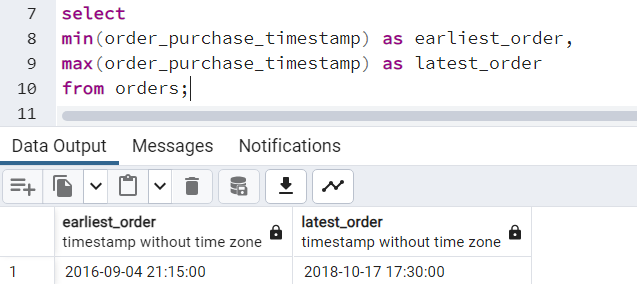
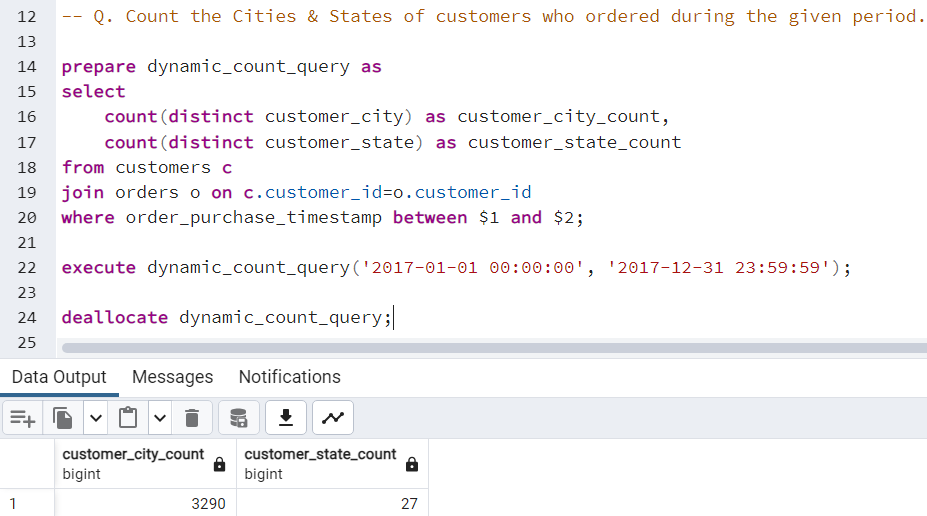
**Problem Statement:**

