COURSE PROJECT PRESENTATION

Group 6

TABLE OF CONTENTS

Skill Extraction

01 **Words/Phrases from Coursera**
Top 3000 most frequent words & phrases

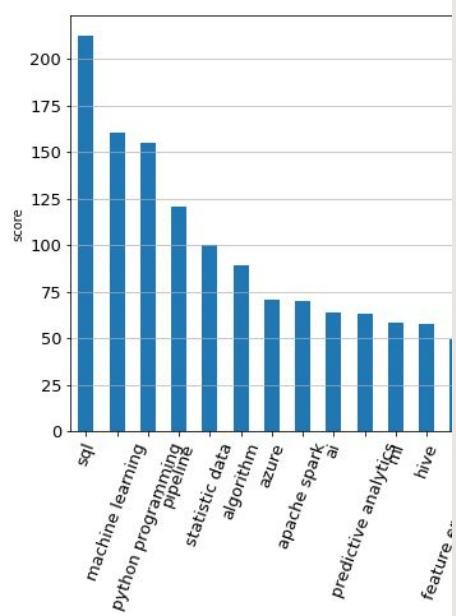
02 **Words/Phrases from Job descriptions**
TF-IDF & N-grams
Cross-filtering

03 **Projection Filtering**
Intersection between Coursera and job description

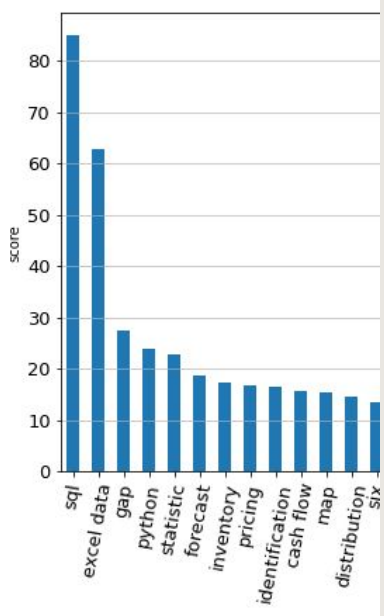
04 **Manual Filtering**
Manually filter redundant words

Skill Extraction

Technical Skills



Business Skills



Python programming

Machine learning

SQL

Course curriculum for MIE 1624

Week 1
(Statistics & Linear algebra)

- Intro to course
- Overview of Linear algebra
- Basic statistics
- Python warmup

Week 2
(Regression & Feature)

- Linear regression
- Logistic regression
- Feature extraction
- Feature visualization

Assignment #1

- Feature selection/visualization
- Statistical measures
- Basic supervised models
- Result analysis

Week 3
(Advanced ML)

- Decision trees
- Naive Bayes, k-NN, SVM
- Ensemble learning

Week 4
(NLP)

- Text Mining/Analysis
- Recommender Systems
- Social Network Analysis

Assignment #2

- Tweet text mining
- User group classification
- Sentiment analysis
- Social network analysis

Week 5
(Deep learning)

- Neural Networks
- CNN
- RNN

Term Project starts

(Open-ended)

Course curriculum for MIE 1624

Week 6(Databases & Basic SQL)

- Database Basics
- Intro to relational database
- Basic SQL queries (CREATE, SELECT, COUNT, INSERT, UPDATE)
- Basic SQL practice

Week 7 (Advanced SQL)

- Advanced SQL queries (Group data, range, sorting data, multi-table query)
- Advanced SQL practice

Week 8 (Cloud computing)

- Cloud computing overview
- Introduction to AWS
- Model hosting and scheduling
- Main service intro: AWS EC2, AWS Data pipeline, AWS Lambda

Assignment #3

- Open topic project require utilizing AWS and SQL
- Apply deep learning technique:
  CNN - image processing
  RNN - text analysis

Week 9(Big Data)

- Big Data overview
- Big data modelling
- Foundations for big data system
- Intro to Hadoop

Week 10 (Term project showcase)

- In-class project demonstration
- In-class project presentation
- Q&A session

Term Project Ends

Program Structure

Core Courses

The core courses are five fundamental courses including aspects of machine learning, data science and statistical analysis. These courses give students a solid foundation for more advanced topics.

