Department of Information Technology – University of the Punjab
Programming for AI – MPhil/PhD (AI) F22
Lab - 07

Max Time: 2.5 hours **Date: 01-02-2023**

Topics: Pandas

Instructions:

- Please provide your own solutions and DO NOT COPY the code from your colleagues or the web.
- You can discuss your problems only with the teachers.

Task 1

Retail Store (20 marks)

Consider the given dataset representing sales data of a retail store:

1. What is the average of the sales column in the dataset?
2. How many unique values are in the product column?
3. How many rows are in the dataset?
4. What is the maximum value in the sales column?
5. How many missing values are in each column?
6. What is the standard deviation of values in the sales column?
7. How to merge two datasets based on common columns? (Assume a second dataset exists).
8. How to group and aggregate values based on values in the product column?
9. How to create a pivot table from the dataset?
10. How to handle duplicate values in the dataset?

Task 2

City Temperature (20 marks)

Consider the given dataset representing daily temperatures of a city:

1. Extract the year, month and day information from the Date column and create separate columns for them.
2. Group the temperatures by year, month and day, and compute the average temperature for each group.
3. Find the highest and lowest temperatures for each day.
4. Calculate the difference in temperature between the maximum and minimum temperature for each day.
5. For each day, find the time when the temperature was the highest and the lowest.

Task 3

Company Sales (20 marks)

Consider the given dataset representing daily temperatures of a city:

1. Plot a line graph to visualize the sales for each region over the months.
2. Plot a bar graph to visualize the total sales for each region.
3. Plot a scatter plot to visualize the relationship between sales and months for each region.