

Hands-on Lab and Demonstration using ChatGPT for Data Analytics

Lab Outline: Using AI-powered Tools To Optimize Data Analytics

Objective:

Participants will develop skills in data analysis and visualization using AI-powered tools (ChatGPT) and Excel. By the end of the lab, participants should be able to:

- Use ChatGPT to analyze data for insights, patterns, and trends.
- Perform data transformations such as filtering, sorting, summarizing, and pivoting
- Create visualizations including bar charts, scatter plots, line graphs, and heat maps

Materials Needed:

- The dataset (superstore training dataset.csv)
- Access to ChatGPT and Excel

Duration:

60 minutes

Step-by-step Instructions:

Step 1: Initial Data Exploration with ChatGPT (15 minutes)

1. Load the Data into ChatGPT:

- Upload the provided dataset (AIE4_Supplement_Dataset.csv¹) to ChatGPT
- Ask ChatGPT to provide a summary of the dataset. Example prompt: "Summarize the key features of this dataset, including the number of records, key columns, and any notable patterns."

2. Basic Statistics:

- Request ChatGPT to calculate basic statistics for key metrics like sales, profit, and discount rates. Example prompt: "Calculate the total sales, average sales, total profit, and standard deviation for the data."

¹Source: <https://www.kaggle.com/datasets/fatihilhan/global-superstore-dataset>

3. Identify Insights:

- Prompt ChatGPT to identify any notable trends or patterns
- Example prompts:
 - **Sales Performance Analysis:** "Can you identify the top 5 most profitable products in the dataset, and explain why they might be the most profitable?"
 - **Customer Segmentation:** "Which customer segment contributes the most to the overall sales, and what are the key characteristics of this segment?"
 - **Regional Sales Analysis:** "Which region has the highest sales, and how does the profit margin compare across different regions?"
 - **Discount Impact:** "How does the discount offered impact the overall profit? Can you identify any patterns or trends?"
 - **Shipping Mode Effectiveness:** "Which shipping mode is the most cost-effective in terms of profit margin? Please provide a breakdown by shipping mode."
 - **Product Category Performance:** "Which product categories are the most and least profitable? What could be the reasons behind their performance?"
 - **High Discount Analysis:** "Which products or categories have the highest discounts applied, and how do these discounts affect their profitability?"
 - **Profitability by City:** "Can you identify the top 5 cities with the highest profitability? What factors might contribute to these cities being more profitable?"

Step 2: Data Transformation with ChatGPT (15 minutes)

1. Filtering and Sorting:

- Use ChatGPT to filter the dataset for specific criteria. Example prompt: "Filter the data to show only transactions with a profit margin greater than 20%."
- Ask ChatGPT to explain what changes were made to the dataset. Example prompt: "Describe what changed."
- Ask ChatGPT to sort the filtered data by sales in descending order

2. Summarizing Data:

- Instruct ChatGPT to group the data by a specific category, such as product category or region, and provide summary statistics. Example prompt: "Group the data by product category and calculate the total sales, average profit, and number of transactions for each category."

3. Pivoting Data:

- Use ChatGPT to create a pivot table–like summary. Example prompt: "Create a pivot summary that shows total sales and average discount by region and year."

Step 3: Data Visualization with ChatGPT and Excel (20 minutes)

1. Bar Charts and Scatter Plots in ChatGPT:

- Ask ChatGPT to generate a bar chart showing sales by product category. Example prompt: "Create a bar chart showing total sales for each product category."
- Generate a scatter plot to visualize the relationship between sales and profit. Example prompt: "Create a scatter plot showing sales on the x-axis and profit on the y-axis."

2. Heat Maps in ChatGPT:

- Request ChatGPT to create a heat map showing profit by region and product category. Example prompt: "Generate a heat map that visualizes sales by region and product category."

Step 4: Reflection and Chat-based Discussion (10 minutes)

1. Discussion Prompt:

- Reflect on the ease of use and accuracy of ChatGPT versus traditional Excel methods
- Discussion prompt: "How did using ChatGPT for data analysis compare to your typical Excel-based approach in terms of efficiency and accuracy? Were there any particular tasks where ChatGPT significantly outperformed or underperformed?"