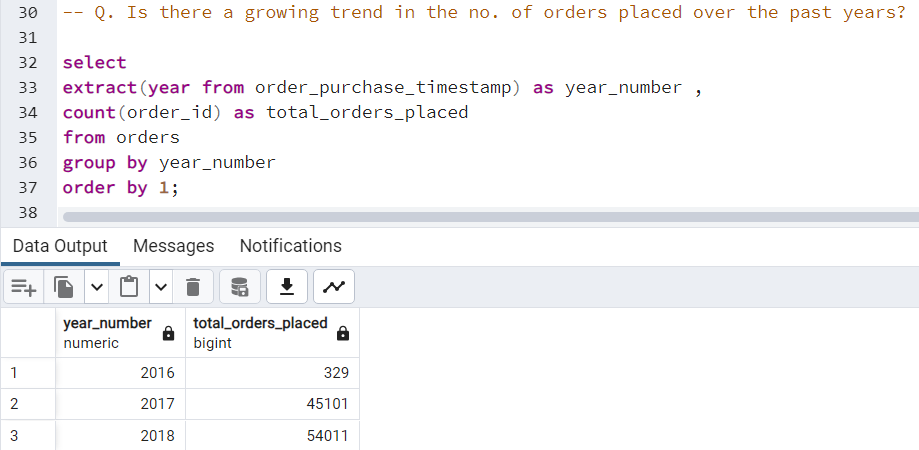
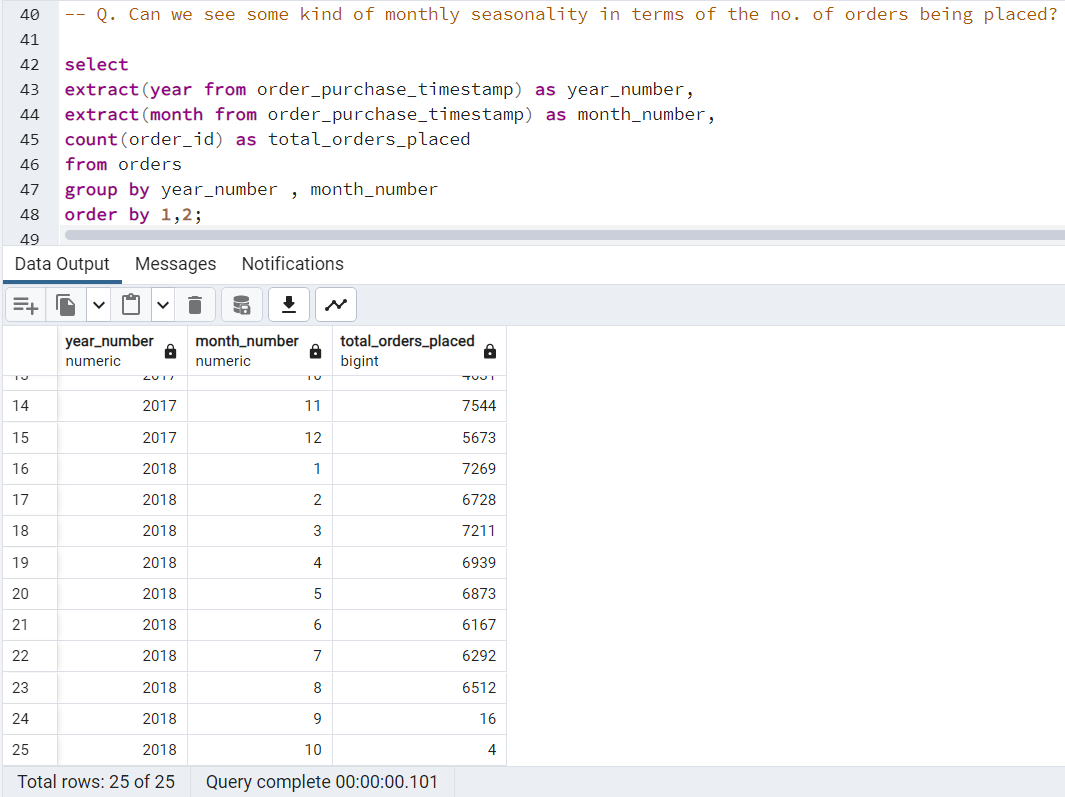
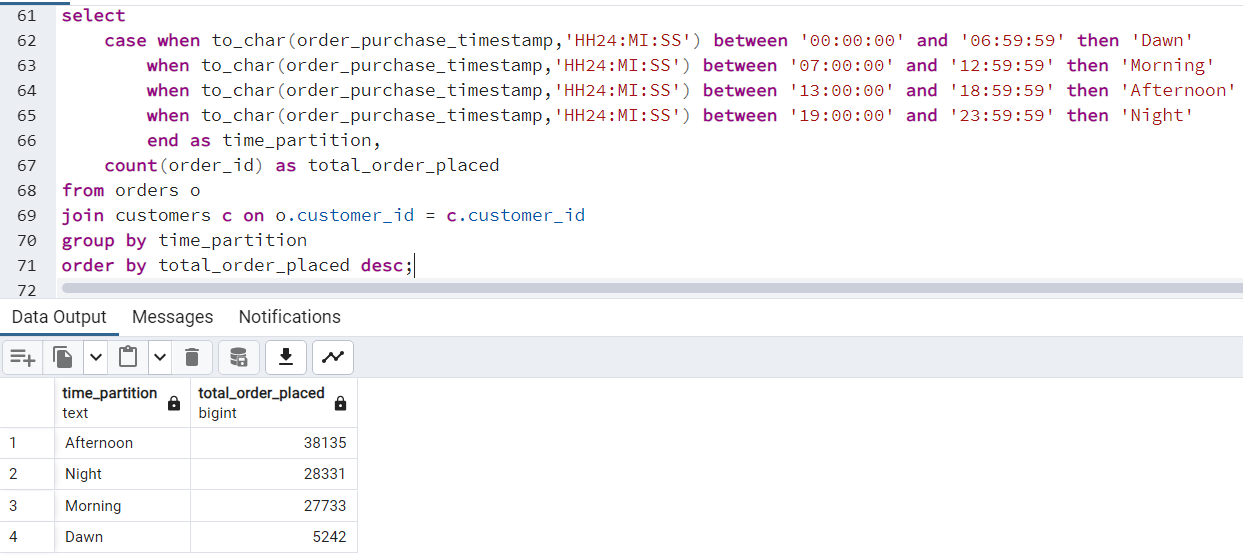
Assuming you are a data analyst/ scientist at Target, you have been assigned the task of analyzing the given dataset to extract valuable insights and provide actionable recommendations.

