Data Science Senior Project Section 01

CS 163

Fall 2024   3 Unit(s)   08/21/2024 to 12/09/2024   Modified 07/16/2024

🖵 Course Description and Requisites

Semester-long individual or team project. Apply knowledge and technology to solve a realistic data science problem, including the entire process of collecting and processing real-world data, applying suitable analytic methods, explaining analysis outcomes, and making appropriate recommendations.

Prerequisite(s): CS 100W, CS 131, and either CS 156 or CS 171.

Letter Graded

✳ Classroom Protocols

Communication with the instructor

Students are requested to use the Canvas message function to contact the instructor. Private messages sent to the instructor's email address gets lost due to the large volume of emails received.

The instructor does not write messages after normal business hours, on weekends or holidays.

Reviewing code for the homework and technical trouble-shooting should be done during the office hours.

▤ Program Information

Diversity Statement - At SJSU, it is important to create a safe learning environment where we can explore, learn, and grow together. We strive to build a diverse, equitable, inclusive culture that values, encourages, and supports students from all backgrounds and experiences.

📊 Course Learning Outcomes (CLOs)

Upon successful completion of this course, students will be able to:

- Carry out a data science project from the beginning of problem analysis to the end of analytical results.
- Construct a literature search and summarize the state of the arts.
- Translate the project objectives into a realistic work plan.
- Design and implement required software using data analysis tools such as Python, R, MatLab, and programming languages.
- Present professionally, in both oral presentation and technical report, including project plan, design, implementation, analysis, final results, and recommendations.

📘 Course Materials

Suggested Reference:

- Peter Bruce, Andrew Bruce, and Peter Gedeck, "Practical Statistics for Data Scientists: 50 + Essential Concepts using R and Python (2nd ed.)", O'Reilly Media, 2020, ISBN: 149207294X. [You can read this book for free with your SJSU account: https://library.sjsu.edu/ebooks/safari-books-online-o-reilly (https://library.sjsu.edu/ebooks/safari-books-online-o-reilly)]
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York, Springer. [You can download this book from the authors' webpage: https://hastie.su.domains/ElemStatLearn/ (https://hastie.su.domains/ElemStatLearn/)]

☰ Course Requirements and Assignments

- This course is conducted in a workshop style, where students bring their proposals and results and discuss them as a group.
- Assignments are designed to provide a timeline to complete an individual project by the end of semester. Making an appropriate schedule is a part of the project task.
- The submitted assignments will be discussed during the lecture time.
- Active participation to the discussion is required to obtain the full credits in each assignment.

Item	Percentage
Git repo	5%
Proposal draft	10%
Proposal document	20%
Analysis plan	10%
Analysis outcome presentation	20%
Visualization plan	10%
Final presentation & Web page	20%

Item	Percentage
In-class exercise	5%

✔ Grading Information

Extra-credits and Reworks
No extra-credit assignments or rework opportunities will be given.

Late Submission
Late submissions within 24 hours will be deducted 10% of its final grade. Submissions over 24 hours late will have 20% grade deducted. Late submissions over 2 days will not be accepted.

Missed Assignments or Exams
When students need to miss an assignment deadline or exam due to health conditions or any other emergency, it should be reported within one week after the due date.

Final Grade Table

Total Grade	Letter Grade
97% and above	A plus
93% to 96%	A
90% to 92%	A minus
87% to 89%	B plus
83% to 86%	B
80% to 82%	B minus
77% to 79%	C plus
73% to 76%	C
70% to 72%	C minus
67% to 69%	D plus
65% to 66%	D
60% to 64%	D minus

59% and below	F

🏛 University Policies

Per [University Policy S16-9 (PDF) (http://www.sjsu.edu/senate/docs/S16-9.pdf)](http://www.sjsu.edu/senate/docs/S16-9.pdf), relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on the [Syllabus Information (https://www.sjsu.edu/curriculum/courses/syllabus-info.php)](https://www.sjsu.edu/curriculum/courses/syllabus-info.php) web page. Make sure to visit this page to review and be aware of these university policies and resources.

📅 Course Schedule

Date	Topic	Reference	Note
8/21	Intro		
8/26	pandas and Git		
8/28	Data collection: RESTful API & Web Scraping		
9/2	Labor Day; No class		
9/4	Data Cleaning		
9/9	Proposal Discussion		Due: Proposal draft
9/11	Exploratory Data Analysis	Practical Statistics Chap 1	
9/16	Data and Sampling	Practical Statistics Chap 2	
9/18	Statistical Experiments	Practical Statistics Chap 3	
9/23	Proposal Presentation 1		Due: Proposal document
9/25	Proposal Presentation 2		
9/30	Dash Basics	https://dash.plotly.com	
10/2	Dash: Plotly Express		
10/7	Dash Implementation		
10/9	Dash: Interactive Graphing		Due: Dash practice implementation
10/14	Regression	Practical Statistics Chap 4	

Date	Topic	Reference	Note
10/16	Classification	Practical Statistics Chap 5	
10/21	Statistical Machine Learning	Practical Statistics Chap 6	
10/23	Hands-on: Analyzing Data		Due: Analysis plan
10/28	Hands-on: Analyzing Data		
10/30	Hands-on: Analyzing Data		
11/4	Analysis Outcome Presentation 1		Due: Analysis outcome presentation
11/6	Analysis Outcome Presentation 2		
11/11	Veterans Day; No class		
11/13	Hands-on: Visualizing Data		Due: Visualization plan
11/18	Hands-on: Visualizing Data		
11/20	Final Presentation Discussion		
11/25	Final Presentation Discussion		
11/27	Fall break; No class		
12/2	Final Presentation 1		Due: Final presentation and Web page
12/4	Final Presentation 2		
12/9	Final Presentation 3		