Time Series Analysis of Olist Marketplace

TOD212: Decision Sciences

Section 1

Prof. Neha Gadhvi

Student Details:

Name	Enrollment Number	Program
Smit Shah	AU1940291	B Tech (CSE)
Pankil Sheth	AU1940280	B Tech (CSE)
Karan Shah	AU1940292	B Tech (CSE)
Saumya Bharadwaj	AU2020	BS CS

Table of Contents

1. Acknowledgement

"It is impossible to prepare a project report without the assistance and encouragement of other people. This one is certainly no exception"

At the very outset of this report, we would like to extend our sincere & heartfelt obligation towards all the personages who helped me in this endeavor. Without their active guidance, help, cooperation & encouragement, we would not have made headway in the project.
We are extremely thankful and pay our gratitude to our faculty Prof. Neha Gadhvi for her valuable guidance and support on completion of this project presently. We extend our gratitude to Ahmedabad University for giving us this opportunity. We would also like to thank Olist for publishing this dataset.

We also acknowledge with a deep sense of reverence, our gratitude towards our parents and members of our families, who have always supported us morally as well as economically. At last, but not least, gratitude goes to all of our friends who directly or indirectly helped us to complete this project report.

Any omission in this brief acknowledgment does not mean lack of gratitude. [[i]]

2. Introduction

Olist is the largest department store in Brazilian marketplaces. Olist connects small businesses from all over Brazil to channels without hassle and with a single contract. Those merchants are able to sell their products through the Olist Store and ship them directly to the customers using Olist logistics partners. It comprises the data about 100k orders slated from 2016 to 2018. This dataset has been put up on Olist website for public exploration. The choice of language here is obviously Portuguese.

The company emphasizes the structuring of the product and focuses on service excellence for the tenants and the end consumers.

3. Data Presentation

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and lacking in certain behaviors or trends, and is likely to contain many errors. Data preprocessing is a proven method of resolving such issues. This technique is used on the data posted by users which frequently contain missing values, spelling errors and incorrect punctuation. After loading our data, we perform data cleaning on the raw data so that we get a proper dataset on which operations can be performed. There were a total 8 datasets published by Olist. We have selected 3 datasets for our project which are: olist_orders_dataset; olist_order_items_dataset; We have merged these datasets using python and cleaned the data with excel. The data we have taken is of 3 years (2016-1019).

4. Observations

We have used the Time series parameter for the analysis of the data. Since the data we have got is a time series data and we are analyzing the trends of the company over the years. We have calculated the maximum and minimum sales of the company year wise and also calculated the percentage of people paying through different modes of payment. We have also performed the regression to find out the most affecting factor on sales.

