Analyzing Expansion Prospects in Brazil for a Major US Retailer: SQL Business Case Study

Sumit Chowdhury

August 19, 2023

[Tool: Google BigQuery]

1 Importing the dataset and Exploratory Analysis

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

QUERY

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

RESULT

Row	column_name	data_type
1	customer_id	STRING
2	customer_unique_id	STRING
3	customer_zip_code_prefix	INT64
4	customer_city	STRING
5	customer_state	STRING

- **Observations**: Based on the provided outcome, it's evident that, with the exception of the 'customer_zip_code_prefix' column in the customer table, all other columns have a data type of string.

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

QUERY

```
with cte as
(select min(order_purchase_timestamp) as min_order_placing_timestamp,
max(order_purchase_timestamp) max_order_placing_timestamp
from
`target_buisness_case.orders`)
select *, date_diff(cte.max_order_placing_timestamp,cte.min_order_placing_timestamp,day) Diff_of_dates
from cte;
```

RESULT

Row	min_order_placing_timestamp	max_order_placing_timestamp	Diff_of_dates
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC	772

- **Observations**: Considering the dataset at hand, information spans from September 4th, 2016, to October 17th, 2018. This implies that the dataset encompasses Brazilian consumer data across a span of 772 days, which offers a substantial basis for thorough analysis. Additionally, the timestamps provided are in UTC, requiring a subtraction of 3 hours to align with the Brazilian time zone.

1.3 Count the number of Cities and States in our dataset.

QUERY

```
with cte as
(select *, min(order_purchase_timestamp ) over() as min_order_placing_timestamp,
max(order_purchase_timestamp) over() max_order_placing_timestamp
from `target_buisness_case.orders` o
left join `target_buisness_case.customers` c
on o.customer_id=c.customer_id)

select count(distinct customer_city) as customer_city_count ,
count(distinct customer_state)as customer_state_count,
from cte
where cte.order_purchase_timestamp between min_order_placing_timestamp and max_order_placing_timestamp ;
```

RESULT

Row	customer_city_count	customer_state_count
1	4119	27

- **Observations**:The query outcome reveals that the dataset encompasses information from 27 different states and incorporates data from 4119 distinct cities situated within Brazil.

2 In-depth Exploration:

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

QUERY

```
with cte2 as
(with cte1 as
(with cte as
(SELECT *,
EXTRACT(year FROM o.order_purchase_timestamp)  as order_year,
EXTRACT(month FROM o.order_purchase_timestamp)  as order_month
from `target_buisness_case.orders` o)

select cte.order_year,cte.order_month,count( distinct order_id) as order_count
from cte
group by cte.order_year,cte.order_month
)
select cte1.order_year,cte1.order_month,cte1.order_count,
lag(cte1.order_count) over(order by cte1.order_year,cte1.order_month) as prev_order_count
from cte1
order by cte1.order_year,cte1.order_month)
select cte2.order_year,cte2.order_month,cte2.order_count,
round(((cte2.order_count-cte2.prev_order_count)/cte2.prev_order_count)*100) as Growth_of_orders
from cte2
;
```

RESULT

Row	order_year	order_month	order_count	Growth_of_orders
1	2016	9	4	*null*
2	2016	10	324	8000.0
3	2016	12	1	-100.0
4	2017	1	800	79900.0
5	2017	2	1780	123.0
6	2017	3	2682	51.0
7	2017	4	2404	-10.0
8	2017	5	3700	54.0
9	2017	6	3245	-12.0
10	2017	7	4026	24.0

- **Observations**:Evidently, there is a noticeable upward trend evident when transitioning from 2016 to 2017. However, the possibility of aggregating data on a yearly basis is hindered due to the limited data availability for only three months in 2016, while comprehensive monthly data is accessible for the entirety of 2017. This prompted a comparison of order quantities on a monthly basis. The dataset clearly highlights a significant surge in order counts during the initial three months of 2017. Notably, a remarkable increase in orders is observed between December and January in both years. Furthermore, an intriguing observation emerges: a negative growth rate in order quantities occurs during the month of September for both the years 2017 and 2018.

