PROG 8460
Midterm Exam
Due: March 27, 2021 on or before 11:59 PM

Total Marks: 25

Instructions for submission:

- Log file containing your notes created during the pre-processing stage,
- Document your answers to the questions OR PowerPoint presentation with your answers,
- Source code used to produce those answers.
- Please submit the packaged Tableau workbook (.twbx) instead of the tableau workbook (.twb) containing visuals/plots.

All these files must be zipped in a folder and submitted in the Midterm dropbox on eConestoga under Assignments folder. When submitting your assignment, put it in a zipped folder (.zip). **Please do not submit your individual files separately**. **The folder should be called {YourName}_{Yourstudentnumber}_Midterm, ex: RadhaChawla_11111_Midterm.zip**

Go to: Assignments -> Midterm and make your submission.

Only submissions until 11:59 PM on 27th March will be considered without penalty. Post 11:59 PM there will be a 5% penalty the first day, 10% penalty the second day, 15% the third day i.e if you submit the midterm 3 days late you will lose 30%.

You are to use R/Python or other methods for data preprocessing and Tableau for visualizations. If you are comfortable using Tableau for preprocessing stage then you can use it, provided you submit a log file documenting the preprocessing steps you have taken. Submit the source code/visuals that you used to complete each question. Any answer that does not have supporting source code/visuals will receive a mark of 0. Ensure to add references.

Download and analyze the dataset located on eConestoga under Midterm content bar.

Your initial task will be pre-processing and with visual examples the tasks performed and manipulation done on data: **[4 marks]**

Task 1-

- You have been approached by a Hedge Fund. The Executives are currently reviewing 1k potentially well doing startups and focusing on the ones they are going to invest in.
- The criterion for selecting investments for this fund is a combination of: If a record is missing a field value, attempt to correct it.
- If you are unable to correct a data record, then you are permitted to remove it.

Once your dataset has been pre-processed, answer the following questions:

1. How many data records are in your final dataset? How many did you remove? **[1 mark]**
2. Create and display a frequency distribution table for the values contained in the "Profit" field using both a standard scale and a log scale. **[2 marks]**
3. Display the growth across each industry and where it has the maximum and least growth. **[3 marks]**
4. Your major task is to assist the Board in identifying which businesses represent the best investment opportunities based on high revenue, low expenses and top growth for year 2015. Explain your answer with the required visuals, all the analysis performed to reach to the conclusion with the respective screenshots.**[9 marks**]

Task 2 -

1. What are four reasons for BI project failure. List them with one liner explanation for each. **[4 marks]**

2. What are the steps to be opted if a project is failing because of lack of planning and technical skills? Explain your reasoning. [**2 marks**]