DCS X Peak AI

Olist

Team Four: Sandbox Analysis

TABLE OF CONTENTS

01 **Our Team**
An overview of our team & backgrounds

02 **Exploring Problem**
Outline of defining an approach with the given information

03 **Hypothesis**
After planning, defining a hypothesis to work with

04 **Analysis**
Preparing the datasets, and analyzing the data

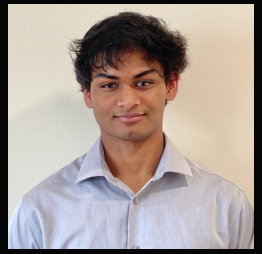
05 **Modeling**
Based on observations, modeling certain trends

06 **Conclusion**
Drawing conclusions from analyses & models

Our Team

Oscar Wan

NYU CAS, class of ‘24
“Adapt and overcome.”



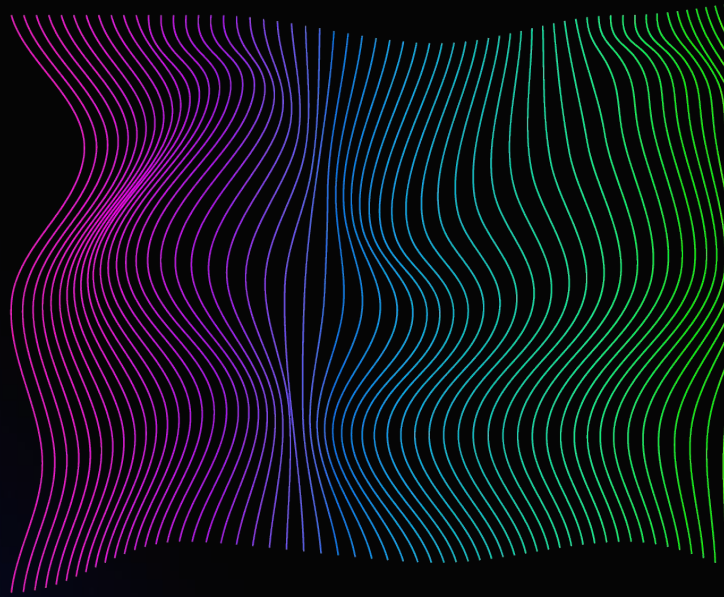
Varun Aravapalli

NYU Tandon, CS, class of ‘24
“I want to use and manipulate data to learn about the world” 🌎

Nikita Pola

NYU Tandon, CS, class of ‘24
“The power of data lies in its ability to analyze the past, and make decisions for the future” 💫

Exploring the Problem



The task at hand

Project

- Sandbox Challenge

Impact

- Understanding the customer preferences

Analysis

- Conducted on orders, products and geolocation information

Planning

1st step

Look into datasets & entity relationship diagrams

2nd step

Define a problem

3rd step

Feature selection/engineering and model building

4th step

Model analysis and recommendations

Support trends with models

01

About the project

PROJECT GOALS

GOAL

How can OList Improve Customer Satisfaction?

FACTOR

“Influence User Reviews the most”

DATA

What factors are most important to customer satisfaction?

Manipulating the Data

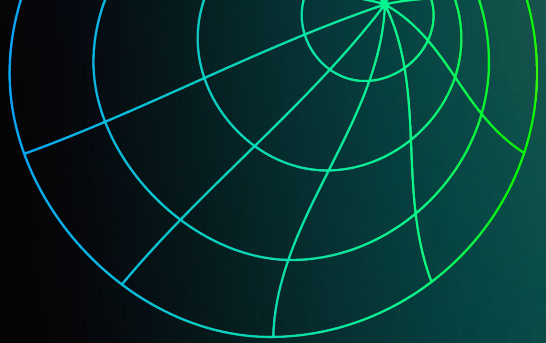
Order_ID

Product_ID

Location_Data

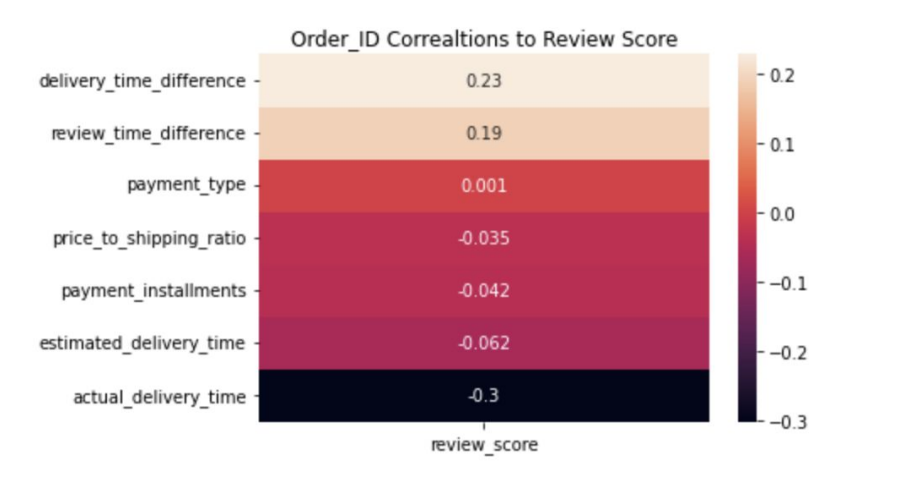
Steps:

1. Merging Datasets
2. Removing Null Values
3. Engineering New Features



Correlation Heatmaps

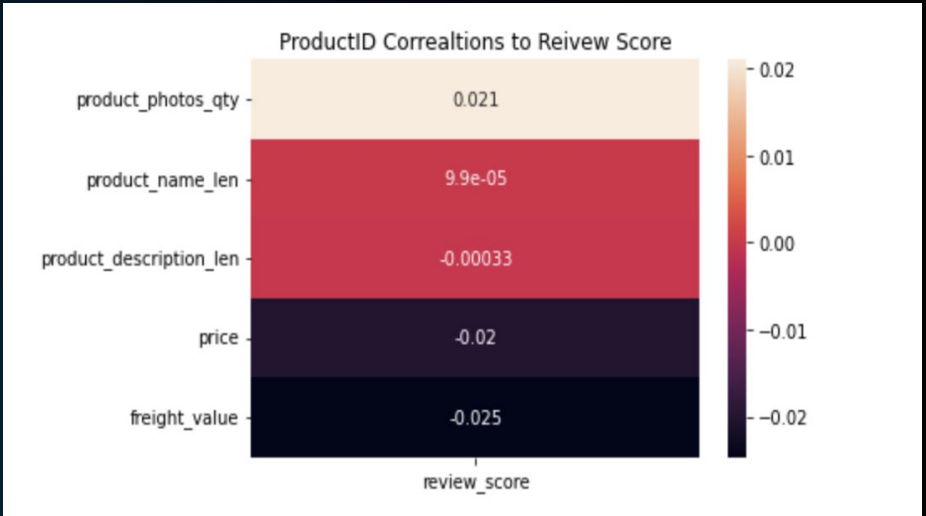
Order Attributes to Review Score



Summary

- Little positive or negative correlation between review scores and location features

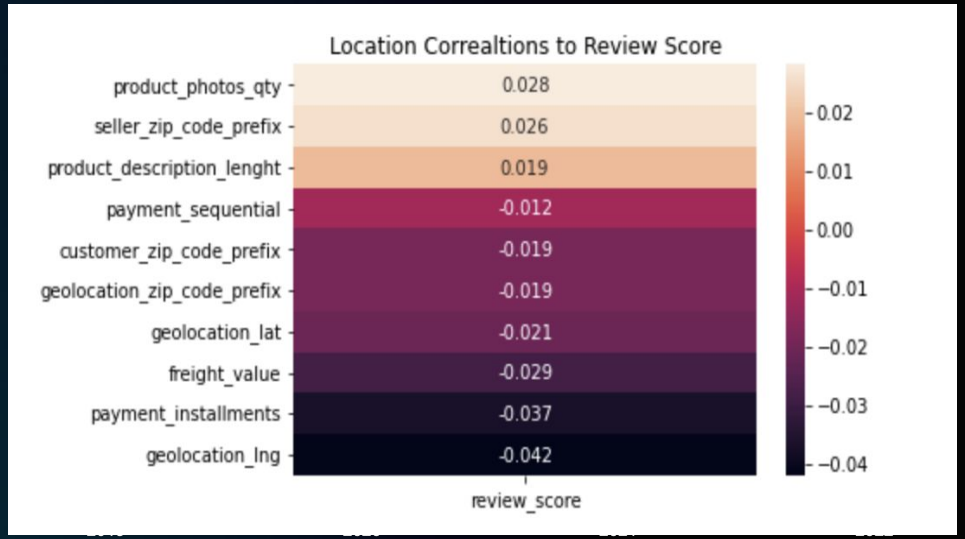
Product Attributes to Review Score



Summary

Little positive or negative correlation between review score and product attributes

Location Attributes to Review Score



Summary

Little positive or negative correlation between review scores and location features

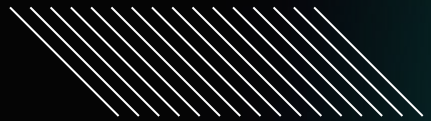
Logistic Regression

Target:

- Whether a customer is satisfied with the order
- If review rating < 4, then no (unsatisfied). Else, yes (satisfied).

