Course Syllabus: Data Science with Python

Program Instruction Team Information

Instruction Team

- **Marcus Zou**: Course Builder and Instructor; MSc in Data Analytics of Georgia Tech., Fmr. Systems Engineer @Emerson; marcus.zou@gmail.com.
- **Lulu Wang**: TA; Computer Science graduate of the University of California Los Angles.
- **Michael Jin**: TA; Information technology guru, Data Engineer of Suncor Energy.
- **Rena Hu**: Program Admin; CCIE7416, Bsc. of Computer Engineering, Principal Consultant to Suncor Energy; techclubcanada@gmail.com.

General Information

Description

This deliberately designed course, **Data Science with Python** (Course ID: DSP2105), starts off with refreshers of Python, R, SQL, Practical Statistics, then dives into the Data Science world, coping with data problems with data analytics processes and toolkits, plus introductions to machine learning technology, ending up with 2 or 3 real-world capstone projects from the industries. During the learning process, data science techniques of EDA, data wrangling, storytelling, and machine learning algorithms of SLR, MLR, SVM, Bayes, Decision Trees, clustering, PCA, feature selection, etc. will be delivered.

Pre-Requisites

* Grade 12+ with Calculus knowledge, current post-secondary scholars preferred.
* Strong interests or experiences in coding, mathematics, and data/computer science.
* Desire to learn cut-edge technologies.
* A PC/ laptop with at least 8GB RAM and 50GB free disk space.
* A decent Internet connection.

Course Overview

This course is intended to prepare students to build data science foundations and make intelligent decisions based on the understanding of data science theory and the utilization of data analytical models, applications plus introduction to machine learning algorithms. Therefore, two broad learning objectives exist of this course:

- **Analytics-based learning objectives** - where students will be required to understand basic data analytics principles, data science processes, toolsets, and applications.
- **Machine Learning objectives** - where students will be required to learn foundations of machine learning theory, algorithms, and how to apply the ML algorithms to the real-world.

Course Requirements & Grading

Assignment	Release Date*	Due Date*	Weight
Homework #1: Explore Weather Trends	**May 18**	**May 28**	**10%**
Homework #2: Exploring US Bikeshare	**May 24**	**June 4**	**10%**
Homework #3: Investigate a Dataset	**May 31**	**June 11**	**10%**
Homework #4: Analyze A/B Test Results	**June 7**	**June 18**	**10%**
Homework #5: Explore and Summarize Data	**June 14**	**June 25**	**10%**
Homework #6: Wrangle and Analyze Data	**June 21**	**July 2**	**10%**
Homework #7: Create a Tableau Story	**June 28**	**July 9**	**10%**
Optional Homework #8 **: Medical Cost Personal or alternative	**July 12**	**July 25**	**0% (10%)**
Capstone Project: 2 or 3 Projects to select with	**July 26**	**August 7**	**30% (20%)**

* The dates above are referred to year 2021.

** **Whether or not we shall assign Homework #8** depends on how practical/experienced the students are in terms of mastering the previous parts/modules/sections. Therefore, the points allocation of Homework #8 and Capstone Project shall be changed accordingly.

Description of Graded Components

All homework should be completed individually. Each assignment will require students to complete a model in Anaconda environment (Python or R) and/or Excel (for the HW#1).

Assignment Turn-In

- Each homework must be completed in accordance with their respective description and submitted by the due date identified within the course syllabus and course outline.
- **If there are any questions or concerns, please contact Marcus/Rena immediately!**

- All Homework will be turned in on Canvas, as per the due dates.

Homework Due Dates

- All homework will be due at the times listed above.
- These times are subject to change (due dates could be changed to a later date, but they will never be moved to an earlier date).
- We will communicate any updated due dates if observed.

Grading Scale

Your final grade will be assigned as a letter grade according to the following scale:

A	90-100%
B	80-89%
C	70-79%
D	60-69%
F	0-59%

Timing Policy

- The Modules follow a logical sequence that includes knowledge-building and experience-building.
- Assignments should be completed by their due dates, in order.
- You will have access to the course content for the scheduled duration of the course.

Course Materials

Course Text

- No required textbook

Additional Materials/Resources

- Excel, preferably Excel 2016. Excel 365 shall work.
- Anaconda environment (2020.1 or above).
- Access to the Internet.

Course Website and Other Classroom Management Tools

- This class will use Canvas to deliver course materials.
- Course e-lectures will be hosted on an external website which is connected to Canvas.
- All other course contents and martials will be placed in Canvas.

Technology/Software Requirements

Internet connection (DSL, LAN, or cable connection desirable)

Course Expectations & Guidelines

Course Integrity

Honesty and transparency are important features of good learning environment.

Additionally, any student is encouraged to find hints to resolve the problem on the Internet, but a simple copy/paste without denoting the source(s) will be treated as a “Cheat”.

Attendance Policy

- This is a fully online course.
- Login on a regular basis to complete your work, so that you do not have to spend a lot of time reviewing and refreshing yourself regarding the content in the last minutes.
- Asking help on the last day is not accepted. Time manage yourself please.

Collaboration & Group Work

- Homework assignments should be done individually.
- Students might be informed by the instructor to form groups for group assignments if such circumstance occurs.

Extensions, Late Assignments, & Re-Scheduled/Missed Exams

- No midterm nor final exams be arranged, but the regular Homework assignments and capstone projects.
- All assignments are expected to be completed and submitted by the due date.
- Late submissions are not accepted unless there are **extenuating circumstances**, which should be communicated with the instructor AND program admin at least one day prior to the due date. Such late submission is **subject to at least 20% penalty** upon whatsoever he/she makes.
- Proper documentation for all extenuating circumstances must be provided and documented properly, preferable in **a new email with a new subject line** (Do not reply to a pre-existing email thread).

Communication

- Discussion boards/forums will be done via MS Teams discussion group.
- Virtual Office hours will be conducted weekly via Zoom Meeting or MS Teams
- Questions for technology & support & logistics should be directed to:
  - techclubcanada@gmail.com

Office Hours and Participation

- The regular Office hours will be held by the Instructor every Saturday morning from 10 AM to 10 PM via online meeting tools, such as, Zoom Meeting, Blue Jeans, or MS Teams.
- The supplementary Office Hours will be held by the lead TA with one hour in the morning of Thursday. The schedule may vary, please visit MS Teams group or email more often.
- The Office Hours are NOT recorded unless arranged, so please secure your time slot if you have something to discuss.
- Participation of the Office Hours is voluntary. You can ignore it if you can progress along others’ pace and manage your Homework very well.
- For the first day of the first week of class I would like for you all to participate in Office Hours so I can “meet you”!!!!

Online Student Conduct and (N)etiquette

Communicating appropriately in the online classroom can be challenging. In order to minimize this challenge, it is important to remember several points of "internet etiquette" that will smooth communication for both students and instructors.

Netiquette refers to etiquette that is used when communicating on the Internet. Review the Core Rules of Netiquette. When you are communicating via email, discussion forums or synchronously (real-time), please use correct spelling, punctuation and grammar consistent with the academic environment.

We expect all participants (learners, teaching assistants, etc.) to interact respectfully. Learners who do not adhere to this guideline may be removed from the course.

Course Schedule

Please visit Time Schedule on Canvas.