**What does 'good' look like?**

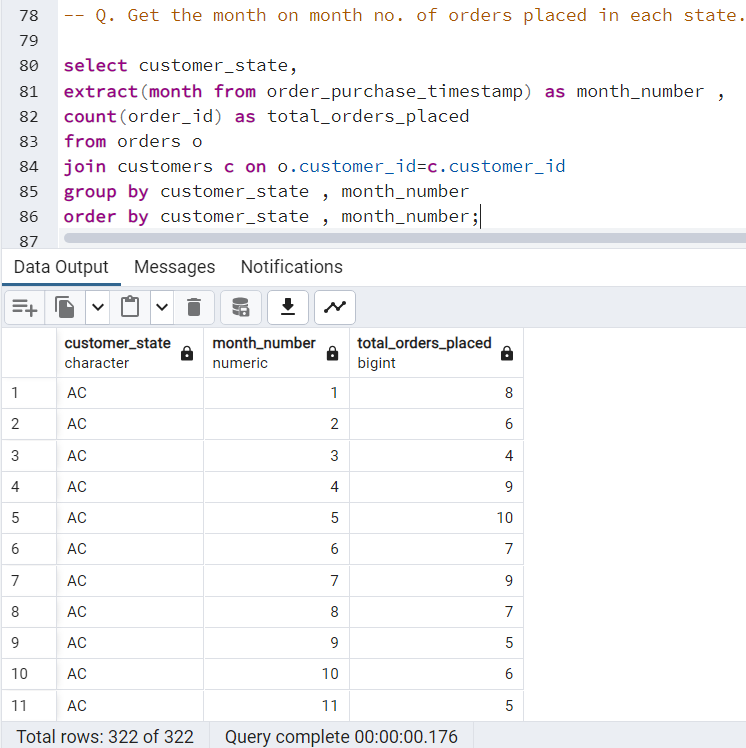
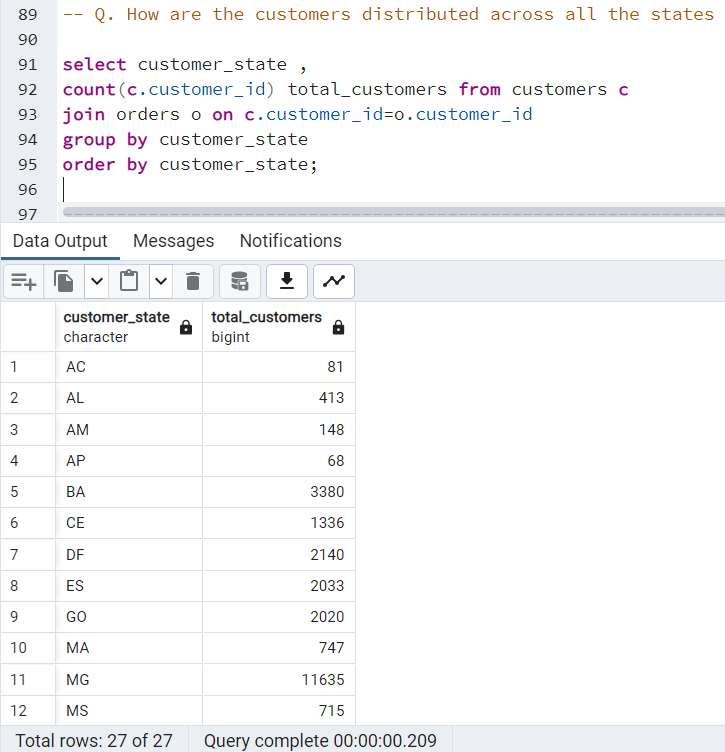
1. **Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:**
   1. **Data type of all columns in the "customers" table**.

* The datatype of customer\_id , customer\_unique\_id are choosen as “text” , “bigint” for customer\_zip\_code , “character(50)” for customer\_city and “character(2)” for customer\_state.
  1. **Get the time range between which the orders were placed.**
* 
  1. **Count the Cities & States of customers who ordered during the given period.**
* 