We have tried to include how the business and sales have flourished over the period of time using plotly. It is clear from the visualizations that in the start of the 2016 when business started, the sales were almost negligible.
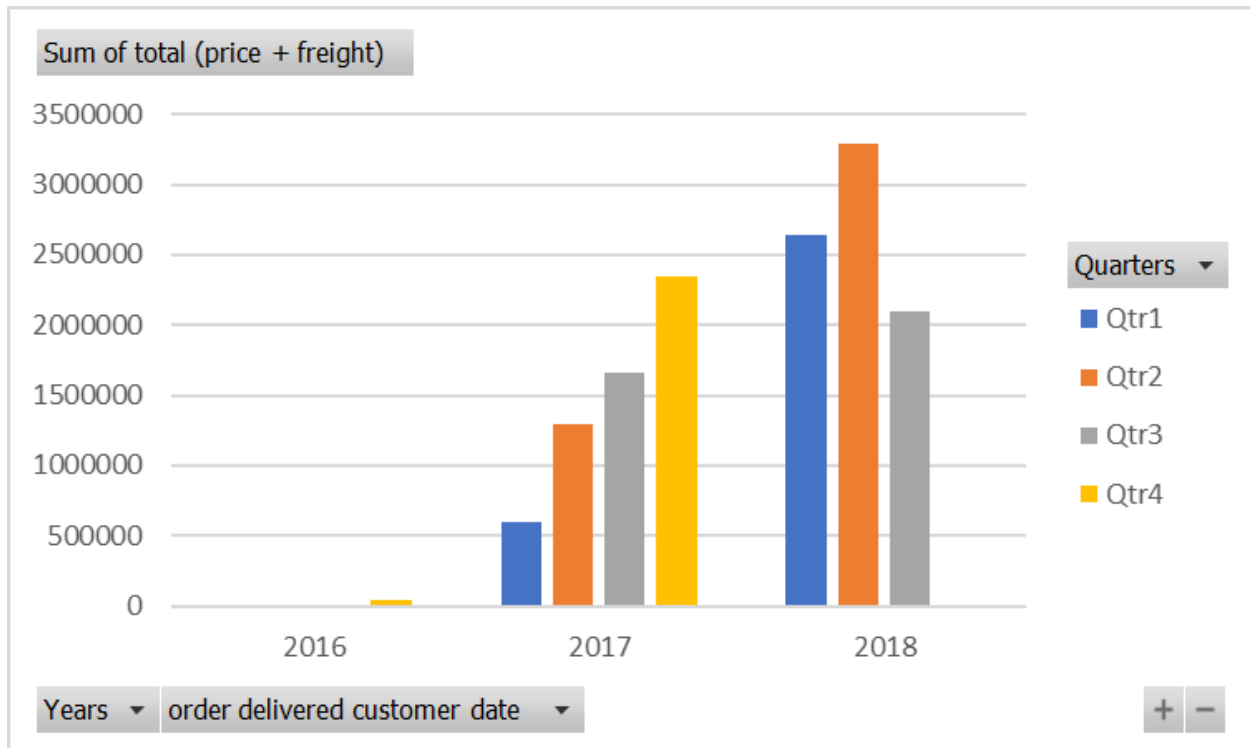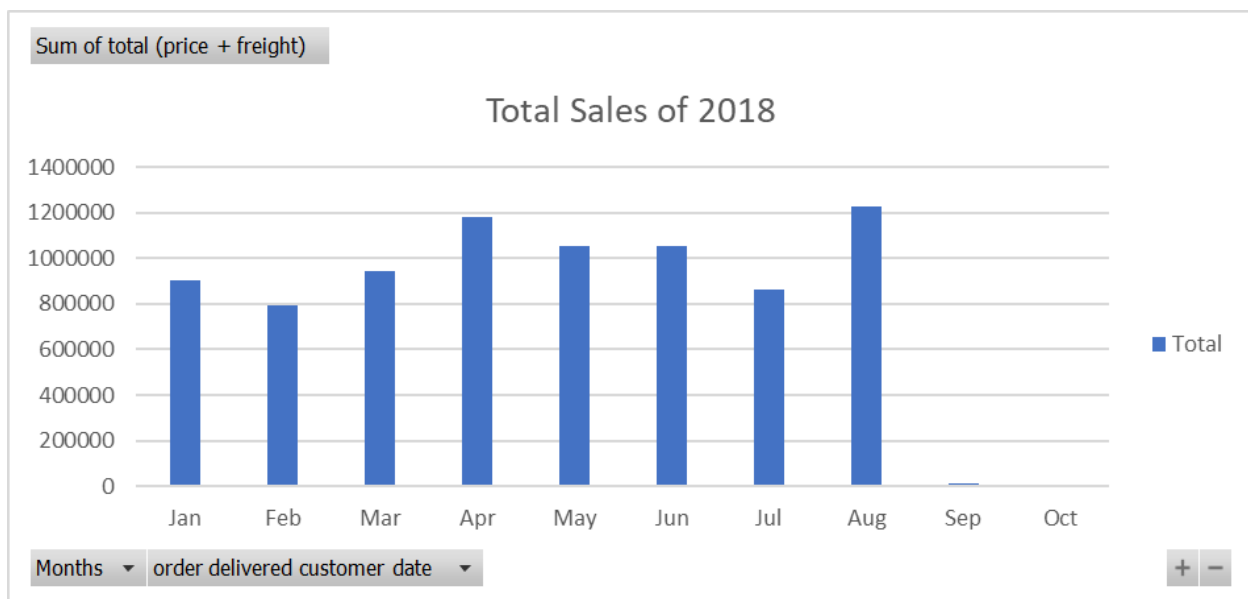
Sum of total (price + freight)	Column Labels				
Row Labels	Qtr1	Qtr2	Qtr3	Qtr4	Grand Total
2016				42064.99	42064.99
