

## ITEC2600 – Group Project

Due on April 12, 2021

About the project:

- Up to three students in one group, **cross section group is NOT allowed**.
- Project report should be electronic, and submitted on crowdmark. See instructions about how to create crowdmark account and upload group assignment attached.
- Your Matlab code should be included. Firstly put a paragraph Matlab code, then display the output. After that, add your explanation of findings based on the output.
- Avoid long output. If your output is too long, just show small part of it.
- One copy for each group, please upload one project report including all your members on crowdmark.

Your project may include the following works (not limited to):

1. Find your own data from online resource. The data need to be official published, not made by yourself. For example, you can find data from UCI Machine Learning Repository. (<https://archive.ics.uci.edu/ml/index.php>)
2. Import the data into Matlab.
3. Give a brief explanation about your data, explain the columns you are interested in. Show the first three rows of your data. If there are too many columns in your data, just list the columns you are going to work on.
4. Clean and organize your data. For example
  - remove lines with missing values;
  - take a subset of data with columns you are interested in, or a subset by rows depends on you analysis goal;
  - ...

Note that if your cleaning work is done by other software (for example: excel), that part of work won't be counted for marks.

5. Try to understand the data by statistic measures. For example, you can choose columns you are interested in, and measure the mean, variance, max, min, ...
6. Try to understand the data pattern by plotting (i.e. 2-d plot, histogram, pie-chart...), and explain your plot and you findings.
7. Add other computing columns to your data for easy analysis

8. Find the correlations between the columns you are interested, and explain your findings.
9. Any other analysis you think helpful to understand and explain the data.

The grade will be determined by the following factors:

- Writing: clear, brief and organized.
- Analysis: sufficient, reasonable, creative.
- Findings: good explanation based on your analysis results.
- The number of effective/reasonable analysis will be an important factor to determine your grade.