UNIVERSITY OF TORONTO

Introduction to Data Science and Analytics

Course Code: MIE1624

A Report on: The Term Project

Semester: Winter of 2019

Instructor: Professor Romanko

Prepared by: Group 7

Group Members: Hossein Nasrazadani
Amir Maghami
Mahdi Mirjalili
Navid Kayhani
Richard Joseph
Amirreza Mahpour

March 2019

Overview

Introduction to Data Science:

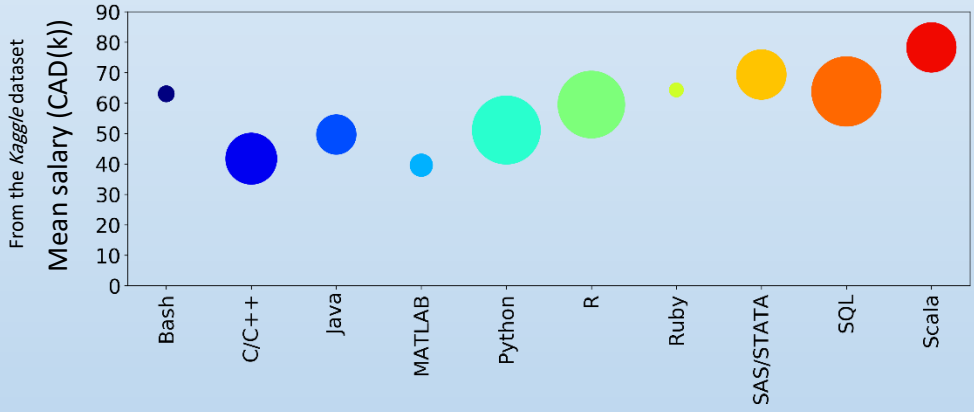
The philosophy of designing this course has been to ensure that the students learn the principles of data science and to enlighten their potential unawareness of its various applications. To maximize the chance of getting their dream data science job, the industry demand for the required skills has also been taken into account in designing the course by analyzing job postings to make sure the students can find a career path to the world of data science this course serving as an impetus. Cutting edge and highly demanded visualization software and cloud platforms have been integrated through assignments and projects to this course to familiarize the students with different analytics potentials which can be unleashed to satisfy a broad spectrum of needs.

Master of Business and Management in Analytics and AI (M.B.A.I.) :

"Once we know something, we find it hard to imagine what it was like not to know it."
Made to Stick, Chip & Dan Heath

Data is changing the way people are going about their day to day life. Smart and successful businesses harness the power of data-driven decision making to tailor their business contributions to the constantly changing needs of people's lives. This program is designed to equip the students with the necessary analytics skills to identify the market needs and translate them into valuable solutions. Almost 90% of market needs for business-analytics skills have been covered in this program.

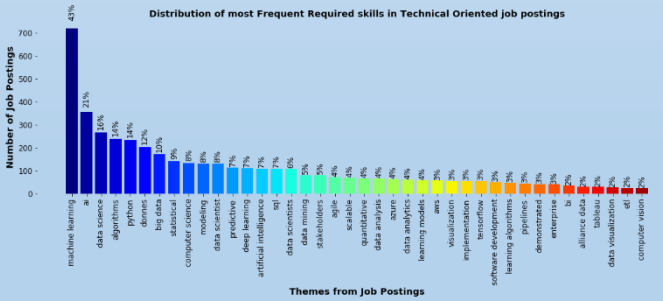
An example of what data can tell us about the market demand for different programming languages

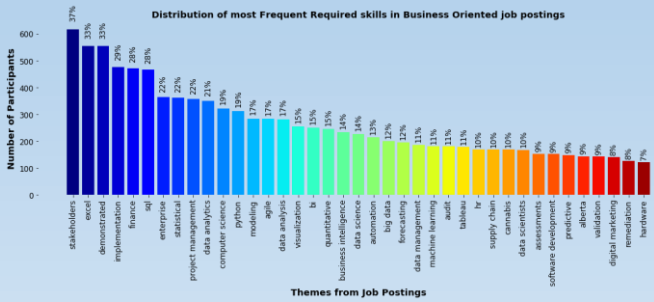


*The size of each bubble represents the frequency of the corresponding word in the job postings on the *indeed webpage*

Master of Data Science and Analytics:

"Our ability to do impactful things with data has the potential to make a difference in the world."
In a more detailed fashion, this program has been designed to offer a deeper journey in the world of data science. This program has been designed to serve students with different professional aspirations. The students will pass some core courses and then specialize in some elective areas based on their own professional purpose. Each elective area has been a highly demanded area in job market.





