

26th August, 2021

E-commerce Sales: Prediction and Recommendation System

By Sahil Chitnis(ssc9983), Phoebe Zhou(yz6729)



CONTENT

1. Introduction
2. Problem Statement (ie Need Analysis)
3. Solution Briefly (ie Concept Defination)
4. Dataset Introduction
5. Technology Stack
6. Detailed Solution
7. Shortcomings
8. Conclusion
9. Gratitude
10. Questions



INTRODUCTION

- The growth and evolution in consumer demand, amalgamated with technological innovations, continue's to drive growth in global e-commerce sales.
- According to Statista, the number of people buying goods and services online is around 2.14 billion in 2021, sky rocketing up from 1.66 billion global buyers in 2016.
- Industry is forecasted to 2 folds in size within the next 2-3 years
- Key drivers of success over the next 10 years will be those who are focused on building a deep understanding of and improved relationship to the empowered consumer



Problem Statement (ie Need Analysis)

- As of 2021, almost all the commodities such as clothes, electronics, everyday use groceries as well as services such as house cleaning, plumbing services etc are sold online on E-commerce websites.
- E-commerce websites too have taken up the task to use existing transaction Big Data (sometimes in size of GB's) and run Big Data technologies on them to analyze the data quickly and make well informed decisions based on sales predictions.
- Recommendation systems have started gaining momentum for a more personalized and enriched experience.



Solution (in brief)

- Month-wise Sales analysis
- Hour-wise Daily Sales analysis
- Map based (ie Geospatial) analysis of sales distribution
- Spending Trend
- Product Review based analysis
- Predicting Sales for next FY
- Customer Segmentation
- Recommendation Engine

