

BSDS 6500: SPECIAL TOPICS IN BIOSTATS

Fall, 2024

[3 credits]

Class start time: 09:00

Class end time: 10:30

Class location: Medical Education Bu Rm# 3220

COURSE INSTRUCTOR

Name: Dr. Anand Paul

Title, Academic Program:

Mailing street address: 2020, Gravier St

Office phone: 504-568-5894

Email: apaul4@lsuhsc.edu

Building/Office Room Number: Lion's Building, LEC 255 Office hours: By appointment

Moodle site: <https://moodle.lsuhs.edu>

COURSE DESCRIPTION

This course offers a comprehensive introduction to the field of Data Science, designed to equip students with the theoretical knowledge and practical skills necessary for analyzing, interpreting, and visualizing data. Students will embark on a journey through the Data Science Life Cycle, gaining insights into fundamental concepts, methodologies, and tools that drive data-driven decision making in contemporary settings.

LEARNING OBJECTIVES

Students successfully completing this course will be able to:

- Understand the Data Science Life Cycle: Grasp each stage of the data science life cycle, from data collection to model deployment, and understand the methodologies applied at each phase.
- Master Data Collection and Preprocessing Techniques: Acquire the skills to gather, clean, and preprocess data effectively, ensuring data quality and readiness for analysis.
- Conduct Exploratory Data Analysis (EDA): Develop the ability to perform exploratory data analysis, utilizing descriptive statistics and visualizations to uncover insights and inform further analysis.
- Apply Data Wrangling Techniques: Employ data wrangling and transformation techniques using tools such as Pandas to manipulate data sets and prepare them for analysis.
- Design and Interpret Data Visualizations: Create meaningful data visualizations with tools like Matplotlib and Seaborn to communicate data insights clearly and effectively.
- Engage in Ethical Data Science Practice: Recognize and address ethical considerations in data science, ensuring responsible use of data and analytics.
- Evaluation metrics for Data Analytics: Evaluate the performance of Data Analysis using standard metrics and techniques, understanding their implications for model selection and improvement.
- Explore Markov : Delve into more Predictive approaches, including Markov models understanding their applications and limitations. Recognizing their potential applications and impact on the field of data science.
- Complete a Capstone Data Science Project: Demonstrate the ability to apply data science techniques to a real-world problem, from problem definition through to solution presentation and evaluation.

- **Communicate Data Science Findings:** Effectively communicate findings and insights from data science projects to a variety of audiences, employing clear visualizations and presentations.

MASTER OF PUBLIC HEALTH KNOWLEDGE AND SKILLS

This course tailored to MPH students, integrating these foundational public health knowledge areas would likely involve:

- *Lecture Content:* Incorporating lectures that relate data science concepts to these core public health areas. For example, using biostatistics as a basis for teaching data analysis techniques or discussing how data science can support epidemiological research.
- *Case Studies and Projects:* Including case studies and projects that require students to apply data science methods to real-world public health issues, such as analyzing epidemiological data, evaluating health policies, or assessing environmental health risks.
- *Interdisciplinary Approach:* Encouraging an interdisciplinary approach by highlighting the intersection between data science and public health, emphasizing how data-driven insights can inform public health decision-making and intervention strategies.

By weaving these foundational public health knowledge areas into the course curriculum, students can develop a well-rounded understanding of how data science techniques can be applied to address public health challenges, ultimately enhancing their competency as public health professionals in a data-driven world

EVALUATION AND GRADING PROCEDURES

Evaluation/Assessment

Students are expected to spend a minimum of 3 hours per hour contact on out-of-class work. All exams will include in-class, closed-book sessions that may also include take-home sections.

Evaluations:

Homework Assignments	25%
Midterm Exam	25%
Research project	25%
Final Exam	25%

Grading:

A=90-100%
B=80-89.9%
C=70-79.9%
D=60-69.9%
F=Below 60%

Homework and assignments throughout the course will assess students' ability to apply data science techniques to real-world health data. This includes data preprocessing, exploratory data analysis, machine learning model implementation, and data visualization. Assignments are designed to reinforce theoretical knowledge through practical application. The project component involves a comprehensive capstone project where students will tackle a public health issue using data science methods. This project will culminate in a presentation, allowing students to demonstrate their ability to conduct a complete data science project from start to finish. The grade for this component will be an average of the instructor's grade and peer grades from presentation evaluations. The midterm exam will cover all material presented in the first half of the course. It will include objective type, short answer and essay questions and some python based solutions, The final exam will primarily focus on the material covered in the latter third of the course but will also require students to integrate knowledge from the entire semester., it has similar pattern as midterm exam.

