Requirements Document

Project Title

Excel-Based Data Analysis Project: E-commerce Consumer Behavior Analysis

Project Overview

E-commerce businesses collect vast amounts of consumer data, but often struggle to derive clear, actionable insights from it. Without a structured analytical approach, marketing and sales teams may miss key trends and opportunities for growth. This project will use a consumer behavior dataset and leverage Microsoft Excel as the sole tool for the entire data analysis lifecycle. The goal is to clean, analyze, and visualize the data to build an interactive dashboard that uncovers customer patterns, identifies high-value segments, and supports data-driven decision-making for business optimization.

Problem Statement

Currently, there is no consolidated view of how different customer demographics and purchasing behaviors impact key business metrics like sales and customer satisfaction. The business lacks the tools to easily analyze which combinations of factors (e.g., age, income, purchase channel) lead to the best outcomes, resulting in generic marketing efforts and missed opportunities. An Excel-based solution is required to integrate the dataset, calculate KPIs, and deliver an interactive dashboard that allows for the identification of high-performing customer segments and purchasing patterns.

Objectives

- Clean and structure the raw dataset within Excel to create a reliable master table for analysis.

- Develop an interactive dashboard on a dedicated Excel sheet using PivotCharts, Slicers, and Timelines for dynamic data exploration.

- Identify which demographic segments (age, gender, income level) drive the most sales and have the highest satisfaction.

- Analyze purchasing behavior by product category, purchase channel, and time-based trends (month, day of the week).
- Analyze the relationship between factors like product ratings, discount usage, and customer satisfaction.
- Document all data cleaning and transformation steps for clarity and reproducibility.
- Deliver a final report summarizing key insights and providing actionable business recommendations.

Functional Scope

- Data Cleaning and Preparation: Utilize Excel functions (e.g., TRIM, IF, VLOOKUP), Power Query, and features like "Remove Duplicates" to clean and structure the data.
- Data Analysis: Employ PivotTables to aggregate, summarize, and analyze the cleaned data across multiple dimensions.
- Data Visualization: Create various PivotCharts (bar charts, line charts, pie charts, histograms) to visualize trends and patterns.
- Interactive Reporting: Develop a centralized Dashboard sheet that integrates charts and KPIs. Implement Slicers and Timelines to allow for interactive filtering and drill-down analysis.

Functional Requirements

- Analyze purchase amounts and frequency by demographic attributes (age group, gender, income level).
- Compare the performance of different purchase channels (Online, In-Store, Mixed) in terms of sales volume and average purchase value.
- Track key metrics like total sales, average purchase amount, and average customer satisfaction.
- Analyze the relationship between `Product_Rating` and `Customer_Satisfaction`.
- Identify peak purchasing times by month and day of the week.

- Enable dynamic filtering of the entire dashboard by demographic groups, product categories, and date ranges.

Non-Functional Requirements

- **Usability:** The dashboard should be intuitive and easy for a non-technical user to filter and interpret.

- **Performance:** The Excel file should remain responsive, with Slicer interactions updating dashboard visuals in under 3 seconds.

- **Accessibility:** The final workbook will be saved in a standard .xlsx format for broad compatibility.

- **Maintainability:** The workbook must be organized with clearly labeled sheets (e.g., "Raw Data," "Cleaned Data," "Pivot Tables," "Dashboard"). A "Documentation" sheet should outline the cleaning steps and define each data field.

Technical Constraints

- Tools: This project must be completed using **only Microsoft Excel**. No other software will be used.

- The project must follow Excel best practices, including separating data, analysis, and presentation into different worksheets.

- **Architecture:** A logical "three-sheet" architecture will be used:
1. **Data Layer:** Sheets for raw and cleaned data.
2. **Analysis Layer:** A sheet to hold the PivotTables driving the dashboard.
3. **Presentation Layer:** The final interactive Dashboard sheet.

Methodology

1. Data Preprocessing

- Import the source CSV dataset into an Excel sheet named "Raw Data".

- Clean, standardize, and transform fields on a new "Cleaned Data" sheet using Excel functions and/or Power Query.

- Engineer new columns as needed (e.g., "Age Group," "Purchase Month") to enrich the data for analysis.

2. Analysis with PivotTables

- Create a new sheet named "Pivot Tables."
- Use the "Cleaned Data" sheet as the source to build multiple PivotTables that aggregate data to meet the functional requirements (e.g., Sales by Age Group, Average Satisfaction by Product Rating).

3. Dashboarding and Reporting

- Create a final sheet named "Dashboard."
- Design and build interactive PivotCharts based on the PivotTables created in the previous step.
- Organize the charts into a logical and visually appealing layout.
- Insert and connect Slicers and Timelines to the PivotCharts to enable full interactivity.
- Add a summary section with key insights and actionable recommendations to guide business strategy.