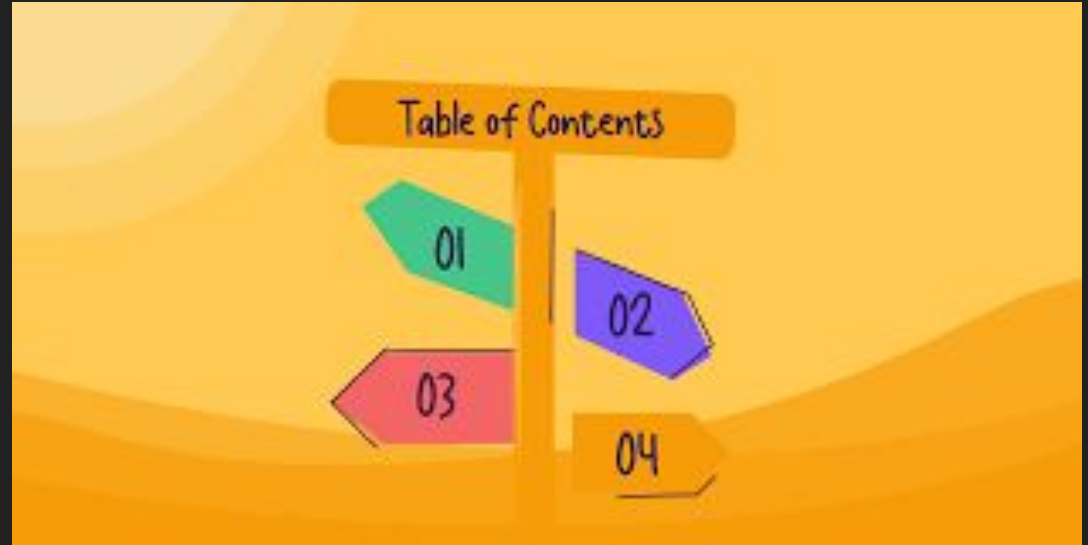


Financial Transactions

By Zaquel Cossa

Table Of Contents

1. Introduction
2. Dataset Overview
3. Data Cleaning
4. Feature Engineering
5. Analysis
6. Dashboard
7. Key Insights
8. Conclusion



Introduction

- Purpose: Clean and analyze a real-world financial transaction dataset
- Objectives:
 - Handle missing or invalid data
 - Fix formatting issues
 - Summarize trends and patterns
 - Visualize through Dashboards

Dataset Overview

Source: Dirty Financial Transaction Dataset From Kaggle -
<https://www.kaggle.com/datasets/alfarisbachmid/dirty-financial-transactions-dataset?resource=download>

Columns: Transaction_ID, Transaction_Date, Customer_ID,
Product_Name, Quantity, Price, Payment Method,
Transaction_Status

Raw Data

Transaction_ID									
	A	B	C	D	E	F	G	H	I
1	Transaction	Transaction	Customer	Product_N	Quantity	Price	Payment	Transaction_Status	
2	T0001	8/2/2024	C2205	Headphor	-5	\$420.21	pay pal		
3	T0002	2/10/2020	C3156	Coffee	469	-445.342	creditcard	Pending	
4	T0003	2025-02-30	C2919	Tablet	-4	810.993	credit card	completed	
5	T0004	8/17/2020	C3009	Tab	-7	868.6083	PayPal	Pending	
6	T0005	2025-02-30	C3488	Coffee Ma	-10	-763.122	PayPal	completed	
7	T0006	#####	C4241	Smartpho	598		PayPal	Completed	
8		2025-02-30	C1313	Laptop	10		credit card	Completed	
9	T0008	2023-13-0	C4736	Headphor	669	-86.9213	Cash		
10	T0009		C3387	Tablet	10	461.702	PayPal		
11	T0010	2025-02-30	C2846	Laptop	-1	404.8907	creditcard	Pending	
12	T0011	2023-13-0	C1004	Laptop	10	-600.839	PayPal	complete	
13	T0012	2023-13-0	C2912	Tablet	608		PayPal	completed	
14	T0013	3/18/2020	C1692	Smartpho	564		Cash	Completed	
15	T0014	6/19/2020	C2992	Coffee Ma	696	905.5147	creditcard	Failed	
16	T0015	2025-02-30	C3879	Headphor	-2	523.8428	credit card	Pending	
17	T0016	2025-02-30	C2837	Laptop	-3		creditcard	Pending	
18	T0017	7/7/2020	C3776	Smartpho	236	-94.5573	credit card	Pending	
19	T0018	2025-02-30	C847	Coffee Ma	7	-108.438	credit card	complete	
20	T0019	2/19/2024	C4871	Smartphone			pay pal	Pending	
21	T0020	2025-02-30	C1924	Headphor	-6		pay pal		
22	T0021	2025-02-30	C1137	Coffee Ma	856	-488.863	Credit Car	Failed	
23	T0022	2023-13-0	C3883	Smartpho	9		creditcard	Completed	
24	T0023	2025-02-30	C4877	Tablet	538	-733.829	PayPal	complete	
25	T0024	11/3/2023	C828	Coffee Ma	487		Cash		
26	T0025	#####	C4985	Tablet	114	276.9363	PayPal		
27	T0026	1/1/2020	C1096	Cof	654		PayPal	complete	
28	T73466	2025-02-30	C2702	Coffee Ma	-5	555.6133	creditcard	complete	

