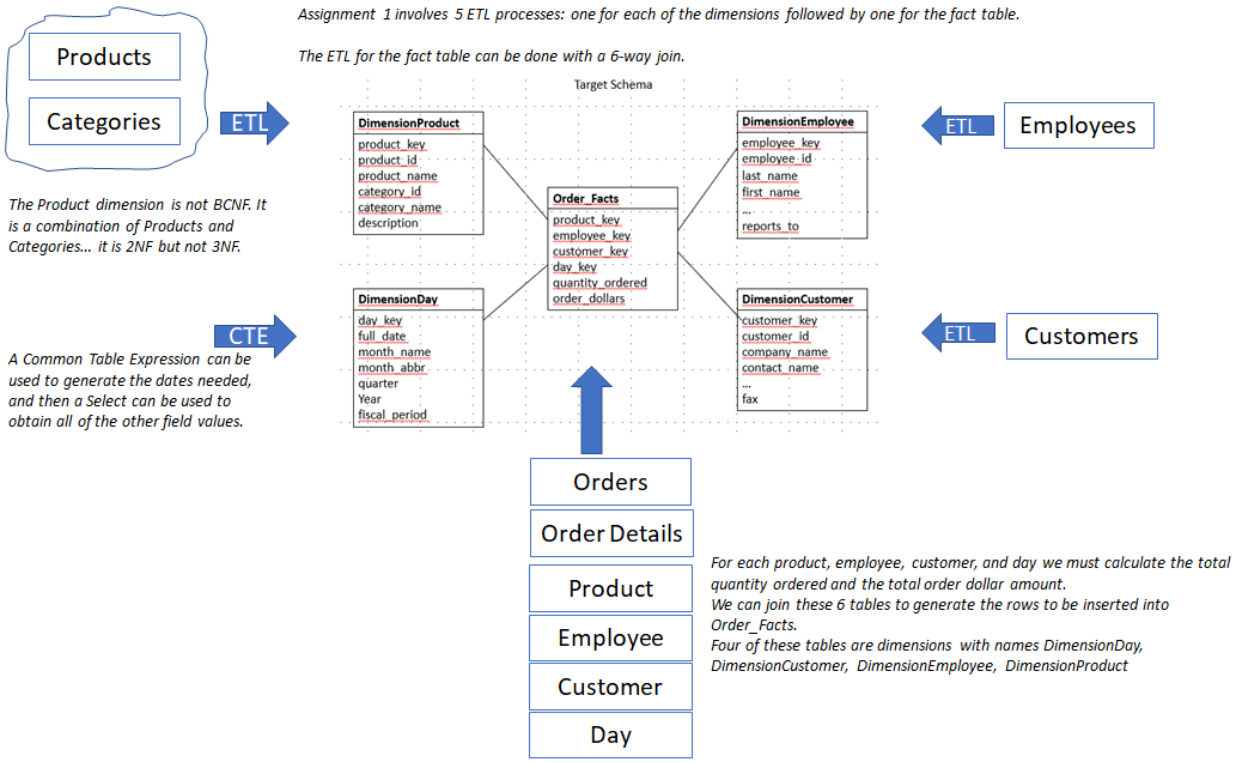
See Nexus to submit your work: Assessments >> Quizzes >> Assignment 1

The purpose of the assignment is to ensure you have code that migrates data from operational tables to a star schema. As you do this, some dimensional design concepts will become clearer. The star schema is modelled after Figure 1-5, but some changes were made to accommodate our sample Northwind database we are using as source.

Regardless of what Nexus may say, you have until September 26 to complete the work. Nexus thinks there is a time limit… all that matters is that you submit by September 26. Any number of attempts is okay, but only the last one is marked.