Business Option

The business stream emphasizes on real-world management analytics and the usage of predictive tools in machine learning and artificial intelligence for the innovation and tech-driven economy.

Technical Option

The technical option course provides students with advanced statistical tools in order to properly analyze complex and large data and how to prepare and interpret visual representation of complex and large data.

Compulsory Course

Statistical Predictive Modeling For Analytics

-Basic statistics, linear algebra
-Probability theorem
-Estimation and prediction
-Classification models
-Programming language R

Visualizations And Business Communications

-Matlab, R, excel: data visualization
-Tableau, PowerBI
-Professional presentation
-Learn visual representation methods and techniques that increase the understanding of complex data and models

Big Data

- Data exploration and cleaning
- Feature engineering
- Spark
- Hadoop, hive
- Cloud Service: AWS, Google, IBM

01 02 03 04 05

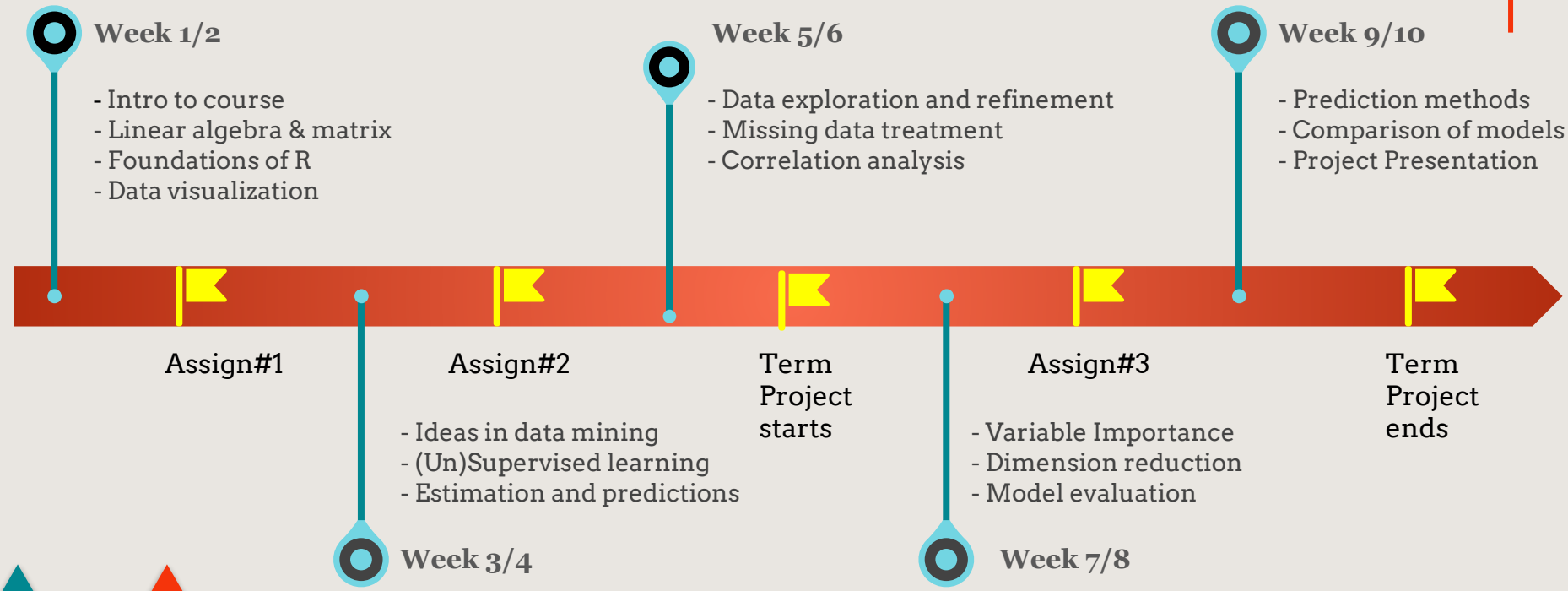
Introduction to File and Database Management

