**[Coffee Order : Who are the Top 5 Customers?](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md" \l "prosper-loan-analysis-who-is-struggling-with-their-loans)**

[**INTRODUCTION**](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md#introduction)

This project is dedicated to the analysing a coffee order dataset from a coffee store. The primary objective is to identify the top 5 coffee customers, spanning across all flavours, in order to award them a loyalty package. Additionally, the project aims to gain valuable insights into customer ordering behaviours.

This is personal project is based on a dataset obtained from Kaggle and this project allowed me to sharpen my data cleaning and preparation skills, ensuring that the data is ready for analysis. It also served as a platform for honing data visualization using Excel for exploratory and explanatory data analysis.

[**OVERVIEW OF THE DATASET**](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md#overview-of-the-dataset)

The dataset contains information of about 1001 customers that patronizes the Coffee shop. It has information about the customers’ names, email, coffee type, size, unit price, amongst other variables. There are 16 fields for the dataset.

[**DATA CLEANING**](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md#data-cleaning)

In this beginner-level project, I worked on cleaning and preparing a dataset to make it ready for visualization. The dataset used was duplicated initially to ensure that no data would be lost in case of any mistakes during the process.

**Project Workflow**

Data Duplication: I started by creating a duplicate of the dataset to safeguard against any potential loss due to errors.

Data Inspection: I applied filtering to each field, allowing me to quickly examine the elements within the dataset and gain a better understanding of its contents.

Duplicate Removal: Next, I checked for and removed any duplicate entries to ensure data consistency and accuracy.

Data Enhancement: To update the main table, I used VLOOKUP and INDEX-MATCH wherever necessary, enhancing the dataset's completeness.

Crosschecking: After crosschecking the new table, I confirmed that the dataset was now ready for visualization.

**Skills Gained**

This project provided an excellent opportunity to gain hands-on experience in data cleaning, data duplication, filtering, duplicate removal, and data enhancement using VLOOKUP and INDEX-MATCH. It prepared the dataset for visualization and helped me develop the foundational skills required for data analysis and presentation.

[**KEY INSIGHTS**](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md#key-insights)

After conducting an analysis of the dataset, here are the key insights and conclusions that were derived:

**Top Customers**: The top 5 customers, across all continents, who qualify for the loyalty package based on the highest spending on coffee are:

* Allis Wilmore
* Brenn Dundredge
* Terri Farra
* Nelson Cutler
* Don Flintiff

**Preferred Coffee Size and Roast Type**: The size of coffee that customers spent the most money on is 2.5kg, and the preferred roast type is Light coffee.

**Top Sales Country**: The United States of America stands out as the country with the highest coffee sales.

**Loyalty Cards and Purchases**: It's noteworthy that customers without loyalty cards make the most purchases.

**Most Purchased Coffee Type**: The coffee variety that was most frequently purchased is Excelsa.

[**DATA VISUALIZATION**](https://github.com/Ebuka456/Prosper-Loan-Analysis/blob/main/README.md#data-visualization)

Data Visualization was done using Microsoft Excel. The Interactive report of the project can be found in the repository.

