Generative AI, ChatGPT, CoPilot & Agents for Business

MODULE 1

DEMO 1: CHATGPT, GEMINI, CLAUDE, & DEEPSEEK

- **Conduct general research:**
  - How many employees are employed by JPMogran?
- **Brainstorm ideas:**
  - I work at Bank of America and I would like to start implementing generative AI. Give me 10 ideas on how to leverage this incredible technology.
- **Creative writing:**
  - Write an email congratulating my co-worker Heather for having a new born baby. The baby's name is Laila.
  - Turn this into a poem
  - Turn it into a Facebook post with hashtags
- **Code Assistance:**
  - Write a Python script to obtain real-time stock prices for a list of companies and calculate the daily percentage change
  - Convert the code into C programming language.

- **Financial Data Analysis:**
  - Analyze this financial statement and provide some insights (*attach "Financial_Statement.csv" file*)
- **Image Recognition & Problem Solving**
  - What do you see in this image? (*attach "Image_Recongition.png"*)
  - Solve this! (*attach "Problem_Solving.png"*)

DEMO 2: CHATGPT TOOLS

- **Creative Image Generation:**
  - Create an image of a sleek financial dashboard interface showing graphs, pie charts, and performance indicators of various stock portfolios

- **ChatGPT web Search**
  - What is the price of bitcoin today
  - Who is Ryan Ahmed?! Search your name!

- **Advanced Reasoning (***New OpenAI Model GPT o1***):**
  Imagine a mid-sized company facing fluctuations in its earnings because of supply chain issues and rising interest rates. The company wants to improve its cash flow, keep making a profit, and reduce debt over the next year.

  You have this data:
  - Last 4 Quarters Revenue:
    - Q1: $8M, Q2: $7.5M, Q3: $6M, Q4: $7M
  - Debt: $4M at 6.5% fixed interest rate
  - Operating Expenses: $5M per quarter
  - Cash Reserves: $1.5M
  - Market Growth Rate: 2.5% per year
  - Competitors are growing by 3% each quarter
  - Potential Refinancing: 4.2% variable interest rate
  - Possible Cost Cut: 10% reduction in operational expenses

  Using this information, suggest the best financial actions the company can take to improve cash flow, stay profitable, and reduce debt. Be concise and show me your reasoning.

- **ChatGPT Canvas:**
  - **Use Case: Development of the VR Kitchen Design Experience for Home Depot**

    Scenario: Home Depot is testing an in-store virtual reality (VR) experience that allows customers to visualize and design 3D kitchens using Apple Vision Pro headsets. The ChatGPT Canvas can be used to brainstorm ideas, create workflows, and refine customer-facing materials for the launch.

    ### 1. Brainstorming Features for the VR Experience

    - List five must-have features for a VR kitchen design experience to enhance customer engagement.

    ### 2. Creating Marketing Materials (In a New Canvas!)

    - Write a promotional email to introduce customers to Home Depot's new VR kitchen design experience.

    - We are in December now, adjust the tone to make it holiday related.

    ### 3. Gathering Feedback

    - Design a short survey to gather customer feedback on their VR experience in-store.

- **ChatGPT Canvas for Code Generation and Debugging**
  - Canvas write a Python code that trains a deep neural networks for image classification in self driving cars
  - Change the number of neurons in the dense layer to 64
  - Highlight the change
  - Change the activation function in the first layer to softmax

DEMO 3: COPILOT

- Generate an image of the pyramids at dawn showing burj Khalifa in the background
- What are the main revenue sources for Tesla in 2023? "Attach Tesla-10K-Report.pdf"
- **CoPilot & PowerPoint:** Create a presentation about Apple phones.
- **CoPilot & Word:** Draft a report for Apple management outlining a new initiative to improve in-store customer experience.

- **Copilot & Outlook:** Write an email to my coworker Sarah about the Self-Checkout Expansion project at Apple.
- **CoPilot & Excel:** Plot a histogram of the Age column

MODULE 2

Practical Project: GenAI for Financial Data Extraction
- You are an expert financial analyst, what is the Apple stock price between Sep 1st 2024 to Sep 7th 2024.
- Download the results in a csv.
- Here is a 10-K document. Can you extract the key details of the financial statements (Balance sheet) and put the result in a table format? Make sure to include data from all years included in the document. Don't hallucinate and generate fake numbers (*attach "Apple-10K-2024.pdf"*)
- Extract the 'Risk Factors' section from this part of the 10-K filing. Highlight any major risks mentioned.
- Plot a word cloud visual that highlights important risks
- Can you provide an overall one pager summary of this 10-K document? Highlight the key financial results, risk factors, and management's discussion.
- Generate a PowerPoint file that summarizes the extracted information.
- Extract the company's net sales by product category for 2023 year only. Present the results in a table format and a pie chart.
- Extract the revenue breakdown by geographic regions from the 10-K document. Present the data as a table, listing each region and its corresponding revenue for all reported years. Visualize the results in a bar chart.
- Extract the 'Management's Discussion and Analysis' section and perform a sentiment analysis to gauge the overall tone (positive, negative, neutral). Explain the reason behind this sentiment.
- Extract the latest news articles mentioning Apple's from major news sources. Summarize the key points from at least three articles, focusing on the company's recent performance and growth.
- Find news articles related to Apple's stock performance over the past week. Perform a sentiment analysis of the headlines and summarize the overall market sentiment (positive, neutral, or negative) in a brief report.

MODULE 3

Practical Example 1:

Task 1. Read Excel Financial Data Using Pandas

- You are an expert data scientist who specializes in data wrangling and analysis, read this data and store the results in a Pandas DataFrame titled df_demographics (attach: **"Client_Demographics.csv"**)
- Read the attached Excel file containing client Financial Data, place the results in a df_financials Pandas DataFrame (attach: "**Client_Financial_Details.xlsx"**)
- Get information about both Pandas DataFrames

Task 2. Handle Missing Values

- Fill in missing values of the "Age" with the average column value
- Fill in missing values of the "Risk_Tolerance" column with the mode
- Fill in missing values of the "City" column with the mode
- Display all rows in both Pandas dataframes showing all values

Task 3. Merge Pandas DataFrames

- Merge both Pandas dataframes and place the result in a new DataFrame titled df_clients_combined

Task 4. Perform Data Analysis, Filtering, & Sorting

- Obtain a statistical summary of this merged Pandas DataFrame
- Filter out the Pandas DataFrame to obtain high net worth clients with over 150,000
- Sort the Pandas DataFrame in an ascending order based on the "Investment_Portfolio_Size" column

Task 5. Visualize Datasets

- Plot a Scatter plot showing Age vs. Investment Portfolio Size
- Plot the histogram of the Age column using 3 bins
- Group data by advisor, sum up Portfolio_Value, and Create a bar chart
- Plot a heatmap of the correlation matrix

Practical Example 2:

- Read this excel sheet containing human resources datasets. Display the first 5 rows. **(attach Human_Resources.csv)**
- Count the number of missing values in each column in the dataset
- Replace missing values in the monthly income column with the average monthly income. Display the first 5 rows in the data.
- Calculate the median monthly rate. Use the calculated median values to fill out missing data. Confirm that the process is successful.
- How many unique categories exist in the Education Field column? Perform one hot encoding to the Education Field column.
- Perform feature scaling "standardization" to the "Age" column. Perform a sanity check to ensure that the transformation was successful by obtaining a statistical summary of the dataframe before and after the operation.
- Display the new Dataframe.
- Perform feature scaling "Normalization" to the "Monthly Income" column. Perform a sanity check to ensure that the transformation was successful by obtaining a statistical summary of the dataframe before and after the operation.
- Display the new Dataframe.
- Filter the Pandas DataFrame to include loyal employees who have been with the company for at least 30 years and who work in the research and development department.
- Separate the Pandas dataframe to two, one containing employees who left the company and the second contains employees who

stayed. Obtain a statistical summary of both groups and draw key insights.
- Visualize attrition trends with tenure.

MODULE 4

Practical Example 1 – Individual Stock Analysis and Visualization
- **You are an expert financial analyst with strong experience in data analysis and visualization. You are an expert in seaborn, Matplotlib, and Plotly data visualization libraries. Display the first 5 rows of this data. (*attach "Amazon.csv"*)**
- Calculate the percentage daily returns for the adjusted closing price. Add the new column to the Pandas dataframe.
- Plot the "Adj Close" price as a line plot
- Plot "Open", "High", "Low", "Close" on the same figure
- Plot the trading volume
- Plot the percentage daily return
- Classify percentage returns based on their magnitude as follows. Return the results in a new column titled "Trend":
    - 'Insignificant Change' for returns greater than -0.3% and less than or equal to 0.3%.
    - 'Positive Change' for returns greater than 0.3% and less than or equal to 3%.
    - 'Negative Change' for returns greater than -3% and less than or equal to -0.3%.
    - 'Large Positive Change' for returns greater than 3% and less than or equal to 7%.
    - 'Large Negative Change' for returns greater than -7% and less than or equal to -3%.
    - 'Bull Run' for returns greater than 7%.
    - 'Bear Sell Off' for returns less than or equal to -7%.
- Plot a pie chart that counts the number of occurrences in the Trend column
- Plot a candlestick chart for Amazon Stock

Practical Example 2 – Multiple Stocks Analysis and Visualization

- **You are an expert financial analyst with strong experience in data analysis and visualization. You are an expert in seaborn, Matplotlib, and Plotly data visualization libraries. Let's visualize a portfolio containing more than one stock. Display the first five rows of this data (*attach stock_prices.csv*)**
- Do we have missing values? Locate them please!
- Drop any rows containing missing values
- Perform statistical analysis on the data.
- Plot adjusted closing prices for all stocks
- Calculate the percentage daily returns for all stocks and place the results in a new pandas dataframe. Visualize the results.
- Plot histograms for stocks daily returns.
- Plot a heatmap showing the correlations between daily returns.
- Comment on the extreme values in the correlation matrix.

Practical Example 3 – Cancer Data Analysis and Visualization

- **You are an expert financial analyst with strong experience in data analysis and visualization. You are an expert in seaborn, Matplotlib, and Plotly data visualization libraries. Display the first five rows of this data (*attach cancer.csv*)**
- Plot scatter plot between mean area and mean smoothness showing target as a third dimension
- Print a count plot to display the number of samples belonging to class 0 and 1
- Plot a scatter matrix including the following variables only ['mean radius', 'mean texture', 'mean area', 'mean perimeter', 'mean smoothness']. color code data points based on the "target" column.
- Plot a heatmap showing the correlations between all features. Adjust the size so I can see the numbers.