2.2 Can we see some kind of monthly seasonality in terms of the no. of orders being placed?[Query]

```
SELECT *
FROM
(SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp)AS order_year,
  EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
  COUNT(distinct order_id) AS num_orders
FROM
 `target_buisness_case.orders`
GROUP BY
  order_year, order_month
HAVING order_year=2016
ORDER BY
  order_year, order_month) tbl1
  FULL OUTER JOIN
  (SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp)AS order_year,
  EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
  COUNT(distinct order_id) AS num_orders
FROM
 `target_buisness_case.orders`
GROUP BY
  order_year, order_month

 HAVING order_year=2017
) tbl2
  using(order_month)
 FULL OUTER JOIN
  (SELECT
  EXTRACT(YEAR FROM order_purchase_timestamp)AS order_year,
  EXTRACT(MONTH FROM order_purchase_timestamp) AS order_month,
  COUNT(distinct order_id) AS num_orders
FROM
 `target_buisness_case.orders`
GROUP BY
  order_year, order_month
 HAVING order_year=2018
) tbl3
  using(order_month)

  ORDER BY
  order_month;
```

RESULT

Row	order_month	order_year	num_orders	order_year_1	num_orders_1	order_year_2	num_orders_2
1	1	null	null	2017	800	2018	7269
2	2	null	null	2017	1780	2018	6728
3	3	null	null	2017	2682	2018	7211
4	4	null	null	2017	2404	2018	6939
5	5	null	null	2017	3700	2018	6873
6	6	null	null	2017	3245	2018	6167
7	7	null	null	2017	4026	2018	6292
8	8	null	null	2017	4331	2018	6512
9	9	2016	4	2017	4285	2018	16
10	10	2016	324	2017	4631	2018	4
11	11	null	null	2017	7544	null	null
12	12	2016	1	2017	5673	null	null

- **Observations**: Comparing the data from the years 2016, 2017, and 2018 alongside each other, it becomes evident that there exists a monthly pattern within the data. As an example, considering the case of March (row 3), a positive increase in order counts is observed. Conversely, for April, a decline in the number of orders is noticeable in both 2017 and 2018. Moreover, variations in growth rates between the two years are apparent in various other months as well. A particularly noteworthy observation is the substantial drop in the number of orders during September across all years. Remarkably, this drop is particularly steep in 2018.

- **Recommendations**: The data reveals a consistent trend where demand experiences a surge during the early months of each year, specifically from January to March. However, this demand tapers off as the year progresses. Consequently, it might be advisable for the company to maintain higher inventory levels during these initial months. Conversely, a notable decline in demand is observed in September and December, indicating the need for reduced inventory during this months.

2.3 During what time of the day, do the Brazilian customers mostly place their orders? 0-6 hrs : Dawn —— 7-12 hrs : Mornings —— 13-18 hrs : Afternoon—— 19-23 hrs : Night

QUERY

```
with cte as
(select *,
case when (EXTRACT(hour FROM order_purchase_timestamp)-3)>=0
then Extract(hour FROM order_purchase_timestamp)-3
else 24+(Extract(hour FROM order_purchase_timestamp)-3)
end as Brazilian_time
from `target_buisness_case.orders`)
select count(distinct order_id) as order_count,
case when Brazilian_time between 0 and 6 then 'Dawn'
 when Brazilian_time between 7 and 12 then 'Morning'
 when Brazilian_time between 13 and 18 then 'Afternoon'
else 'Night'
end as time_of_the_day
from cte
group by time_of_the_day
order by order_count desc;
```

RESULT

Row	order_count	time_of_the_day
1	38291	Morning
2	36986	Afternoon
3	14013	Night
4	10151	Dawn

- **Observations**: Analyzing the query outcome, it becomes apparent that the peak number of orders occurs during the morning hours, specifically between 7 AM and 12 PM Brazilian time. Additionally, a substantial volume of orders is observed in the afternoon period (1 PM - 6 PM), which is reasonably expected. However, a contrasting trend emerges during the night and dawn hours, with significantly fewer orders being placed. This suggests that consumers exhibit reluctance towards making purchases during these times.

