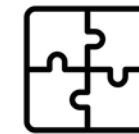
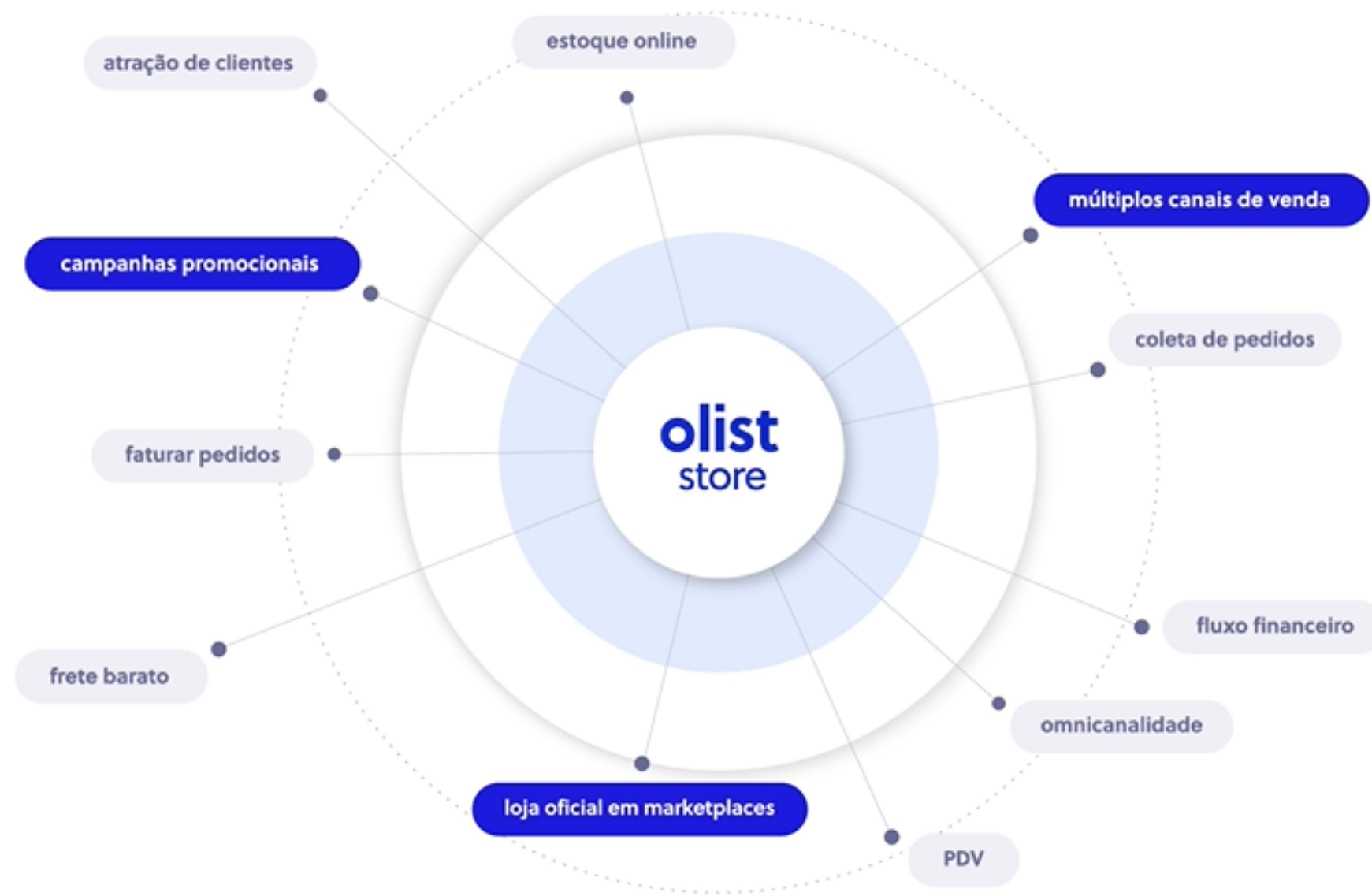


# Analyzing Factors Affecting Review Scores in Brazilian E-commerce Marketplaces: A Big Data Project

# Context



Integrated services: client attraction and orders management.