2017	594916.66	1297960.71	1662868.66	2342274.85	5898020.88
2018	2641881.45	3292066.6	2101349.57	347.95	8035645.57
Grand Total	3236798.11	4590027.31	3764218.23	2384687.79	13975731.44
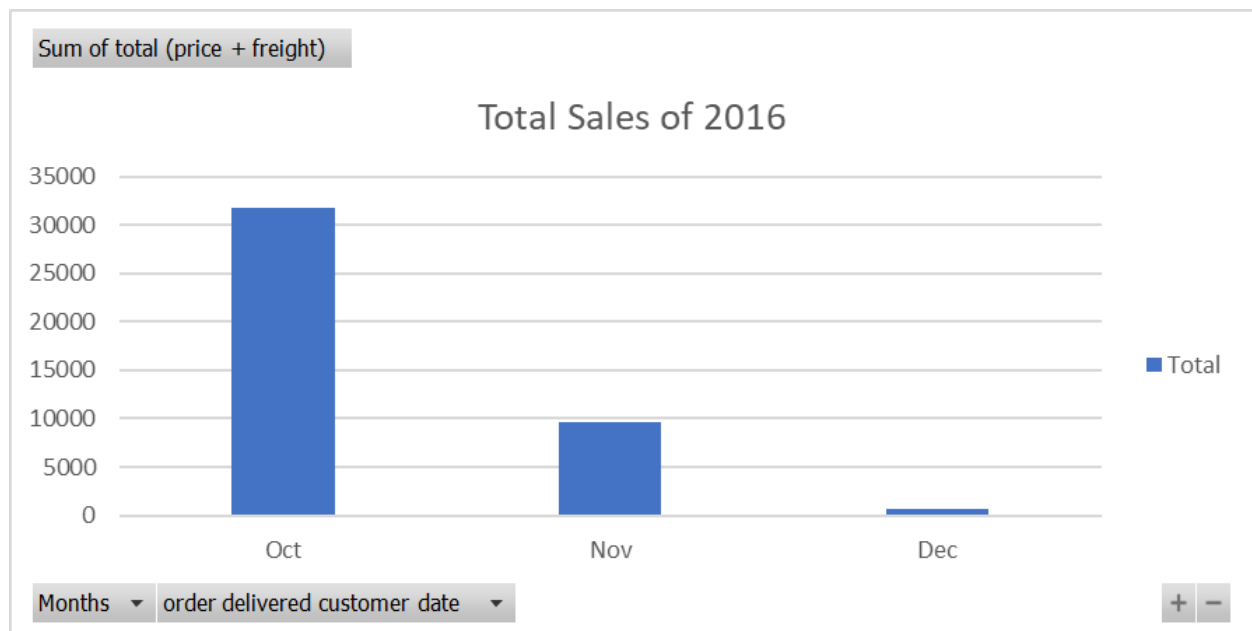
It can be seen that for the year 2016, Olist has done maximum sales in October and their least sales has been recorded in the month of December. The total amount obtained by having sales in october is close to 32,000.