<

>

Data

+

Ready

Accessibility: Investigate

Data Cleaning

In this project, I applied several data cleaning techniques to prepare the dataset for analysis. Each step focused on improving data quality and consistency. Below, I walk through the key actions I took during the cleaning process.



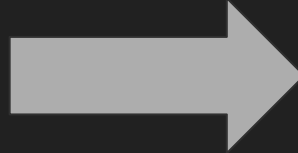
1. Formatted Data as a Table & Removed Duplicates

	A	B	C	D	E	F	G	H
1	Transaction_ID	Transaction_Date	Customer_ID	Product_Name	Quantity	Price	Payment_Method	Transaction_Status
2	T0001	8/2/2024	C2205	Headphones	-5	\$420.21	pay pal	
3	T0002	2/10/2020	C3156	Coffee	469	-445.3420253	creditcard	Pending
4	T0003	2025-02-30	C2919	Tablet	-4	810.9930124	credit card	completed
5	T0004	8/17/2020	C3009	Tab	-7	868.6083413	PayPal	Pending
6	T0005	2025-02-30	C3488	Coffee Machine	-10	-763.122449	PayPal	completed
7	T0006	10/26/2021	C4241	Smartphone	598		PayPal	Completed
8		2025-02-30	C1313	Laptop	10		credit card	Completed
9	T0008	2023-13-01	C4736	Headphones	669	-86.92126929	Cash	
10	T0009		C3387	Tablet	10	461.7019844	PayPal	
11	T0010	2025-02-30	C2846	Laptop	-1	404.8907066	creditcard	Pending
12	T0011	2023-13-01	C1004	Laptop	10	-600.8393094	PayPal	complete
13	T0012	2023-13-01	C2912	Tablet	608		PayPal	completed
14	T0013	3/18/2020	C1692	Smartphone	564		Cash	Completed
15	T0014	6/19/2020	C2992	Coffee Machine	696	905.5147299	creditcard	Failed
16	T0015	2025-02-30	C3879	Headphones	-2	523.8427582	credit card	Pending
17	T0016	2025-02-30	C2837	Laptop	-3		creditcard	Pending
18	T0017	7/7/2020	C3776	Smartphone	236	-94.55726527	credit card	Pending
19	T0018	2025-02-30	C847	Coffee Ma	7	-108.4378241	credit card	complete
20	T0019	2/19/2024	C4871	Smartphone			pay pal	Pending
21	T0020	2025-02-30	C1924	Headphones	-6		pay pal	
22	T0021	2025-02-30	C1137	Coffee Machine	856	-488.8626491	Credit Card	Failed
23	T0022	2023-13-01	C3883	Smartphone	9		creditcard	Completed
24	T0023	2025-02-30	C4877	Tablet	538	-733.8292285	PayPal	complete
25	T0024	11/3/2023	C828	Coffee Machine	487		Cash	
26	T0025	11/18/2022	C4985	Tablet	114	276.9363204	PayPal	
27	T0026	1/1/2020	C1096	Cof	654		PayPal	complete
28	T73466	2025-02-30	C2702	Coffee Machine	-5	555.6133484	creditcard	complete

2. Used Power Query to Delete Rows with a missing Transaction_ID , Customer_ID, Quantity, or Price.

Explanation: These columns are essential data, and without these values we can't use the rest of the data for that transaction.