Connects businesses to different online sales channels and logistic solutions

# Data analyzed



**Customers Dataset:** information about the customer and its location. Each order is assigned to a unique `customer_id`. This means that the same customer will get different ids for different orders. The purpose of having a `customer_id` on the dataset is to allow you to identify customers that made repurchases at the store. Otherwise, you would find that each order had a different customer associated with it.

**Order Items Dataset:** data about the items purchased within each order.

**Payments Dataset:** data about the orders payment options.

**Order Reviews Dataset:** data about the reviews made by the customers

**Order Dataset:** the core dataset. From each order you might find all other information.

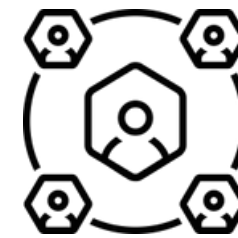
**Products Dataset:** data about the products sold by Olist.



“what factors impact  
the review score of  
the purchases?”

---

# Why reviews are so important?



## Social influence

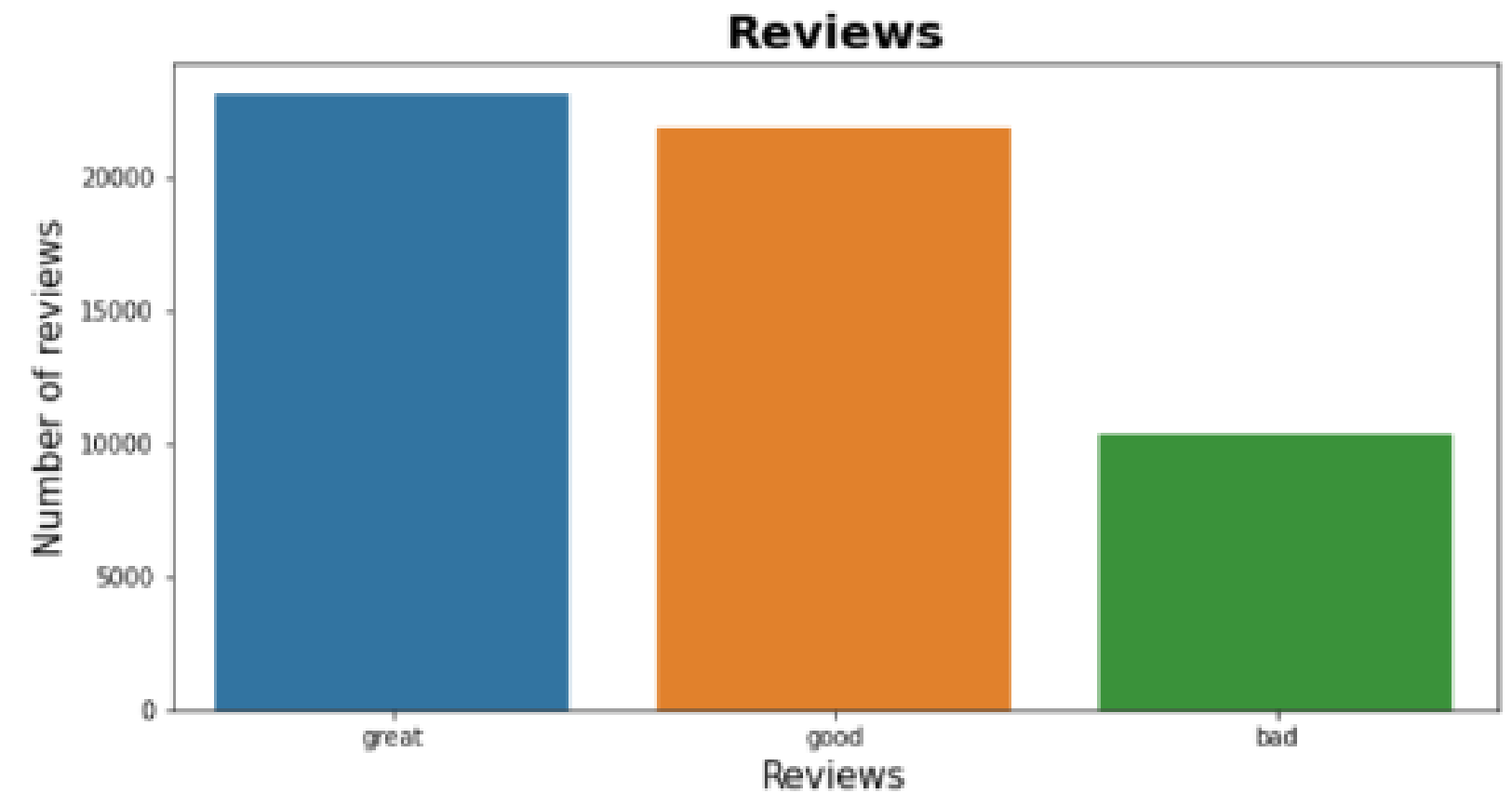
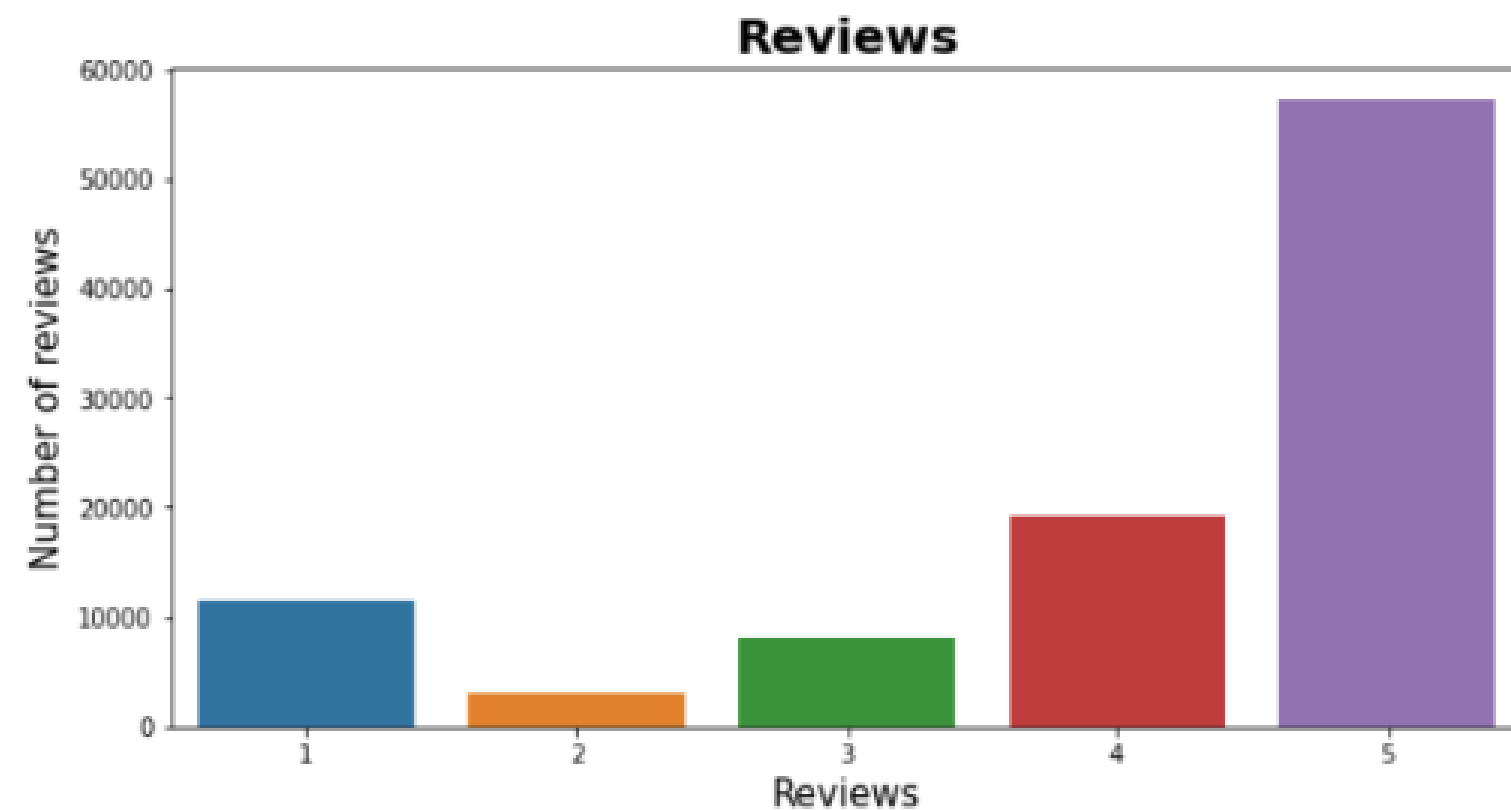
impact the product experiences and the reviewer's own product rating"