-MySQL/ NoSQL
- CRUD (create, read, update and delete)
- Python, web scraping
- Database Modeling, Design, Normalization, Security

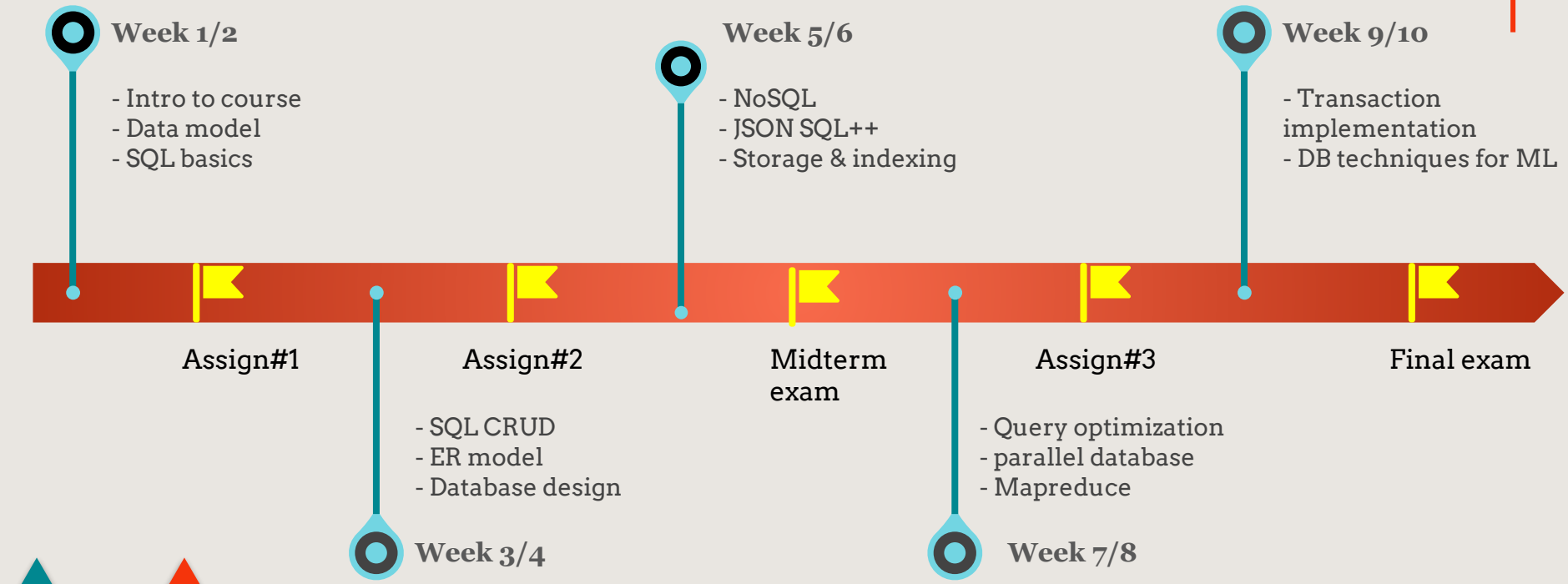
Introduction to Machine Learning

- Python
- Cloud API
- Data Cleaning and exploration
- Regression algorithms
- Classification algorithms

