



Syllabus

SCHOOL OF TECHNOLOGY & COMPUTING
CS 612 Data Exploration and Visualization

3 Credits
Effective: Summer 2020

*Access to the Internet is required.
All written assignments must be in Microsoft-Word-compatible formats.
See the library's APA Style Guide tutorial for a list of resources that can help you use APA style.*

FACULTY

Faculty Name: FACULTY NAME

Contact Information: CONTACT INFORMATION

[INSTRUCTOR MAY INSERT PERSONAL MESSAGE IF DESIRED]

COURSE DESCRIPTION

As the base of data science, data should be acquired, integrated, preprocessed, analyzed, and visualized. The data acquisition is a crucial step to ensure both the quantity and quality of data and improve the effectiveness of the following steps of data processing. For the data scientist, it is also important to be aware of the range of options and possibilities and to be able to deploy the analyses as appropriate. Thus, a data scientist must understand concepts and approaches of data acquisition, including data shaping, information extraction, information integration, data reduction and compression, data transformation, as well as data cleaning. Through the use of graphs and other forms of diagrams, visualization can be used in providing readily understood summaries but can also greatly assist in guiding such activities as clustering and classification.

COURSE RESOURCES

- McKinney, W. (2017). [*Python for Data Analysis : Data Wrangling with Pandas, NumPy, and IPython*](#). O'Reilly. (ISBN: 9781491957660)
- Han, J., Kamber, M. & Pei, J. (2011). [*Data Mining: Concepts and Techniques \(3rd ed.\)*](#).Morgan & Kaufmann. (ISBN: 9780123814791)

COURSE OUTCOMES

In this course, learners:

- Understand techniques for data acquisition according to the features of data sources and applications.
- Understand the adoption of a user-centered approach to analysis and presentation.
- Apply data reduction and compression steps for a given data.
- Analyze rules for data cleaning according to the requirement of applications and data semantics.
- Evaluate the effectiveness of data transformation.
- Create an effective approach to visualization on a set of data used for a particular purpose.

CORE CONCEPTS

Topics include:

- Getting Started with Python
 - Python Language Basics
- Working with Data Structures, Functions, and Files
 - Data Structures and Sequences
 - Functions
 - Files and the Operating System
- Using NumPy
 - The NumPy ndarray
 - Fast Element-Wise Array Functions
 - Array-Oriented Programming with Arrays

- File Input and output with Arrays
- Data Objects and Attribute Types
- Getting Started with pandas
 - pandas Data Structures
 - Essential Functionality in pandas
 - Summarizing and Computing Descriptive Statistics
- Reading and Writing Data
 - Data Preprocessing
 - Reading and Writing Data in Text Format
 - Binary Data Formats
 - Interacting with Web APIs
 - Interacting with Databases
- Cleaning and Preparing Data
 - Data Cleaning
 - String Manipulation
- Combining, Joining, and Rearranging Data
 - Data Integration
 - Hierarchical Indexing
 - Combining and Merging Datasets
 - Reshaping and Pivoting
- Transforming Data
 - Data Transformation
 - Data Discretization
- Visualizing Data
 - Data Visualization
 - Matplotlib API Primer
 - Plotting with pandas and seaborn
 - Other Python Visualization Tools
 - Data Visualization Tools
- Aggregating Data and Applying Group Operations
 - Data Reduction
 - GroupBy Mechanics
 - Data Aggregation
 - Apply: General split-apply-combine
 - Pivot Tables and Cross-Tabulation

OVERVIEW OF COURSE GRADING

The grades earned for the course will be derived by using City University of Seattle's decimal grading system, based on the following:

OVERVIEW OF REQUIRED ASSIGNMENTS	% OF FINAL GRADE	POINTS
The Muddiest Point (MP)	5%	50 = 5 points * 10 modules
Concept Test (CT)	5%	50 = 5 points * 10 modules
Discussion Board (DB)	10%	100 = 10 points * 10 modules
Hands-On Practice (HOP)	20%	200 = 20 points * 10 modules
Programming Exercise (PE)	30%	300 = 30 points * 10 modules

