

# Proyecto – Tablero Auto Partes

## Instrucciones

El siguiente es el enunciado del proyecto de Análisis de Datos para la creación de un tablero de análisis OLAP de Ordenes y Cotizaciones de la empresa Auto Partes, el cual permita visualizar diferentes indicadores de desempeño en ventas y sus métricas asociadas.

El sistema contará con las siguientes funcionalidades

- Procesos y tareas ETL (o ELT) que permitan cargar la información de archivos y hojas de cálculo a bases de datos relacionales.
- Procesos y tareas ETL (o ELT) que creen la estructura de un modelo DWH (estrella o copo de nieve) desnormalizado y orientado al análisis de información.
- Procesos y tareas para crear/actualizar un cubo OLAP (Jobs)
- Implementación de reportes y tablero de control sobre Power BI

Para el proyecto descrito a continuación deberá crear los paquetes de Integration Services, la base de datos de Integration Services (SSISDB) en dicho servidor, crear y programar un “job” (tarea programada) que permita su ejecución según una calendarización definida y crear y diseñar las bases de datos que contendrán la información de los indicadores que se definen en el enunciado

El detalle de cada uno de estos se encuentra a continuación:

- Los paquetes de Integration Services creados por medio del presente proyecto deben encargarse de ejecutar procesos ETL (o ELT, justifique cuál de los dos tipos de procesos elegirá, y por qué), en los cuales se realizará la carga de los datos que se proporcionan por medio de la base de datos en la nube.
- Deberá diseñar un cubo usando SQL Analysis Services que se encontrará también en este servidor, dicho cubo tendrá como medidas aquellos valores numéricos que se desean analizar y el cubo debe también crear jerarquías correspondientes a dimensiones como las localizaciones y las fechas.
- Debe crear un paquete de SSIS que sirva para actualizar el modelo del DWH por medio de SCDs y Merge hacia la tabla de hechos, así como también la actualización del cubo OLAP en SSAS.
- Por último, los reportes y dashboards de Power BI se conectarán a las fuentes de datos para mostrar las métricas e información por medio de visualizaciones interactivas.

## Fuentes de datos

La fuente de datos será la base de datos “RepuestosWeb”, cada grupo tendrá acceso a una copia en un servidor en la nube de la cual le entregaran credenciales para acceder a la misma.

También pueden descargar un backup de dicha base de datos en el enlace [“RepuestosWebProyecto.bak”](#) esto con el objetivo de que cada grupo también pueda restaurar esta base de datos en su instancia local para trabajar de forma “offline”