Sum of total (price + freight)

Total Sales of 2017
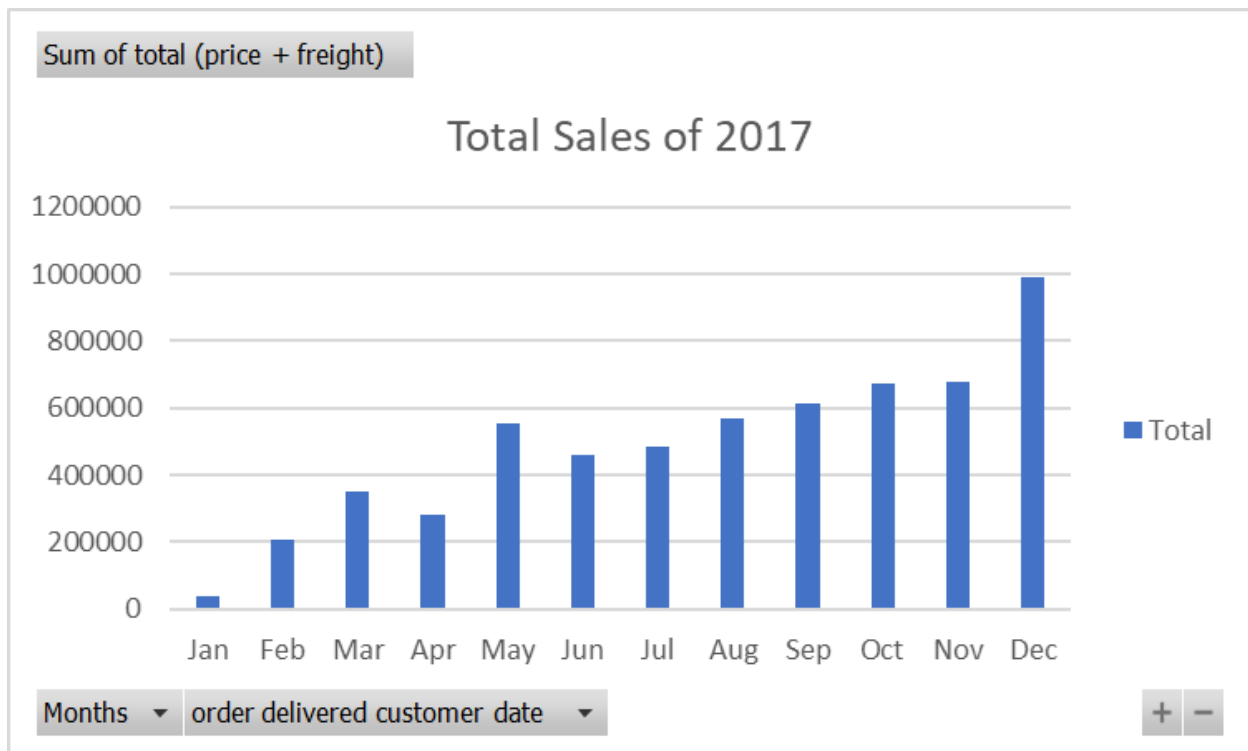
For the year 2017, December has recorded the highest sales. The sales were close to 10,00,000.

Sum of total (price + freight)

Total Sales of 2018

For 2018, August has the highest sales recorded of Olist. It was recorded about 12,00,000.

For all the 3 years we have concluded that Olist has recorded least sales on weekends. Potential reasons can be people moving out for small vacations or going to clubs/football games.

Count of payment type

Total
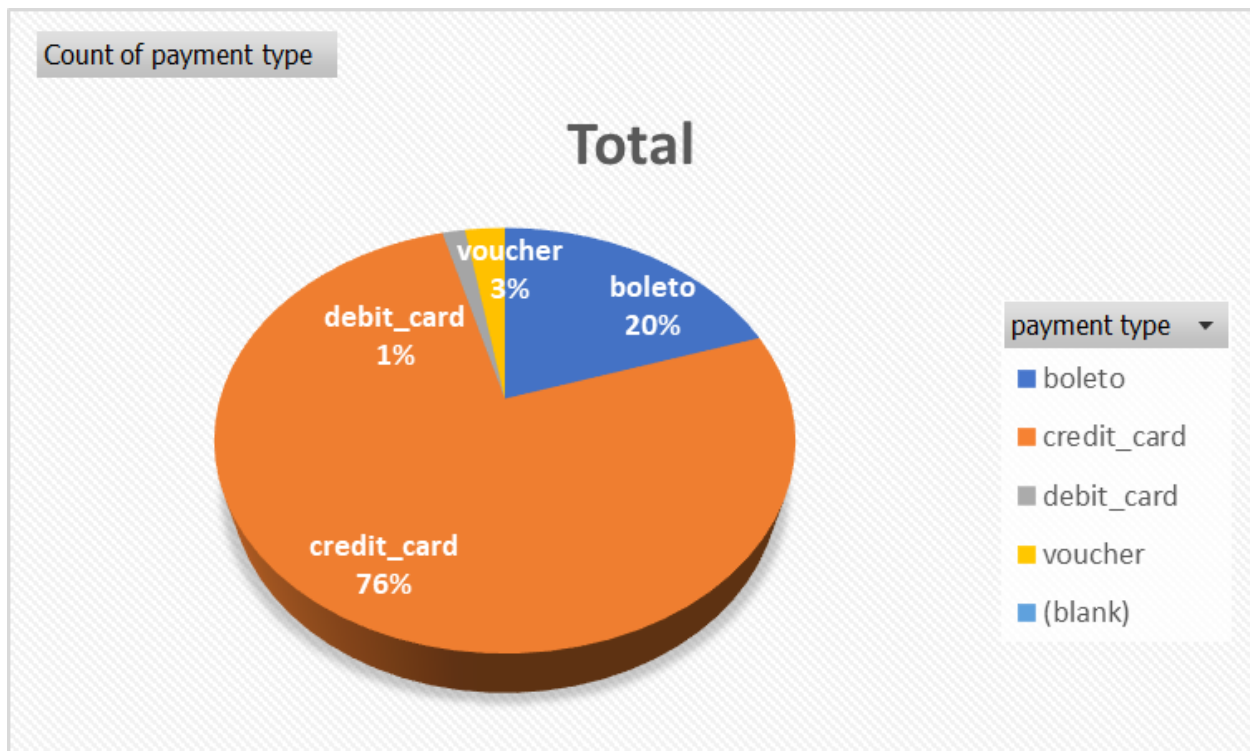
It becomes very important to analyze how the customer chooses to pay. From that information we can figure the background of the customers and choose the right target products for particular customer segments.