Assignment	%
Homework 01	5
Homework 02	5
Homework 03	5
Homework 04	5
Homework 05	5
Total	25

Late assignments: No Score

POLICIES AND EXPECTATIONS

Accessibility Statement: LSUHSC-New Orleans is committed to creating an inclusive and accessible environment according to the ADA, ADAA, and Section 504 of the Rehabilitation Act of 1973. Students in need of classroom accommodations should contact the Office of Disability Services (ODS) at ods@lsuhsc.edu. Please keep in mind that accommodations take effect when an accommodation letter has been generated; they are not retroactive. New accommodation letters need to be requested every academic term in which you are enrolled. More information can be found on the [ODS website](#).

Diversity Statement: LSUHSC-New Orleans welcomes and respects individuals from diverse backgrounds and viewpoints. All faculty and students should be treated with dignity and empathy. Discrimination and harassment of any kind will not be tolerated and can be reported via the [Sexual Misconduct & Power-Based Violence Incident Report Form](#).

Names & Pronouns: All students should be treated with professional consideration and respect. While our university data system requires the use of legal names and gender markers on official documents, students are welcome to inform instructors of the pronouns and name that they use. Preferred names can be modified on ID badges as well. Students who have questions concerning this policy are encouraged to contact the [university registrar](#).

Title IX Statement: LSUHSC-NO promotes integrity, civility, and mutual respect in an environment free from harassment and discrimination based on sex, gender, sexual orientation, gender identity, sexual misconduct, and power-based violence. As your course director and/or instructor, I am mandated to report to the [Title IX Coordinator](#) any incident of Title IX or power-based violence told directly to me. Exceptions to this required reporting include disclosures that are shared in the course of academic work such as a class discussion, group work, etc. Confidential resources that do not have the reporting requirement are available through CAP and the Ombuds.

[All gender restrooms](#) can be found in several locations on campus.

LSUHSC-NO does not discriminate against anyone with pregnancy or parenting status. Reasonable accommodations are available to anyone who requests them. [Lactation spaces](#) are available across campus as well.

Academic Dishonesty and Professional Behavior: The faculty and administration in the School of Public Health are committed to maintaining the integrity of its academic requirements and procedures and fostering an environment of honest intellectual inquiry. The [LSUHSC SPH Student Handbook](#) establishes the School's definition of student misconduct, provides examples of the categories of unacceptable misconduct, describes the procedures to be followed in suspected cases and outlines the penalties, which

may be appropriate in proven cases. Students are expected to acquaint themselves with school-wide policies and procedures found in the [LSUHSC SPH Student Handbook](#) along with the LSUHSC Student Rights and Responsibilities outlined in [Chancellor's Memorandum-56](#). Chancellor's Memorandum-56 describes the rights, roles, and responsibilities of LSUHSC-NO students. It also defines the processes and procedures for both student complaints and disciplinary proceedings, including those required for compliance with the "Louisiana Student Due Process and Protection Act" (La. R.S. 17:3394), to sustain a professional, ethical, and equitable learning community at LSUHSC-NO.

Artificial intelligence (AI) language models: The use of AI language models, such as ChatGPT, is determined by the faculty and if allowed, any use of AI-generated material must have appropriate citation. Examples of citing AI language models are available at libguides.umn.edu/chatgpt. You are responsible for fact checking the accuracy of statements composed by AI language models.

- * [How to cite generative AI output \(MLA\)](#)
- * [How to cite generative AI output \(APA\)](#)
- * [How to cite generative AI output \(Chicago\)](#)

Failure to cite is considered a violation of the LSUHSC SPH [policy](#) on academic misconduct in the Student Handbook. Violations could result in failure of the assignment or failure of the course.

READINGS

Suggested:

1. The Data Science Handbook, 2017 first edition (John Wiley & Sons) by Field Cady
2. Data science from scratch first principles with python 2nd edition (O'Reilly) 2nd Edition by Joel Grus
3. Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython (2017, O'Reilly) 2nd edition by Wes McKinney.

Most of the materials are given as PPT / Notes/ Videos developed by the professor, and some are taken from different sources (appropriate citations are provided).

CALENDAR/SCHEDULE

Tue.	Thu.	Topics*
	8/22	Class Introduction
8/27	8/29	Data Science Life Cycle, Data Collection, Data Preparation
9/3	9/5	Exploratory Data Analytics (python) (Homework 01)
9/10	9/12	Graphical Representation/ Data Wrangling (python) (Homework 02)
9/17	9/19	Principles of Data Visualization (tableau)
9/24	9/26	Orange Software (demonstration) (Homework 03)
10/1	10/3	Capstone Project Discussion / Midterm Exam (10/3)
10/8	10/10	Public Health – Data Analytics Case Study - I
10/15	10/17	Public Health – Data Analytics Case Study - II
10/22	10/24	Midterm Exam
10/29	10/31	Supervised/Unsupervised Learning Introduction (Homework 04)
11/5	11/7	K-Means Clustering Introduction
11/12	11/14	Random Forest Introduction (Homework 05)
11/19	11/21	Ethics of Data Science

11/26	11/28	Final Presentation
12/3	12/5	Final Presentation
12/10		Final Exam
