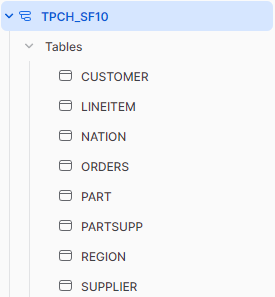
**LAB – 1: DATA PREPARATION FOR STAR SCHEMA**

**DATA WAREHOUSE:**

* A data warehouse is a system used for reporting and data analysis, and is considered a core component of business intelligence.
* It is a central repository of integrated data from one or more disparate sources, and is designed to enable and support business intelligence (BI) activities, especially analytics.
* Key characteristics of Data Warehouse are:
  1. Integration
  2. Subject-oriented
  3. Time-variant
  4. Non-volatile

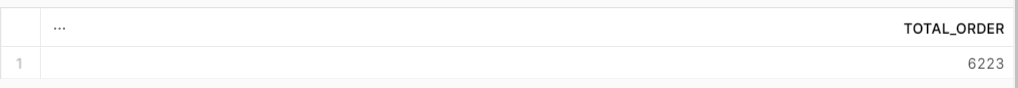
**LAB PROCEDURES:**

1. **Create account in Snowflake.**
   1. Snowflake is a data platform that would harness the immense power of the cloud.
   2. Simply, it is a Data Warehouse.
   3. We created our account in this platform.
2. **Understand TPCH\_SF10 Schema.**
   1. TPCH\_SF10 Schema is a sample data provided by Snowflake.
   2. This schema contains 8 tables.
   3. The table is provided below:



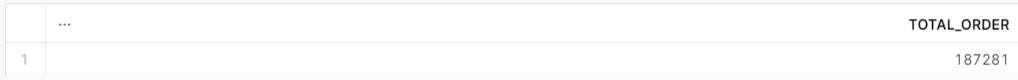
1. **Write SQL statements for the following:**
   1. **What is the total number of orders received in “1995-9-16”?**

SELECT COUNT(\*) AS TOTAL\_ORDER FROM ORDERS WHERE O\_ORDERDATE = '1995-09-16';



* 1. **What is the total number of orders received in September 1995?**

SELECT COUNT(\*) AS TOTAL\_ORDER FROM ORDERS WHERE MONTH(O\_ORDERDATE) = '9' AND YEAR(O\_ORDERDATE) = '1995';

