CMPT 459 - D100 Special Topics in Database Systems

**Instructor:** Martin Ester

**Calendar Objective/Description:**
Current topics in database and information systems depending on faculty and student interest.

**Instructor's Objectives:**
This course introduces data mining, an area that plays a key role in Big Data analytics. The goal of data mining is the efficient discovery of useful patterns in large datasets. This course emphasizes data-driven thinking and focuses on fundamental data mining algorithms and key applications.

Students taking this course are expected to have taken an algorithms course and to have an understanding of basic statistics equivalent to an entry-level course.

Programming assignments will be in R, and students are expected to be (or to make themselves) familiar with the R programming language.

**Prerequisites:**
CMPT 354.

**Topics:**
- Introduction
- Frequent Pattern Mining
- Cluster Analysis
- Outlier Analysis
- Classification
- Social Network Analysis
- Recommender systems

**Grading:**
The grading scheme will be discussed in the first week of the class. Evaluation will be based on paper and pencil assignments, programming assignments, and a final exam.

Students must attain an overall passing grade on the weighted average of exams in the course in order to obtain a clear pass (C- or better).

**Required Books:**
Data Mining: The Textbook, Charu Aggarwal, Springer, 2015, 9783319141411

**Academic Honesty Statement:**
Academic honesty plays a key role in our efforts to maintain a high standard of academic excellence and integrity. Students are advised that ALL acts of intellectual dishonesty will be handled in accordance with the SFU Academic Honesty and Student Conduct Policies (http://www.sfu.ca/policies/gazette/student.html).