TARGET CASE STUDY

Q1. **Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset**

Q1.1 **Data type of all columns in the "customers" table**

<u>Query</u>

```
SELECT
    column_name,
    data_type
FROM Target_case_study.INFORMATION_SCHEMA.COLUMNS
WHERE table_name = 'customers'
```

<u>Output</u>

column_name ▼	data_type ▼
customer_id	STRING
customer_unique_id	STRING
customer_zip_code_prefix	INT64
customer_city	STRING
customer_state	STRING

<u>Insights</u>

➢ In the "customers" table, columns namely customer_id, customer_unique_id, customer_city and customer_state are having string data type whereas customer_zip_code_prefix column is having integer data type.

➢ Customer_id is the one created in the backend to register any new customer in any company's database, whereas customer_unique_id is the id which a customer creates while signing up.
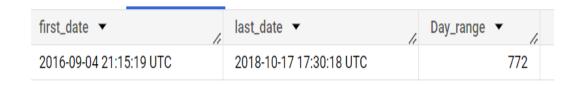
Q1.2.   Get the time range between which the orders were placed.

<u>Query:</u>

```sql
SELECT
  MIN(order_purchase_timestamp) AS first_date,
  MAX(order_purchase_timestamp) AS last_date,
  DATE_DIFF(MAX(order_purchase_timestamp),MIN(order_purchase_timestamp),day) AS Day_range
FROM `Target_case_study.orders`
```

<u>Output</u>

first_date ▼	last_date ▼	Day_range ▼
2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC	772

<u>Insights</u>

➢ The dataset pertains to the data wherein first order was placed on 4th september, 2016 and last order was placed on 17th october, 2018.

➢ So its a study made for over a range of 772 days.

Q1.3. Count the number of Cities and States in our dataset

Query

```sql
SELECT
      count (DISTINCT c.customer_state) AS States,
      count (DISTINCT c.customer_city) AS Cities
FROM `Target_case_study.customers` AS c
JOIN `Target_case_study.orders` AS o
      ON c.customer_id = o.customer_id
```

Output

States ▼	Cities ▼
27	4119

Insights

➢ The dataset is pertaining to 27 unique states and 4119 cities of Brazil

Q2.   In-depth Exploration

Q2.1   Is there a growing trend in the no. of orders placed over the past years?

<u>Query</u>

```sql
WITH extracted_year_and_orders AS(
      SELECT
            EXTRACT(year FROM order_purchase_timestamp) AS _year,
            EXTRACT(month FROM order_purchase_timestamp) AS _month,
            COUNT(order_id) AS current_orders
      FROM `Target_case_study.orders`
      GROUP BY
            EXTRACT(year FROM order_purchase_timestamp),
            EXTRACT(month FROM order_purchase_timestamp))
SELECT
      *,
      LAG(current_orders,1) OVER(ORDER BY _year, _month) AS prev_orders,
      round(((current_orders - LAG(current_orders,1) OVER(ORDER BY _year,
      _month))*100/LAG(current_orders,1) OVER(ORDER BY _year, _month)),2) AS
      percent_change_in_odr,
      CASE
      WHEN current_orders > LAG(current_orders,1) OVER(ORDER BY _year, _month)
            THEN "orders are rising"
      WHEN current_orders < LAG(current_orders,1) OVER(ORDER BY _year, _month)
            THEN "orders are falling"
      END AS status
FROM extracted_year_and_orders
ORDER BY _year, _month
```

Output

_year	_month	current_orders	prev_orders	percent_change_in_o	status
2016	9	4	*null*	*null*	*null*
2016	10	324	4	8000.0	orders are rising
2016	12	1	324	-99.69	orders are falling
2017	1	800	1	79900.0	orders are rising
2017	2	1780	800	122.5	orders are rising
2017	3	2682	1780	50.67	orders are rising
2017	4	2404	2682	-10.37	orders are falling
2017	5	3700	2404	53.91	orders are rising
2017	6	3245	3700	-12.3	orders are falling
2017	7	4026	3245	24.07	orders are rising

Insights

➢ The dataset we have is from September 2016 till October 2018.

➢ In 2016 there was a period of just 3 to 4 months, among which October had the highest peak of 8000% as compared to September, in terms of the number of orders placed.

➢ In 2017 the orders placed were increased over the entire year except in June, May and December. There were dips of 10.37%, 12.3%, 24.8% respectively

➢ In 2018, the number of order placed increased by 28.13% in January after that there was a drastic decline over the year having a decline of 99.75% in the month of September

➢ So overall we can see a rising trend in the number of orders placed till March 2018, after that there was a declining trend.

