**Team Number 1**

**Data Analytics Final Project Proposal**

| **Proposed by:** | Team Number 1  Cyril Fofana  Akriti Dhandhania  Charlton D’Souza  Chi Ezeogu  Becky Leach  Enrique Molero |
| --- | --- |
| **Timeframe:** | Completion by 2nd of October 2023.  Presentation on 2nd of October 2023 |

**Overview**

*[add an overview of your intended approach here. Which topics from the Olist dataset will you analyse? Which questions will you ask?]*

Holistically we will be looking at three core areas of the Olist dataset the sales, customer satisfaction and logistics fields to determine how each unique field relate to each other and also how they differ and impact the sales of a product and service. In the order\_payment table dataset it would be interesting to see how the types of payment can have an influence over the installment\_plan and payment\_value (interest rate). On the other hand we’ll see how the reviews directly correlate to repurchases and create a cluster to gain a better insight. And finally look at how logistics is affected and try to understand maybe the correlation between good, average and bad reviews in relation to delivery time.

**Plan of action**

*[List the concrete steps your group will take to complete this work. Include descriptions of how the work will be divided between team members].*

1. **Sales** *- Becky / Cyril*

* Cards: No of Orders per year, Top products, Total Sales
* Line Charts: Predictions - how we will use the purchase dates to predict future sales
* Analyse the correlation between Public Holidays and Sale peaks
* Types of payment (CC, Bolero & Voucher) > Instalment Plan
* Product category > Product Amount > Repurchases
* Seller who is the most profitable > City
* Recommendation

1. **Customer Satisfaction** *- Chi / Charlton*

* *Cards: Total no of customers, sellers, % of good and bad reviews, Avg Review Time, Avg rating*
* *Sensitivity Analysis*
* *Customer > Reviews > Repurchases*
* *Matrix: Unique Customer > Customer ID or Order ID > (All dates, Review)*
* *Cluster - scatter plot of positive, negative and neutral reviews*
* *States vs Ratings*
* *Avg time taken to drop a review vs rating*
* *Recommendation*

1. *Logistics - Akriti / Enrique*

* *Cards: Avg Delivery time, Avg processing time*
* *Product categories > Products > (Avg Delivery Time, Avg Ratings)*
* *Estimated delivery time vs Actual delivery time*
* *Delivery time vs Customer review*
* *Sellers’ distance vs buyers’ distance*
* *Gantt Chart*
* *Recommendation*