**A Micro Project report on**

**Marketing Analytics Platform Using Predictive Models.**

Submitted to the **CMR Institute of Technology** in partial fulfillment of the requirement of the

**AUTOMATED WRITING TOOLS - ChatGPT**

**Of**

**III-B.Tech. I-Semester**

In

**Computer Science Department**

Submitted by

| **AKASH YAMAGANI** | **22R01A0562** |
| --- | --- |
| **PRANAY YASA** | **22R01A0563** |
| **AKANKSHA** | **22R01A0565** |
| **SHRAVANI AKITI** | **22R01A0566** |

Under the Guidance of

**MR. KUNAL GAIKWAD**

(ASSISTANT PROFESSOR)



**CMR INSTITUTE OFTECHNOLOGY**

(UGC AUTONOMOUS**)**

**(Approved by AICTE and Affiliated to JNTU, Kukatpally**

**Kandlakoya, Medchal Road**

**CMR INSTITUTE OF TECHNOLOGY**

**(UGC-AUTONOMOUS)**

(Approved by AICTE, Affiliated to JNTU, Hyderabad) Kandlakoya, Medchal Road, Hyderabad.

**Department of Computer Science and Engineering**



**CERTIFICATE**

This is to certify that a Micro Project entitled with **“Marketing Analytics Platform Using Predictive Models”** is the bonafide work carried out and submitted by

| **AKASH YAMAGANI** | **22R01A0562** |
| --- | --- |
| **PRANAY YASA** | **22R01A0563** |
| **AKANKSHA** | **22R01A0565** |
| **SHRAVANI AKITI** | **22R01A0566** |
|  |  |

In partial fulfillment of the requirement for completion of the  **AUTOMATED WRITING TOOLS – ChatGPT** lab of III– B.Tech I – Semester is a record of Bonafide work carried out under our guidance during the academic year 2024-2025.

.

PROJECT GUIDEHEAD OF THE DEPARTMENT

**Mr. Kunal Gaikwad Mr. A. Prakash**

Assistant Professor Head of Department

**ACKNOWLEDGEMENT**

We are extremely grateful to **Dr. M. Janga Reddy , Director , Dr. B. Satyanarayana ,Principal and Mr. A. Prakash , Head of Department** , Dept of Computer Science and Engineering , CMR Institute of Technology for their inspiration and valuable guidance during entire duration.

We are extremely thankful to our **AUTOMATED WRITING TOOLS - ChatGPT** faculty in-charge **Mr. Kunal Gaikwad** , **Assistant Professor** , Computer Science and Engineering department , CMR Institute of Technology for their constant guidance , encouragement and moral support throughout the project.

We express our thanks to all staff members and friends for all the help and coordination extended in bringing out this Project successfully in time.

Finally, we are very much thankful to our parents and relatives who guided directly or indirectly for successful completion of the project.

| **AKASH YAMAGANI** | **22R01A0562** |
| --- | --- |
| **PRANAY YASA** | **22R01A0563** |
| **AKANKSHA** | **22R01A0565** |
| **SHRAVANI AKITI** | **22R01A0566** |

**1.INTRODUCTION**

Predictive analytics has become a cornerstone in modern data-driven decision-making, especially within the marketing sector. It involves the use of **historical data** and **statistical algorithms** to forecast future trends and behaviours, thereby providing valuable insights for businesses. Through predictive models, organizations can enhance decision-making, optimize marketing strategies, and drive operational efficiencies. This methodology is particularly beneficial for predicting **customer behaviour**, **sales performance**, and **market trends**, enabling businesses to proactively adjust their strategies and allocate resources more effectively.

In marketing, predictive analytics plays a critical role in understanding how various factors, such as **promotional campaigns product pricing**, and **market condition** impact sales performance. By applying the appropriate predictive models, companies can identify which marketing strategies are most likely to generate positive sales outcomes and customer engagement. This capability gives businesses a competitive advantage in a dynamic market environment.

The primary objective of this project is to develop a Marketing Analytics Platform that employs predictive models to forecast sales outcomes based on multiple features, such as promotion type, market size, and sales campaigns. Key goals of the platform include:

- Predicting future sales trends and values

- Identifying significant factors that influence sales performance

- Providing actionable insights to refine marketing strategies

By utilizing machine learning algorithms, this platform aims to deliver more accurate sales forecasts and help businesses make informed, data-backed decisions. The following sections will provide a detailed walkthrough of the steps involved in building and evaluating the predictive model, as well as the insights generated from the dataset.

**2.DATA OVERVIEW**

The dataset titled WA\_Marketing-Campaign.csv provides a comprehensive overview of sales data influenced by various marketing promotions across different markets. It captures crucial information regarding sales performance and market characteristics, which are essential for understanding the impact of marketing strategies. Below is a detailed description of the dataset, its structure, and the key features it contains.

1. **Description of the Dataset**

This dataset captures detailed sales performance data that is segmented by market characteristics and promotional strategies. It encompasses the following key components:

- SalesInThousands: This is the target variable, representing the sales figures in thousands of dollars for each market in a given week. It serves as the dependent variable in predictive modeling, reflecting the outcome of marketing efforts.

- MarketSize: A categorical variable that indicates the size classification of the market. The possible values include Small, Medium, and Large markets. The size of the market may influence the overall sales performance, as larger markets could have more customers and demand.

- LocationID: A unique numeric identifier for each market location. This identifier ensures that each market’s sales data can be tracked and analyzed individually. It provides insight into location-specific sales patterns.

- AgeOfStore: This numerical variable represents the age of the store in years. It gives insight into the establishment’s longevity and may reflect the store’s brand recognition or customer loyalty, both of which can influence sales outcomes.

- Promotion: A categorical variable indicating different types of promotions applied during the sales period. The variable is encoded numerically, with different integers corresponding to different types of promotions (e.g., discounts, limited-time offers, etc.). Understanding the effect of promotions on sales is crucial for businesses to optimize marketing strategies.

- Week: A numeric representation indicating the week number in which the sales were recorded. This allows for time-series analysis, enabling the identification of trends or patterns in sales across different weeks.

1. **Structure of the Dataset**

The dataset is structured in a tabular format with multiple rows and columns. Each row represents a specific record of sales data for a given market during a particular week. The structure of the dataset is as follows:

- Columns:

- MarketID: Integer – A unique identifier for each market entry.

- MarketSize: Categorical (String) – Classification of the market

- LocationID: Integer – Identifier for specific market locations.

- AgeOfStore: Integer – Number of years since establishment of the store.

- Promotion: Integer – Categorical representation of the promotion type applied.

- Week: Integer – Week number corresponding to the sales data.

- SalesInThousands: Float – Sales figures in thousands.

- Rows: Each row corresponds to a sales record for a particular market and week. This allows for comprehensive analysis over time, as data is segmented by week and market.

1. **Data Types**

The data types for each column are as follows:

- MarketID: Integer

- MarketSize: Categorical (String)

- LocationID: Integer

- AgeOfStore: Integer

- Promotion: Integer

- Week: Integer

- SalesInThousands: Float

The mix of numerical and categorical data types enables various types of analyses, including regression, time-series forecasting, and classification.

1. **Key Features Overview**

1. Promotion Type: The Promotion feature represents different promotional strategies employed during sales periods. These strategies may include discounts, special offers, or seasonal promotions. It is crucial to understand how each promotion type impacts sales performance.

2. Market Size: The MarketSize feature categorizes markets into Small, Medium, and Large. Market size plays a key role in determining the potential customer base and the overall sales volume. Larger markets may exhibit different sales trends compared to smaller markets.

3. Weekly Sales: SalesInThousands records the sales performance for each market, measured on a weekly basis. This feature is critical for time-series analysis and for identifying sales trends across different weeks, promotions, and market conditions.

4. Store Age: The AgeOfStore feature provides insight into how long the store has been operational. Older stores might have more loyal customers and established brand recognition, potentially leading to more stable sales patterns. Conversely, newer stores might show higher growth rates as they establish their customer base.

1. **Significance of the Dataset**

This structured dataset serves as a valuable resource for businesses and analysts seeking to understand the relationship between various marketing strategies and sales outcomes. Through comprehensive analysis of these key features, organizations can identify:

- The effectiveness of different promotional campaigns in driving sales.

- How market size influences overall sales performance.

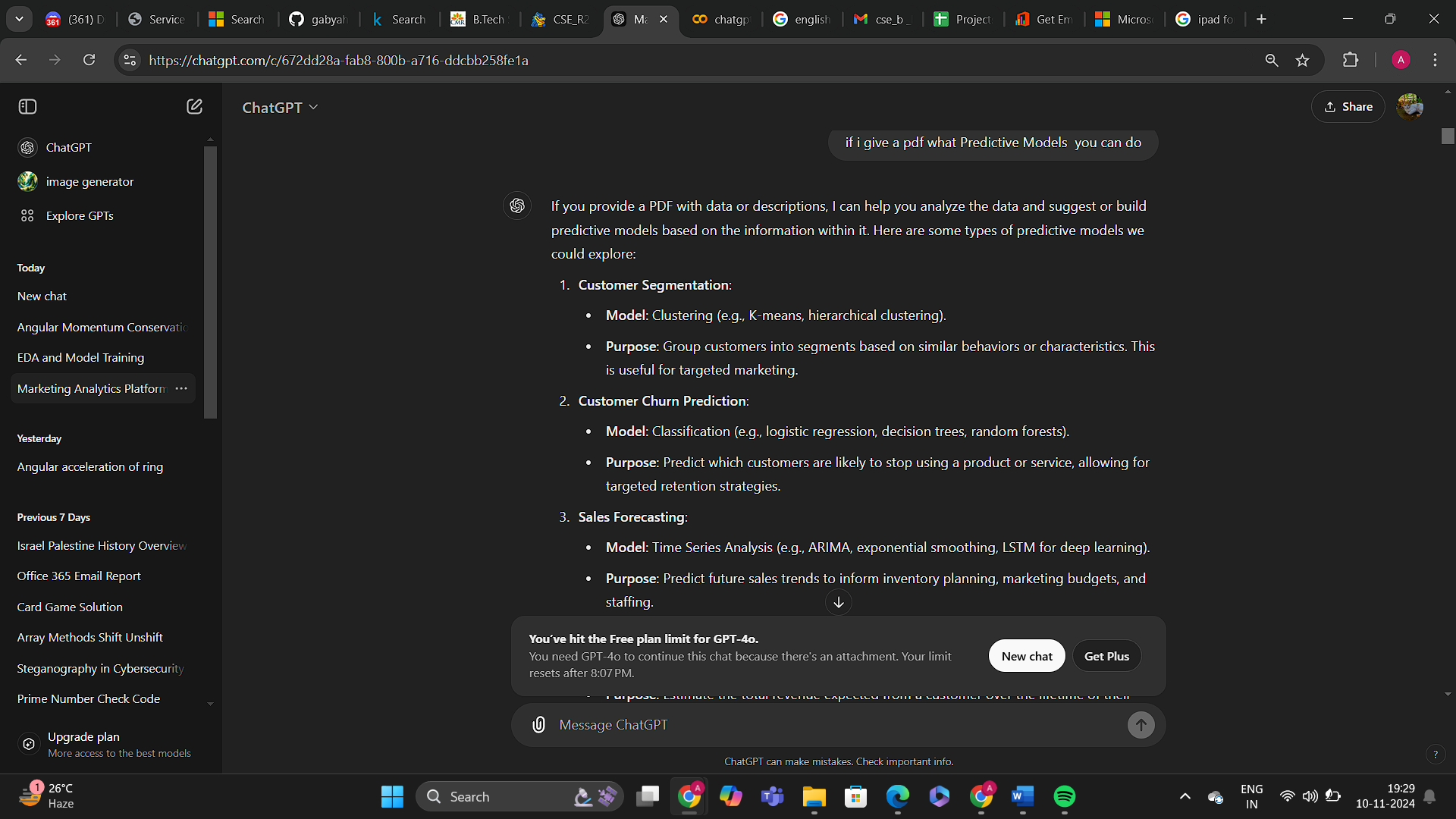
- How the age of a store affects its sales and customer retention.