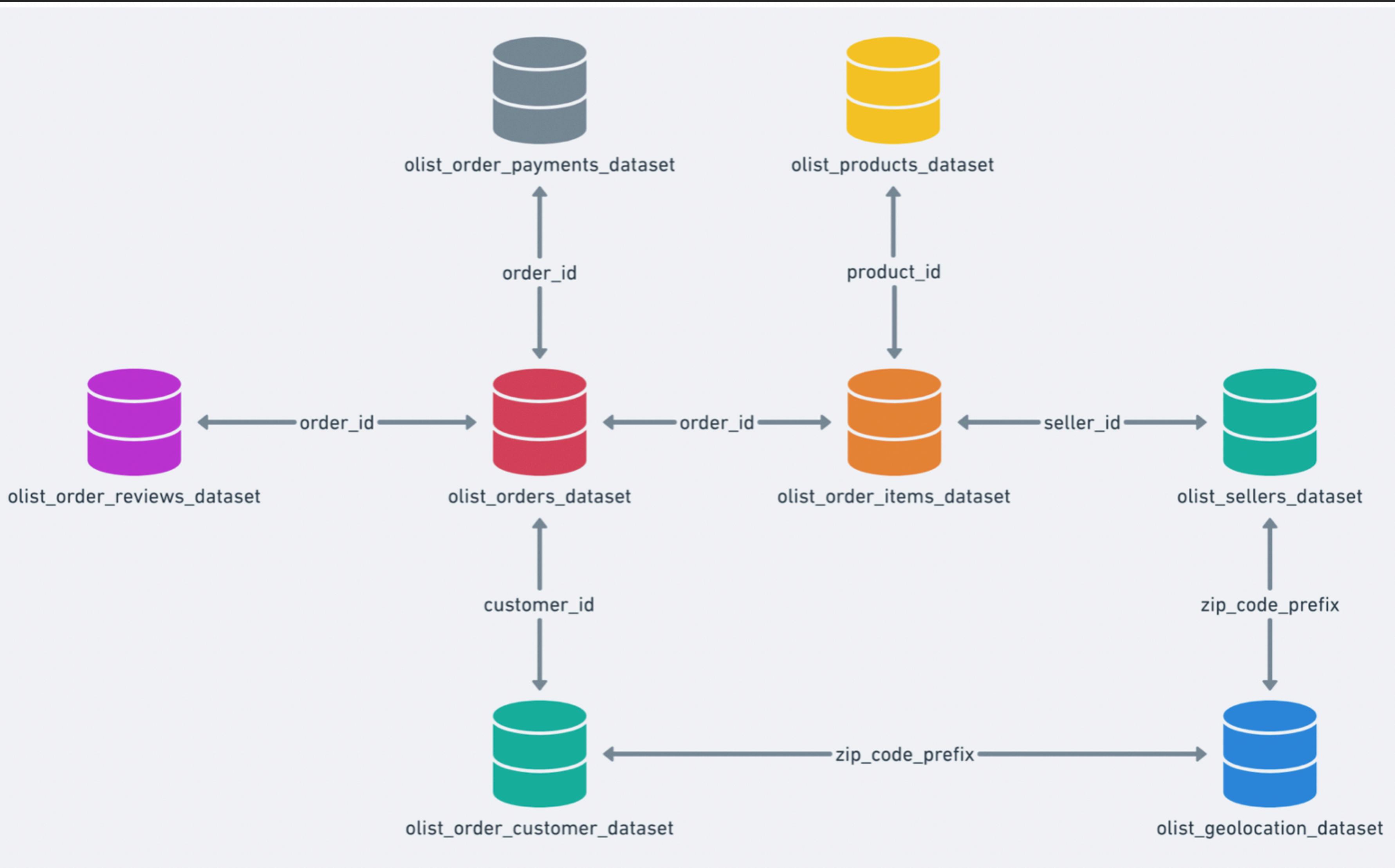


Dataset Introduction

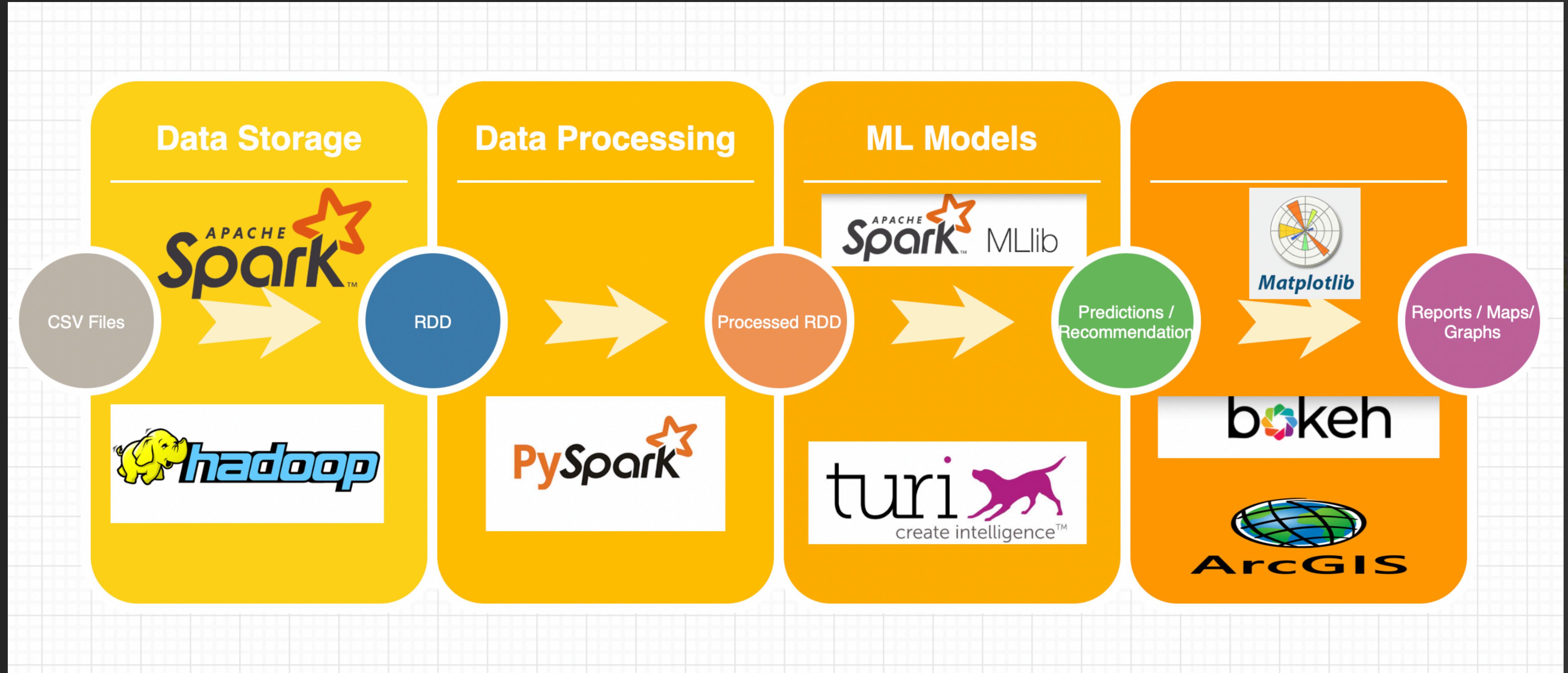
- We have used the Brazilian e-commerce public dataset of orders made at **Olist** Store. This dataset can be found on Kaggle here : <https://www.kaggle.com/olistbr/brazilian-ecommerce>.
- **The dataset has information of 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil.**
- **Olist** is a Brazilian departmental store (marketplace) that operates in e-commerce segment