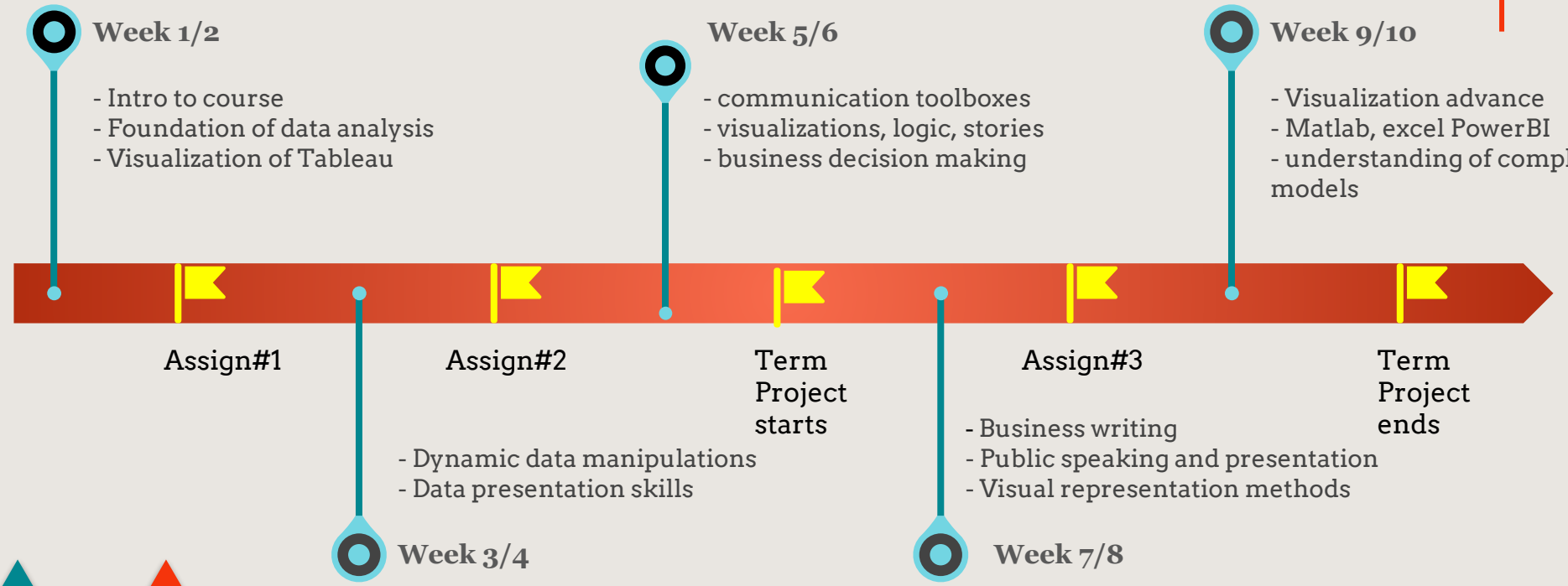
Visualization of Compulsory Course 1



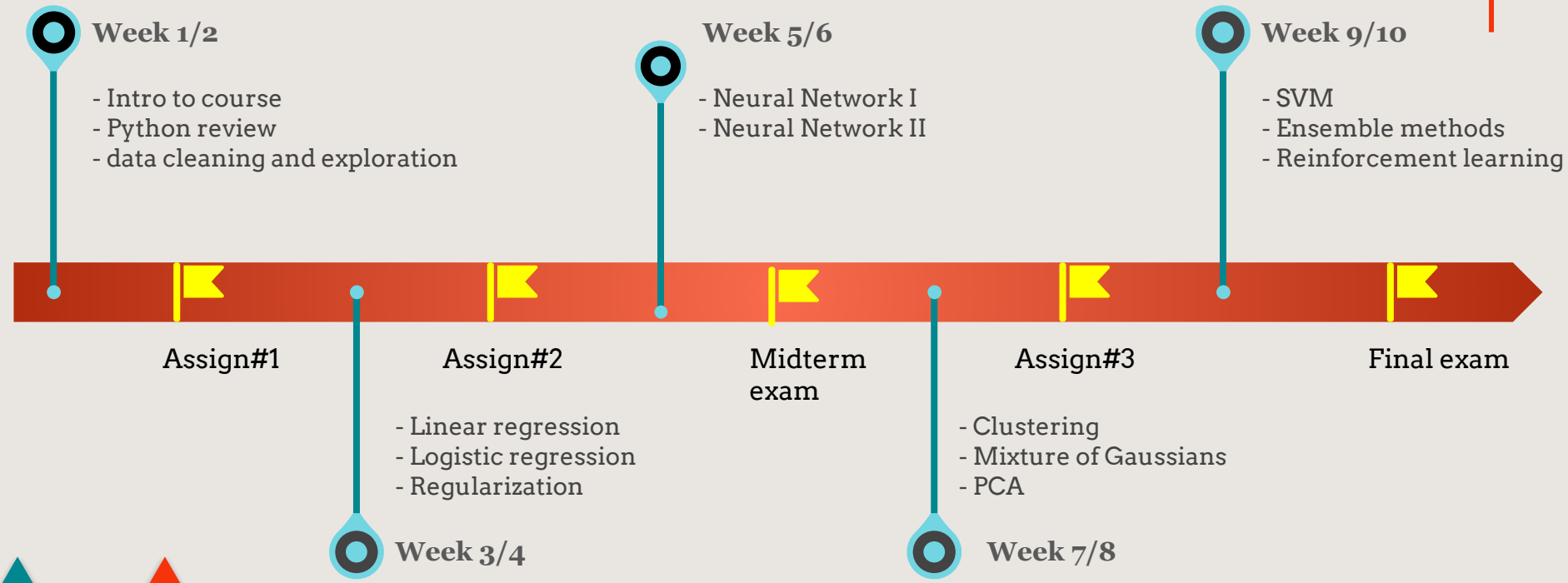
Visualization of Compulsory Course 2



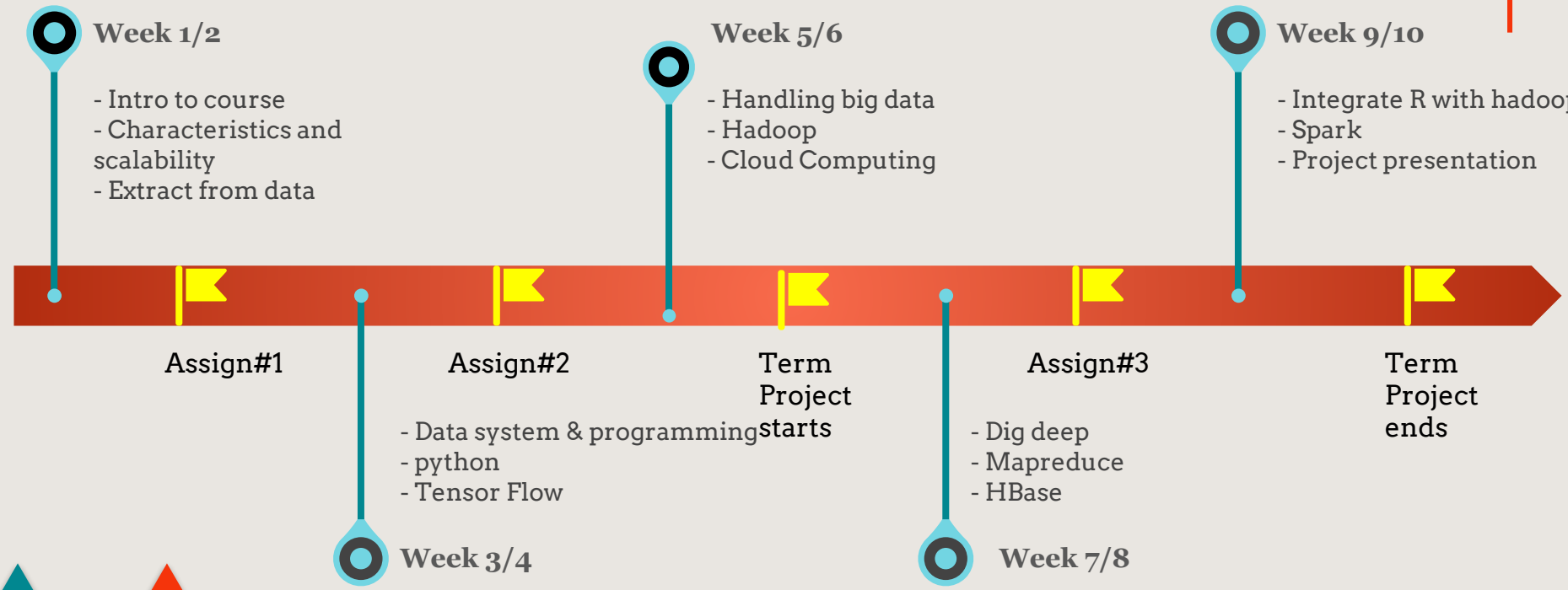
Visualization of Compulsory Course 3



Visualization of Compulsory Course 4



Visualization of Compulsory Course 5



Optional Technical Course

Neural network & Deep learning
-Deep Unsupervised Learning
-Convolutional Neural Networks
-Non-convex optimization for deep networks
-Stochastic Optimization

