Excel Portfolio Project

Coffee Bean Sales - Data Analysis

Elaborated by:
Fernando Andrés Chajud Vásquez

Project duration:
one day

Feb - 2024

Abstract

This project aims to demonstrate my proficiency in the use of Excel through the analysis and transformation of a dataset obtained from Kaggle. The dataset consists of coffee bean sales records. The goal is to create an interactive dashboard to track historical sales behavior, using specific segmenters to visualize the following key aspects:

Total sales.
Sales by country.
Top 5 customers

The dashboard will provide a clear and understandable visual representation of this data, which will facilitate the identification of trends and patterns in coffee bean sales.

Introduction

In the current context of data analysis, the ability to manipulate and extract valuable information from complex data sets is critical. In this project, a detailed analysis of coffee bean sales behavior is presented, using advanced Excel tools to effectively transform and visualize the data.

The main objective of this project is to demonstrate mastery of Excel as a data analysis tool, through the implementation of an interactive dashboard to track the historical behavior of coffee sales. This analysis is based on a dataset obtained from Kaggle, which provides detailed information on sales transactions, including data such as country of origin, quantity sold and main customers.

The importance of this project lies in the ability to identify trends, patterns and opportunities for improvement from the available data. By better understanding the behavior of coffee sales over time and in different geographic regions, informed strategic decisions can be made to optimize operational efficiency and maximize revenues.

Methodology

The methodology used in this project is divided into several stages, ranging from data collection and preparation to the creation of the interactive dashboard in Excel. The steps followed are detailed below:

1. Data collection: A data set of coffee bean sales data was obtained from Kaggle, which includes information on the transactions made, such as country of origin, quantity sold and main customers.

2. Data cleaning: A thorough data cleaning was performed to remove any duplicate or inconsistent entries. In addition, possible errors in the data were identified and corrected, ensuring the integrity and quality of the information.

3. Data transformation: Several transformations were performed on the data to prepare it properly for analysis. This included data types formatting, the aggregation of columns in the Orders table and the filling of that data through the relationships between the other tables using foreign keys.

4. Dashboard design: An interactive dashboard was designed in Excel using the advanced functionalities of the tool, such as pivot charts, pivot tables and filter controls. Multiple views were created to allow a detailed exploration of the data from different perspectives.

5. Dashboard implementation: The dashboard design was implemented in Excel, ensuring its functionality and usability. Extensive testing was carried out to verify that all the functions of the dashboard worked as intended.

Project development

This section presents the steps taken from the preparation of the data to the presentation of the final results of the analysis and the preparation of the dashboard.

Data

The dataset was obtained from kaggle: https://www.kaggle.com/datasets/saadharoon27/coffee-bean-sales-raw-dataset

The dataset contains 3 tables:

Table orders:

Order ID	Order Date	Customer ID	Product ID	Quantity
QEV-37451-860	05/09/2019	17670-51384-MA	R-M-1	2
QEV-37451-860	05/09/2019	17670-51384-MA	E-M-0.5	5
FAA-43335-268	17/06/2021	21125-22134-PX	A-L-1	1
KAC-83089-793	15/07/2021	23806-46781-OU	E-M-1	2
KAC-83089-793	15/07/2021	23806-46781-OU	R-L-2.5	2
CVP-18956-553	04/08/2021	86561-91660-RB	L-D-1	3
IPP-31994-879	21/01/2022	65223-29612-CB	E-D-0.5	3
SNZ-65340-705	20/05/2022	21134-81676-FR	L-L-0.2	1
EZT-46571-659	02/01/2019	03396-68805-ZC	R-M-0.5	3
NWQ-70061-912	05/09/2019	61021-27840-ZN	R-M-0.5	1
BKK-47233-845	08/03/2021	76239-90137-UQ	A-D-1	4
VQR-01002-970	28/10/2020	49315-21985-BB	E-L-2.5	5

Table customers:

Customer ID	Customer Name	Email	Phone Number	Address Line 1	City	Country	Postcode	Loyalty Card
17670-51384-MA	Aloisia Allner	aallner0@lulu.com	+1 (862) 817-0124	57999 Pepper Wood Alley	Paterson	United States	7505	Yes
73342-18763-UW	Piotr Bote	pbote1@yelp.com	+353 (913) 396-4653	2112 Ridgeway Hill	Crumlin	Ireland	D6W	No
21125-22134-PX	Jami Redholes	jredholes2@tmall.com	+1 (210) 986-6806	5214 Bartillon Park	San Antonio	United States	78205	Yes
71253-00052-RN	Dene Azema	dazema3@facebook.com	+1 (217) 418-0714	27 Maywood Place	Springfield	United States	62711	Yes
23806-46781-OU	Christoffer O' Shea		+353 (698) 362-9201	38980 Manitowish Junction	Cill Airne	Ireland	N41	No
86561-91660-RB	Beryle Cottier		+1 (570) 289-7473	2651 Stoughton Place	Scranton	United States	18505	No
65223-29612-CB	Shaylynn Lobe	slobe6@nifty.com	+1 (937) 954-4541	7005 Mariners Cove Place	Dayton	United States	45440	Yes
21134-81676-FR	Melvin Wharfe		+353 (507) 574-3034	7 Straubel Road	Kill	Ireland	P24	Yes
03396-68805-ZC	Guthrey Petracci	gpetracci8@livejournal.com	+1 (310) 868-1842	949 Paget Parkway	Los Angeles	United States	90045	No
61021-27840-ZN	Rodger Raven	rraven9@ed.gov	+1 (213) 263-0288	1 Reinke Avenue	Los Angeles	United States	90065	No
76239-90137-UQ	Ferrell Ferber	fferbera@businesswire.com	+1 (408) 383-5302	68 High Crossing Court	San Jose	United States	95160	No
49315-21985-BB	Duky Phizackerly	dphizackerlyb@utexas.edu	+1 (408) 533-6012	28643 Bluejay Crossing	San Jose	United States	95194	Yes
34136-36674-OM	Rosaleen Scholar	rscholarc@nyu.edu	+1 (804) 420-0420	80915 Montana Park	Richmond	United States	23285	No
39396-12890-PE	Terence Vanyutin	tvanyutind@wix.com		331 Bunting Hill	Migrate	United States	41905	No

Table products:

Product ID	Coffee Type	Roast Type	Size	Unit Price	Price per 100g	Profit
A-L-0.2	Ara	L	0.2	3.885	1.9425	0.34965
A-L-0.5	Ara	L	0.5	7.77	1.554	0.6993
A-L-1	Ara	L	1.0	12.95	1.295	1.1655
A-L-2.5	Ara	L	2.5	29.785	1.1914	2.68065
A-M-0.2	Ara	M	0.2	3.375	1.6875	0.30375
A-M-0.5	Ara	M	0.5	6.75	1.35	0.6075
A-M-1	Ara	M	1.0	11.25	1.125	1.0125
A-M-2.5	Ara	M	2.5	25.875	1.035	2.32875
A-D-0.2	Ara	D	0.2	2.985	1.4925	0.26865
A-D-0.5	Ara	D	0.5	5.97	1.194	0.5373
A-D-1	Ara	D	1.0	9.95	0.995	0.8955
A-D-2.5	Ara	D	2.5	22.885	0.9154	2.05965
R-L-0.2	Rob	L	0.2	3.585	1.7925	0.2151
R-L-0.5	Rob	L	0.5	7.17	1.434	0.4302

Data transformation

Having viewed and reviewed the data, the following columns were added to the *orders* table: Customer Name, Email, Country, Coffee Type, Roast Type, Size, Unit Price, Sales and Loyalty card.

Then, the Excel function XLOOKUP() is used to fill in the data of the columns related to the customer information added in the *orders* table, using the foreign key Customer ID of the *orders* table.

Subsequently, to dynamically fill in the data of the columns related to the product sold information added in the *orders* table, the Excel functions INDEX() and MATCH() are used, using the foreign key Product ID of the *orders* table. The function is used:

```
INDEX(
        Array = $table products$;
        Row_num = MATCH(
                Lookup value = Product ID;
                Lookup array = $products!Product ID$;
                Match type = exact match
                );
        Column_num = MATCH(
```

Lookup value = orders!column_names;
Lookup array = products!column_names;
Match type = exact match
);
)

Finally, to fill in the Sales column, a multiplication is used.

To change the abbreviation of the coffee type to the full name of the coffee type, a nested IF() function is used.

Then, the date format is changed to: dd-mmm-yyyy. This, to help understand better the final information.

Afterwards, the number format for column Size is costumed so it shows the scale (kg).

As a last step before going into the pivot tables, we are going to insert a table.

Pivot tables and Pivot charts + Formatting

Total Sales:

We insert a pivot table called Total Sales:

- Columns: Coffee Type Name
- Rows: Years and Months from Order Date
- Values: Sum of Sales (whole number)

We insert a pivot chart:

- Lines chart
- Hide field buttons

We format the chart.

We insert the timeline and format it.

We insert the slicers and format it.

Sales by Country:

We insert a pivot table called Sales by Country:

- Axes: Country
- Values: Sum of Sales

We insert a pivot chart:

- Bar chart
- Hide field buttons

We format the chart.

Top 5 customers:

We insert a pivot table called Sales by Country:

- Axes: Customer name
- Values: Sum of Sales

We insert a pivot chart:

- Bar chart
- Hide field buttons
- Sort data

We format the chart.

Final Dashboard