**<Data Science Toolbox: Python Programming>**

**PROJECT REPORT**

(Project Semester January-April 2025)

***(Candy Distributor Sales)***

Submitted by

(Hanumanthu Yugandhar)

Registration No 12318361

Programme and Section – B. Tech CSE, K23VM

Course Code INT-375

Under the Guidance of

**Name of faculty:** Dr. Manpreet Singh Sehgal

**UID:** 32354

**Discipline of CSE/IT**

**Lovely School of Computer Science**

**Lovely Professional University, Phagwara**

**CERTIFICATE**

This is to certify that Hanumanthu Yugandhar bearing Registration no. 12318361 has completed INT 375 project titled, **“Candy Distributor Sales”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor**

**Designation of the Supervisor**

**School of …………………………………………….**

Lovely Professional University

Phagwara, Punjab.

Date:

**DECLARATION**

I, Hanumanthu Yugandhar, student of Computer Science under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: Signature

Registration No. 12318361 Hanumanthu Yugandhar

1. **Problem Statement:**
2. Calculate key statistics (mean, median, mode, total, count, and range) for candy sales data provided in the CSV file using Pandas. Display the results.
3. Extend the analysis of candy sales data from the CSV file by calculating and displaying additional key statistics: maximum and minimum sales, standard deviation, and variance of the sales data using Pandas.
4. Analyse candy sales trends over time using the "Candy\_Sales.csv" file. Clean the data, convert 'Order Date' to a date format, and plot a line chart to show daily sales patterns during the period.
5. Plot a scatter graph to show the relationship between candy prices and sales from the "Candy\_Sales.csv" file. This helps uncover any patterns or correlations between price and sales, enabling better decision-making.
6. Compare the total sales generated by randomly assigned marketing campaigns from the "Candy\_Sales.csv" file using a bar chart to identify the best-performing campaign.
7. Determine the monthly candy sales trends using the "Candy\_Sales.csv" file. Process the data to extract and group sales by month, then visualize the results with a bar chart to identify inventory needs for each month.
8. Identify the most popular candy products by analysing total sales for each product and visualizing the results with a bar chart.
9. Analyse sales across different flavour categories in the "Candy\_Sales.csv" file. Extract the flavour from product names and use a boxplot to visualize how sales vary by flavour.
10. Analyse repeat customer purchases and display their frequency using a histogram.
11. Evaluate regional candy sales performance by grouping sales data from the "Candy\_Sales.csv" file by region and visualizing the results using a bar chart.
12. Datasetlink:<https://app.mavenanalytics.io/datasets?search=US+Candy+Distributor>
13. **Implementation: -**

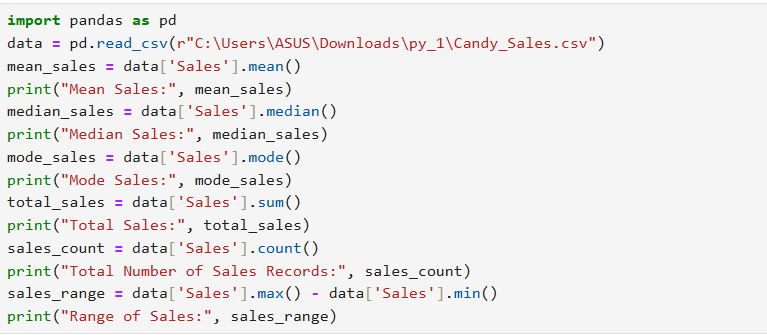
**Objective 1:** Calculate key statistics (mean, median, mode, total, count, and range) for candy sales data provided in the CSV file using Pandas. Display the results.

**Purpose:** To summarize candy sales data and gain insights through key statistical metrics.

**Approach**: Use Pandas to read the CSV file and calculate mean, median, mode, total, count, and range for sales. Display the calculated results.

**Outcome:** A clear overview of candy sales performance through essential statistics, aiding further analysis or decision-making.

