|  |
| --- |
| ***NAME : P. UMESH CHANDRA***  ***EMAIL ID : [umeshpathipka27@gmail.com](mailto:umeshpathipka27@gmail.com)***  ***MOBILE NO : 7207469952*** |

**Business Case: Target SQL**

***Target is a globally renowned brand and a prominent retailer in the United States. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation and an exceptional guest experience that no other retailer can deliver.***

***This particular business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018.***

***The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.***

***By analyzing this extensive dataset, it becomes possible to gain valuable insights into Target's operations in Brazil. The information can shed light on various aspects of the business, such as order processing, pricing strategies, payment and shipping efficiency, customer demographics, product characteristics, and customer satisfaction levels.***

***FREIGHT AND SHIPPING EFFICEIENCY***

***FREIGHT Cost refers to Expenses Increase in Transporting Goods***

***Costs Related to - Shipping, Handling, Fuel Costs, Ware House Costs.***

***In Target Brazil Operations Freight cost is a key metric to analyze, - Profit Margin, Pricing strategies, customer Satisfaction.***

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

**Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset:**

**Data type of all columns in the "customers" table.**

**QUERY:**

SELECT

table\_name,

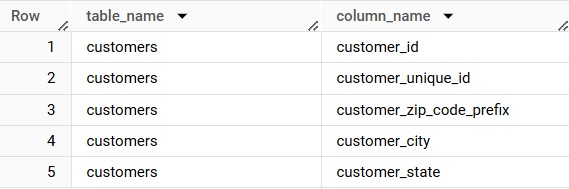
column\_name

FROM

`target-brazil-analysis.target\_brazil.INFORMATION\_SCHEMA.COLUMNS`

WHERE table\_name= 'customers'

**OUTPUT:**

****

**Get the time range between which the orders were placed.**

***Determine the Earliest and latest timestamps of when orders were placed in the dataset.***

***The earliest date (minimum timestamp)***

***The Latest date (Maximum Timestamp)***

**QUERY:**

SELECT

MIN(order\_purchase\_timestamp)

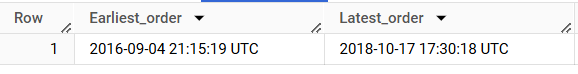
AS Earliest\_order,

MAX(order\_purchase\_timestamp)

AS Latest\_order

FROM`target-brazil-analysis.target\_brazil.orders`

**OUTPUT:**

****

* ***Earliest order date: September 4, 2016***
* ***Latest Order date: October 17, 2018***
* ***Time range is 2 Years, 1 Month, and 13 days***

**Count the Cities & States of customers who ordered during the given period.**

**QUERY:**

SELECT

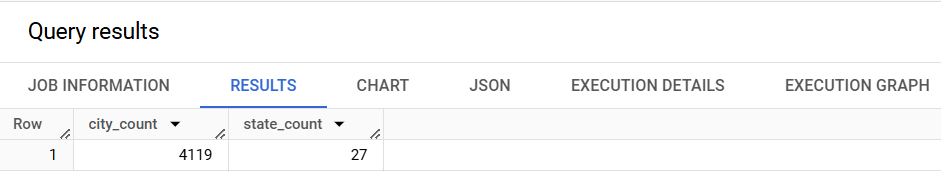
COUNT(DISTINCT customer\_city)

AS city\_count,

COUNT(DISTINCT customer\_state)

AS state\_count

FROM`target-brazil-analysis.target\_brazil.customers`



*OBSERVATION:*

* *Unique cities are 4,119 and Unique states are 27*

**IN-DEPTH EXPLORATION**

**Is there a growing trend in the no. of orders placed over the past years?**

**QUERY:**

SELECT

EXTRACT(YEAR FROM order\_purchase\_timestamp) AS order\_years,

COUNT(order\_id) AS total\_orders

FROM`target-brazil-analysis.target\_brazil.orders`

GROUP BY order\_years

ORDER BY order\_years

**OUTPUT:**

****

**OBSERVATION:**

* **Number of orders increases every year, It indicates a positive growth trend in e-commerce orders.**
* **The Number of orders increased from 329 to 45,101 in one year, suggests a significant expansion in customer reach, marketing efforts or platform popularity.**
* **Orders increased further from 45,101 to 54,011(~20% Growth).**
* **Massive growth from 2016 to 2017 and sustained Growth from 2017 to 2018.**

**Can we see some kind of monthly seasonality in terms of the no. of orders being placed?**

**QUERY:**

SELECT

EXTRACT(MONTH FROM order\_purchase\_timestamp)

AS Monthly\_Orders,

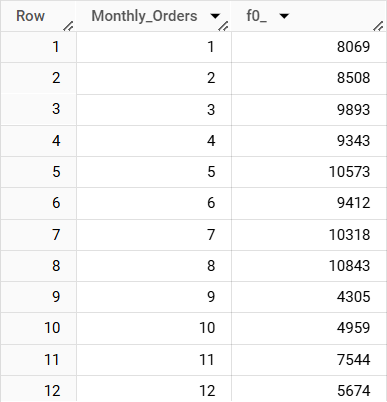
COUNT(order\_id)

FROM`target-brazil-analysis.target\_brazil.orders`

GROUP BY Monthly\_Orders

ORDER BY Monthly\_Orders;

**OUTPUT:**

****

**OBSERVATIONS:**

* **Peak months: May, July, August**
* **Orders peak in May(10,573), July(10,318) and August(10,843) indicates that seasonal shopping trends, promotional events or holidays.**
* **Drop in orders: September, October, December**
* **December(5674) is also lower, which is surprising given holiday shopping.**
* **Steady Growth from January to August**
* **Orders gradually increase from January to August, Indicating strong Demand in the first 8 months.**

**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**

**QUERY:**

SELECT

CASE

WHEN EXTRACT(HOUR FROM order\_purchase\_timestamp)

BETWEEN 0 AND 6 THEN 'Dawn'

WHEN EXTRACT(HOUR FROM order\_purchase\_timestamp)

BETWEEN 7 AND 12 THEN 'Morning'

WHEN EXTRACT(HOUR FROM order\_purchase\_timestamp)

BETWEEN 13 AND 18 THEN 'Afternoon'

ELSE 'Night'

END AS time\_of\_day,

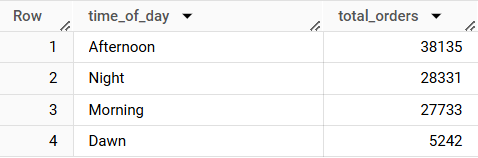
COUNT(order\_id) AS total\_orders

FROM`target-brazil-analysis.target\_brazil.orders`

GROUP BY time\_of\_day

ORDER BY total\_orders DESC

**OUTPUT:**



**OBSERVATION:**

* **Peak order time:Afternoon, Most orders are placed between 1PM to 6PM, indicating that customers prefer shopping during launch breaks**
* **Night and Morning orders are similar(28K)**
* **Very few orders come in between midnight and 6AM**

**Evolution of E-commerce orders in the Brazil region**

**Analyzing the Evolution of E-Commerce orders in Brazil**

1. **Monthly Orders by state**
2. **Customer Distribution across states**

**Get the month on month no. of orders placed in each state.**

**QUERY:**

SELECT

c.customer\_state,

EXTRACT(MONTH FROM o.order\_purchase\_timestamp)

AS order\_month,

COUNT(o.order\_id) AS total\_orders

FROM`target-brazil-analysis.target\_brazil.orders` AS o

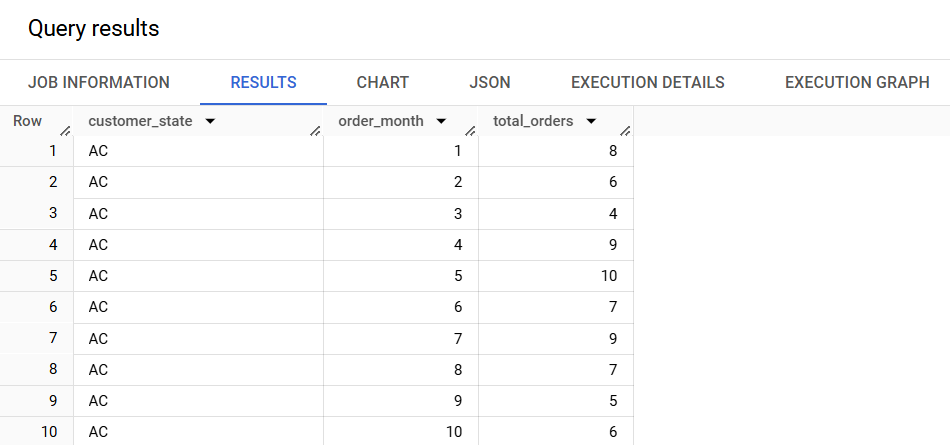
JOIN`target-brazil-analysis.target\_brazil.customers` AS c

ON o.customer\_id= c.customer\_id

WHERE o.order\_status NOT IN('canceled','unavailable')

GROUP BY c.customer\_state, order\_month

ORDER BY c.customer\_state, order\_month

****

**How are the customers distributed across all the states?**

**QUERY:**

SELECT

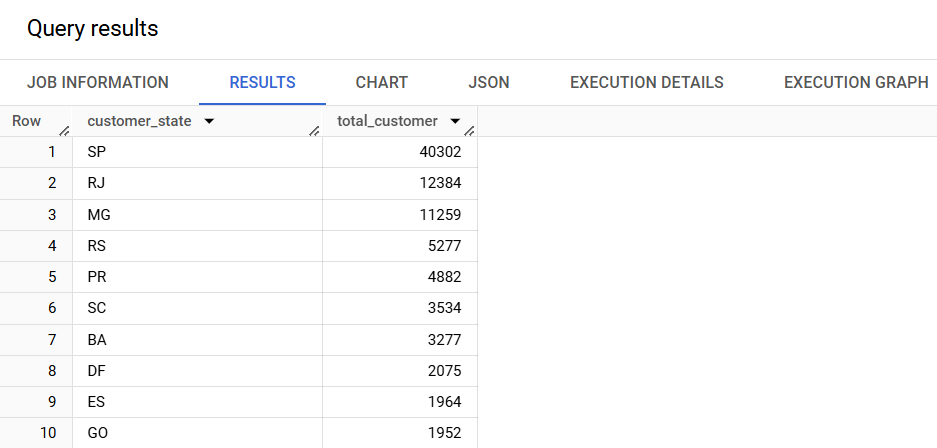
customer\_state,

COUNT(DISTINCT customer\_unique\_id)AS total\_customer

FROM`target-brazil-analysis.target\_brazil.customers`

GROUP BY customer\_state

ORDER BY total\_customer DESC;