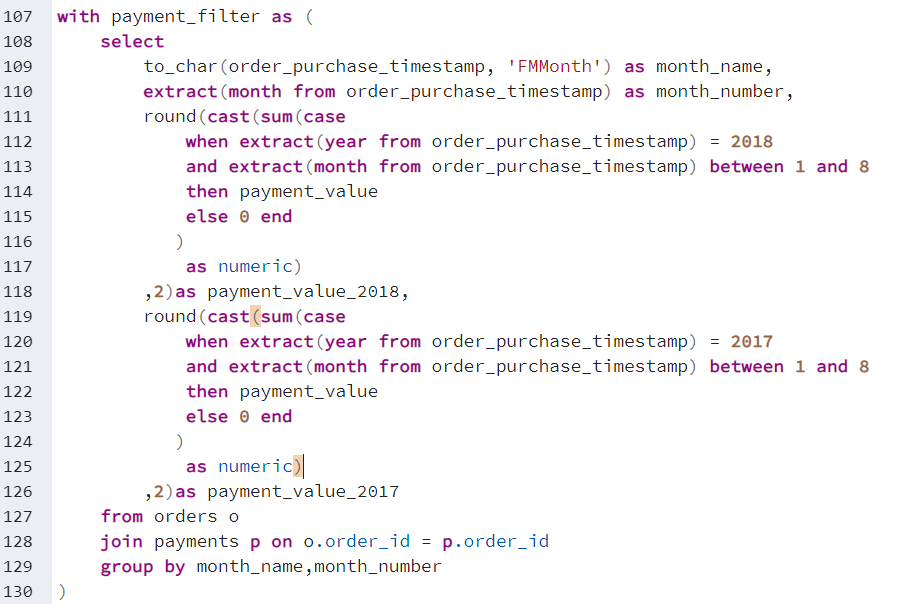
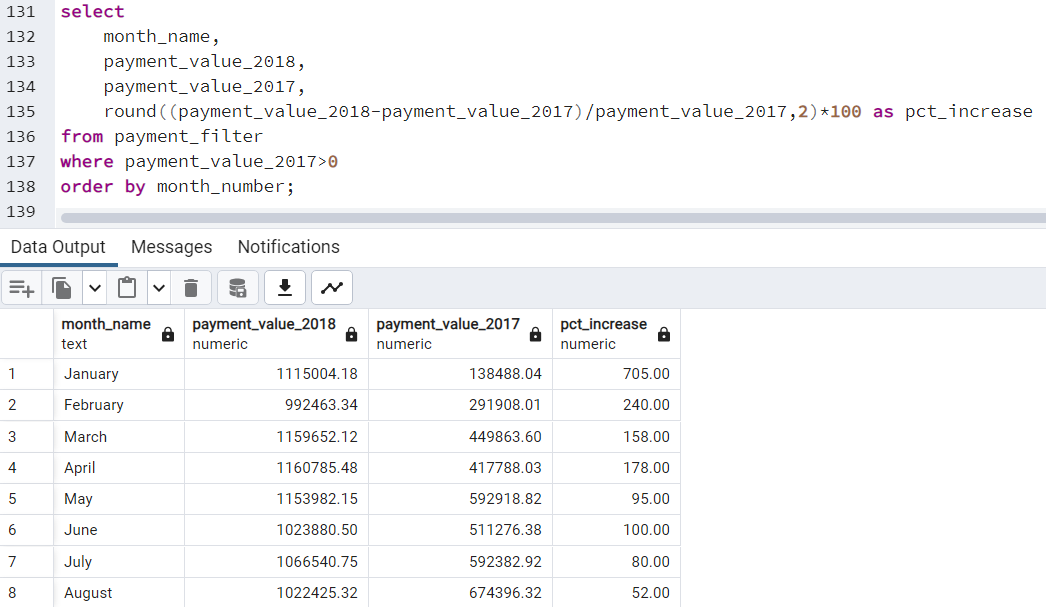
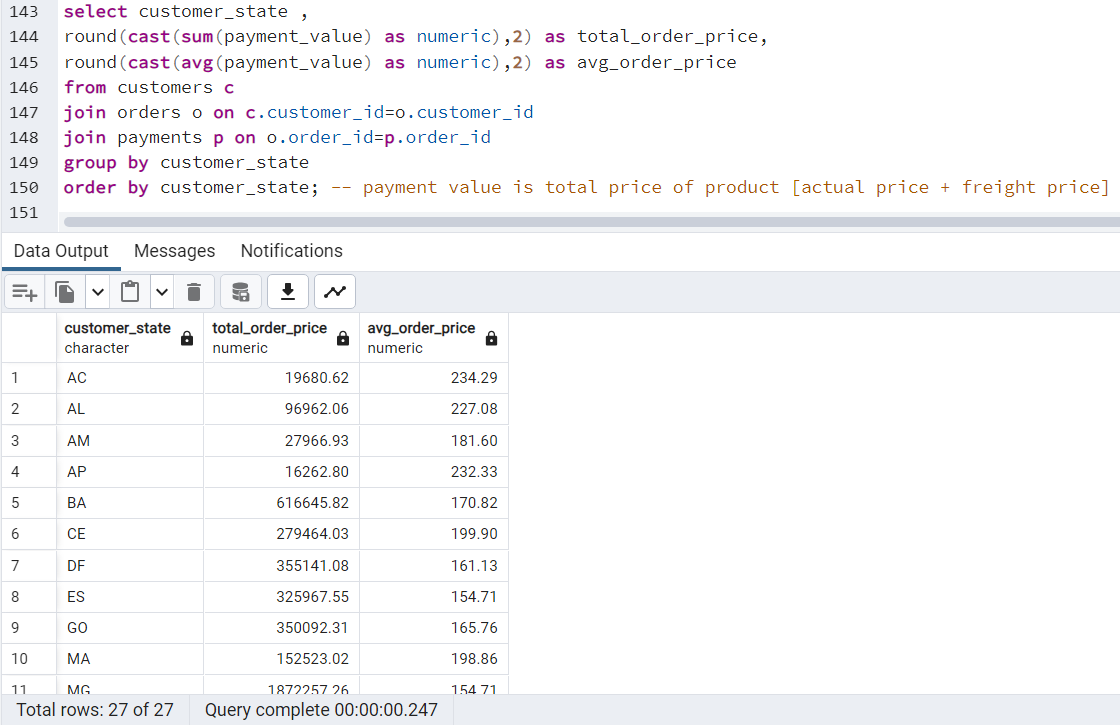
