



PROYECTO NO. 1

1. OBJETIVOS

- **General**
 - Utilizar consultas para presentar la información requerida de una base de datos relacional.
- **Específicos**
 - Desarrollar la habilidad de generar consultas avanzadas en lenguaje SQL.
 - Mostrar información específica de una base de datos.
 - Aprender a realizar cargas masivas desde archivos separados por coma (csv) a una base de datos.
 - Desarrollar el modelo entidad relación a partir de un conjunto de datos.

2. DESCRIPCIÓN DEL PROYECTO

El centro de datos "El Eje del Mundo" lo contrata a usted para realizar reingeniería al proceso de almacenamiento que actualmente manejan para guardar la información de las diferentes compañías que tiene a cargo ya que actualmente hay demasiada inconsistencia que puede provocarles serios problemas legales. El centro de datos hace el esfuerzo de exportar toda la información actual a un archivo delimitado por comas (csv) que usted tendrá a disposición para realizar el análisis correspondiente con el objetivo de eliminar redundancia de los datos y crear un nuevo modelo de base de datos estable para almacenar la información.

3. ESTRUCTURA DEL ARCHIVO CSV

A continuación se muestra cómo está estructurado el archivo separado por comas que contiene los datos a cargar. Con esta información se debe crear el modelo de datos que permita realizar los reportes de la sección 5. El archivo puede ser descargado desde el UEDI en la sección Proyectos.

Columna	Descripción
nombre_compania	Nombre de la compañía <i>Tipo: texto</i>
contacto_compania	Nombre de la persona con la que se tiene contacto dentro de la empresa. <i>Tipo: texto</i>
correo_compania	Correo electrónico de la compañía. <i>Tipo: texto</i>
telefono_compania	Teléfono de la compañía. <i>Tipo: texto</i>
tipo	Carácter que representa si los siguientes campos (nombre, correo, teléfono y campos de dirección) corresponde a un Cliente [C] o a un proveedor [P]. <i>Tipo: carácter</i>
nombre	Nombre del cliente ó proveedor. <i>Tipo: texto</i>
correo	Correo electrónico del cliente ó proveedor. <i>Tipo: texto</i>
teléfono	Teléfono del cliente ó proveedor. <i>Tipo: texto</i>
fecha_registro	Fecha en la que se registro al sistema el cliente ó proveedor.

	<i>Tipo: fecha</i>
dirección	Dirección del cliente ó proveedor. <i>Tipo: texto</i>
ciudad	Ciudad de donde es el cliente ó proveedor. <i>Tipo: texto</i>
codigo_postal	Código postal de la dirección del cliente ó proveedor. <i>Tipo: numérico</i>
región	Región correspondiente de la ciudad de donde es el cliente ó proveedor. <i>Tipo: texto</i>
producto	Nombre del producto que se compró/vendió según corresponda. <i>Tipo: texto</i>
categoria_producto	Categoría a la que pertenece el producto. <i>Tipo: texto</i>
cantidad	Número que representa la cantidad de producto adquirido/vendido. <i>Tipo: numérico</i>
precio_unitario	Costo que representa el adquirir una única unidad de un producto en específico. <i>Tipo: decimal</i>

4. SOLICITUD

Una vez analizado el archivo anterior, se solicita que usted diseñe un nuevo modelo de datos para almacenar toda esa información. El proceso es el siguiente:

1. Debe crear una tabla temporal en donde debe cargar toda la información del CSV. La herramienta que debe utilizar es SQL*Loader, la cual es propia de Oracle y será enseñado su funcionamiento en el laboratorio.
2. Haciendo uso de consultas a la tabla temporal debe de llenar el modelo que usted propone para almacenar la información de forma correcta de modo que no exista redundancia de los datos.
3. Una vez lleno todo el modelo, debe realizar los reportes de la sección 5 los cuales serán evaluados para verificar que la información mostrada sea consistente y verídica.