1	Transaction_ID	Customer_ID	Product_Name	Quantity	Price
2	T0026	C1096	Cof	654	
3	T7685	C257	Coffee Machine	9	945.8729674
4	T8720	C3400	Headphones	-9	596.1991538
5	T15153	C2165	Smartphone		
6	T22312	C687	Tablet	973	848.4864727
7	T23672	C2215	Smartphone	-3	
8	T25919	C3308	Coffee Machine	939	
9	T33592	C3850	Smartphone	604	-201.4210649
10	T39156	C4666	Coffee Machine	1	800.9409012
11	T46858	C3054	Coffee Machine	-2	952.3268604
12	T59619	C4785	Smartphone	5	
13	T62679	C1912	Tablet	7	
14	T66678	C570	Smartphone		-514.0365804
15	T66966	C4761	Headphones	943	495.2697687
16	T68111	C2744	Smartphone	7	857.7359015
17	T74138	C2910	Laptop	-1	837.1273828

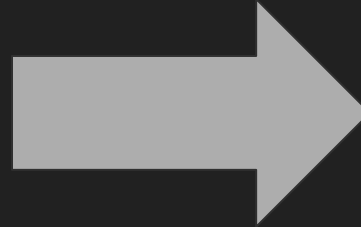


1	Transaction_ID	Customer_ID	Product_Name	Quantity	Price
2	T7685	C257	Coffee Machine	9	945.87
3	T22312	C687	Tablet	973	848.49
4	T39156	C4666	Coffee Machine	1	800.94
5	T66966	C4761	Headphones	943	495.27
6	T68111	C2744	Smartphone	7	857.74
7	T19057	C832	Coffee Machine	913	270.64
8	T43708	C3518	Headphones	5	346.06
9	T96200	C1860	Coffee Machine	747	551.78
10	T99760	C779	Headphones	1	145.72
11	T24601	C2008	Smartphone	9	992.19
12	T61388	C2081	Coffee Machine	9	733.07
13	T62952	C720	Headphones	149	891.75
14	T10772	C3103	Smartphone	3	879.61
15	T29333	C2555	Smartphone	870	597.76
16	T66896	C4993	Headphones	270	939.22
17	T89018	C405	Tablet	8	847.95

3. Spell Check and Standardize Product_Name

Product Name

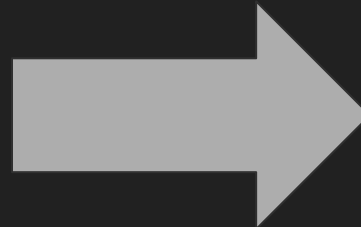
Product_Name
Head
Headphones
Headphones
Coffee Machin
Smar
Coffe



Product_Name
Coffee Machine
Coffee Machine
Headphones
Smartphone
Tablet

Payment Method

Payment_Method
PayPal
creditcard
pay pal
PayPal



Payment_Method
Credit Card
Cash
Credit Card
PayPal
Credit Card

4. Find Mode Value for Product Name and Payment Method Using Pivot Table

Product Pivot

Mode Product	Mode Payment
Tablet	Credit Card
Product Name ▼	Count of Product_Name
Tablet	5778
Laptop	5717
Coffee Machine	5690
Smartphone	5652
Headphones	5573
Grand Total	28410

Payment Pivot

	A	B
1	Mode Product	Mode Payment
2	Tablet	Credit Card
3		
4	Payment Meth ▼	Count of Payment_Method
5	Credit Card	12308
6	PayPal	12148
7	Cash	3954
8	Grand Total	28410