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

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

QUERY

```
select c.customer_state as State,
extract (year from o.order_purchase_timestamp) as Year,
extract(month from o.order_purchase_timestamp) as Month,

count(distinct o.order_id) as No_of_orders
from `target_buisness_case.orders` o
join
`target_buisness_case.customers` c
on o.customer_id=c.customer_id
group by c.customer_state,Year,month
order by c.customer_state,Year,month;
```

RESULT

Row	State	Year	Month	No_of_orders
1	AC	2017	1	2
2	AC	2017	2	3
3	AC	2017	3	2
4	AC	2017	4	5
5	AC	2017	5	8
6	AC	2017	6	4
7	AC	2017	7	5
8	AC	2017	8	4
9	AC	2017	9	5
10	AC	2017	10	6

- **Observations**: The results above illustrate the monthly count of orders placed for every state and year. This information could be valuable for delving into the specific order quantities per state in each month.

3.2 How are the customers distributed across all the states?

QUERY

```
with cte as
(select customer_state,count(distinct customer_id) as No_of_customers
from `target_buisness_case.customers`
group by customer_state
)
select  *, round((cte.No_of_customers/(SELECT sum(No_of_customers) from cte))*100,2) as percentage_of_customers
from cte
order by No_of_customers desc;
```

RESULT

Row	customer_state	No_of_customers	percentage_of_custo
1	SP	41746	41.98
2	RJ	12852	12.92
3	MG	11635	11.7
4	RS	5466	5.5
5	PR	5045	5.07
6	SC	3637	3.66
7	BA	3380	3.4
8	DF	2140	2.15
9	ES	2033	2.04
10	GO	2020	2.03

- **Observations**: The query outcome demonstrates that Sao Paulo contributes nearly 42 percent of our organization's customer base. This is unsurprising considering Sao Paulo's status as the most densely populated state in Brazil. Aside from Sao Paulo, Rio de Janeiro and Minas Gerais are also noteworthy for generating substantial order volumes. Moreover, Rio Grande do Sul and Paraná are among the top five states, each contributing around 5 percent of the total orders.

- **Recommendations**: Consequently, based on these findings, it is imperative to direct significant attention towards **Sao Paulo, Rio de Janeiro, and Minas Gerais**. Enhancements to customer amenities, post-sales services, and other provisions should be prioritized in these states. Additionally, the organization can strategize a marketing campaign tailored to the preferences and requirements of the inhabitants in these three states. In situations of parity, Sao Paulo should be given the highest priority.

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

4.1 Percentage increase in the cost of orders from year 2017 to 2018 (including months between Jan to Aug only).

QUERY

```
with cte2 as
(with cte1 as
(with cte as(
select extract(year from o.order_purchase_timestamp) as order_year,
extract(month from o.order_purchase_timestamp) as order_month,
round(sum(p.payment_value),2) as cost_of_order
from `target_buisness_case.payments` p
join
 `target_buisness_case.orders` o
 on p.order_id=o.order_id
 group by order_year,order_month
 having order_month between 1 and 8 #include months between Jan to Aug only
 order by order_year,order_month)
 select order_year,round(sum(cost_of_order),2) as cost_order_per_year
 from cte
 group by order_year
 order by order_year)
 select order_year,cte1.cost_order_per_year, lag(cte1.cost_order_per_year) over(order by order_year) as prev_cost
 from cte1)
 select order_year,cte2.cost_order_per_year, round(((cost_order_per_year - prev_cost)/prev_cost)*100,2) as percentage_increase_of_cost
 from cte2
 order by order_year;
```

RESULT

Row	order_year	cost_order_per_year	percentage_increase_of_cost
1	2017	3669022.12	*null*
2	2018	8694733.84	136.98

- **Observations**: The query outcome indicates that when contrasting January to August, there has been a surge of around 137 percent in the expenses associated with orders. This indicates an upward trend in people's expenditures on Target's items compared to the previous year. This phenomenon could be attributed to either an expansion in the customer base or an increase in spending per individual customer.

- **Recommendations**: Consequently, the Brazilian market for Target is displaying growth compared to the previous year. Given the assumption of consistent spending per individual and an expanding customer base, Target could consider extending its operations within this region.