Files for ETL are available in Nexus

1. Using PostgreSQL, create an operational Northwind database

   https://github.com/pthom/northwind_psql

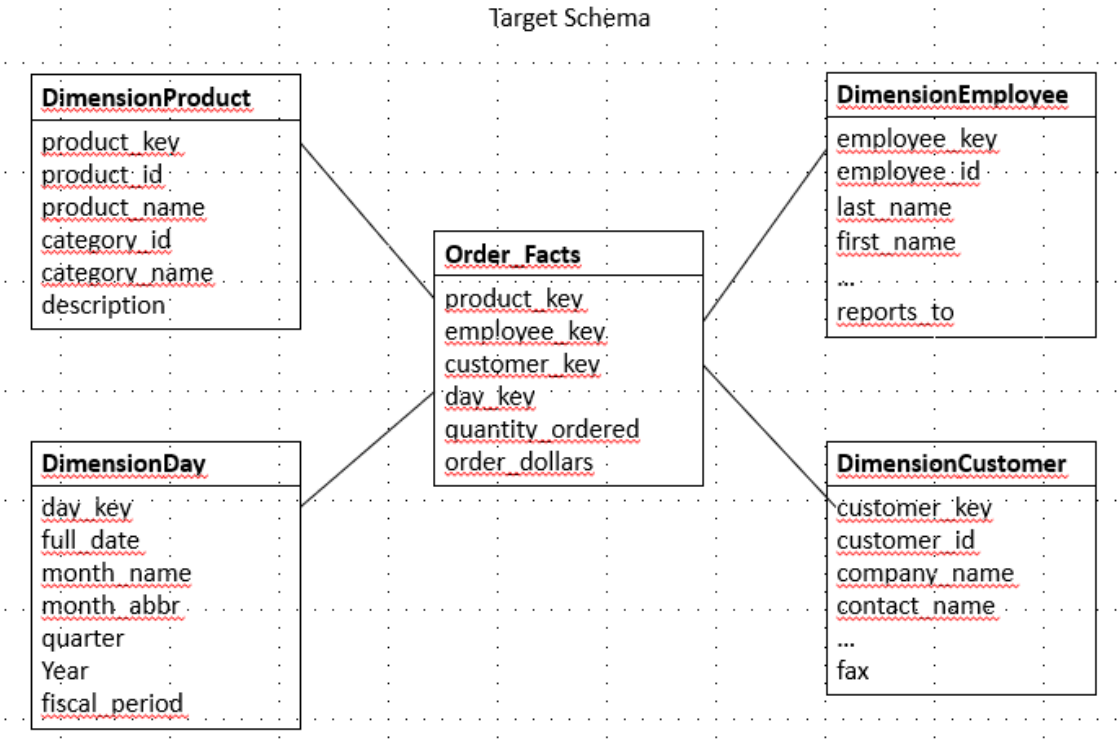
   Note there is no data for a few of the tables.

Verify you have loaded the correct number of rows in the tables. You can use:

```
select sum(NumberOfRows) from (
select 'customers' as Table, count(*) as NumberOfRows from customers
union
select 'employees' as Table, count(*) as NumberOfRows from employees
union
select 'employee_territories' as Table, count(*) as NumberOfRows from
employee_territories
union
select  'order_details' as Table, count(*) as NumberOfRows from order_details
union
select  'orders' as Table, count(*) as NumberOfRows from orders
union
select  'products' as Table, count(*) as NumberOfRows from products
union
select  'region' as Table, count(*) as NumberOfRows from region
union
select  'shippers' as Table, count(*) as NumberOfRows from shippers
union
select  'suppliers' as Table, count(*) as NumberOfRows from suppliers
union
select  'territories' as Table, count(*) as NumberOfRows from territories
union
select 'us_states' as Table,  count(*) as NumberOfRows from us_states
union
select 'categories' as Table,  count(*) as NumberOfRows from categories
        ) x;
```

The result you get should be 3362.

2. Use/modify the sample ETL (see files provided) to extract data from Northwind and populate the target star schema design (similar to Figure 1-5, but adapted for Northwind as the source schema). Store the star schema tables in the same database as Northwind (not ideal, but this simplifies your work).



3. Run this SQL Select for January 1997 facts on your star schema:

```
SELECT  category_name, product_name,
SUM( order_dollars ) AS "ORDER DOLLARS"
FROM DimensionDay, DimensionProduct, order_facts
WHERE
TRIM(month_name) = 'January' AND
year = 1997 AND
order_facts.day_key = DimensionDay.day_key AND
order_facts.product_key = DimensionProduct.product_key
GROUP BY DimensionProduct.category_name, DimensionProduct.product_name
ORDER BY DimensionProduct.category_name, DimensionProduct.product_name
```

4. Run a similar statement on your Northwind database to verify you have the correct results above. Consider:

```
--Northwind query
SELECT  c.category_name, p.product_name,
SUM( od.unit_price*od.quantity ) AS "ORDER DOLLARS"
FROM orders o inner join order_details od on (o.order_id = od.order_id)
inner join products p on (od.product_id = p.product_id)
inner join categories c on (p.category_id = c.category_id)
WHERE
extract(MONTH FROM order_DATE) = 1 and
extract(YEAR FROM order_DATE) = 1997
GROUP BY c.category_name, p.product_name
ORDER BY c.category_name, p.product_name
```

5. Code a query for the star schema to list the total sales for category Beverages, in month of April, by employees who are Sales Representatives, and for customers in France or Germany.

6. Code a query for the original Northwind database to list the total sales for category Beverages, in month of April, by employees who are Sales Representatives, and for customers in France or Germany.

See Nexus for support materials

Contact your instructor if you need assistance

If you have problems with postgreSQL on your computer, there is a version on the ACS lab facility. Contact ACS technical support for assistance.