

Mathematical Methods for Data Analytics Section 11

DATA 220

Fall 2023 3 Unit(s) 08/21/2023 to 12/06/2023 Modified 08/17/2023

Class Days & Times: Th 6:00PM - 8:45PM

Classroom: Boccardo Business Center (BBC) 104

Contact Information

Instructor: Dr. Mohammad Masum

Email: mohammad.masum@sjsu.edu

Office: Clark Hall 418C

Phone: 408-924-2736

Office Hours

Thursday, 4:00 PM to 5:30 PM, Clark Hall 418C

or by appointment (via zoom)

Course Description and Requisites

Mathematical and Statistical methods for Data Analytics. Selected topics from probability, statistics, linear algebra, and mathematical optimization. Programming for numerical implementation of mathematical and statistical algorithms.

Prerequisite: Instructor Consent

Letter Graded

* Classroom Protocols

- Classes will start on time, so do not be late.
- If a student misses a class, the student is responsible for finding out what's said/done in that class (such as new announcement, deadline change, etc.) and responding accordingly.
- Please make sure to turn off or mute the cell phone during class.
- Please do not perform irrelevant or distracting activities in class.

Students Are Not Allowed to Record

Students are prohibited from recording class activities (including class lectures, office hours, advising sessions, etc.), distributing class recordings, or posting class recordings. Materials created by the instructor for the course (syllabi, lectures and lecture notes, presentations, etc.) are copyrighted by the instructor. This university policy (S12-7) is in place to protect the privacy of students in the course, as well as to maintain academic integrity by reducing the instances of cheating. Students who record, distribute, or post these materials will be referred to the Student Conduct and Ethical Development office. Unauthorized recording may violate university and state law. It is the responsibility of students that require special accommodations or assistive technology due to a disability to notify the Accessible Education Center (AEC) at <http://www.sjsu.edu/aec>. AEC will in turn contact the instructor as needed.

Course Format

- Canvas will be used to distribute lecture notes (and other class materials), schedule class activities, host online discussions, provide a dropbox for assignment submittal, and detect plagiarism.
- A student must actively check and follow the postings of this site.

Program Information

MS Data Analytics Program Learning Outcomes

PLO 1 - Integrate multidisciplinary knowledge to engage in practical data analytics projects, from analyzing requirements to managing data, building models, presenting results, and assessing societal impacts.

PLO 2 - Identify, use, and evaluate current and emerging multidisciplinary data analytics technologies and tools.

PLO 3 - Apply effective oral and written communication skills necessary to professional work including collaboration and presentation to multidisciplinary audiences.

PLO 4 - Perform independent research in applying data analytics to specific domains.

PLO 5 - Utilize quantitative skills from a rich foundation of essential knowledge to solve complex, dynamic, and practical data analytics problems in various professional domains.

PLO 6 - Work productively as individuals and in teamwork to perform data analytics tasks as part of a multidisciplinary team.

PLO 7 - Identify ways in which data analytics professionals can contribute to the cultural, economic, educational and social well-being in diverse and multicultural local, national and global contexts.

Course Learning Outcomes (CLOs)

Note: PLO= Program Learning Outcome. Upon successful completion of this course, students will be able to:

1. Demonstrate the capability to understand mathematical/statistical programming and implement it in diverse applications. [PLO 1,2]
2. Demonstrate the capability to build optimization models with various programming languages for real applications. [PLO 1, 2, 5]
3. Demonstrate the capability to analyze data through mathematical/statistical methods. [PLO 1, 2]

Course Requirements and Assignments

Students are expected to actively participate in classroom discussions, which often lead to a deeper understanding of the concepts. Students may collaborate on homework but must write independent code/solutions. Copying and other forms of cheating will not be tolerated and will result in a zero score for the homework (minimal penalty) or a failing grade for the course, possibly combined with other disciplinary actions from the university.

Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the syllabus.

Late Policy

- All Canvas deliverables are due at 11:59 pm.
- First late assignment (until the next day): 20% penalty
- Second late assignment (until the next day): 30% penalty
- No late assignment will be accepted after the second late assignment.
- Please do not ask for date extensions.
- Exceptions may be accommodated only if a student has serious and compelling reasons that can be proven by an independent authority (e.g., doctor's note if the student has been sick).
- If you have trouble uploading on Canvas, please email me with attached your homework.

Required Texts/Readings

Textbook

No required textbook for this course

Other technology requirements / equipment / material

Python (Pandas, NumPy, Scikit-Learn)

✓ Grading Information

Grade Distributions

Homework (HW)

- 5 HomeWorks (20 X 5 = 100).
- Theory: All word files, PDF, scanned documents, Jupyter Notebook are accepted.
- Practice: Jupyter Notebook

Midterm Exam

50 points

Final Exam

50 points

Breakdown

Grade	Points
A	188.00-200.00
A minus	178.00-187.99
B Plus	172.00-177.99
B	166.00-171.99
B minus	160.00-165.99
C plus	154.00-159.99
C	146.00-153.99
C minus	138.00-145.99
D plus	132.00-137.00
D	120.00-131.99
F	<120.00

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

The schedule (and related dates/assignments) is tentative and subject to change with the fair notice on Canvas as well as in the

classroom in due time.

Week	Date	Topics
1	Aug 24	Course Introduction: Syllabus and other materials, Python for Data Analysis
2	Aug 31	Probability and Statistics
3	Sep 7	Probability and Statistics
4	Sep 14	Probability and Statistics HW1 Due by Sep 14
5	Sep 21	Probability and Statistics
6	Sep 28	Probability and Statistics
7	Oct 5	Probability and Statistics HW2 Due by Oct 5
8	Oct 12	Midterm
9	Oct 19	Linear Algebra
10	Oct 26	Linear Algebra
11	Nov 2	Multivariate Calculus and Optimization HW3 Due by Nov 2
12	Nov 9	Multivariate Calculus and Optimization
13	Nov 16	Information Theory & Bayesian Probability HW 4 Due Nov 16
14	Nov 23	Campus Closed - Thanksgiving Day
15	Nov 30	Regression Analysis HW5 Due by Dec 6
Final Exam	Dec 14	5.15-7.30 PM

Probability and statistics topic covers - Descriptive and Inferential Statistics, Exploratory Data Analysis, Probability, Probability Distributions, Hypothesis Tests

Linear Algebra topic covers - Fundamentals of linear algebra, vector, matrix, Factorization, Dimension Reduction

Multivariate Calculus and Optimization topic cover - Fundamentals of multivariate calculus, partial derivatives, chain rule, gradients, fundamentals of optimization, Lagrange multipliers, gradient descent, Linear Regression with gradient descent,

Information Theory & Bayesian Probability topic covers - Fundamentals of Information theory, Entropy, Information gain, Gini Index, KL divergence, Decision Tree, Bayes Theorem, Naïve Bayes Classifier

Final Exam Schedule - <https://www.sjsu.edu/classes/final-exam-schedule/fall-2023.php>