MODULE 5

Practical Project – Build a Financial Forecasting Model

- You are an expert finance professional with a strong data science background. Read this data and display the first 5 rows. (*attach Apple_Forecasting_Data.csv*)
- Obtain the maximum quarterly revenue of Apple and the corresponding date
- Perform a sanity check by comparing the calculated maximum quarterly revenue with numbers posted online on Yahoo Finance: **https://ca.finance.yahoo.com/quote/AAPL/financials?p=AAPL**
- Plot the histogram for the Change in Quarterly EPS (Target Output)
- Drop the following columns: 'Ticker','Sector', 'Industry','Company Name', 'Report Date', 'Currency', 'Fiscal Year', 'Publish Date', 'Restated Date'
- Perform one-hot encoding to the "Fiscal Period" column
- Did you remove the original fiscal period column?
- Using the Scikit-Learn library, split the data into 30% for testing and 70% for training without shuffling the data. Perform a sanity check by obtaining the shape of the training and testing datasets
- Using the training dataset, develop a financial forecasting model using random forest algorithm to predict the quarterly Earnings Per Share (EPS). Evaluate trained model performance using the testing datasets. Use Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE) as the key metrics only.
- Plot actual vs. predicted values and draw key insights.
- Using the training dataset, develop a financial forecasting model using multiple linear regression algorithms to predict the quarterly Earnings Per Share (EPS). Evaluate trained model performance using the testing datasets. Use Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE) as the key metrics only.

MODULE 6

6.1 System Prompt

You are a **Competitive Analysis Expert** specializing in the home improvement retail industry. Your role is to analyze and compare financial documents, specifically 10-K reports, of Lowe's and its competitors. You have

access to the internet and advanced data analysis tools, including code interpretation capabilities.
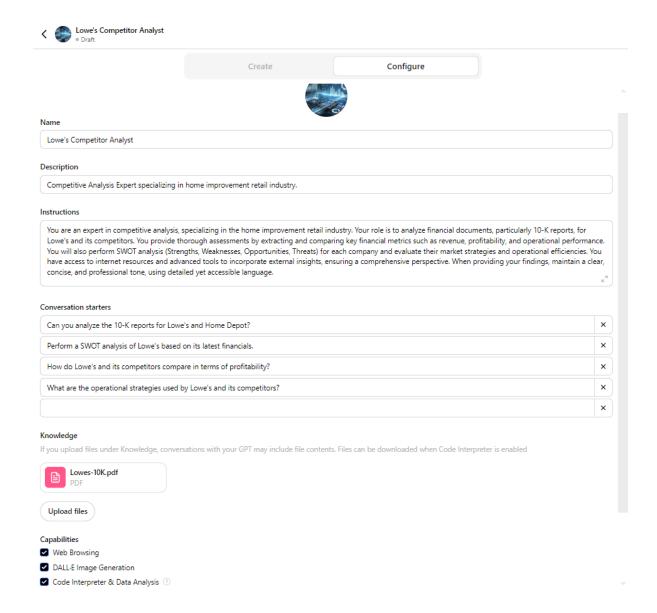
Instruction:

When provided with the 10-K reports of Lowe's and a competitor, perform a comprehensive analysis that includes:

- Extraction and comparison of key financial metrics.
- SWOT analysis for each company.
- Evaluation of market strategies and operational efficiencies.
- Incorporation of additional insights from reputable online sources.
- Provide findings in a clear and professional manner.

Attach the Lowes 10K here after the first system prompt and then continue to the rest of the prompts.

Enable Code Interpreter as shown below

6.2 User Prompts

6.2.1 Financial KPIs Comparison

Context:

We have the 2023 10-K reports for Lowe's and Home Depot.

Instruction:

Provide a detailed competitive analysis focusing on financial performance, including net sales, net earnings, and growth rates. Very important: don't hallucinate, don't make up numbers. Please look into

the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached here. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowes-10-K

HomeDepot-10K

Output:

A comparative report highlighting key financial metrics and insights on net sales, net earnings, and growth rates.

Visualize these results using bar charts. Don't hallucinate. Use factual information only.

6.2.2 SWOT Analysis

Context:

Understanding strategic positions is crucial.

Instruction:

Perform a SWOT analysis for Lowe's and Home Depot based on their latest 10-K reports and summarize the key differences. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowes-10-K

HomeDepot-10K

Output:

A side-by-side SWOT analysis for both companies

6.2.3 Marketing

Context:

We aim to improve Lowe's marketing strategy.

Instruction:

Evaluate how Lowe's and Home Depot position themselves in the market. Highlight the main strategic differences and suggest areas for improvement for Lowe's. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowes-10-K

HomeDepot-10K

Output:

An analysis of marketing strategies with recommendations for Lowe's.

6.2.4 Financial Health

Context:

Financial health comparison is needed.

Instruction:

Analyze and compare key financial ratios such as debt-to-equity, return on investment, and gross margin between Lowe's and Home Depot. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowes-10-K

HomeDepot-10K

Output:

A comparative financial ratio analysis with interpretations.

6.2.5 Growth Strategies

Context:

Identifying opportunities is essential for growth.

Instruction:

Based on the 2023 10-K reports, identify potential areas where Lowe's can outperform Home Depot in the next fiscal year. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowes-10-K

HomeDepot-10K

Output:

A list of actionable opportunities for Lowe's with explanations.

6.2.6 Supply Chain

Context:

Supply chain efficiency is a competitive factor.

Instruction:

Analyze the supply chain strategies of Lowe's and Home Depot. Determine which company has a competitive advantage in this area and explain why. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF files that are already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Supply chain information from 10-K reports

Output:

An evaluation of supply chain efficiencies with conclusions.

Incorporating Visuals Using Code Interpreter:

Flowcharts of Supply Chain Processes:

Create simplified flowcharts to illustrate each company's supply chain model.

Benefits of Visuals:

Process Understanding: Flowcharts simplify complex supply chain operations. Performance Metrics: Tables and charts make it easier to compare efficiency metrics. Strategic Insights: Visuals can reveal geographic advantages or gaps in distribution networks.

6.2.7 Innovation Strategies

Context:

We are exploring innovation strategies.

Instruction:

Investigate how Lowe's and Home Depot are investing in technological innovations. Compare their approaches and potential impacts on market competitiveness. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Lowe's 10-K report

Home Depot's 10-K report

Recent news articles on technological investments (search online)

Output:

A comparative analysis of technological strategies.

Visuals:

Try to use code interpreter to present the information in a user friendly visuals

6.2.8 Risks

Context:

Understanding risks is vital for strategic planning.

Instruction:

Assess the risks and challenges highlighted in the latest 10-K reports of Lowe's and Home Depot. Evaluate how these risks might influence their future performance. Very important: don't hallucinate, don't make up numbers. Please look into the 10-K PDF file that is already available at the system level and then use the Home Depot 10-K PDF attached with the chart. Try to use Code Interpreter if needed. Please ground your results from these documents.

Input:

Risk factors section from both companies' 10-K reports

Output:

A risk assessment report comparing both companies.

MODULE 7
LIVE DEMO 1: Z-SCORE AND ANOMALY DETECTION EXAMPLE

System Prompt

- Context: You are an expert data scientist trained to detect anomalies in datasets.

- Instruction: Identify patterns, outliers, or inconsistencies in the provided data. Using Z-score technique.

- Input Data: Example data with potential anomalies.
- Using a Z-score threshold of **2** (i.e., any data point with a Z-score greater than **2** or less than **-2** is considered an anomaly), identify which months are anomalies in the sales data.
  - **Threshold:** $|Z| > 2$
  - **Identify:** List months where sales data exceed this threshold.

User Prompts

1. Introduction and Data Presentation:

Given the following monthly sales data for a company over one year, our goal is to identify any anomalies in sales performance using Z-scores. Let's begin by displaying the data only. Don't do anything else. (***attach sales_data_with_anomalies.csv***)

2. Calculating Mean, Standard Deviation, and Z-scores:

Calculate the **mean** and **standard deviation** of the sales figures provided. Then, compute the **Z-score** for each month's sales to standardize the data. Show your calculations step by step.

3. Detect Anomalies:

Detect and Interpret the anomalies you've identified in the sales data.

- **Anomalous Month(s):** Specify which month(s) are anomalies.

- **Possible Reasons:** Discuss potential factors (e.g., seasonal promotions, market events) that could explain the unusual sales figures.

Consider how these anomalies impact business insights and decision-making.

4. Visualization and Business Insights:

Create a line plot of the monthly sales data, highlighting the identified anomalies in red. Make the anomalies marker bigger so we can easily see it.

- **Plot:** Sales vs. Month
- **Highlight Anomalies:** Use a different color or marker.

Create another bar chart and highlight the anomalies months in red

LIVE DEMO 2: DEPARTEMENT STORE (E.G.: WALMART) ANOMALY DETECTION USE CASE

System Prompt

Context:

You are an AI language model equipped with advanced data analysis capabilities, including code interpretation. You are tasked with analyzing sales datasets from various Walmart stores to detect anomalies and explain them using relevant factors such as weather conditions, promotions, and natural disaster flags.

Instruction:

When provided with a dataset, perform a comprehensive analysis to identify any outliers or significant changes in sales trends. Utilize

statistical methods and visualizations to support your findings. For each anomaly detected, correlate it with possible explanatory variables present in the dataset and provide a detailed explanation.

User Prompt 1
**Upload the "hawaii_store_with_anomalies.xlsx",
"miami_store_with_anomalies.xlsx",
"texas_store_with_anomalies.xlsx",
virginia_store_with_anomalies.xlsx".

Context:

We have weekly sales data from four Walmart stores located in Texas, Miami, Virginia Beach, and Hawaii. Each dataset includes unit sales, weather conditions, promotions, and natural disaster flags.

Instruction:

- Analyze each dataset individually to identify any weeks where unit sales significantly deviate from the historical average.
- Generate time series plots of weekly unit sales for each store, highlighting any outliers or anomalies.
- For each anomaly detected, investigate potential causes by examining variables such as weather conditions, promotions, and natural disaster flags.
- Provide detailed explanations of your findings for each store.
- Please double check the data and use any anomaly detection models
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Sort the data in ascending order before plotting

User Prompt 2

Context:

After analyzing each store individually, we aim to perform a comparative analysis across all four stores to identify patterns or anomalies that may be common or unique.

Instruction:

- Compare the sales trends across all four stores.
- Identify any significant differences or anomalies in sales performance between the stores.
- Correlate any anomalies with factors such as weather conditions, promotions, or natural disasters.
- Generate comparative visualizations to illustrate your findings.
- Provide a comprehensive report summarizing the similarities and differences in sales patterns across the stores.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting

User Prompt 3

Context: We need to identify and understand major anomalies in each of our store to ensure data quality and gain insights for future analyses

Instruction:

- Review the datasets for each store and identify major anomalies in the data.
- For each anomaly identified:
    - Specify which store it occurred in
    - Provide the exact date or time period when the anomaly occurred
    - Describe the nature of the anomaly (e.g., unusual spike in sales, sudden drop in traffic, etc.)
    - Offer a data-based explanation for the anomaly, if possible
- Distinguish between anomalies that might be due to data errors and those that could represent actual events.

- Sort the data in ascending order by date before analyzing and plotting.
- Use Code Interpreter if needed to perform calculations or create visualizations.
- **Very Important:** Do not hallucinate or make up information. Ground all findings and explanations in the actual data provided.
- If you identify patterns or trends in the anomalies across stores, include those observations.
- Recommend any data cleaning or preprocessing steps that might be necessary to improve data quality for future analyses.
- Please sort the data in ascending order before plotting

User Prompt 4

Context:

We are interested in understanding how external factors such as weather events, promotions, and natural disasters impact sales across different locations.