Olist has recorded major payments in terms of e payments i.e. through credit cards, debit cards. They may consider having full e payments in upcoming years.

Order Count in a Category vs Category Name
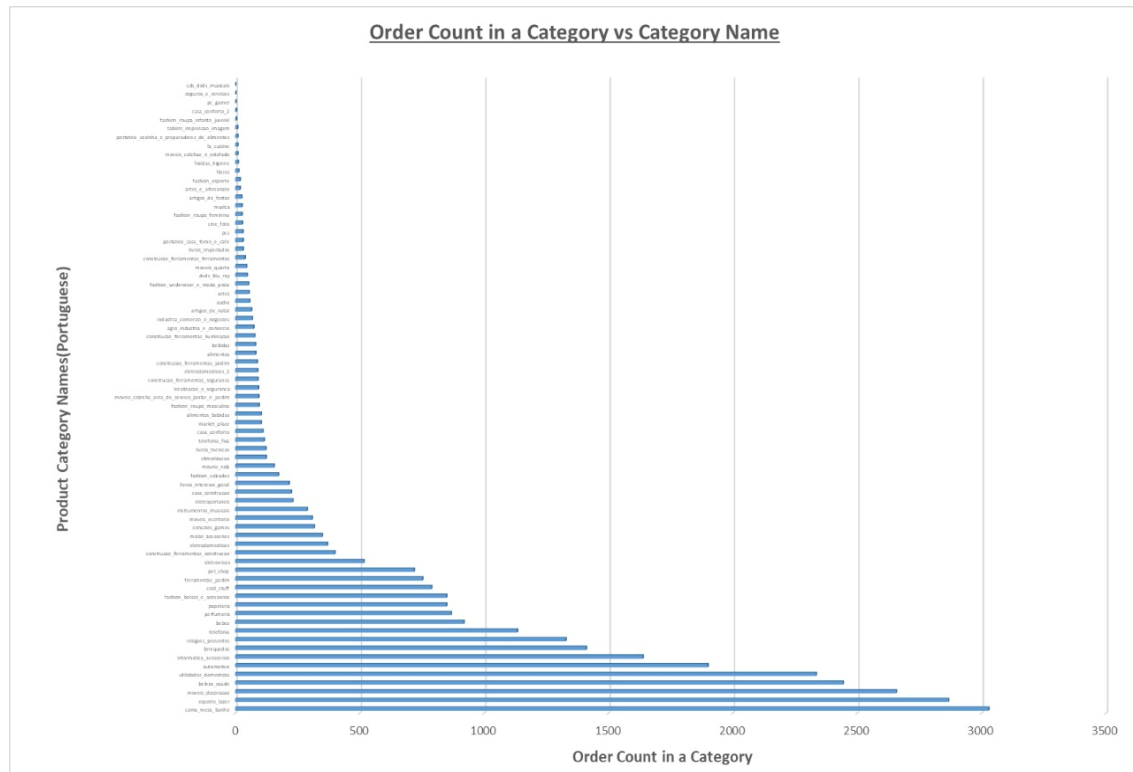
15 of the most ordered products are showcased in the table shown below :

