States

En el apartado de productos se encuentra la lista de activos que la empresa distribuye a sus clientes y sus diferentes detalles



<input type="checkbox"/>	Internal Reference	Name	Sales Price	Customer Taxes	Vendor Taxes
<input type="checkbox"/>	EST-8274	Escritorio Gerencial	300.75		
<input type="checkbox"/>	FURN_7800	Escritorio combinado	450.00		
<input type="checkbox"/>	FURN_1118	Escritorio con espejo	85.00		
<input type="checkbox"/>	FURN_7888	Escritorio con vidrio templado	2,100.00		
<input type="checkbox"/>	FURN_8220	Escritorio de 4 personas	2,350.00		
<input type="checkbox"/>	E-COM06	Escritorio de esquina	147.00		
<input type="checkbox"/>	E-COM007	Escritorio doble nivel	150.00		
<input type="checkbox"/>	FURN_9001	Escritorio ejecutivo	1,950.00		
<input type="checkbox"/>	E-COM11	Escritorio gris	140.00		
<input type="checkbox"/>	FURN_6741	Escritorio largo de reuniones	4,000.00		
<input type="checkbox"/>	FURN_8900	Escritorio negro	25.00		
<input type="checkbox"/>	FURN_0789	Escritorio privado de trabajo	885.00		
<input type="checkbox"/>	FURN_8855	Mini escritorio	110.50		
<input type="checkbox"/>	FURN_8999	Silla Gamer Pro	1,500.00		
<input type="checkbox"/>	FURN_8888	Silla con colchon extra afeipado	40.00		
<input type="checkbox"/>	FURN_7777	Silla con rodos	70.00		
<input type="checkbox"/>	ECOM-99823	Silla de madera	12.00		

PRINT LABELS

0 Extra Prices

0.00 Units Sold

Product Name
☆ **Escritorio de esquina**

Can be Sold Can be Purchased

General Information | Attributes & Variants | Sales | Purchase | Accounting

Product Type ?	Consumable	Sales Price ?	\$ 147.00
Invoicing Policy ?	Delivered quantities	Customer Taxes ?	
Re-Invoice Expenses ?	<input checked="" type="radio"/> No <input type="radio"/> At cost <input type="radio"/> Sales price	Cost ?	\$ 600.00
		Product Category	All / Saleable / Office Furniture
		Internal Reference	E-COM06
		Barcode	
		Product Tags	

Consumables are physical products for which you don't manage the inventory level: they are always available.

Invoice after delivery, based on quantities delivered, not ordered.






También se tiene el apartado de tiendas que la empresa posee con sus detalles

Vendors

Vendor Bills Search...

NEW

Filters Group By Favorites 1-5 / 5 < >

 Zona Central Chalatenango chalatenango@inversionesDigitales.com \$1	 Zona Occidental Sonsonate sonsonate@inversionesDigitales.com	 Zona Oriental San Miguel zonaOriental@inversionesDigitales.com	 Zona Central La Libertad lalibertad@inversionesDigitales.com
 Zona Central San Salvador sansalvador@inversionesDigitales.com			

En la parte de ventas realizadas a los clientes es donde se encuentran los detalles de esas facturas de a que cliente se le vende y que productos del inventario

Invoices Search...

NEW **UPLOAD**

▼ Filters Group By ★ Favorites

<input type="checkbox"/>	Invoice ID	Customer	Date	Status	Icon	Price	Total
<input type="checkbox"/>	INV/2023/00094	The Jackson Group, Gordon Owens	08/09/2023	4 days ago	🕒	\$ 8,500.00	\$ 9,775.00
<input type="checkbox"/>	INV/2023/00093	The Jackson Group	08/07/2023	4 days ago	🕒	\$ 1,575.00	\$ 1,811.25
<input type="checkbox"/>	INV/2023/00092	ecu	08/05/2023	68 days ago	🕒	\$ 1,547.00	\$ 1,779.05
<input type="checkbox"/>	INV/2023/00091	Deco Addict, Addison Olson	08/01/2023	98 days ago	🕒	\$ 760.00	\$ 874.00
<input type="checkbox"/>	INV/2023/00090	CTGuyane	07/28/2023	88 days ago	🕒	\$ 3,640.00	\$ 4,186.00
<input type="checkbox"/>	INV/2023/00089	Wood Corner	07/25/2023	88 days ago	🕒	\$ 87,750.00	\$ 100,912.50
<input type="checkbox"/>	INV/2023/00088	Reishunger	07/19/2023	88 days ago	🕒	\$ 15,045.00	\$ 17,301.75
<input type="checkbox"/>	INV/2023/00087	Wood Corner, Tom Ruiz	07/15/2023	88 days ago	🕒	\$ 39,000.00	\$ 44,850.00
<input type="checkbox"/>	INV/2023/00086	Deco Addict	07/13/2023	08/12/2023	🕒	\$ 6,615.00	\$ 7,607.25
<input type="checkbox"/>	INV/2023/00085	Framer	07/10/2023	08/09/2023	🕒	\$ 8,000.00	\$ 9,200.00
<input type="checkbox"/>	INV/2023/00014	Cleveland Clinic	07/08/2023	98 days ago	🕒	\$ 1,912.50	\$ 2,199.38
<input type="checkbox"/>	INV/2023/00084	Abigail Peterson	07/06/2023	88 days ago	🕒	\$ 1,125.00	\$ 1,293.75
<input type="checkbox"/>	INV/2023/00083	Eli Lambert	07/03/2023	4 days ago	🕒	\$ 1,560.00	\$ 1,794.00
<input type="checkbox"/>	INV/2023/00080	Ready Mat, Billy Fox	06/30/2023	08/14/2023	🕒	\$ 800.00	\$ 920.00
<input type="checkbox"/>	INV/2023/00079	indi	06/24/2023	4 days ago	🕒	\$ 2,099.50	\$ 2,414.43
<input type="checkbox"/>	INV/2023/00078	Ready Mat	06/20/2023	08/04/2023	🕒	\$ 76,050.00	\$ 87,457.50
<input type="checkbox"/>	INV/2023/00077	Tina Williamson	06/18/2023	08/10/2023	🕒	\$ 900.00	\$ 1,035.00
<input type="checkbox"/>	INV/2023/00076	Ledcor	06/15/2023	08/14/2023	🕒	\$ 8,619.00	\$ 9,911.85