Q2.2. Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

Query

```sql
WITH extract_seasonality_over_month AS(
    SELECT
            EXTRACT(month FROM order_purchase_timestamp) AS _month,
            COUNT(order_id) AS current_orders
    FROM `Target_case_study.orders`
    GROUP BY
        EXTRACT(month FROM order_purchase_timestamp))

SELECT
    *,
    LAG(current_orders) OVER(ORDER BY _month) AS prev_orders,
    ROUND(((current_orders - LAG(current_orders,1) OVER(ORDER BY _month)) * 100
    / LAG(current_orders,1) OVER (ORDER BY _month)),2) AS
    percent_change_in_odr,
    CASE
            WHEN current_orders > LAG(current_orders,1) OVER(ORDER BY _month)
                THEN "orders are rising"
            WHEN current_orders < LAG(current_orders,1) OVER(ORDER BY _month)
                        THEN "orders are falling"
    END AS status
FROM extract_seasonality_over_month
ORDER BY _month
```

Output

_month	current_orders	prev_orders	percent_change_in_o	status
1	8069	null	null	null
2	8508	8069	5.44	orders are rising
3	9893	8508	16.28	orders are rising
4	9343	9893	-5.56	orders are falling
5	10573	9343	13.16	orders are rising
6	9412	10573	-10.98	orders are falling
7	10318	9412	9.63	orders are rising
8	10843	10318	5.09	orders are rising
9	4305	10843	-60.3	orders are falling
10	4959	4305	15.19	orders are rising
11	7544	4959	52.13	orders are rising
12	5674	7544	-24.79	orders are falling

Insights

➢ We can see a rise and fall in the number of orders being placed over the months.

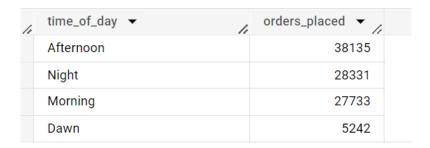➢ There is a 60% drop in the month of September and a 52% rise in November.

Q2.3 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

```sql
WITH day_time AS (
    SELECT
        *,
        CASE
        WHEN EXTRACT (time FROM order_purchase_timestamp) BETWEEN "00:00:00" AND
    "06:59:59" THEN "Dawn"
        WHEN EXTRACT (time FROM order_purchase_timestamp) BETWEEN "07:00:00" AND
    "12:59:59" THEN "Morning"
        WHEN EXTRACT (time FROM order_purchase_timestamp) BETWEEN "13:00:00" AND
    "18:59:59" THEN "Afternoon"
        WHEN EXTRACT (time FROM order_purchase_timestamp) BETWEEN "19:00:00" AND
    "23:59:59" THEN "Night"
        END AS time_of_day
    FROM Target_case_study.orders`)

SELECT
    time_of_day,
    COUNT(order_id) AS orders_placed
FROM day_time
GROUP BY time_of_day
ORDER BY COUNT(order_id) DESC
```

Output

time_of_day ▼	orders_placed ▼
Afternoon	38135
Night	28331
Morning	27733
Dawn	5242

Insights

➢ Most orders are placed in the afternoon from 1 pm to 7 pm.

➢ Least orders are placed at dawn from 12 am to 7 am.

Q3 Evolution of E-commerce orders in the Brazil region:

Q3.1 Get the month on month no. of orders placed in each state.

<u>Query</u>

```sql
SELECT *
FROM (
      SELECT
            DISTINCT c.customer_state,
            EXTRACT(month FROM o.order_purchase_timestamp) AS _month,
            count (o.order_id) AS orders_placed
      FROM `Target_case_study.customers` AS c
      JOIN `Target_case_study.orders` AS o
            ON c.customer_id = o.customer_id
      GROUP BY
            EXTRACT(month FROMo.order_purchase_timestamp),
            c.customer_state)AS x
ORDER BY Customer_state, _month
```

Output

customer_state	_month	orders_placed
SP	1	3351
SP	2	3357
SP	3	4047
SP	4	3967
SP	5	4632
SP	6	4104
SP	7	4381
SP	8	4982
SP	9	1648
SP	10	1908
SP	11	3012
SP	12	2357

customer_state	_month	orders_placed
AP	1	11
AP	2	4
AP	3	8
AP	4	5
AP	5	11
AP	6	4
AP	7	7
AP	8	5
AP	9	2
AP	10	3
AP	11	4
AP	12	4

<u>Insights</u>

➢ Month on month, the number of orders placed for each state of Brazil is analyzed and we can say that, total orders placed month over month for the state of Sao Paulo (SP) is highest across all states.