Category (Portuguese)	Category	Products in category ordered
cama_mesa_banho	Bed Table Bath	3029
esporte_lazer	Sport Leisure	2867
moveis_decoracao	Decoration Furniture	2657
beleza_saude	Beauty Health	2444
utilidades_domesticas	Housewares	2335
automotivo	Automotive	1900
informatica_acessorios	Computer Accessories	1639
brinquedos	Toys	1411
relogios_presentes	Gifts Watches	1329
telefonia	Telephones	1134
bebes	Baby products	919
perfumaria	Perfumes	868
papelaria	Stationary	849
fashion_bolsas_e_acessorios	Fashion Bags and Accessories	849
cool_stuff	Cool Stuff	789

| product_id	count	product_category_name
aca2eb7d00ea1a7b8ebd4e68314663af	527	moveis_decoracao
99a4788cb24856965c36a24e339b6058	488	cama_mesa_banho
422879e10f46682990de24d770e7f83d	484	ferramentas_jardim
389d119b48cf3043d311335e499d9c6b	392	ferramentas_jardim
368c6c730842d78016ad823897a372db	388	ferramentas_jardim
53759a2ecddad2bb87a079a1f1519f73	373	ferramentas_jardim
d1c427060a0f73f6b889a5c7c61f2ac4	343	informatica_acessorios
53b36df67ebb7c41585e8d54d6772e08	323	relogios_presentes
154e7e31ebfa092203795c972e5804a6	281	beleza_saude
3dd2a17168ec895c781a9191c1e95ad7	274	informatica_acessorios
2b4609f8948be18874494203496bc318	260	beleza_saude
7c1bd920dbdf22470b68bde975dd3ccf	231	beleza_saude
a62e25e09e05e6faf31d90c6ec1aa3d1	226	relogios_presentes
5a848e4ab52fd5445cdc07aab1c40e48	197	
bb50f2e236e5eea0100680137654686c	195	beleza_saude

5. Inferences & Suggestions

Selling Products:

- Max products ordered are for category bed table and bath, followed by sports leisure.
- Top selling product categories are bed-table-bath, sports leisure, decoration furniture, beauty health, and housewares
- Least selling product categories include perfumes, stationary, fashion bags and accessories, and cool stuff.
- Highest selling products include decoration furniture, bed table and bath, and garden tools. While least selling products include that of beauty health.

- This should indicate the company to stop diversifying their product range and focus on the top and mid-range selling products.

Payment Modes:

- Credit card is the most used method of payment done by customers followed by boleto (Brazilian currency) . However, the company may look forward to accepting digital payments or payments via net-banking more as compared to boleto as the payment value for most top selling products such as furniture is quite high, and it would be beneficial for them to have all of the data related to payments on a more streamlined front.

	price	freight value	payment installment	payment value	review score
price	1				
freight value	0.419363336	1			
payment installment	0.312589498	0.208103954	1		
payment value	0.90796159	0.452951212	0.323162231	1	
review score	-0.005467956	-0.035057668	-0.030336203	-0.041378937	1

- It can also be inferred from the above table that payment value is obviously highly dependent upon the price of the product, followed by the freight value, and payment installment. Although, more focus should also be placed on the customer review scores as higher scores will fetch higher sales in the long run for Olist.