Invoices / INV/2023/00094 Print Action 41 / 80 **New**

INV/2023/00094

<p>Customer The Jackson Group, Gordon Owens 1611 Peony Dr Tracy CA 95377 United States</p>	<p>Invoice Date 08/09/2023 Payment Reference ? INV/2023/00094 Due Date 12/03/2023 Journal Customer Invoices in USD</p>
--------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------

Invoice Lines			Journal Items	Other Info			
Product	Label	Account	Analytic	Quantity	Price	Taxes	Subtotal
[FURN_1118] Escritorio con espejo	[FURN_1118] Escritorio con espejo	400000 Product Sales		100.00	85.00	Tax 15%	\$ 8,500.00

Terms and Conditions

Untaxed Amount:	\$ 8,500.00
Tax 15%:	\$ 1,275.00
Total:	\$ 9,775.00

Amount Due: **\$ 9,775.00**

Para finalizar tenemos el apartado de compras de las tiendas en donde adquieren los productos que les distribuyen a los clientes

Bills Search...

NEW [UPLOAD](#)

Filters Group By Favorites

<input type="checkbox"/>	Number	Vendor	Bill Date	Due Date	Reference	Activities	Tax Excluded	Total
<input type="checkbox"/>	BILL/2023/12/0002	Zona Central Chalatenango	12/05/2023	In 65 days			\$ -80,000.00	\$ -92,000.00
<input type="checkbox"/>	BILL/2023/12/0001	Zona Central La Libertad	12/02/2023	In 65 days			\$ -735.00	\$ -845.25
<input type="checkbox"/>	BILL/2023/11/0009	Zona Occidental Sonsonate	11/29/2023	In 34 days			\$ -630.00	\$ -724.50
<input type="checkbox"/>	BILL/2023/11/0008	Zona Central Chalatenango	11/25/2023	In 34 days			\$ -320.00	\$ -368.00
<input type="checkbox"/>	BILL/2023/11/0007	Zona Oriental San Miguel	11/20/2023	In 34 days			\$ -22,400.00	\$ -25,760.00
<input type="checkbox"/>	BILL/2023/11/0006	Zona Occidental Sonsonate	11/18/2023	In 34 days			\$ -1,092.00	\$ -1,255.80
<input type="checkbox"/>	BILL/2023/11/0005	Zona Occidental Sonsonate	11/15/2023	In 34 days			\$ -15,200.00	\$ -17,480.00
<input type="checkbox"/>	BILL/2023/11/0004	Zona Central Chalatenango	11/10/2023	In 34 days			\$ -14,400.00	\$ -16,560.00
<input type="checkbox"/>	BILL/2023/11/0003	Zona Central Chalatenango	11/08/2023	In 34 days			\$ -6,000.00	\$ -6,900.00
<input type="checkbox"/>	BILL/2023/11/0002	Zona Occidental Sonsonate	11/05/2023	In 34 days			\$ -980.00	\$ -1,127.00
<input type="checkbox"/>	BILL/2023/11/0001	Zona Central, La Libertad	11/03/2023	11 days ago			\$ -75.00	\$ -86.25
<input type="checkbox"/>	BILL/2023/10/00009	Zona Central Chalatenango	10/30/2023	In 3 days			\$ -600.00	\$ -690.00
<input type="checkbox"/>	BILL/2023/10/00008	Zona Central San Salvador	10/28/2023	In 3 days			\$ -2,400.00	\$ -2,760.00
<input type="checkbox"/>	BILL/2023/10/00007	Zona Central La Libertad	10/23/2023	In 3 days			\$ -380.00	\$ -437.00
<input type="checkbox"/>	BILL/2023/10/00006	Zona Oriental San Miguel	10/20/2023	In 3 days			\$ -4,699.50	\$ -5,404.43
<input type="checkbox"/>	BILL/2023/10/00005	Zona Central La Libertad	10/15/2023	In 3 days			\$ -98,724.00	\$ -113,532.60
<input type="checkbox"/>	BILL/2023/10/00004	Zona Occidental Sonsonate	10/11/2023	In 3 days			\$ -460.00	\$ -529.00

Bills / BILL/2023/12/0002

Print Action 1 / 80 [New](#)

[REGISTER PAYMENT](#) [ADD CREDIT NOTE](#) [RESET TO DRAFT](#) [DRAFT](#) [POSTED](#)

Vendor Bill
BILL/2023/12/0002

Vendor Zona Central Chalatenango	Bill Date 12/05/2023
Bill Reference	Accounting Date 12/07/2023
	Payment Reference ?
	Recipient Bank ?
	Payment terms 30 days End of Month on the 10th
	Journal Vendor Bills in USD

[Invoice Lines](#) | [Journal Items](#) | [Other Info](#)

Product	Label	Asset Categ...	Account	Analytic	Quantity	Price	Taxes	Subtotal
[E-COM07] Silla ger [E-COM07]	Silla gerencial		600000 Expenses		100.00	800.00	Purchase Tax 15%	\$ 80,000.00