➢ Total orders placed month over month for the state of Amapa (AP) is least across all states

Q3.2   How are the customers distributed across all the states

Query

```
SELECT
      customer_state,
      COUNT(customer_id) AS cus_across_states
FROM `Target_case_study.customers`
GROUP BY customer_state
ORDER BY cus_across_states desc
```

Output

	customer_state	cus_across_states
1	SP	41746
2	RJ	12852
3	MG	11635
4	RS	5466
5	PR	5045
6	SC	3637
7	BA	3380
8	DF	2140
9	ES	2033
10	GO	2020

Insights

➢ By looking at the customer distribution over all the 27 states of Brazil, we can say that "Target" has the highest customers located in the state of Sao Paulo(SP), followed by the state of RIO de Janeiro(RJ), state of Minas Gerais(MG) and so on.

Q4. Impact on Economy: Analyze the money movement by e-commerce by looking at order prices, freight and others

Q4.1 Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

<u>Query</u>

```sql
WITH extract_year_month AS(
    SELECT
        EXTRACT (year FROM order_purchase_timestamp) AS _year,
        EXTRACT (month FROM order_purchase_timestamp) AS _month,
        ROUND(SUM(p.payment_value),2) AS current_month_cost
    FROM `Target_case_study.orders` AS o
    JOIN `Target_case_study.payments` AS p
        ON o.order_id = p.order_id
    GROUP BY
        EXTRACT (year FROM order_purchase_timestamp),
        EXTRACT (month FROM order_purchase_timestamp) ),
lag_cost AS (
    SELECT *,
        LAG(current_month_cost) OVER (ORDER BY _year, _month) AS
prev_month_cost
    FROM extract_year_month
    WHERE
        _month BETWEEN 1 AND 8 AND
        _year IN (2017, 2018)
    ORDER BY _year, _month )
SELECT *,
    ROUND((current_month_cost - prev_month_cost)*100/(prev_month_cost),2) AS
percent_increase
    FROM lag_cost
```

Output

_year	_month	current_month_cost	prev_month_cost	percent_increase
2017	1	138488.04	null	null
2017	2	291908.01	138488.04	110.78
2017	3	449863.6	291908.01	54.11
2017	4	417788.03	449863.6	-7.13
2017	5	592918.82	417788.03	41.92
2017	6	511276.38	592918.82	-13.77
2017	7	592382.92	511276.38	15.86
2017	8	674396.32	592382.92	13.84
2018	1	1115004.18	674396.32	65.33
2018	2	992463.34	1115004.18	-10.99

Insights

➢ Cost of order in feb 2017 was 110% higher as compared to jan 2017 and gradually it decreases as we move ahead in the year.

➢ There was a 13.77% fall in the prices in June 2017.

➢ In 2018, there's a hike of 65% in Jan 2018 as compared to Dec 2017.

➢ There was a 11.27% fall in the prices in June 2018.

Q4.2   Calculate the Total & Average value of order price for each state.

```
WITH analysis AS (
      SELECT
            c.customer_state,
            ROUND(SUM(p.payment_value),2) AS Total_order_price,
            ROUND(AVG(p.payment_value),2) AS Avg_order_price
      FROM `Target_case_study.customers` AS c
      JOIN `Target_case_study.orders` AS o
            ON c.customer_id = o.customer_id
      JOIN `Target_case_study.payments` AS p
            ON p.order_id = o.order_id
      GROUP BY c.customer_state)

SELECT *
FROM analysis
ORDER BY analysis.Total_order_price desc
```

Output

customer_state	Total_order_price	Avg_order_price
SP	5998226.96	137.5
RJ	2144379.69	158.53
MG	1872257.26	154.71
RS	890898.54	157.18
PR	811156.38	154.15
SC	623086.43	165.98
BA	616645.82	170.82
DF	355141.08	161.13
GO	350092.31	165.76
ES	325967.55	154.71

Insights

➢ Total order price is highest for SP (São Paulo) and least for RR (Roraima)

➢ Average order price is highest for PB (Paraiba) and least for SP (São Paulo)

Q4.3   Calculate the Total & Average value of order freight for each state.