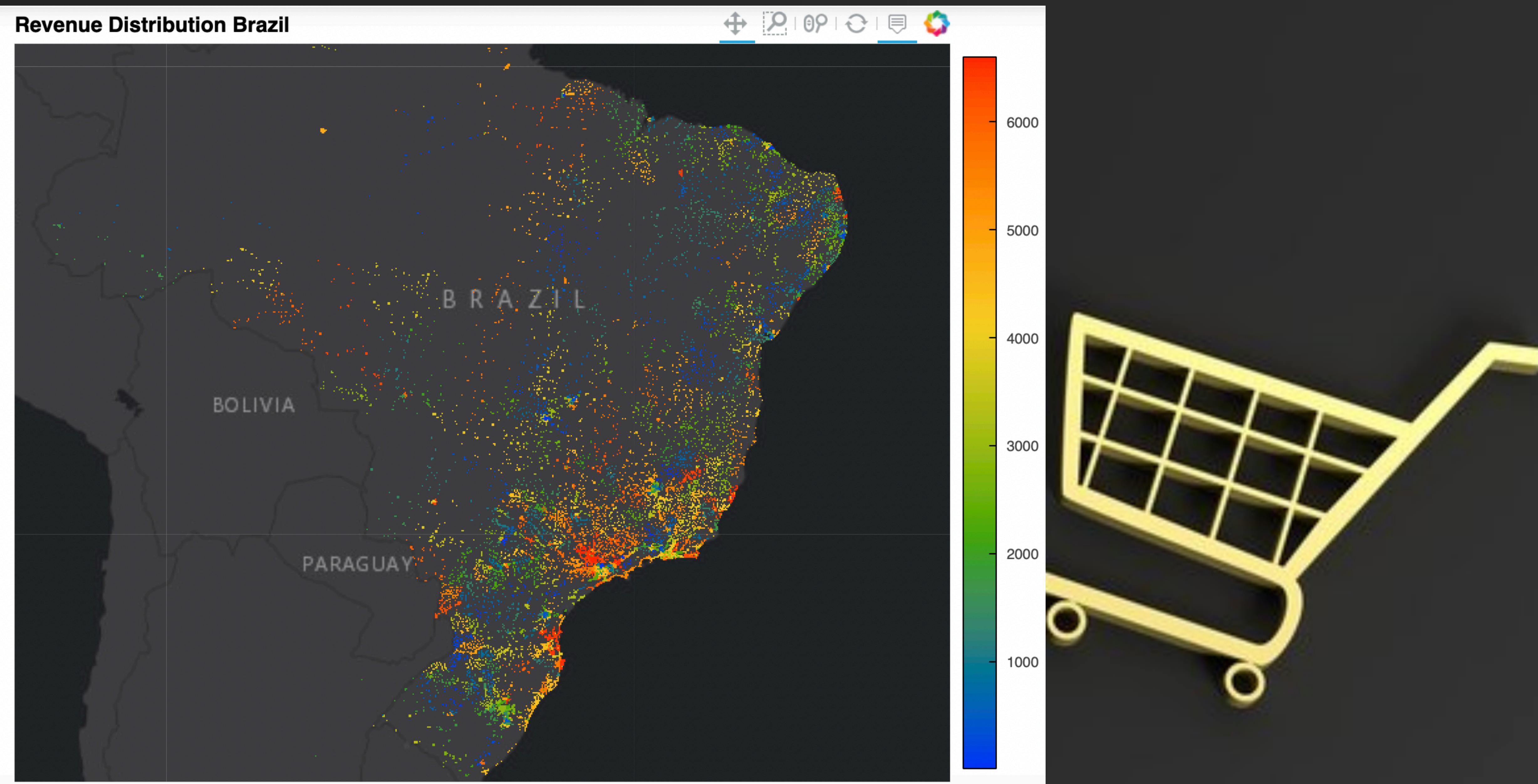
Dataset Introduction



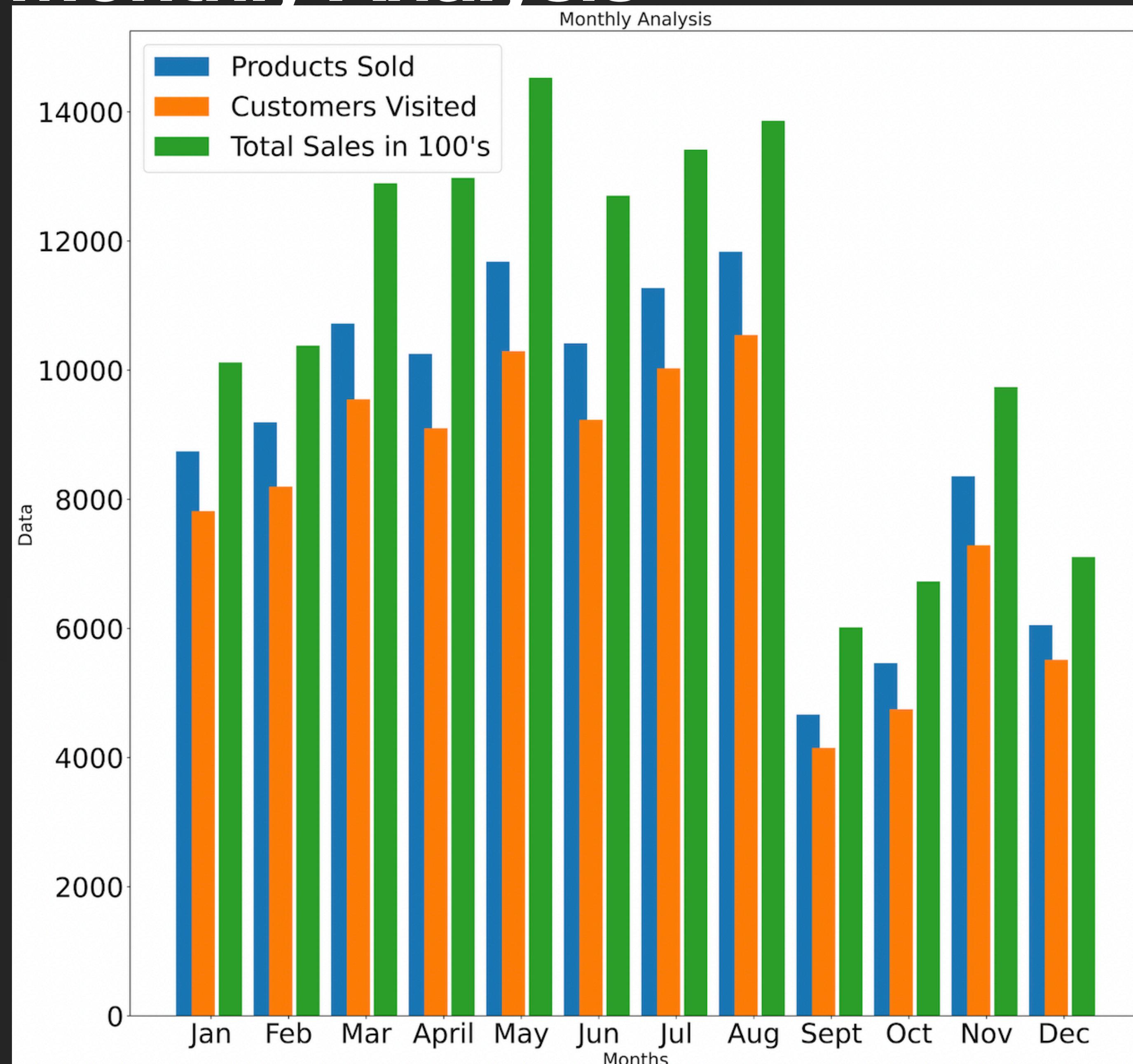
Technology Stack



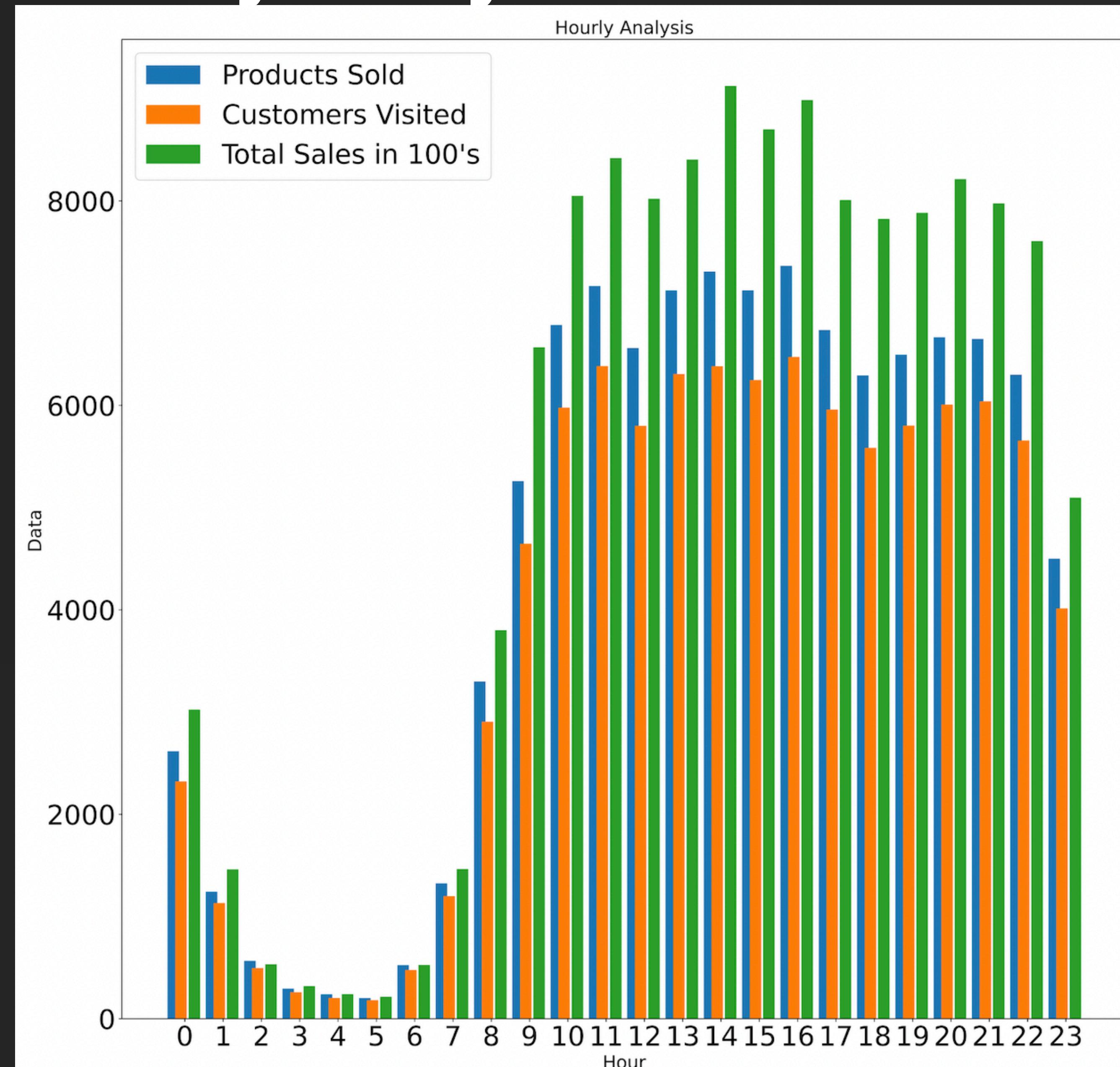