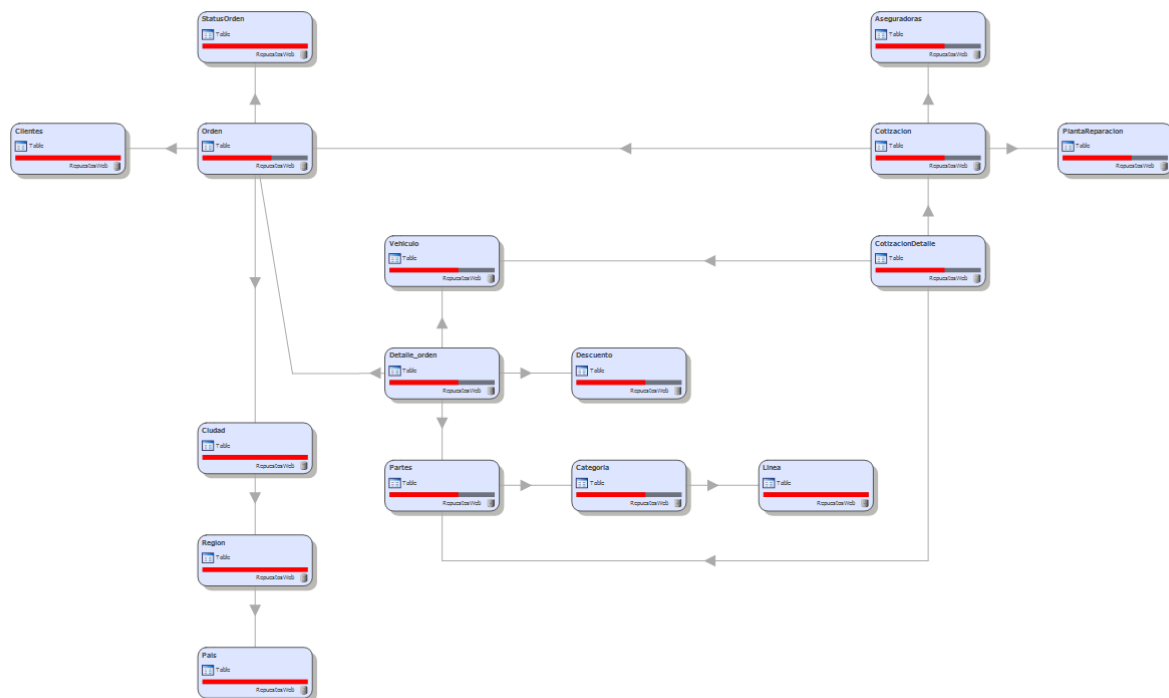
### Nota Importante

La principal fuente de datos será la base de datos “RepuestosWeb” la cual se ha visto en clase anteriormente, sin embargo, la copia que se entrega para la elaboración del presente proyecto provee de objetos adicionales, así como también datos de miles de ordenes que servirán para las pruebas y realización del proyecto.

### Descripción de los datos

La base de datos “RepuestosWeb” es usada por la empresa Auto Partes para llevar un registro de Ordenes y Cotizaciones realizadas por clientes y Aseguradoras para la adquisición de partes de vehículos desde sus plataformas de comercio electrónico.

A continuación, se presenta el diagrama de objetos OLTP de la base de datos encargada de procesar cotizaciones y ordenes:



Auto Partes posee tres canales principales para la venta y generación de las órdenes de compra:

- Aseguradoras – Canal B2B en la cual empresas de seguros realizan cotizaciones y ordenes
- Clientes registrados – Canal B2C en la cual personas que se registras por medio de una cuenta realizan compras.
- Clientes No registrados – Igual que el anterior, pero en este las personas pueden optar para comprar como “Invitado” sin registrarse ni brindar muchos datos.

## Descripción de objetos

A continuación, encontrará una breve descripción de cada uno de los objetos de la base de datos OLTP:

### Tablas catalogo

- ☐ Partes
  - Una de las tablas principales ya que contiene la información de los productos, así como sus precios. Actualmente la gerencia no tiene forma de llevar un registro histórico de sus precios para consultas históricas.
- ☐ Vehículo
  - Información del vehículo del cual se cotiza o vende una parte, la información de esta es bastante relevante porque se busca encontrar que Marcas o modelos son los más rentables
- ☐ Aseguradora
  - Información de empresas de seguros de vehículos que cotizan con Auto Partes, dichas cotizaciones son gestionadas por medio de plantas de reparación de asegurados que buscan reparar sus vehículos.
- ☐ Clientes
  - Listado de clientes que se han registrado desde el sitio web de ventas para ordenar productos
- ☐ Ciudad/Region/Pais
  - posee información geográfica de las diferentes regiones en las cuales se han enviado productos anteriormente
- ☐ Descuento
  - Posee información de descuentos promocionales que se pueden aplicar a una orden, una de las cosas que busca la gerencia es evaluar cuantos descuentos se han aplicado históricamente
- ☐ Categoria/Linea
  - Estas tablas poseen categorizaciones jerárquicas de la tabla de Partes, ya que agrupan los productos individuales según diferentes categorías tales como “Llantas”
- ☐ PlantaReparacion
  - Información de plantas de reparaciones que envían cotizaciones a Auto Partes en representación de una empresa de seguros
- ☐ StatusOrden
  - Información del status de la orden

## Tablas transaccionales

### ☐ Cotizacion

- Las cotizaciones son generadas únicamente por las aseguradoras, las cuales antes de generar una orden de compra con nosotros cotizan nuestros precios para evaluar con la competencia, esto implica que no todas las cotizaciones generan una Orden con nosotros, sin embargo, se desea evaluar el desempeño de que tantas cotizaciones resultan en órdenes de compra.

