SJSU SAN JOSÉ STATE UNIVERSITY

College of Science · Computer Science

Introduction to Data Visualization

CS 133

Spring 2026 In Person 3 Unit(s) 01/22/2026 to 05/11/2026 Modified 01/24/2026

Contact Information

Instructor: Jelena Segan

- **Email:** jelena.segan@sjsu.edu
- **Office:** DH 282

Class Days: Tuesday and Thursday
Time: 10:30AM - 11:45AM
Classroom: Health Building 407

Office Hours

- Tuesday 1:30 PM to 2:30 PM
- Thursday 1:30 PM to 2:30 PM

Course Information

Topics in data analysis and visualization. Covers tools and techniques to efficiently analyze and visualize large volumes of data in meaningful ways to help solve complex problems in fields such as life sciences, business, and social sciences.

Prerequisite(s): CS 146 with a grade of C- or better or CS 22B and graduate standing.
Grading: Letter Graded

Lecture

- Tuesday, Thursday 10:30 AM to 11:45 AM, Health Building 407

Course Description and Requisites

Topics in data analysis and visualization. Covers tools and techniques to efficiently analyze and visualize large volumes of data in meaningful ways to help solve complex problems in fields such as life sciences, business, and social sciences.

Prerequisite(s): CS 146 with a grade of "C-" or better, or CS 22B and graduate standing. Computer Science or Software Engineering majors only.

Letter Graded

Classroom Protocols

Students are expected to adhere to the Student Conduct Code (https://www.sjsu.edu/studentconduct/).

This semester, we'll be using a private Discord server as our main space for class communication and collaboration.

Discord is where we can:

- Ask and answer course-related questions.
- Share ideas, resources, and helpful tips.
- Learn from one another's perspectives and approaches.
- Continue discussions beyond class time.
- Organize and coordinate with your project team.

Instead of sending most questions by email, please post them in the appropriate Discord channel so the whole class can see and benefit from the responses. Your classmates may have the same question, and sometimes, a peer's explanation will click for you even faster.

A direct link to join will be given in Canvas.

Private or sensitive matters should be addressed through direct messages or discussed during office hours.

Office hours are the best time to get help with assignments, conceptual questions, or technical troubleshooting.

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 Goals

Understand:

- How to access data
- How to format data for visualizations

- Different data types and how to visualize them
- Data visualization as a tool for problem solving

Course Learning Outcomes (CLOs)

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

CLO 1 Manipulate large datasets and handle missing or inconsistent values in datasets.

CLO 2 Perform statistical analysis using packages such as Numpy and Scipy.

CLO 3 Analyze and visualize datasets using packages such as seaborn and matplotlib.

CLO 4 Develop interactive visualization using packages such as Plotly and Panel.

CLO 5 Recognize and reduce data and spatial biases.

Course Materials

Recommended readings:

Biological data exploration book with Python, Pandas and Seaborn

Author: Martin Jones
Year: 2020
ISBN: ISBN-13: 979-8612757238

https://pythonforbiologists.com/site_stuff/bde_book.html (https://pythonforbiologists.com/site_stuff/bde_book.html)

Hands-On Data Visualization: Interactive Storytelling From Spreadsheets to Code

Author: ack Dougherty and Ilya Ilyankou
Year: 2021
ISBN: ISBN-13: 978-1492086000

Free open-access web edition at https://HandsOnDataViz.org (https://HandsOnDataViz.org)

Course Requirements and Assignments

This course is designed to be hands-on and cumulative, preparing students to access data, format data for visualizations, utilize different data types and know how to visualize them, and use visualization as a tool for problem solving. Success in this course requires regular engagement, thoughtful experimentation, and a consistent time investment both inside and outside the classroom.

* Homework Assignments (15%) *

Weekly homework assignments will reinforce topics covered in class. All homework is individual and submitted via Canvas.

Collaboration is encouraged in concept, but your code must be your own. Academic integrity violations will be referred to the Office of Student Conduct.

* Quizzes (5%) *

Short quizzes will be administered in class 1-2 times per week. These are designed to assess retention of recent material and promote consistent engagement.

I understand that life happens. Students are allowed up to two make-up quizzes/homework during the semester, provided they notify the instructor in advance (when possible) or within a reasonable time after the missed class. Make-up quizzes/homework assignments must be completed within one week of the original date.

* Project Assignments (40%) *

As part of your CS 133 experience, you will join the SJSU DataVis Lab - a simulated, hands-on startup environment where you will work in rotating roles on a data-driven team. This project is designed to give you real-world exposure to data engineering and visualization, team collaboration, agile workflows, and data storytelling.

