Assignment-2

Data Analytics and Visualization (CS/IT312)

19 January, 2022

Instructions

1. Create"Assignment2_Branch_YourId.py" file. *YourId* will be the student ID and *Branch* will be "CS" or "IT". For example, if student ID is 202018001 and Branch is CS, then file name will be "Assignment2_CS_202018001.py".
2. A student has to write the code for following tasks.
3. Upload your "Assignment2_Branch_YourId.py" file into the Google form shared on Google class room. Fill the Google form.
4. Please be present during lab session. The deadline of the submission is 11:59 pm, 28 January, 2022.

Get help from:
Book: chapter 6,7,8: Fundamentals of Data Visualization
Data Source: https://data.gov.in/ and https://www.kaggle.com/datasets?tags=13208-Data+Visualization

Data:

- Patient attendance in various district wise allopathic facility of J& K state (JK-Allopathic-Outpatient_attendance-May-2019.csv). The performance of each facility is divided into 4 range groups, 1 to 100, 101 to 500, 501 to 1000 and ¿1000 number of patients.
- Fifa player profile (Fifa_player_football_data.csv).

Note: The italic font type of word represent the column name of the data frame

Tasks

Patient Attendance Data:

1. Compute total patient attendance for all *district* for all four range group and plot the bar diagram. Set the bar plot parameters for better visualization.
2. Compute total patient attendance for all *district* for each *Facility Type* (DH, CHC and SC) for all four range groups and plot the staked bar diagram of three. Set the bar plot parameters for better visualization.
3. Plot group bar plot for *Performance - Overall Average* of different *Facility Type* (DH, CHC and SC) of Anantnag, Jammu, Poonch, Reasi and Udhampur.
4. Present dot plot for *Performance - Maximum* of any 20 different district. *Performance - Maximum* for different *Facility Type* should be combined appropriately using a aggregation function for each district.

Patient Attendance Data:

1. Present *Age* of various football players as histogram and kernel density plots. Set appropriate parameters of the plot.
2. Present *Age* of various Football players as Kernel Density plots for each of FC Barcelona, Chelsea, Juventus and Real Madrid *Club*s. Set appropriate parameters of the plot.
3. Plot *Value* of players as Stacked Histogram *Preferred Foot* wise (right and left).
4. Check distribution of *International Reputation* using Q-Q plot.