

# **Ahmad Dzikra Fatahillah**

**Data Scientist Assessment Test** 

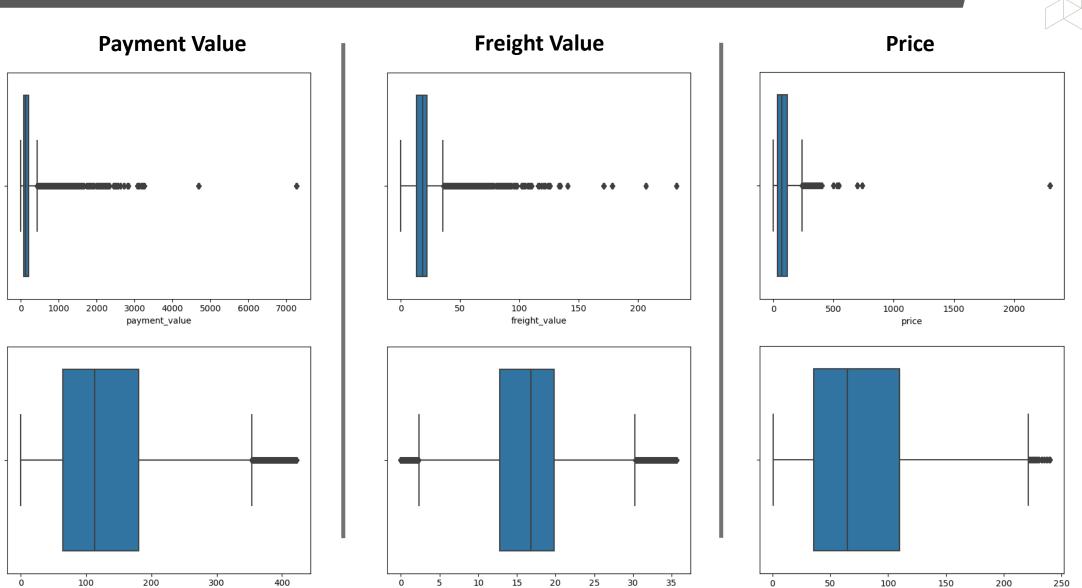


**Before** 

**After** 

## **Data Preparation – Handling Outliers**

payment\_value



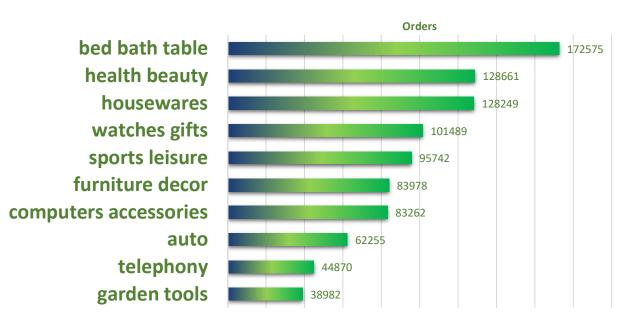
freight\_value

price



### **Exploratory Data Analysis**

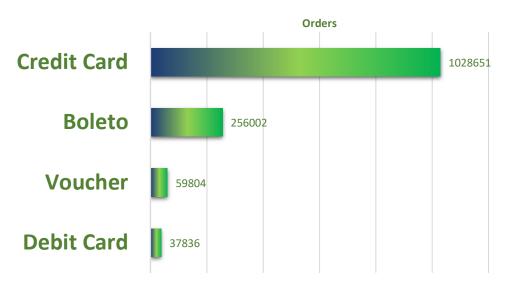
#### **Categories by Orders**



#### **Payment Installments by Orders**



#### **Payment Types by Orders**



#### **Review Scores by Orders**

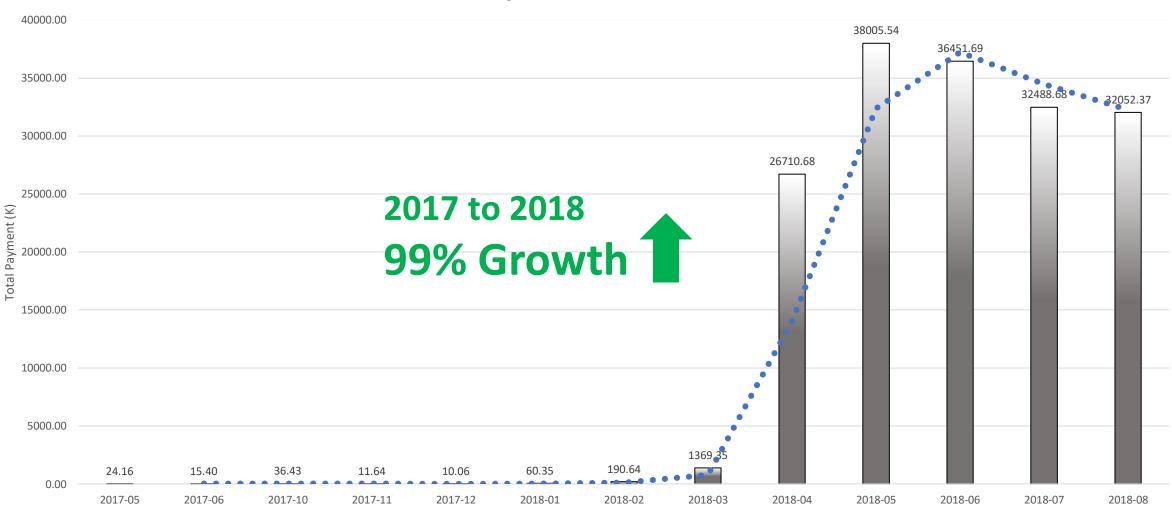






# Growth Payment (2017 to 2018)

### **Payment Growth**





#### **Customer City by Orders**



#### **Seller City by Orders**



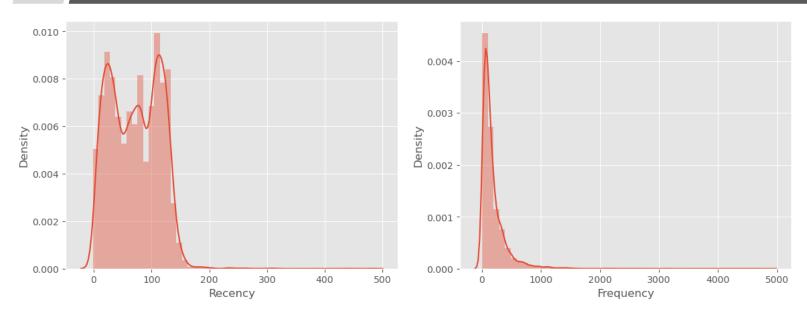
#### Brazil view of events

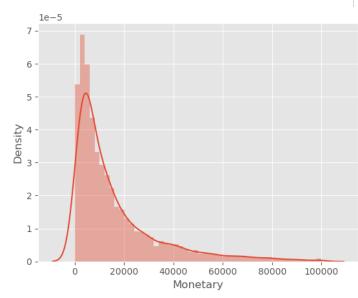






## **RFM – Customer Segmentation**





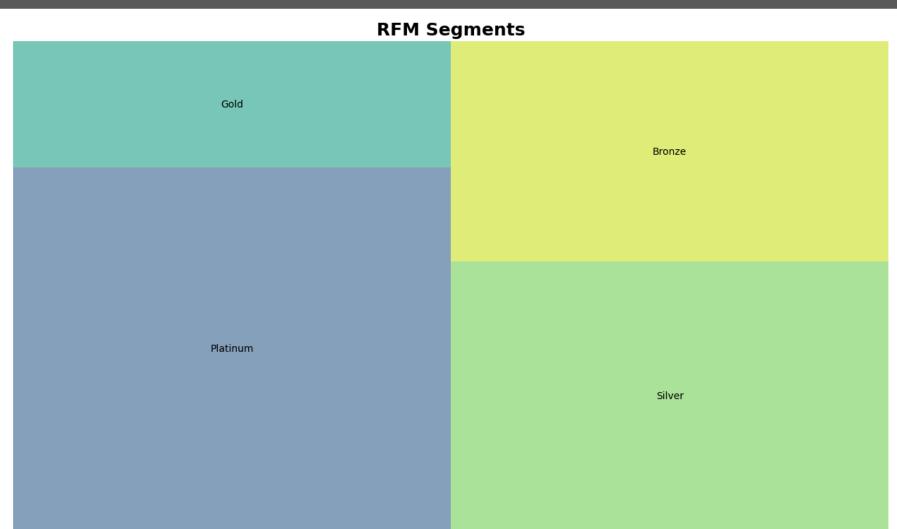
**Top 5 Customer by Monetary** 

Customer Id	Recency	Frequency	Monetary	R	F	М	RFM Group	RFM Score	RFM Loyalty