Geospatial Analysis - Revenue



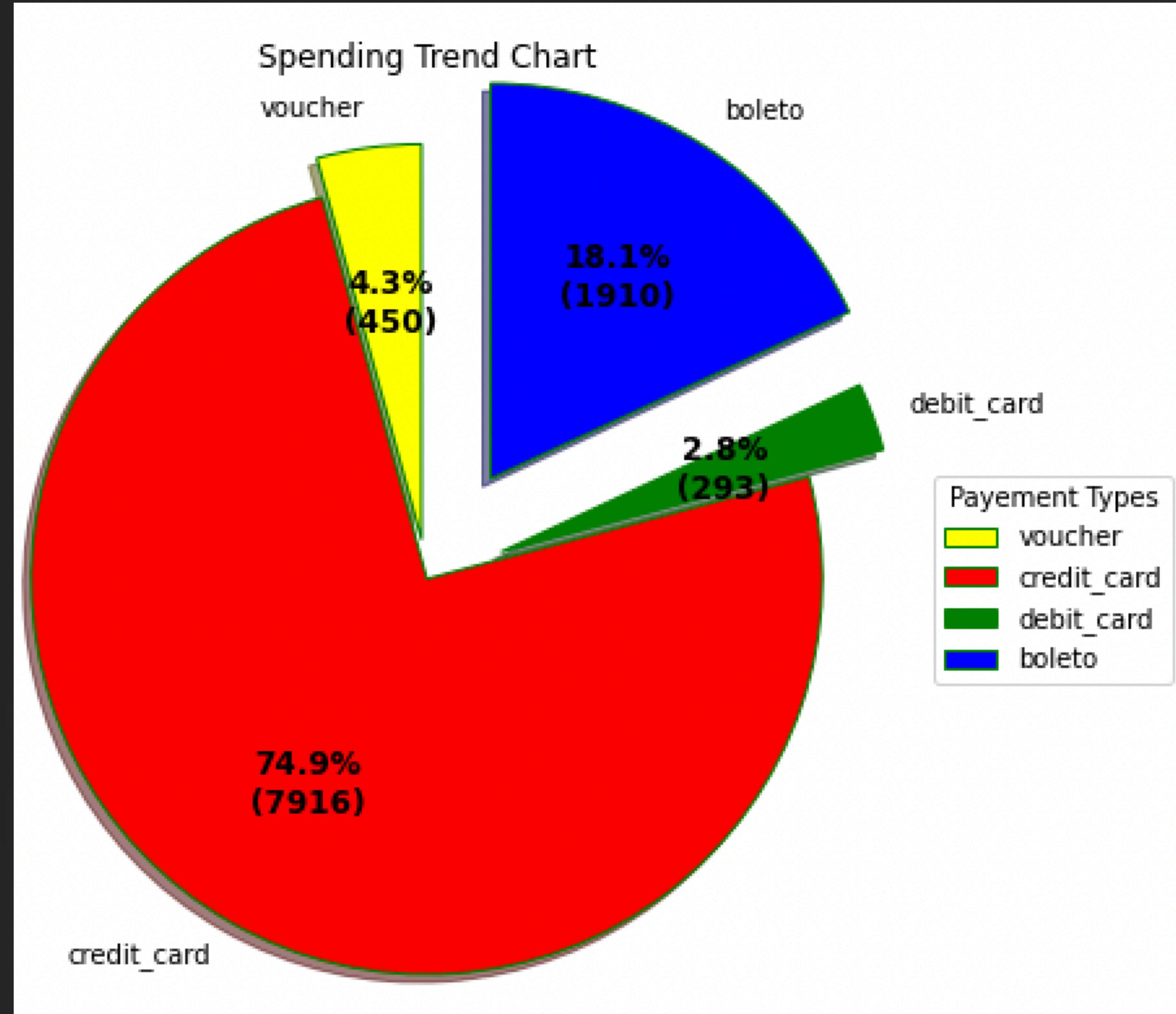
Monthly Analysis



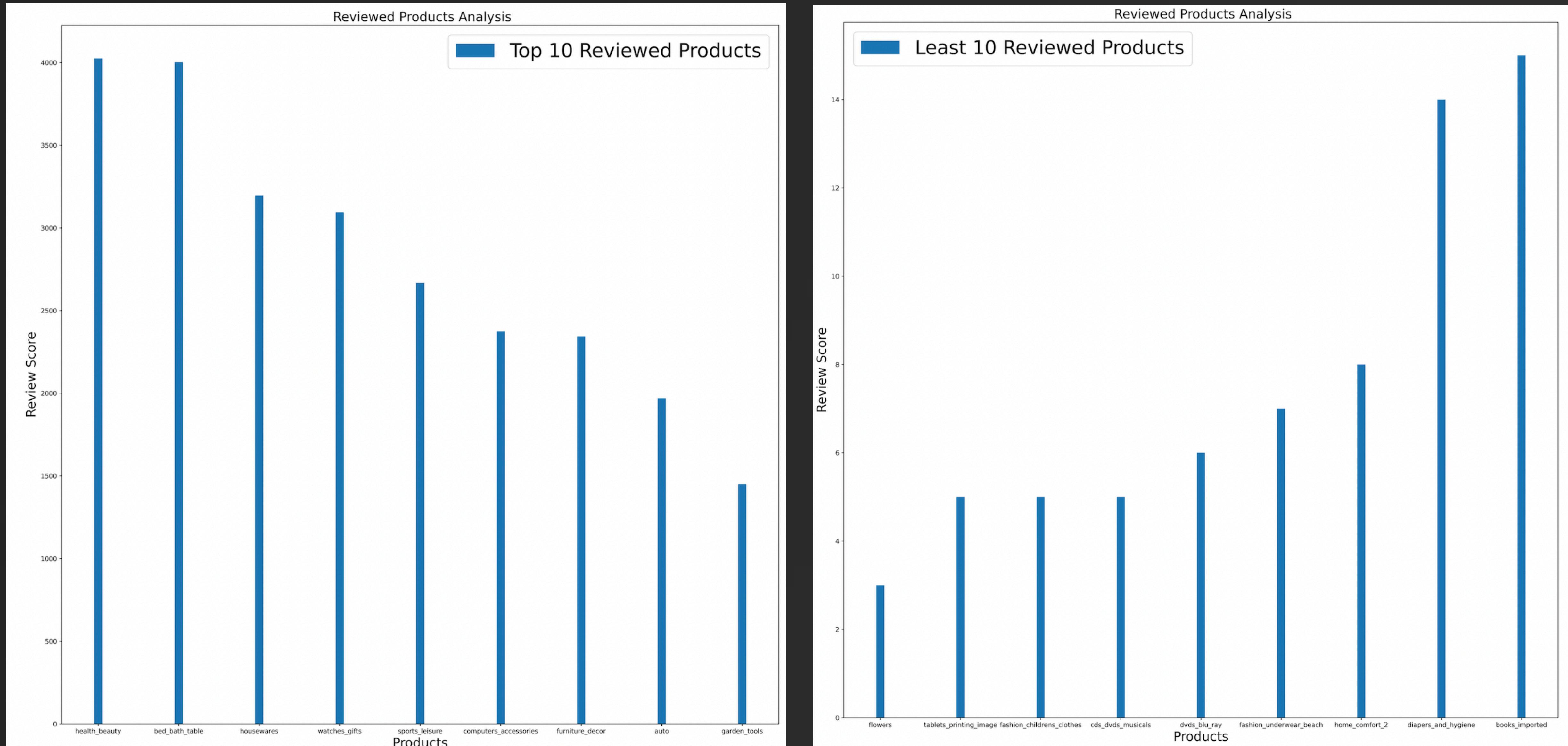
Hourly Analysis



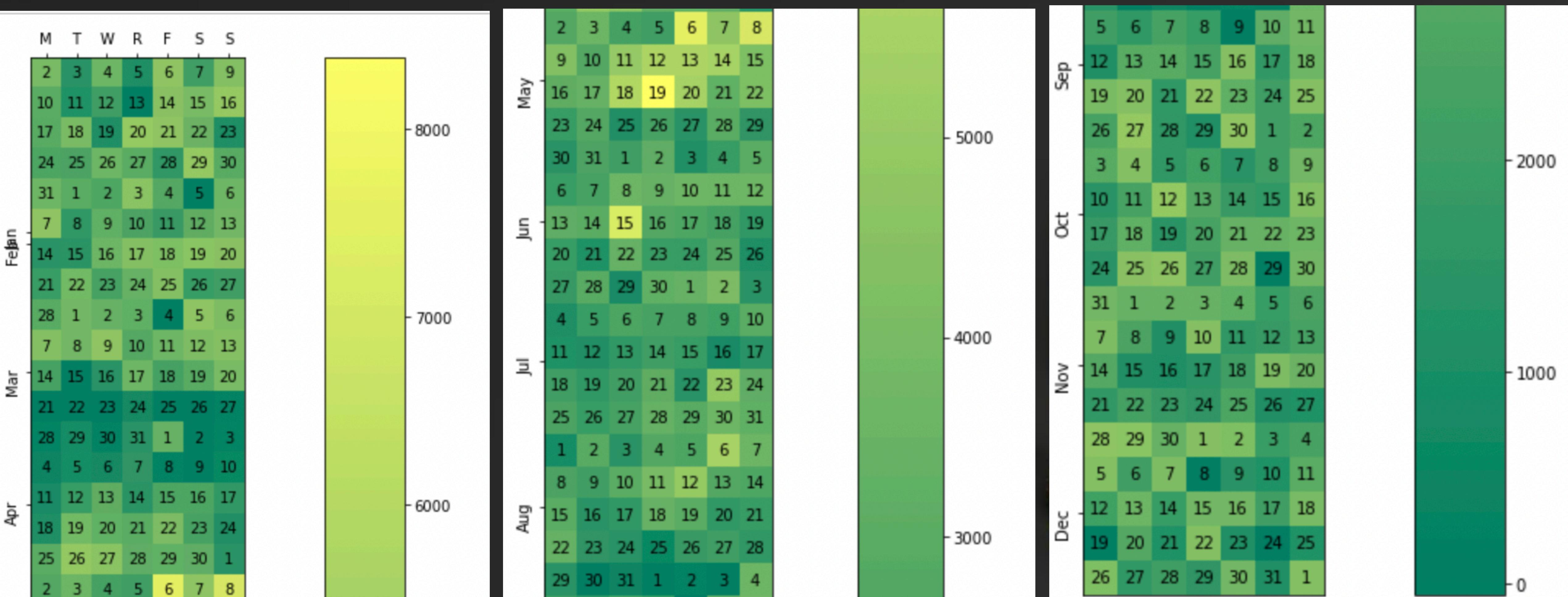