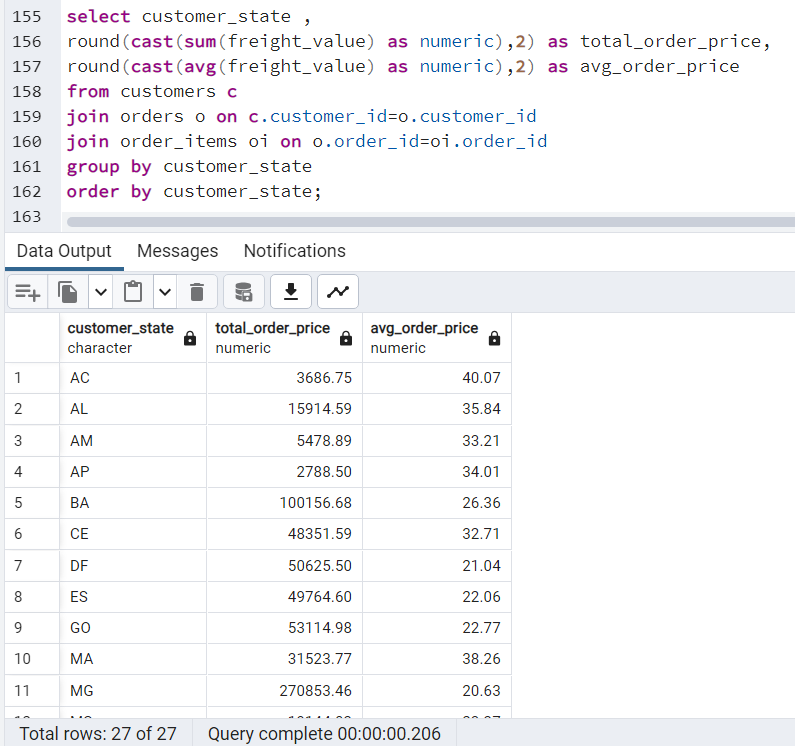
1. **In-depth Exploration:**  
   1. **Is there a growing trend in the no. of orders placed over the past years?**