**OBSERVATION**

* **Top contributing states:-**
* **SP has the highest number of customers, with 40,302. Followed by RJ with 12,384, and MG with 11,259 indicating strong customers presence**
* **Moderate Contributing States:-**
* **States like RS, PR, SC shows a moderate number of customers**

**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.**

**QUERY:**

WITH order\_details AS(

SELECT

order\_id,

EXTRACT(YEAR FROM order\_purchase\_timestamp)

AS order\_year,

EXTRACT(MONTH FROM order\_purchase\_timestamp)

AS order\_month

FROM`target-brazil-analysis.target\_brazil.orders`),

payment\_2017 AS (

SELECT SUM(payment\_value) AS total\_payment\_2017,

FROM`target-brazil-analysis.target\_brazil.payments`

JOIN order\_details USING(order\_id)

WHERE order\_year= 2017 AND order\_month<=8),

payment\_2018AS(

SELECT SUM(payment\_value)

AS total\_payment\_2018,

FROM`target-brazil-analysis.target\_brazil.payments`

JOIN order\_details USING(order\_id)

WHERE order\_year= 2018 AND order\_month<=8)

SELECT ROUND(

(total\_payment\_2018-total\_payment\_2017)/total\_payment\_2017\*100,2) AS Percentage\_Increase

FROM payment\_2017 CROSSJOIN payment\_2018

**OUTPUT:**

****

**EXPLAINATION:-**

* **To calculate the percentage increase in order cost from 2017 to 2018(Including January-August)**
* **The formula for Percentage Increase is**

**(TOTAL COST IN 2018 - TOTAL COST IN 2017) \* 100 / (TOTAL COST IN 2017)**

**OBSERVATION:-**

* The total cost of orders increased by 136.98% from 2017-2018(Jan-Aug)
* The revenue more than doubled in just one year

**2.Calculate the Total & Average value of order price for each state.**

**QUERY:**

SELECT

customer\_state,

ROUND(SUM(price),2)

AS Total\_order\_price,

ROUND(AVG(price),2)

AS Average\_order\_price

FROM `target-brazil-analysis.target\_brazil.order\_items`

JOIN `target-brazil-analysis.target\_brazil.orders`

USING(order\_id)

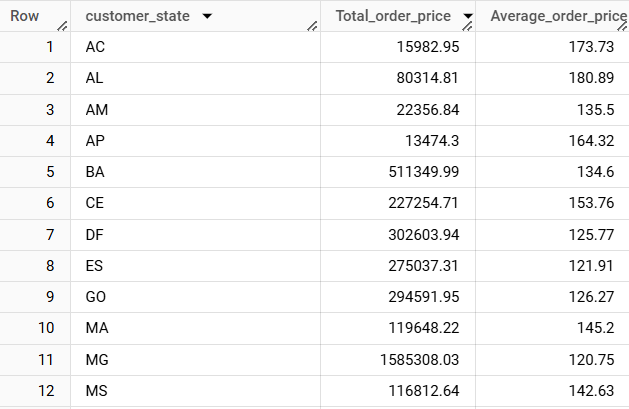
JOIN`target-brazil-analysis.target\_brazil.customers`

USING(customer\_id)

GROUP BY customer\_state

ORDER BY customer\_state

**OUTPUT:**



**Calculate the Total & Average value of order freight for each state.**

**QUERY:**

SELECT

customer\_state,

ROUND(SUM(freight\_value),2) AS Total\_freight\_value,

ROUND(AVG(freight\_value),2) AS Average\_freight\_value

FROM `target-brazil-analysis.target\_brazil.order\_items`

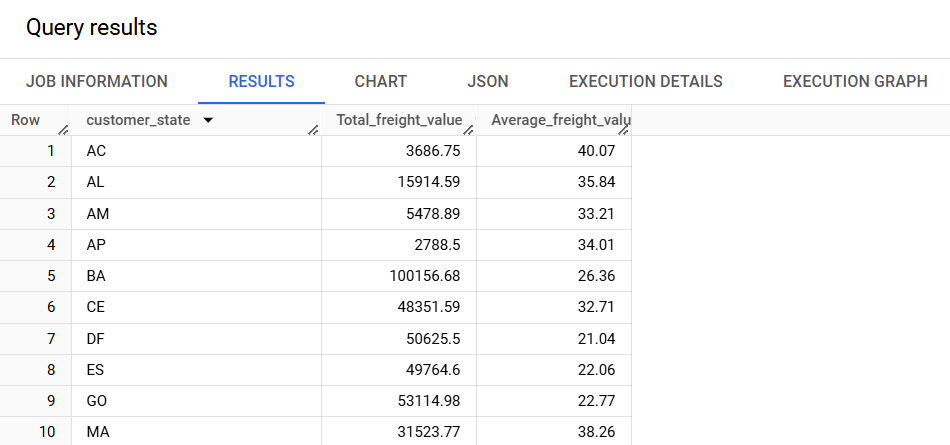
JOIN `target-brazil-analysis.target\_brazil.orders`

USING(order\_id) JOIN`target-brazil-analysis.target\_brazil.customers`

USING(customer\_id)

GROUP BY customer\_state

ORDER BY customer\_state

****

**Analysis based on sales, freight and delivery time.**

**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**

**QUERY:**

SELECT

order\_id,

DATE\_DIFF(order\_delivered\_customer\_date,

order\_purchase\_timestamp,day)

AS time\_to\_deliver,

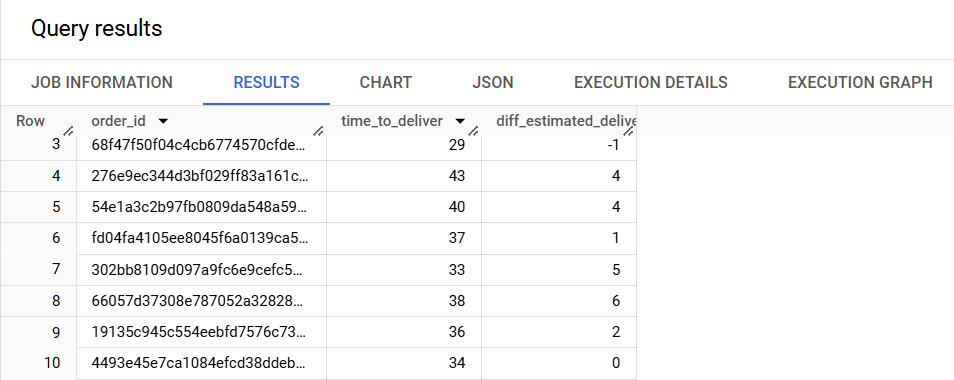
DATE\_DIFF(order\_delivered\_customer\_date,

order\_estimated\_delivery\_date,day)

AS diff\_estimated\_delivery

FROM`target-brazil-analysis.target\_brazil.orders`

WHERE order\_status= 'delivered'



**Find out the top 5 states with the highest & lowest average freight value.**

**QUERY:**

SELECT

customer\_state,

average\_freight\_value,

rank\_order

FROM(

SELECT c.customer\_state,

ROUND(AVG(oi.freight\_value),2) AS average\_freight\_value,

DENSE\_RANK()

OVER(ORDER BY ROUND(AVG(oi.freight\_value),2)DESC)

AS rank\_order

FROM`target-brazil-analysis.target\_brazil.customers`

AS c JOIN`target-brazil-analysis.target\_brazil.orders`AS o

ON c.customer\_id= o.customer\_id

JOIN`target-brazil-analysis.target\_brazil.order\_items`AS oi

ON o.order\_id= oi.order\_id

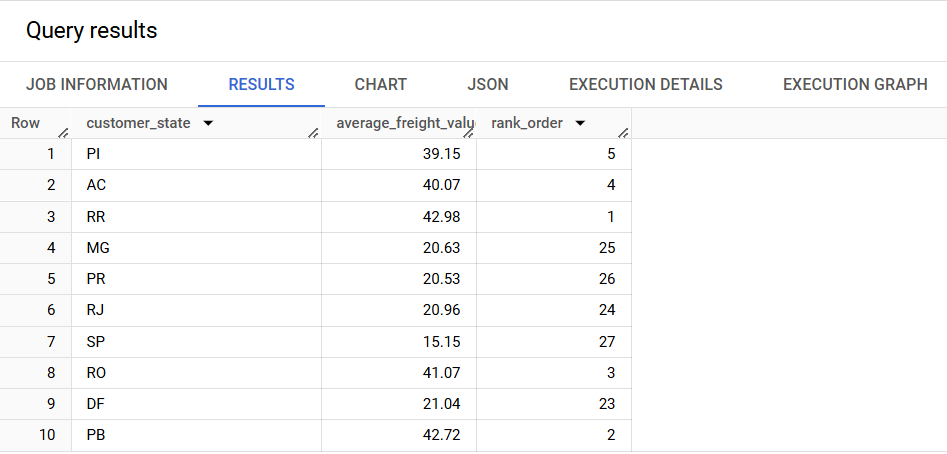
GROUP BY c.customer\_state

)AS ranked\_states

WHERE rank\_order<=5 OR rank\_order>=(

SELECT COUNT(DISTINCT customer\_state)-4

FROM`target-brazil-analysis.target\_brazil.customers`);

****

**Find out the top 5 states with the highest & lowest average delivery time.**

**QUERY:**

WITH DeliveryTime AS(

SELECT

o.order\_id,

c.customer\_state,

DATE\_DIFF(o.order\_delivered\_customer\_date,

o.order\_purchase\_timestamp,day)

AS time\_to\_deliver\_in\_days,

FROM`target-brazil-analysis.target\_brazil.orders`

AS o JOIN`target-brazil-analysis.target\_brazil.customers`

AS c

ON o.customer\_id= c.customer\_id

WHERE o.order\_status= 'delivered'),

RankedStates AS(

SELECT customer\_state,

ROUND(AVG(time\_to\_deliver\_in\_days),2)

AS avg\_delivery\_time,

DENSE\_RANK()

OVER(ORDER BY ROUND(AVG(time\_to\_deliver\_in\_days),2)DESC)

AS rank\_highest

FROM DeliveryTime

GROUP BY customer\_state)

