

**DataSpark: Illuminating Insights for Global Electronics**

Retail Analytics in the Electronics Industry

SUBHIKSHA | MDTM28 |

# CONTENT

1. OBJECTIVE
2. DATA CLEANING AND PREPARATION
3. LOAD DATA
4. POWER BI VISUALIZATION
5. SQL QUERIES
6. SCREENSHOTS – POWER BI

1.OBJECTIVE:

From Global Electronics datasets, focusing on enhancing customer satisfaction, optimizing operational efficiency, and driving business growth. Specifically, we aim to:

1. Identify Key Customer Segments: Analyze customer data to tailor marketing strategies that resonate with diverse demographic and behavioral profiles.
2. Optimize Inventory Management: Evaluate product sales patterns to inform better inventory practices, minimizing stockouts and overstock situations.
3. Enhance Sales Forecasting: Utilize historical sales data to develop predictive models that improve the accuracy of sales forecasts, aiding in strategic planning.
4. Evaluate Marketing Campaign Effectiveness: Assess the impact of past promotions and marketing initiatives on sales performance to refine future campaigns.
5. Strategize Store Operations: Analyze store performance metrics to identify opportunities for expansion, optimization, or closures based on local market dynamics.
6. Assess Currency Exchange Impact: Investigate how fluctuations in currency exchange rates affect sales, enabling informed pricing strategies for international markets.

By achieving these objectives, we will equip Global Electronics with the insights necessary to make data-driven decisions that enhance customer experiences and foster sustainable growth

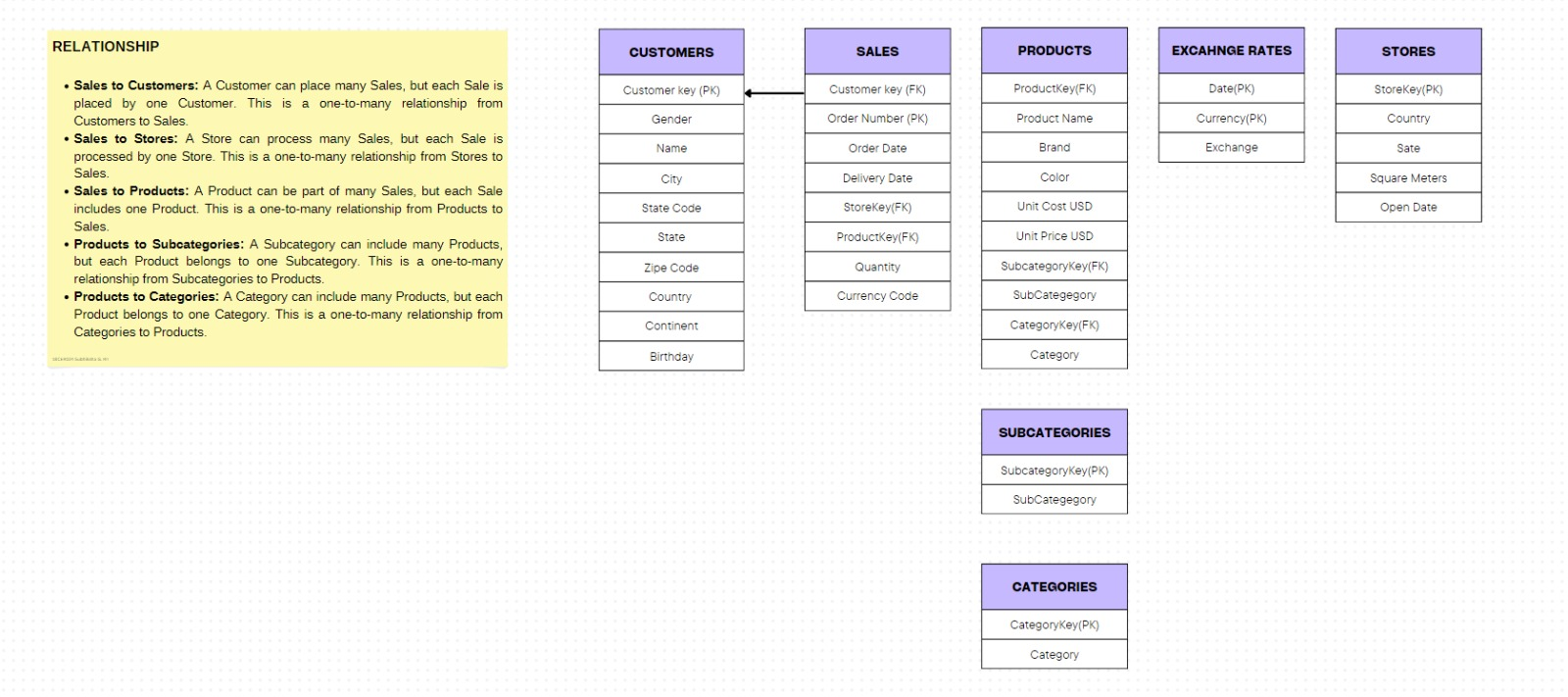
2.DATA CLEANING AND PREPARATION:

At first we have 5 CSV as data set

**DATA SET:**

<https://drive.google.com/drive/folders/1Ag9pAbXmPtBVz6pDHa_yGI_wzK61TTao>

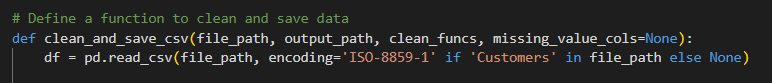
* Customers.csv – 10 columns
* Products.csv
* Sales.csv
* Stores.csv
* Exchange\_rates. csv



**DATA SET PROCESSING:**

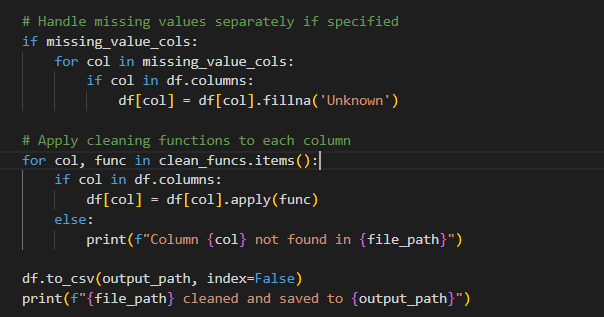
1. Analyzed the data and converted into DataFrame:

Analyzed the structure of the dataset.



This line reads the CSV file specified by file\_path and creates a DataFrame named “df”.

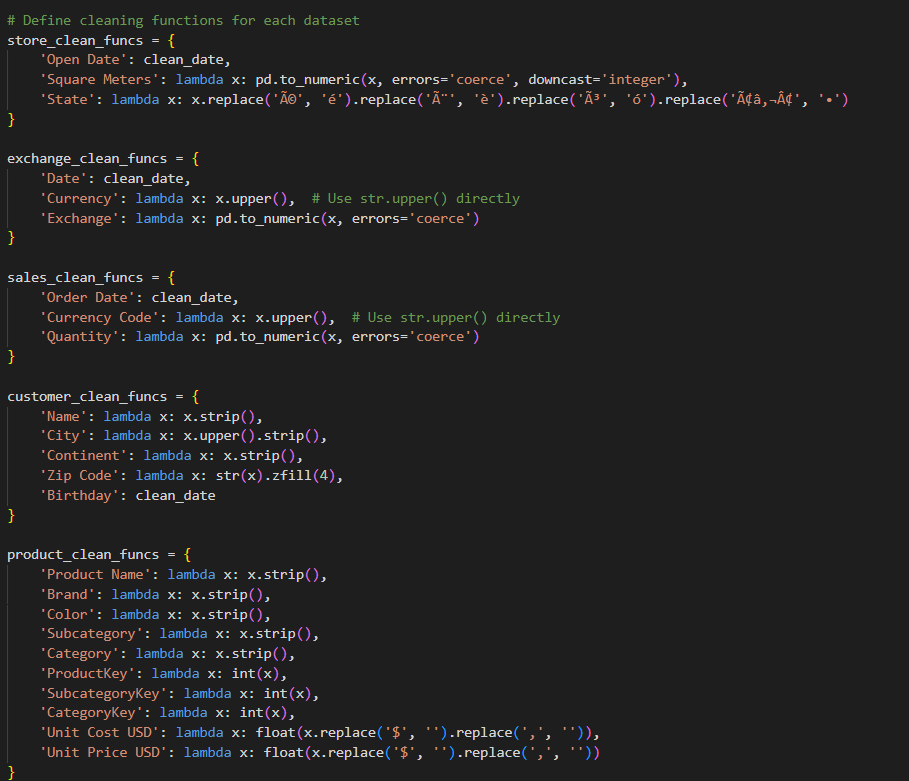
1. Handled the missing values – “Unknown” and saved that file:



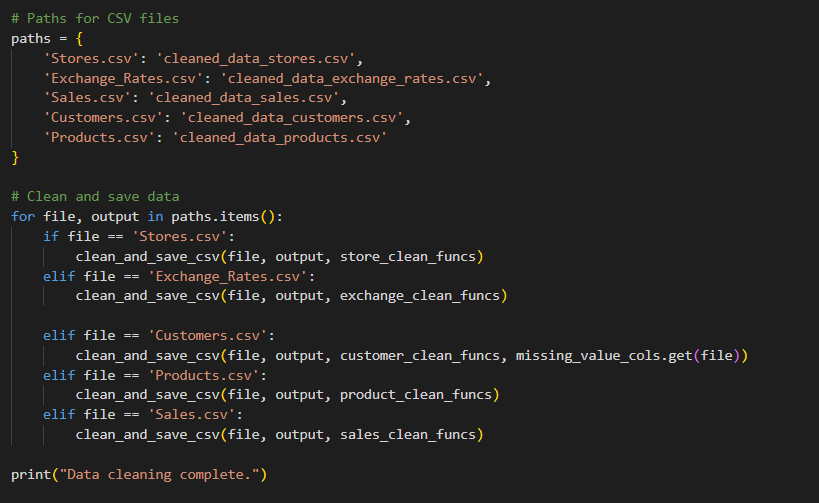
In customer.csv – The gender and the State columns are handled with missing values.

1. Formatting particular columns:

I have divided the formatting into three stages

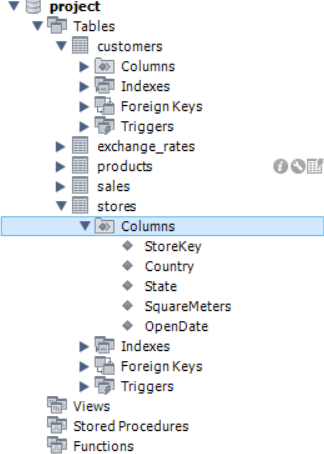
1. Formatted accordingly – States, Date, Currency, Country Code, Birthday, Unit Cost USD, Unit Price USD.
2. Formatted to numeric – Exchange, Sq.mts, SubcategoryKey, Category, Quantity, Productkey and Category Key
3. Formatted using strip – Name, City, ProductName, Continent, Brand, Color, Category and Subcategory.
4. Formatted using strip and replace – ZipCode
5. Cleaned and saved to new CSV files:

After performing the various cleaning operations on this DataFrame it is saving it back to a new CSV file



3.LOAD DATA:

By using create\_engine from the sqlalchemy the preprocessed cleaned datafiles are loaded into the MySQL Database with the database name as “project”