Features:

- Whether the order is delivered late compared to the estimated delivery time (is_late)
- Shipping fee to item price ratio (shipping_cost)
- Amount of time it took for Olist to ask for review on an order (review_time)
- Number of photos of the product ordered (num_photos)



Reasoning on features selected

- **Whether order is delivered late:** *if there are delays in delivery, customer frustration would be greater*
- **Review time difference:** if the survey is sent shortly after delivery, users would more likely respond accurately (recency bias)
- ***Product Photos Quantity:*** *more information available (more photos) to the user, leads to more transparency in the transaction*
- ***Price to Shipping Ratio:*** *customers might be unhappy with a high shipping cost if they buy something cheaper*

Logistic regression

- Is_late is the only strong factor
- The coefficient of -1.786 means that an order being late makes a customer about 6 time **LESS** likely to be satisfied with the order than if the order were on time!
- Olist should try to keep the orders delivered on time!

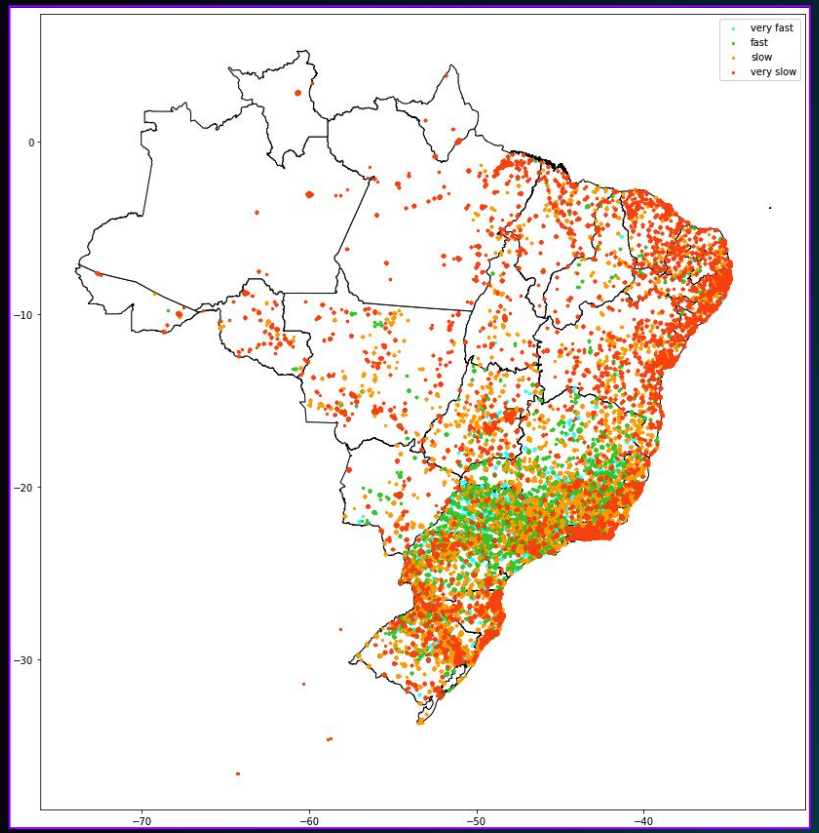
Feature	Coefficient
is_late	-1.786
shipping_cost	-0.298
review_time	0.014
num_photo	0.047

Delivery time across zip code

- Some statistics of average delivery time across zip codes
- On average, it takes almost 2 weeks to deliver an order to a customer (some packages took over 100 days, but they were shipped to outside of brazil)!

	actual_delivery_time
count	14869.000000
mean	13.101967
std	7.796964
min	1.000000
25%	8.000000
50%	11.666667
75%	16.000000
max	194.000000

Delivery Time across Zip Codes (Cont.)



- Red and orange indicate very slow (over 16 days), and slow (over 11 days) delivery
- Delivery is very slow in most of the locations
- We recommend Olist to add more shipping centers in these areas with high demand!