5. Replace Missing Product_Name and Payment Method with Modes

1	Transaction_ID	Transaction_Date	Customer_ID	Product_Name	Quantity	Price	Payment_Method	Transaction_Status
2	T7685	1/1/2020	C257	Coffee Machine	9	945.87	Credit Card	complete
3	T22312	1/1/2020	C687	Tablet	973	848.49	Cash	UNKNOWN
4	T39156	1/1/2020	C4666	Coffee Machine	1	800.94	Credit Card	Failed
5	T66966	1/1/2020	C4761	Headphones	943	495.27	PayPal	completed
6	T68111	1/1/2020	C2744	Smartphone	7	857.74	Credit Card	completed
7	T19057	1/2/2020	C832	Coffee Machine	913	270.64	Cash	Completed
8	T43708	1/2/2020	C3518	Headphones	5	346.06	Credit Card	completed
9	T96200	1/3/2020	C1860	Coffee Machine	747	551.78	PayPal	complete
10	T99760	1/3/2020	C779	Headphones	1	145.72	PayPal	Completed
11	T24601	1/4/2020	C2008	Smartphone	9	992.19	Credit Card	Pending
12	T61388	1/4/2020	C2081	Coffee Machine	9	733.07	Credit Card	completed
13	T62952	1/4/2020	C720	Headphones	149	891.75	Credit Card	Completed
14	T10772	1/5/2020	C3103	Smartphone	3	879.61	PayPal	Completed
15	T29333	1/5/2020	C2555	Smartphone	870	597.76	PayPal	Completed
16	T66896	1/5/2020	C4993	Headphones	270	939.22	Cash	Failed
17	T89918	1/5/2020	C405	Tablet	8	847.25	Cash	Completed

6. Replace Missing Transaction_Status with Unknown

H
Transaction_Status ▼
Pending
completed
Pending
completed
Completed
Pending



K
Transaction_Status ▼
Completed
UNKNOWN
Failed
Completed
Completed

Feature Engineering

Added new columns Day, Month, and Year Using Formulas.

	A	B	C	D	E	F	G	H	I	J	K
1	Transaction_ID ▾	Transaction Date ▾	Day ▾	Month ▾	Year ▾	Customer_ID ▾	Product Name ▾	Quantity ▾	Price ▾	Payment Method ▾	Transaction Status ▾
2	T7685	1/1/2020	Sunday	January	2020	C257	Coffee Machine	9	945.87	Credit Card	Completed
3	T22312	1/1/2020	Sunday	January	2020	C687	Tablet	973	848.49	Cash	UNKNOWN
4	T39156	1/1/2020	Sunday	January	2020	C4666	Coffee Machine	1	800.94	Credit Card	Failed
5	T66966	1/1/2020	Sunday	January	2020	C4761	Headphones	943	495.27	PayPal	Completed
6	T68111	1/1/2020	Sunday	January	2020	C2744	Smartphone	7	857.74	Credit Card	Completed
7	T19057	1/2/2020	Monday	January	2020	C832	Coffee Machine	913	270.64	Cash	Completed
8	T43708	1/2/2020	Monday	January	2020	C3518	Headphones	5	346.06	Credit Card	Completed
9	T96200	1/3/2020	Tuesday	January	2020	C1860	Coffee Machine	747	551.78	PayPal	Completed
0	T99760	1/3/2020	Tuesday	January	2020	C779	Headphones	1	145.72	PayPal	Completed
1	T24601	1/4/2020	Wednesday	January	2020	C2008	Smartphone	9	992.19	Credit Card	Pending
2	T61388	1/4/2020	Wednesday	January	2020	C2081	Coffee Machine	9	733.07	Credit Card	Completed
3	T62952	1/4/2020	Wednesday	January	2020	C720	Headphones	149	891.75	Credit Card	Completed
4	T10772	1/5/2020	Thursday	January	2020	C3103	Smartphone	3	879.61	PayPal	Completed
5	T29333	1/5/2020	Thursday	January	2020	C2555	Smartphone	870	597.76	PayPal	Completed
6	T66896	1/5/2020	Thursday	January	2020	C4993	Headphones	270	939.22	Cash	Failed
7	T89918	1/5/2020	Thursday	January	2020	C405	Tablet	8	847.25	Cash	Completed
8	T91910	1/5/2020	Thursday	January	2020	C188	Laptop	4	611.21	Credit Card	Failed
9	T46447	1/6/2020	Friday	January	2020	C1392	Laptop	3	249.10	Cash	Failed
0	T63088	1/7/2020	Saturday	January	2020	C4407	Tablet	456	591.39	Credit Card	Completed
1	T82327	1/7/2020	Saturday	January	2020	C3719	Smartphone	297	462.77	Credit Card	Pending
2	T41989	1/8/2020	Sunday	January	2020	C2864	Tablet	1	92.33	Credit Card	Completed
3	T43374	1/8/2020	Sunday	January	2020	C2792	Tablet	766	78.23	PayPal	Failed
4	T47954	1/8/2020	Sunday	January	2020	C551	Headphones	2	662.65	PayPal	Completed
5	T62244	1/8/2020	Sunday	January	2020	C2776	Smartphone	794	817.24	Credit Card	Completed
6	T81016	1/8/2020	Sunday	January	2020	C2016	Headphones	5	227.42	PayPal	Completed

