Loading data into Northwind Data Warehouse from Northwind Database/ PostgreSQL

April 6, 2023

Didem B. Aykurt

Colorado State University Global

MIS541; Data Warehousing in Enterprise Environments

Dr.Peter Salemi

Loading Dimension Data from the Northwind OLTP Database into the NW_Customer_DIM and NW_Employee_DIM Dimension Table

In this project, I will insert data into the two-dimensional table. Before jumping into the technical part, I would like to talk about what kind of business data is contained in the Northwind_DW and how to process the design of this business dimension table. Using the Northwind retail business database and our textbook's fourth-step dimensional design process helps me understand the dimensions. The first step is to select a business process that helps to understand business performance and activities and the dimension describing each business process to identify the organization's business process. For example, the customer dimension explains the customer process, and the employee contains performance metrics. Most business processes affect results by input, which is why most organizations put first process output and then the metrics of the following information. The second is to declare the grain answer question of what each fact table row points to regarding the business process's events or "How to describe a single row in the fact table such as one row containing per sales transaction of a single product. The fact table's primary key shows the grain report business terms. If the grain isn't apparent at the beginning of the design process, all data models rest on a trap. If steps 3 or 4 fail, you may turn to step 2 to create grain. The third step is to identify the dimension answer to how businesspeople explain the data from the business measurement events. If step 2 is clear about grain, it is easy to identify the dimensions. For example, I created a customer and employee dimension table. The last step, identifying the facts, helps to answer the business process measures. Fact table contains different grains from separate dimension tables and the numeric result of data like quantity order or dollar cost. In the previous project, I created a two-dimensional table; now, I will load data from the OLTP Northwind database. The Bellow SQL comment will explain the new function or syntax I will use on this project.

The 'INSER INTO table_name VALUE (....,...)' statement helps to insert or add values for all the columns or specific columns of the table.

The 'SELECT * FROM table_name' statement displays all the available fields in the table.

Figure 1: Loading data into NW_Customer_DIM with PostgreSQL.

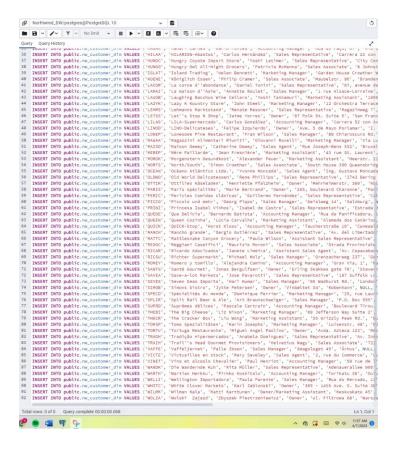


Figure 2: Display the fields in the NW_Customer_DIM dimension table.

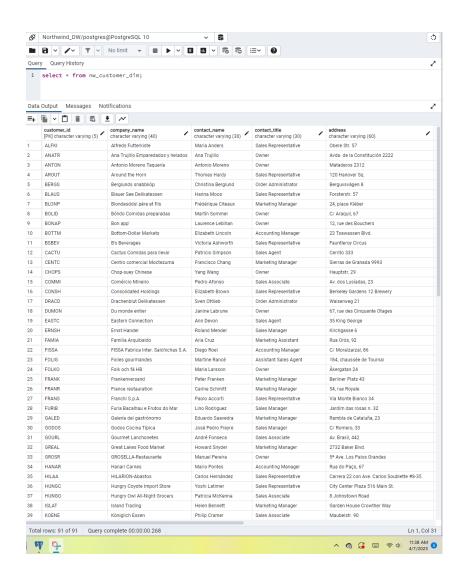
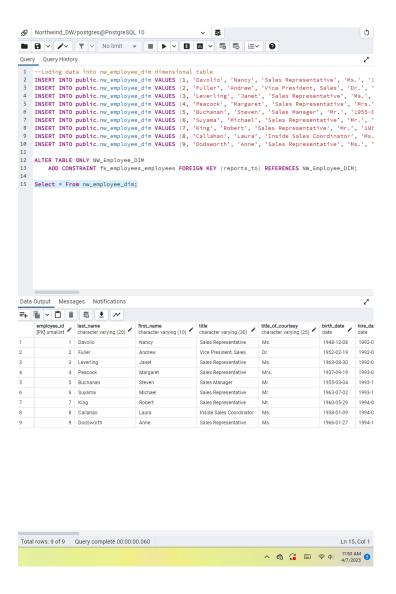


Figure 3: Insert and display the NW_Employee_DIM dimension table



Concern:

When loading data from OLTP into OLAP, should we insert data manually, as I did in this project, for each value? OLTP database contains data from the direct App on time; the business user wants to see the report, and the current data warehouse can't support report needs. Can I create a new data warehouse with new dimensions and a fact table to insert a million fields into the new data warehouse? There should be an automatic update of the data system from OLTP to OLAP system.

References

Sebastian Insausti, 2019. Running a Data Warehouse on PostgreSQL.

https://severalnines.com/blog/running-data-warehouse-postgresql/

PostgreSQL.org https://www.postgresql.org

Ralph Kimball & Margy Ross, 2013. The Data Warehouse Toolkit The Definitive Guide to Dimensional Modeling. 3rd edition.