* 
* Yes , we can conclude the growing trend in orders over past years.
  1. **Can we see some kind of monthly seasonality in terms of the no. of orders being placed?**
* 
* strong monthly seasonality in 2017, with specific peaks around the end of the year, but weakened seasonality in 2018 after mid-year.
  1. **During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)**
     + **0-6 hrs : Dawn**
     + **7-12 hrs : Mornings**
     + **13-18 hrs : Afternoon**
     + **19-23 hrs : Night**
* 
* Afternoon is the most preferred order time for Brazilian customers.

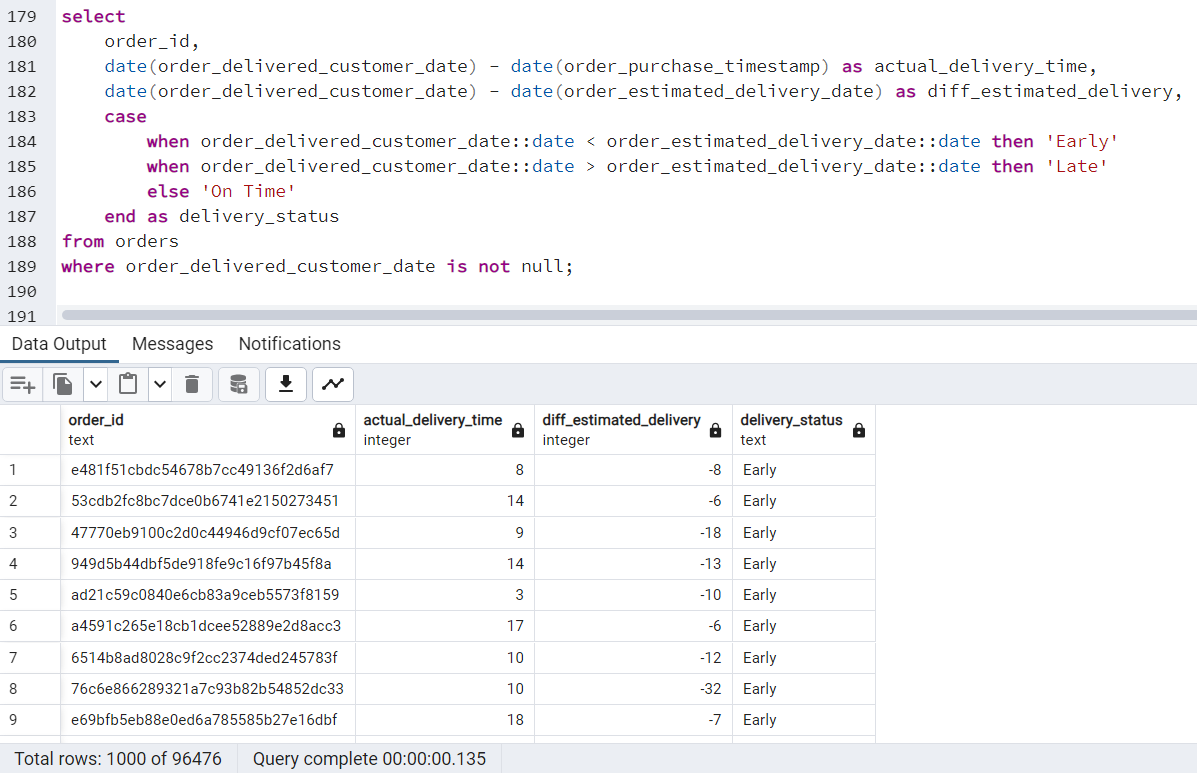
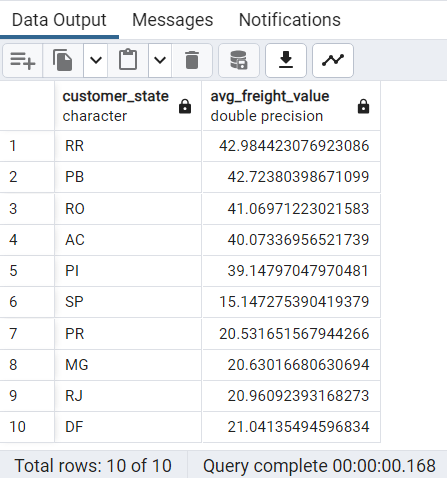
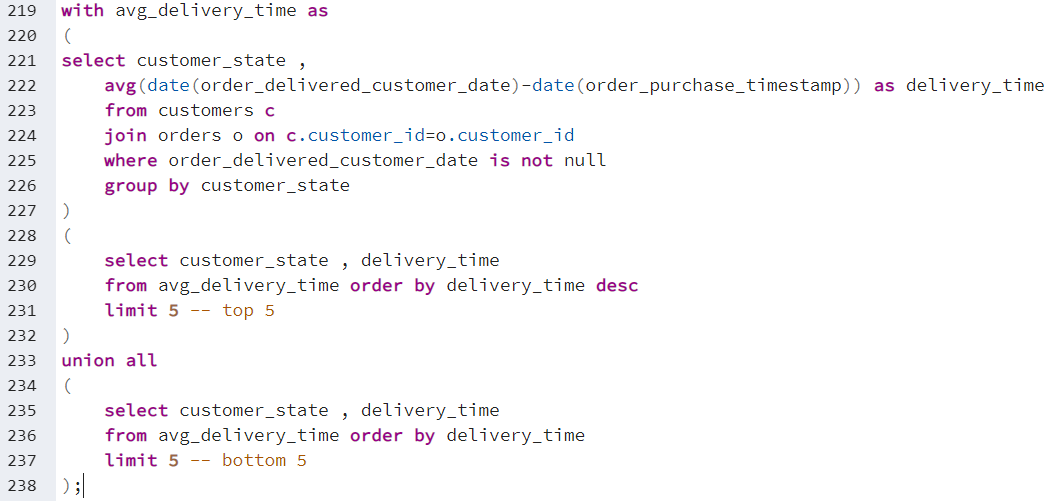
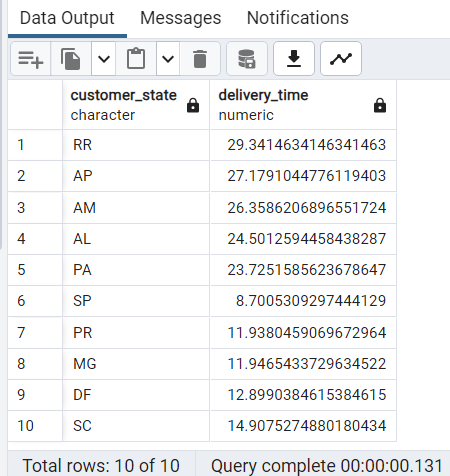
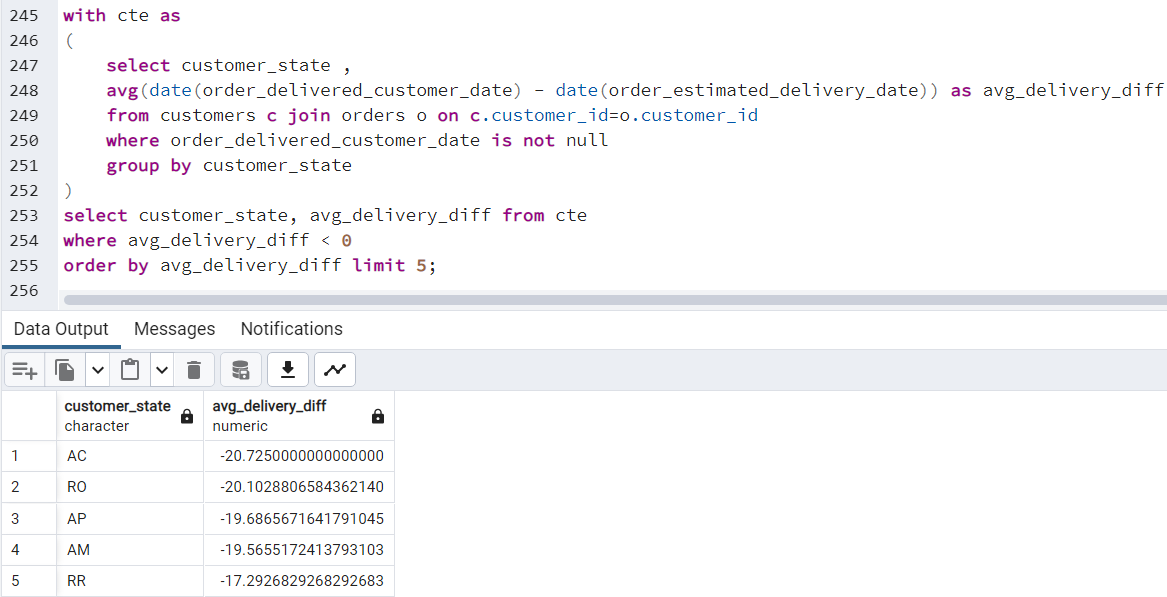
1. **Evolution of E-commerce orders in the Brazil region:**
   1. **Get the month on month no. of orders placed in each state.**