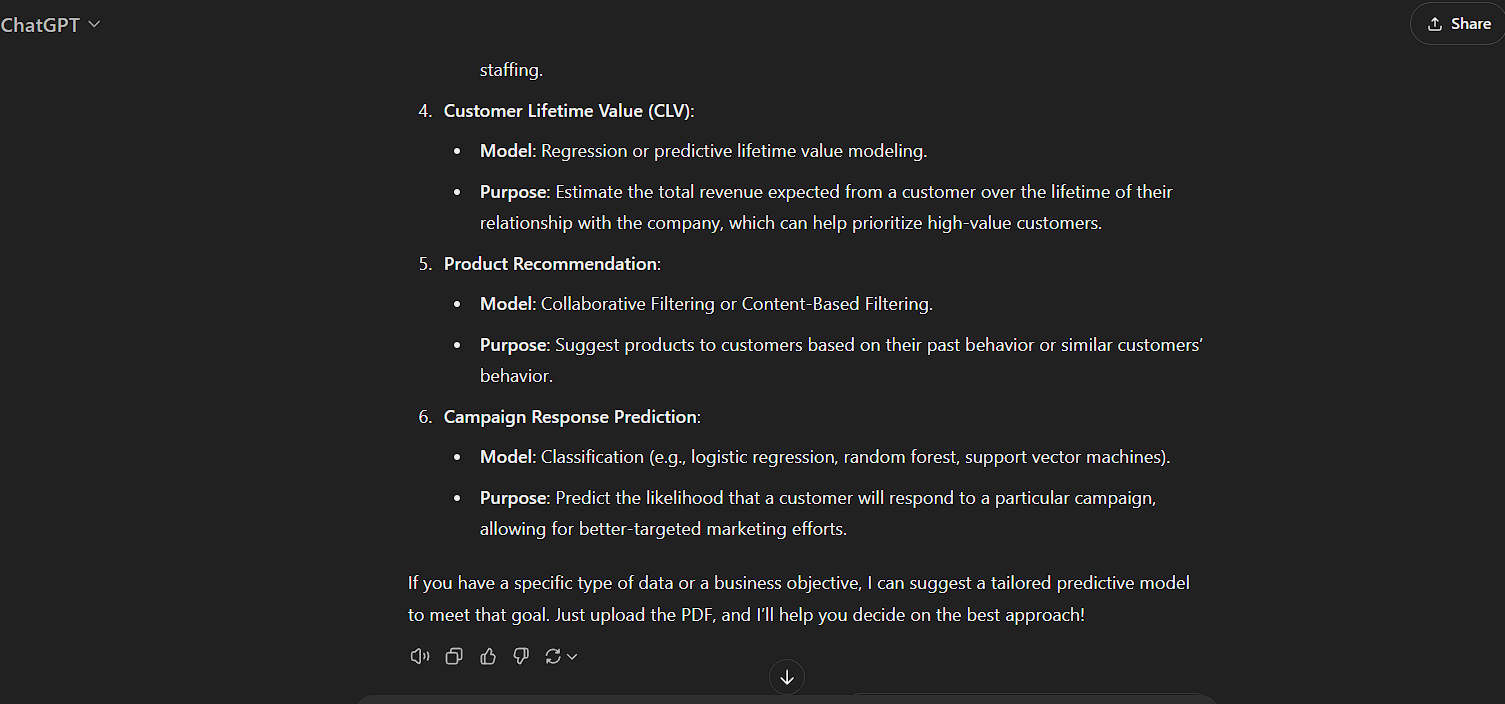
By analyzing these features in combination, businesses can derive actionable insights to refine their marketing strategies, forecast future sales, and optimize promotional efforts across various markets.

**3.PROMPT**

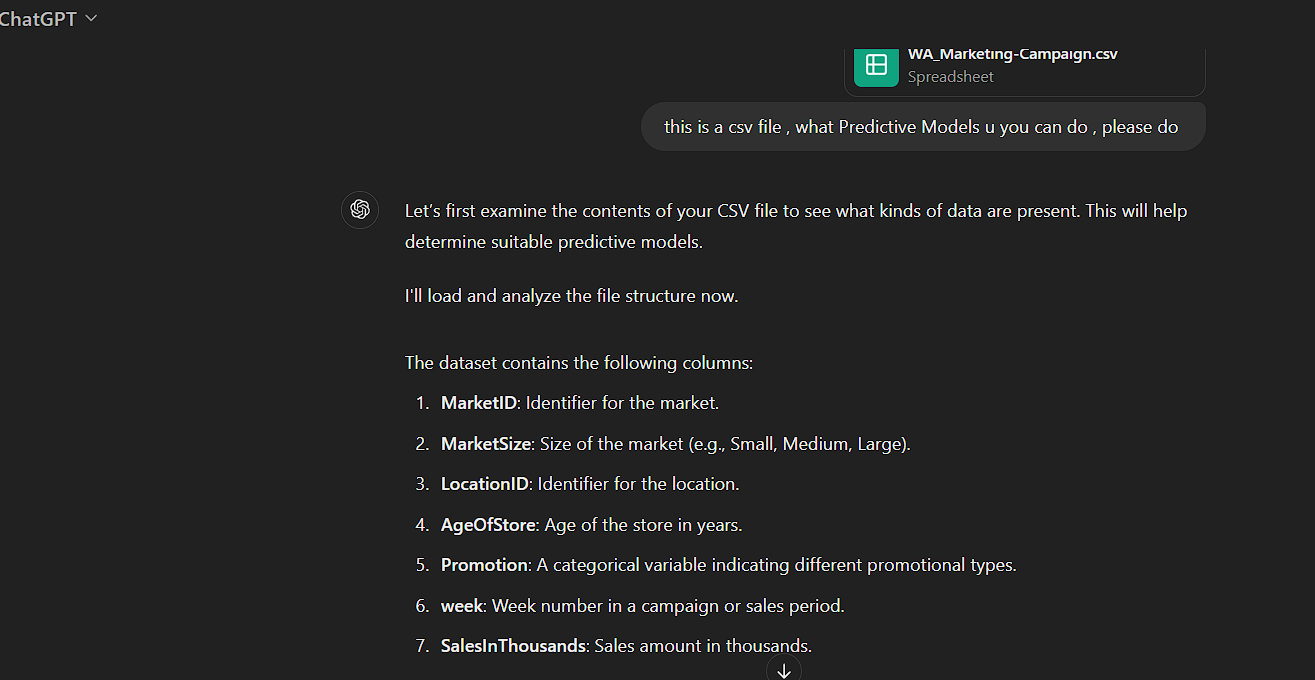
In this section we will see the prompt that we have used to create Marketing Analytics Platform Using Predictive Models and the result that ChatGpt provided us.

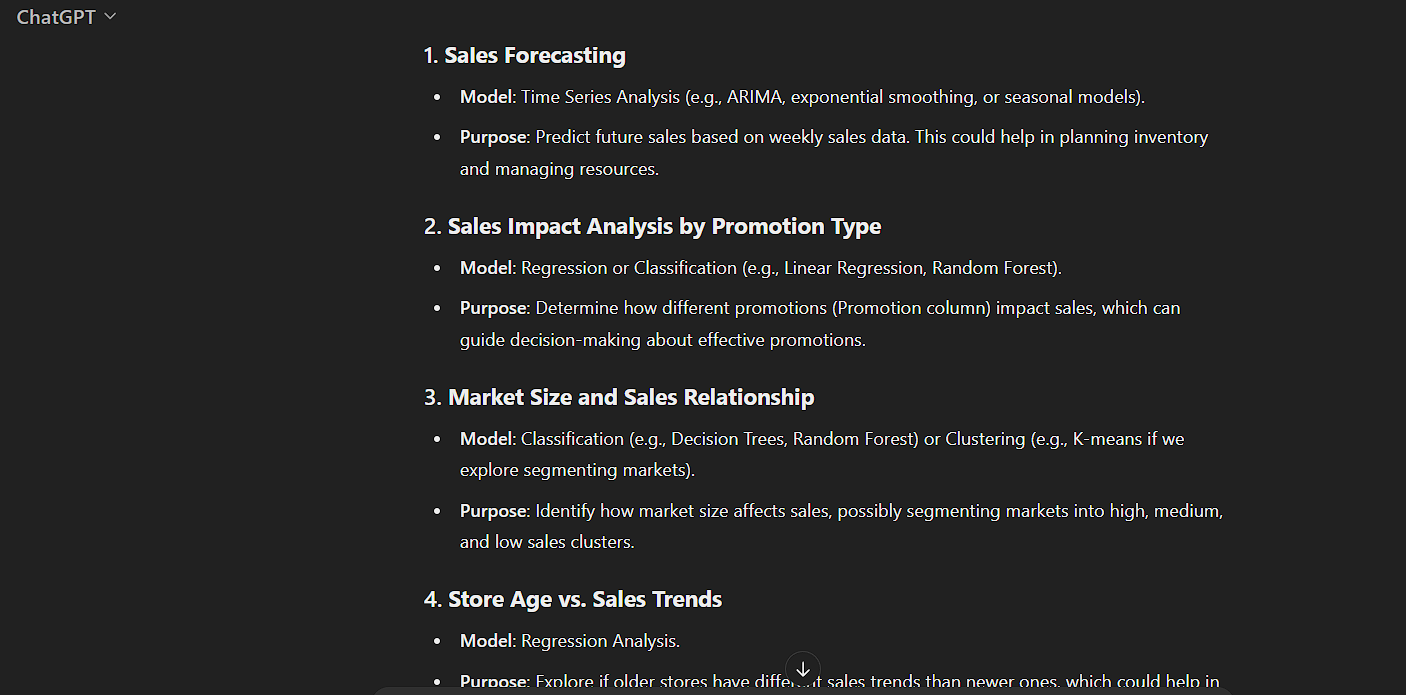
1. Result of Prompt 1:

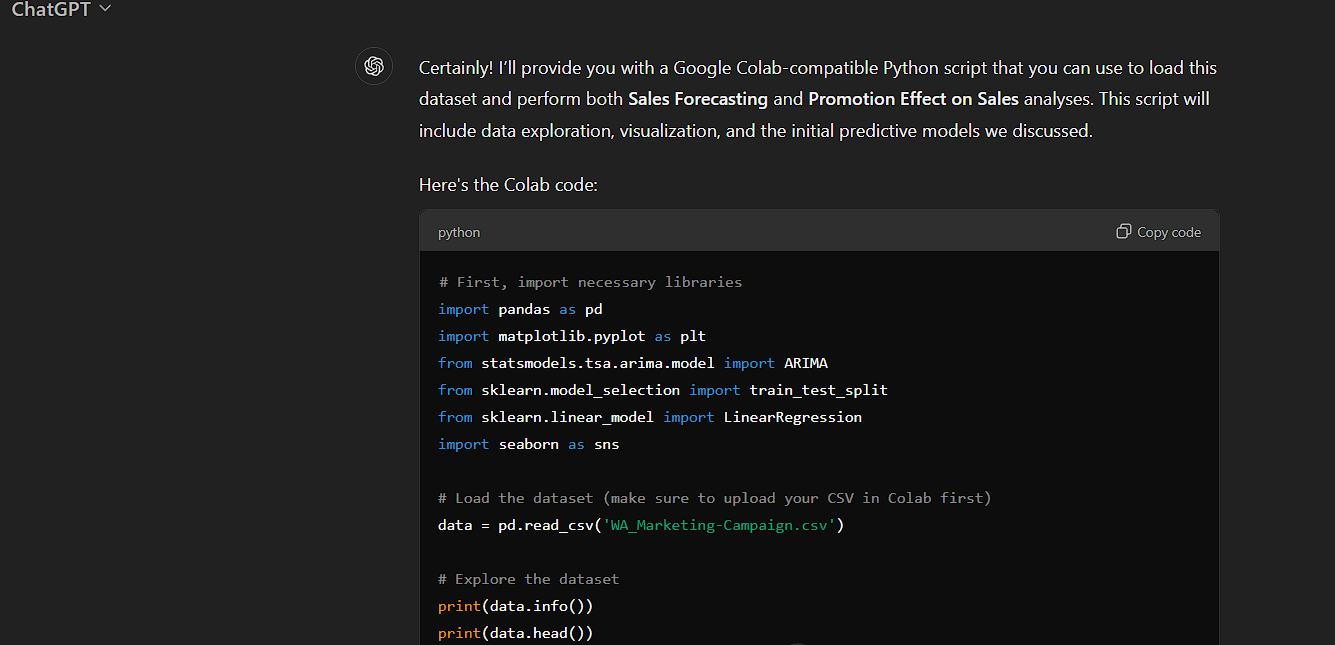
****

****

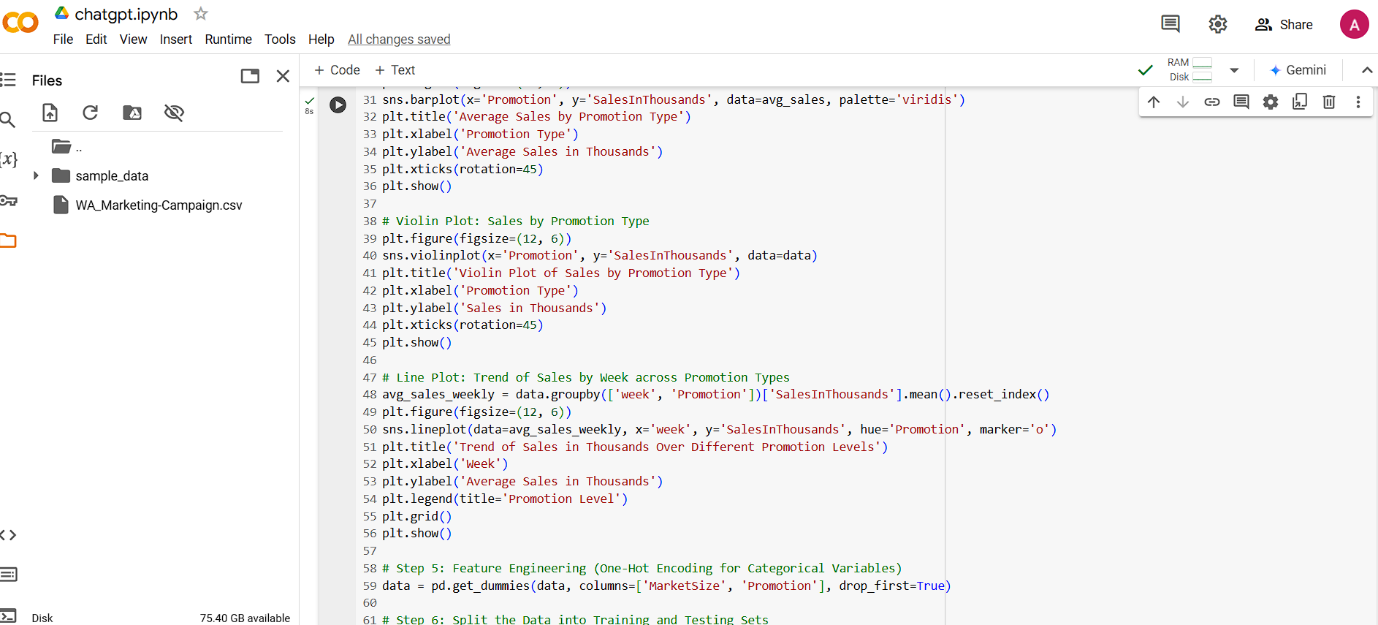
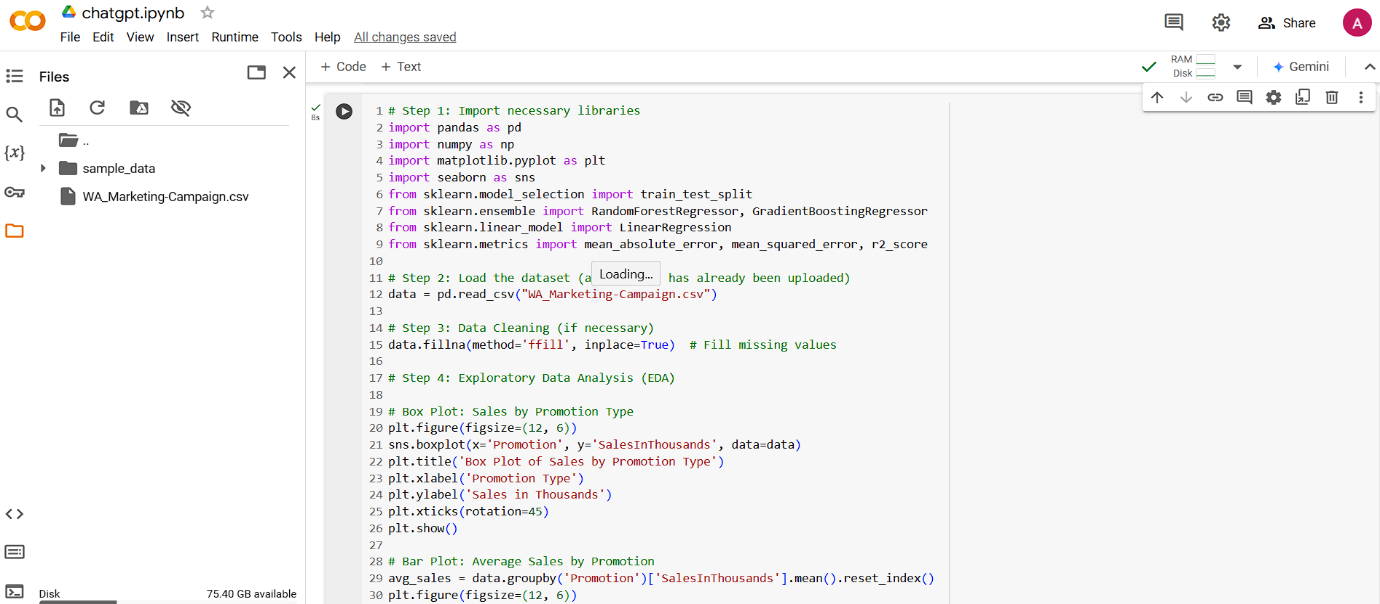
1. Result of Prompt 2:

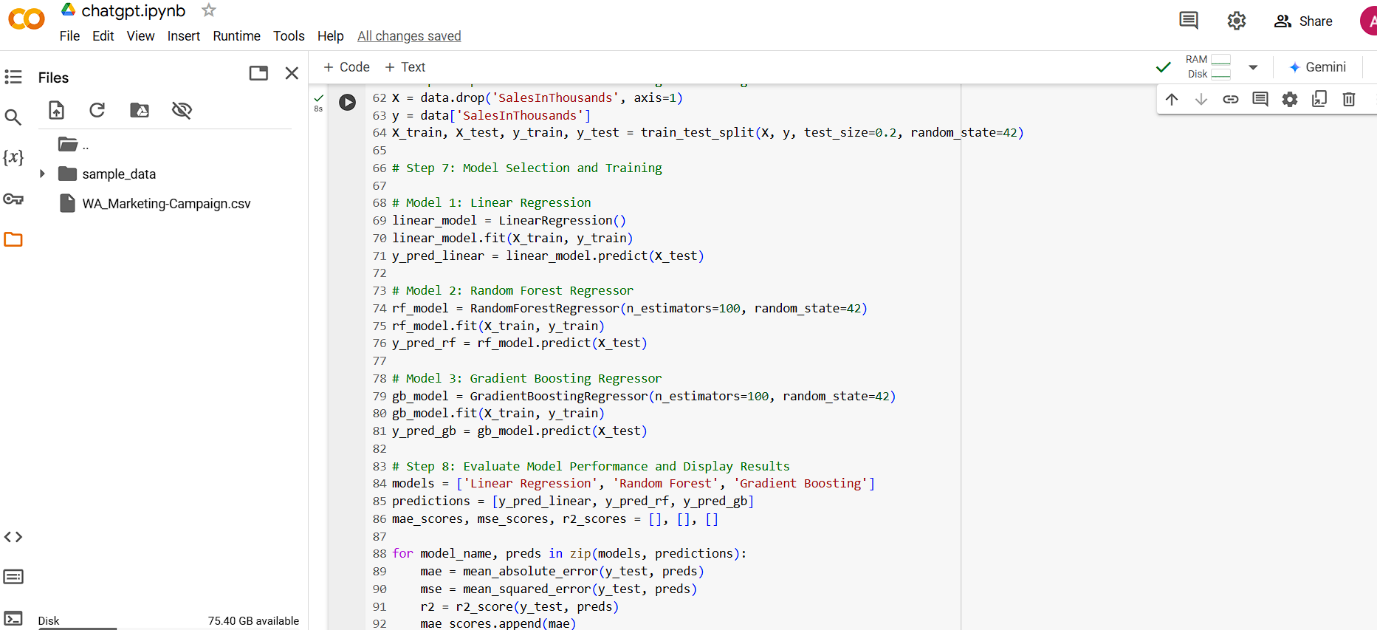
****

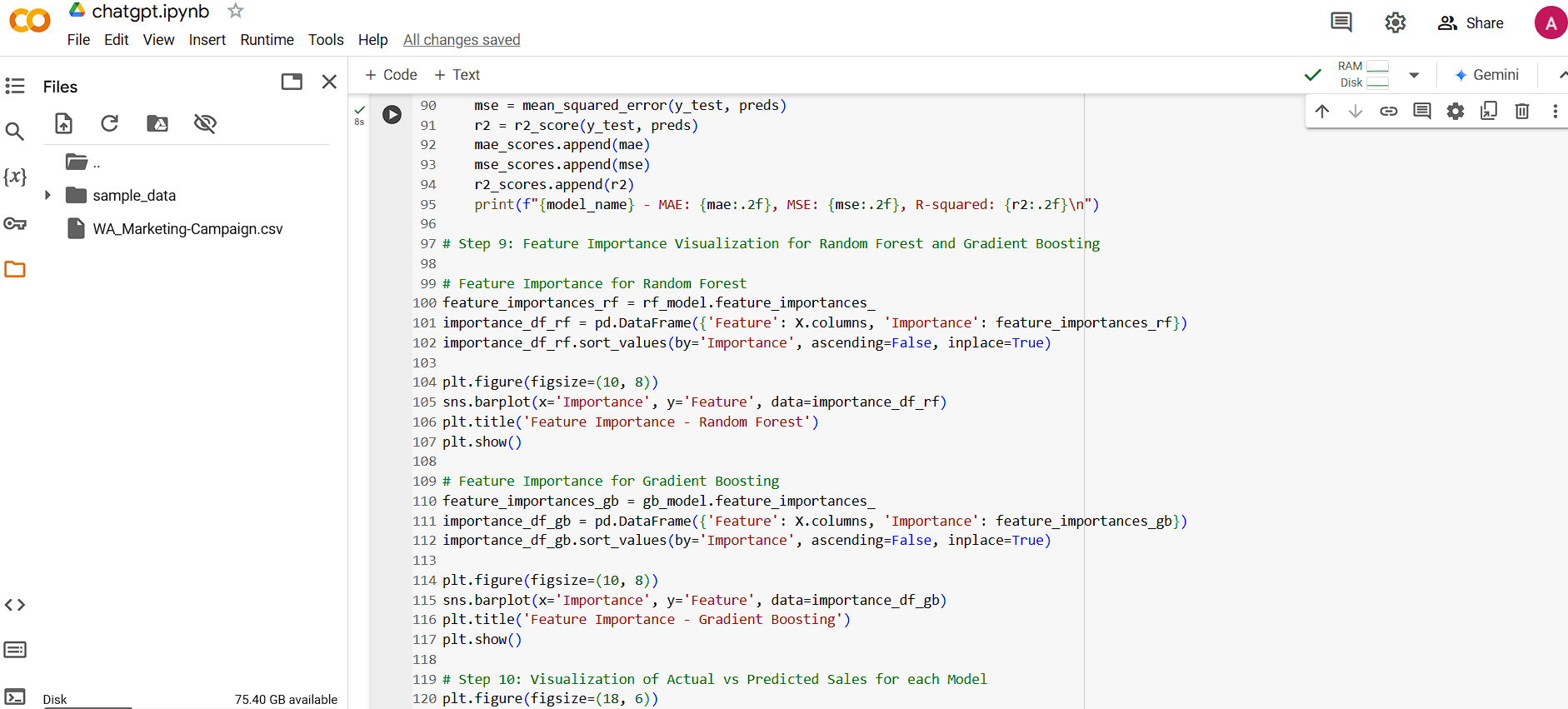
****

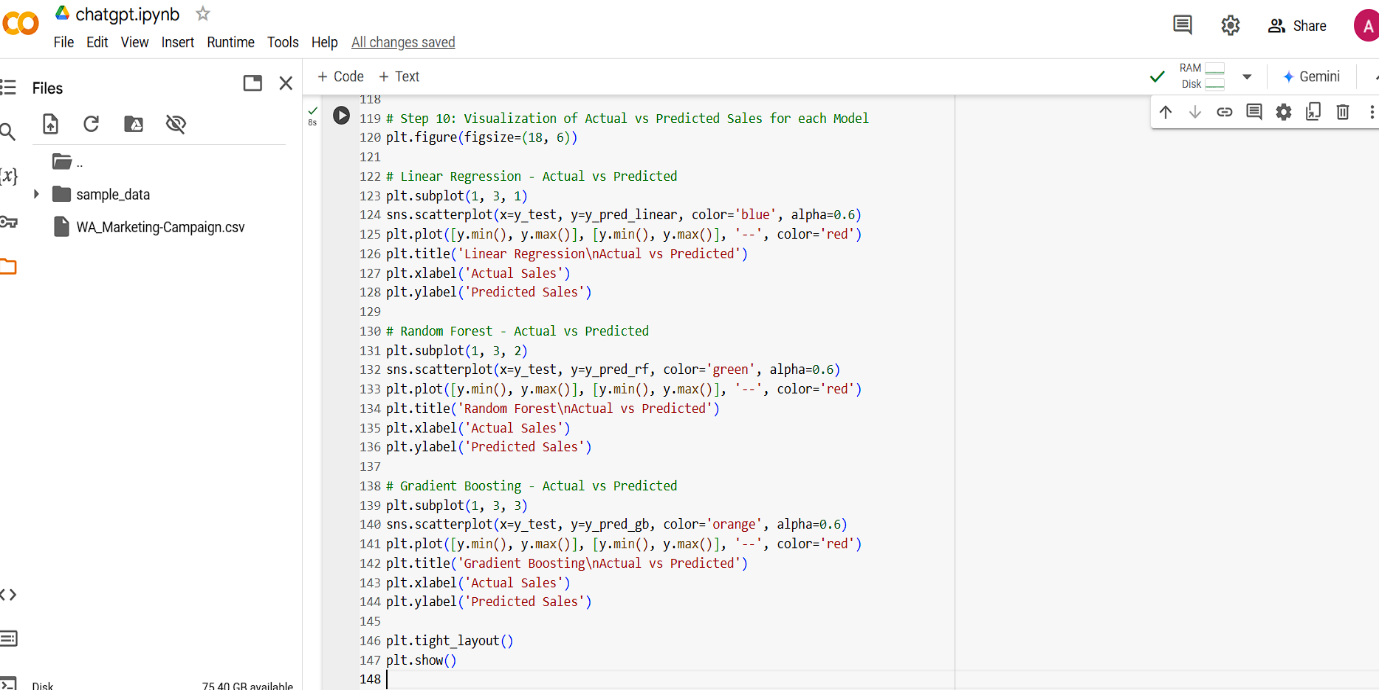


**Here is Google Colab Code generated By ChatGpt:**

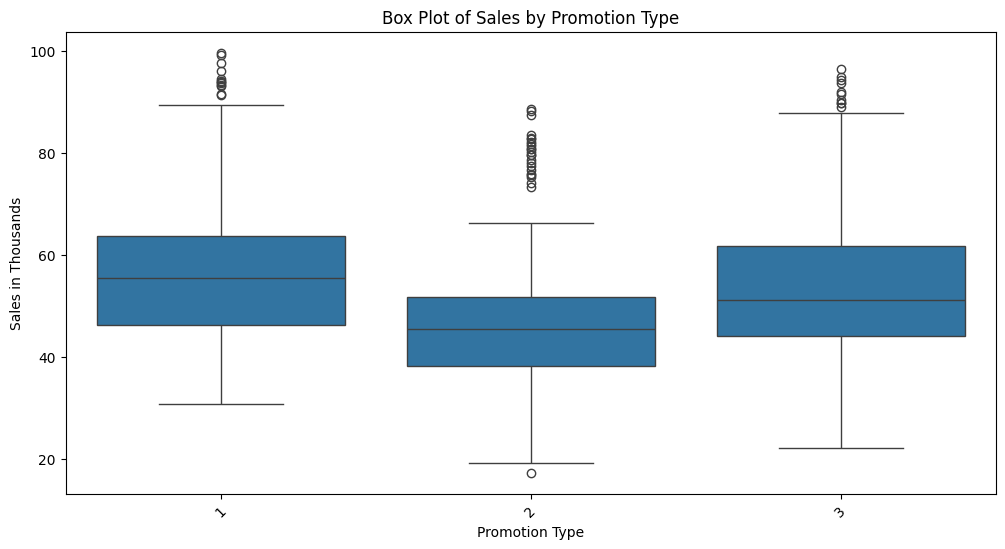