Recommendations

Potential considerations based on findings and analysis

POSSIBLE LOCATIONS

Natal

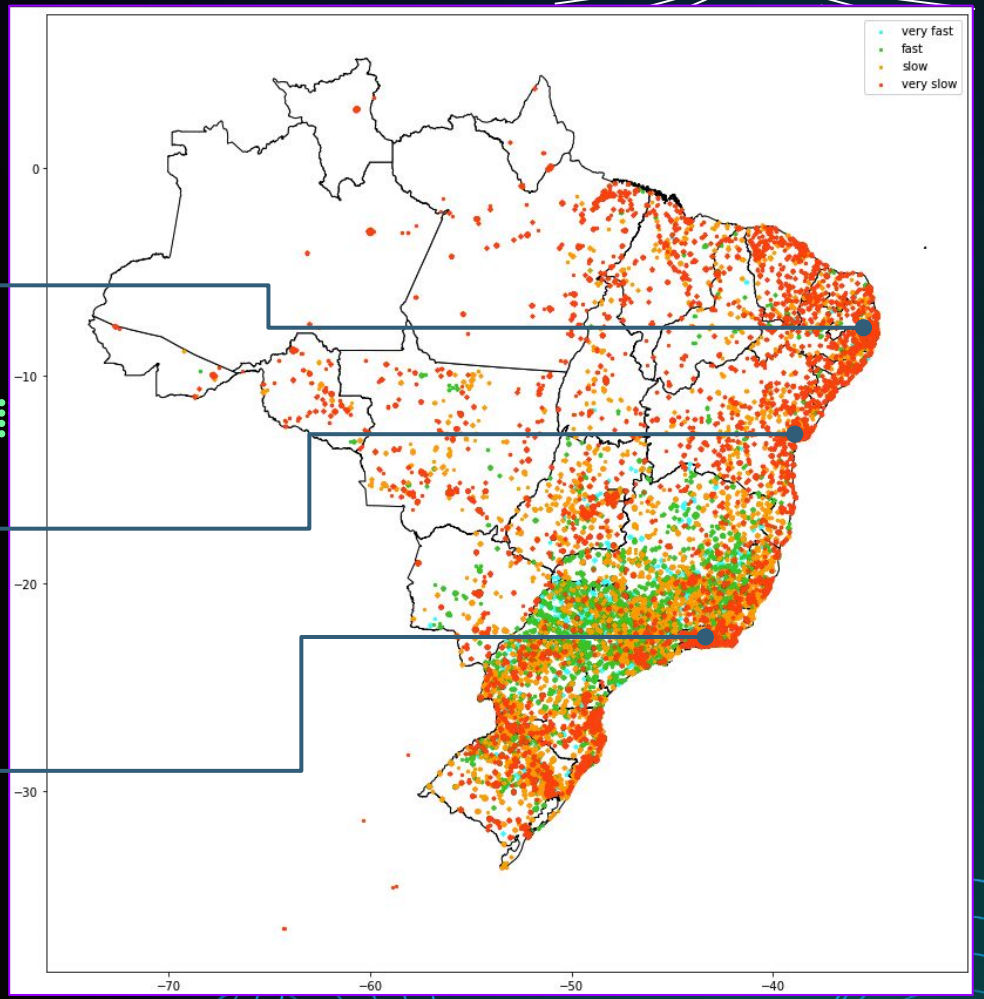
Capital city of Rio Grande

Salvador

Capital city of Bahia

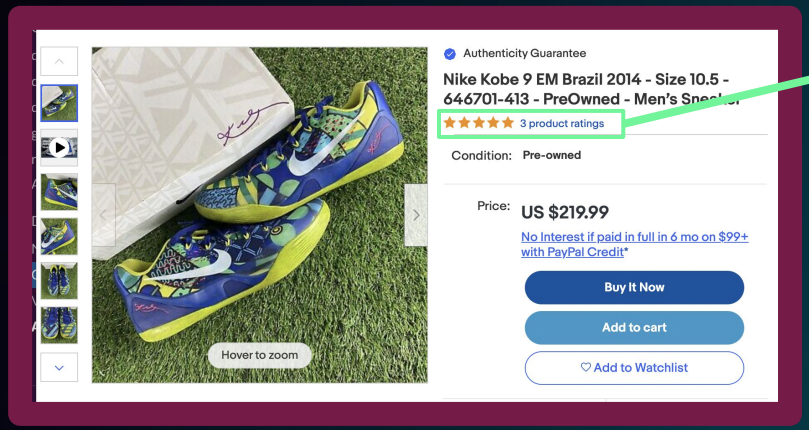
Rio de Janeiro

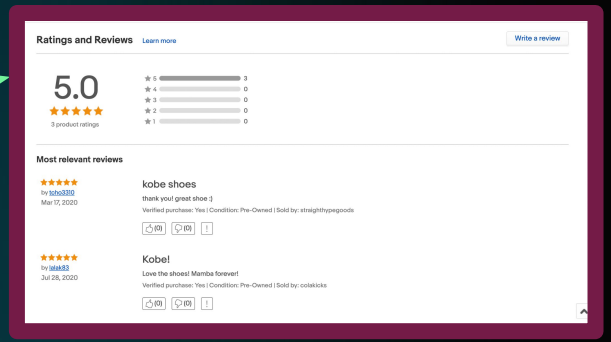
Capital city of Brazil



Alternative Recommendations

- Add features for product reviews and seller reviews
  - Currently there is only order reviews (without mandated entries)
- Obtain customer-specific data to find trends among customers and lead to customer-driven solutions based on targeted customer segments





Thanks for listening!