Query

```sql
WITH analysis AS (
    SELECT
        c.customer_state,
        ROUND(SUM(oi.freight_value),2) AS Total_order_freight,
        ROUND(AVG(oi.freight_value),2) AS Avg_order_freight
    FROM `Target_case_study.customers` AS c
    JOIN `Target_case_study.orders` AS o
        ON c.customer_id = o.customer_id
    JOIN `Target_case_study.order_items` As oi
        ON o.order_id = oi.order_id
    GROUP BY c.customer_state)

SELECT *
FROM analysis
ORDER BY analysis.Total_order_freight desc
```

Output

customer_state	Total_order_freight	Avg_order_freight
SP	718723.07	15.15
RJ	305589.31	20.96
MG	270853.46	20.63
RS	135522.74	21.74
PR	117851.68	20.53
BA	100156.68	26.36
SC	89660.26	21.47
PE	59449.66	32.92
GO	53114.98	22.77
DF	50625.5	21.04

Insights

➢ Total order freight value is highest for SP (São Paulo) and least is for RR (Roraima)

➢ Average order freight value is highest for RR (Roraima) and least for SP (São Paulo)

Q5   Analysis based on sales, freight and delivery time.

Q5.1   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.

Query

```
SELECT
      order_id,
      customer_id,
      DATE_DIFF(order_delivered_customer_date, order_purchase_timestamp, day) AS
time_to_deliver,
      DATE_DIFF(order_estimated_delivery_date, order_delivered_customer_date,
day) AS diff_estimated_delivery
FROM `Target_case_study.orders`
ORDER BY
      time_to_deliver DESC,
      Diff_estimated_delivery
```

Output

order_id ▼	customer_id ▼	time_to_deliver ▼	diff_estimated_delive
ca07593549f1816d26a572e06…	75683a92331068e2d281b11a…	209	-181
1b3190b2dfa9d789e1f14c05b…	d306426abe5fca15e54b645e4…	208	-188
440d0d17af552815d15a9e41a…	7815125148cfa1e8c7fee1ff79…	195	-165
285ab9426d6982034523a855f…	9cf2c3fa2632cee748e1a59ca9…	194	-166
0f4519c5f1c541ddec9f21b3bd…	1a8a4a30dc296976717f44e78…	194	-161
2fb597c2f772eca01b1f5c561b…	217906bc11a32c1e470eb7e08…	194	-155
47b40429ed8cce3aee9199792…	cb2caaaead400c97350c37a3f…	191	-175
2fe324febf907e3ea3f2aa9650…	65b14237885b3972ebec28c0f…	189	-167
2d7561026d542c8dbd8f0daea…	8199345f57c6d1cbe9701f924…	188	-159
c27815f7e3dd0b926b5855262…	f85e9ec0719b16dc4dd0edd43…	187	-162

Insights

➢ The orders shown in the output need more attention because those are the ones which took more than 187 days to deliver the order from the date it was placed.

➢ These orders are late by more than 160 days from the estimated date of order delivery.

Q5.2 Find out the top 5 states with the highest & lowest average freight value.

Query

```
WITH analysis AS (
      SELECT
            c.customer_state,
            ROUND(AVG(oi.freight_value),2) AS Avg_order_freight
      FROM `Target_case_study.customers` AS c
      JOIN `Target_case_study.orders` AS o
            ON c.customer_id = o.customer_id
      JOIN `Target_case_study.order_items` AS oi
            ON o.order_id = oi.order_id
      GROUP BY c.customer_state)


SELECT *
FROM analysis
ORDER BY analysis.Avg_order_freight DESC
LIMIT 5
```

Output

States having highest avg freight value		States having lowest avg freight value	
customer_state ▼	Avg_order_freight ▼	customer_state ▼	Avg_order_freight ▼
RR	42.98	SP	15.15
PB	42.72	PR	20.53
RO	41.07	MG	20.63
AC	40.07	RJ	20.96
PI	39.15	DF	21.04

<u>Insights</u>

➢ States having highest average freight value are RR (Roraima), PB (Paraiba), RO (Ronodonia), AC (Acre), PI (Piaui)

➢ States having lowest average freight value are SP (São Paulo), PR (Paraná), MG (Minas Gerais), RJ (Rio de Janeiro), DF (Distrito Federal)

Q5.3   Find out the top 5 states with the highest & lowest average delivery time.

