INT 233 G2

Data Visualization

CA-2

Roll No: \_\_\_\_\_\_\_\_\_\_\_\_ Section:\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Note:

1. All questions are compulsory.
2. It is mandatory to write the analysis for every question. Write at least 4-5 valid points.
3. Make a single pdf and upload it on the UMS.
4. Marks will depend on your presentation. More the informative your visualization will, more the marks you will get.
5. Each part is of 5 marks.

Q1. Download Death from Covid-19 (Vaccinated People)/Poland from Kaggle dataset and perform the following. Add title and caption to every worksheet.

1. Compare number of comorbidities of age group 65-74 and 75-84 with remaining age groups. Create two groups. One for 65-74 and 75-85 and other for the remaining age groups.
2. In which district number, highest number of deaths occurred? Also display Top N district based on death occurred.
3. Find which age category and sex (combined) have the highest number of deaths. Also represent different age category and sex (combined) using colors.

Q2. Download Student Performance dataset from Kaggle and perform the following. Add Title and caption to every worksheet.

1. Create a hierarchy of Gender and Race/ethnicity and find the sum of math score.
2. Which two parental level of education have the highest reading score and how much is the reading score?
3. Which race/ethnicity is having the lowest sum of writing score. Use color card to show race/ethnicity with different colors.