

## **Bonus points, Plotly application development report**

**Paulo Martins 2015469**

**Beatriz Fonseca 20201599**

### **# Introduction**

This application makes use of the Plotly library to produce dynamic and interactive data visualizations, in conjunction with the Bootstrap framework to build a responsive and accessible HTML interface, ensuring compatibility with modern web browsers. All necessary files are contained within the accompanying directory, and users can easily get started by launching the application.

It provides end-users with a simple and robust tool to gain a comprehensive understanding of the characteristics of the ABCDEats dataset. To further enable decision-makers to explore the data freely, the application leverages column filtering and row drill-down capabilities. Allowing for insights into individual feature characteristics, their interrelationships, and actionable customer segment insights derived through advanced clustering techniques.

### **# Structure**

The visualizations included in the dashboard belong to one of four types, Exploratory Analysis and Clustering, below in more detail.

#### **Overview (to be implemented):**

- introductory page, with the main findings

#### **Exploratory Analysis:**

- Histogram (or Barplots) and Boxplot of single variables
- Stacked Histograms, by categorical variables.
- Pairplots, of pairwise variables
- Three Way ANOVA of select variables

#### **Clustering:**

- By RFM - using RFM analysis and KMeans
- By City - using Spectral Clustering and KMeans
- By Cuisines - sparse PCA and Gaussian Mixture Models
- By Time - using Non-Negative Matrix Factorization and Self-Organizing Maps

## Customer Profiling (to be implemented):

- Allows the user to view a customers profiling and segmentation, based on their inputs.

### # Features

Noteworthy features of this implementation include:

#### # Column Selector

Allows the user to select the columns they wish to include in visualizations.

#### # Slicer based Drill-Down

Allows the user to slice the rows of selected columns based on valid set operations for the variable type.

#### # Conditional Filter

Allows the user to define conditions to apply over the observations.

#### # Customer behavior prediction

Allows the user to provide as inputs values for clustering variables, and receive as output a comprehensive prediction of customer behavior based on cluster predictions.

### # Results - Insights

Building an application with Plotly was incredibly intuitive, and easy, for developers with experience in both back end and front end. We were able to create a very aesthetic and immersive experience, that allows users to dive deep into the data.

Even though we have been ourselves working with this dataset for quite some time not, in truth, when we were able to slice our clusters using the tools we built, many patterns became evident. For instance, a user can compare how the Spending Behavior Cluster changes when we limit our analysis to either, only new, or only older customers. And in simpler terms even finding relations between variables using pair plots became much easier, being so easy to iterate through feature combinations.

### # Results Images



Figure 1: Histogram and Boxplot with conditional filter



Figure 2: Spending CCluster pairwise plots with cluster labels

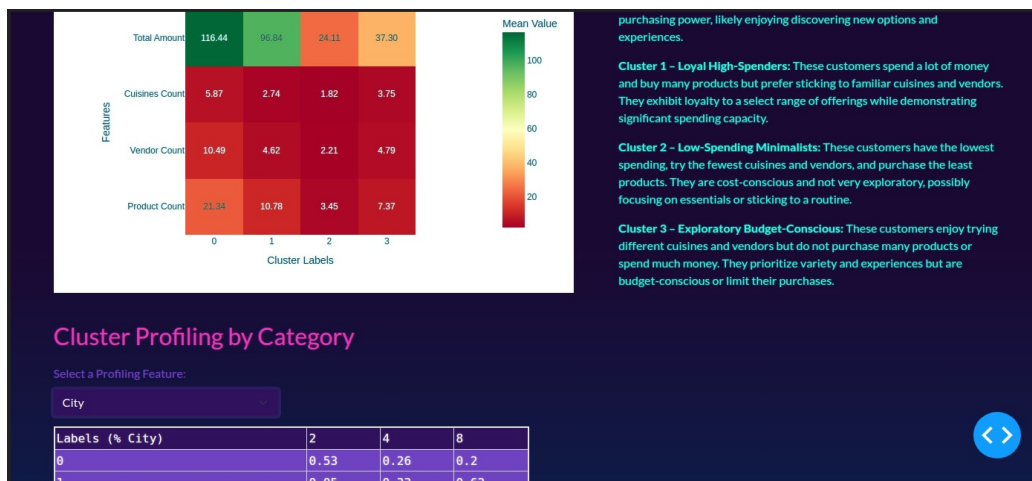


Figure 3: Heatmap of average behaviour, and profiling(bottom)

## # ANNEX

### # Implementation Details:

This application make heavy use of **joblib** to create **“.pkl”** objects that save trained models and other previously trained, making the process of profiling customer segments much less computationally and time intensive.

To run the application, the user needs only to run “main.ipynb”, which runs on HTML, in the near future we hope to have this code fully deployed as an executable to HTML.

**Link to GitHub:** <https://github.com/M4rtinsFons3caConsulting/DataMining>