4.2 Calculating the Total and Average value of order price for each state

QUERY

```
select c.customer_state as State, ROUND(sum(oi.price),2) as Total_price,
ROUND(avg(oi.price),2)as Average_price
from `target_buisness_case.order_items` oi
join
`target_buisness_case.orders` o
on o.order_id=oi.order_id
join `target_buisness_case.customers` c
on o.customer_id=c.customer_id
group by c.customer_state
order by Average_price DESC
;
```

RESULT

Row	State	Total_price	Average_price
1	PB	115268.08	191.48
2	AL	80314.81	180.89
3	AC	15982.95	173.73
4	RO	46140.64	165.97
5	PA	178947.81	165.69
6	AP	13474.3	164.32
7	PI	86914.08	160.36
8	TO	49621.74	157.53
9	RN	83034.98	156.97
10	CE	227254.71	153.76

- **Observations**: The data extracted from the query reveals that Paraibas, Acre, and Alagoas are the states exhibiting higher average prices. Conversely, Sao Paulo boasts the lowest average price across all states. An essential influencer for the states like Paraná,Rio Grande do Sul with lower average prices is their predominant presence in the densely populated South and Southeast regions. Notably, Acre stands out due to its location in the tropical Amazon region, implying potentially elevated transportation costs for these states. As discernible from the subsequent query outcome, states with the highest average prices also tend to possess elevated average freight values, contributing to the overall price increase.

- **Recommendations**: Therefore, Target should prioritize its attention on states characterized by a substantial consumer base and lower average prices. This strategy would encourage greater purchasing due to the affordability of products. Moreover, considering that freight costs are lower in southern states like Sao Paulo,Paraná,Rio Grande do Sul,Rio de Janeiro there is a favorable environment for efficient penetration and expansion.

4.3 Calculating the Total and Average value of order freight for each state

QUERY

```
select c.customer_state as State, ROUND(sum(oi.freight_value),2) as Total_freight_value,
ROUND(avg(oi.freight_value),2)as Average_freight_value
from `target_buisness_case.order_items` oi
join
`target_buisness_case.orders` o
on o.order_id=oi.order_id
join `target_buisness_case.customers` c
on o.customer_id=c.customer_id
group by c.customer_state
order by Average_freight_value desc
;
```

RESULT

Row	State	Total_freight_value	Average_freight_value
1	RR	2235.19	42.98
2	PB	25719.73	42.72
3	RO	11417.38	41.07
4	AC	3686.75	40.07
5	PI	21218.2	39.15
6	MA	31523.77	38.26
7	TO	11732.68	37.25
8	SE	14111.47	36.65
9	AL	15914.59	35.84
10	PA	38699.3	35.83

- **Asumptions**: We've sourced the geographical details of the states from the website: https://brazil-help.com/brazilian_states.htm

- **Observations**:In this context, it becomes evident that northern states such as Roraima, Paraibas, Rondônia, and Acre exhibit higher average freight costs, while states like Sao Paulo, Paraná, and Rio de Janeiro have comparatively lower average freight expenses. This factor significantly influences the overall average prices of these states.

5 Analysis based on sales, freight and delivery time

5.1 Finding the delivery time and the difference (in days) between the estimated and actual delivery date of an order.

QUERY

```
select order_id,customer_id,order_status,
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_buisness_case.orders`
where order_status='delivered'
order by diff_estimated_delivery  ,time_to_deliver desc;
```

RESULT

Row	order_id	customer_id	time_to_deliver	diff_estimated_delivery
1	0607f0efea4b566f1eb8f7d3c2...	a5fbb6579eacbeb02752a143b...	3	-146
2	c72727d29cde4cf870d569bf6...	964253ff0e4e08180064764a4...	6	-139
3	eec7f369423b033e549c02f3c...	32cef4bdd6bfa50612d81dc77...	20	-134
4	c2bb89b5c1dd978d507284be...	6357fffb5704244d552615bbfc...	16	-123
5	40dc2ba6f322a17626aac6244...	6210a37f9d6a265a4f3fbe2c21...	7	-108
6	1a695d543b7302aa9446c8d5f...	b882cbae40f34e60a3ec3efefa...	12	-83
7	39e0115911bf404857e14baa7...	816642e9995c2461f2172469e...	11	-82
8	559eea5a72341a4c82dbce988...	5fd5cb96f515996e88d84c610...	13	-77
9	c5132855100a12d63ed4e8ae0...	6e5b6ba2e8de70d9e920b541a...	12	-77
10	38930f76efb00b138f4d632e4d...	0f9043c635f86f7eb1e8fc0770...	11	-77