## Word of mouth

recomendations of other customers shape individual decision making

# Initial data x Training data



# Selected features:

payment with multiple installments

freight value

customer region

payment type

product average score

estimated vs delivered



Score  
prediction

# Results:

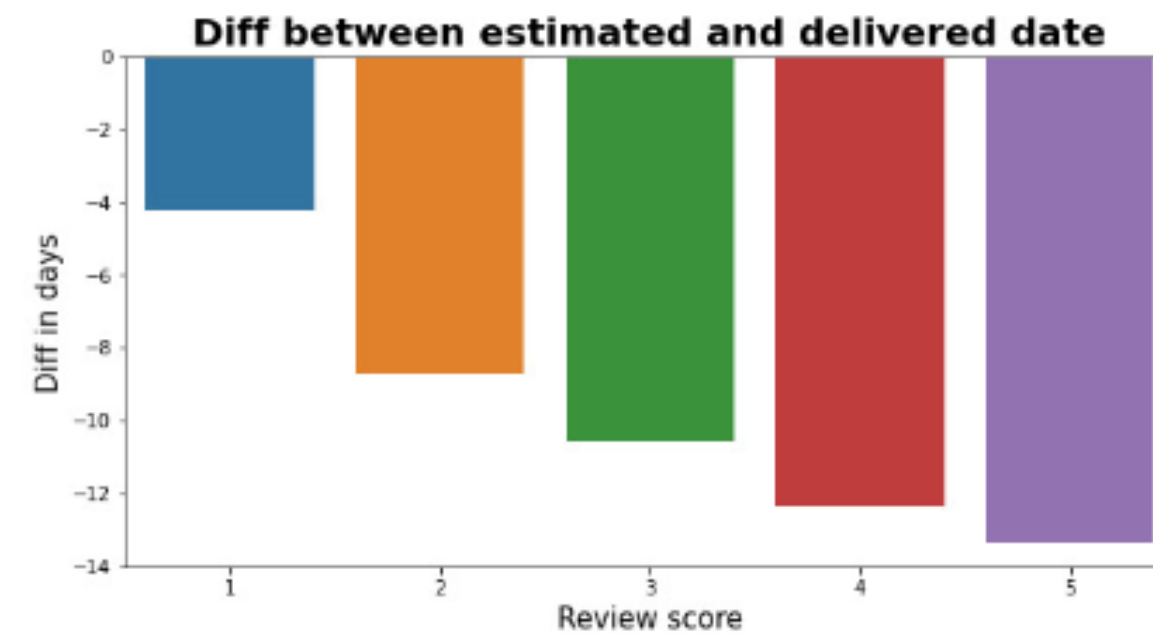
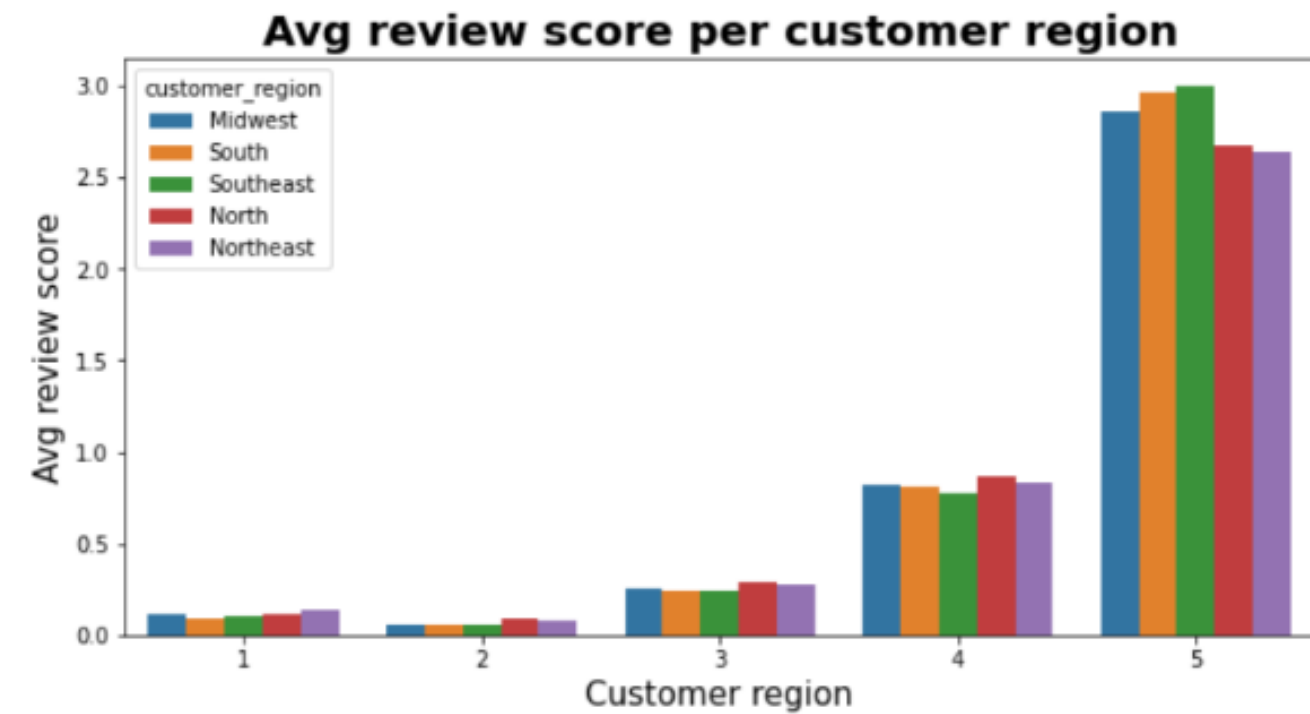
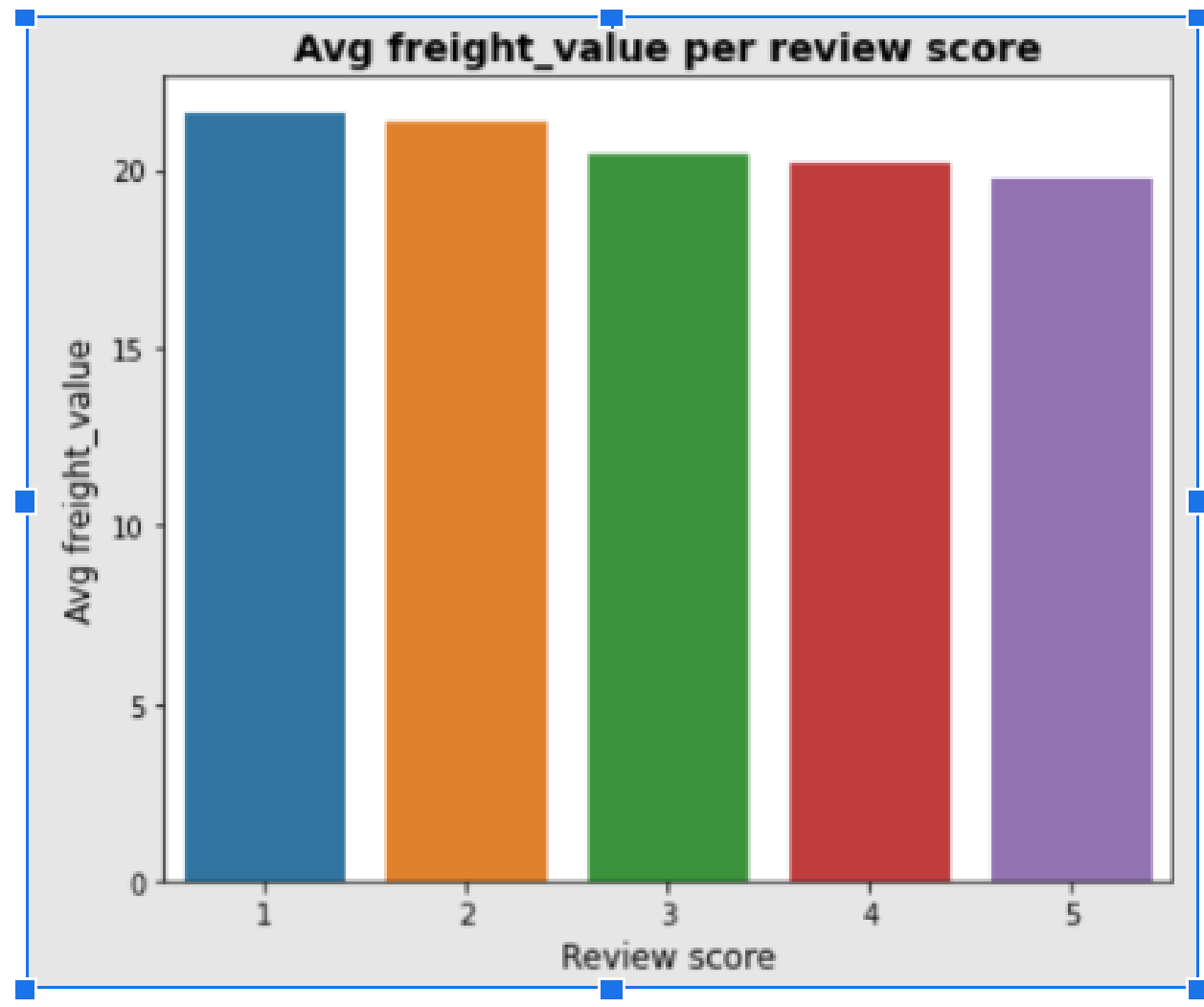
```
Metric: accuracy = 0.6281084386465552
Metric: f1 = 0.639348763369143
Metric: weightedPrecision = 0.6676417850046765
Metric: weightedRecall = 0.6281084386465552
truePositiveRateByLabel for label 0 = 0.683686715707486