You won’t just learn about tools - you’ll build something meaningful with them.

The setup

You are part of the SJSU DataVis Lab, a fictive company where your professor serves as the CEO.

You will work in teams of 4-5 students. Each team becomes a Product Team responsible for one dataset-driven product of their own choosing. You will choose a dataset and theme of interest (e.g., sports analytics, transportation, music, education, social trends, etc.)

You will rotate roles every two-three weeks to gain experience across product, engineering, and storytelling functions, and gain better understanding of your preference in the types of role you like most.

Team Roles (Rotated every 4 weeks)

1. 1. Product Manager (PM)
    - Represents the team in meetings with the CEO (professor) during sprint planning (biweekly).
    - Translates strategic goals into actionable tasks.
    - Coordinates team responsibilities.
    - Owns the vision for the product and helps scope the sprint.

1. 2. Data Engineers (1-2 per sprint)
    - Implement the core technical work using course tools.
    - Follow Git best practices (branches, pull requests, peer review).
    - Attend weekly stand-up meetings with the CEO to share progress and raise blockers.

1. 3. Data Storyteller
    - Creates the biweekly project report.
    - Translates technical output into a clear narrative.
    - Connects CEO goals → PM tasks → engineering work → final insights.

Final Deliverables

- ○ A cleaned, well-documented Git repo with 3-4 sprints of development.
  ○ 3-4 professional reports from the storyteller (you'll rotate through this role).
  ○ A final 10-minute team presentation.
  ○ Peer Feedback Form

* Exams: Midterms and Final (40%) *

There will be two midterm exams and a comprehensive final exam.

- ○ Midterm 1 (10%)
  ○ Midterm 2 (10%)

- Final Exam (20%)

Exams will test:

- Conceptual understanding of data visualization tools
- Practical problem-solving
- Interpretation and optimization of code snippets

Make-up exams will only be considered for emergencies with proper documentation.

✔ Grading Information

Grading breakdown:

- Homework Assignments 15%
- Project Assignments 40%
- Quizzes 5%
- Midterm 1 10%
- Midterm 2 10%
- Final 20%
- Final grades:

Grade	Points
A plus	> 96
A	93 - 95.99
A minus	90 - 92.99
B plus	86 - 89.99
B	83 - 85.99
B minus	80 - 82.99

C plus	76 - 79.99
C	73 - 75.99
C minus	70 - 72.99
D plus	66 - 69.99
D	63 - 65.99
D minus	60 - 62.99
F	< 60

University Policies

Per University Policy 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) web page. Make sure to visit this page to review and be aware of these university policies and resources.

Course Schedule

The course schedule is subject to change with fair notice. Changes will be announced on Canvas.

Week	Day	Date	Topics
W1	Th	January 22	Syllabus. Introductions. Course Expectations.
W2	Tu	January 27	Introduction to Pandas
W2	Th	January 29	Pandas, Series, and DataFrame Objects

W3	Tu	February 3	Data Exploration Using Pandas
W3	Th	February 5	Data Exploration Using Pandas
W4	Tu	February 10	Advanced Features in Pandas
W4	Th	February 12	Intro to Seaborn
W5	Tu	February 17	Representing Categorical Data
W5	Th	February 19	Reshaping Data
W6	Tu	February 24	Reshaping Data
W6	Th	February 26	Handling Complicated Data Files
W7	Tu	March 3	Case Study: Applying Pandas and Seaborn for Problem Solving
W7	Th	March 5	Matrix Charts and Heatmaps
W8	Tu	March 10	Midterm Review
W8	Th	March 12	Midterm #1
W9	Tu	March 17	Introduction to Interactive Plots with Plotly
W9	Th	March 19	Creating Maps with Geopy and Folium

W10	Tu	March 24	Creating Interactive Dashboards with Panel
W10	Th	March 26	Handling Large Datasets with Polars
W11	Tu	April 7	Visualizing High-dimensional Data in a Low-dimensional Space
W11	Th	April 9	Relational Databases - SQLite, Accessing data in Google Drive and Google Sheets
W12	Tu	April 14	Integrating Data Visualization in Web App
W12	Th	April 16	Midterm #2 Review
W13	Tu	April 21	Midterm #2
W13	Th	April 23	Integrating Data Visualization in Web App
W14	Tu	April 28	Intro to Machine Learning
W14	Th	April 30	How We Use Data Visualization to Help Inform Machine Learning
W15	Tu	May 5	Final Project Due. Project Presentations

W15	Th	May 7	Project Presentations
W16	Wed	May 13	Final Exam 10:45 AM-12:45 PM