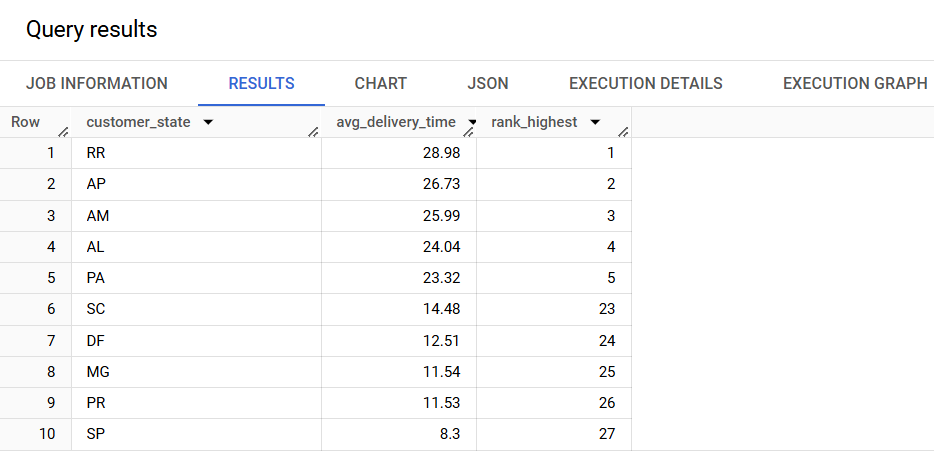
SELECT customer\_state,

avg\_delivery\_time, rank\_highest

FROM RankedStates

WHERE rank\_highest<=5 OR rank\_highest>=23

ORDER BY avg\_delivery\_time DESC;

****

**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.**

**QUERY:**

WITH DeliveryPerformance AS(

SELECT o.order\_id,c.customer\_state,

DATE\_DIFF(o.order\_delivered\_customer\_date,

o.order\_estimated\_delivery\_date,day)

AS early\_delivery\_days,

FROM`target-brazil-analysis.target\_brazil.orders`

AS o JOIN`target-brazil-analysis.target\_brazil.customers`

AS c ON o.customer\_id= c.customer\_id

WHERE o.order\_status= 'delivered'

AND o.order\_delivered\_customer\_date IS NOT NULL),

RankedStates AS (

SELECT customer\_state,

ROUND(AVG(early\_delivery\_days),2)AS avg\_early\_delivery\_days,

DENSE\_RANK()

OVER(ORDERBYROUND(AVG(early\_delivery\_days),2)ASC)

AS early\_delivery\_rank

FROM DeliveryPerformance

GROUP BY customer\_state)

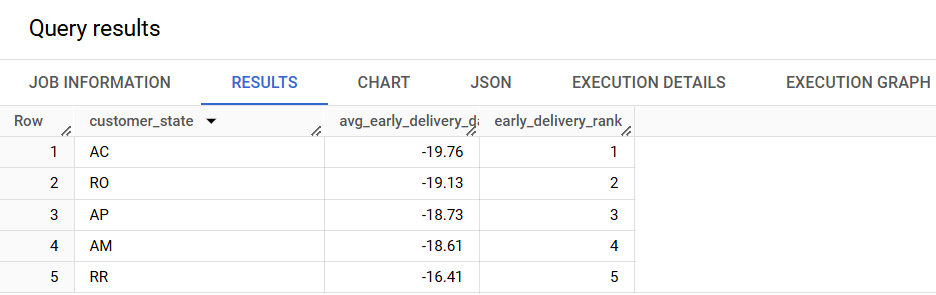
SELECT customer\_state,

avg\_early\_delivery\_days, early\_delivery\_rank

FROM RankedStates

WHERE early\_delivery\_rank<=5

ORDER BY avg\_early\_delivery\_days ASC;

****

**Find out the top 5 states where the order delivery is really Slowest Deliveries 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.**

**QUERY:**

WITH DeliveryDifferenceAS(

SELECT o.order\_id,c.customer\_state,

DATE\_DIFF(o.order\_delivered\_customer\_date,

o.order\_estimated\_delivery\_date,day) AS delay\_days,

FROM`target-brazil-analysis.target\_brazil.orders`

AS o JOIN`target-brazil-analysis.target\_brazil.customers`

AS c

ON o.customer\_id= c.customer\_id

WHERE o.order\_status= 'delivered'

AND o.order\_delivered\_customer\_date IS NOT NULL),

RankedStates AS(

SELECT customer\_state,

ROUND(AVG(delay\_days),2) AS avg\_delivery\_delay\_days,

DENSE\_RANK()

OVER(ORDER BY ROUND(AVG(delay\_days),2)DESC)

AS slowest\_delivery\_rank

FROM DeliveryDifference

GROUP BY customer\_state)

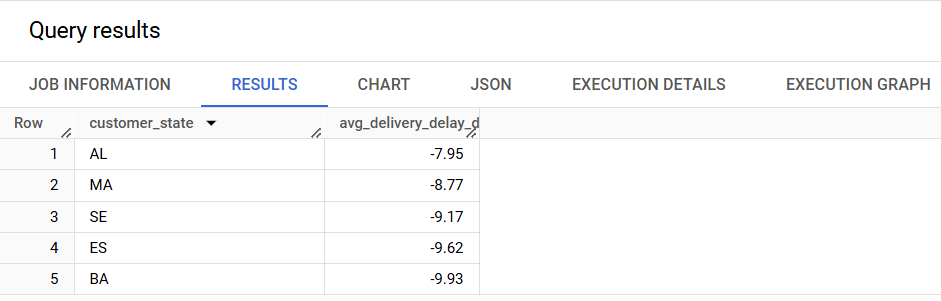
SELECT customer\_state,

avg\_delivery\_delay\_days

FROM RankedStates

WHERE slowest\_delivery\_rank <=5

ORDER BY avg\_delivery\_delay\_days DESC;



**Analysis based on the payments:**

**Find the month on month no. of orders placed using different payment types.**

**Find the no. of orders placed on the basis of the payment installments that have been paid.**

**QUERY:**

SELECT

EXTRACT(MONTH FROM order\_purchase\_timestamp)

AS order\_month,

payment\_type,

COUNT(order\_id) AS Total\_orders

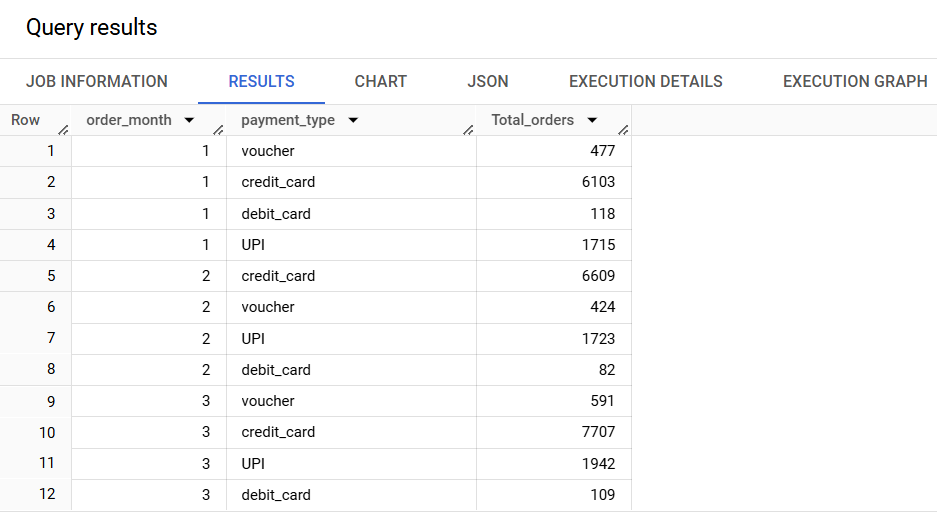
FROM`target-brazil-analysis.target\_brazil.payments`

JOIN`target-brazil-analysis.target\_brazil.orders`

USING(order\_id)

GROUP BY payment\_type, order\_month

ORDER BY order\_month ASC;

****

**OBSERVATION:-**

* **Across all months, CREDIT CARD is the most frequently used payment method.**
* **The highest number of credit card orders was in May(8,350) Peak in Transactions**
* **UPI shows a steady rise reaching its highest in August(2,077 orders)**
* **Debit card usage is Minimal, Voucher Usage is consistent but Low.**