truePositiveRateByLabel for label 1 = 0.571766418046207
truePositiveRateByLabel for label 2 = 0.4988335925349922
```

## Coefficients:

Name	Weight	Name	Weight	Name	Weight
more_than_one_ins...	62.382786	customer_region_vec	7.555803	more_than_one_ins...	55.173805
freight_value	62.24596	payment_type_vec	7.377288	freight_value	54.965374
customer_region_vec	61.644405	freight_value	7.280585	payment_type_vec	54.243603
payment_type_vec	61.62089	more_than_one_ins...	7.208983	customer_region_vec	54.0886
average_product_s...	34.027622	average_product_s...	1.6371744	average_product_s...	35.6648
estimated_vs_deli...	1.5876719	estimated_vs_deli...	0.41306403	estimated_vs_deli...	2.000736



# Insights that led to recommendations



# Recommendations



Negotiate favorable shipping rates with logistics partners on behalf of its clients



Renegotiate with logistic partners to make sure the delivery timings are respected and delivery time is predicted precisely.



Assist its clients, optimizing their inventory for regional preferences, providing localized support, and ensuring efficient delivery

# Thank you.

Authors:  
Yramaia Salviano 20220033  
Antonina Filatova 20221104  
Marta Manevska 20220056  
Ludovico Toscano 20220044