SJSU SAN JOSÉ STATE UNIVERSITY

College of Science · Computer Science

Data Science Senior Project Section 01 CS 163

Spring 2025 In Person 3 Unit(s) 01/23/2025 to 05/12/2025 Modified 01/22/2025

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.

E 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 that involves data collection, data cleaning, analysis, and visualization.

- 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, and MatLab.
- 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: <u>https://library.sjsu.edu/ebooks/safari-books-online-o-reilly</u> (<u>https://library.sjsu.edu/ebooks/safari-books-online-o-reilly</u>]
- 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/ (https://hastie.su.domains/ElemStatLearn/ (https://hastie.su.domains/ElemStatLearn/ (https://hastie.su/ (<a href="h

⇐ 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
Proposal draft	5
Finalized proposal	15
EDA summary	10
Analysis and visualization plan	10
Analysis summary	10
Visualization results	10
Final presentation	15
Website	15
Midterm Exam	10

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%	А
90% to 92%	A minus
87% to 89%	B plus
83% to 86%	В
80% to 82%	B minus
77% to 79%	C plus
73% to 76%	С
70% to 72%	C minus
67% to 69%	D plus
65% to 66%	D
60% to 64%	D minus
59% and below	F

1 University Policies

Per <u>University Policy S16-9 (PDF) (http://www.sjsu.edu/senate/docs/S16-9.pdf)</u>, 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 <u>Syllabus Information</u>

(<u>https://www.sjsu.edu/curriculum/courses/syllabus-info.php</u>) web page. Make sure to visit this page to review and be aware of these university policies and resources.

📅 Course Schedule

Date	Торіс	Note	Reference
1/27	Intro		
1/29	Project samples		
2/3	Data collection: RESTful API & Web Scraping	Due: Group formation	
2/5	Data Cleaning		
2/10	Proposal Discussion	Due: Proposal draft	
2/12	Exploratory Data Analysis		Practical Statistics Chap 1
2/17	Sampling and Statistical Experiments		Practical Statistics Chap 2&3
2/19	Proposal Presentation 1	Due: Finalized proposal	
2/24	Proposal Presentation 2		
2/26	Dash: Basics		https://dash.plotly.com
3/3	Dash: Interactive Graphing and themes		
3/5	Dash: Exercise	Due: Link to your website	
3/10	Time Series Analysis	Due: EDA Summary	https://www.oreilly.com/library/view/time-series- forecasting/9781617299889/
3/12	Regression		Practical Statistics Chap 4
3/17	Statistical Machine Learning		Practical Statistics Chap 6

3/19	Interpretable ML		https://christophm.github.io/interpretable-ml-book/
3/24	Analysis and visualization plan presentation 1	Due: Analysis and visualization plan	
3/26	Analysis and visualization plan presentation 2		
3/31	Spring Recess (No class)		
4/2	Spring Recess (No class)		
4/7	Midterm review		
4/9	Midterm exam		
4/14	Analysis discussion	Due: Analysis summary	
4/16	Analysis discussion		
4/21	Visualization discussion	Due: Visualization results	
4/23	Visualization discussion		
4/28	Final presentaiton preparation		
4/30	Final Presentation 1	Due: Presentation slides	
5/5	Final Presentation 2		
5/7	Final Presentation 3		
5/12	Final Presentation 4		