Course Curriculum
Intro to Data Science and Analytics (MIE 1624)

This course is designed to prepare the students for pursuing a fruitful path to the world of data science. Through this course, the students will learn the basics and develop highly demanded skills in the job market through lectures, assignments and group projects. The course aims at leveraging the most in-demand skill in the job market and provides the students with the opportunity to find their favorite job in the field of data science. According to the job postings and the Kaggle dataset, databases such as SQL and cloud platforms like AWS are important which are not covered in the current syllabus.

Prerequisites: Statistics (MIE237H1); Linear algebra (MAT188H1); Python

Fundamentals
Overview
Statistics
Linear algebra

Data Analysis & Visualization
Data cleaning
Data exploration
Visualization; Tableau

Data Mining
SQL; Query Formulation
HTML, Web-scraping
Text mining

Machine Learning
Algorithms
- Supervised (Regression, Classification)
- Unsupervised (Clustering)

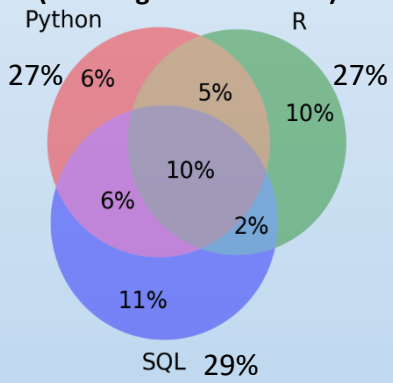
Python library
- Sci-kit learn

Modelling
Predictive modeling
Optimization
Visual analytics and storytelling

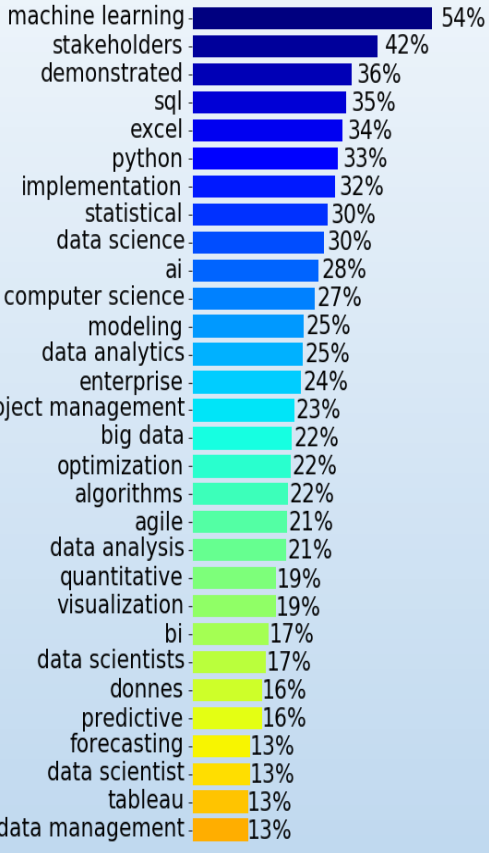
Cloud Analytics
Cloud computing
AWS
IBM

Extra Topics
Artificial Intelligence
Big data
Deep learning

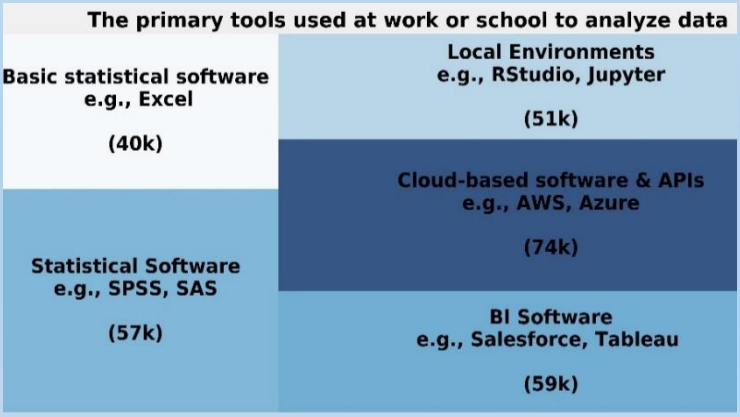
The percentage of job postings containing each coding language based on occurrence (showing unions as well)