9eb3d566e87289dcb0acf28e1407c839	81	4872	1188865.44	3	1	1	311	5	Platinum
57255fab455feb9858e2ef1490150eb0	126	2139	725826.87	4	1	1	411	6	Platinum
5c664af0ac60101f1ca9f331b338e66b	109	3144	647412.48	3	1	1	311	5	Platinum
bab47e7a1b74c94a99ea042acfa7ea66	36	2220	634964.4	2	1	1	211	4	Platinum
932289aae7b7195bc53d4c640ea071f4	51	2093	621662.86	2	1	1	211	4	Platinum





# RFM – Customer Segmentation







### **Total Order Prediction & Modeling using Linear Regression**

#### **Linear Regression Total Order Prediction using Payment Value**

We can input the data:

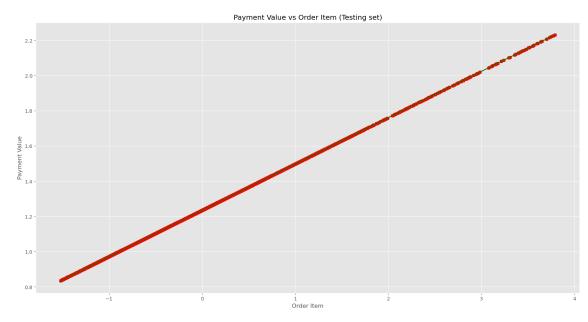
For example

Payment Value: 121

**Review Score: 3** 

Price: 77

The Total Order Prediction Result: 29,88



#### Prediction Result

```
user_input = [[121,3,77]] #[payment_value','review_score','price']
user_pred = linear.predict(user_input)
print("Item Ordered by The Customer are:",user_pred)
```

Item Ordered by The Customer are: [29.88669288]



<sup>\*</sup>adjustable data input

<sup>\*</sup>adjustable parameter input