**Find the no. of orders placed on the basis of the payment installments that have been paid.**

**QUERY:**

SELECT

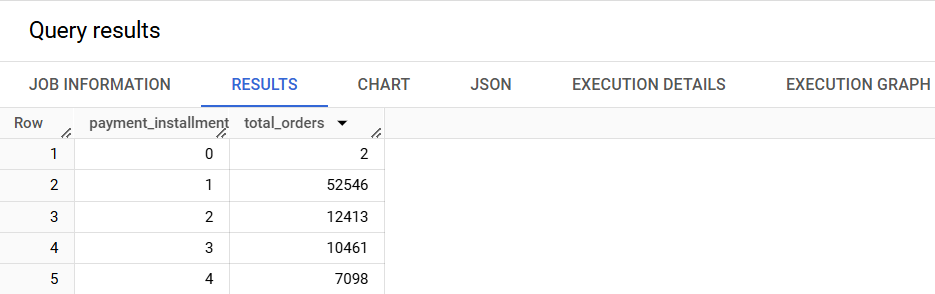
payment\_installments,

COUNT(order\_id) AS total\_orders

FROM`target-brazil-analysis.target\_brazil.payments`

GROUP BY payment\_installments

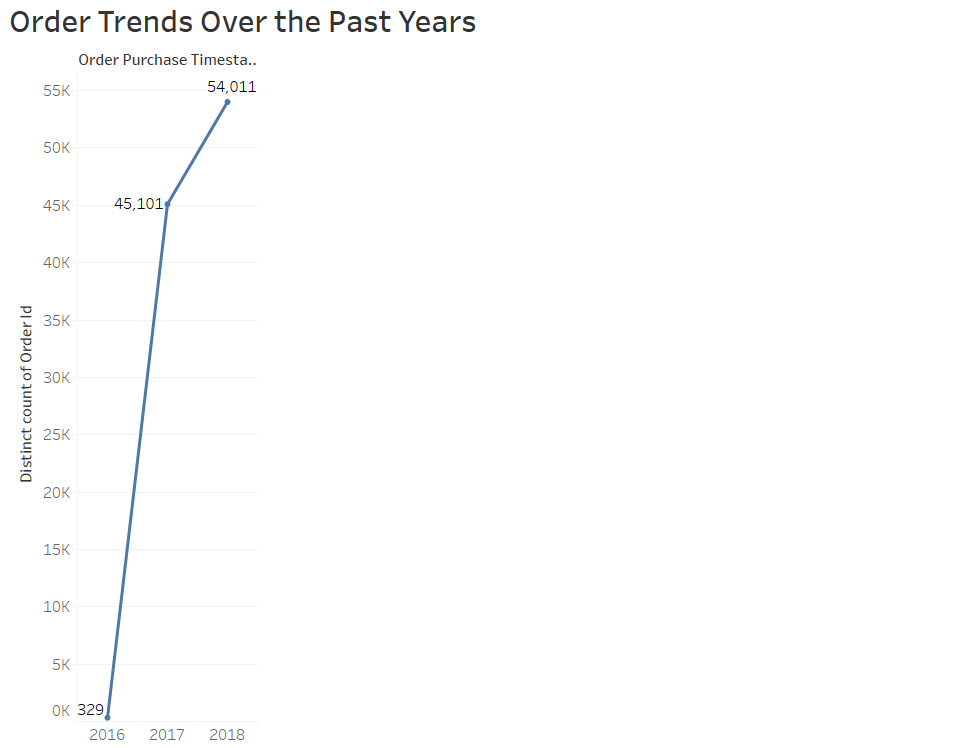
ORDER BY payment\_installments



**OBSERVATION:-**

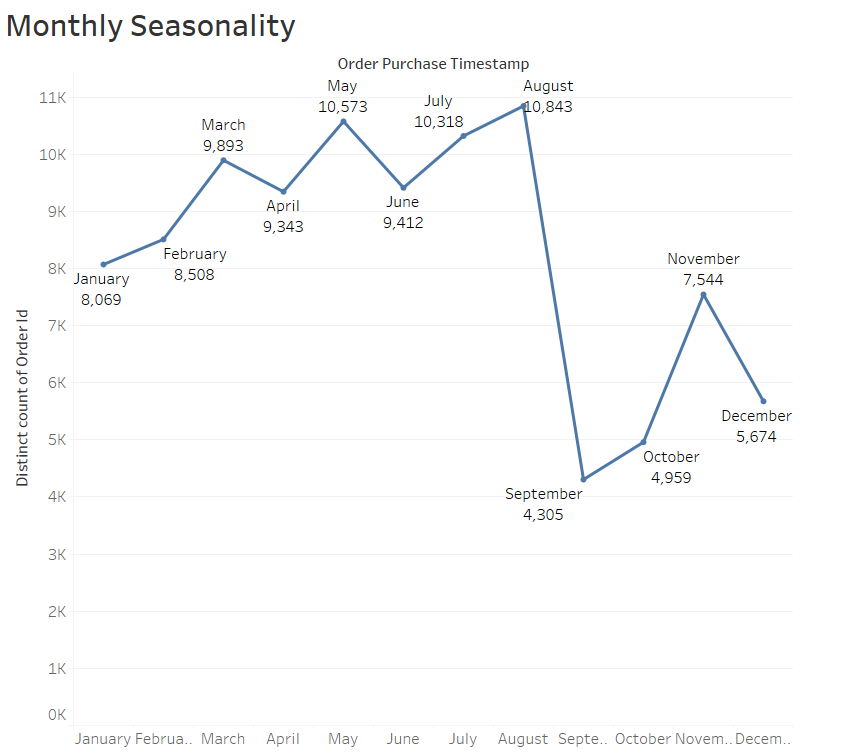
* **The highest number of orders(52,546) were placed using one installment(Full Payment)**
* **12,413 orders were placed using two installments**
* **As the number of installments increases, the number of orders declines**

**VISUALIZATION USING TABLEAU**

****

**OBSERVATION:**

* **Rapid Growth : Orders increased sharply from 329 in 2016 to 45,101 in 2017**
* **Continued increase order grew to 54,011 in 2018, but slower rate than before**

****

**OBSERVATION:**

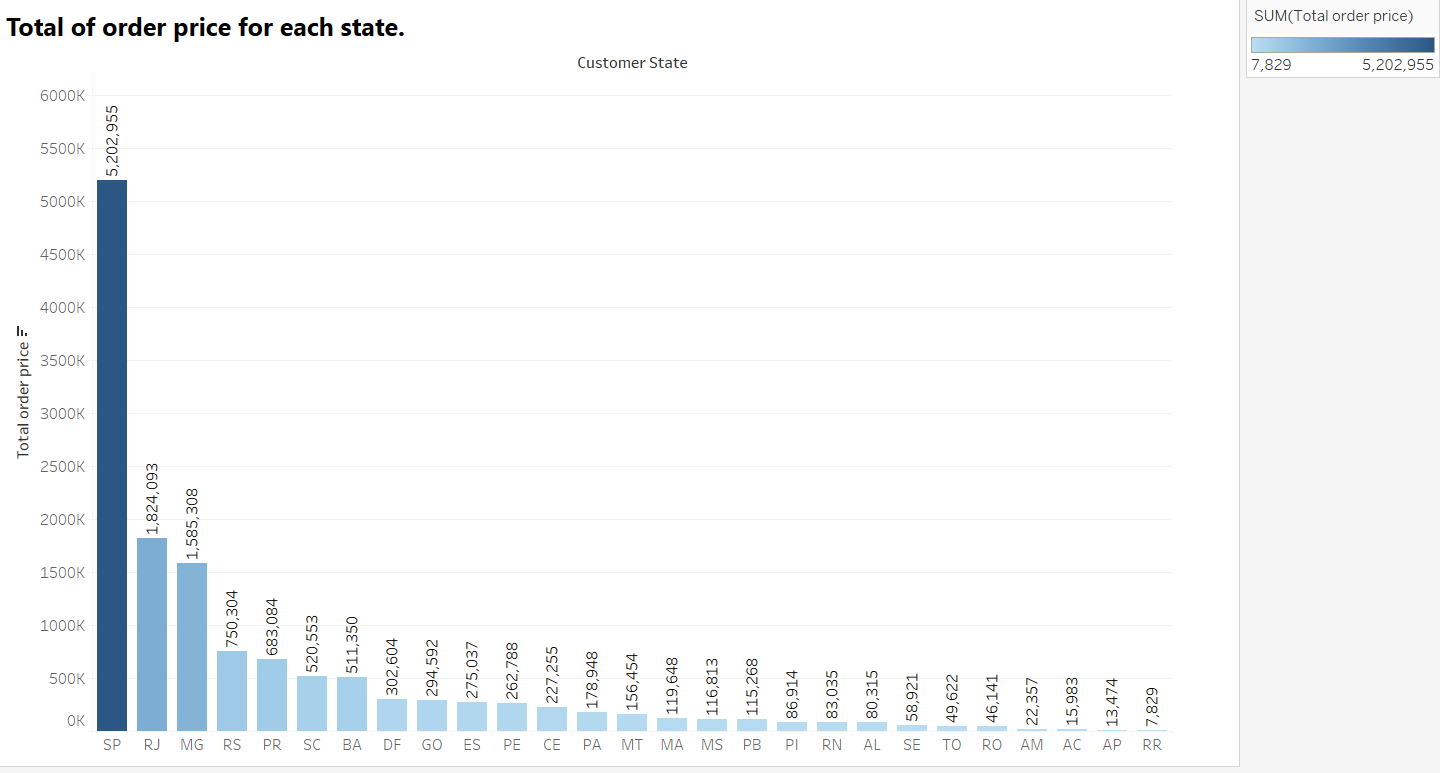
* **Orders are highest in May, July, and August indicating strong demand in these months**
* **September and October have the least orders showing a seasonal dip**
* **After a rise in November, orders decline in December**



**OBSERVATIONS:-**

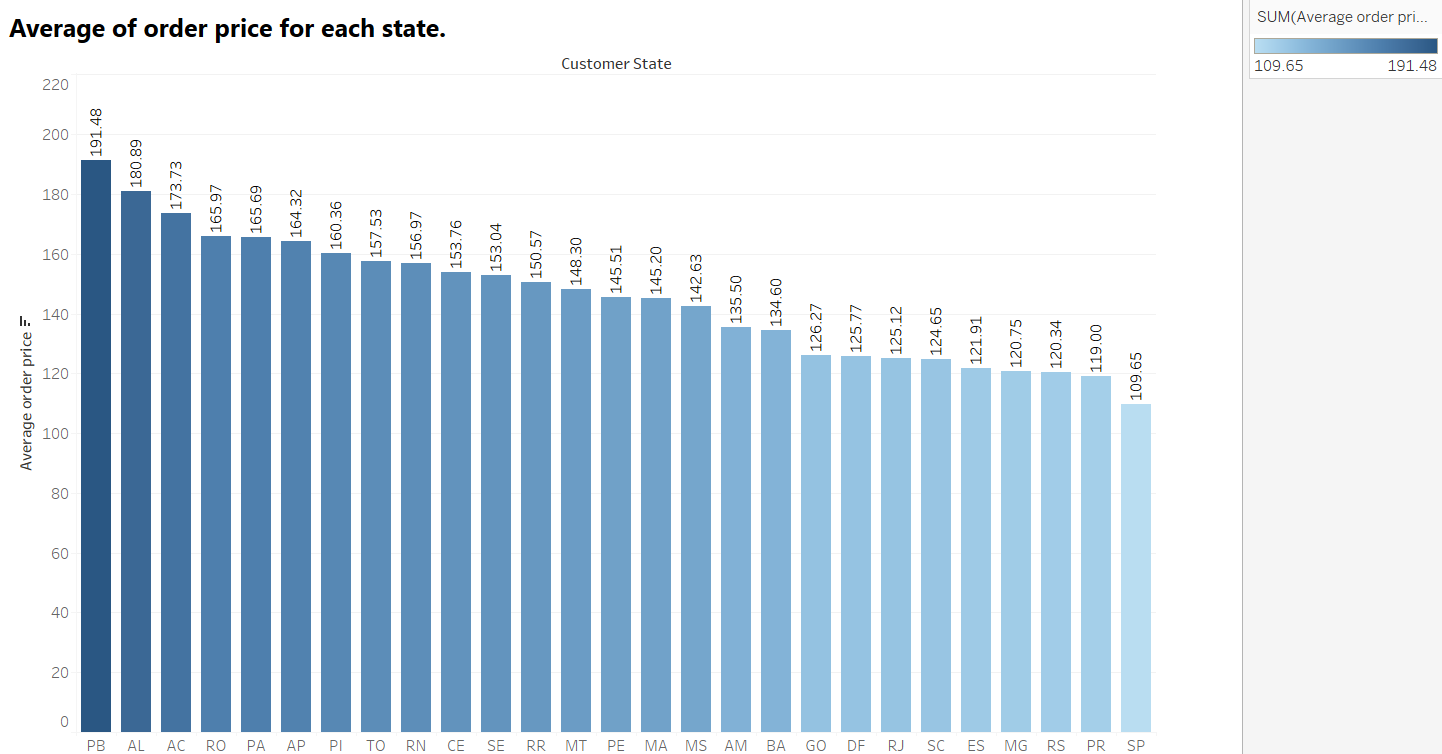
* **Most Orders placed in the afternoon, Making it the busiest time.**
* **Night and morning also see high order volumes**
* **Least orders at Dawn, Indicating minimal shopping**