**Code:**

****

**Output:**

**A screenshot of a computer

AI-generated content may be incorrect.**

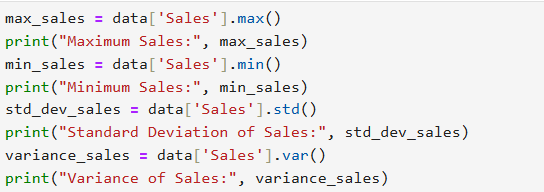
**Objective 2:** Extend the analysis of candy sales data from the CSV file by calculating and displaying additional key statistics: maximum and minimum sales, standard deviation, and variance of the sales data using Pandas.

**Purpose:** To deepen insights into candy sales data by exploring additional statistical measures.

**Approach:** Use Pandas to calculate maximum, minimum, standard deviation, and variance from the CSV file. Display the results for a comprehensive understanding.

**Outcome:** A detailed view of sales variability and extremes, aiding in identifying patterns or anomalies.

**Code:**

****

**Output:**

A black text on a white background

AI-generated content may be incorrect.

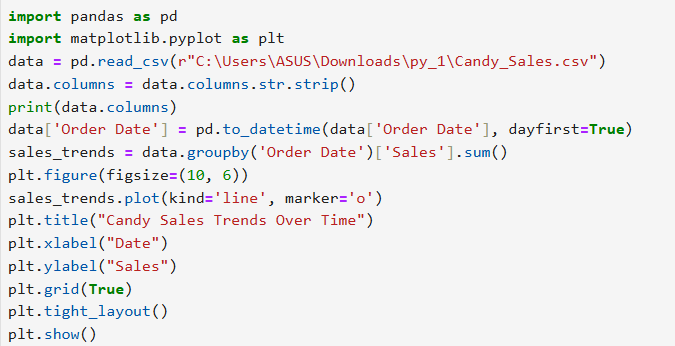
**Objective 3:** Analyse candy sales trends over time using the "Candy\_Sales.csv" file. Clean the data, convert 'Order Date' to a date format, and plot a line chart to show daily sales patterns during the period.

**Purpose:** To uncover daily candy sales trends and patterns over a specified time frame.

**Approach:** Clean and process the data from the CSV file, convert the 'Order Date' column to a date format, and group sales data by day. Use the processed data to create a line chart visualizing daily sales fluctuations.

**Outcome:** A clear visualization of candy sales trends over time, enabling better understanding and strategic planning.

**Code:**



**Output:**

A graph of blue lines

AI-generated content may be incorrect.

**Objective 4:** Plot a scatter graph to show the relationship between candy prices and sales from the "Candy\_Sales.csv" file. This helps uncover any patterns or correlations between price and sales, enabling better decision-making.

**Purpose:** To explore how candy prices influence sales and identify potential correlations or trends.

**Approach:** Use the data from the "Candy\_Sales.csv" file to create a scatter plot comparing candy prices to their respective sales.

**Outcome:** A visual representation of the price-sales relationship, helping in making informed pricing and sales strategies.

**Code:**

**A computer screen shot of a code

AI-generated content may be incorrect.**

**Output:**

A graph with green and black dots

AI-generated content may be incorrect.

**Objective 5:** Compare the total sales generated by randomly assigned marketing campaigns from the "Candy\_Sales.csv" file using a bar chart to identify the best-performing campaign.

**Purpose:** To evaluate the effectiveness of marketing campaigns based on candy sales performance.

**Approach:** Randomly assign campaigns to the sales data from the "Candy\_Sales.csv" file, calculate total sales for each campaign, and display the results using a bar chart.

**Outcome:** A clear comparison of campaign performance, highlighting the most successful marketing strategy.

**Code:**

A computer screen shot of a program code

AI-generated content may be incorrect.

