# Proyecto de Curso: ETL para las dimensiones

El proyecto del curso busca desarrollar una bodega de datos para AdventureWorks, una multinacional dedicada a la venta de bicicletas.

## Actividades

En grupos, realice las siguientes actividades:

**Codifique el ETL para las dimensión tiempo**

* Cree primero una consulta sobre la base de datos transaccional que le permita obtener todas las fechas relacionadas con las transacciones o eventos que son de interés en la bodega de datos. Incluya el nombre de la base de datos con el nombre de la tabla.
  + Por ejemplo:  
    **SELECT DISTINCT(fecha)   
    FROM oltp.transacción**
* Modifique la consulta para obtener la información (por ejemplo, mes, año) que se desea almacenar en las dimensiones. Revise el diseño de la dimensión y determine cuales son las funciones de SQL que les permite obtener esta información. Use la consulta anterior como un subquery o un CTE.
  + Por ejemplo:  
    SELECT

TO\_DAYS(fecha) AS date\_key,

fecha,

DATE\_FORMAT(fecha, '%d %M %Y') AS date\_medium,

MONTH(fecha) AS month,

MONTHNAME(fecha) AS month\_name,

YEAR(fecha) AS year4,

DATE\_FORMAT(fecha, '%Y-%m') AS year\_month\_number

FROM (

**SELECT distinct DATE(fecha) as fecha**

**FROM oltp.transacción**

) AS fechas

* + O, como otro ejemplo:  
    WITH fechas AS (

**SELECT distinct DATE(fecha) as fecha**

**FROM oltp.transacción**

)

SELECT

TO\_DAYS(fecha) AS date\_key,

fecha,

DATE\_FORMAT(fecha, '%d %M %Y') AS date\_medium,

MONTH(fecha) AS month,

MONTHNAME(fecha) AS month\_name,

YEAR(fecha) AS year4,

DATE\_FORMAT(fecha, '%Y-%m') AS year\_month\_number

FROM fechas

* Modifique la consulta para insertar los datos en la dimensión tiempo. Incluya el nombre de la base de datos en el nombre de la tabla
  + Por ejemplo,

INSERT INTO **datawh.dim\_tiempo** (

date\_key,

date\_value,

date\_medium,

month\_number,

month\_name,

year4,

year\_month\_number

)

SELECT …

* Incluya el código en el repositorio Github.

**Codifique el ETL para las otras dimensiones**

* Cree una consulta sobre la base de datos transaccional que le permita obtener los datos de las entidades que corresponden en las otras dimensiones de la bodega de datos. Identifique cuáles son las tablas que corresponden con cada entidad y determine qué columnas corresponden con los datos de las dimensiones.
  + Por ejemplo:  
    **SELECT  
     entidad\_id,  
     entidad\_nombre,  
     clasificacion\_nombre  
    FROM oltp.entidad  
     JOIN oltp.clasificacion USING (clasificacion\_id)**
* Si se requiere, modifique la consulta para hacer cálculos y obtener la información (por ejemplo, días de demora, utilidad de la venta) que se desea almacenar en las dimensiones. Revise el diseño de la dimensión y determine cuales son las funciones de SQL que les permite obtener esta información. Use la consulta anterior como un subquery o un CTE.
  + Por ejemplo:  
    SELECT

entidad\_id,

entidad\_nombre,

FUNCION\_CUALQUIERA(entidad\_nombre) AS nombre\_modificado

clasificacion\_nombre

FROM (

**SELECT**

**entidad\_id,**

**entidad\_nombre,**

**clasificacion\_nombre**

**FROM oltp.entidad**

**JOIN oltp.clasificacion USING (clasificacion\_id)**

) AS datos

* + O, como otro ejemplo:  
    WITH datos AS (

**SELECT  
 entidad\_id,  
 entidad\_nombre,  
 clasificacion\_nombre  
 FROM oltp.entidad  
 JOIN oltp.clasificacion USING (clasificacion\_id)**

)

SELECT

entidad\_id,

entidad\_nombre,

FUNCION\_CUALQUIERA(entidad\_nombre) AS nombre\_modificado

clasificacion\_nombre

FROM datos

* Modifique la consulta para insertar los datos en la dimensión.
  + Por ejemplo,

INSERT INTO datawh.dim\_xxx (

...

)

SELECT ...

* Incluya el código en el repositorio Github.
* Repita este procedimiento para cada dimensión

## Elementos mínimos del informe

(No aplica)