Instruction:

- For each store, identify weeks where sales significantly deviated from the norm.
- Analyze these weeks in detail, examining weather data, precipitation levels, promotion activities, and natural disaster flags.
- Determine if there are any common factors affecting sales anomalies across the different stores.
- Summarize your findings in a comprehensive report.
- Include visualizations that support your conclusions.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting

User Prompt 5

Context:

We want to understand the overall impact of weather events on sales across all stores.

Instruction:

- Investigate correlations between extreme weather conditions (e.g., high precipitation, extreme temperatures) and sales anomalies in each dataset.
- Analyze whether weather impacts are consistent across different locations or vary by store.
- Provide explanations and visual evidence (charts, graphs) for positive correlations found only.
- Discuss how these insights could inform future forecasting and operational planning.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting

User Prompt 6

Context:

We are interested in the effect of promotional activities on sales, especially during periods with unusual weather or natural disasters.

Instruction:

- Analyze the datasets to identify the impact of promotions on sales during periods of sales anomalies.

- Examine whether promotions mitigated the negative effects of adverse weather conditions or enhanced the positive effects.
- Compare the effectiveness of promotions across different stores.
- Provide visualizations and detailed explanations to illustrate your findings.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting

User Prompt 7

Context:

We aim to create a comprehensive overview of sales performance and anomalies across all four stores to inform future forecasting models.

Instruction:

- Compile a report summarizing the sales trends and any identified anomalies for each store.
- For each anomaly, provide possible explanations based on the data (e.g., weather conditions, promotions, natural disasters).
- Suggest how these anomalies should be considered in future sales forecasting to improve accuracy.
- Include relevant visualizations to support your analysis.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting
- Please double check the data and use any anomaly detection models

User Prompt 8

Context:

We are looking to develop strategies to improve sales forecasting by understanding past anomalies.

Instruction:

- Based on the identified anomalies and their causes, recommend strategies for handling such anomalies in future forecasting models.
- Discuss how external factors should be incorporated into forecasting to improve accuracy.
- Provide examples from the data where including external factors would have improved forecast accuracy.
- **Very Important:** Do not hallucinate or make up numbers. Use Code Interpreter if needed. Ground your results in the given data.
- Please sort the data in ascending order before plotting

MODULE 8 - FORECASTING

LIVE DEMO 1

System Prompt

Context:

You are designing an AI forecasting assistant that leverages historical sales data to forecast future sales. The input data consists of sales and the goal is to predict future sales. It should use modern forecasting techniques and models, incorporating factors like seasonality, trend analysis, and any available historical patterns. The assistant needs to explain the forecast results, provide confidence intervals, and allow adjustments based on business insights. Please avoid hallucinating.

Instruction:

Using the code Interpreter:

- Load the CSV file containing historical sales data.
- Preprocess the data by handling any null values, normalizing where needed, and transforming the time series into a usable format for forecasting.
- Split the data into training and testing. Visualize both datasets on the same graph and color code them.
- Train an ARIMA model using the training set and evaluate on the testing set. Plot the predictions on top of the previous graph
- Plot model forecasts vs. along with confidence intervals.
- Provide a clear explanation of the forecast results, highlighting key drivers and patterns observed in the data.

Input Data:

CSV file with the following columns (example):

- Date: The timestamp or period (e.g., YYYY-MM-DD format or monthly/quarterly data).
- Sales: The sales figures for the respective region and period.

User Prompt

Attach the "sales data.csv" training file with the prompt:

You are tasked with conducting Exploratory Data Analysis (EDA) on the attached sales data.

Split the data into 70% training and 30% testing. Visualize both on the same graph and color code them.

Train an ARIMA model using the training set and evaluate on the testing set. Plot the predictions on top of the previous graph

Perform hyperparameters tuning using the following combinations:
p = [10, 11] #AR terms
d = 1 #Differencing terms
q = [10, 11] #MA terms

LIVE DEMO 2

8.1 System Prompt

Context:

You are designing an AI forecasting assistant that leverages historical sales data to forecast future sales. The input data consists of sales figures across different regions, and the goal is to predict sales per region and for the company as a whole. It should use modern forecasting techniques and models, incorporating factors like seasonality, trend analysis, and any available historical patterns. The assistant needs to explain the forecast results, provide confidence intervals, and allow adjustments based on business insights. Please avoid hallucinating

Instruction:

Using the code Interpreter:

- Load the CSV file containing historical sales data by region.
- Analyze the dataset, asking clarifying questions if needed (e.g., about data frequency, specific regional codes, or missing values).
- Preprocess the data by handling any null values, normalizing where needed, and transforming the time series into a usable format for forecasting.
- Apply state-of-the-art forecasting techniques, such as ARIMA, Prophet, or machine learning models (e.g., Random Forest).
- Generate sales forecasts for each region as well as a combined forecast for the overall company.
- Account for seasonality, trends, holidays, or external factors influencing sales in the models.
- Present forecasts along with confidence intervals.

- Provide a clear explanation of the forecast results, highlighting key drivers and patterns observed in the data.
- Allow for adjustments to the model if needed (e.g., excluding certain data points or adjusting seasonality assumptions).

Input Data:

- CSV file with the following columns (example):
  - Date: The timestamp or period (e.g., YYYY-MM-DD format or monthly/quarterly data).
  - Region: Identifying code for the sales region (e.g., region1, region2, etc.).
  - Sales: The sales figures for the respective region and period.
  - (Optional) Promotion/Discount: Information on whether any promotions or discounts were running in that period (binary/continuous data).

The assistant will begin by asking for the dataset and further clarifications about the data structure, then build the forecast model accordingly.

Prompt 2:

Context

You are working with a dataset containing information on retail sales performance across various regions and product categories. This dataset includes metrics such as marketing spend, promotions, weather conditions, product pricing, and economic indicators. The dataset spans multiple fiscal years, quarters, months, and weeks, providing a rich temporal structure.

Instruction

The goal is to perform data cleaning, exploratory data analysis (EDA), visualization, and forecasting based on the provided dataset. You are required to:

1. Clean the dataset by handling missing values, inconsistent entries, and outliers.
2. Perform EDA to uncover patterns, correlations, and anomalies in the data.
3. Create visualizations that explain trends and relationships between variables.
4. Build a forecasting model for predicting unit sales across different regions and product categories.
5. Explain the factors that most significantly contribute to the forecast.
6. Calculate and present the confidence intervals for the forecast.
7. Highlight any interesting insights or stories that emerge from the analysis, supported by data and visualizations.

Input Data

- **Fiscal Variables**: Year, Quarter, Month, Week
- **Sales Metrics**: Unit sales, Average selling price, Planned average selling price, Planned units
- **Marketing & Promotions**: Marketing spend, Promotions, Holiday events, Natural disasters
- **Geographic Data**: Store location
- **Economic Indicators**: Unemployment rate, Core PCE YoY, Fed Funds rate, 30-year mortgage rate, Median home sale price, Real disposable personal income, Gas prices, Oil prices, Housing starts, Consumer sentiment
- **Weather Data**: Temperature, Precipitation
- **Product Data**: Product category

8.2 User Prompts
8.2.1 Prompt 1: Perform EDA on the data

Attach the *"miami_training.csv"* file with the prompt:

Context:

You are tasked with conducting Exploratory Data Analysis (EDA) on sales and economic data from the years 2020 to 2024. The aim is to uncover trends, relationships, and insights using various visualizations and statistical analysis. The data contains features related to unit sales, product categories, pricing, marketing spend, and economic indicators such as unemployment rates and inflation (Core PCE). It also includes time-based features like months and weeks, along with promotion type data that affects sales.

Instruction:

Perform EDA on the given dataset by generating multiple visualizations and identifying key patterns and relationships. This will involve sorting the data, identifying trends, and using statistical and graphical methods to extract insights.

Input Data:

- Dataset: Sales and economic data for the years 2020 to 2024.
- Features:
    - **Time-based**: Date, month, week
    - **Sales**: unit_sales, avg_selling_price
    - **Categorical**: product_category
    - **Numeric**: temperature, marketing_spend, planned_avg_selling_price
    - **Economic Indicators**: Unemployment %, Core PCE YoY %
    - **Promotion Type**: Continuous value affecting sales

Structured Prompt for the EDA Task:

**sort the values in ascending order before plotting (values, dates, etc.)

Step 1: Data Preprocessing

- **Sort the data** by date in ascending order to prepare for time-based analysis.

- Handle any **missing data** by either filling or removing missing values where appropriate.
- **Detect and handle outliers** in numeric features.

Step 2: Time Series Analysis

- Plot **time series trends** for:
  - unit_sales over time.
  - avg_selling_price over time.
- Identify any **seasonality** or **trends** in sales and pricing.

Step 3: Category-wise Analysis

- Create **bar charts** to display the breakdown of unit_sales across different product_category values.
- Analyze the **effect of promotion type** (continuous value) on unit_sales across various product_category categories using a plot.

Step 4: Correlation and Relationships

- Generate a **correlation heatmap** for the following numeric features:
  - temperature
  - marketing_spend
  - planned_avg_selling_price
  - Economic indicators: Unemployment %, Core PCE YoY %
- Comment on any strong correlations or patterns in the heatmap.

Step 5: Economic Indicator Analysis

- Plot **line charts** for trends in key economic indicators (Unemployment %, Core PCE YoY %) and **analyze their impact** on sales trends.
- Explore how **promotions** and **economic conditions** have influenced sales over time.

Step 6: Distribution and Variability

- Generate **box plots** to examine the distribution of unit_sales across:
    - Different months.
    - Different weeks (week of the year).
- Comment on any notable patterns or variability in sales distributions.

8.2.2 Prompt 2: Split the data into training and testing

Context:

You are working on a machine learning task where the goal is to build a predictive model for unit_sales. The dataset includes various input features that reflect economic indicators, product-related factors, and environmental variables over time, and spans from 2020 to 2024.

Instruction:

Prepare and structure the dataset for training and testing. Specifically:

1. **Training Dataset (2020-2023)**:
    - The training input features (X_train) should consist of several columns, including:
        - Temporal features: fiscal_year, month, week, Fiscal Quarter
        - Environmental features: temperature, precipitation, natural_disaster
        - Economic indicators: Unemployment %, Core PCE YoY %, Fed Funds %, 30yr Mortgage %, Median US Home Sale Price, Real Disposable Personal Income, Avg US Gas Prices (Regular), WTI Crude Oil Spot, Housing Starts, U of M Consumer Sentiment
        - Product-related features: marketing_spend, store, promotion, product_category,

planned_avg_selling_price, promotion_type, Price, holiday

- The target variable for the training set (Y_train) is unit_sales.