Lookup Table

Created Transaction Status Lookup Table by Transaction_ID Using XLOOKUP.

Transaction ID:	T7685		
Transaction Status:	=XLOOKUP(N2,A:A,K:K,"Not Found")		



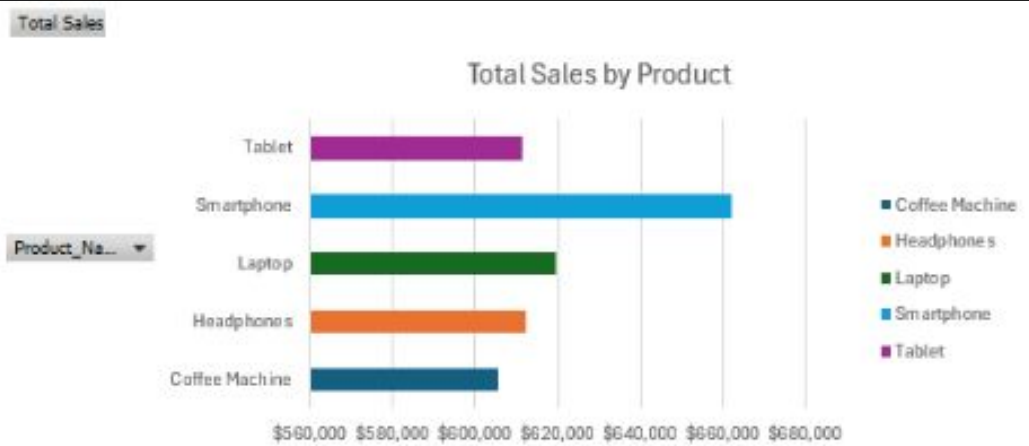
Transaction ID:	T7685
Transaction Status:	Completed