The percentage of job postings containing each n-gram (based on word count)



Salary analysis of people using different tools



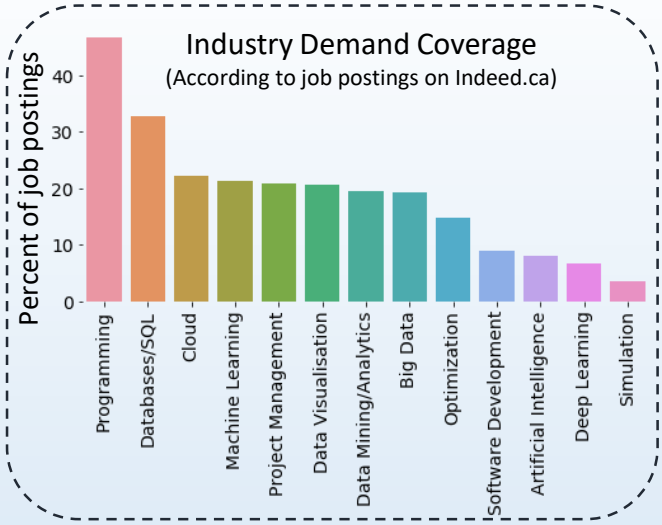
HW 1: Data analysis and visualization using a SQL database
HW 2: Machine learning algorithms implementation

Course project: Data analytics in the cloud

Master of Data Science and Analytics

The 18-month **Technical Data Science** program is a perfect fit for U of T by providing a program that covers a full range of the most in-demand areas of data science. This will prepare students for both local and international markets.

The program and the skills developed through are depicted below. It contains four core courses, a capstone project, a programing course, a workshop, and five elective courses. The students who are not proficient in programming are advised to take the programming course. The project management workshop is to provide students with soft skills to better do their capstone project. The students have to complete the core courses, the capstone project, and three units of elective courses to earn the degree. The capstone project enables students to gain experience by working on real-world problems, and makes them ready to enter the industry.



Core Units

Statistics
- Measurement
- Inferential statistics
- Causal inference
- Statistical tests
- Regression
- Sampling
- R

Data Analytics
- SQL: Query
- Web scraping
- ETL
- Data cleaning
- Data exploration

Applied Machine Learning
- Linear algebra
- Feature selection
- Learning models
- Regression
- Classification
- Clustering
- Neural network
- Python

Big Data and Cloud Computing
- Hadoop/Spark/Hive
- AWS/Azure
- ML with Big data
- Python

Capstone Project
- Enterprise problem
- Team work with advisor
- Research design
- Communication
- Reporting/presenting
- Software development

START

FINISH

Seminars
Workshops

Programming
- Python
- SQL
- Scala

Project Management Workshop
- Team building
- Scope management
- Planning/ control
- Effective cooperation

Elective Units

Software Development
- Agile tools
- Software design
- Implementation
- Testing
- Maintenance
- Scala

Data Visualization and Communication
- Visualization tools Tableau/ Python libraries
- Analytics written/oral communication

Advanced Machine Learning & Artificial Intelligence
- Reinforcement learning
- Neural network (RNN)
- Ensemple methods
- Intelligent analytics
- TensorFlow

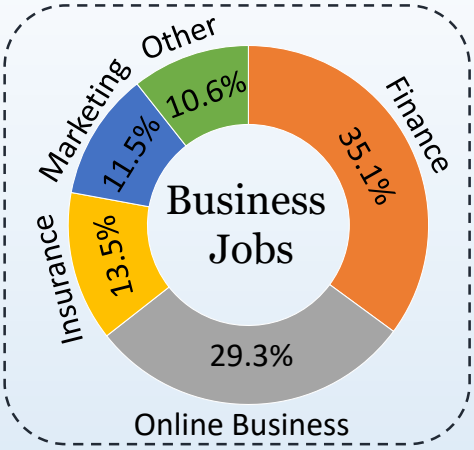
Optimization & Simulation
- Convex optimization
- Stochastic optimizaiton
- Monte-carlo methods
- Simulation models
- Simulation optimization

Deep Learning
- Natural Language Processing
- Image processing
- Neural network (CNN)
- Keras
- Pytorch

Master of Business Data Analytics and AI

The 18-month **Master of Business Data Analytics and AI** program is broadly applicable and provides graduates from quantitative backgrounds with both technical and industry skills. The four business pillars for the industry streams align with the strongest marketplace demand for business analytics experts in four areas of finance, online business, insurance, and marketing.