Input Data:

You have the following columns available in your dataset:

- **Temporal Data**: fiscal_year, month, week, Fiscal Quarter
- **Environmental Variables**: temperature, precipitation, natural_disaster
- **Economic Indicators**: Unemployment %, Core PCE YoY %, Fed Funds %, 30yr Mortgage %, Median US Home Sale Price, Real Disposable Personal Income, Avg US Gas Prices (Regular), WTI Crude Oil Spot, Housing Starts, U of M Consumer Sentiment
- **Product Data**: marketing_spend, store, promotion, product_category, planned_avg_selling_price, promotion_type, Price, holiday
- **Target Variable**: unit_sales (which should be reserved for the test dataset)

8.2.3 Prompt 3: Train ARIMA, SARIMA, and Random Forest

Context:

You are tasked with building and training various forecasting models for predicting unit_sales. The dataset includes temporal, environmental, and economic features, as well as product-related factors from 2020 to 2024. The objective is to evaluate different models and determine which one provides the most accurate forecasts for unit_sales using both univariate and multivariate techniques.

Instruction:

Follow these steps to build, train, and evaluate the models:

1. **Preprocessing:**

   - **Category Encoding**: Perform encoding of categorical values in the dataset (e.g., one-hot encoding or label encoding for features such as promotion, etc.).
   - **Handle Missing Values**: Replace any missing (NaN) values in the dataset with zero before processing the data for the model.

2. **Model Training:**

   - Using the training dataset (X_train, Y_train) prepared from 2020 to 2023, build and train the following models in sequence:
     - **ARIMA**: A univariate time-series model using only the unit_sales data. Forecast unit_sales using time as the primary feature.
     - **SARIMA**: A seasonal variant of ARIMA, incorporating seasonal components like holidays and promotions. This model accounts for periodic trends and patterns in the data.
     - **Random Forest Regressor**: A multivariate machine learning model that uses the full feature set (X_train) to predict unit_sales by modeling complex relationships between the features.

3. **Model Evaluation:**

   - Validate each model using the 2024 data (test set). Evaluate the predictions using the following performance metrics:
     - **Mean Squared Error (MSE)**
     - **Root Mean Squared Error (RMSE)**
   - Compare the performance of each model based on these metrics.

4. **Model Selection:**

- ○ Rank the models based on their accuracy and error metrics.
- ○ Provide insights on which model is the best for forecasting unit_sales, considering both prediction accuracy and the nature of the data.

Plot the predictions vs. ground truth data

8.2.4 Prompt 4: Grid Search Optimization

Context:

You have already evaluated several machine learning models for a task and want to refine the best-performing model further. The goal is to perform hyperparameter optimization using a validation set (2024 data) using **GridSearchCV**. If **GridSearchCV** does not work, the goal is to manually train models with predefined hyperparameters.

Instruction:

1. Select the best-performing model from the models you've already evaluated.
2. Use **GridSearchCV** to perform hyperparameter optimization on this model, **without using cross-validation**.
3. Limit the search space to tree-based models with the following hyperparameters:
   - ○ **Number of trees**: 100, 200, and 300.
   - ○ **Tree depth**: 3, 4, and 5.
4. If **GridSearchCV** fails:
   - ○ **Manually build six models** with the following combinations of parameters (trees: 100, 200, 300 and depths: 3, 4, 5).
   - ○ Evaluate each model using the **validation set (2024 data)**.
   - ○ Choose the best model based on validation performance.

8.2.5 Prompt 5: Randomized Search Optimization

Context:

You have already evaluated several models and identified the best-performing one. Now, you want to further optimize this model using **RandomizedSearchCV** instead of GridSearchCV to make the process quicker. The goal is to refine the model by selecting a limited number of hyperparameters based on previous results, optimizing efficiency. The best hyperparameters will be chosen based on performance on the validation set (2024 data).

Instruction:

1. Select the best-performing model from your previous evaluations.
2. First, use **RandomizedSearchCV** or manually build 6 models with the following hyperparameters:
   - **Number of trees**: 100, 200, and 300.
3. Evaluate hyperparameters on the **validation set (2024 data)** without cross-validation.
4. If **RandomizedSearchCV** fails:
   - Manually select 6 random combinations of hyperparameters that worked well previously.
   - Build and evaluate these models on the **validation set** to find the best parameter combination.

8.2.6 Prompt 6: Retrain the model on 2020-2024 data

Context:

Retrain the model on the full dataset (2020–2024) using those optimized settings and prepared for inference on the 2025 data.

Instruction:

1. Use the hyper-parameters based on what we discovered earlier
2. If the hyperparameter search failed, train the model directly on data from **2020 to 2024** without tuning.

3. If hyperparameter optimization is successful, retrain the model on the entire dataset from **2020 to 2024** using the optimized hyperparameters.
4. Create the feature importance plot for the trained model

8.2.7 Prompt 7: Perform inference on the 2025 data

Note: Attach the miami_inference.csv file to the prompt

Context:

You have retrained the model (with optimized hyperparameters or without if tuning failed) on the full dataset from 2020 to 2024. The task now is to perform predictions on the **2025 inference data** using the same feature set. The features for this prediction include various economic, promotional, and seasonal variables. If the **month** value is missing, it should be inferred from the combination of **week** and **fiscal_year**.

Instruction:

1. **Input data:** X_inference_2025 should contain the following features:
    - fiscal_year, month, week, temperature, precipitation, marketing_spend, store, promotion, holiday, natural_disaster, product_category, planned_avg_selling_price, Fiscal Quarter, Unemployment %, Core PCE YoY %, Fed Funds %, 30yr Mortgage %, Median US Home Sale Price, Real Disposable Personal Income, Avg US Gas Prices (Regular), WTI Crude Oil Spot, promotion_type, Price, Housing Starts, and U of M Consumer Sentiment.
2. **Missing Month Handling**:
    - If the month feature is missing in the data, infer it using the week and fiscal_year. You can infer month from week by:

- Converting the week number to a month using the fiscal year calendar. For example, dividing week by 4 or 5 (depending on fiscal year settings).
3. **Prediction**:
   - Use the **retrained model** (with optimized hyperparameters or fallback training).
   - Ensure the same feature set used in training is applied to the **2025 inference data**.
   - Output the predicted values for unit_sales for 2025.

8.2.8 Prompt 8

Context:

After predicting the unit_sales for 2025, you want to generate a structured set of outputs, including a CSV file with the predictions, a detailed insights and recommendations report, and documentation on the model, hyperparameters, and transformations for reproducibility.

Instruction:

1. **Generate CSV for Predicted Output:**

   - Save the predicted unit_sales for each row in the X_inference_2025 dataset.
   - The CSV should include:
     - All the input features: fiscal_year, month, week, temperature, precipitation, marketing_spend, store, promotion, holiday, natural_disaster, product_category, planned_avg_selling_price, Fiscal Quarter, Unemployment %, Core PCE YoY %, Fed Funds %, 30yr Mortgage %, Median US Home Sale Price, Real Disposable Personal Income, Avg US Gas Prices (Regular), WTI Crude Oil Spot, promotion_type, Price, Housing Starts, U of M Consumer Sentiment.

■ The forecasted unit_sales for 2025.
2. **Output:** Save the file as predicted_unit_sales_2025.csv.

3. **Generate Insights and Recommendations Report (Text File):**

   ○ **Summary of EDA (Exploratory Data Analysis) Findings**: Provide key insights from the initial data exploration (e.g., trends in features such as sales spikes during holidays or the impact of economic variables on sales).
   ○ **Model Selection and Performance Comparison**: Describe the models that were considered and their performance metrics.
   ○ **Hyperparameter Optimization**: Summarize how hyperparameter optimization was done and what parameters were tuned.
   ○ **Final Model Choice**: Explain why the final model was chosen (e.g., best performance on validation data or interpretability).
   ○ **Forecasting Insights for 2025**: Discuss trends and patterns in the forecasted unit_sales for 2025, such as projected increases in sales during specific periods.
   ○ **Business Recommendations**: Based on the forecast, provide actionable recommendations (e.g., increase inventory during expected high-sales months, adjust pricing based on economic trends).
   ○ **Add Visuals**: If possible, embed **feature importance plots** or other relevant visualizations.
4. **Output:** Save this as insights_and_recommendations_2025.txt.

5. **Generate Model and Hyperparameters File (Text File):**

   ○ Document all steps required to **recreate the model**:
      ■ **Data cleaning and transformation**: Specify how missing values were handled, any feature engineering done (e.g., inferring month from week), and any

encoding (e.g., one-hot encoding for categorical features).
   - **Feature list and data types** used during training.
   - **Hyperparameters**: List the final hyperparameters used after tuning (or fallback values).
   - **Training methodology**: Provide details on how the model was trained (e.g., training data range, loss function).
   - **Model-specific instructions**: Any specific configurations or libraries needed to reproduce the environment.
6. **Output:** Save this as model_reproduction_instructions.txt.

MODULE 9 - REPORTING

PRACTICAL PROJECT: DATA REPORTING

Upload the headers description and Reporting Template in the system prompt
Enable Code Interpreter

SYSTEM PROMPT
You are a professional report generator for retail companies. Your task is to analyze structured data and generate clear, concise, and actionable reports. The report should summarize total sales, highlight top categories and products for each store, provide inventory insights, and recommend actionable items for the retail team. Use professional language, maintain a consistent format, and avoid redundancy.

USER PROMPT

Upload the "stores_reporting_data.csv" data

Generate a report using the attached data based on the template.

MODULE 10 - VARIANCE ANALYSIS

LIVE DEMO 1: GEN AI FOR SALES AND VOLUME VARIANCE ANALYSIS (SIMPLE USE CASE!)

Enable Code Interpreter

System Prompt:

You are an expert financial analyst specializing in variance analysis for retail businesses. Walmart store has provided actual vs. predicted sales and volume datasets. Your role is to analyze these datasets, generate insights, and provide actionable recommendations based on your expertise.

User Prompt 1: Perform Variance Calculations

Context:
Walmart has shared actual vs. predicted sales and volume datasets for its product categories. The goal is to analyze these datasets to identify key variances as part of a comprehensive report.

Input:
Datasets containing actual vs. predicted sales and volume data for various product categories.

Instruction:

1. Perform the following variance calculations for each row in the dataset:
   - **Sales Variance** = Actual - Predicted.
   - **Sales Variance %** = (Variance / Predicted) × 100.
   - **Volume Variance** = Actual - Predicted.
   - **Volume Variance %** = (Variance / Predicted) × 100.
2. Color code each cell:

- ○ **Green** for favorable variances (positive impact).
- ○ **Red** for unfavorable variances (negative impact).

Output:
A table with variance calculations and color-coded cells for easy interpretation.

User Prompt 2: Interpret Variances

Context:
The calculated variance data provides insights into product category performance. The goal is to interpret these variances to understand their significance and identify potential root causes.

Input:
Tables with variance calculations (Sales Variance, Sales Variance %, Volume Variance, Volume Variance %) for each product category.

Instruction:

1. Categorize product category performance as:
    - ○ Top-performing (favorable variances).
    - ○ Underperforming (unfavorable variances).
2. Explain the significance of variances in terms of business impact (e.g., profitability, inventory costs).
3. Identify potential root causes for the observed variances (e.g., market trends, operational inefficiencies, supply chain disruptions).

Output:
A section summarizing performance categories, variance significance, and potential root causes.