Data Science Capstone
-Machine learning
-Data mining
-Preparing, analyzing and visualizing data
-Building and testing models
-Communication and presentation

Data structure and Algorithm
-Data types : list, stacks, queues, trees, traversal, binary trees, etc.
-Data structures for coding and compression
-Searching, merging and sorting
-Dynamic programming, Greedy methods
- Graph algorithms

Natural language processing (NLP)
-N-gram Language Models
-Part Of Speech Tagging and Sequence Labeling
-LSTM Recurrent Neural Networks
-Syntactic parsing
-Semantic Analysis
-Information Extraction (IE)
-Machine Translation (MT)

Optional Business Course

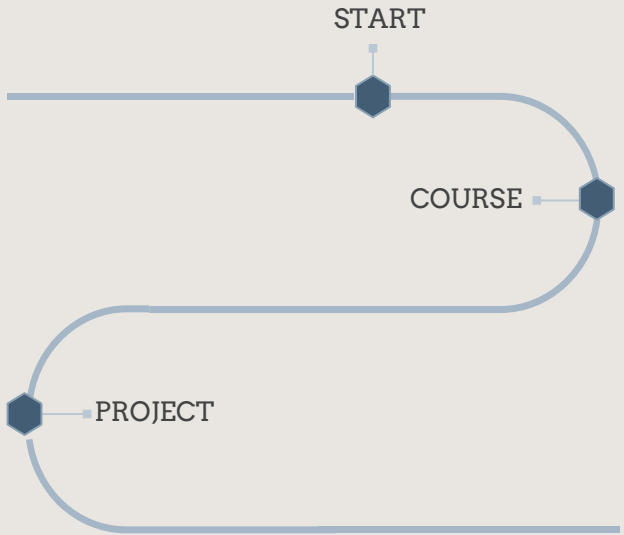
Business analysis
- Statistics
- Probability models
- Simulation models and Hypothesis testing
- key processes, exploratory and predictive analytics

Revenue Management and Pricing
-Capacity allocation
-Markdown management
-Dynamic pricing for e-commerce
-Customized pricing
-Demand forecasts under market uncertainty

Business Decision-Making Through Advanced Analytics
-Structure, analyze, and solve business decision problems
-Decision-making analysis: systematic, critical, and logical thinking)
-Basic techniques and modeling approaches (interpret results of the analysis in the context of a decision-making objective).
-Optimization
-Decision trees, and simulation.

Forecasting Models
-Managerial decision making
-Forecasting techniques
-ARIMA
-ARCH techniques
-Toolkit of techniques in Econometric Views (EVIEWS)

EDTECH STARTUP



Our EdTech program, **Data Science for Education**, aims at helping students approach internships of Canadian companies in data science.

It will connect the students with the courses in universities as well as the program in large corporations that are in huge demand of help.

While attending this program, students need to take three courses as well as one capstone project.

COURSE SELECTION SYSTEM

Questionnaire
Use input function to create a temporary questionnaire in python to get student's personal interest data.

Analyze Data
This system subtracts keywords from descriptions of targeted courses and personal data.

Find Similarity
A designed system scores every course based on their similarity.

Print Output
Recommend the top three courses based on the rank of scores.

Capstone project directions

Financial Analysis Field
RBC
CIBC
Scotia

Information Techniques
Microsoft
Pinterest
TELUS

Academic Research
Queen's University
McGill University
University of Toronto

A.I in Big Data
Yelp
Uber

THANKS!

Reference

This is where you give credit to the ones who are part of this project.

Did you like the resources on this template? Get them for free at our other websites.

- Presentation template by Slidesgo
- Icons by Flaticon
- Infographics by Freepik
- Author introduction slide photo created by Freepik
- Text & Image slide photo created by Freepik.com