

Spring 2019

CSCI 381/780 Data Analytics

Kumar Ramansenthil

CUNY Queens College, Kumar.Ramansenthil@qc.cuny.edu

NYC Tech-in-Residence Corps

rdomanski@sbs.nyc.gov

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/qc_oers

 Part of the [Computer Sciences Commons](#)

Recommended Citation

Ramansenthil, Kumar and NYC Tech-in-Residence Corps, "CSCI 381/780 Data Analytics" (2019). *CUNY Academic Works*.
https://academicworks.cuny.edu/qc_oers/2

This Syllabus is brought to you for free and open access by the Queens College at CUNY Academic Works. It has been accepted for inclusion in Open Educational Resources by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

SYLLABUS

COURSE	CS 381/780		
INSTRUCTOR	Kumar Ramansenthil	E-MAIL	Kumar.Ramansenthil@qc.cuny.edu
Classroom	SB D133	Contact Method:	Email (expect response in 1-2 business days)

LEARNING OBJECTIVES:

At the end of this course, students will be able to:

1. Use technical skills in SQL databases to pull data from various sources and perform analysis
2. Use analytical thinking and tools such as Spreadsheets, Python and Tableau.
3. Draw insights about the data and make recommendations
4. Visually present analysis and data driven recommendations using advanced visualization

DESCRIPTION:

Every industry currently uses data driven analytical solution to make business decisions such as whether to launch a product, who to market it to etc. Data Analytics is omnipresent from the moment an idea is conceived and till the product is launched. This course is designed to prepare students for analytical careers through a combination of relevant fundamental knowledge and technical training in industry standard tools and hands on experience in data driven analytics on real-life data. Student will explore different type of data like transactional and behavioral data, data cleaning, predictive modeling, introduction to API and sentiment analysis. The course covers state of the art tools such as Tableau, Python and Hive which have become essential for careers in Data Analytics.

The prerequisite for this course are CSCI 313, Math 241. **Students are expected to complete various assignments using Python, SQL and Tableau**

TEXTBOOK:

Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking, 1st Edition, by Foster Provost, Tom Fawcett
ISBN-13: 978-1449361327
ISBN-10: 1449361323

METHOD OF ASSESSMENT:

- 2 midterms and Final exams ---60%
- Homework ---10%
- 2 Projects ---20%
- Class Participation ---10%

QUIZZES:

During the semester, students may be given quizzes from the topics covered during the previous lectures. This is done to encourage students to cultivate the habit of preparing for the classes weekly, and hence be ready to learn new concepts the following weeks. This also helps them to prepare for the examination. Generally, a quiz will be given at the beginning of a class and students who are absent or late for the class will not be allowed to take that quiz at a later time.

Quizzes will be counted toward class participation grades

RESPONSIBILITIES:

Students are expected to attend the class lectures and take the exams at the scheduled times. Assigned readings, exercises and projects must be completed on time.

ACADEMIC INTEGRITY

Cheating on exams or copying projects will not be tolerated. Please review the College's policies on Plagiarism and Cheating. Moreover, if you copy projects or assignments, you will not be able to answer questions on the exams.

PROPOSED SCHEDULE OF LECTURES:

Week	TOPICS	Required Reading	Assessment
1	Introduction to the course Introduction to Analytical Thinking Activity: Hands on dealing with data with Python and SQL	Chapter 1 and 2	Written Homework (Short answer type question) Lab (Python, SQL)
2	Introduction to Predictive Modeling Supervised Segmentation Visualizing Segmentation	Chapter 3	Written Homework (Short answer type question)
3	Fitting Model to Data Activity: Pre-Project Assignment	Chapter 4	Project 1 (Excel and Python)
4	Data overfitting and its avoidance Activity: Project 1 Presentation	Chapter 5	Homework (Short Conceptual Question), Project, Presentation
	MID TERM		
5	Model Performance Analytics Activity: Project 1 Final Presentation	Chapter 7	Homework (Short Conceptual Question), Project, Presentation
6	Visualizing Model Performance Activity: Pre-Project Assignment	Chapter 8	Homework (Short Conceptual Question)
7	Prediction via evidence combination Representing and Mining text Introduction to API Activity: Project 2 Presentation	Chapter 9,10	Homework (Short Conceptual Question)
	MID TERM		
8	Similarity, Distance, Nearest, Neighbor, Clustering Case Study: TBD Activity: Project 2 Final Presentation	Chapter 5	HW Report
10	Other Data Science and Technique Visualization Activity: Tableau Introduction and Setup	Chapter 12	Lab: Tableau. Presentation
11	Big Data: Towards Analytical Engineering Introduction to Hive Activity: Hands on Hive coding	Chapter 11	Lab
12	Data Science and Business Strategy Conclusion	Chapter 13	Final Project Presentation