Concordia Institute for Information Systems Engineering

INSE 6220 — **Fall 2025**

Advanced Statistical Approaches to Quality

Final Course Project — Due Date: December 8, 2025

Purpose: The purpose of the INSE 6220 term project is for students to apply advanced statistical quality control techniques that are discussed in class throughout the semester using Python. The term project must be done individually.

Instructions: You must submit your project report in PDF format electronically through the Moodle learning management system: `https://moodle.concordia.ca`

- The written report should be between 7 and 10 pages in length, formatted in two columns per page.
- Do NOT submit a ZIP file. Only the PDF file of your project report must be submitted.
- Write your name and ID number on the first page of your project report.
- Include the GitHub link to your Jupyter Notebook on the first page of your project report.
- E-mail submissions of the project will NOT be accepted.
- The due date is `December 8,` before midnight.

Failure to follow these instructions will result in deduction of marks.

Potential Data Sets:

You should be able to identify a good potential data set based on other work within your research group, or within your laboratory, or from the literature, or previous experience. This will be an excellent way to have relevant data for use in the project. Some data sets are available at:

`https://mldata.org/repository/data/`
`http://archive.ics.uci.edu/ml/datasets.html`
`https://www.kaggle.com/datasets`
`http://www.statsci.org/datasets.html`
`https://github.com/awesomedata/awesome-public-datasets`
`https://github.com/plotly/datasets`

Final Written Report:

You will turn in a `written comprehensive report` about your work with the `GitHub link to your Jupyter Notebook` you produced for your project. The written report should be between 7 and 10 pages in length, formatted in two columns per page. You may use IEEE template for LaTeX available at:

`https://www.ieee.org/conferences/publishing/templates.html`

Sample Project Report and GitHub link to Jupyter Notebook are available on Moodle.

Evaluation:

The project report consists of two parts + GitHub link, and constitutes **30%** of your course grade as follows:

- **10%** for applying and analyzing the results of principal component analysis on your data set.
- **10%** for applying and analyzing the results of any machine learning algorithm on your data set:
  `https://github.com/pycaret/pycaret/tree/master/tutorials`
  `https://pythonprogramming.net/machine-learning-tutorial-python-introduction/`
  `https://machinelearningmastery.com/machine-learning-in-python-step-by-step/`
  `https://towardsdatascience.com/predict-employee-turnover-with-python-da4975588aa3`
  `https://github.com/susanli2016/Machine-Learning-with-Python`
- **10%** for the `GitHub link to your Jupyter Notebook`.