Cleaned Spreadsheet

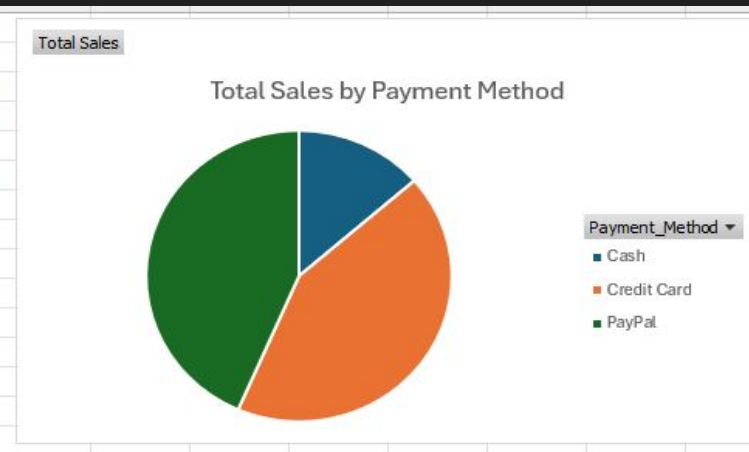
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Transaction_ID ▾	Transaction_Date ▾	Day ▾	Month ▾	Year ▾	Customer_ID ▾	Product_Name ▾	Quantity ▾	Price ▾	Payment_Method ▾	Transaction_Status ▾			
2	T7685	1/1/2020	Sunday	January	2020	C257	Coffee Machine	9	945.87	Credit Card	Completed		Transaction ID:	T7685
3	T22312	1/1/2020	Sunday	January	2020	C687	Tablet	973	848.49	Cash	UNKNOWN		Transaction Status:	Completed
4	T39156	1/1/2020	Sunday	January	2020	C4666	Coffee Machine	1	800.94	Credit Card	Failed			
5	T66966	1/1/2020	Sunday	January	2020	C4761	Headphones	943	495.27	PayPal	Completed			
6	T68111	1/1/2020	Sunday	January	2020	C2744	Smartphone	7	857.74	Credit Card	Completed			
7	T19057	1/2/2020	Monday	January	2020	C832	Coffee Machine	913	270.64	Cash	Completed			
8	T43708	1/2/2020	Monday	January	2020	C3518	Headphones	5	346.06	Credit Card	Completed			
9	T96200	1/3/2020	Tuesday	January	2020	C1860	Coffee Machine	747	551.78	PayPal	Completed			
0	T99760	1/3/2020	Tuesday	January	2020	C779	Headphones	1	145.72	PayPal	Completed			
1	T24601	1/4/2020	Wednesday	January	2020	C2008	Smartphone	9	992.19	Credit Card	Pending			
2	T61388	1/4/2020	Wednesday	January	2020	C2081	Coffee Machine	9	733.07	Credit Card	Completed			
3	T62952	1/4/2020	Wednesday	January	2020	C720	Headphones	149	891.75	Credit Card	Completed			
4	T10772	1/5/2020	Thursday	January	2020	C3103	Smartphone	3	879.61	PayPal	Completed			
5	T29333	1/5/2020	Thursday	January	2020	C2555	Smartphone	870	597.76	PayPal	Completed			
6	T66896	1/5/2020	Thursday	January	2020	C4993	Headphones	270	939.22	Cash	Failed			
7	T89918	1/5/2020	Thursday	January	2020	C405	Tablet	8	847.25	Cash	Completed			
8	T91910	1/5/2020	Thursday	January	2020	C188	Laptop	4	611.21	Credit Card	Failed			
9	T46447	1/6/2020	Friday	January	2020	C1392	Laptop	3	249.10	Cash	Failed			
0	T63088	1/7/2020	Saturday	January	2020	C4407	Tablet	456	591.39	Credit Card	Completed			
1	T82327	1/7/2020	Saturday	January	2020	C3719	Smartphone	297	462.77	Credit Card	Pending			
2	T41989	1/8/2020	Sunday	January	2020	C2864	Tablet	1	92.33	Credit Card	Completed			
3	T43374	1/8/2020	Sunday	January	2020	C2792	Tablet	766	78.23	PayPal	Failed			
4	T47954	1/8/2020	Sunday	January	2020	C551	Headphones	2	662.65	PayPal	Completed			
5	T62244	1/8/2020	Sunday	January	2020	C2776	Smartphone	794	817.24	Credit Card	Completed			
6	T91016	1/8/2020	Sunday	January	2020	C2016	Headphones	5	227.42	PayPal	Completed			

Analysis

Product	Total Sales
Coffee Machine	605,457.14
Headphones	612,402.51
Laptop	619,303.17
Smartphone	662,036.11
Tablet	611,134.64
Grand Total	3,110,333.57