Knowledge Check (KC)	10%	100 = 10 points * 10 modules
Team Project (TP)	20%	Proposal: 30 points Progress: 70 points Final Report: 70 points Final PPT: 30 points Subtotal: 200 points
TOTAL	100%	1,000 points

The following approaches are used for developing this course content:

Assessment

- Summative Assessment. https://en.wikipedia.org/wiki/Summative_assessment
- Formative Assessment. https://en.wikipedia.org/wiki/Formative_assessment

Classroom Assessment Techniques

- The Muddiest Point. https://en.wikipedia.org/wiki/Classroom_Assessment_Techniques

Active Learning. https://en.wikipedia.org/wiki/Active_learning

- Flipped Classroom. https://en.wikipedia.org/wiki/Flipped_classroom
- Just-in-time Teaching (JiTT). https://en.wikipedia.org/wiki/Just-in-time_teaching
- Peer Instruction. https://en.wikipedia.org/wiki/Peer_instruction

Learning Theory

- Learning-by-doing. <https://en.wikipedia.org/wiki/Learning-by-doing>
- Project-Based Learning (PBL). https://en.wikipedia.org/wiki/Project-based_learning
- Social Learning. [https://en.wikipedia.org/wiki/Social_learning_\(social_pedagogy\)](https://en.wikipedia.org/wiki/Social_learning_(social_pedagogy))

Evidence-Based Practice (EBP). https://en.wikipedia.org/wiki/Evidence-based_practice

- Pair Programming. https://en.wikipedia.org/wiki/Pair_programming
- Stand-up Meeting. https://en.wikipedia.org/wiki/Stand-up_meeting
- Agile Software Development. https://en.wikipedia.org/wiki/Agile_software_development

SPECIFICS OF COURSE ASSIGNMENTS

The instructor will provide grading rubrics that will provide more detail as to how this assignment will be graded.

The Muddiest Point (MP)

Before class, students are required to finish the muddiest point activity. This activity is designed to stimulate student engagement in class. Also, the instructor uses feedbacks from the Muddiest Point in preparation for the classroom lecture to implement Just-in-Time Teaching (JiTT). This activity consists of writing a brief reflective essay (<= 50 words) in which students identify the most confusing part (i.e., the muddiest point) of the content covered in the upcoming module. If you do not have an MP, you can explain the most exciting aspect. Also, students will answer one multiple choice question from the required reading to determine students' grasp of core concepts.

<i>Components</i>	<i>% of Grade</i>
Quality Participation: Meets requirements on time.	60%
Writing: Is clear, concise, and grammatically correct.	20%

Accuracy: Answers quizzes correctly.	20%
TOTAL	100%

Concept Test (CT)

In class, students may be required to answer questions called Concept Tests, which allows peers to teach others, i.e., Peer Instruction. 1) The instructor poses a problem based on students' responses to their pre-class reading. 2) Students reflect on the question. 3) Students commit to a definite answer. 4) Instructor reviews student responses without giving the correct answer to the students. 5) Students discuss their thinking and solutions with their peers. 6) Students then commit again to a specific answer. 7) The instructor back reviews responses and decides whether more explanation is needed before moving on to the next concept. Any participating students will earn their 100% grade.

<i>Components</i>	<i>% of Grade</i>
Quality Participation: Meets requirements on time.	100%
TOTAL	100%

Discussion Board (DB)

All classes are required to use the Discussion Board. Participation through DB is an integral part of this course. It is defined as active engagement in a discussion or other activity. Instructors will determine the type of activities and their due dates; moreover, different DB activities will have different substance and length guidelines. The instructor will provide specific instructions to students.

A discussion question or topic from the instructor appears weekly in the Discussion Board. Students post their answers and responses to two other students' ones in the DB by the end of each module. The DB is to help promote student to student engagement. The instructor may not respond to each posting.