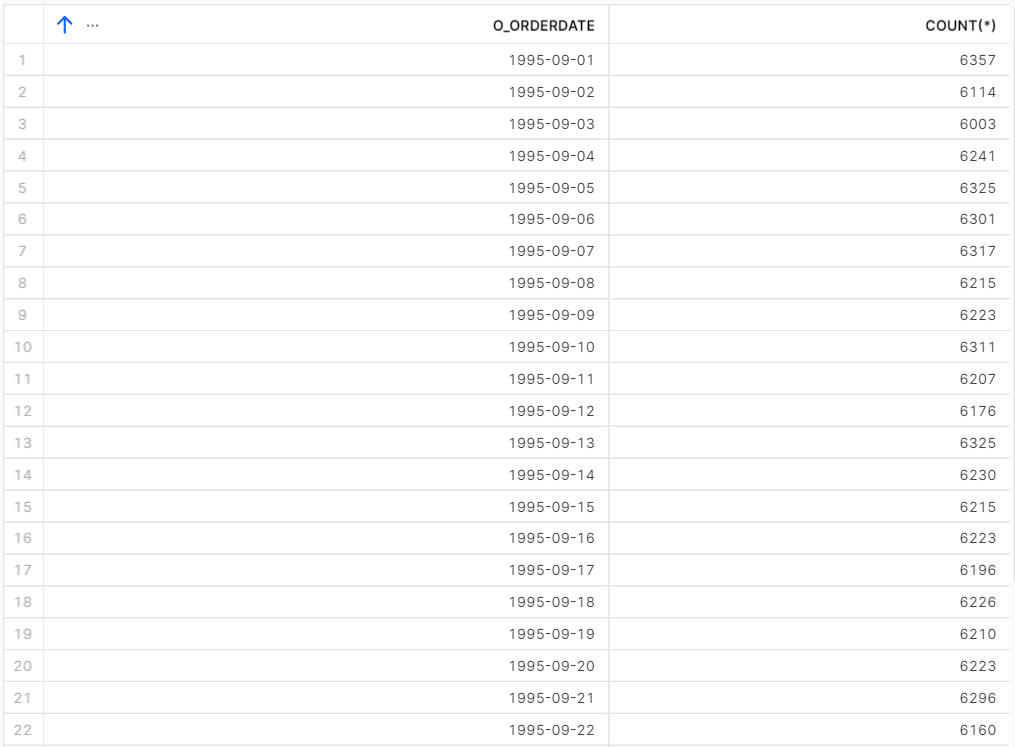


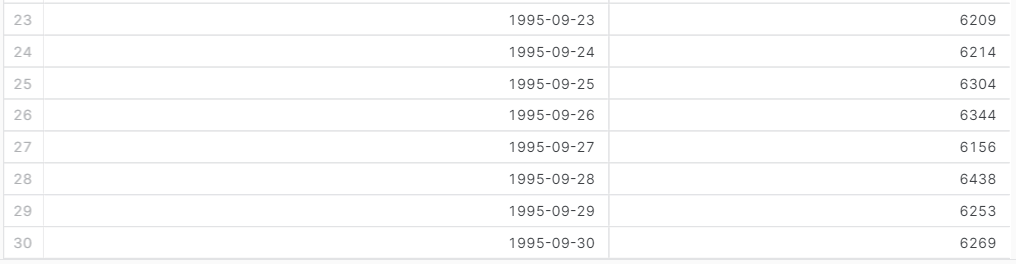
* 1. **What is the total number of orders received on each day of September?**

SELECT O\_ORDERDATE, COUNT(\*) FROM ORDERS

WHERE MONTH(O\_ORDERDATE) = '9' AND YEAR(O\_ORDERDATE) = '1995'

GROUP BY O\_ORDERDATE;



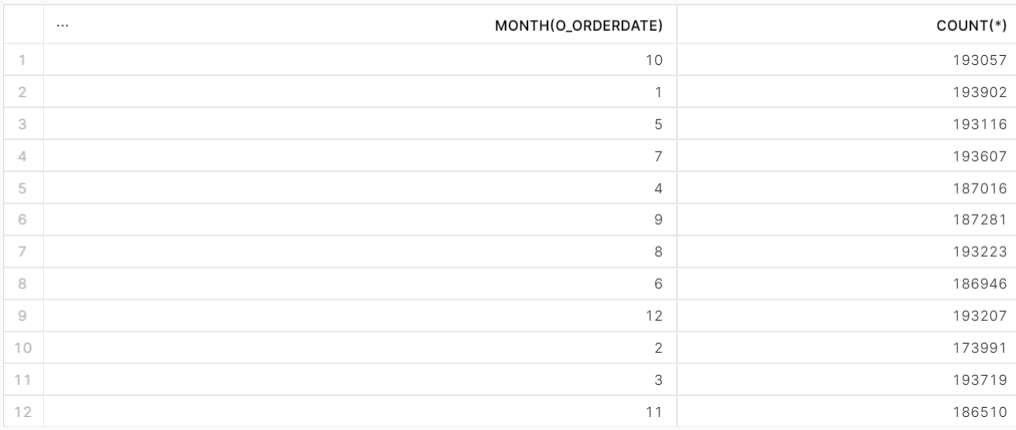


* 1. **What is the total number of orders received in each month of 1995?**

SELECT MONTH(O\_ORDERDATE), COUNT(\*) FROM ORDERS

WHERE YEAR(O\_ORDERDATE) = '1995'

GROUP BY MONTH(O\_ORDERDATE);



* 1. **What is the total number of orders received in all years?**

SELECT YEAR(O\_ORDERDATE), COUNT(\*) FROM ORDERS

GROUP BY YEAR(O\_ORDERDATE);



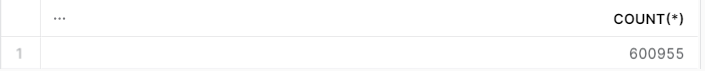
* 1. **What is the total number of orders received in all years?**

SELECT COUNT(\*) FROM ORDERS

JOIN CUSTOMER ON ORDERS.O\_CUSTKEY = CUSTOMER.C\_CUSTKEY

JOIN NATION ON CUSTOMER.C\_NATIONKEY = NATION.N\_NATIONKEY

WHERE NATION.N\_NAME = 'BRAZIL';



**CONCLUSION:**

In this lab, we made a Fact table of ORDERS and for Dimensions we used Time and Location.