Spending Trend



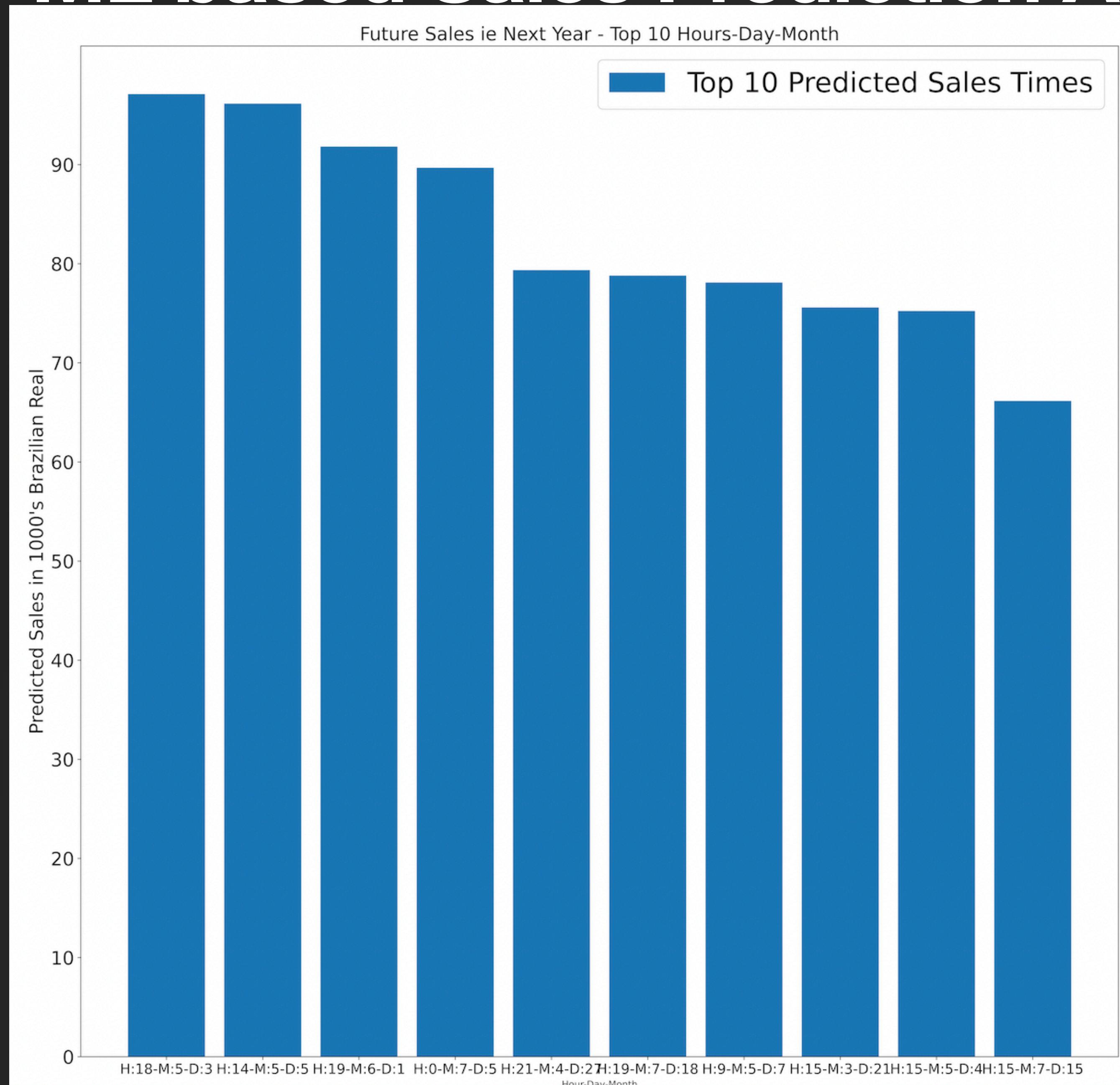
Product Review Analysis



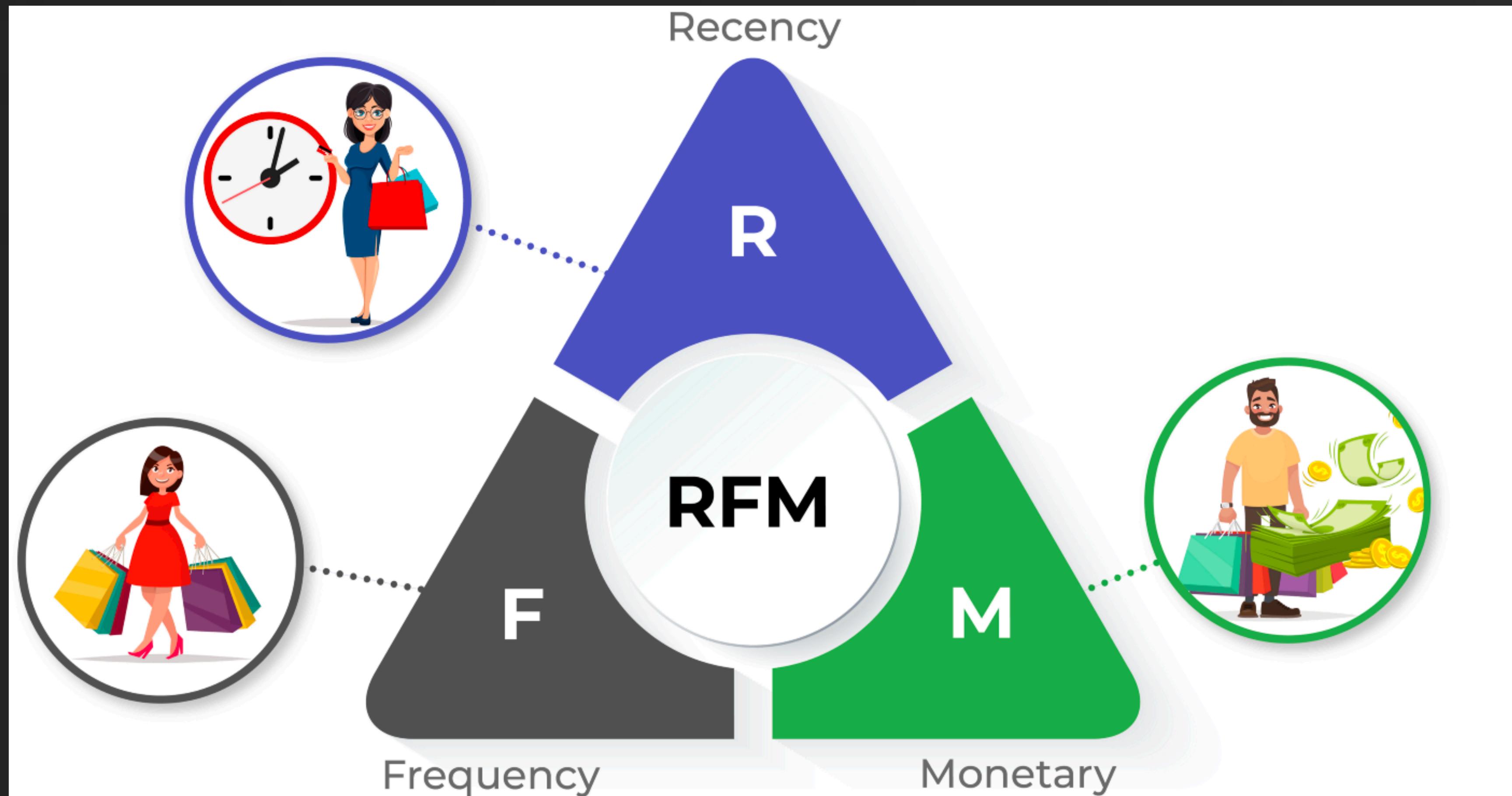
ML based Sales Prediction Analysis (Yearly Calendar)



ML based Sales Prediction Analysis



Customer Segmentation



RFM Analysis

- Assigned a score from 1 to 4 to each customer, where 1 denotes the lowest score while 4 is the highest score, based on ***statistical quartiles***.
- Grouped customers to low value customers, mid value customers and high value customers.