Query

```sql
WITH analysis AS (
    SELECT
        c.customer_state,
        AVG(DATE_DIFF(order_delivered_customer_date,
order_purchase_timestamp, day)) AS time_to_deliver
    FROM `Target_case_study.customers` AS c
    JOIN `Target_case_study.orders` AS o
        ON c.customer_id = o.customer_id
    JOIN `Target_case_study.order_items` AS oi
        ON o.order_id = oi.order_id
    GROUP BY c.customer_state)

SELECT
    customer_state,
    ROUND(analysis.time_to_deliver,2)
FROM analysis
ORDER BY analysis.time_to_deliver
LIMIT 5
```

Output

States having highest avg delivery time	States having lowest avg delivery time
customer_state	avg_delivery_time	customer_state	avg_delivery_time
RR	27.83	SP	8.26
AP	27.75	PR	11.48
AM	25.96	MG	11.52
AL	23.99	DF	12.5
PA	23.3	SC	14.52

Insights

➢ States having highest average delivery time are RR (Roraima), AP (Amapa), AM (Amazonas), AL (Alagoas), PA (Para)

➢ States having lowest average delivery time are SP (São Paulo), PR (Paraná), MG (Minas Gerais), DF (Distrito Federal), SC (Santa Catarina)

Q5.4 Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

Query

```
WITH analysis AS (
    SELECT
        C.customer_state,
        AVG (DATE_DIFF(order_estimated_delivery_date,
order_delivered_customer_date, day)) AS diff_estimated_delivery
    FROM `Target_case_study.customers` AS c
    JOIN `Target_case_study.orders` AS o
        ON c.customer_id = o.customer_id
    JOIN `Target_case_study.order_items` AS oi
        ON o.order_id = oi.order_id
    GROUP BYc.customer_state)

SELECT
    customer_state,
    ROUND(analysis.diff_estimated_delivery,2) AS avg_delivery_time
FROM analysis
ORDER BY analysis.diff_estimated_delivery DESC
LIMIT 5
```

Output

customer_state	avg_delivery_time
AC	20.01
RO	19.08
AM	18.98
AP	17.44
RR	17.43

Insights

➢ States where order delivery is very fast as compared to the estimated delivery date are, AC (Acre), RO (Ronodonia), AM (Amazonas), AP (Amapa), RR (Roraima)

Q6 Analysis based on the payments

Q6.1 Find the month on month no. of orders placed using different payment

Query

```
SELECT *
FROM (
    SELECT
        EXTRACT (month FROM o.order_purchase_timestamp) AS _month,
        p.payment_type,
        COUNT(o.order_id) as no_of_orders
    FROM `Target_case_study.orders` AS o
    JOIN `Target_case_study.payments` AS p
        ON o.order_id = p.order_id
    GROUP BY
        EXTRACT (month FROM o.order_purchase_timestamp),
        p.payment_type) AS x
ORDER BY
    X._month,
    x.no_of_orders desc
```

Output

_month	payment_type	no_of_orders
1	credit_card	6103
1	UPI	1715
1	voucher	477
1	debit_card	118
2	credit_card	6609
2	UPI	1723
2	voucher	424
2	debit_card	82
3	credit_card	7707
3	UPI	1942
3	voucher	591
3	debit_card	109

Insights

➢ While analyzing the month on month number of orders placed, every month
the most used payment types were credit_card and UPI

Q6.2 Find the no. of orders placed on the basis of the payment installments that have been paid.

Query

```sql
SELECT
    p.payment_installments AS installments,
    p.payment_sequential AS paid_installments,
    COUNT(DISTINCT o.order_id) AS no_of_orders
FROM `Target_case_study.orders` AS o
JOIN `Target_case_study.payments` AS p
    ON o.order_id = p.order_id
WHERE payment_installments > 1
GROUP BY
    p.payment_installments,
    p.payment_sequential
ORDER BY installments
```

Output

installments	paid_installments	no_of_orders
2	1	12360
2	2	53
3	1	10422
3	2	38
3	3	1
4	1	7066
4	2	32
5	1	5221
5	2	18
6	1	3904

<u>Insights</u>

➢ Installments greater than one is considered as EMI, in the dataset there are orders for which customers have taken installments starting from 2 till 24.

➢ There are very few orders for which all the installments were paid. Out of total orders placed taking two installments only customers for 53 orders have paid both the installments. Rest customers pertaining to 12360 orders have paid only one installment.