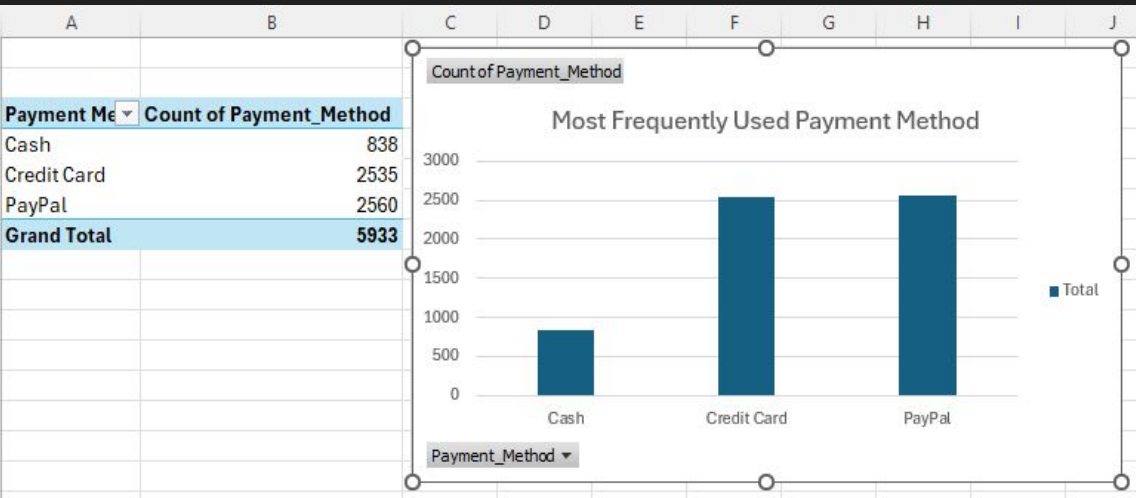


Payment Method	Total Sales
Cash	423,656.56
Credit Card	1,334,556.34
PayPal	1,352,120.68
Grand Total	3,110,333.57



Analysis

Month	Total Sales
January	312,621.92
February	241,344.74
March	236,232.82
April	245,124.25
May	274,730.26
June	266,782.84
July	244,233.10
August	243,052.55
September	256,557.37
October	261,235.66
November	252,373.29
December	276,044.79
Grand Total	3,110,333.57

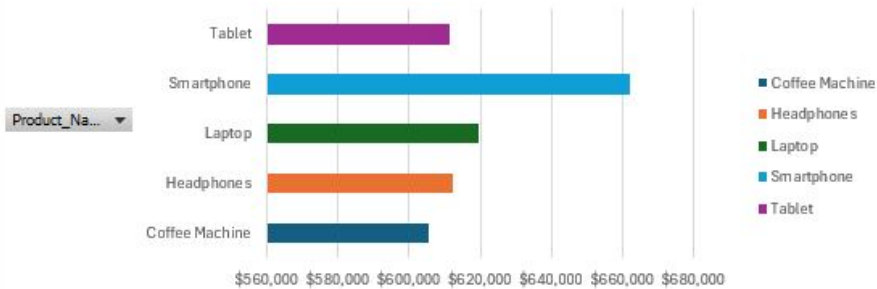


Dashboard

Financial Transactions Dashboard

Total Sales

Total Sales by Product



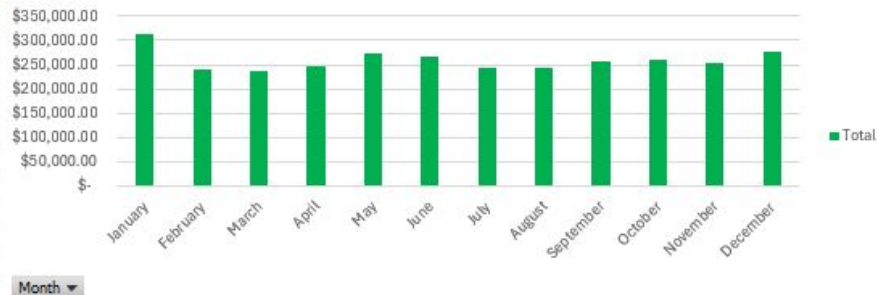
Total Sales

Total Sales by Payment Method



Total Sales

Total Sales By Month



Count of Payment_Method

Most Frequently Used Payment Method



Key Insights

- Smartphones had the highest Total Sales by a large margin, showing strong customer demand.
- Credit Card and PayPal were the dominant payment methods in both total sales and frequency, indicating preference for digital transactions over cash.
- January and December saw the highest total sales, suggesting seasonal shopping trends or holiday effects.

Conclusion

Throughout this project, I cleaned and transformed a messy dataset using tools and techniques such as Power Query, formulas like SUM, DAY, MONTH, and YEAR. Used other tools such as XLOOKUP, Pivot Tables, sorting/filtering, and Find & Replace. These steps improved data quality and allowed me to uncover clear trends in sales and customer behavior. The final dashboard highlights valuable insights, such as product popularity, preferred payment methods, and seasonal sales patterns. This process shows how effective data cleaning and analysis can turn raw data into useful business intelligence.