**Impact on Economy**

****

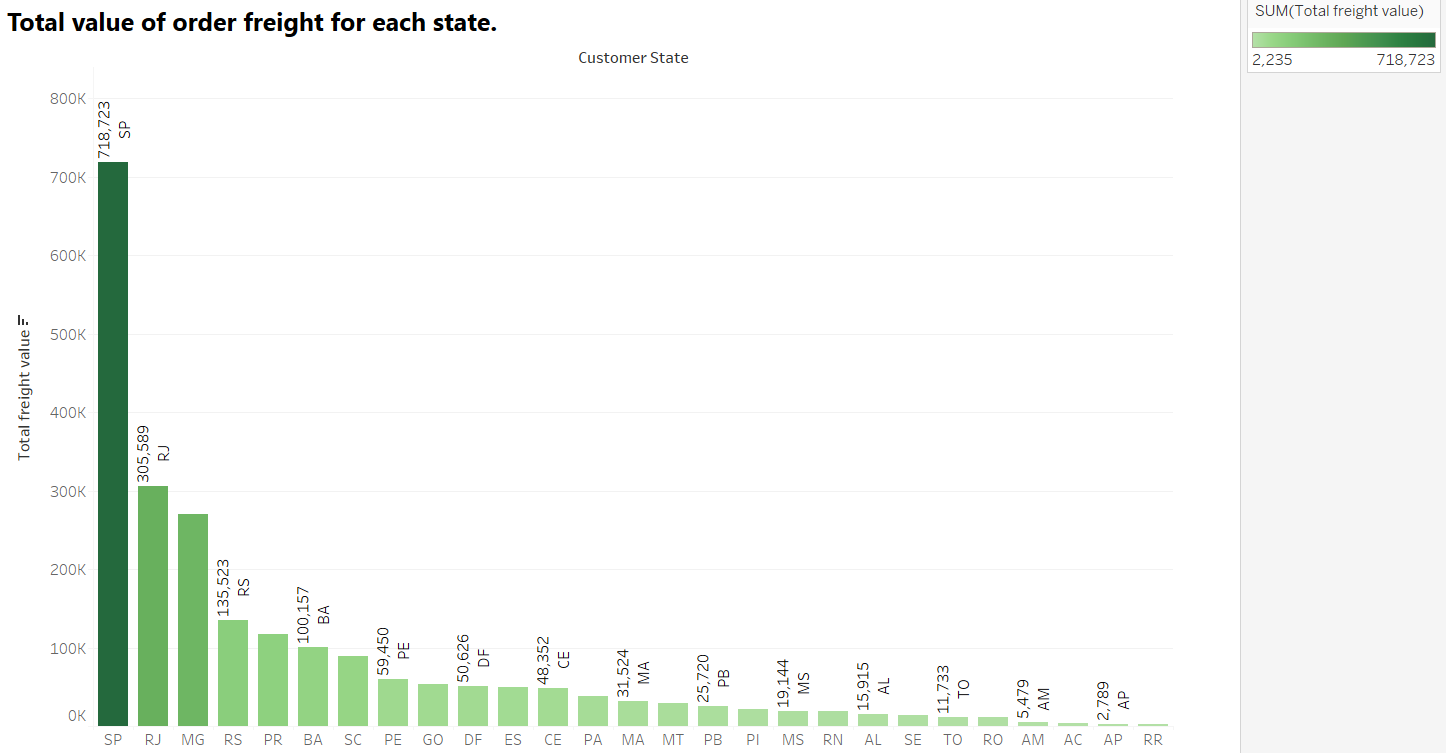
**OBSERVATION:-**

* **SP has the highest total orders price at 5,202,955**
* **Top 3 high-performing states are SP, RJ, MG**
* **RR has the lowest total order price Followed by AP**

****

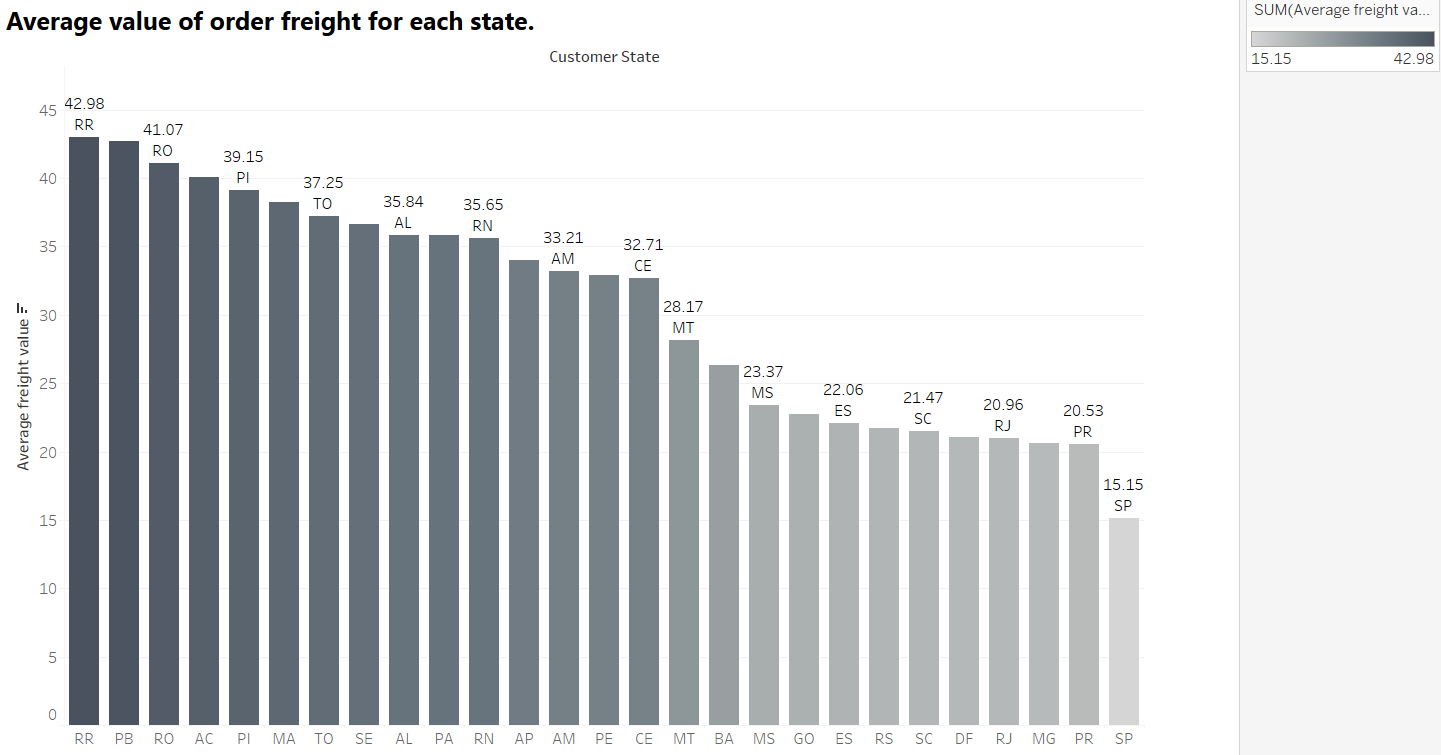
**OBSERVATION:-**

* **Top 3 states with Highest Average Order Price PB, AL, AC.**
* **These state have customers who, on average, spend more per order**
* **3 States with the lowest Average Order Price are SP, PR, RS.**

****

**OBSERVATION:-**

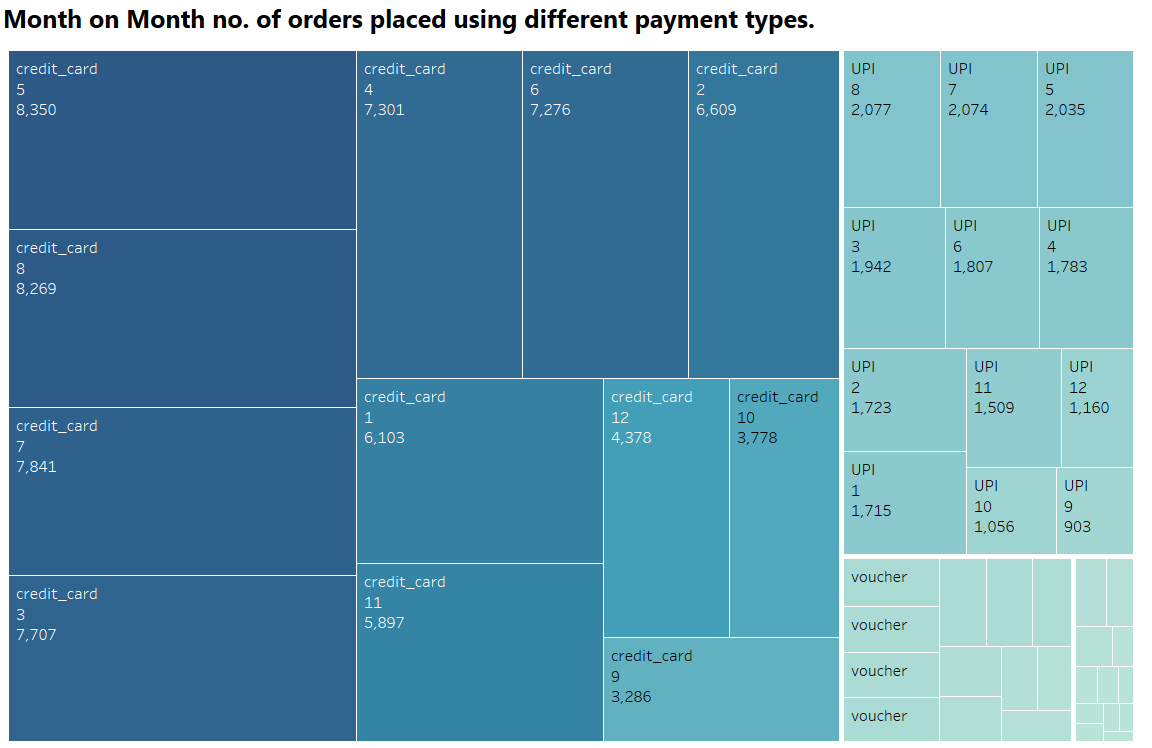
* **Top 3 states with the Highest Freight value are SP, RJ, MG**
* **These states contributes overall freight costs, due to high order volumes**
* **States with the lowest freight value are RR, AP, AM**
* **SP has the Highest freight value but the lowest average order price**

