



Tarea 2

Transport Operation Management

Introducción

Una empresa con operación en todo el país, esta enfrentando retos para poder optimizar y eficientizar su operación logística. La información disponible se encuentra en una base de datos dentro de un servidor en Azure:



SERVER: server-training.database.windows.net

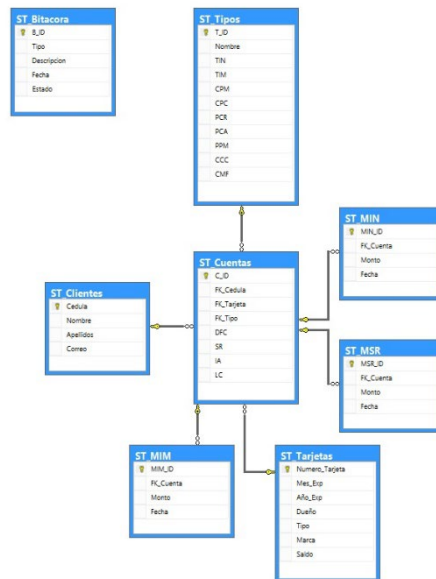
DB: db_unam

USER: unam_reader

PSWD: d@tamining_iim@s22

Base de Datos

Las tablas que se encuentran en la base de datos son las siguientes:



Customer: Contiene la información en conjunto de los clientes a los cuales se distribuye y la información de los proveedores de servicios de transporte (la empresa no cuenta con vehículos propios por lo que tiene que tercerizar el servicio de envío)

Equipment: Contiene la información de los tipos de transportes que usa la empresa para los envíos

Shipment: Contiene la información detallada de cada envío con un número significativo de variables, de las cuales vamos a describir las siguientes:

- **shipmentXid:** ID del shipment
- **totalWeight:** Total de peso en la carga del envío
- **totalWeight_uom_code:** Unidad de medida de peso
- **totalNetVolume:** Total de volumen en la carga del envío
- **totalVolume_uom_code:** Unidad de medida de volumen
- **weightUtilization:** % Utilización del peso
- **volumeUtilization:** % Utilización del volumen
- **totalShipUnitCount:** % Utilización del volumen
- **servprovGid:** ID del proveedor de servicio
- **destLocationGid:** ID del cliente al cual se mandó el envío
- **plannedCost:** Costo base del envío
- **totalDeclaredValue:** Monto del total de ganancia que genera el envío
- **firstEquipmentGroupGid:** ID del tipo de transporte
- **totalActualCost:** Costo real del envío (algunas veces hay ajustes y se tiene que tomar este campo con el valor del costo para todos los cálculos que se van a realizar)
- **insert_date:** Fecha en la que se genero la orden de envío
- **LastUpdatedDateTime:** Ultima fecha/hora de modificación del envío (En clase vimos porque es muy importante esta columna)

Invoice: Contiene la información de las facturas generadas de cada envío, algo importante a considerar es que aquí es donde se detalla el costo de los *Cargos Adicionales* de cada envío, los cuales son gastos que ocurren durante el transporte de la mercancía y la factura se genera una vez llegado el envío a su destino final, de las cuales vamos a describir las siguientes:

- **invoiceXid:** ID de la factura
- **consolidationType:** Tipo de factura del pedido (Standard: Completo, se genera a 1 cliente, Parent/Child: Consolidado, se genera a 2 clientes o más, ¿Cuál es la diferencia entre c/u?)
- **otherCharge:** Costo total de los cargos adicionales
- **attribute5:** ID del envío
- **dateReceived:** Fecha de recibo de la factura
- **LastUpdatedDateTime:** Ultima fecha/hora de modificación de la factura

Invoice Status: Contiene la información del status de la factura, desde la solicitud hasta la entrega (Esta tabla es opcional, puede alimentar tu análisis, pero no es necesario usarla)

Order Release: Información detallada de la orden de cada cliente, un envío puede llevar más de una orden y como ya vimos a más de un cliente, aquí se encuentra el **valor real de la ganancia** del envío. Hay que tomar en cuenta que cada orden va a llevar un número de productos los cuales se encuentran en la tabla de *s_ShipunitLine*. La suma de los valores de cada orden debería de dar el total que se presenta en la tabla de *Shipment*, pero esto no siempre es así, esta tabla es la que da el **valor correcto**. A continuación, vamos a describir los siguientes campos:

- **orderReleaseXid:** ID de la orden
- **totalDeclaredValue:** Monto de la ganancia total del envío
- **totalItemPackageCount:** Total de productos en la orden

s_ShipunitLine: Contiene la información de los productos dentro de cada orden

- **shipUnitGid:** ID del producto
- **orderReleaseGid:** ID de la orden
- **itemPackageCount:** Total de productos ordenados
- **packagedItemGid:** ID del producto (solo tenemos el ID, no la descripción del producto)
- **declared_value:** Monto de la ganancia por producto, si sumas este monto por orden debe cuadrar con la tabla de *order_release*

shipmentStopD: Contiene la información de los productos que se envían en un Shipment (Esta tabla es un puente entre las tabla de productos, la tabla de ordenes y la tabla de envíos)

- **shipmentGid:** ID del envío
- **shipUnitGid:** ID del producto
- **declared_value:** Monto de la ganancia por producto, si sumas este monto por orden debe cuadrar con la tabla de *order_release*
- **LastUpdatedDateTime:** Ultima fecha/hora de modificación de la factura

Indicadores Principales

- **Cost to Serve:** Representa % del Costo total entre el valor total de Ganancia por envío, es la métrica que nos servirá para medir el performance.
- **Truck Fill Rate:** Representa el % de llenado de los vehículos de transporte para cada uno de los envíos
- **¿Qué otros indicadores podríamos construir?**

Reto

1. Trazar el diagrama entidad relación de la base de datos
2. Dar un status general del sistema de logística actual, es decir, estadística descriptiva por estado, por cliente, por mes, etc. Puedes usar un Dashboard y conectarlo a SQL :O!! o si prefieres puedes usar la herramienta que mejor te convenga :). **¿Dónde están las áreas de oportunidad? y que recomendaciones y sugerencias darías.**
3. El trabajo es por equipos de **4 integrantes** y se debe presentar los entregables con la propuesta, estos pueden ser una PPT, un Dashboard, un Jupyter, etc. etc. Esta vez no pasaremos a exponer los resultados equipo por equipo, ya que, como la tarea anterior, es la misma base para todos :(. Pero discutiremos en clase los hallazgos encontrados. **Fecha de entrega miércoles 28 de septiembre del 2022**

Recomendaciones

- Las tablas contienen muchos más campos de los descritos aquí, trata de enfocarte solo en lo que se está describiendo
- Tienes que conectar las tablas para traer la información que necesitas de cada una de ellas
- Habrá mucha información que podría estar en NULL o vacía, esto es un tema directo de calidad de datos, no te preocupes por ello, es algo a lo que siempre te vas a enfrentar. Pero trata de ver de que manera lo puedes manejar, es buena practica solo trabajar con la información que necesites
- Trata de construir un **perfil de envío (shipment)** con toda su información (a que Cliente fue, a que Destino, cual Proveedor de servicio se usó, Fecha, Costo Total, Ganancia, Número de Ordenes, Número de Productos
- A veces usar solo SQL es muy complicado, puedes apoyarte de otras herramientas ;)