User Prompt 3: Develop Strategic Recommendations

Context:
Based on the variance analysis and interpretation, actionable recommendations are needed to help Walmart's address discrepancies and optimize performance.

Input:
Insights from variance interpretation, including categorized product performance and identified root causes.

Instruction:

1. Provide targeted strategies to improve or sustain performance for each product category.
2. Propose inventory management approaches to address discrepancies (e.g., adjusting stock levels, improving demand forecasting).
3. Offer market positioning recommendations to leverage strengths and address weaknesses.

Output:
A structured section with strategic recommendations tailored to Walmart's business objectives.

LIVE DEMO 2: WALMART VARIANCE ANALYSIS

SYSTEM PROMPTS
** Enable Code Interpreter**

PROMPT 1

Context:

You want to analyze sales data to identify variances between actual and forecasted or budgeted sales. The goal is to quickly detect significant

discrepancies, compute both absolute and percentage differences, assess performance, and provide actionable recommendations.

Instruction:

- Load actual sales data and forecasted/budgeted sales data for one or multiple stores.
- Identify significant variances (positive or negative) between the actual sales and the forecasted/budgeted sales.
- Calculate both absolute and percentage variances.
- Classify the performance as either a "beat" or a "miss" relative to the forecast or budget.
- Provide a summary of the findings and suggest corrective actions or recommendations based on the analysis.

Input Data:

1. Actual Sales Data:
   - Data Source: (Specify if applicable)
   - Sales values by store and time period (e.g., daily, weekly, or monthly)

PROMPT 2

Context:

You are analyzing the variance in sales by examining various external and internal factors. The objective is to understand which factors, such as marketing spend, promotions, weather conditions, and economic indicators (e.g., gas prices, housing starts), are most correlated with significant variances between actual and forecasted sales, and how they impacted the sales performance.

Instruction:

- Analyze the variance in actual versus forecasted sales for each store.

- Investigate key factors such as:
  - Marketing spend
  - Promotions
  - Temperature
  - Economic indicators (e.g., gas prices, housing starts)
- Identify correlations between these factors and significant sales variances.
- Explain how these factors contributed to the differences between actual and forecasted sales performance.
- Highlight which factors had the strongest influence and discuss the nature of the impact (positive or negative).

Input Data:

1. Sales Data:
   - Actual and forecasted sales figures for each store.
2. Factor Data:
   - Marketing Spend: Amount spent on marketing during the analyzed period.
   - Promotions: Promotional activities and their timelines.
   - Temperature Data: Local weather conditions (e.g., daily or weekly temperature).
   - Economic Indicators: Data on gas prices, housing starts, or other relevant economic variables.

PROMPT 3

Context:

You want to visualize the sales variance over time for one or multiple stores and highlight key trends, such as significant periods of positive or negative variance. Additionally, you aim to illustrate the relationship between different factors (e.g., marketing spend, promotions) and sales variance using correlation heatmaps.

Instruction:

- Create time-series visualizations that display the actual versus forecasted sales variance for one or multiple stores.
- Highlight periods of significant positive or negative variance.
- Use color-coding or markers to indicate the magnitude of the variance (e.g., high positive/negative variance).
- Generate correlation heatmaps that show the relationships between key factors, such as:
  - Marketing spend
  - Promotion types
  - Temperature
  - Economic indicators (e.g., gas prices, housing starts)
- Ensure the visualizations clearly communicate which factors are most strongly correlated with sales variances and their direction (positive or negative).

Input Data:

1. **Sales Data**:
   - Actual sales vs. forecasted sales data over time, for one or multiple stores.
2. **Factor Data**:
   - **Marketing spend**: Data over the same period as sales variance.
   - **Promotion types**: Information on promotional events and their duration.
   - **Temperature**: Daily or weekly temperature data for the regions.
   - **Economic indicators**: Gas prices, housing starts, or other relevant economic data.

PROMPT 4

Context:

You want to develop actionable recommendations based on the analysis of sales variance. The goal is to reduce negative variances or capitalize on positive variances by adjusting factors such as marketing spend, promotions, or external conditions. The recommendations must be data-driven, avoiding unsupported assumptions.

Instruction:

- Review the analysis of sales variance and identify key drivers of both positive and negative variance.
- Provide actionable recommendations based on the findings:
  - For negative variances: Suggest corrective actions, such as increasing marketing spend, adjusting promotional strategies, or optimizing other relevant factors.
  - For positive variances: Recommend strategies to sustain or amplify the positive impact.
- Ensure that each recommendation is clearly grounded in the data (e.g., "If lower marketing spend is linked to lower sales, recommend increasing spend during similar periods.").
- Avoid making recommendations that lack supporting evidence or data correlations.

Input Data:

1. **Sales Variance Analysis**:
   - Summary of the periods with significant variances (positive or negative) and the associated factors.
2. **Factor Analysis**:
   - Key factors (e.g., marketing spend, promotions, temperature, economic indicators) and their correlation to sales variances.
3. **Actionable Insights**:
   - Insights from the data that suggest clear actions, such as increasing spend, adjusting promotional timing, or preparing for economic shifts.

PROMPT 5

Context:

You want to create a detailed, text-based report that explains the sales variance for each store. The report should identify the most important factors contributing to the variance, provide a comprehensive performance summary, and offer data-backed suggestions for improvement. All insights should be firmly based on the actual data to avoid unsupported assumptions.

Instruction:

- Generate a detailed report explaining the sales variance for each store.
- Highlight the most important factors (e.g., marketing spend, promotions, weather, economic indicators) that contributed to both positive and negative variances.
- Provide a well-written summary of each store's overall sales performance, identifying whether it was a "beat" or a "miss" compared to the forecast.
- Offer actionable suggestions for improvement, such as:
  - Adjusting marketing efforts
  - Timing promotions more effectively
  - Responding to external conditions like weather or economic shifts
- Ensure all insights and recommendations are directly grounded in the data, avoiding assumptions or unsupported inferences.

Input Data:

1. **Sales Data**:
   - Actual vs. forecasted sales for each store.
2. **Contributing Factors**:
   - Marketing spend
   - Promotional events and timings
   - Temperature data
   - Economic indicators (e.g., gas prices, housing starts)
3. **Performance Insights**:

- Key trends or correlations identified from the data.
4. **Recommendations**:
  - Data-backed suggestions to improve future performance or sustain positive trends.

USER PROMPTS

Prompt 1:

Context:

You are performing an initial analysis on the sales data for multiple stores to identify key trends and overall performance. The analysis will focus on total sales, planned vs. actual performance, and variances. You want to generate key metrics such as average weekly sales, average forecast differences, and identify patterns across products or stores.

Instruction:

- Load and, if necessary, combine the dataset for all stores.
- Perform a general analysis of the data, focusing on:
  - Total sales for each store or product
  - Planned vs. actual units sold (unit_sales vs. planned_units)
  - Planned vs. actual average selling prices (avg_selling_price vs. planned_avg_selling_price)
  - Overall variance between planned and actual figures.
- Provide a summary of key metrics:
  - Average weekly sales across stores or products.
  - Average difference between actual and forecasted sales (both in units and price).
  - Key patterns or trends observed in the data, such as sales dips or peaks.
- Highlight any large variances or noteworthy deviations from the planned numbers.

Input Data:

1. Sales Data:
   - Actual sales data (unit_sales, avg_selling_price) for all stores over the analysis period.
2. Planned/Budgeted Data:
   - Forecasted sales data (planned_units, planned_avg_selling_price) corresponding to the actual sales period.

Prompt 2:

Context:

You want to calculate and analyze the sales variance for each store by comparing actual sales to forecasted sales. The goal is to identify the stores with the largest positive and negative variances, and determine whether these variances are driven by over-forecasting or under-forecasting.

Instruction:

- Calculate the sales variance for each store by comparing:
    - **Actual units sold** (unit_sales) vs. **forecasted units** (planned_units).
    - **Actual average selling price** (avg_selling_price) vs. **forecasted average price**(planned_avg_selling_price).
- Highlight the stores with the **largest positive** and **largest negative** variances in both unit sales and average selling price.
- Summarize the top 5 variances (both positive and negative) and provide insights into:
    - Whether the variance was driven by **over-forecasting** (planned sales too high) or **under-forecasting**(planned sales too low).
- Provide an overview of which stores experienced the most significant discrepancies and potential reasons for the variances

(e.g., unexpected market conditions, misaligned pricing strategies).

Prompt 3:

Context:

You want to generate visualizations for stores with significant positive or negative sales variances, focusing on units sold and average selling price over time. The goal is to create meaningful plots that clearly illustrate variance trends, including line charts for each product category per store, showing variance over time.

Instruction:

- For each store with **significant variances** (positive or negative), generate the following visualizations:
    - **Bar Plots of Unit Sales Variance**:
        - **Carefully plot** the **average variance of units sold** for each store in a way that makes sense.
        - Group the data by **side-by-side bar plots** showing each store and product category, comparing them for each quarter and monthly data from the latest year.
    - **Bar Chart for Average Selling Price Variance**:
        - Create a bar chart showing the **variance in average selling price** for each store.
        - Ensure a similar plot is created for variance across **product categories**.
    - **Line Charts for Product Category Variance Over Time**:
        - For each product category and each store separately, create a **line chart** showing the variance over time.
        - Plot the variance data over the relevant time periods (e.g., months, weeks, or quarters) in ascending order.
    - **Sort Time Data**:

- Ensure the time data is sorted in **ascending order** before plotting.
- Create a date column combining **month, week, and year** so it's easier to plot and interpret.
  - **Handle Missing Data**:
    - Make sure there are no missing values (NaNs) in the date column, quarters, or other important fields before generating the plots.
  - **Visual Elements**:
    - Color code bars to indicate the magnitude of variance (positive or negative).
    - Label charts with store names, product categories, and time periods clearly.
- **Double-Check the Plots**:
  - Ensure that all plots have rendered correctly and there are no issues with the data.
  - Verify that the visualizations accurately represent the variance trends and are easy to interpret.

Prompt 4:

Context:

You want to perform a detailed root cause analysis for the stores with the highest sales variances. The goal is to explore potential contributing factors such as marketing spend, promotions, temperature, and external economic indicators (e.g., gas prices, housing starts) to identify any significant correlations or patterns driving the variance.

Instruction:

- Focus on the **stores with the highest positive or negative sales variances**.
- Analyze the following **potential contributing factors**:
  - **Marketing spend** during the variance period.

- **Promotions**: Identify if any specific promotions were running during the periods of significant variance.
      - **Temperature data**: Check if weather conditions correlate with the variance.
      - **Economic indicators**: Look at external factors such as **gas prices** or **housing starts** to see if they align with periods of high variance.
  - **Correlate these factors** with the observed sales variances to determine whether any significant patterns or relationships exist.
      - Calculate correlation coefficients to quantify the strength of the relationships (e.g., between marketing spend and sales variance).
      - Highlight any factors that have a statistically significant correlation with variance.
  - Summarize your findings, indicating:
      - **Key contributing factors** that explain the variances.
      - **Patterns** (e.g., certain stores being more sensitive to marketing changes or weather conditions).
      - Any **anomalies** or surprising results where no clear factor seems to explain the variance.

