Course Information
Course Number: MIS 6V99 Sec 010 – 28019
Title: Special Topics in MIS-Programming for Data Science
Term: Spring 2017
Class Hours: Saturday 9:00 AM– 11:45 AM
Class Room: JSOM 12.210

Instructor Information
Name: Syed Ali Nabeel
Email: snabeel@utdallas.edu
Office Hours: By Appointment

Prerequisite
Some programming background, linear Algebra, elementary statistics and probability knowledge would be beneficial.

Course Description

This course intends to cover the programming aspects required for the solution architecture,   implementation and presentation of data science problems using predictive modeling and analytics. Python will be used as a programming language. Though concepts and theoretical aspects are addressed, more emphasis will be on solving a real world data science problem by performing descriptive and inferential statistics, applying advanced analytical techniques, building machine learning models and presenting results for adaptive and smart business decision making by simply connecting the dots from business gains to human behaviors and from data generation to dollars spent.

Learning Objectives

- To understand the fundamental programming concepts of data science environments including their solution design, implementation and presentation by mapping technology strategy of an enterprise to its business requirements.
- Data Extraction and filtering from a large dataset stored in a variety of different formats.
- Perform complex statistical analysis on datasets.
- Predictive modeling on datasets using advanced analytical techniques as well as using machine learning algorithms and natural language processing techniques.
- Interactively visualize large datasets and present the results as a driver for adaptive business intelligence.

Textbook
    Not Required. Course Materials, Presentations and Internet Resources will be used in the class.

Recommended Books:

1. Learn Python the Hard Way: A Very Simple Introduction to the Terrifyingly Beautiful World of Computers and Code (3rd Edition) (Zed Shaw's Hard Way Series) 3rd Edition.The book is electronically available for free at https://learnpythonthehardway.org/book/
2. Python Data Science Essentials:Alberto Boschetti, Luca Massaron April 2015,Packt Publishing,ISBN 9781785280429 , https://www.packtpub.com/big-data-and-business-intelligence/python-data-science-essentials
3. Python Data Analytics:Fabio Nelli,APress,ISBN 978-1-4842-0958-5
4. Introduction to programming using Python by Schneider
5. Introduction to programming using Python by Liang

Technical Requirements
    Each student is required to bring a laptop computer to download, install and run the following

    1. Download and install the Anaconda distribution for scientific computing from https://www.continuum.io/downloads which provides all the necessary packages for Linux, Mac, and Windows platforms.

    2. Sign up for a free kaggle account at https://www.kaggle.com/

Course Assignments
    Details will be discussed in the class.

Academic Calendar

	Date	Descriptions	Notes
1	1/14	Analytics Platform Architecture, Python Overview, Variables, Operations and Assignments, Introduction to Data types	
2	1/21	Lists, Functions and Packages, Functions, Tuples, Lists, Sets and Files Control flow statements and loops	
3	1/28	NumPy and SciPy Introduction, Dictionaries and data frames.	

4	2/4	NumPy and SciPy,Data Processing with Dictionaries	Assignment 1 Posted
5	2/11	NumPy and SciPy(Contd). Classes and Objects	
6	2/18	Data Ingestion-Cleaning, Wrangling and Munging, Exploring Pandas	
7	2/25	Data Ingestion-Cleaning, Wrangling and Munging, Exploring Pandas(Contd.)	Assignment 1 Due.
8	3/4	Exploring Relational Databases with SQL and Python	
9	3/11	Mid Term	
10	3/18	No Class	Spring break
11	3/25	Statistical Analysis and Probability with Python	Assignment 2 posted
12	4/1	Introduction to Machine Learning, Exploring Python Machine Learning using Scikit-Learn	
13	4/8	Introduction to TensorFlow Natural Language Processing and Text Analytics.	Assignment 2 due, Assignment 3 posted
14	4/15	Natural Language Processing and Text Analytics	
15	4/22	Data Visualization using MatplotLib Exploring Bokeh	
16	4/29	Data Visualization using MatplotLib Exploring Bokeh	Assignment 3 due

(The schedule may change.)

Grading Policy

3 Assignments (5% + 5% + 5%) =15%
Pop Quizzes 25%
Midterm 30%
Final 30%

Scale	Letter Grades
90-100	A
85-89	A-
80-84	B+
75-79	B
70-74	B-
65-69	C+
60-64	C
55-59	C-
Below 55	F

Course and Instructor Policy

Late Assignment: No late assignment will be accepted.

Makeup Exams: There will be no make-up exams, except for medical emergency (written statement justifying the situation from a physician required). The written statement should include the physician's address and phone number for the verification purpose.

UT Dallas Syllabus Policies and Procedures

The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus.

Please go to **http://go.utdallas.edu/syllabus-policies** for these policies.

The descriptions and timelines contained in this syllabus
are subject to change at the discretion of the Professor.