****

**OBSERVATION:-**

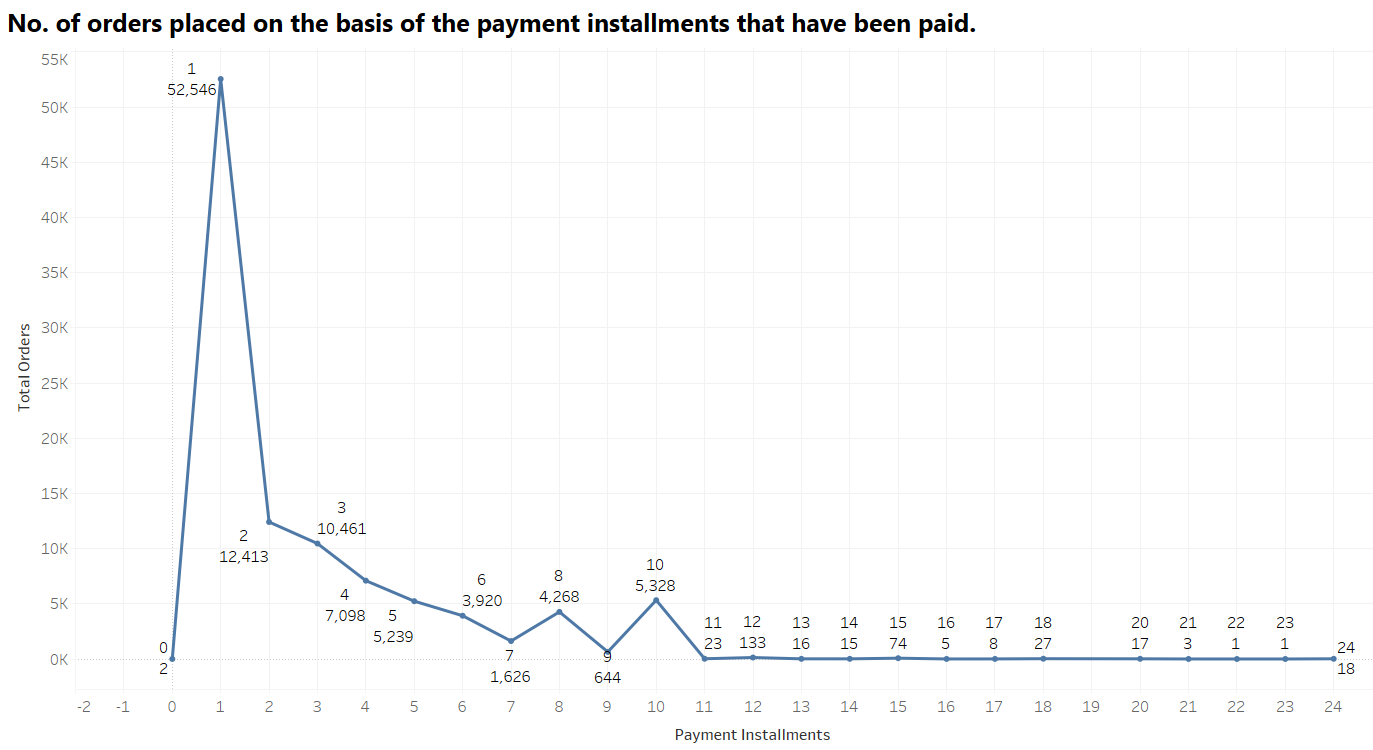
* **States with the highest average freight value are RR, RO, PI, Due to remote locations**
* **States with lowest average freight value are SP,PR, RJ**

**Analysis based on the payments:**

****

**OBSERVATION:-**

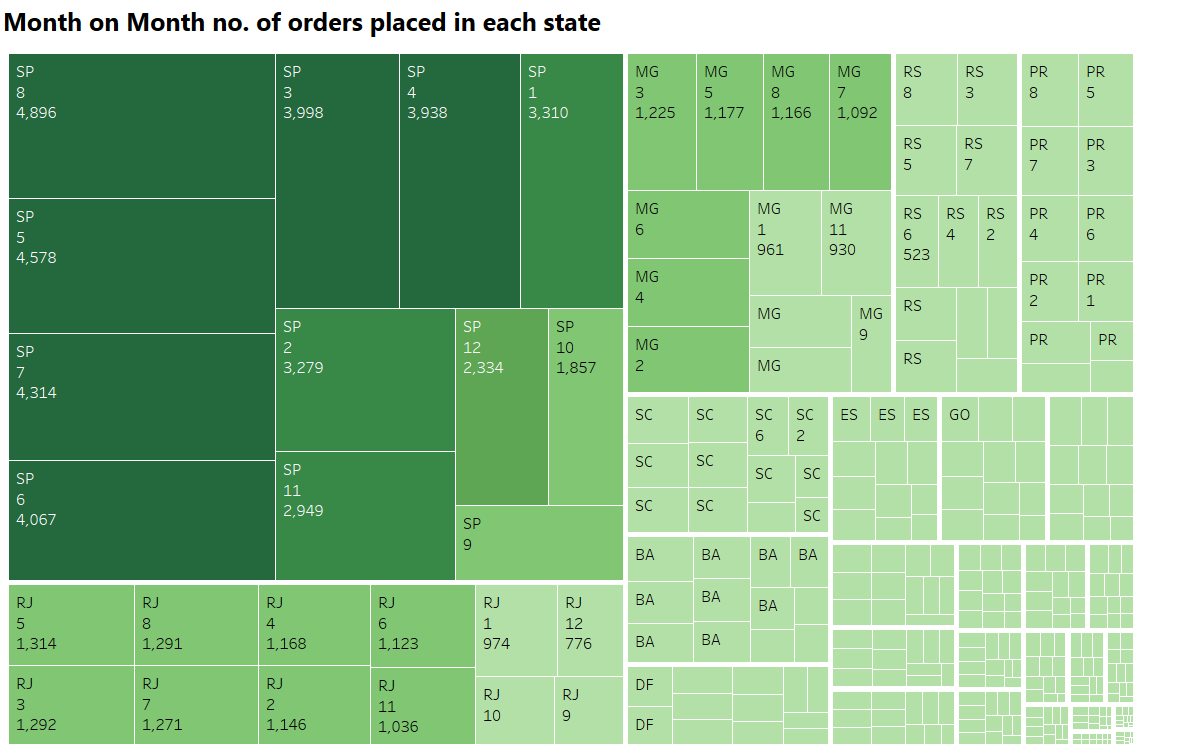
* **The majority of orders are placed using Credit cards across all months**
* **The highest order count using credit card appears in Month 5, Followed closely by month 8**
* **The highest number of UPI is in Month 8**
* **UPI orders remain stable but significantly lower than credit card transactions**
* **Voucher-based transactions appear in the smallest segments, indicating limited edition**



**OBSERVATION:-**

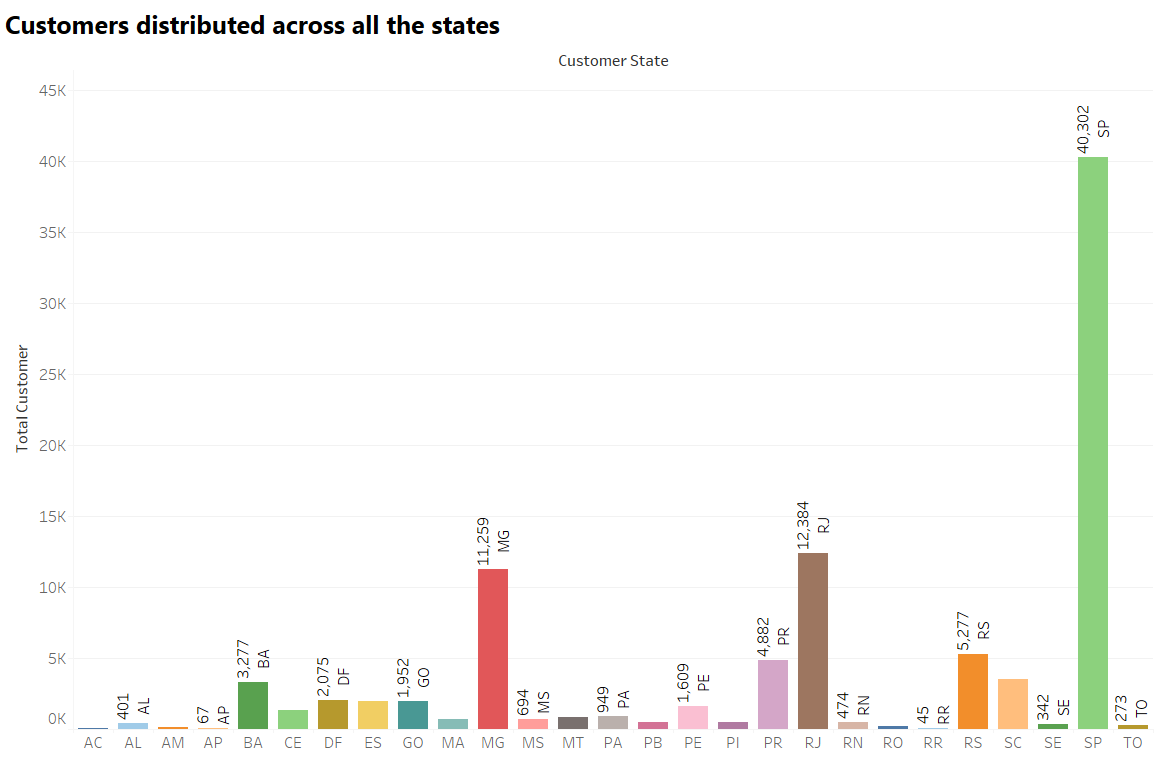
* Majority of orders were placed with a single payment instalment
* Orders decline sharply as the number of payment instalments increases
* Most customers prefer full-payment method
* Instalment 2 & 3 have moderate adoption
* While the overall trend decreases, slight peaks at instalments 6,8, and 10