5. REPORTES

Se deben de entregar los siguientes reportes:

1. Mostrar el nombre del proveedor, número de teléfono, número de orden, total de la orden por la cual se haya pagado la mayor cantidad de dinero.
2. Mostrar el número de cliente, nombre, apellido y total del cliente que más productos ha comprado.
3. Mostrar la dirección, región, ciudad y código postal hacia la cual se han hecho más solicitudes de pedidos y a cuál menos (en una sola consulta).
4. Mostrar el número de cliente, nombre, apellido, el número de órdenes que ha realizado y el total de cada una de los cinco clientes que más han comprado productos de la categoría 'Cheese' .
5. Mostrar el número de mes de la fecha de registro, nombre y apellido de todos los clientes que más han comprado y los que menos han comprado (en dinero) utilizando una sola consulta.
6. Mostrar el nombre de la categoría más y menos vendida y el total vendido en dinero (en una sola consulta).
7. Mostrar el top 5 de proveedores que más productos han vendido (en dinero) de la categoría de productos 'Fresh Vegetables' .
8. Mostrar la dirección, región, ciudad y código postal de los clientes que más han comprado y de los que menos (en dinero) en una sola consulta.
9. Mostrar el nombre del proveedor, número de teléfono, número de orden, total de la orden por la cual se haya obtenido la menor cantidad de producto.

10. Mostrar el top 10 de los clientes que más productos han comprado de la categoría **'Seafood'** .
11. Mostrar el porcentaje de clientes que le corresponden a cada región.
12. Mostrar las ciudades en donde más se consume el producto **"Tortillas"** de la categoría **"Refrigerated Items"** .
13. Mostrar la cantidad de clientes que hay en las ciudades agrupadas por su letra inicial, es decir, agrupar las ciudades con A y mostrar la cantidad de clientes, lo mismo para las que inicien con B y así sucesivamente.
14. Mostrar el porcentaje de las categorías más vendidas de cada ciudad de la siguiente manera: Ciudad, Categoría, Porcentaje De Mercado
15. Mostrar los clientes que hayan consumido más que el promedio que consume la ciudad de Frankfurt.

REQUERIMIENTOS Y RESTRICCIONES

- El modelo debe ser desarrollado en la versión de Oracle 12c o mayor en un sistema operativo basado en Linux o utilizar docker.
- Deben utilizar Data Modeler para diseñar el modelo relacional.
- No se debe separar en diferentes archivos csv el archivo proporcionado. Debe ser cargado únicamente el que fue proporcionado.
- Los reportes deben ser realizados a la entidad relación creada por el estudiante, reportes hechos a la tabla temporal serán anulados.
- La práctica es individual.
- Copias totales o parciales tendrán nota de 0 puntos y serán reportados a escuela.
- Entrega tarde tendrán una penalización del 25% de la nota total.

ENTREGABLES

- Los entregables deben de ser los siguientes:
 - El modelo relacional en notación baker propuesto en formato de imagen.
Formato:
[BD1]ModeloPropuesto.png
 - Script (con extensión .sql) que contiene las instrucciones DDL para crear el modelo relacional. A este script debe de aplicarle todas las reglas de integridad vistas en clase y laboratorio. Formato:
[BD1]InstruccionesDDL.sql
 - Archivo de control utilizado para realizar la carga masiva desde el archivo CSV a la tabla temporal. Formato: **[BD1]ArchivoControl.ctf**
 - Script (con extensión .sql) que contiene todas las consultas necesarias para llenar el modelo relacional propuesto a partir de la tabla temporal. Formato:
[BD1]CargaDeDatos.sql
 - Script (con extensión .sql) que contiene todas las consultas necesarias para realizar los reportes. Formato: **[BD1]Consultas.sql**
- Todos los archivos anteriormente mencionados deben de ir dentro de una carpeta comprimida con el siguiente formato:

[BD1]Proyecto1_#carnet.rar

FECHA DE ENTREGA

- Viernes 2 de Octubre de 2020, antes de las 11:59 pm.
- La entrega será a través de la plataforma de UEDl en el post correspondiente al Proyecto No. 1.
- Los horarios de calificación serán notificados en el transcurso de las semanas siendo Sabado 3 y Domingo 4 la calificación.