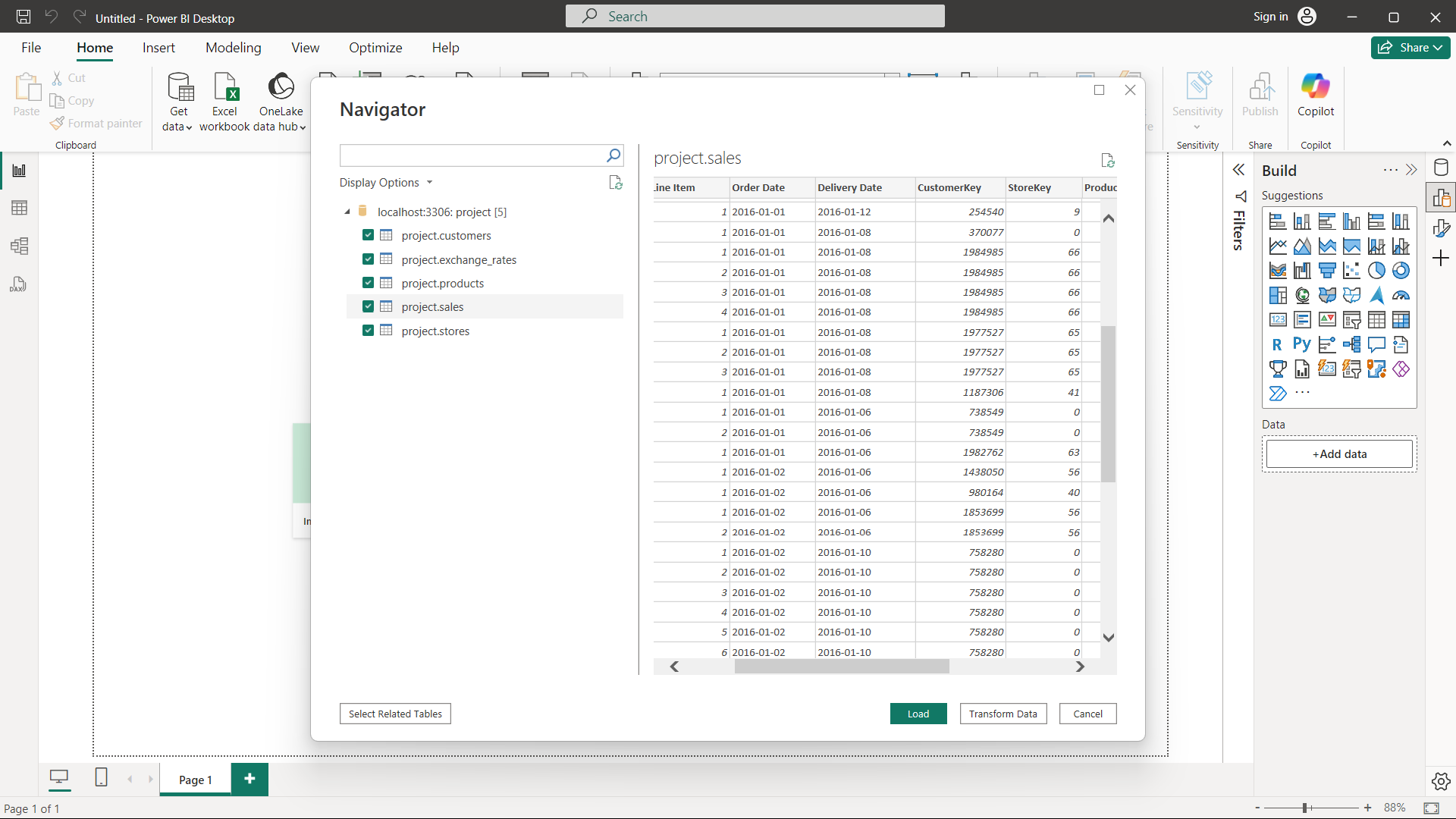