**E-commerce orders**



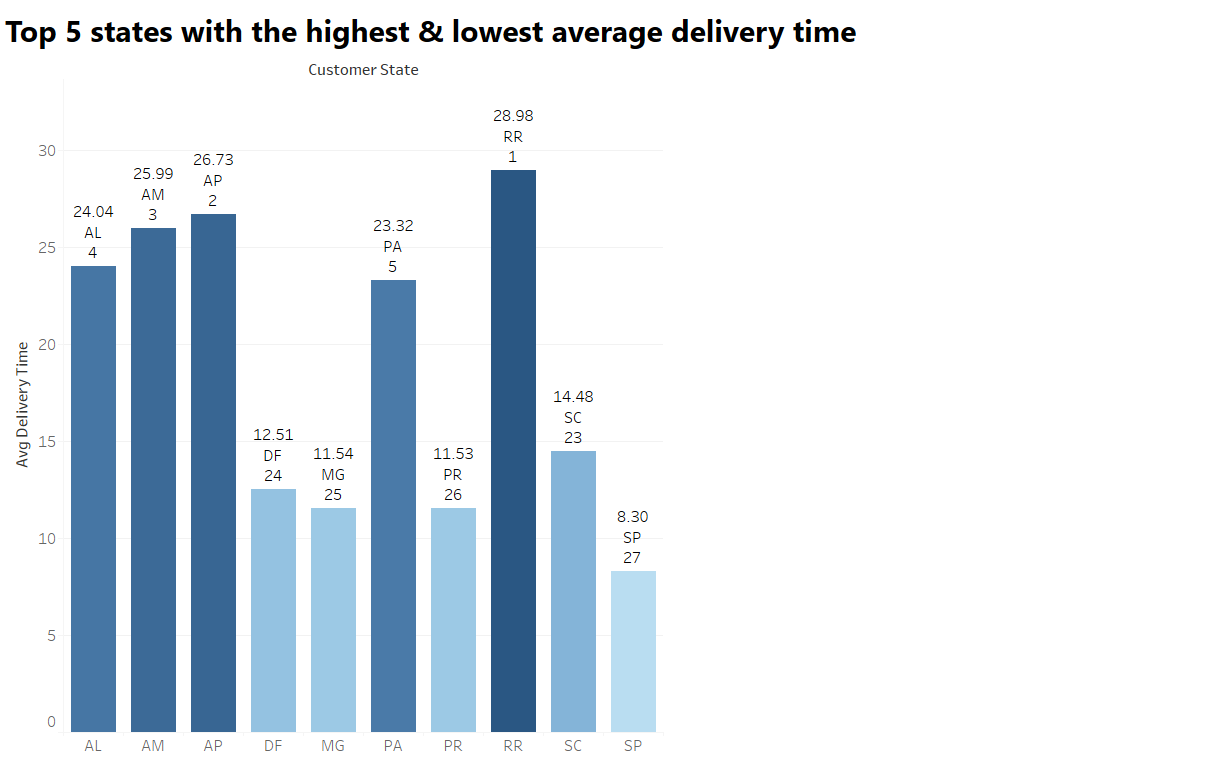
**OBSERVATION:-**

* SP dominates the chart with multiple months showing high order counts
* The highest order count is 4896 in SP in 8th Month, Followed by other months like SP-5, SP-7, and SP-6
* Moderate Activity in MG and RJ
* Lower order volume in States are RS, PR, SC



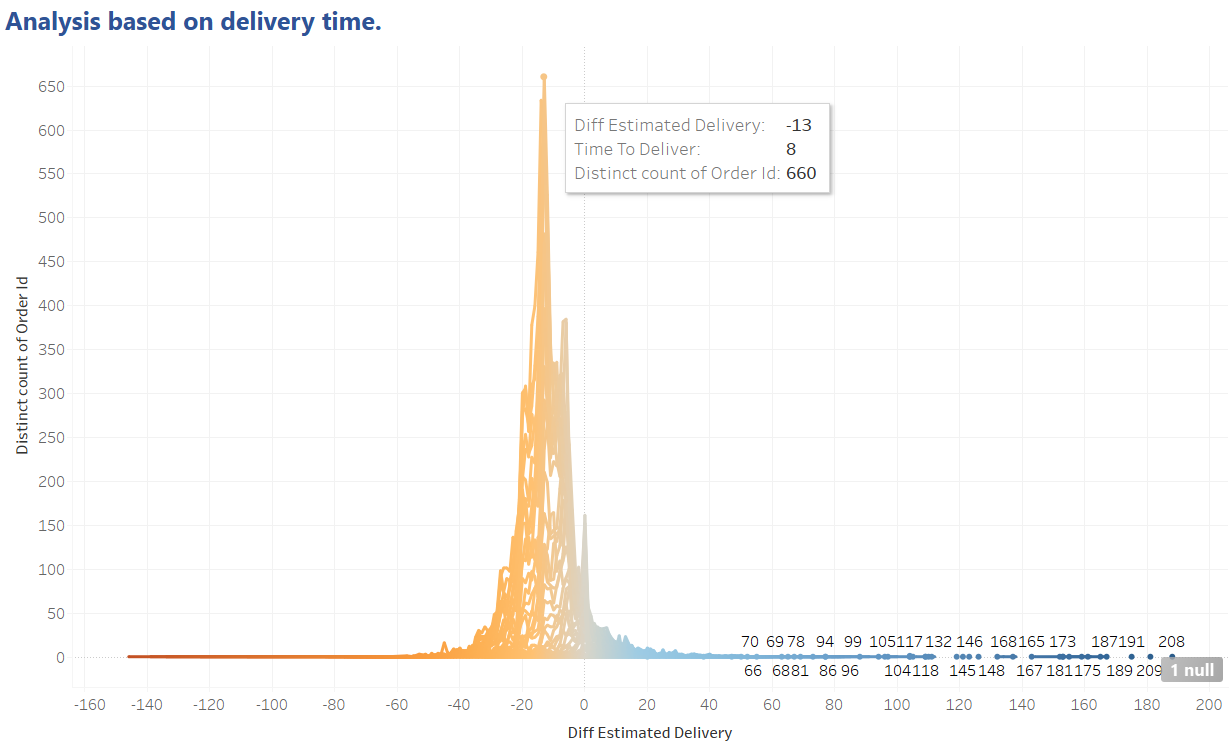
**OBSERVATIONS:-**

* With 40,302 customers, SP has the highest number of customers
* RJ, And MG are the Next Major Markets RJ-12,384 and MG-11,259 Customers
* Moderate customer base in PR, RS, BA



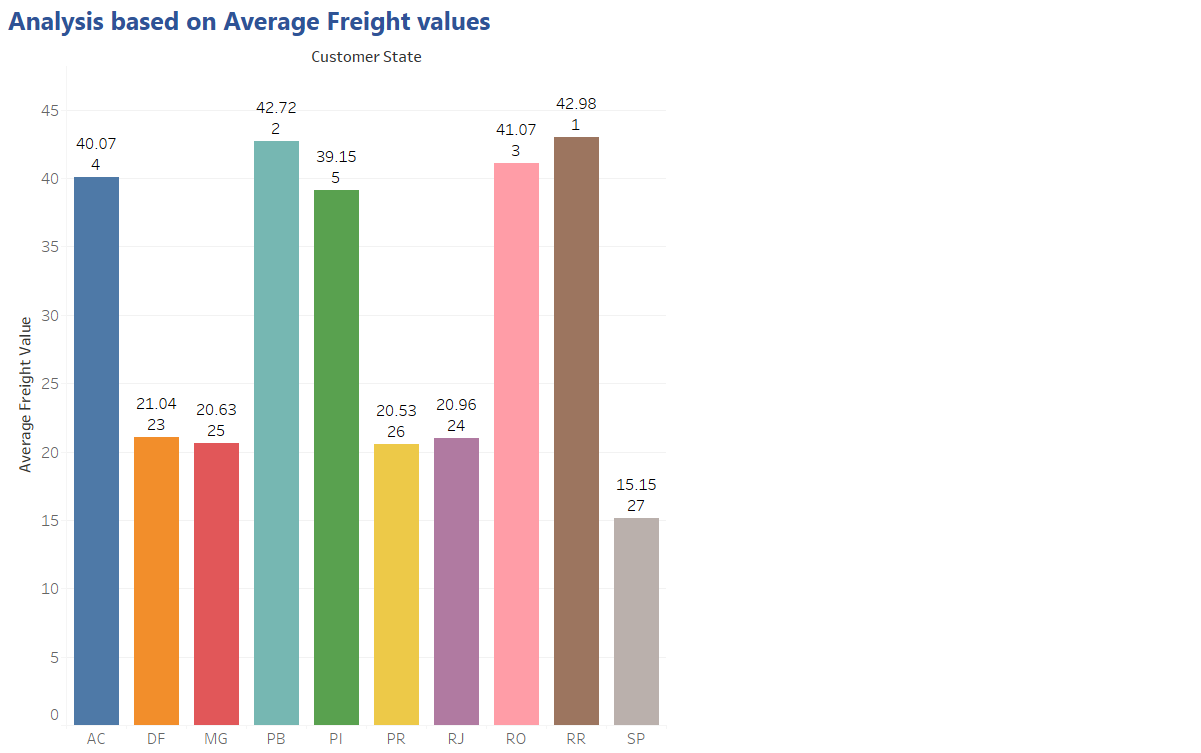
OBSERVATION:

* States with highest average delivery time are RR-28.98 days, AP-26.73 days, AM-25.99 days, AL-24.04 days, PA-23.32 days
* Long delivery times can lead to customer dissatisfaction and higher order cancellations
* States with Lowest average delivery time are SP,PR,MG, DF, SC
* Shorter delivery times contribute to higher customer satisfaction.



