Amazon Sales Data Analysis Project Overview

This document provides a concise overview of the **Amazon Sales Data Analysis Project**, which leverages **Microsoft Excel** for robust data preparation and **Microsoft Power BI** for interactive visualization and insightful dashboard creation.

1. Project Goal

The primary objective of this project was to analyze a dataset of Amazon sales to identify key trends, performance metrics, and actionable insights that could inform business strategies and decision-making.

2. Data Source

The raw sales data for this project was sourced from **Kaggle**, a prominent platform for data science datasets.

3. Methodologies & Tools Used

3.1. Data Cleaning and Preparation (Microsoft Excel)

A crucial phase of this project involved extensive **data cleaning and transformation performed entirely within Microsoft Excel**. This focused on:

- **Handling Missing Values:** Identifying and addressing incomplete data entries.
- **Data Consistency:** Ensuring uniform formatting and removing discrepancies.
- **Feature Engineering:** Creating new variables or modifying existing ones to enhance analytical depth.

The use of Excel for these tasks not only strengthened advanced spreadsheet skills but also demonstrated an ability to work effectively with various data manipulation tools. Furthermore, **ChatGPT was utilized as an AI assistant** during this phase, proving invaluable for generating complex Excel formulas and optimizing data cleaning workflows, thereby increasing efficiency.

3.2. Data Visualization and Dashboard Creation (Microsoft Power BI)

Following thorough data preparation, the cleaned dataset was imported into **Microsoft Power BI**. Here, an **interactive and dynamic dashboard** was created to visually represent key aspects of the Amazon sales data. The dashboard includes detailed visualizations for:

- **Sales Performance Metrics:** Overall **sum of sales** and **sum of profit by state**, alongside other critical volume indicators.

- **Product & Category Performance: Sum of sales by category** to highlight top-performing product groups.
- **Geographical Analysis:** Insights into sales distribution across different regions.
- **Time-Series Analysis:** Comprehensive sales trends broken down by **year, quarter, month, and specific dates** to show granular performance over time.

The dashboard is designed to allow users to explore data dynamically, providing a comprehensive view of sales operations and opportunities.

4. Key Outcomes & Skills Demonstrated

- **Proficiency in Data Cleaning:** Demonstrated advanced skills in data manipulation and validation using Excel.
- **Interactive Dashboard Design:** Ability to create intuitive and insightful dashboards in Power BI.
- **Data-Driven Insights:** Extracted meaningful business intelligence from complex datasets.
- **Workflow Optimization:** Showcased adaptability and efficiency by integrating AI tools (ChatGPT) into the data preparation process.
- **End-to-End Analysis:** Completed a full data analysis lifecycle from raw data acquisition to final visualization.