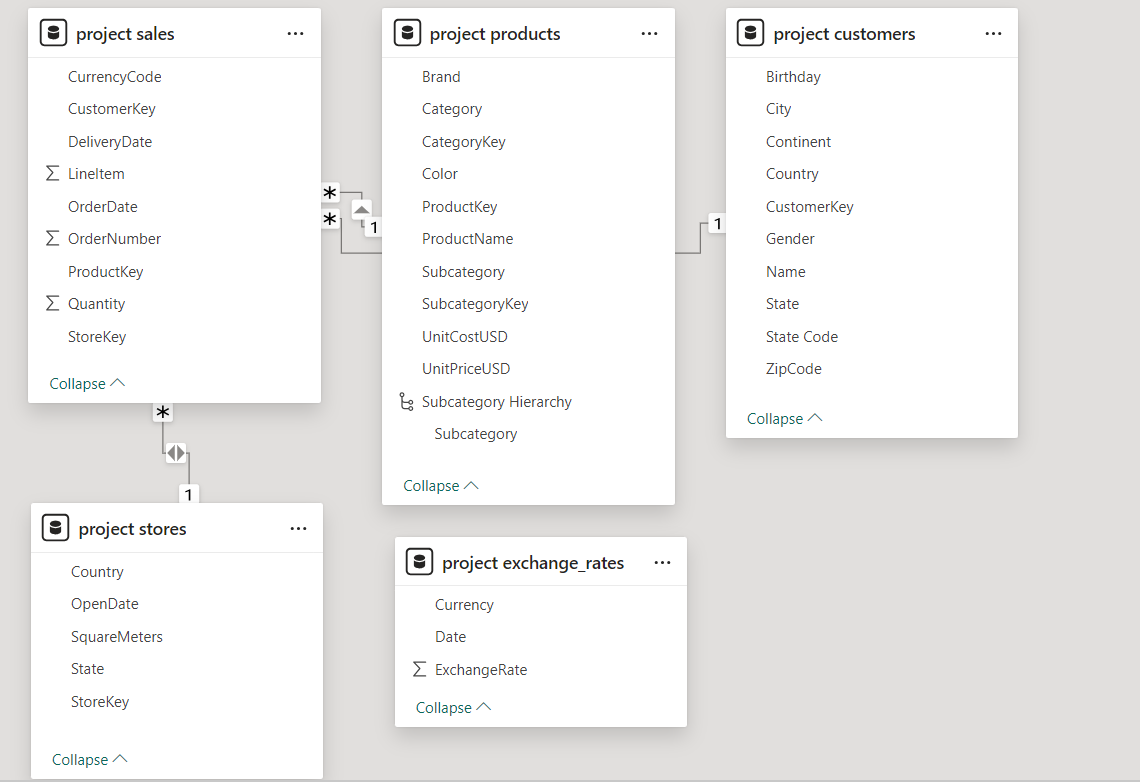
4. POWER BI VISUALIZATION:

Inserted the preprocessed data into an SQL database by creating relevant tables

GET DATA ----- DATABASES ----- MYSQL DATABASE ----- LOAD THE DATA ------

By using the Server name and the Database name





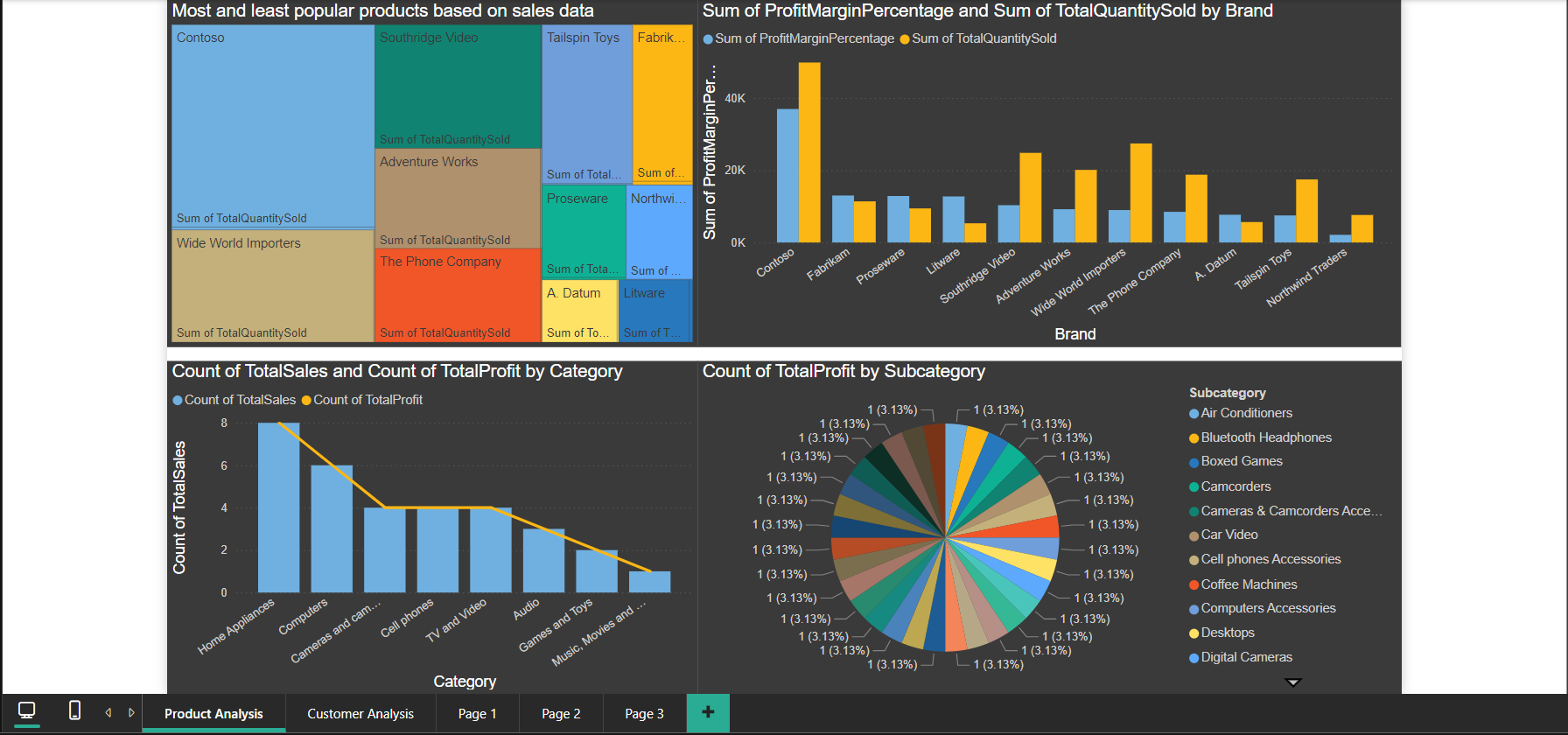
5. SQL QUERIES:

[SQL QUERIES IN EXCEL.xlsx](SQL%20QUERIES%20IN%20EXCEL.xlsx)

6. SCREENSHOTS:

PRODUCT ANALYSIS: - BUSINESS OUTCOMES

* Product Popularity: Identify the most and least popular products based on sales data.
* Profitability Analysis: Calculate profit margins for products by comparing unit cost and unit price.
* Category Analysis: Analyze sales performance across different product categories and subcategories.



CUSTOMER ANALYSIS: - BUSINESS OUTCOMES

* Demographic Distribution: Analyze the distribution of customers based on gender, age (calculated from birthday), location (city, state, country, continent).
* Purchase Patterns: Identify purchasing patterns such as average order value, frequency of purchases, and preferred products.
* Segmentation: Segment customers based on demographics and purchasing behavior to identify key customer groups.

