



Institute for Management & Innovation
UNIVERSITY OF TORONTO
MISSISSAUGA

IMix Course Overview

Title: Exploratory Data Analysis and Data Visualization

Length: 1 day

Rationale/purpose of the course:

- What does the course/module seek to achieve?
 - In this advanced training, using the Python programming language we will be analyzing, summarizing, and reporting data, allowing you to extend your skills to industry strength analytics and to develop capabilities to make better business decisions to help reduce costs, increase profits, or better manage operations.
- Why is the course important?
 - Especially when confronted with big data sets we often do not have a clear understanding how different attributes are related or even distributed, so we use exploratory data analysis (EDA), or graphical data analysis. EDA allows us to delve into the data, to discover relationships among the attributes, and to identify interesting subsets of the observations.
- Who should attend?
 - Managers, data analysts, and others who need to keep abreast of the latest methods for enhancing return on investment.

Instructor: Gerhard Trippen, Associate Professor, Operations Management

Issues/Topics to be covered:

- Investigate the variables, examine the distributions of the categorical variables, look at the histograms of the numeric variables, and explore the relationships among sets of variables.
- The overall objective for the data mining project as a whole (not just the EDA phase) is to develop a model for the prediction of the target variable. So, which predictive variables are associated with the target variable? We want to explore the data while keeping an eye on our overall goal.

Modes of instruction:

- Discussion
- Interactive Sessions including Coding

*This course is a requirement for the Certificate in Data Analytics