➢ Orders placed using 10 installments, maximum two of them are paid

➢ Installments larger than 10, for this kind of orders only one installment is paid and rest are yet to complete

Overall Insights & Actionable Recommendations:

"Target" has given 772 days of data pertaining to its business in 27 states and 4119 cities of Brazil. On analyzing the dataset in SQL, below are a few insights and recommendations which will help the growth of business in Brazil.

1. Trend in number of orders placed:
   - Overall there was a rising trend in the number of orders placed from September 2016 till March 2018, after that there was a declining trend.

2. Monthly Seasonality:
   - A 60% drop in the month of September and a 52% rise in November was seen, in terms of the number of orders being placed over the span.

3. Peak Hours where maximum orders were placed:
   - Most orders (38,135) were placed in the afternoon from 1 pm to 7 pm.
   - Hence, when "Target" offers some new product releases or any seasonal sales, it should prioritize this time slot as a larger audience will notice

4. Month on month number. of orders placed in each state:
   - It is seen that the state of Sao Paulo (SP) has placed the highest number of orders over all the months,
   - On the other hand, states like Amapa (AP), Roraima (RR) have the least number of orders placed over the months.

5. Customer distribution across all states:
   - "Target" has the highest number of customers located in the state of Sao Paulo(SP), followed by the state of RIO de Janeiro(RJ), state of Minas Gerais(MG) and so on

- ○ "Target" can introduce products that are particularly pertaining to the culture of these states, because already a good number of the audience is inclined towards the firm. New product launches will succeed in these regions.

6. Percentage increase in the cost of orders from year 2017 to 2018:
   - ○ The financial year of the Brazilian government begins on 1st January and ends on 31st December. Hence that could be one of the reasons for a 110% hike in the cost of orders in February 2017 and similarly a hike of 65% in January, 2018.
   - ○ As government's fiscal planning and budget releases in this part of the year affect the costs of orders in all businesses, due to price hike chances of less orders getting placed might be higher.
   - ○ Also, at the start of 2017, Brazil was suffering from many downfalls like recession, declining GDP, political instability etc so all of this may have impacted the costs of orders for "Target".

7. Order price value for each state:
   - ○ Total order price was higher for the state of Sao Paulo (SP) so to maintain the business in this region, the workforce should be increased. And according to the demand, adequate supply is to be provided to maintain a good customer satisfaction.
   - ○ In state (RR) Roraima, the total order price is the lowest. In these regions, reasons are to be figured out as to why the customers are less interested and their concerns are to be addressed by increasing the supply of products as per the demand. And the items which are having lesser demand should be pulled out. Also keeping in mind that the regular customers' needs should not get affected by these changes.

8. Order Delivery Time Analysis:
   - Efficiency in order delivery needs to be improved, as there were many orders which took more than a 100 days to deliver the product from the date of order purchased. And they were more than 70 days late from the estimated date of delivery.
   - "Target" took the highest average delivery time to deliver orders in the states of RR (Roraima), AP (Amapa), AM (Amazonas), AL (Alagoas), PA (Para).
   - "Target" took the lowest average delivery time to deliver orders in the states of SP (São Paulo), PR (Paraná), MG (Minas Gerais), DF (Distrito Federal), SC (Santa Catarina)
   - States where order delivery is very fast as compared to the estimated delivery date are, AC (Acre), RO (Ronodonia), AM (Amazonas), AP (Amapa), RR (Roraima)
   - States which are getting the deliveries faster as compared to the estimated delivery date but getting late deliveries from the date of purchase are, AM (Amazonas), AP (Amapá), RR (Roraima)

9. Average Freight Value Analysis:
   - States having highest average freight value are RR (Roraima), PB (Paraiba), RO (Ronodonia), AC (Acre), PI (Piaui)
   - States having lowest average freight value are SP (São Paulo), PR (Paraná), MG (Minas Gerais), RJ (Rio de Janeiro), DF (Distrito Federal)
   - To grow the business in regions having the highest average freight value, nearby warehouses and outlets are needed to be set up, so that cost and time of delivery can be reduced.

10. Month on month number of orders placed using different payment methods:
   - Every month the most used payment methods were credit cards and UPI.
   - To increase the sales, discounts can be offered to customers who are using these payment methods.

11. Payment installments paid and its relation with number of orders placed:
   - There are very few orders for which all the installments were paid. Out of total orders placed taking two installments only customers for 53 orders have paid both the installments. Rest customers pertaining to 12360 orders have paid only one installment, whereas for orders placed using ten installments, maximum two of them are paid
   - Installments larger than ten, for this kind of order only one installment is paid and rest are yet to complete.