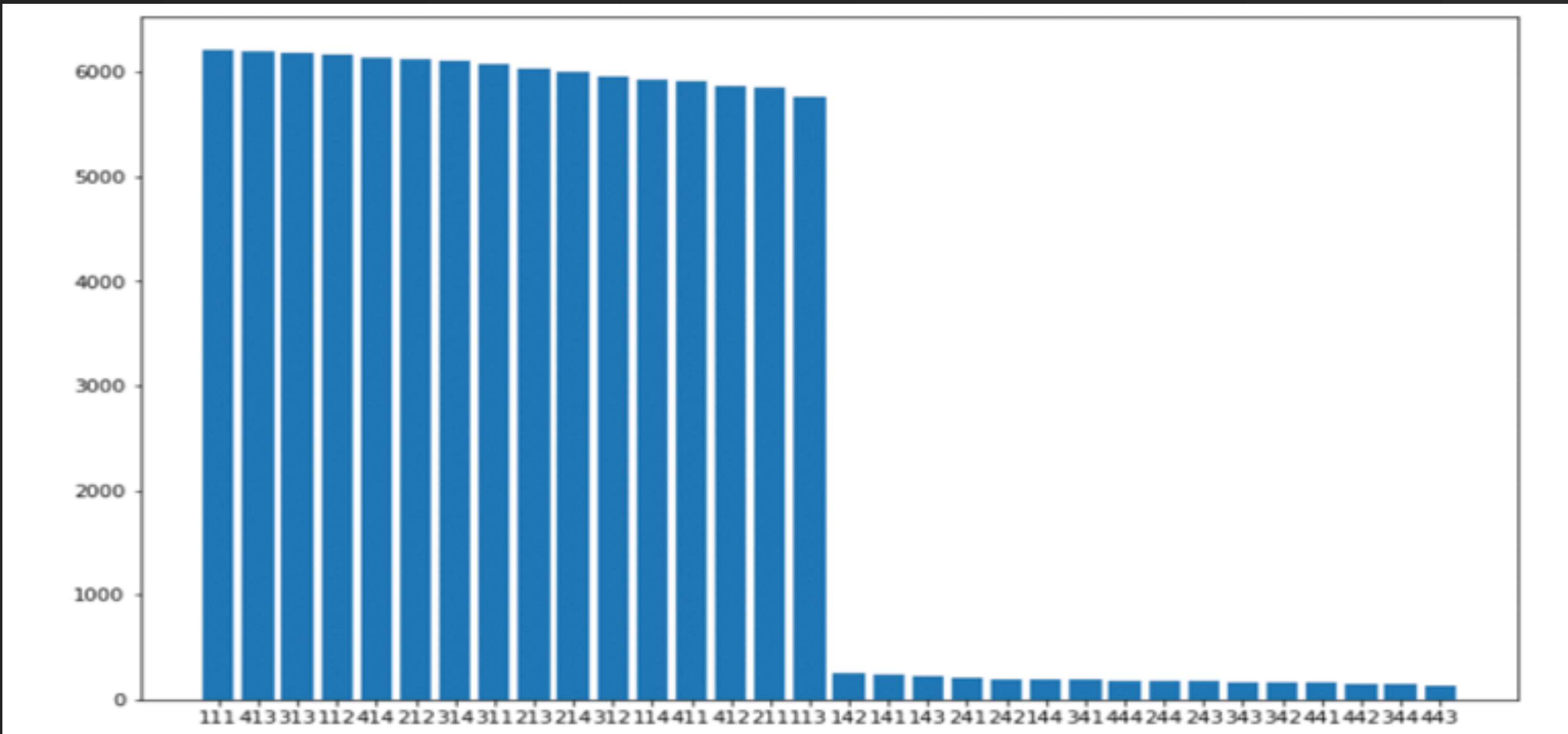
customer_id	Recency	Frequency	Monetary	R_Quartile	F_Quartile	M_Quartile	RFM_Score
4f55f44c9af248a6c...	589	1	396.97	1	1	4	114
e2b3cf60eee891a56...	67	1	171.75	4	1	3	413
d550ac7fe9688fbc8...	324	1	228.6	2	1	4	214
e154e499a4edf9f98...	508	1	105.28	1	1	2	112
6bafcfc687acf7e1c...	292	1	57.78	2	1	1	211
18f6ca10777417c93...	272	1	61.36	2	1	1	211
860ac166573be76ff...	396	1	111.02	2	1	3	213
316fc927e9216ff78...	466	1	117.94	1	1	3	113
e3c7e245a96d7fa33...	367	1	415.97	2	1	4	214
4ba1bde676ff918af...	165	1	117.26	4	1	3	413

Marketing Strategies

For Low Value Customers(111):

- Created Customer Profiles
- Marketing Penetration
- Competitive Analysis: contribution margin and profit points

High Spending New Customers(413): target customers



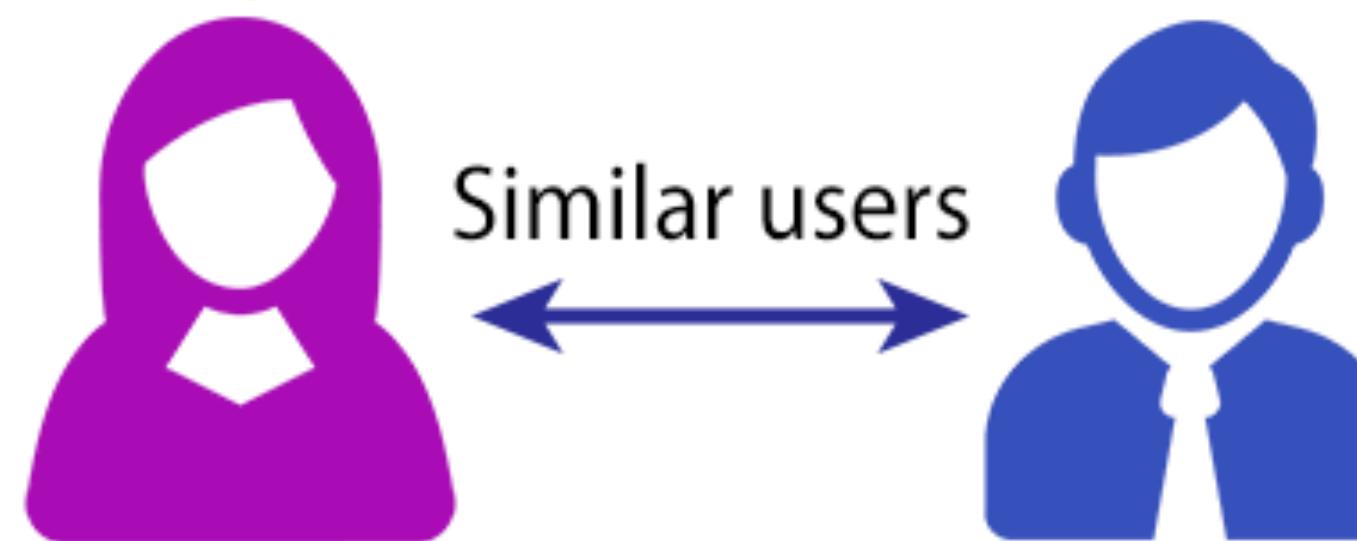
Recommendation Engine

COLLABORATIVE FILTERING

Read by both users



Similar users



Read by her,
but not him

CONTENT-BASED FILTERING

Read by user



Similar articles



Recommended
to user



Recommendation Engine

- Model Pipeline: StringIndexer, ALS
 - Evaluation : RMSE
 - Results

Shortcomings

- Linear regression: with 54% R²
- Use better visualizations tools such as TABLEAU
- Recommendation System: with 1.04 RMSE
- Customer Segmentation: Competitive Analysis



Conclusion

- Pricing Strategies
- Marketing Strategies
- Sales Strategies



Gratitude

- We want to extend sincere thanks to Professor Juan Rodriguez and TA Vishal Vanam for their teachings and guidance.



Thank You



Questions ???