Sum of total (price + freight)
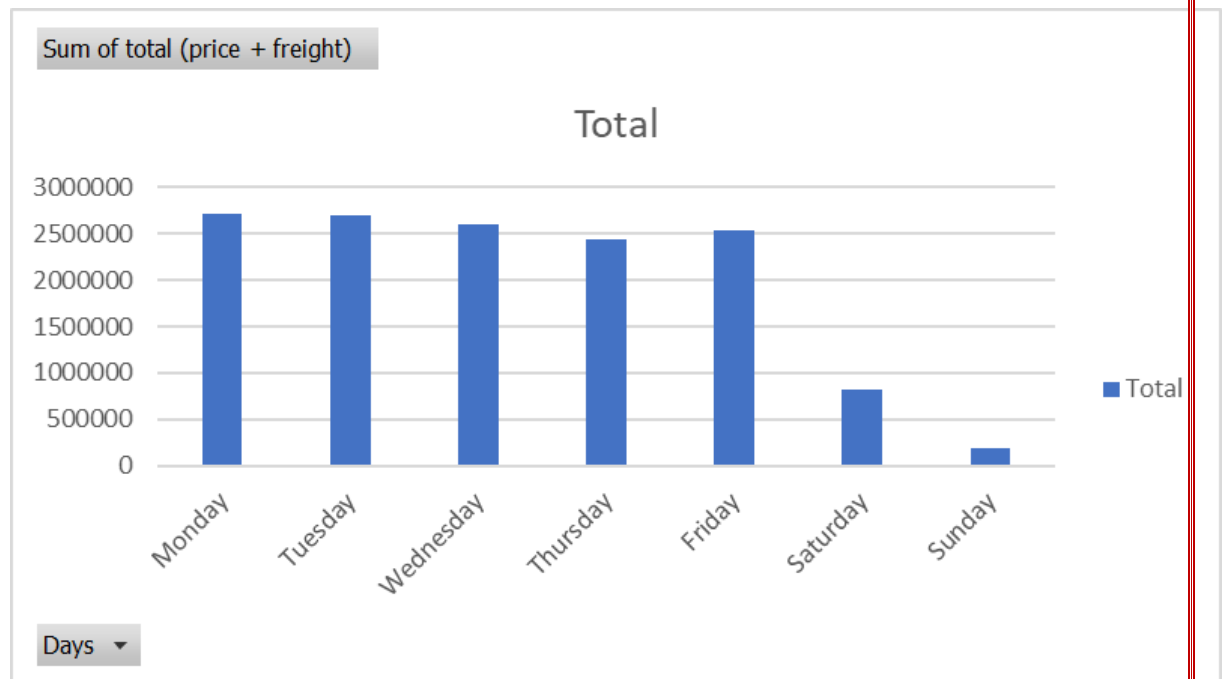
- Olist receives most payments from orders on weekdays as compared to weekends. A probable reason might be the freight/delivery staff is also on break over the weekend and the product is temporarily stored in a warehouse.

6. Conclusion

While Olist has generated a substantial amount of revenue via sales since its inception, it is clear that the company could grow and flourish further given that they follow more data-backed advice.
Moreover, recently JPMorgan's 2020 e-commerce payment trend has found that the e-commerce market in Brazil has been growing at double digit rates since 2017, and is predicted to grow in excess of 9% annually through 2023, reflecting economic pressure caused by COVID-19 (Moeser, 2021). Olist is an e-commerce aggregator company that has a good market capture in Brazil. Due to the current pandemic situation the company has made good progress in revenue generation however this is modern day news the data for which was not made available to us on Kaggle.

The analysis also focuses on the product categories with respect to the price and the review score of the customer. So, if the company focuses on the product's pricing strategy and makes product improvements on the basis of consumer opinion, the company can make better revenue and can increase the range of products under each category. Albeit the organization has caught the worldwide market, through a wide scope of items it can catch the immaculate district also. The market and its clients choose the organization's development and its maintainability later on. Consequently, the organization needs to purposely address the purchaser's survey and continue to adjust and change the item range. Future work in this space might incorporate recognizing how connections revealed in this investigation might be applied to further develop income inside comparative internet business organizations in Brazil.

7. Bibliography

References

Brazilian E-Commerce Public Dataset by Olist. (2022). Retrieved 20 April 2022, from
https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce

Hall, C. (2021, December 15). Brazil's Olist gets its horn with new $186M funding round.
TechCrunch. Retrieved April 20, 2022, from

https://techcrunch.com/2021/12/15/brazils-olist-gets-its-horn-with-new-186m-funding-round/

8. Working Part

We as a group have been very robust and active for this project. We all have contributed significantly to this project. We have been part of all the different verticals of the project visually coding part, excel part, report part and the presentation part.

Major responsibilities:
Smit: Coding, Excel
Karan: Excel, Report
Pankil: Excel, Report
Saumya: PPT, Report