- **Observations**: The data retrieved from the aforementioned query provides a clear view of the delivery duration for individual products and the variance between their anticipated and actual delivery times. The outcomes reveal a spectrum of deliveries occurring notably early, as well as instances where orders experience substantial delays. These inconsistencies could likely be attributed to the geographical context of certain northern states, situated in the Amazon tropical region or hilly terrains, which contrasts with the smoother delivery process in the southern states.

5.2 Finding out the top 5 states with the highest and lowest average freight value.

QUERY

```
(select c.customer_state as State, round(avg(oi.freight_value),2) as Avg_freight_value,
row_number() over(order by avg(oi.freight_value) desc) as state_rank
from `target_buisness_case.order_items` oi
join `target_buisness_case.orders` o
on oi.order_id=o.order_id
join `target_buisness_case.customers` c
on c.customer_id=o.customer_id
group by c.customer_state
order by avg_freight_value desc
limit 5)
union distinct
(select c.customer_state as State, round(avg(oi.freight_value),2) as Avg_freight_value,
row_number() over(order by avg(oi.freight_value) desc) as state_rank
from `target_buisness_case.order_items` oi
join `target_buisness_case.orders` o
on oi.order_id=o.order_id
join `target_buisness_case.customers` c
on c.customer_id=o.customer_id
group by c.customer_state
order by avg_freight_value
limit 5)
order by state_rank;
```

RESULT

Row	State	Avg_freight_value	state_rank
1	RR	42.98	1
2	PB	42.72	2
3	RO	41.07	3
4	AC	40.07	4
5	PI	39.15	5
6	DF	21.04	23
7	RJ	20.96	24
8	MG	20.63	25
9	PR	20.53	26
10	SP	15.15	27

- **Asumptions**: We've made the assumption that we're interested in precisely the top five states with both the highest and lowest average freight values. Furthermore, the uniqueness of the average freight values aligns with our assumption.

- **Observations**:Evidently, the foremost five states with the most elevated and least average freight values stand out prominently. As previously noted, northern states such as Roraima, Paraibas, Rondônia, and Acre demonstrate elevated average freight expenses, while states like Sao Paulo, Paraná, and Rio de Janeiro showcase lowest average freight costs. This dynamic substantially shapes the overall average pricing trends within these states.

5.3 Find out the top 5 states with the highest and lowest average delivery time.

QUERY

```
(select c.customer_state as State,
ROUND(avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day)),2) as avg_delivery_time,
row_number() over(order by avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day)) desc) state_rank
from `target_buisness_case.order_items` oi
join `target_buisness_case.orders` o
on oi.order_id=o.order_id
join `target_buisness_case.customers` c
on c.customer_id=o.customer_id
group by c.customer_state
order by avg_delivery_time desc
limit 5)
union distinct
(select c.customer_state as State,
ROUND(avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day)),2) as avg_delivery_time,
row_number() over(order by avg(date_diff(order_delivered_customer_date,order_purchase_timestamp,day)) desc) state_rank
from `target_buisness_case.order_items` oi
join `target_buisness_case.orders` o
on oi.order_id=o.order_id
join `target_buisness_case.customers` c
on c.customer_id=o.customer_id
group by c.customer_state
order by avg_delivery_time
limit 5)
order by state_rank;
```

RESULT

Row	State	avg_delivery_time	state_rank
1	RR	27.83	1
2	AP	27.75	2
3	AM	25.96	3
4	AL	23.99	4
5	PA	23.3	5
6	SC	14.52	23
7	DF	12.5	24
8	MG	11.52	25
9	PR	11.48	26
10	SP	8.26	27

- **Observations**: The relationship between the five states with the shortest average delivery times and the lowest average freight costs is evident. This correlation is particularly notable as four out of these five states overlap. The states located in the southern region, namely Sao Paulo, Paraná, and Minas Gerais, exhibit notably swift average delivery times. In contrast, northern states like Roraima, Amapá, and Amazonas experience extended average delivery times due to the challenging geographical conditions they face.