****

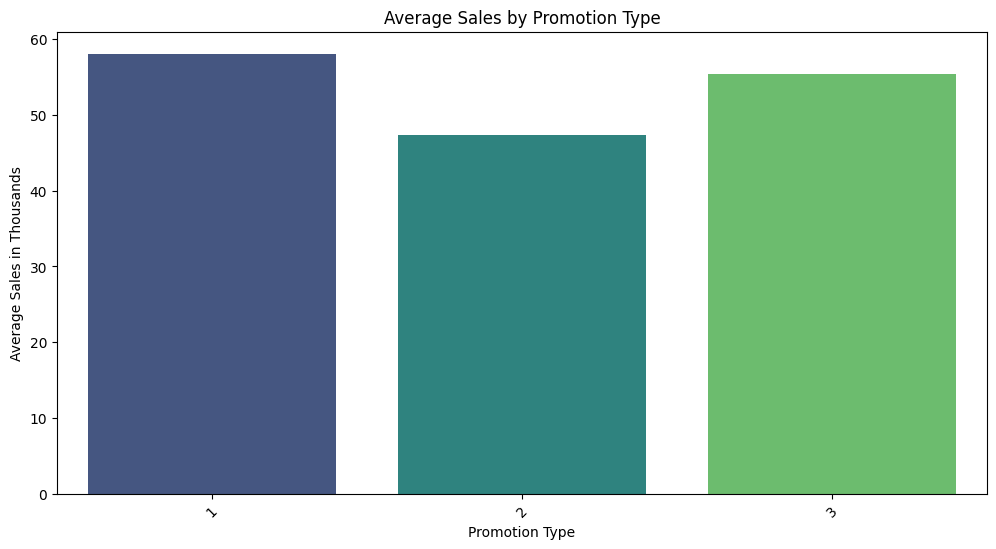
****

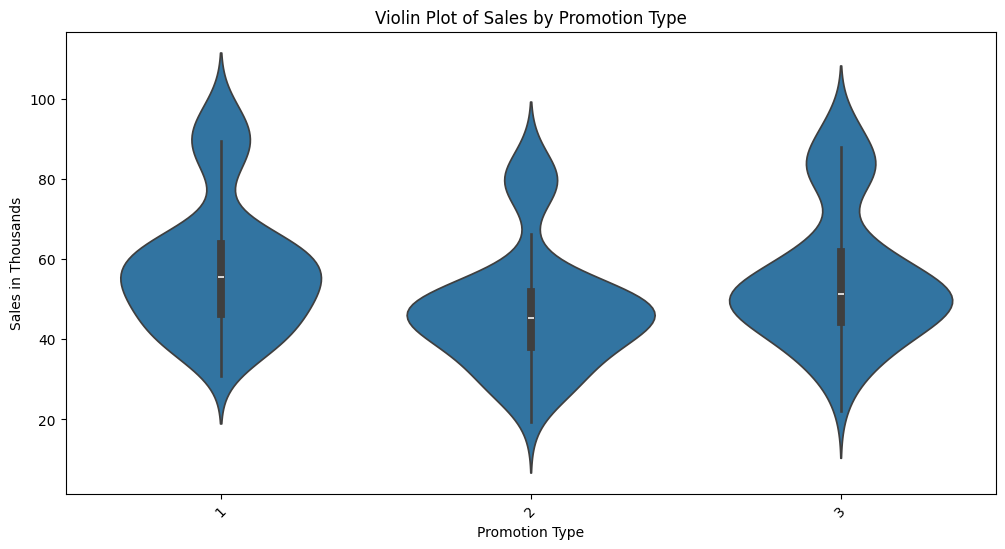
****

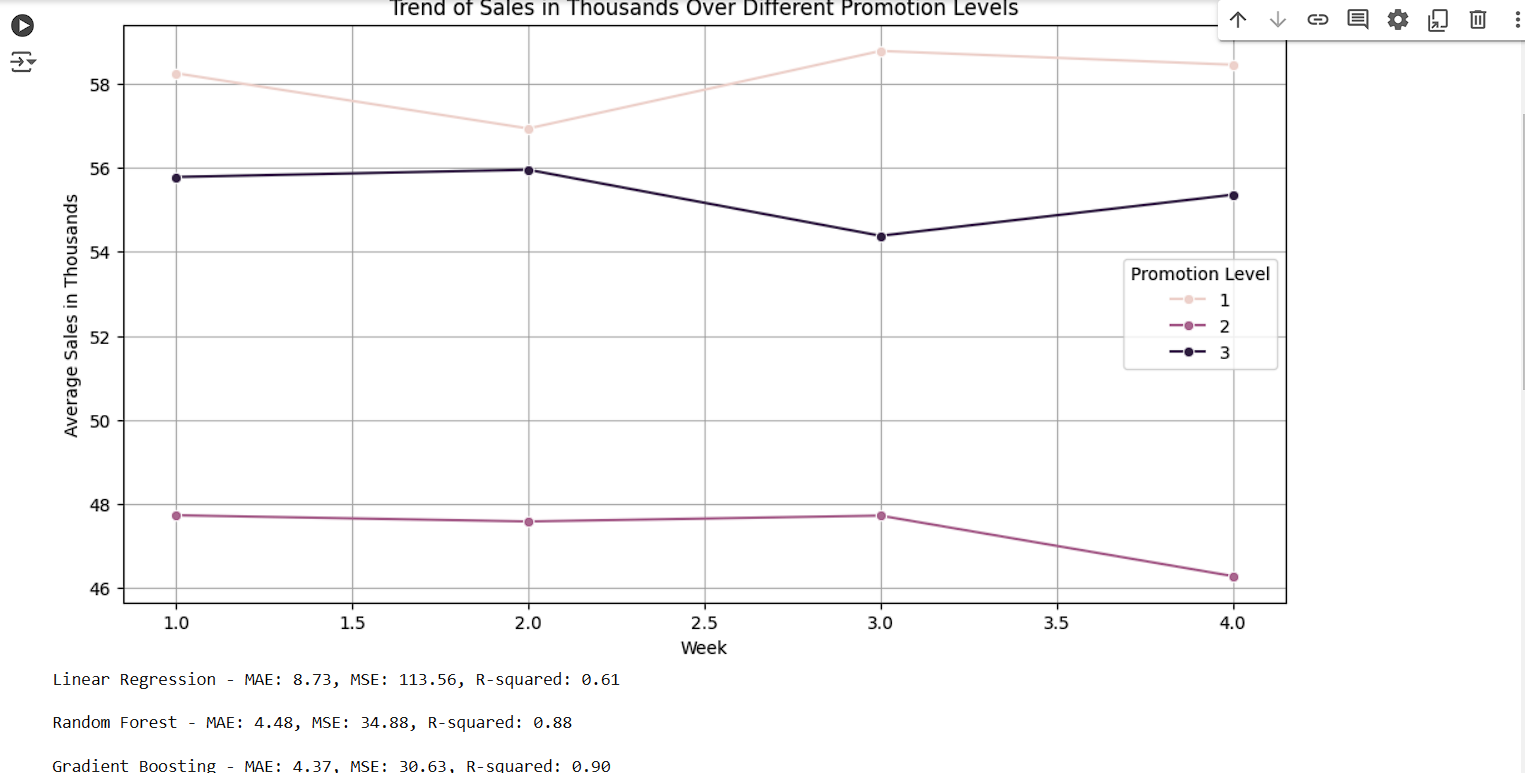
****

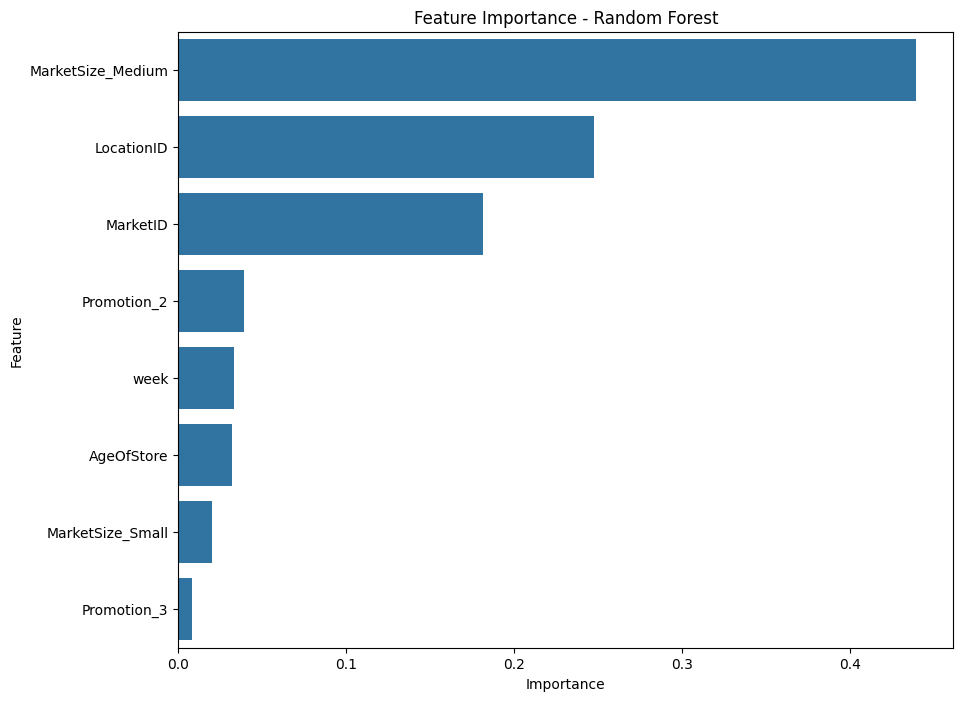
Result of the Google Colab Code

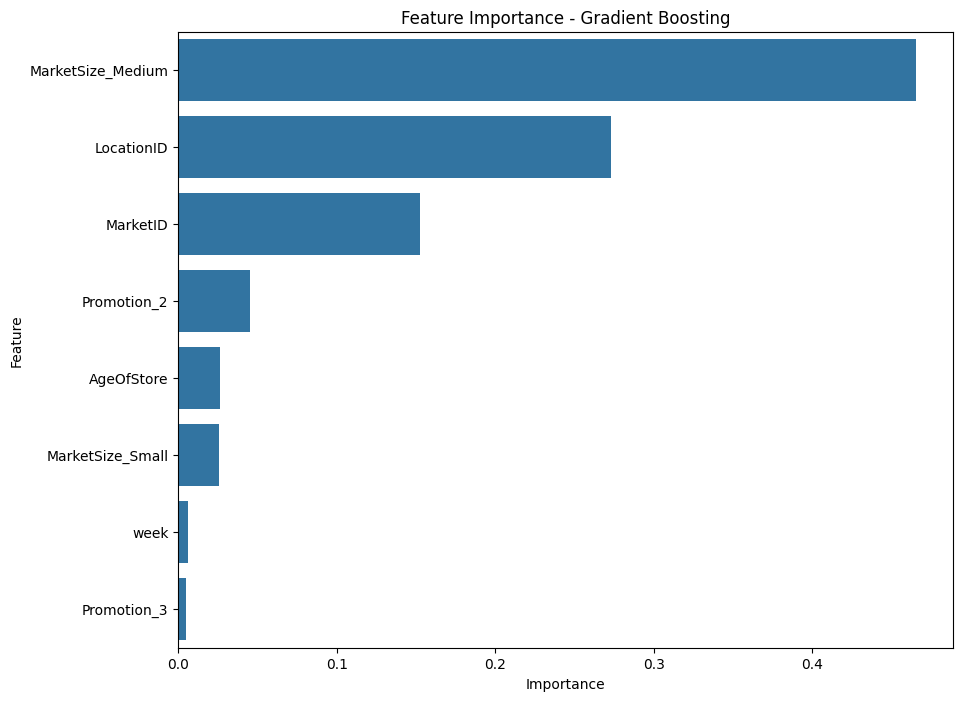