The program and the skills developed through are depicted below. The program is consist of four core courses, one capstone project, four elective courses, and four business-focused courses. The students are required to pass the core courses, the capstone project, two of the elective courses, and two courses from the business streams in order to complete the program degree requirements. The students have determine their business streams after completing the first semester of the program.



(From Kaggle survey 2018)

Core Units

Business Analytics Math
- Calculus
- Linear algebra
- Statistics
- Optimization
- Bayes theorem

Business Data Analytics
- Business Intelligence
- Data management
- Database: SQL
- Predictive analytics
- Modeling

Data-based Strategic Project Management
- Data ethical issues
- Agile management
- Stakeholder management
- Data-based Communication

Machine Learning and AI
- Business strategy with ML and AI
- Learning algorithms
  Supervised
  Unsupervised
- Artificial Intelligence

Capstone project
- Enterprise project
- Supervised project
- Team work
- Business planning
- Start-up development

Elective Units

Analytical Tools
- Excel
- SQL
- R
- SAS

Visualization
- Tableau
- ggplot2

Time series analytics
- Forecasting
- ARIMA Models

Deep learning
- Social media Analytics
- Neural Network (CNN)

Business Streams

Financial Analytics
- Financial markets
- ML and AI for Finance:

Examples:
1) Modeling market price
2) Volatility prediction
3) Portfolio analysis

Online Business Analytics
- Business intelligence
- Web-based analytics
- Cloud computing
- Big data platforms

1) Spark
2) Hadoop

Insurance Risk Analytics
- Insurance risks
- ML and AI for Insurance:

Examples:
1) Credit risk analytics
2) Customer classification
3) Pricing

Marketing Analytics
- Agile Marketing
- ML and AI for marketing

Examples:
1) Market classification
2) Pricing
3) Promotion design

100% HIRING RATE

LOW TUITION

REMOTE LEARNING

COSTMIZABLE SKILLS

Portfolio Booster

Beyond Skills

Portfolio Booster is a 12-week low-tuition and project-based data science bootcamp that not only provides you with the latest required technical and soft skills to get your dream job, but also **guarantees the job!**

WHAT CLASSIC PROGRAMS DO NOT PROVIDE!

"Your network is your net worth..."

The required skills to get your dream job is changing quickly nowadays. So, your training path should be dynamically designed based on **actual jobs**.

All the proposed programs still suffer from some deficiencies in maximizing the chance of getting a job for the trainees. This startup aims to bridge this gap by involving trainees in real projects from huge companies. Our experts approach companies, identify any source of inefficiency in their system, define the project, bring it to **Portfolio Booster** , and our trainees will solve the problem under the mentorship of our technical experts. As a compensation, the company will hire from the team of trainees. The number of people whom will be hired depends on the complexity of the project.

"We have the data, we don't know how to use it though..."

Applicants interested in this bootcamp need to apply online. Our examiners will assess their skills and capabilities through an evaluation process including an online basic test and virtual/in-person interviews. Successful applicants attend a free weekend prep-workshop. Based on the evaluation processes, mentors will be assigned to those who are elected as finalist . The training path is customized by the recommendation system according to the trainees interests and their mentors opinion, based on the available projects. During the training process, they will meet experts from the industry, establish connections, and learn those technical and soft skills that the industry actually is looking for. Once the mentor finds the team members for the project, they will solve the problem together. Based on our contract with our industry partner, the team will be hired, and **Portfolio Booster** will get a percentage of their salary.

UNSOLVED PROBLEMS WILL BE SOLVED FOR FREE

HIRE OUR TRAINEES

STRONG NETWORK

CUSTOMIZED SKILLS

EXPERIENCE IN REAL PROJECT

LOW TUITION

Portfolio Booster
Beyond Skills

FIND YOUR PROBLEM

SOLVE THE PROBLEM

IBM Bell

amazon