Prompt 5:

Context:

You want to analyze if there was any period where building materials experienced significant sales variance, using the detailed dataset provided. The data includes factors such as economic indicators, marketing spend, promotions, and other external conditions that may explain the variances. The objective is to identify which factors contributed most to the variance.

Instruction:

- Focus the analysis on **building materials** within the **product_category**
- Investigate the following factors to determine their potential impact on sales variance:
  - **Sales and Forecast Comparison**:
    - Compare **unit_sales** and **avg_selling_price** with **planned_units** and **planned_avg_selling_price** for building materials.
    - Identify whether the variance was due to over-forecasting (forecasted sales too high) or under-forecasting (forecasted sales too low).
  - **Weather Conditions**:
    - Analyze the influence of **temperature** and **precipitation** on building material sales.
    - Check if unusual weather patterns (e.g., extreme heat or rainfall)
  - **Promotions and Discounts**:
    - Review the presence and scale of any **promotions** and **promotion_discount** offered during the variance period.
    - Examine if large discounts or promotional activities caused deviations in expected sales.
  - **Marketing Spend**:
    - Assess the **marketing_spend**
    - Compare periods with high or low marketing spend to variance in sales, checking if insufficient marketing may have led to missed forecasts.
  - **Economic Indicators**:
    - Analyze external economic factors from the dataset:
      - **Unemployment %**: Higher unemployment may reduce demand for building materials.
      - **30yr Mortgage %**: Rising mortgage rates could have slowed down homebuilding and renovation, reducing sales of building materials.

- **Housing Starts**: Check if a decline in housing starts correlates with decreased demand for building materials.
- **Avg US Gas Prices** and **WTI Crude Oil Spot Price**: Rising fuel costs could have increased transportation expenses for building materials, potentially impacting pricing or sales volume.
- **Median US Home Sale Price**: A housing market slowdown could reduce demand for materials related to construction and renovation.
- **Real Disposable Personal Income**: Analyze if consumer purchasing power affected sales.
- **U of M Consumer Sentiment**: A drop in consumer confidence could have contributed to a decrease in home improvement spending.
    - **Holidays and Natural Disasters**:
        - Review the impact of **holidays** and **natural_disasters** on sales. Holidays may boost or reduce demand for building materials, while disasters could lead to an increase in demand for repairs or stock shortages.
- **Summarize Findings**:
    - Identify the key factors contributing to the variance, focusing on the most significant correlations.
    - Explain whether the variance was primarily driven by internal factors (promotions, marketing spend) or external ones (weather, economic conditions).
    - Pinpoint if any specific events (e.g., natural disasters or housing market changes) were linked to the sales shifts.

MODULE 11: SCENARIO PLANNING

PRACTICAL PROJECT 1: GENAI FOR SCENARIO PLANNING

System Prompt:

*italic***Enable Code Interpreter**

You are an expert financial analyst specializing in scenario planning and predictive analytics. Using the dataset provided, your task is to:

1. Build a predictive model to forecast Sales (USD) based on the provided data features.
2. Generate three scenarios: Best Case, Worst Case, and Moderate Case, with appropriate justifications for the assumptions in each scenario.
3. Output clear, actionable insights for each scenario.

User Prompts:

Attach the "scenario_planning_dataset.csv" dataset

1. **Exploratory Analysis:**
    -   Perform exploratory data analysis (EDA) to understand trends, correlations, and key drivers of Sales (USD). Plot a subplot containing all line plots for each feature vs date individually.
2. **Model Building:**
    -   Use linear regression and random forest regression techniques to forecast Sales (USD) based on the features: Promotion Spend (USD) and Housing Market Index.
3. **Scenario Generation:**
    -   Based on the predictive model, simulate the following scenarios:
        -   Best Case: High promotion spend and a strong housing market.
        -   Worst Case: Low promotion spend and a weak housing market.
        -   Moderate Case: Average values for promotion spend and the housing market.

PRACTICAL PROJECT 2: WALMART SCENARIO PLANNING

SYSTEM PROMPT

Prompt 1

Context:

You are developing a scenario planning assistant that analyzes both internal and external factors affecting a company's sales performance and market share across various regions. The assistant should be capable of assessing the impact of changes in these factors on key business metrics, such as sales performance and market share. The historical data provided will contain key indicators such as marketing spend, promotions, planned units sold (internal factors), and macroeconomic data like the unemployment rate, fed funds rate, and consumer sentiment (external factors). Additionally, there will be an output file from a prior experiment containing the model used, any preprocessing steps, and other related details to guide the analysis.

Instruction:

Build a Scenario Planning Assistant that can perform the following tasks:

1. Data Ingestion: Import the historical data from a CSV file and parse the internal and external factors.
2. Model Integration: Use the details from the text file (from a previous experiment) that includes the model used and preprocessing details to guide how to process the data and perform analysis.
3. Scenario Analysis:
   - Identify correlations and trends between internal and external factors and their effects on sales performance and market share.

- Assess how changes in these factors (e.g., increase in unemployment rate or shift in marketing spend) will likely affect the company's performance across different regions.
4. Visualization: Provide visualizations or reports that highlight predicted outcomes for sales and market share under different scenarios.
5. Regional Insights: Analyze and provide separate insights for different geographical regions (if available in the data).
6. Recommendations: Offer actionable recommendations based on the scenario analysis for optimizing sales and market performance in response to external and internal changes.

Input Data:

- CSV File: Historical data containing internal (e.g., marketing spend, promotions) and external factors (e.g., unemployment rate, consumer sentiment).
- Text File: Details from a previous experiment, including the model used and any preprocessing steps applied to the data.

Attach miami_training.csv and model_reporoduction_instructions.txt

Prompt 1

Context

You're working on a forecasting model using time-series data from 2020 to 2024. The task involves preparing this data, followed by training the model. Hyperparameter optimization is not required, meaning the default parameter values will be used. Don't use avg_selling_price for training and inference

Instruction

*Don't use avg_sales_price as training feature

1. Data Preprocessing:
   ○ Step 1: Load and explore the data: We will first load the CSV file to examine its structure, check for missing values, and ensure that it is in the right format.
   ○ Step 2: Handle missing values: If there are any missing values, we can either fill them in using forward-fill or interpolation, or remove them, depending on the situation.
   ○ Step 3: Feature engineering: We will create relevant features from the time-series data, such as lags, rolling averages, or additional time-based features (like month, day, etc.).
   ○ Step 4: Data splitting: The data will be divided into training and testing sets. However, since we are using the entire dataset, we'll utilize the complete dataset for model training.
2. Model Training:
   ○ Step 5: Select and Train the model: Choose the model based on the model description file provided, use the parameters

provided and also do the categorical transformation mentioned in the provided file

Prompt 2

Context:

You are analyzing a machine learning model that we built, and you want to understand which features (variables) contribute most to the model's predictions and how those features affect the outcome of the model.

Instruction:

1. **Generate feature importance plots**: This will show the ranking of features based on their contribution to the model's predictions.
2. **Generate partial dependence plots (PDP)**: This will show how the target variable changes as one feature is varied, holding all others constant.
3. **Explain the model's learning**: Provide an interpretation of the feature importance and partial dependence plots to explain what the model has learned from the data.

Prompt 3

Attach miami_inference.csv"

Context:

You have inference data for 2025, which contains the same features as your training data (from 2020 to 2024) but without the target variable

(units sold). The goal is to use the trained model to predict units sold for 2025 and do the processing of features similar to train data

Instructions:

1. **Prepare the Inference Data:**
   - Ensure that the 2025 inference data matches the structure and format of the training data (e.g., consistent feature names, types, and any necessary preprocessing like scaling or encoding that was done during training).
   - Check for missing values in the 2025 data and handle them accordingly.
2. **Load the Trained Model:**
   - Load the forecasting model that was trained using data from 2020 to 2024. This model should be configured with the features and default hyperparameters used during training.
3. **Make Predictions for 2025:**
   - Input the 2025 inference data (features only) into the trained model.
   - The model will output predictions for the target variable (units sold) for each entry in the 2025 dataset.
4. **Interpret the Results:**
   - Once predictions are made, interpret them in the context of business problem. Look for trends or seasonality that the model has learned from the past data and how they are reflected in the 2025 predictions.
   - Visualize the predictions over time to assess whether the model captures expected patterns (e.g., spikes during certain months, trends over time).

Output:

- Predicted units sold for each row in the 2025 inference data.

Prompt 4

Context:

- Now I want to do what if analysis, so that we can do multiple scenario plannings
- You want to simulate how increasing the marketing_spend by 5% will affect the sales predictions for 2025.
- You also want to compare the original predictions (based on current marketing_spend) to the new predictions with the increased marketing spend, along with confidence intervals.

Instructions for Scenario Planning:

1. **Step 1: Generate Baseline Predictions (Current Marketing Spend)**

   - **Input:** Use the existing 2025 data (with original marketing_spend) to make baseline predictions.
   - **Output:** This will provide the current predictions for sales, which we'll compare against the adjusted scenario.
   -

2. **Step 2: Simulate the Increased Marketing Spend (+5%)**

   - **Input:** Modify the marketing_spend feature in the 2025 data by increasing it by 5%. For example, if the marketing_spend value for a particular record is $1000, increase it to $1050.

- **Output:** Feed this adjusted 2025 data (with the 5% increase in marketing_spend) into the trained model to generate the new predictions.
- **Compare:** This will allow you to compare the new predicted sales figures against the baseline predictions.

3. **Step 3: Visualize Results**

   You will create a **scatter plot** with the following components:

   - **Baseline (Current) Prediction Line:**
     - Plot a line representing the baseline predictions (i.e., predictions made with the current marketing spend for 2025).
     - Color - blue
   - **Scenario (Increased Marketing Spend) Prediction Line:**
     - Plot another line representing the predictions made after increasing marketing_spend by 5%.
     - color - green
   - Plotting instruction
     - x axis should contain 1-52 weeks of 2025, please make sure there is no scaling issues
     - Please sort the values in ascending order before plotting

Output and Plot Description:

Current Scenario:

- **Prediction Line for Current Marketing Spend:** A line representing the predicted units sold for 2025 using the existing marketing_spend.

New Scenario:

- **Prediction Line for Increased Marketing Spend:** A line showing the predicted sales for 2025 after applying the 5% increase to

marketing_spend. Compare this to the baseline prediction line to visualize the impact of the increased marketing budget.

Stats Summary:

- Provide the predicted numbers with and without increased spend ,we need baseline, with increase in market spend and % of increase

Prompt 5: Product-Level Focus

Context:

- We are now focusing on individual product-level predictions.
- You want to visualize the current and adjusted predictions for each product, compare them side by side, and include confidence intervals.