**Output:**

**A graph of marketing campaigns

AI-generated content may be incorrect.**

**Objective 6:** Determine the monthly candy sales trends using the "Candy\_Sales.csv" file. Process the data to extract and group sales by month, then visualize the results with a bar chart to identify inventory needs for each month.

**Purpose:** To analyse monthly candy sales patterns and identify inventory needs based on sales trends.

**Approach:** Process the "Candy\_Sales.csv" file by grouping sales data by month and summing up monthly totals. Visualize these trends using a bar chart.

**Outcome:** A clear understanding of monthly sales fluctuations, aiding effective inventory planning.

**Code:**

**A screen shot of a computer code

AI-generated content may be incorrect.**

**Output:**

**A graph of blue bars

AI-generated content may be incorrect.**

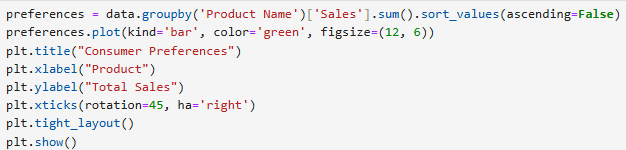
**Objective 7:** Identify the most popular candy products by analysing total sales for each product and visualizing the results with a bar chart.

**Purpose:** To determine which candy products are the most popular based on total sales.

**Approach:** Analyse the "Candy\_Sales.csv" file to calculate total sales for each product. Visualize the results in a bar chart for a clear comparison of product popularity.

**Outcome:** A ranked visualization of the most popular candy products, helping to focus on top-performing items.

**Code:**

****

**Output:**

**A graph with green bars

AI-generated content may be incorrect.**

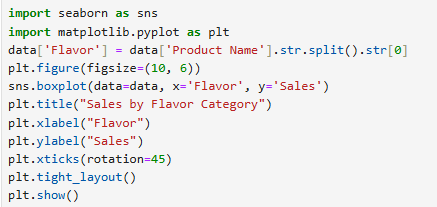
**Objective 8:** Analyse sales across different flavour categories in the "Candy\_Sales.csv" file. Extract the flavour from product names and use a boxplot to visualize how sales vary by flavour.

**Purpose:** To study how sales differ across various candy flavour categories.

**Approach:** Extract flavour names from product titles in the "Candy\_Sales.csv" file and visualize sales variations for each flavour using a boxplot.

**Outcome:** A clear comparison of sales performance across flavour categories, aiding in identifying the most successful Flavors.

**Code:**

****

**Output:**

A graph with blue and black squares

AI-generated content may be incorrect.

**Objective 9:** Analyse repeat customer purchases and display their frequency using a histogram.

**Purpose:** To understand repeat customer behaviour by analysing the frequency of their purchases.

**Approach:** Group sales data from the "Candy\_Sales.csv" file by customer ID to count their purchases. Use a histogram to visualize how frequently customers make repeat purchases.

**Outcome:** A clear visualization of customer purchase patterns, aiding in assessing loyalty and targeting frequent buyers.

**Code:**

A screenshot of a computer code

AI-generated content may be incorrect.

**Output:**

A graph with purple bars

AI-generated content may be incorrect.

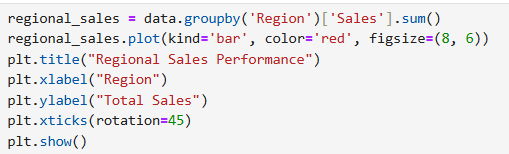
**Objective 10:** Evaluate regional candy sales performance by grouping sales data from the "Candy\_Sales.csv" file by region and visualizing the results using a bar chart.

**Purpose:** To assess candy sales performance in different regions for better geographical insights.

**Approach:** Group sales data from the "Candy\_Sales.csv" file by region and calculate the total sales for each. Use a bar chart to display the performance of each region.

**Outcome:** A clear comparison of regional sales, helping identify high-performing and underperforming areas.

**Code:**



**Output:**

A graph with red rectangles

AI-generated content may be incorrect.