

COFFEE SALES ANALYSIS

Hello, I'm Clifford and today, I will walk you through my process of completing this excel project, the "Coffee Sales Analysis". I will cover how I went by this project from the earliest stage of obtaining the data, to data visualization.

Project Overview

The objective of the project was to develop a Coffee Sales Dashboard for McKeezy Coffee Shop, focusing on finding the following:

1. Sales over time for each coffee type.
2. Sales by country.
3. Top customers.
4. Loyalty Subscription.

Other Key Metrics like Total Sales, Total Profit, Total Orders, Top Selling Product, that will be deduced.

The dashboard will be interactive, so you will be able to filter by time period, type of coffee roast, coffee sizes, and whether or not the customers are loyalty members.

Getting the Data

First, I needed to get the data. I downloaded this dataset from Kaggle: [Coffee Bean Sales Raw Dataset](#). The dataset had three sheets:

1. Orders
2. Customers
3. Products

These sheets contained all the details about the coffee shop's sales transactions, customer info, and product specifics.

After downloading the dataset, I performed the following steps which are:

1. Data Cleaning
2. Data Analysis
3. Data Visualization

Methodology

1. Data Cleaning

a. Data Exploration:

I conducted a thorough examination of the dataset in Microsoft Excel to enhance comprehension and identify areas requiring cleaning. The Orders and Customers sheet had a total of 1000 records each, and the Products sheet had 48 records. There were no duplicates, errors or blank values.

b. Data Cleaning:

The data cleaning process started with copying the entire workbook into another workbook. I always do this and save the original dataset so I can come back to it for any reference. Now, I converted each sheet into a table so I can easily refer and work with it. The following are some cleaning procedures I performed:

1. Converting all the price related fields to currency with the \$ sign and two decimal places.
2. Creating a custom format and adding “kg” to the size field.
3. The coffee types were then changed to their full meaning to avoid confusion. (Ara = Arabica, Rob = Robusta, Lib = Liberica, Exc = Excelsia)
4. The above naming was done for the Roast Types (L = Light, M = Medium, D = Dark)

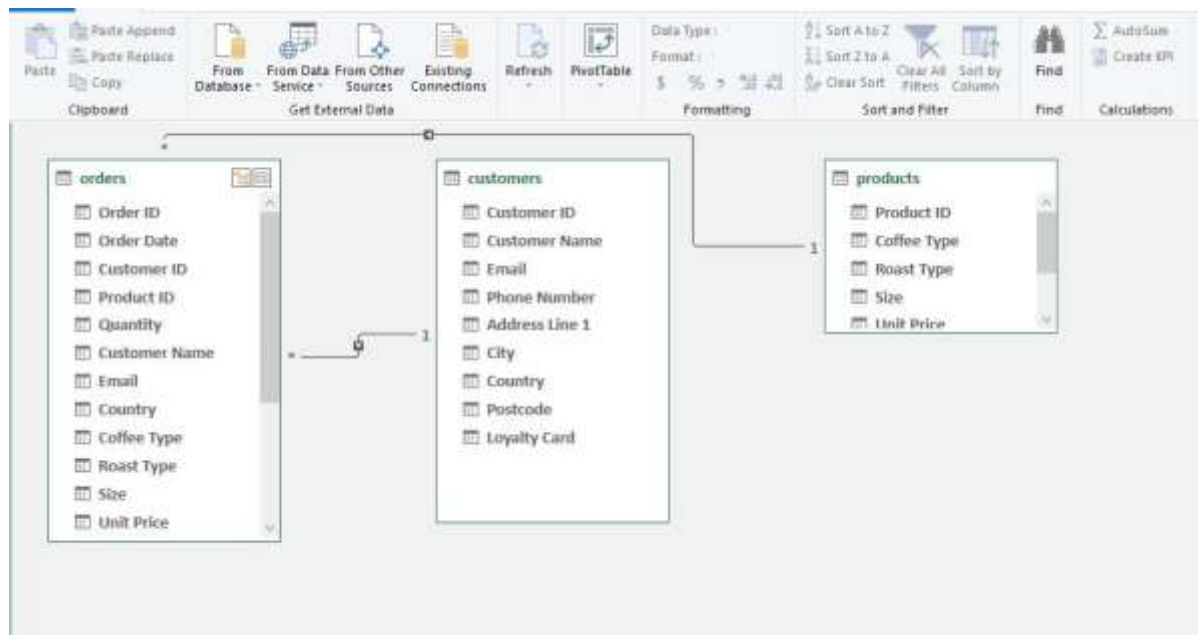
*Step 3 and 4 were done with the **Find and Replace** method in Excel.*

2. Data Analysis

a. Data Modelling:

I added the three tables to the Data Model so I could use Power Pivot to do data modelling. I created a relationship between the three tables. Using the Customer ID as a key, I created a relationship between the Orders table and the Customers table. Similarly, I created a relationship between the Orders table and the Products table using the Product ID.

Instead of creating relationships, I could have used **XLOOKUP** and **Index Match** to get the data from the Products and Customers table into the Orders table so it will be one whole table. I thought about how it might affect efficiency if I have the same data in multiple sheets.



I added the count of orders for each coffee type to get the total orders.