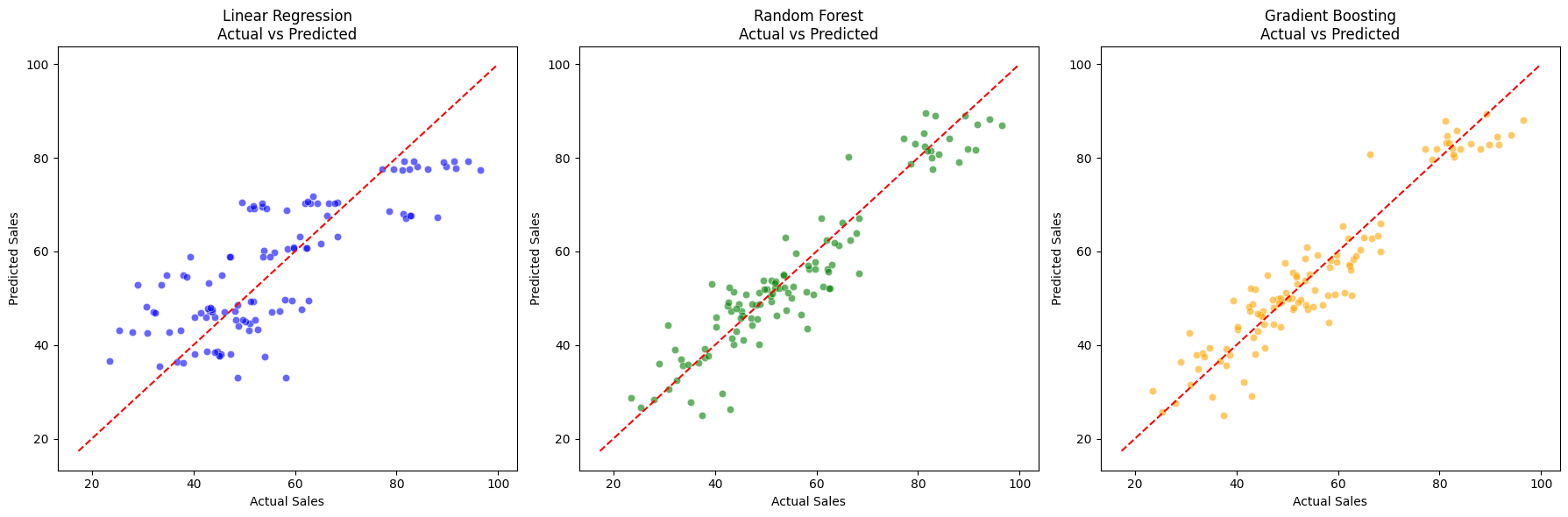




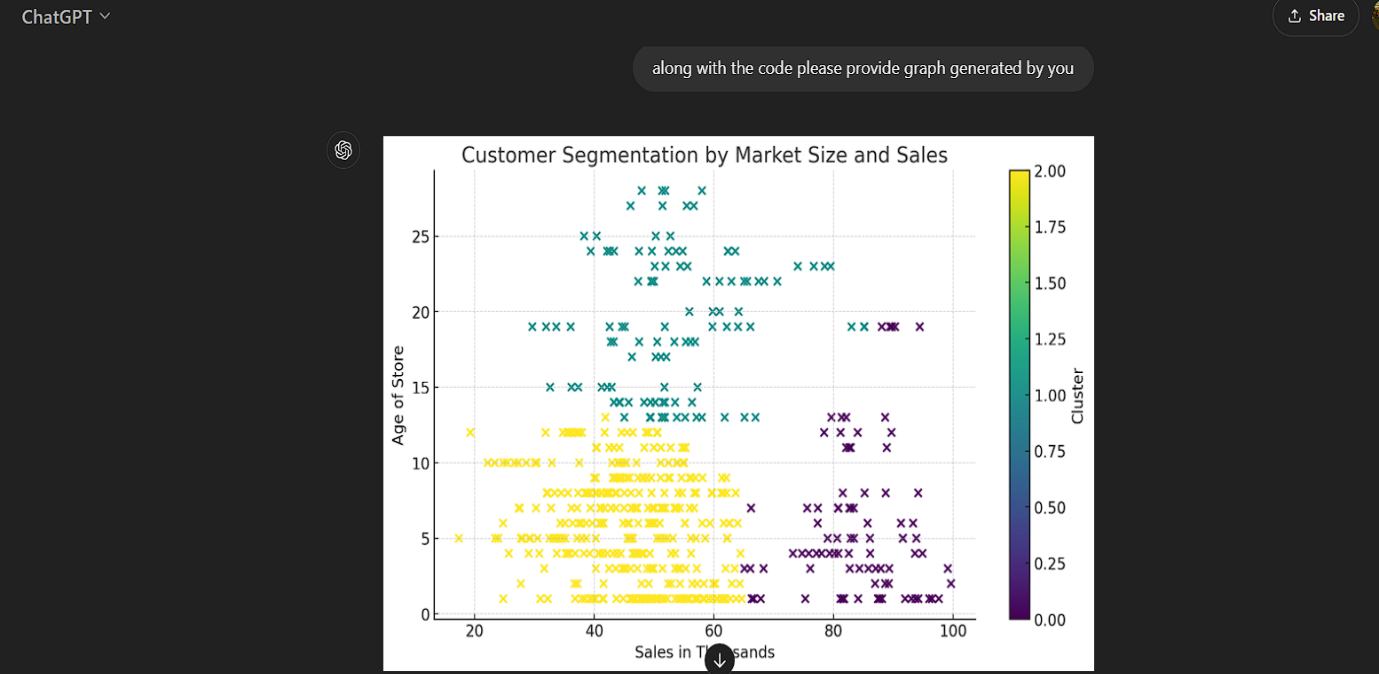


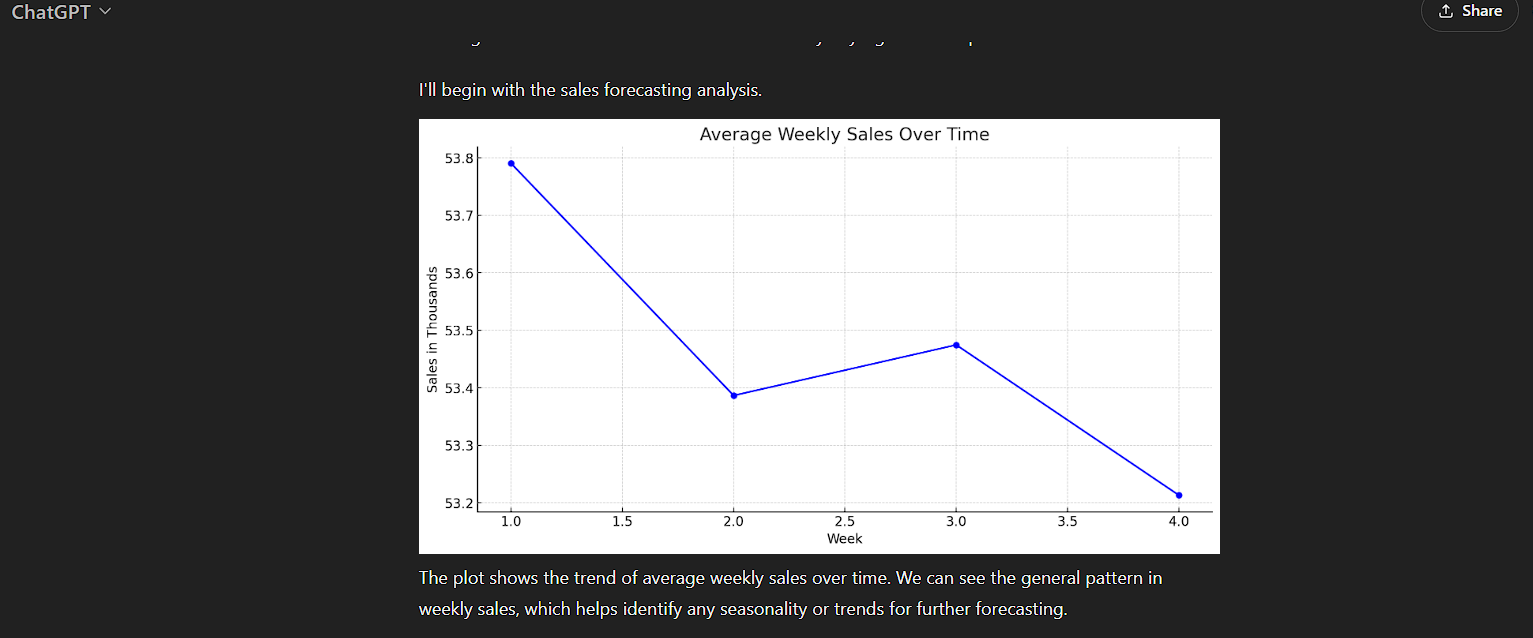




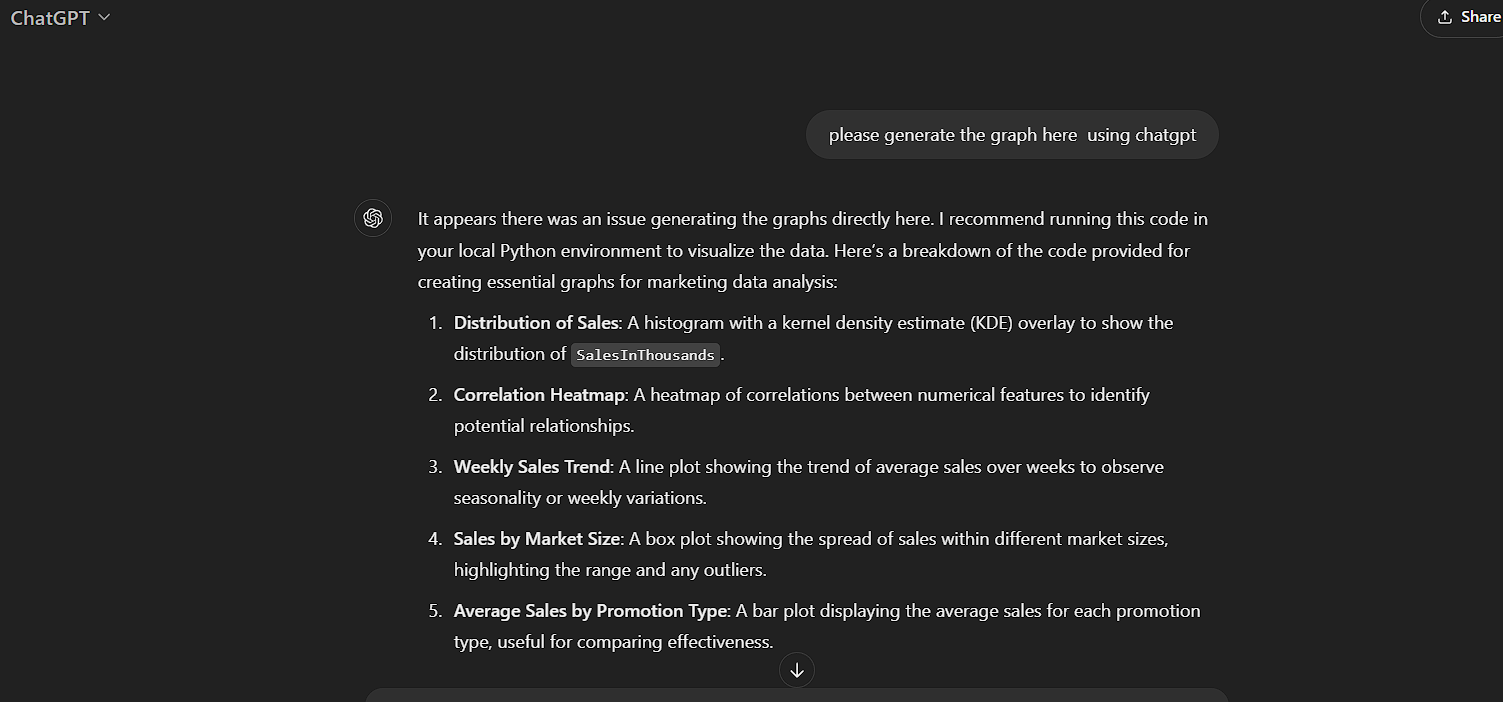


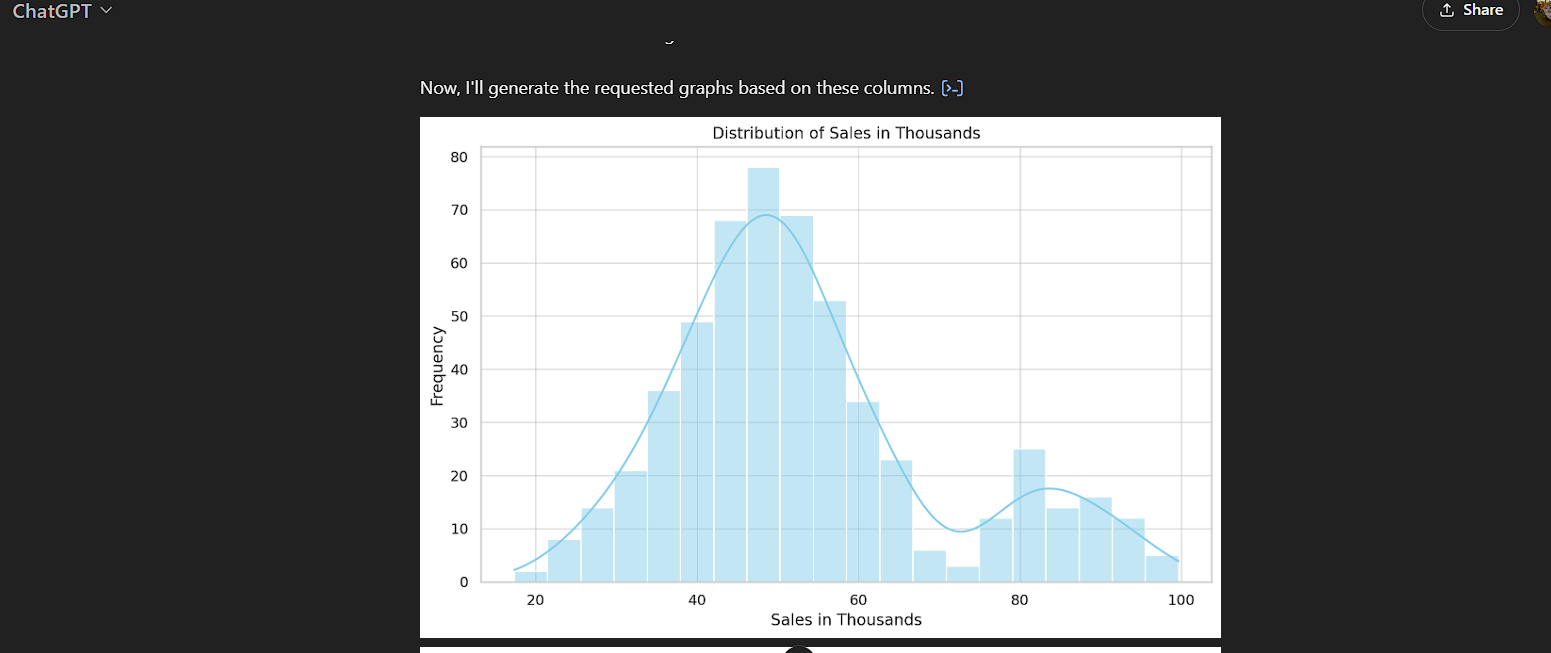
1. Result of Prompt 3:

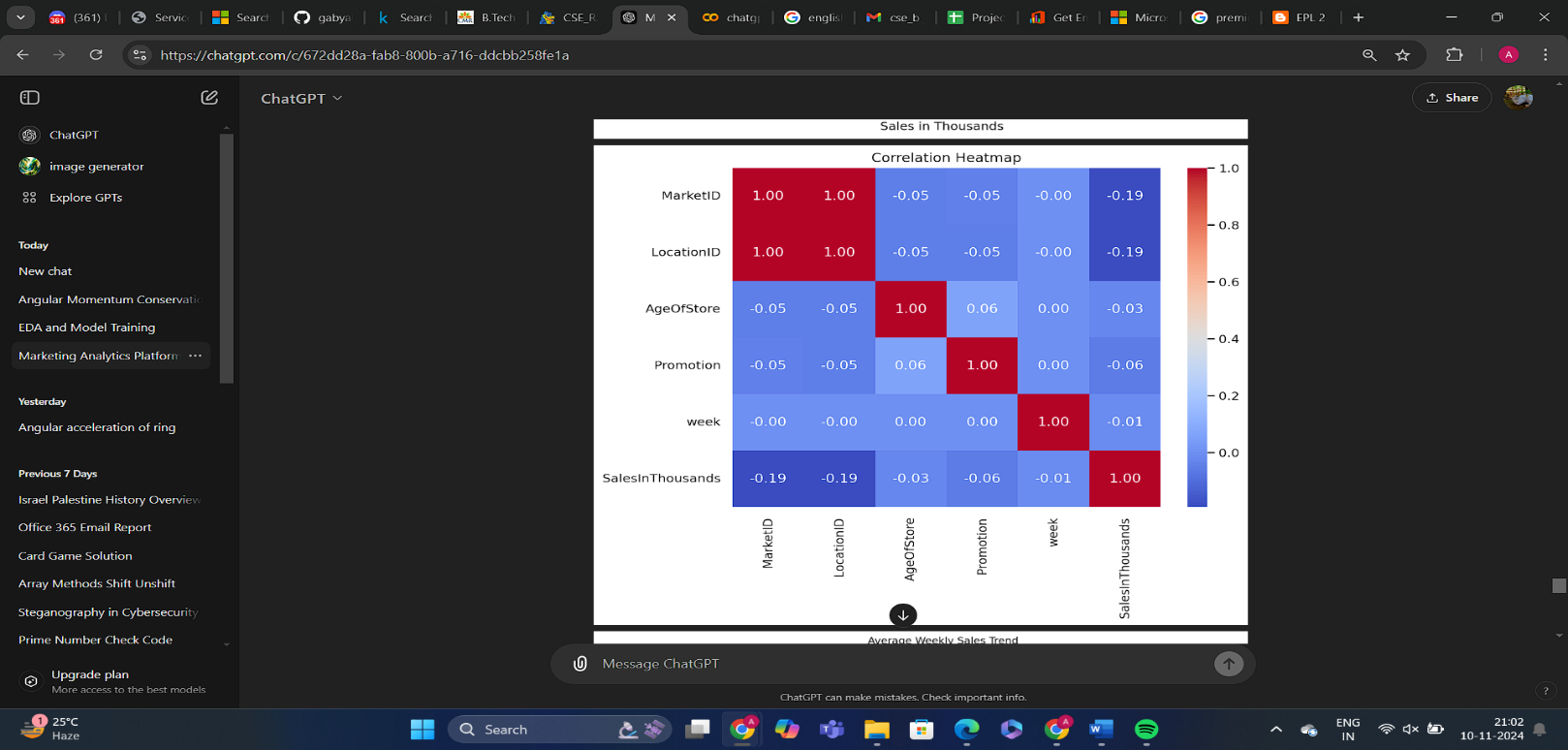


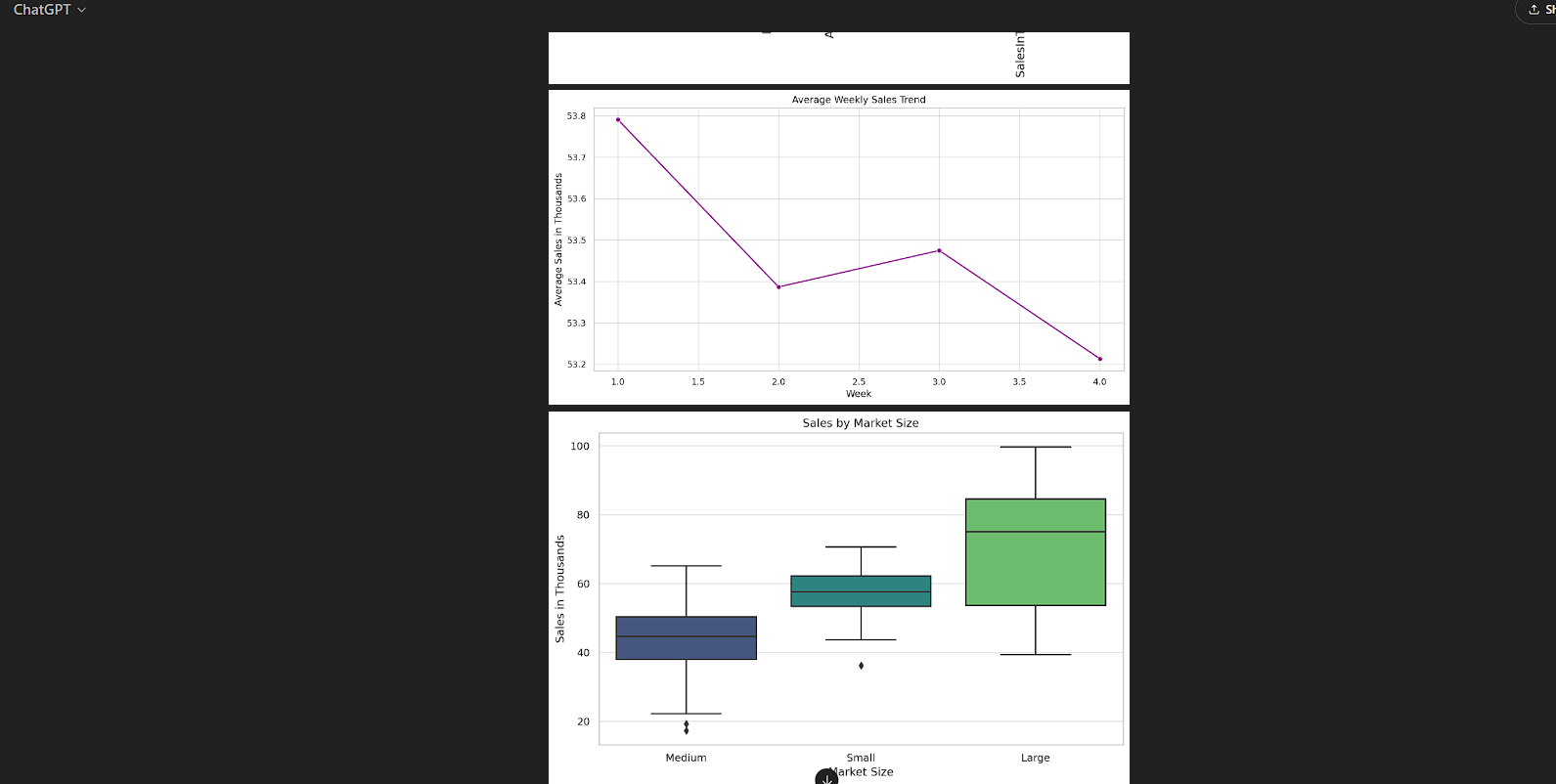


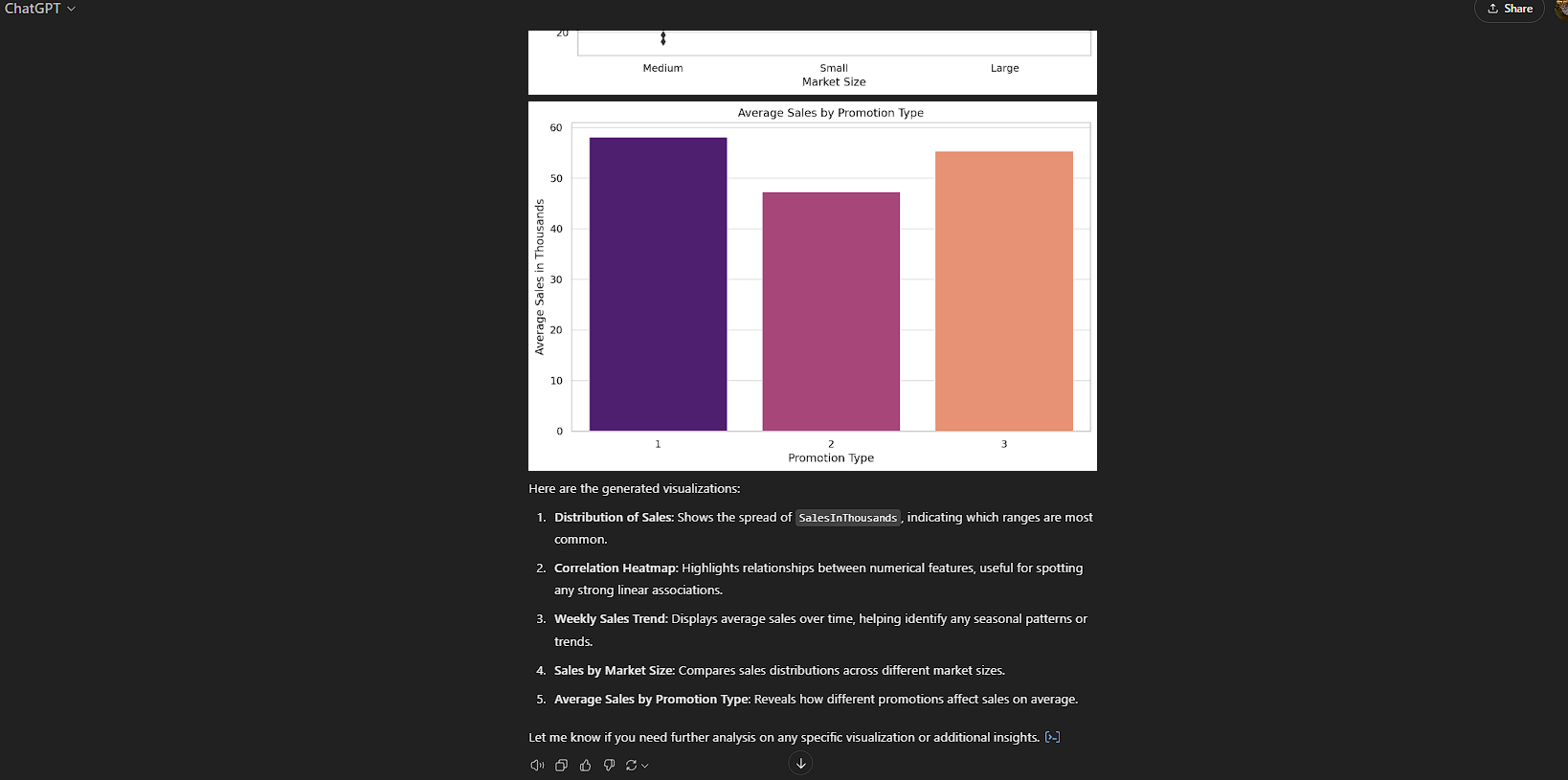
1. Result of Prompt 4:



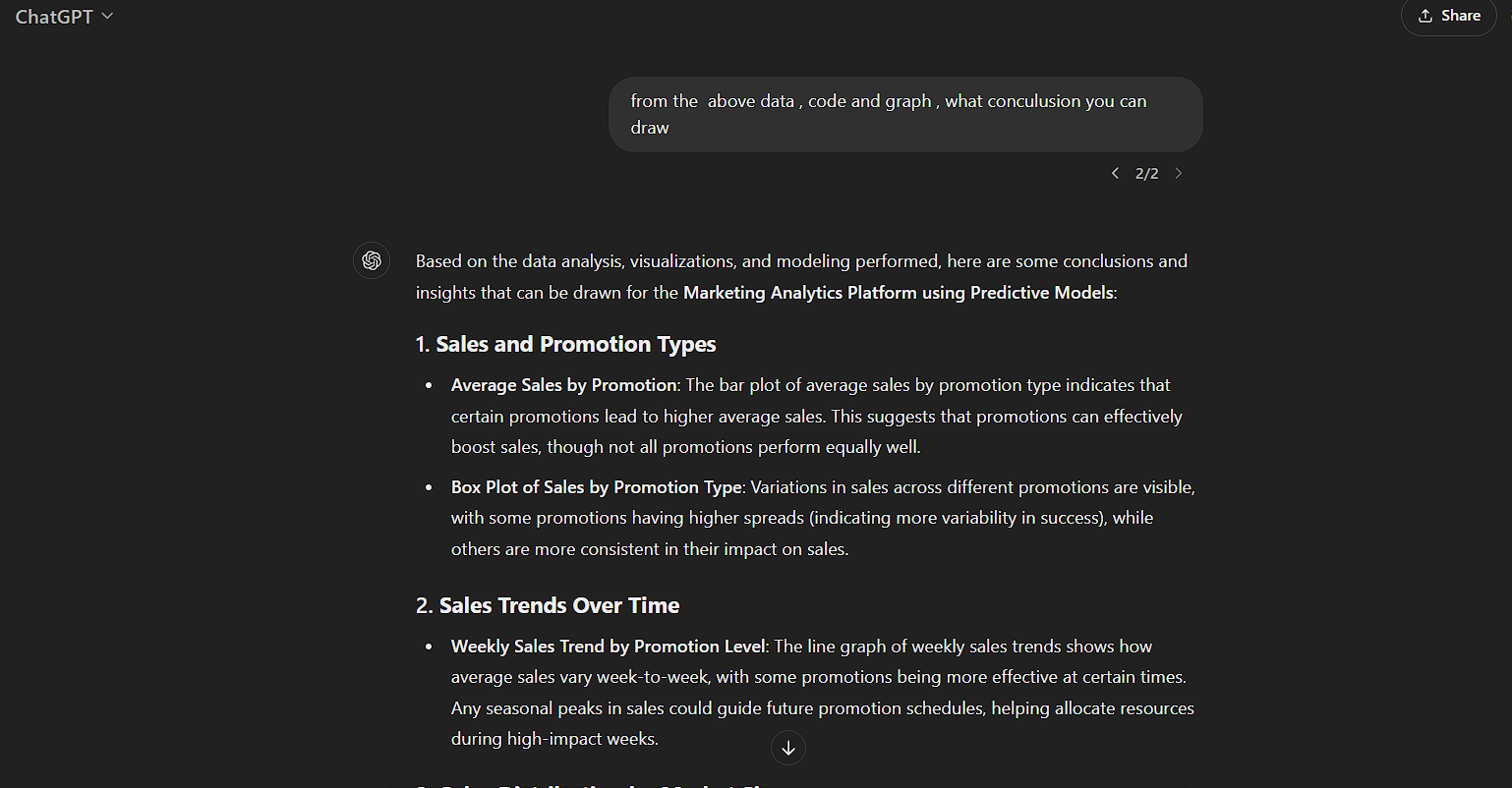


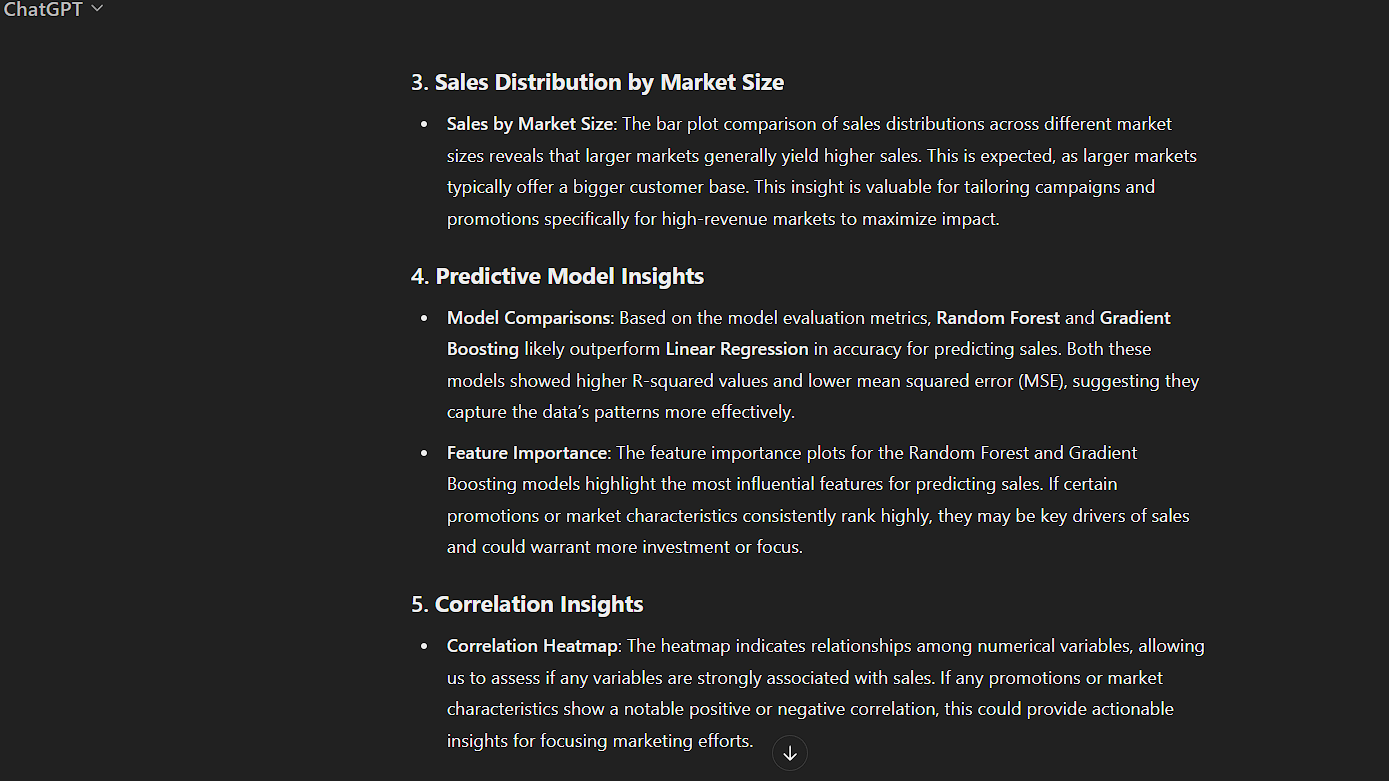


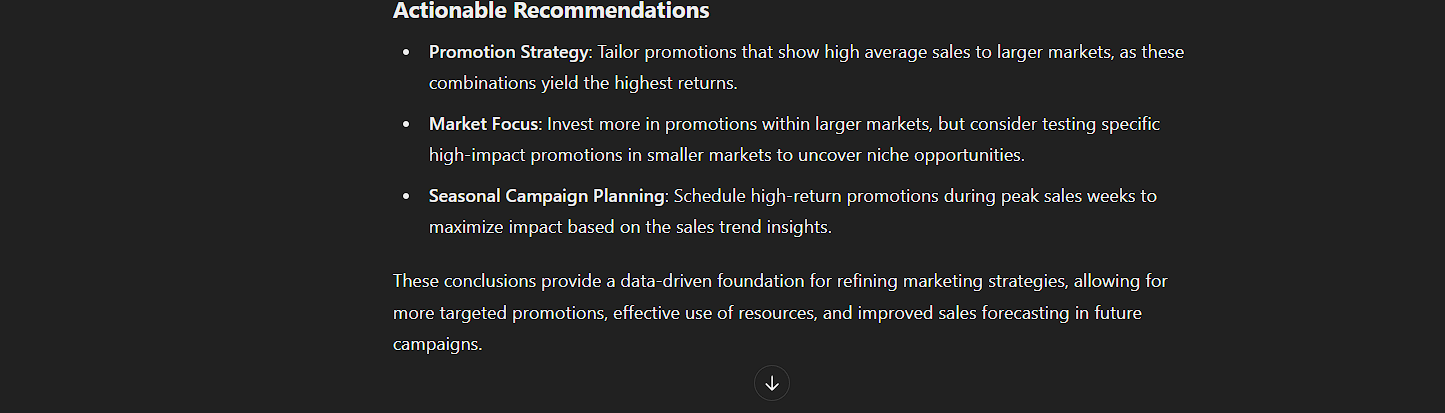




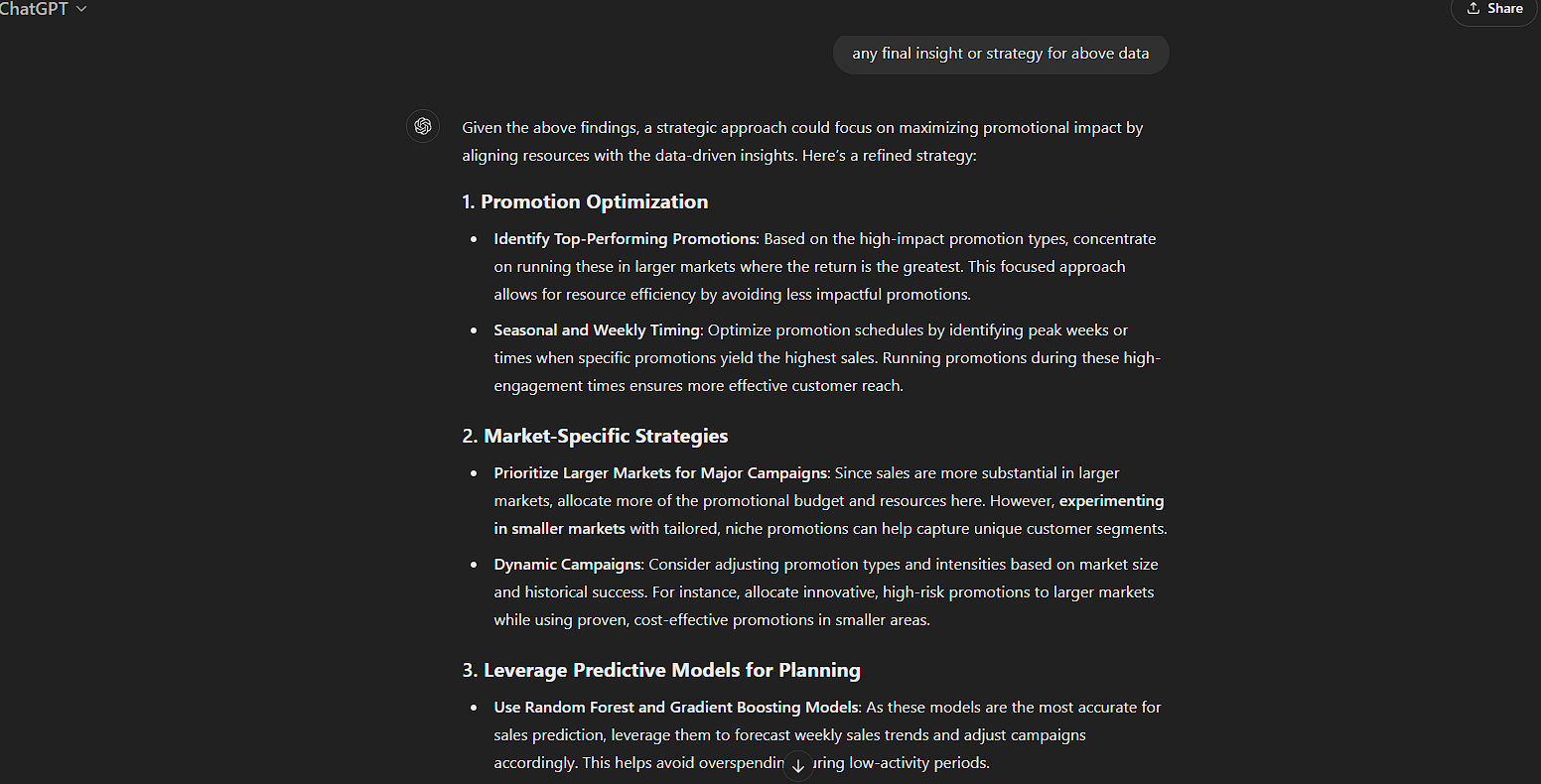
1. Result of Prompt 5:

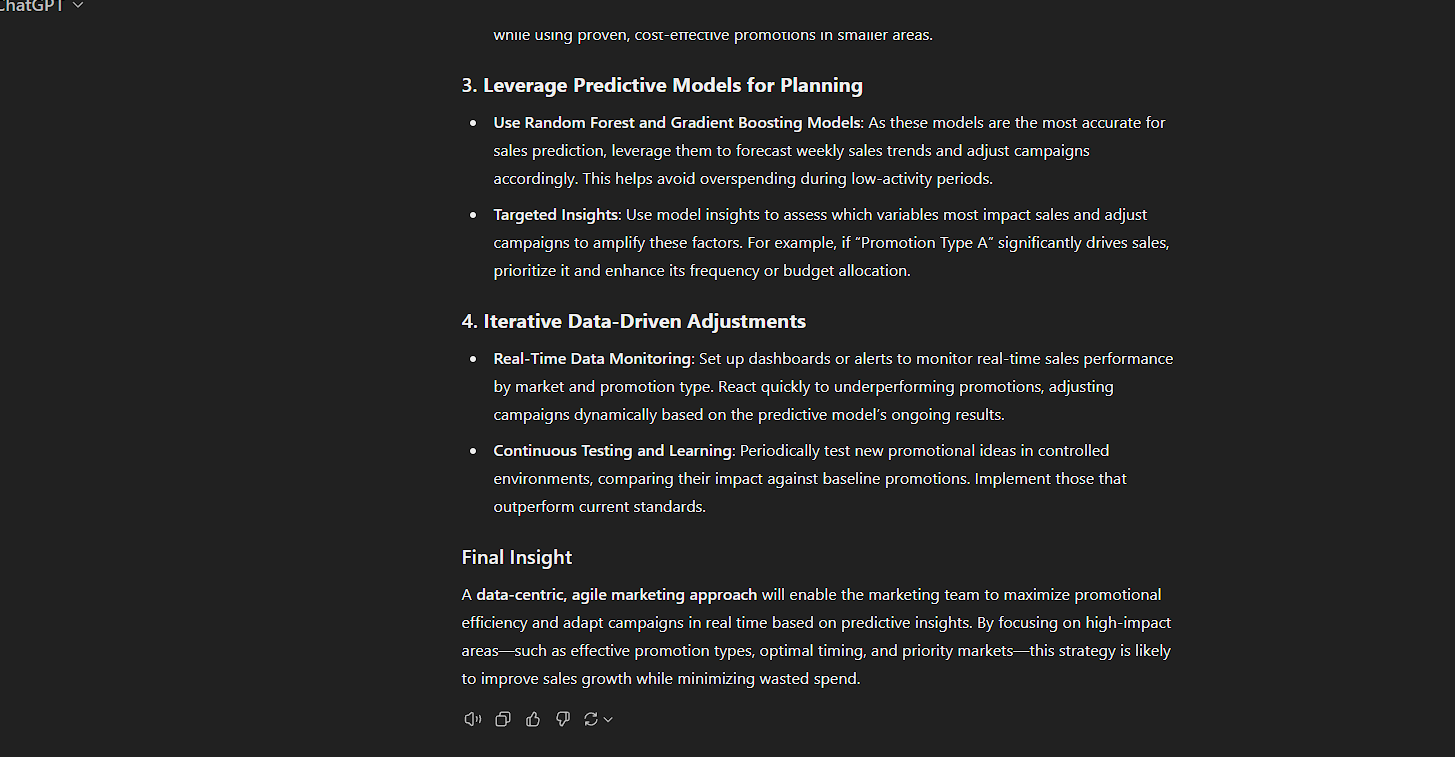






1. Result of Prompt 6:





**5.CONCLUSION**

In this project, we developed a *Marketing Analytics Platform* to analyze and predict sales outcomes using the *WA\_Marketing-Campaign.csv* dataset, with assistance from ChatGPT. This involved data exploration, feature engineering, predictive modeling, and insights generation, leading to actionable recommendations.

**Data Preparation and Exploration**:

The dataset was cleaned by handling missing values and exploring key features like *Promotion*, *MarketSize*, and *SalesInThousands*. Graphical analyses, including box plots and line graphs, revealed patterns in sales trends and promotion effectiveness across market sizes.

**Modeling and Evaluation**:

After preprocessing the data (including one-hot encoding for categorical variables), we trained multiple regression models—*Linear Regression*, *Random Forest Regressor*, and *Gradient Boosting Regressor*—to predict sales outcomes. These models were evaluated using metrics such as *Mean Absolute Error (MAE)* and *R-squared* scores. Gradient Boosting outperformed others, with high accuracy in predicting *SalesInThousands* and indicating strong predictive power for marketing strategies.

**Strategic Recommendations**:

To optimize marketing efforts, we recommend deploying targeted promotions aligned with market size to drive sales effectively. Resources should focus on larger markets where higher returns are likely. Implementing the Gradient Boosting model as part of an ongoing platform will allow for predictive monitoring and adjustment of marketing strategies based on data-driven insights.

Overall, with ChatGPT’s guidance, we created a predictive platform that offers valuable insights into optimizing marketing campaigns, aligning promotions with market dynamics, and leveraging predictive analytics for better sales outcomes.