Cuando una cotización produce una Orden de compra, esta se refleja en la columna “IDOrden”

### ☐ CotizacionDetalle

- El detalle de la cotización que posee la información del vehículo y listado de productos de los cuales la aseguradora solicitó información. Uno de los aspectos que interesa es encontrar que tipo de Vehículos son los mas rentables (por ejemplo, los vehículos Toyota)

### ☐ Orden

- Esta tabla almacena las Ordenes realizadas a nuestra empresa, las cuales se traducen en ventas, motivo por el cual es una de las principales objeto de análisis para la gerencia.

Una de las cosas que necesita la empresa es conocer que tantas ordenes provienen de los diferentes canales (Aseguradora, Clientes Registrados, Invitados).

Cuando una Orden está relacionada con un registro en la tabla de “Cotización” esta proviene de una “Aseguradora”, de lo contrario esta posee un valor en el campo “IDCliente” cuando proviene de un “Cliente Registrado”.

Si ninguna de las anteriores condiciones es verdadera entonces esta orden proviene de un Invitado.

### ☐ Detalle\_Orden

- El detalle de la orden es parecido a la tabla “Cotización\_Detalle” sin embargo esta contiene la información de todas las ordenes de las diferentes fuentes (Aseguradora, Clientes Registrados, Invitados).

Es importante notar que las ordenes que tienen una cotización no necesariamente tienen la misma cantidad de productos, por ejemplo, una cotización pudo haber solicitado información de 3 partes pero la orden resulto en la compra de 2.

## Partes del proyecto

Para integrar el de Ordenes y Cotizaciones de principio a fin, se deben de crear tareas y procesos que procesen la data desde la base de datos fuente en la nube, hasta llegar a los reportes/dashboards y para esto se desarrollaran las siguientes partes del proyecto en un orden secuencial, integrando SSIS, SSAS y Power BI:

- Procesos ETL (SSIS) que cargaran los datos diariamente desde la base de datos en la nube, estos procesos ETL pueden contener (pero no limitado) a las siguientes características:
  - Implementación de proyecto de SSIS sobre el SSIS Catalog en la instancia SQL Server
  - ETL que actualice (únicamente de forma incremental) la data nueva recibida (por ejemplo, usando tablas “staging” para recibir la data y luego sincronizar solo la data nueva)
  - Tarea calendarizada en el SQL Agent que se encargue de actualizar una vez al día los datos nuevos
  - Tablas de control/log que especifican información tal como “Registro de ejecuciones”, “Fechas de ejecución”, “Cantidad de registros importados” etc..

**Importante:** Realizar un proceso que permita una carga Incremental es **sumamente recomendado**, de lo contrario los procesos de actualización diarios serán poco eficientes lo que repercutirá en las pruebas realizadas sobre dichos procesos.

- Procesos ETL (SSIS) para carga de la base de datos relacional a una base de datos de modelado multidimensional lógico (RepuestosWeb\_DWH) que contenga (pero no limitado) a las siguientes características:
  - Implementación de proyecto de SSIS sobre el SSIS Catalog en la instancia SQL Server
  - Implementación de dimensión fecha
  - ETL que actualice (únicamente de forma incremental) la data nueva recibida (por ejemplo, SCDs para las dimensiones y MERGE para el Fact)
  - Selección de SCDs tipo 0, 1 y 2, el enunciado describe ciertos atributos de los cuales se quiere mantener registro histórico, pero aplique también criterio para seleccionar que tipo de SCD aplicar.
  - Tarea calendarizada en el SQL Agent que se encargue de actualizar una vez al día los datos nuevos
  - Tablas de control/log que especificación información tal como “Registro de ejecuciones”, “Fechas de ejecución”, “Cantidad de registros importados/nuevas” etc..
- Carga de la base de datos multidimensional lógica a una implementación física, es decir a un cubo OLAP (SSAS), que contenga (pero no limitado) a las siguientes características:
  - Al menos un cubo que integre estadísticas de casos reportados con medidas tales como (pero no limitado a)