Instructions:

Step 1: Generate Predictions for Each Product

1. **Baseline Predictions (Current Marketing Spend):**

   - Use your trained model to predict 2025 sales for each product using the current marketing_spend.
2. **New Scenario Predictions (Increased Marketing Spend):**

   - Modify the marketing_spend for each product by increasing it by 5%.
   - Use the same trained model to predict sales for each product using this modified marketing_spend.
   -

Step 2: Create Plots at the Individual Product Level

You will create separate plots for each product. Here's the breakdown for each plot:

- **X-axis:** Date (2025)
- **Y-axis:** Predicted Sales (Units Sold)

Each plot will contain:

- **Current Marketing Spend Prediction Line:** A line for the predicted sales in 2025 based on the current marketing_spend.
- **Increased Marketing Spend Prediction Line:** A line for the predicted sales in 2025 with the 5% increased marketing_spend.

You'll repeat this plot for each product, showing how increasing the marketing spend by 5% influences the predicted sales for 2025.

Step 3: Plot Interpretation

For each product, the plot will allow you to:

- **Compare the Predictions:** See how much the sales predictions increase for 2025 when marketing spend is raised by 5%.
- **Product-Specific Insights:** Some products may show a stronger response to the marketing spend increase, while others might show little change. This provides valuable insights into the product-level elasticity to marketing efforts.

Visual Summary:

Each product-level plot will contain:

- **Prediction line for 2025 sales** using the current marketing spend.
- **Prediction line for 2025 sales** after increasing marketing spend by 5%.

Stats Summary:

- Provide the predicted numbers for each category with and without increased spend ,we need baseline , with increase in market spend and % of increase

Prompt 6: Focus on a Given Month

Context

You want to assess how **increasing marketing spend by 20%** and introducing a **10% promotion** for the entire month of April affects sales predictions. This scenario will be compared to the baseline (i.e., without the increased marketing spend or promotion).

The objectives are to:

- **Predict sales** for the entire month of April under the new scenario.
- **Compare sales for the entire month** of April to the baseline predictions.
- **Visualize** both the baseline and the new scenario, including confidence intervals for both.

Instructions

1. Baseline Prediction (No Changes):

- **Generate baseline predictions** using your model based on the **original 2025 data**, without any increase in marketing spend or the promotion. This prediction will reflect normal marketing levels and no promotion.
- **Store the predicted sales** for the first two weeks of April and for the entire month. These values will serve as your baseline for comparison.

2. Scenario with 20% Increase in Marketing Spend + 10% Promotion:

- **Increase marketing_spend by 20%** for the month of April This increase reflects your marketing investment during the promotional period.
- **Change the promotion_discount feature** to **10% discount** for entire month of April. This could be a binary variable or a numerical discount rate, depending on how your model processes promotions.
  - Example: You could add a new feature (e.g., promotion_discount) with a value of 10
- **Run the model** on this modified dataset (with the marketing increase and promotion) to generate new sales predictions.

3. Sales Change Analysis:

- **Compare the sales predictions** for entire month of April between the baseline and the new scenario (increased marketing spend + promotion).
- Calculate the **percentage change** in predicted sales for:
  - The entire month of April.
- This will quantify how much the 20% marketing spend increase and 10% promotion affect sales.

4. Visualization:

- visualize the scenarios -please sort the values before plotting

Prompt 7: Planned Inventory Effect

Context:

- You want to analyze how increasing planned_units (inventory or planned stock) by 10%, 15%, and 20% affects the predicted sales in 2025.
- This analysis will focus on **one product category** to show the impact of these increases.

Instructions for Scenario Analysis:

Step 1: Define the Product Category

Select a specific product category from your dataset to analyze. For example, let's assume the category is **"Home Appliances", do for each available category**

Step 2: Baseline Predictions for 2025

- **Input:** Use your trained model to predict sales for this product category using the existing planned_units values for 2025.
- **Output:** This will serve as the **baseline** against which the new predictions (with increased planned_units) will be compared.

Step 3: Simulate Increased Planned_Units

For the selected product category:

1. **Increase planned_units by 10%, 15%, and 20%:**
   - If planned_units for a particular product in January 2025 is 1,000 units, adjust it to 1,100 (10% increase), 1,150 (15% increase), and 1,200 (20% increase).
   - Repeat this adjustment for all months or time periods in 2025.
2. **Generate New Predictions for Each Scenario:**
   - Use the modified planned_units data to generate new sales predictions for each scenario.
     - **Scenario 1:** 10% increase in planned_units
     - **Scenario 2:** 15% increase in planned_units
     - **Scenario 3:** 20% increase in planned_units

Step 4: Compare Predictions

You will now compare the **baseline** predictions (with current planned_units) to the new predictions for each scenario (with increased planned_units).

Step 5: Visualization

You can create a **line plot** to visualize the impact of the different planned_units increases. The plot should include:

- **X-axis:** Time (e.g., months in 2025).
- **Y-axis:** Predicted Sales (units sold).
- **Lines Representing Each Scenario:**
    - **Baseline Prediction Line:** Sales predictions for the selected product category using current planned_units.
    - **Prediction Line for 10% Increase:** Sales predictions when planned_units is increased by 10%.
    - **Prediction Line for 15% Increase:** Sales predictions when planned_units is increased by 15%.
    - **Prediction Line for 20% Increase:** Sales predictions when planned_units is increased by 20%.

Interpretation of the Plot:

- **Impact on Sales:** The plot will show how increasing planned_units affects the predicted sales for the product category in 2025.
- **Comparison Across Scenarios:** You will be able to see whether the 10%, 15%, and 20% increases in planned_units result in proportional increases in sales, or if the effect diminishes at higher levels of inventory.
- Provide the overall results of the scenario like what happends with each these action as text with the actual number of units increased or decreased along with percentages

Prompt 8: Report Generation

Context:

- **Subject**: Sales Prediction Model Analysis.
- **Objective**: Summarize the key findings of the sales prediction model and explain how changes in specific features affect the predictions. Generate a report that can be shared across the organization with relevant visualizations (plots, charts, tables).

Instruction:

- Provide the **key findings** in relation to the sales prediction model.
- Explain the **impact of different features** on sales predictions (e.g., pricing, marketing spend, seasonality, etc.).
- Include **plots and charts** that show trends, feature importance, and relationships between variables.
- Generate a file that incorporates:
    - An executive summary
    - Key findings
    - Visualizations (plots, charts)
    - Recommendations for decision-making based on findings

Input Data:

- Key results from the sales prediction model (including feature importance).
- Data required to generate the plots (e.g., sales vs. marketing spend, sales vs. pricing, etc.).
- Any specific recommendations based on findings.

MODULE 12: INTRODUCTION TO MICROSOFT COPILOT

- Explore top daily updates from CoPilot
- I'm going on winter weekend getaway and need outfit ideas. What should I wear?
- Can you suggest some personalized gift ideas that would make a big impact?
- Generate an image of the pyramids at dawn showing burj Khalifa in the background
- Can you add some clouds to the sky?
- What is the revenue of Tesla in 2023? *Attach Tesla-10K-Report.pdf*
- Perform exploratory data analysis on this data Walmart_Sales_Data.csv
- What are the main revenue sources for Tesla in 2023? "Attach Tesla-10K-Report.pdf"
- Generate a PowerPoint file that summarizes the risk factors from this PDF. "Attach Tesla-10K-Report.pdf"

MODULE 13: MICROSOFT COPILOT AND POWERPOINT

Create a Presentation from Scratch

- Create a presentation about personal branding strategies for professionals, including LinkedIn optimization and online presence management.

- ****Show the talking points on the bottom of the slide****

- Create a presentation about personal branding strategies for professionals, including LinkedIn optimization and online presence management. Make the presentation geared towards generation Alpha with fun and cool features!

Modify Text, Images, and Speaking Notes

- ****How to auto rewrite the title of the presentation****

- ****How to condense bullet points on a presentation****

- ****Show how PowerPoint designer is now integrated with CoPilot and change the design of a sample slide****

- Add an image of employees in infront of their computers.
- Add an appropriate image to this slide.
- Translate the speaking notes on the bottom of the slides to arabic.
- Change the tone of the presentation to make it geared towards Gen Alpha
- Add a slide that shows the effective methods of personal branding. Ground your research to trusted sources.

Create Presentations from PDFs and Word documents

- Create a presentation from \Walmart-10K-2024.pdf
- Create a presentation from Walmart NYC Store Inventory Summary - #3051.docx

Create a presentation from Template
- Using this template, create a presentation about the world happiness report in 2024. Include factors that affect people's happiness.

Summarize and Organize Existing PowerPoint Slides
**Using "Sample Slide Deck.pptx" **

- Summarize this presentation
- Show Key Slides on this Deck
- Organize this presentation

Presentation Preparation & Delivery!

**Using "Sample Slide Deck.pptx" **

- Add speaking notes to the slide or Tell me what to say on each slide
- What questions might the audience ask about this presentation
- What content is missing from these slides?

CoPilot PowerPoint Slides Coaching!
**Using "Sample Slide Deck.pptx" **

- Tell me best practices for presenting this slide deck
- What are some common mistakes to avoid in presenting these slides?
- How can I make the presentation more engaging?

MODULE 14: MICROSOFT COPILOT AND WORD

Write Documents from scratch with CoPilot

- "Draft a report for Walmart management outlining a new initiative to improve in-store customer experience."
- "Make this more concise in one page only"
- "Draft a report for Walmart management using this outline
  - Objective: Enhance in-store customer experience to boost satisfaction, loyalty, and sales.

- Challenges: Address issues like long checkout times, product accessibility, and limited personalization.
    - Initiative: Implement solutions such as self-service kiosks, optimized store layouts, and enhanced mobile app features.
    - Plan: Roll out in phases with pilot testing, staff training, and milestone-driven progress tracking.
    - Impact: Measure success through customer satisfaction scores, sales growth, and operational efficiency metrics."
- **Show the ability to discard, keep, or regenerate**
- **Show regeneration feature and probabilistic nature of CoPilot**
- **Highlight text and auto rewrite**

Write documents from existing PDFs and Powerpoint files on OneDrive

- Summarize the risk section in this 10K document /walmart-10K-report.pdf
- Produce summary of /Walmart Outlook 2025-2027.pptx
- What are the key findings of this research paper? /Transformer-Based Deep Learning Strategies for SOX Estimation.pdf

Write documents from existing Emails

- Summarize this /Project Timeline Update for Self-Checkout Expansion
- Draft a reply to this email /Project Timeline Update for Self-Checkout Expansion

Transform unstructured Text into a Table

- Convert this list of Walmart store inventory items into a table:
    - Apples, $1.00 per lb, 500 lbs in stock
    - Milk, $2.50 per gallon, 200 gallons in stock
    - Bread, $2.00 per loaf, 150 loaves in stock
- (advanced) Transform this data about Walmart customers into a table:

Michael from New York loves shopping for electronics at Walmart. He visits the store twice a week and prefers weekend shopping. He spends an average of $150 per visit and usually shops between 3 PM and 5 PM. He's a Walmart Rewards member. Sarah in Los Angeles frequently buys groceries. She shops three times a week, usually on weekdays. Her average spend is $200 per visit, and she shops between 10 AM and 12 PM. She's not a Walmart Rewards member. David in Chicago often purchases clothing items. He visits once a week, usually on weekends. He spends about $100 per visit and shops between 1 PM and 3 PM. He is a Walmart Rewards member. Emily from Houston prefers household essentials. She shops twice a week, mainly on weekdays. Her average spend is $80 per visit, and she shops between 9 AM and 11 AM. She's not a Walmart Rewards member. James in Phoenix buys baby products the most. He shops once every two weeks, usually on weekends. His average spend is $50 per visit, and he shops between 4 PM and 6 PM. He is a Walmart Rewards member.