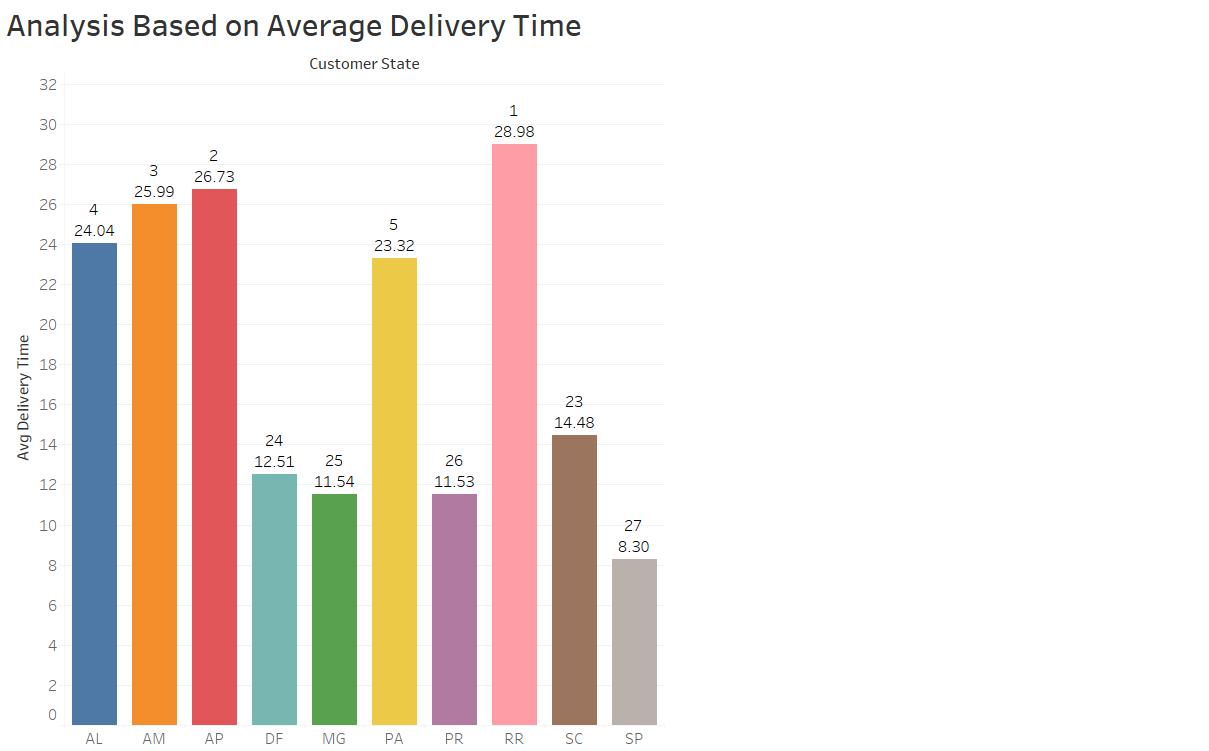
**OBSERVATION:**

* The peak of distribution occurs at negative values means most orders were delivered earlier than the estimated date
* Some orders experienced delays exceeding 100+ days.
* This could be due to remote locations and failed delivery attempts



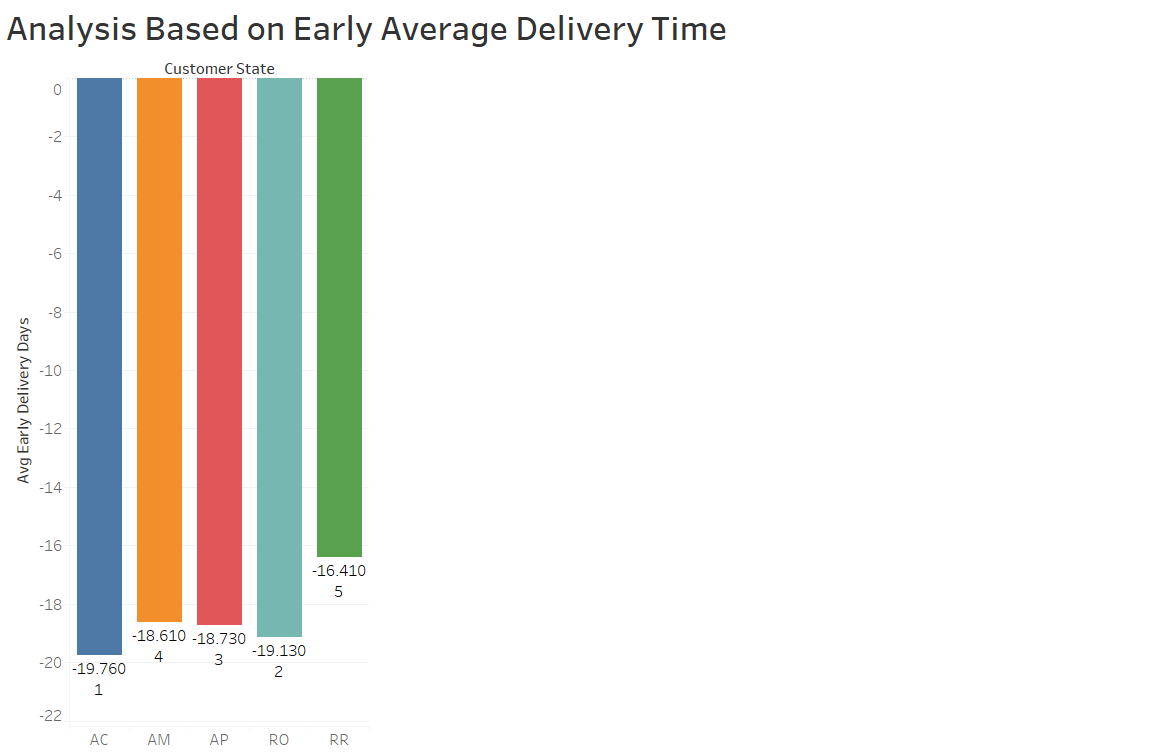
**OBSERVATION:**

* The states with higher freight values may be more remote or have less developed logistics network
* The states with lower freight values are likely to be urban centres with well-established transportation systems.



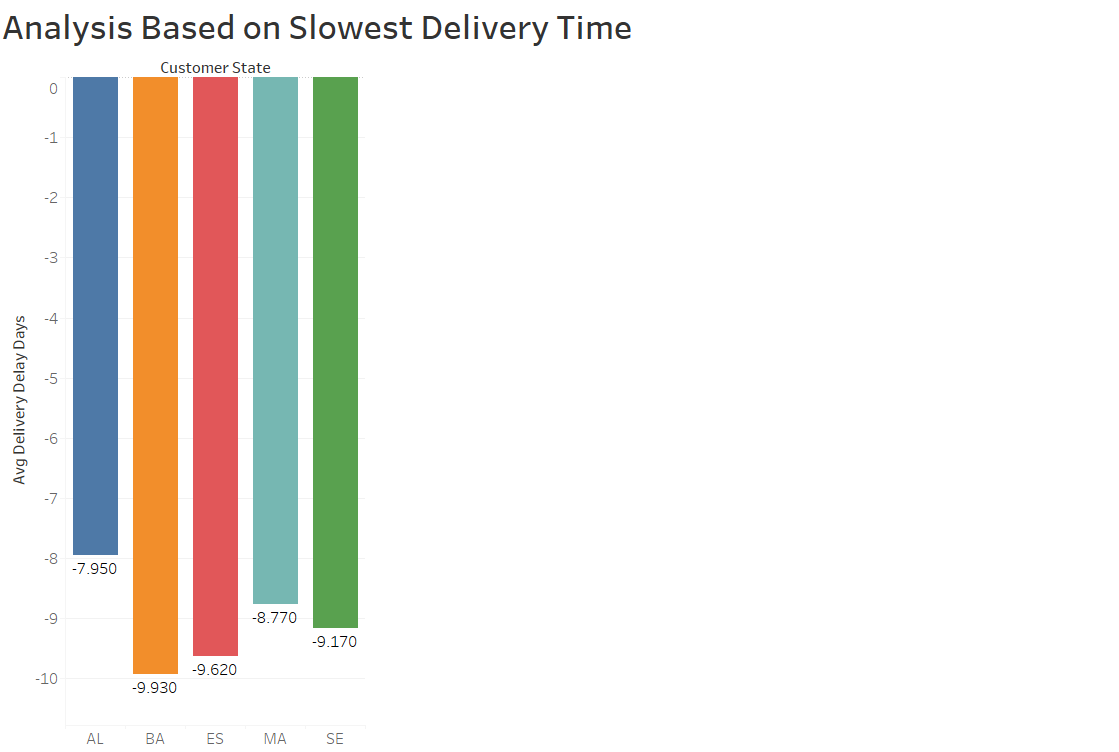
**OBSERVATION:**

* Top 5 states with the Longest delivery Time are RR, AP, AM, AL, PA. These states may have longer shipping times due to greater distances from distribution centres.
* Bottom 5 states with the shortest Delivery time are SP, PR, MG, DF, SC. These states likely benefit from well-developed transportation.



**OBSERVATION:**

* AC has the highest average delivery time, followed by AM and AP
* RR has the shortest delivery time is 16 days Early arriving orders



OBSERVATION:

* The States like BA, ES, SE have Slowest Average delivery times
* State like AL much better than remaining states Average delivery time.

***INSIGHTS:-***

***1.ORDER TRENDS***

***Growing Trend:***

***Orders have shown a steady increase over the years, with 2017, 45,101 orders and 2018, 54,011 orders. seeing higher order volumes than 2016, 329 orders***

***Monthly seasonality:***

***a spike in orders is observed during key months, like August, May, July, and Followed by March, June, April***

***Peak Order Time:***

***Most orders are placed in the Afternoon(1pm - 6pm) and Night(7pm - 11pm), suggesting consumers preferred shopping hours.***

***2.Where Customers Are Buying From:***

***Top States From Orders:Most Orders come from SP, MG, And RJ***

***Low orders in some areas like RR, AP have very few orders possibly due to delivery issues***

***3.Revenue and freight cost insights***

***More Money spent Over Time:***

***A significant increase in total order value from 2017 to 2018(January-august) was observed, showing a rise in consumer spending.***

***People are spending more each year, and total revenue increasing.***

***Expensive Deliveries in Remote areas:***

***The farther the customer, the higher the shipping cost.***

***Top spending state:***

***SP contributes the highest in total order value, while smaller states show lower average spending.***

***4.Delivery Performance and Challenges:***

***States with the longest average delivery times include RR, AP, AM with delivery durations exceeding 25 days.***

***SP, MG, PR exhibit the shortest delivery times, averaging around 8-12 days.***

***Top states for early delivery are AC, RO, AP***

***5.Payment insights:***

***Most preferred payment method:Credit card dominate payments***

***A significant portion of orders is purchased through instalment plans***

***CONCLUSION***

***Target Brazil’s e-commerce business is growing, but it faces challenges with slow deliveries, high shipping costs, and low orders in remote areas***

***Most sales come from big cities, while rural areas still have logistical problems.***

***Improving delivery speed, offering better payment options, and attracting customers from less popular areas can help grow business***

***Recommendations:***

***1. Freight Cost Reduction Strategies:***

***Implement regional warehouses to minimize long-distance shipping costs.***

***Offer free or discounted shipping for high-value orders to incentivize larger purchases.***

***2. Targeted Marketing in Low-Sales States:***

***Focus advertising and promotions in states with low order volumes to drive customer engagement.***

***Provide incentives like first-time purchase discounts in underperforming regions.***

***3. Optimize Product Pricing & Shipping Policies:***

***Introduce bundled pricing strategies where customers can combine items to get free shipping.***

***Adjust pricing models to accommodate varying freight costs across regions.***

***4. Improve Delivery Performance:***

***Analyse supplier-to-customer routes and optimize for efficiency.***

***Explore partnerships with regional delivery services for faster and cheaper shipping.***