CSC-3154DSIA: Industrial Applications of Data Science

Course Description

The purpose of this course is to give a cursory overview of how modern data analytics tools are used in the workplace. Topics include discussion of python modeling, SQL database management, data visualization in Tableau, linear regression, advanced regression, modern business problems, and full scale data science project development.

Inspired by the Data-X course at UC-Berkeley https://datax.berkeley.edu/berkeley-course/

All code for lectures can be found here: https://github.com/NolanSmithSolutions/Lectures

Learning Outcomes

- Build relevant programs individually and in groups using Python/Tableau or similar languages/programs that can be used at work
- Interpret results from statistical packages and methods
- Use standard packages and data sources from the data science community

Prerequisites

A general interest in diving into data. Prior experience with excel, SQL, probability and statistics will be helpful but not necessary. Rudimentary understanding of a scripting language is required.

Grading

Homework (weekly) - 35%
Midterm Exam - 15%
Final Exam - 20%
Project - 30%

Contact Information

Ben Smith
Email: bensmith.nss@gmail.com
Cell: (513) 212-1910

Matthew Nolan
Email: mattnolan.nss@gmail.com
Cell: (202) 868-2853

Schedule

Week	**Lecture 1**	**Lecture 2**
1	Syllabus, introductions, finding project members	Python Intro ● For loops ● Function design ● If / Then statements
2	Intro to Pandas & Numpy	Finish Pandas
3	SQL Basics ● Basic pulls ● Joins & Merges	Advanced SQL ● Modifying tables ● Table aggregations ● Creating a SQL database using a CSV
4	Web Scraping in Python ● Beautiful Soup ● Selenium ● Parsing HTML	Finish Web Scraping with examples from online data sources ● BEA/BLS/FRED/YahooFinance API's/Kaggle
5	TD Ameritrade API ● Pulling stock price data ● Pulling fundamental data ● Applying pandas and numpy to cleanse data	Data visualization in Tableau ● Basic dashboard design ● Charting ● Linking to external data sources
6	Midterm Exam	Programming languages, other courses and future developments in tech
7	Application 1: Inventory Management ● Discuss how inventory flows through a company ● Why it matters to manage inventory closely	Application 1: Inventory Management ● Use SQL & Python to create database from provided CSV files ● Use Python and new database to develop inventory forecast
8	Application 2: Statistical Modeling - Stock Prices ● Review dummy variables ● Review regression concepts ● Use python to read in fundamental and technical data	Application 2: Statistical Modeling - Stock Prices ● Create basic machine learning model using SciPy

9	Application 2: Statistical Modeling - Stock Prices ● Create more advanced machine learning models with sklearn	Application 2: Statistical Modeling - Stock Prices ● Interpret outcomes using Tableau
10	Application 3: Risk Management ● Review of Confidence intervals Understand Value at Risk (VaR)	Application 3: Risk Management ● Use TD API & Python to calculate VaR for 2 stock portfolio Create dashboard summarizing findings
11	Project workshop	Project presentations
12	Project presentations	Final Exam