Day 2 : 25.03.2024

1. Suppose you have a dataset containing historical weather data for a city for recent 5 years. Your task is to use pandas to read and manipulate the data for analysis for following

- Load and preprocess the weather dataset.
- Visualize the average temperature trends over different seasons.
- Identify and visualize any patterns or anomalies in precipitation.
- Plot temperature distributions and highlight extreme weather events.

2. Explore correlations between temperature, humidity, and other weather parameters.You are working on a project that involves analyzing a dataset containing information about houses in a neighborhood. The dataset is stored in a CSV file, and you have imported it into a NumPy array named `house_data`. Each row of the array represents a house, and the columns contain various features such as the number of bedrooms, square footage, and sale price. Using NumPy arrays and operations, how would you find the average sale price of houses with more than four bedrooms in the neighborhood?

3. You are a data analyst working for a car manufacturing company. As part of your analysis, you have a dataset containing information about the fuel efficiency of different car models. The dataset is stored in a NumPy array named `fuel_efficiency`, where each element represents the fuel efficiency (in miles per gallon) of a specific car model. Your task is to calculate the average fuel efficiency and determine the percentage improvement in fuel efficiency between two car models. How would you use NumPy arrays and arithmetic operations to calculate the average fuel efficiency and determine the percentage improvement in fuel efficiency between two car models?

4. You are a cashier at a grocery store and need to calculate the total cost of a customer's purchase, including applicable discounts and taxes. You have the item prices and quantities in separate lists, and the discount and tax rates are given as percentages. Your task is to calculate the total cost for the customer. Use arithmetic operations to calculate the total cost of a customer's purchase, including discounts and taxes, given the item prices, quantities, discount rate, and tax rate?

5. You are working as a data analyst for an e-commerce company. You have been given a dataset containing information about customer orders, stored in a Pandas DataFrame named `order_data`. The DataFrame has columns for customer ID, order date, product name, and order quantity. Your task is to analyze the data and answer specific questions about the orders. Using Pandas DataFrame operations, how would you find the following information from the `order_data` DataFrame:

- The total number of orders made by each customer.
- The average order quantity for each product.
- The earliest and latest order dates in the dataset.

6. You are a data scientist working for a company that sells products online. You have been tasked with analyzing the sales data for the past month. The data is stored in a Pandas data frame. How would you find the top 5 products that have been sold the most in the past month?

7. You work for a real estate agency and have been given a dataset containing information about properties for sale. The dataset is stored in a Pandas DataFrame named `property_data`. The DataFrame has columns for property ID, location, number of bedrooms, area in square feet, and listing price. Your task is to analyze the data and answer specific questions about the properties. Using Pandas DataFrame operations, how would you find the following information from the `property_data` DataFrame:

- The average listing price of properties in each location.
- The number of properties with more than four bedrooms.
- The property with the largest area.

8. You are working on a data visualization project and need to create basic plots using Matplotlib. You have a dataset containing the monthly sales data for a company, including the month and corresponding sales values. Your task is to develop a Python program that generates line plots and bar plots to visualize the sales data.

- How would you develop a Python program to create a line plot of the monthly sales data?
- How would you develop a Python program to create a bar plot of the monthly sales data?

9. You are working on a data analysis project that involves analyzing the monthly temperature and rainfall data for a city. You have a dataset containing the monthly temperature and rainfall values for each month of a year. Your task is to develop a Python program that generates line plots and scatter plots to visualize the temperature and rainfall data.

- Develop a Python program to create a line plot of the monthly temperature data.
- Develop a Python program to create a scatter plot of the monthly rainfall data.

10. You are a data scientist working for a company that sells products online. You have been tasked with creating a simple plot to show the sales of a product over time.

- Write code to create a simple line plot in Python using Matplotlib to predict sales happened in a month?
- Write code to create a scatter plot in Python using Matplotlib to predict sales happened in a month?

Develop a Python program to create a bar plot of the monthly sales data.