Questions or comments specifically for the instructor should be emailed directly to the instructor or posted in the Question and Answer Forum. Students who want to talk with other students about issues unrelated to the discussion forums should use the Coffee Talk Forum.

Although the tone of your DB postings can be informal, your instructor will expect the content to be on a professional level. Your comments and questions for discussion should be clear and thoughtful, with correct grammar, spelling, and punctuation. As with written assignments, the quality of your discussion postings will be graded on both content and presentation.

<i>Components</i>	<i>% of Grade</i>
Quality Participation: Meets requirements on time.	80%
Writing: Is clear, concise, and grammatically correct.	20%
TOTAL	100%

Hands-on Practice (HOP)

The instructor may assign hands-on practice exercises to a pair of students in class or individually online. Students will learn and practice either specific tools or languages pertinent to their course. Each activity will be graded by pass or fail to encourage collaboration among students. (Pair programming can be used for the generation of more diverse solutions to problems.)

<i>Components</i>	<i>% of Grade</i>
Quality Participation: Meets requirements on time.	80%
Accuracy: Answers questions correctly.	20%
TOTAL	100%

Programming Exercise (PE)

The students must individually perform the programming exercise, which is based on the topics and Hands-on Practice. No code sharing or copying from other sources are allowed. Non-executable programs will not be graded. The programs in poor coding styles will be asked to be resubmitted. **Please note that copying a segment of code from the Internet and submitting it as your work is considered as plagiarism.**

<i>Components</i>	<i>% of Grade</i>
Accuracy: Answers questions correctly.	80%
Writing: Is clear, concise, and grammatically correct.	20%
TOTAL	100%

Knowledge Check (KC)

Students will complete weekly quizzes from the course content to reflect on what they have learned in the course. Completing all KCs will help ensure that you successfully master the concepts in this course. The best way for you to gain a thorough understanding of the underlying concepts is to apply those concepts to solve the quizzes. You should focus on the underlying principles, rather than just memorizing information.

<i>Components</i>	<i>% of Grade</i>
Accuracy: Answers questions correctly.	100%
TOTAL	100%

Team Project (TP)

Each student can select his or her team. Each team consists of three students. A group of fewer than three students requires the instructor's approval. Each team will use an instructor approved topic relevant to the course.

The paper must be no less than 6-7 pages. We required you to use the paper template from [EDSIG/CONISAR](#), the international conference standard. *The best papers in this course may be submitted to conferences with your team's approval and the instructor's recommendation and revisions. Paper submission is optional and has nothing to do with your course grade.*

The report template is provided. The file name consists of team project number, team number, and list of first names of your team members:

- TP01 for the proposal
- TP02 for the progress report
- TP03 for the final report
- TP04 for the final presentation slide

As in any scholarly writing, students should not merely copy information from another author. Students should use evidence to support the contentions they have drawn from their findings and critically analyze related literature. In essence, each paper needs to be an analytical paper, not a summary of readings.

In addition, a team presentation slide is required.

- The presentation consists of 15+4 slides: 15 slides for content, 4 slides for cover, agenda, key reference, and Q&A.
- The PPT template is provided. Your team can change design and color for your team's purpose.
- If necessary, a presentation video (15 minutes) may be requested for an online class or a missing member.

- If necessary, a demo video (a maximum of 1-2 minutes) may be requested. But, the total presentation time (15 minutes) will not be changed.

