COSC 311 Lab 2 (10 points)

Due Friday, 10 March 2023 at 11:59 PM

Please finish the following tasks and submit your lab report via MyClasses. Your submission must contain your source code file (one ".py" file for all code) and a PDF document. For each task, please include the output/results of your program (may use screenshot).

Dataset: Please download the "Algerian Forest Fires Dataset Data Set" from https://archive.ics.uci.edu/ml/datasets/Algerian+Forest+Fires+Dataset++

- read the description of this dataset on this webpage
- click "Data Folder" and download file named "Algerian_forest_fires_dataset_UPDATE.csv"
- open the file in EXCEL and check the data, you will see there are two parts. One is the "Bejaia Region Dataset" and the other is the "Sidi-Bel Abbes Region Dataset"
- We split the dataset into two separate CSV files named "Bejaia_Region.csv" and "Sidi-Bel_Abbes_Region.csv". You will use these two files for the following analysis.

Task 1: Using ***pandas*** package to:

- use info() function to show the column information of these two datasets separately
- use describe() function to show the statistics (count, mean, std, min, ...) of these two datasets separately
- show the unique values of the "Ws" (wind speed) column of these two datasets separately
- count how many samples there are in each of these two datasets separately

Task 2: draw a line figure to show the temperature change with time for the "Bejaia Region Dataset"

Task 3: draw a scatterplot figure to show the relationship between the temperature and the Fire Weather Index (FWI) for the "Sidi-Bel Abbes Region Dataset"

Task 4: draw a histogram to show the average Relative Humidity (RH) for each month for the "Bejaia Region Dataset"

Task 5: draw a bar figure to show the maximum Rain amount in a day for each month for the "Bejaia Region Dataset"

Task 6: draw a histogram to show the Wind speed (Ws) distribution in 5 bins for the "Sidi-Bel Abbes Region Dataset" in June, 2012

Task 7: draw a line figure to show the correlation between temperature (Temp) and Relative Humidity (RH) for the "Sidi-Bel Abbes Region Dataset" in July, 2012

Task 8: draw a bar figure to show the distribution of Relative Humidity (RH) for the "Bejaia Region Dataset". The x-axis is the decile of Relative Humidity (20s, 30s, ..., 90s), and y-axis is the number of days

Task 9: draw a figure (any type you want) to show the average temperature for each month when there is "no fire" and there is "fire" for the "Bejaia Region Dataset"

Note: for each figure, you need to add necessary description information (such as title, x-label, y-label, etc.) and select appropriate figure properties (such as marker, line size, line type, etc.).

Policy

1. Each student MUST finish this lab independently. NO TEAM WORK and DISCUSSION are allowed. If you need any help, please feel free to contact the instructor.

2. You need to write your whole program in an editor and save your source code as a ".py" file, which will be submitted to MyClasses together with your PDF report.