- Cantidad de Ordenes
  - Cantidad de Cotizaciones
  - Total de las Ordenes
  - Total de Cotizaciones
  - Promedio de Orden
  - Promedio de Cotización
  - Total de Partes vendidas
  - Promedio de Partes vendidas
  - Etc...
- o Dimensiones que permitan filtrar la data tales como (pero no limitado a)
  - País/Región/Ciudad
  - StatusOrden
  - Origen de orden (Aseguradora, Cliente Registrado y no registrado)
  - Vehiculos
  - Etc... (prácticamente cualquier tabla catlogo)
- o KPIs que muestren indicadores relevantes tales como (pero no limitado a)
  - Cantidad de productos por Orden
  - Total de la Orden vs Promedio de Orden
  - Etc...
- o Jerarquías relevantes en las dimensiones tales como (pero no limitado a)
  - Año/Trimestre/Mes
  - País/Region/Ciudad
  - Año/Marca/Modelo
  - Etc...
- Reportes y Tablero de control de visualizaciones de estadísticas de Ordenes y Cotizaciones, que contenga (pero no limitado) a las siguientes características:
  - o Gráficos de medidas de casos
    - Métricas totales de (pero no limitado) a
      - Ordenes
      - Cotizaciones
      - Descuentos
      - Etc..
    - Métricas de promedios de (pero no limitado) a
      - Ordenes
      - Cotizaciones
      - Cantidad de partes
      - Etc..
    - Gráficos con métricas agrupadas por (pero no limitado) a
      - Vehículos
      - Partes
      - País/Región/Ciudad
      - Etc..
    - Gráficos que muestren información a través del tiempo con datos tales como:
      - Cantidad de Ordenes por año/mes
      - Cantidad de Cotizaciones año/mes
      - Etc...

Agregue pronósticos para los gráficos de fechas

- Gráficos que muestren información geográfica en mapas tales como (pero no limitado) a
  - Cantidad de Ordenes
  - Totales de Ordenes
  - Etc...
- Dashboards (tableros de control) en el servicio de la nube de Power BI que integren visualizaciones de los diferentes reportes descritos anteriormente

Sean creativos con la creación de reportes y dashboards, no existe un número mínimo o máximo de reportes/dashboards sin embargo se espera que use la mayoría de visualizaciones disponibles en Power BI para mostrar estadísticas de diferentes formas (fechas, conteos, mapas, etc..)

## Entregables

Los entregables esperados son:

- Folder de proyecto en SSISDB
  - o Es importante incluir Scripts para detener e iniciar los servicios de integration run time para no incurrir en costos innecesarios
- Proyectos SSIS
- Base de datos relacional y modelo DWH (Esquema estrella o copo de nieve)
- Proyectos SSAS
- Scripts de Jobs (tareas calendarizadas) sobre SQL Server Agent para su creación en otro servidor, que permitan actualizar todas las bases de datos (RepuestosWeb\_DWH y Cubo OLAP)
- Listado de reportes Power BI publicados en servidor
- Tareas calendarizadas para actualizar reporte de Power BI por medio de un Gateway.
- Scripts de creación de objetos (tales como base de datos DWH, Jobs SQL Agent, etc..)
- Manual técnico – explicando el diseño de cada objeto (paquete SSIS, proyecto SSAS, reportes de Power BI, etc..) así como también como se actualiza la información de los reportes por medio de los “Jobs” que refrescan la información
- Manual de usuario – explicando a un usuario no técnico como interactuar con los objetos (los reportes principalmente) y la información de estos.

## Fecha de entrega

Lunes 25 de Octubre del 2021 a las 23:55 horas, para dicha fecha deberá de proveer lo siguiente:

- Subir al portal URL por medio del link llamado “Entregables Proyecto 1” un archivo comprimido con:
  - Carpeta Proyectos SSIS -> contiene el proyecto de Visual Studio con todos los paquetes ETL o ELT que actualizan los datos e información.  
  
Favor parametrizar los connection strings para su fácil implementación.
  - Carpeta Proyectos SSAS -> contiene el proyecto de Visual Studio con el/los cubos y todas sus partes
  - Carpeta reportes -> contiene los archivos de PowerBI .pbix o .pbix con todos y cada uno de los reportes/dashboards creados
  - Manual técnico
  - Manual de usuario
- Base de datos relacional, Modelo DWH y Cubo OLAP en el servidor SQL de la nube.

El proyecto puede realizarse en grupos de un máximo de 4 integrantes que ya fueron previamente creados.