* 
  1. **How are the customers distributed across all the states?**
* 
* Here , we had find customer distribution across all states for only those customers that had ordered something.
* So by this , we can get idea about which state’s people are our frequent buyer or least buyer.

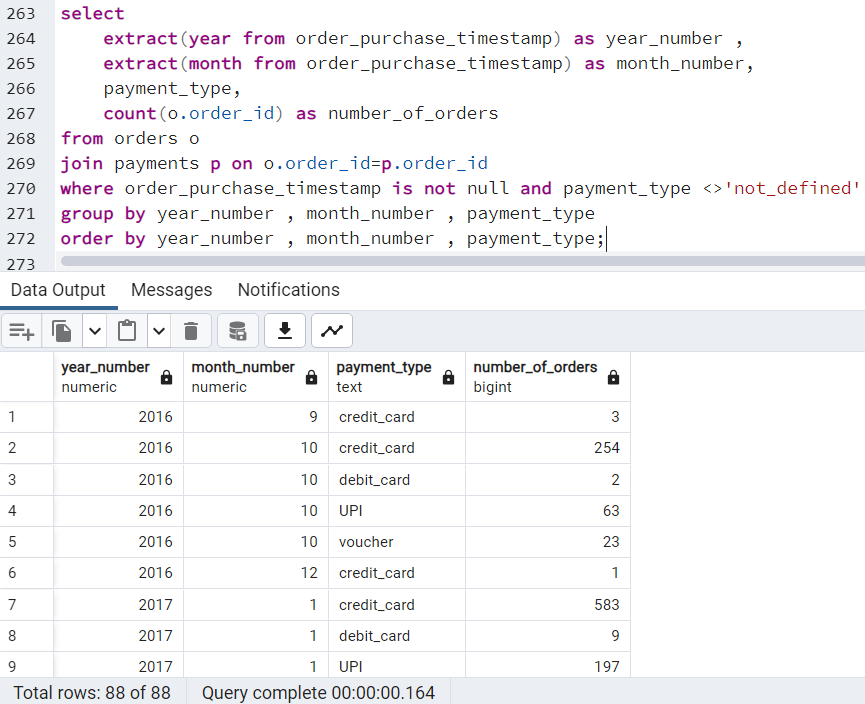
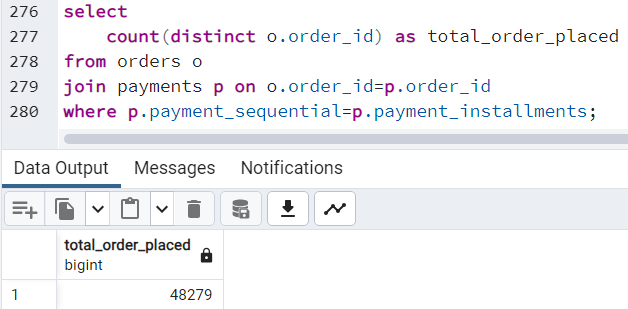
1. **Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others.**
   1. **Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).  
      You can use the "payment\_value" column in the payments table to get the cost of orders.**

* 
* 
* There is an excellent increase in the cost of product from last year (2017).
* Above we had quoted downfall in number of orders placed in 2018 compared to 2017 , but from this we can conclude that the company isn’t facing any loss
* Infact they earned way more than what they earned in 2017.
* People are purchasing costlier products which obviously reduces quantity.
  1. **Calculate the Total & Average value of order price for each state.**
* 
* By this , we can find state-wise contribution in the economy.
* i.e top n / least n states contributing in the economy.
  1. **Calculate the Total & Average value of order freight for each state.**
* 
* Freight is basically a price rate at which order is delivered from one place to another.
* We can work on the supply chain of low freight value state for more growth of the company.

1. **Analysis based on sales, freight and delivery time.**
   1. **Find the no. of days taken to deliver each order from the order’s purchase date as delivery time.  
      Also, calculate the difference (in days) between the estimated & actual delivery date of an order.  
      Do this in a single query.  
        
      You can calculate the delivery time and the difference between the estimated & actual delivery date using the given formula:**
      * **time\_to\_deliver = order\_delivered\_customer\_date - order\_purchase\_timestamp**
      * **diff\_estimated\_delivery = order\_delivered\_customer\_date - order\_estimated\_delivery\_date**

* 
* We could work on the supply chain where delivery status is late.
  1. **Find out the top 5 states with the highest & lowest average freight value.**
* 
* 
  1. **Find out the top 5 states with the highest & lowest average delivery time.**
* 
* 
* We may work on supply chain to minimize delivery time of the state which are among the top most states having highest average delivery time.
  1. **Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.  
     You can use the difference between the averages of actual & estimated delivery date to figure out how fast the delivery was for each state.**
* 
* Here -ve sign represents how many days earlier the product was delivered from expected delivery time.

1. **Analysis based on the payments:**
   1. **Find the month on month no. of orders placed using different payment types.**

* 
* We can get top top payment method preferred by customers.
* This can be useful on applying discount % on various payment method which attracts customer and thus increase company sale.
  1. **Find the no. of orders placed on the basis of the payment installments that have been paid.**
* 
* This data will help the company to crack the deal with those customers who aren’t capable to pay wholesome amount at a time.
* Showcasing this will not only motivate people and help them to pay in installments but will also result in the benefit of the company.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Evaluation Criteria (100 points):**

1. Initial exploration like checking the structure & characteristics of the data (15 points)
2. In-depth Exploration (15 points)
3. Evolution of E-commerce orders in the Brazil region (10 points)
4. Impact on Economy (20 points)
5. Analysis on sales, freight and delivery time (20 points)
6. Analysis based on the payments (10 points)
7. Actionable Insights & Recommendations (10 points)

Refer README.txt file