5.4 Finding the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

QUERY

```
with cte as
(select c.customer_state fastest_state_for_delivery,
round(avg(date_diff(order_delivered_customer_date,order_estimated_delivery_date,day)),2) as avg_diff_in_expected_time
from `target_buisness_case.customers` c
join
`target_buisness_case.orders` o
on o.customer_id=c.customer_id
group by c.customer_state
order by avg_diff_in_expected_time
limit 5)
select *,
row_number() over(order by avg_diff_in_expected_time) as state_rank
from cte
order by cte.avg_diff_in_expected_time;
```

RESULT

Row	fastest_state_for_delivery	avg_diff_in_expected_time	state_rank
1	AC	-19.76	1
2	RO	-19.13	2
3	AP	-18.73	3
4	AM	-18.61	4
5	RR	-16.41	5

- **Observations**: It's puzzling that in the geographically challenging northern Brazilian states (Acre, Rondônia, Amazonas, Amapá), despite longer average delivery times, actual delivery is remarkably swift compared to estimates. This discrepancy can be explained by strategically setting conservative estimated delivery times to account for complex logistics, like harsh terrain or remote locations. This approach builds in a buffer for possible delays. When deliveries outpace estimates, it enhances customer satisfaction. This strategy aligns with prudent risk management and customer-centric objectives.

6 Analysis based on the payments:

6.1 Finding the month on month no. of orders placed using different payment types

QUERY

```
SELECT
  EXTRACT(YEAR FROM O.order_purchase_timestamp)AS order_year,
  EXTRACT(MONTH FROM O.order_purchase_timestamp) AS order_month,
  P.payment_type,
  COUNT(distinct O.order_id) AS num_orders
FROM `target_buisness_case.orders` O
JOIN `target_buisness_case.payments` P
ON O.order_id=P.order_id
 GROUP BY
  order_year, order_month, P.payment_type
ORDER BY
  order_year, order_month,P.payment_type;
```

RESULT

Row	order_year	order_month	payment_type	num_orders
1	2016	9	credit_card	3
2	2016	10	UPI	63
3	2016	10	credit_card	253
4	2016	10	debit_card	2
5	2016	10	voucher	11
6	2016	12	credit_card	1
7	2017	1	UPI	197
8	2017	1	credit_card	582
9	2017	1	debit_card	9
10	2017	1	voucher	33

- **Observations**:The analysis depicts the monthly variation in the number of orders made through various payment options. The data makes it clear that credit card is the most preferred choice every month across all years. Following closely is UPI, which stands out as a noteworthy payment method due to its transaction convenience. In contrast, debit card usage as a payment method appears infrequent.

- **Recommendations**:Given the regular utilization of credit cards, a strategy to encourage increased purchases could involve implementing a discount scheme for transactions reaching a specific threshold or higher. Similarly, a comparable approach can be adopted for UPI transactions by collaborating with prominent UPI platforms to offer incentives for certain spending thresholds.

6.2 Finding the no. of orders placed on the basis of the payment installments that have been paid.

QUERY

```
select payment_installments,count(order_id) as number_of_orders
from `target_buisness_case.payments`
where payment_value>0
group by payment_installments
having payment_installments>=1
order by payment_installments ;
```

RESULT

Row	payment_installment	number_of_orders
1	1	52537
2	2	12413
3	3	10461
4	4	7098
5	5	5239
6	6	3920
7	7	1626
8	8	4268
9	9	644
10	10	5328

- **Assumptions**: We've defined payment installments as orders with a payment value above zero, and for EMI purchases, the number of installments must be one or more. By focusing on payment values above zero, we've filtered out orders where at least one installment has been settled.

- **Observations**: The outcome reveals that the majority of orders are placed with a single payment installment. This preference could be attributed to the potential increase in interest payments for customers if they opt for multiple installments, motivating them to choose a single payment option.

- **Recommendations**:There appears to be a correlation between higher interest rates and payment plans with more than one installment. To enhance customer attraction, particularly since credit card usage is predominant, strategies like introducing frictionless credit options such as no-cost EMI or similar schemes could be implemented to streamline the credit process.