The following two resources are useful to improve your presentation slides:

- St. George International School of English. (2013, Nov 14). [Steve Jobs Presentation Skills](https://www.youtube.com/watch?v=iJq-thyDF9Q) (7:34). Retrieved from <https://www.youtube.com/watch?v=iJq-thyDF9Q>
- Gyaantastic. (2017, Feb 5). [7 Presentation Skills to learn from Steve Jobs](https://medium.com/@gyaantastic/7-presentation-skills-to-learn-from-steve-jobs-8fbfdebc4fc4). Retrieved from <https://medium.com/@gyaantastic/7-presentation-skills-to-learn-from-steve-jobs-8fbfdebc4fc4>

Four submissions are required according to the following schedule:

- Proposal (1 page; 30 points) - Starting (Module 1) & Ending (Module 3)
- Progress Report (3-4 pages; 70 points; graded after the proposal has been submitted) - Starting (Module 4) & Ending (Module 7)
- Final Report (6-7 pages; 70 points; graded after the progress has been submitted) - Starting (Module 8) & Ending (Module 10)
- Final PPT (15+4 slides, 30 points; graded after the final report has been submitted) - Starting (Module 8) & Ending (Module 10)

Students are expected to use the assigned readings, videos, and other materials throughout the quarter. Students will need to utilize additional sources that were not assigned by the professor. While stylized after an industry report, nonetheless, students are expected to employ APA formatting of citations, footnotes, and bibliography. Students must cite the sources of all ideas, facts, and information used that are not their own, even if they have put the information into their own words. Failure to do so is plagiarism, although the oversight is unintentional.

<i>Components</i>	<i>% of Grade</i>
Structure: Consists of the required report elements.	10%
Content: Demonstrates critical analysis & synthesis of concepts.	30%
Reference: Is pertinent to the topic and cited appropriately.	10%
Writing: Is clear, concise, and grammatically correct.	10%
Visual Presentation: Is well designed, legible, and persuasive.	20%
Team Collaboration: Is based on peer review.	20%
TOTAL	100%

COURSE POLICIES

Late Assignments

The late assignment without advanced notice will receive an -10% penalty per day for a maximum of 7 days.

Participation

Students will participate in activities and discussions as defined by the instructor. Whether in class, online, or a mixed-mode setting, students will be graded on the following things: their participation in classroom discussions; their ability to present, explain, or defend alternative viewpoints; and the degree to which they have mastered the concepts and principles inherent in the study of the subjects.

Professional Writing

Assignments require error-free writing that uses Standard English conventions and logical flow of organization to address topics thoroughly and concisely. CityU requires the use of APA style.

UNIVERSITY POLICIES

You are responsible for understanding and adhering to all of City University of Seattle's academic policies. The most current versions of these policies can be found in the [University Catalog](#) linked to the CityU Website.

Scholastic Honesty

Academic honesty in students requires the pursuit of scholarly activity that is free from fraud, deception, and unauthorized collaboration with other individuals. You are responsible for understanding CityU's policy on academic honesty and adhering to its standards in meeting all course requirements. A complete copy of this policy can be found in the [University Catalog](#) in the section titled *Scholastic Honesty* under *Student Rights & Responsibilities*.

Attendance

Students taking courses in any format at the University are expected to be diligent in their studies and to attend class regularly.

Regular class attendance is essential in achieving learning outcomes in the course and maybe a valid consideration in determining the final grade. For classes where a physical presence is required, a student has attended if s/he is present at any time during the class session. For online courses, a student has attended if s/he has posted or submitted an assignment. A complete copy of this policy can be found in the [University Catalog](#) in the section titled *Attendance Policy for Mixed Mode, Online and Correspondence Courses*.

SUPPORT SERVICES

Disability Resources

If you are a student with a disability and require an accommodation, please contact the Disability Resource Office as soon as possible. For additional information, please see the section in the [University Catalog](#) titled *Students with Special Needs* under *Student Rights & Responsibilities*.

Library Services

CityU librarians help you find the resources and information you need to succeed in this course. Contact a CityU librarian through the [Ask a Librarian](#) service, or access [library resources and services online](#), 24 hours a day, seven days a week.

Smarthinking

As a CityU student, you have access to 10 free hours of online tutoring offered through Smarthinking, including writing support from certified tutors 24 hours a day, seven days a week. Contact CityU's Student Support Center at help@cityu.edu to request your user name and password.