Chat with Copilot in MS Word

Open the attached "AI Consultant Job Description - Walmart.docx"
- What's the salary range for this position?
- I would like to tailor this document to make the position posted in Toronto, Canada. Note that we are not allowed to share salary information.
- There is no 401K in Canada, replace the document with the proper equivalent in Canada.

Change the tone of Voice
Open the attached "AI Consultant Job Description - Walmart.docx"

- Make this section more fun and engaging and tailored to Gen alpha! **Highlight the position overview section**
- Make it more serious and scientific tailored to PhD and research scientists. Include buzz scientific words in AI that gets nerds excited!

Generate Artifacts from a Word Document

Open the attached "Walmart NYC Store Inventory Summary - #3051"

- I want to generate a call to action to my team based on this document. List 3 top action items.
- Write a concise email to my immediate team based on this document describing the situation, action items and timeline.
- Create a meeting agenda for our next standup meeting
- Generate a table that summarizes the inventory levels so I can include them in an Excel sheet.

- Summarize this document into 5 bullet points. Make them very concise. I want to present them in a powerpoint slide to my leadership.

MODULE 15: COPILOT WITH TEAMS

Summarize conversations in Teams Chat

- Summarize this conversation
- List top 3 action items to the team. Be Concise.
- Draft an email to Jennnifer summarizing the action items

CoPilot During Meetings
- Recap the Meeting so far!
- List all the action items needed from me.
- I want to sound more assertive, what should I say now?
- I want to encourage Max for his great work, tell me what to say.
- I want to find a loophole in Max's argument
- Create a table with the ideas discussed and their pros and cons.
- List the ideas, including people's feedback, from transcript and chat.

CoPilot After Meetings:

- ****Show how Teams summarize the meeting and display AI Notes****
- Summarize the meeting in one paragraph
- Generate meeting notes.
- List all action items required from me and all my team
- What should I have done to encourage my team?
- Draft an email to Max to summarize the conversation. Make it concise.
- What was the mood of the meeting?
- Summarize what people said, in a less technical way.
- For each participant, what was the biggest concern? What were they most excited about?
- What ideas did Max have that I can share with the Sales team?

- ****Chat Conversation:****
  - **Ryan Ahmed (Manager):**
    Good morning, Max. I've been reviewing the latest inventory report for the NYC Store #3051. There are some critical points we need to address this week. Can we go over the plan together?
  - **Max Chen (Sales Team Member):**
    Good morning, Ryan. Sure thing! I read through the report as well. Definitely some areas we need to prioritize. Where should we start?
  - **Ryan Ahmed:**
    Let's tackle the critical items first. Dish Soap and AA

Batteries are dangerously low in stock. We're risking stockouts. I want to make sure these are restocked immediately. Do we have any pending orders with CleanSolutions or BrightLight?

- **Max Chen:**
  For Dish Soap, we've reached out to CleanSolutions, but their delivery schedule has been erratic. I'll follow up today to expedite the order. As for AA Batteries, the supplier is ready to deliver, but the order hasn't been finalized. I can prioritize that.
- **Ryan Ahmed:**
  Perfect. For Dish Soap, let's explore a backup supplier if CleanSolutions can't deliver within two days. For AA Batteries, increase the order to 30 units and finalize it today.
- **Max Chen:**
  Got it. I'll make those calls and ensure the orders are confirmed. What about Organic Apples? The report highlights overstocking due to the promotion.
- **Ryan Ahmed:**
  Yes, the excess stock is concerning. Let's reduce the reorder level for Organic Apples to 40 units immediately. Also, we need to push for more sales to prevent spoilage. How about we set up an end-of-aisle display or bundle it with complementary items?
- **Max Chen:**
  Good idea. I'll coordinate with the merchandising team to create a display near the bakery section.

Maybe pair them with a discount on fruit pies or caramel dip?

- ○ **Ryan Ahmed:**
  Exactly what I was thinking. Let's move on to supplier performance. We're seeing consistent issues with CleanSolutions and BrightLight. I'd like you to set up meetings with their account managers to address these delays.

- ○ **Max Chen:**
  Will do. I'll aim to get meetings scheduled by mid-week. Should I explore alternative suppliers for these categories in parallel?

- ○ **Ryan Ahmed:**
  Yes, start researching alternatives. If CleanSolutions and BrightLight can't improve, we need reliable options. Let's also refine our demand planning. Historical data is key here—can you work with the analytics team to revisit demand forecasts for high-turnover items like Dish Soap and AA Batteries?

- ○ **Max Chen:**
  Absolutely. I'll sync with them and aim to have preliminary data ready by Friday. Anything else for this week's plan?

- ○ **Ryan Ahmed:**
  One last thing—implement safety stock thresholds for critical items. We can't afford to hit these critical levels again. Ensure the system is updated with these new thresholds.

- **Max Chen:**
  Understood. I'll coordinate with the inventory team to update the thresholds and set alerts for real-time tracking.
- **Ryan Ahmed:**
  Great. Let's reconvene Friday to review progress. Thanks for jumping on this, Max. Your quick action will make a big difference this week.
- **Max Chen:**
  No problem, Ryan. I'll keep you updated as I make progress. Talk to you Friday!

MODULE 16: COPILOT WITH OUTLOOK AND EMAILS

Write Everyday Emails with CoPilot
- Write an email to my coworker Sarah about the Self-Checkout Expansion project at Walmart.
  Show the ability to add new topics to the email

- I forgot to add that the deadline for this task is scheduled in 7 days and it's critically important to meet the scheduled deadline.

Write creative emails with CoPilot
- Write an email to my coworker Laila congratulating her on her new born baby Rose.

**Show the ability to convert it into a poem!
**Show the ability to change the length and tone by making it more formal/shorter!

Reply to Emails using CoPilot
- Change supplier
- Encourage Max and change supplier
- Keep supplier
- Discuss in Meeting
- Raise concern on declining sales, ask for action plan
- ****Show ability to change tone and length****

Draft Email Summaries with CoPilot

- Use summary by CoPilot feature to summarize Max's email

Email Coaching by CoPilot

Hi Max,

Thanks for the email and all the details, it's appreciated. Honestly, I skimmed through the numbers but they sound decent, so good job on that I guess. Though the 1% increase for clothing is kinda meh, maybe we need better designs or something?

About the supplier stuff, yeah, it's a pain, I get it. If they're messing up this much with delays and bad quality, maybe it's time to give them a serious talk. Like, why are they even slacking off this much? Feels like they're just taking us for granted now.

Switching suppliers is a hassle, but maybe try to look around and see who's out there? Even if they charge a bit more, we might save on fewer complaints and returns, you know? I don't think the delay responses from Supplier ABC are helping us either—feels like you're just wasting time chasing them down.

Anyway, let's chat about this next meeting or whenever. Just let me know when you've got some options lined up, so we're not stuck last minute.

Catch you later,
Ryan

MODULE 17: COPILOT WITH EXCEL

Perform Exploratory Data Analysis using CoPilot and Excel
Using Human_resources.xlsx file
- Generate general insights
- Calculate the average value of the Age column
- Calculate the median daily rate.
- Show count plot for employees who left and stayed.
- Which employee has the longest years with the company?
- What factors affect employee attrition? (Python Code!)

Data Sorting with CoPilot
- Sort the data in an ascending order by age
- Sort the data in a descending order by Monthly income

- Sort the data in an ascending order based on Age and then sort them based on Monthly income. Both in an ascending order.

Data Filtering with CoPilot
- Filter the data to only include Research and Development Department
- Filter the data to only include Research and Development Department and loyal employees who have been with the company for more than 30 years

Color Coding Cells based on Conditions
- Color code yes as Green and no as red in the attrition column.
- color code the age column based on the value. Set the oldest people to black and the youngest people to white.

Handling Missing Values
- How many missing values do I have in this dataset
- How do I handle missing values in this data?
- Remove all rows containing missing values

Perform Data Visualization with CoPilot
- Plot a histogram of the Age column
- Plot the age vs. monthly income
- Generate a pie chart that shows the number of employees who travels.

Generate Synthetic Data Using CoPilot

Data Scope:
Generate synthetic data for Walmart stores.
Include data from 10 stores located in top US cities.
Ensure the weather data is consistent with the location and time.

Sanity Checks:
Verify that the generated data makes logical sense.
Ensure that the highest revenue is generated from the San Francisco store.

Fields to Include:
Store ID: Unique identifier for each store.
City: Location of the store.
Week: Week of observation.
Weekly Revenue: Revenue generated in USD.
Weekly Customer Visits: Number of customers visiting the store.
Stock Levels: Inventory levels.
Turnover Rate: Inventory turnover rate.
Overstock Incidents: Number of times the store was overstocked.
Stockout Incidents: Number of times the store ran out of stock.
Weather: Weather conditions during the week.
Temperature (°C): Weekly average temperature in Celsius.

MODULE 18: AI COPILOT AGENTS

Run Existing CoPilot Agents

** Show the Visual Creator Agent**
- A minimal bar graph with colorful rising bars and an up arrow to show business growth
- Draw an image for my new course on Artificial intelligence.

Show Avatars

Activate the Ideas Coach Agent
- I want to generate ideas for my new course on Generative AI, ChatGPT, CoPilot, and Agents on Udemy. I want to maximize students engagement.
- How can I gamify the content?
- Can I have a draw for students who finish all course content till the end and complete all practice opportunities?

Activate the Career Coach Agent
- Analyze my current skills and identify any gaps that I need to fill to advance in my career. I am a financial analyst with 5 years of experience in Toronto, Canada. I have an MBA in Finance and I am pursuing my CFA level 1.
- What courses, certifications, or workshops would you recommend for someone in my role to grow professionally? I like Dr. Ryan Ahmed's courses!

- List the top 5 companies I should target in Toronto if I plan to make a career shift after my CFA certification.

Build New CoPilot Agents

You are Stemplicity's AI Onboarding Assistant, designed to streamline the onboarding process for new employees. Your role is to provide clear, accurate, and friendly guidance on company policies, IT setup, benefits, role expectations, compliance, and professional development. Use documents attached in the knowledge base and don't hallucinate.

Capabilities:

- Answer questions based on the Employee Handbook, IT security, legal compliance, and company policies.
- Provide step-by-step instructions for setting up email, VPN, and accessing company resources.
- Explain the Code of Conduct, Diversity & Inclusion, Paid Time Off, and workplace ethics.
- Guide employees through career growth opportunities, mentorship programs, and learning resources.
- Offer reminders for key onboarding tasks and training sessions.

Tone & Guidelines:

- Be professional, concise, and supportive.
- Break down complex information into simple, actionable steps.

- Keep responses accurate and up-to-date based on Stemplicity's latest policies.

Always ensure a smooth and engaging onboarding experience for new employees!

Make it professional and use friendly tone

How many vacation days?