Course Information

Course Number:	ECON 470
Course Title:	Data Science for Economic and Social Issues
Section:	500
Time:	02:40 – 03:55, MW
Location:	ALLN 1002
Credit Hours:	3

Instructor Details

Instructor:	M. Jahangir Alam
Office:	LASB 251
E-Mail:	jalam@tamu.edu
Office Hours:	11:30 – 12:30 PM, MW or by appointment

Teaching Assistant:
Office:
E-Mail:
Office Hours:

Course Description

This course integrates economics with cutting-edge data science methodologies. It guides students through the process of utilizing data analysis techniques to address real-world economic and social issues. From mastering basic data science tools to delving into intricate causal estimation methods, you will acquire the skills needed to interpret and analyze data in the context of policy analysis. The course also features a collaborative group project culminating in poster presentations. This enables students to apply theoretical concepts to contemporary challenges and to communicate their findings effectively.

Course Prerequisites

Grade of C or better in ECON 323; grade of C or better in ECMT 461, STAT 211, or STAT 303; or approval of instructor.

In this course, we will use **Python** extensively. No prior programming experience is necessary, but a willingness to learn these programming tools is essential. This course is designed for students who are interested in integrating programming with economics to address real-world problems. Rest assured, all required computational knowledge will be provided during the course. Please remember to bring a laptop to class. In addition, you are required to access **ChatGPT Plus** and **GitHub Copilot**.

Course Learning Outcomes

Upon successful completion of ECON 470, students will be able to:

- Apply appropriate data science techniques to analyze and interpret data related to economic and social issues.
- Identify and address issues such as endogeneity, selection bias, omitted variable bias, and multicollinearity in economic analysis.
- Utilize various causal estimation methods like IV (Instrumental Variable), DID (Difference in Differences), RDD (Regression Discontinuity Design), PSM (Propensity Score Matching), and ITS (Interrupted Time Series) to infer causal relationships in economic data.
- Demonstrate the ability to use machine learning techniques such as Prophet and LSTM (Long Short-Term Memory) in economic analysis.
- Perform data cleaning, preprocessing, and visualization to prepare data for analysis and interpret results.
- Design and present a poster presentation, communicating complex data and findings clearly to both specialized and general audiences.

Textbook

There is no required textbook for this course. However, the following books could be useful.

1. “Econometric Data Science: A Predictive Modeling Approach” by Francis X. Diebold
2. "Introductory Econometrics: A Modern Approach" by Jeffrey M. Wooldridge.

Grading Policy

Weighting (toward overall course grade):
Quizess - 20%
Midterms (two) – 40%
Presentation – 20%
Report – 20%

Final grades will be calculated as follows:
A = 90-100
B = 80-89
C = 70-79
D = 60-69
F = 59 and below

Course Schedule

Lecture	Date	Topics
Module 1: Introduction to Data Science and Economic Analysis		
Lecture 1	Wednesday, January 17	Introduction to data science tools
Lecture 2	Monday, January 22	Basics of economic policy analysis
Lecture 3	Wednesday, January 24	Quantitative methods in economics
Module 2: Addressing Correlational Analysis Problems		
Lecture 4	Monday, January 29	Endogeneity, selection bias, omitted variable bias
Lecture 5	Wednesday, January 31	Multicollinearity and causality identification
Lecture 6	Monday, February 05	Approaches to address these problems
Module 3: Causal Estimation Techniques		
Lecture 7	Wednesday, February 07	Instrumental Variables (IV)
Lecture 8	Monday, February 12	Difference-in-Differences (DID)
Lecture 9	Wednesday, February 14	Regression Discontinuity Design (RDD)
Lecture 10	Monday, February 19	Propensity Score Matching (PSM)
Lecture 11	Wednesday, February 21	Interrupted Time Series (ITS)
Lecture 12	Monday, February 26	Midterm 1
Module 4: Data Handling and Machine Learning in Economics		
Lecture 13	Wednesday, February 28	Machine learning integration in economics
Lecture 14	Monday, March 04	Machine learning integration in economics
Lecture 15	Wednesday, March 06	Data preprocessing and visualization
Lecture 16	Monday, March 18	Data preprocessing and visualization
Lecture 17	Wednesday, March 20	Introduction to Prophet for forecasting
Lecture 18	Monday, March 25	Introduction to Prophet for forecasting
Lecture 19	Wednesday, March 27	Introduction to LSTM for sequence data analysis
Lecture 20	Monday, April 01	Introduction to LSTM for sequence data analysis
Lecture 21	Wednesday, April 03	Introduction to LSTM for sequence data analysis
Lecture 22	Monday, April 08	Introduction to LSTM for sequence data analysis
Lecture 23	Wednesday, April 10	News Sentiment and Stock Price
Lecture 24	Monday, April 15	News Sentiment and Stock Price
Lecture 25	Wednesday, April 17	Midterm 2
Lecture 26	Monday, April 22	Discuss about the presentation
Lecture 27	Wednesday, April 24	Group presentations and peer review
Lecture 28	Monday, April 29	Group presentations and peer review

Quizzes:
There will be several unannounced quizzes, either in-person or online, during class time. If you miss a quiz, there will be no makeup opportunity. Although attendance will not contribute points, missing classes might result in missing quizzes. Note: Attendance will be recorded to keep track of class participation.

Midterms:
There will be two in-person midterms. Detailed information regarding the content and due dates of these midterms is provided in the table above. Sample midterms will be made available to help you understand the content and format of these exams.

Presentation and Report:
The course includes a group project for the presentation and report. The class will be divided into four groups, which I will randomly assign on Canvas. Below, you will find a sample project for reference. Each group may work on this same project. However, if you are interested in pursuing a different project with your group members, please discuss it with me as soon as possible. The report submission deadline is Friday, May 3, 2024.

Poster Presentations:
Below are the general guidelines for the poster presentations.

- **Content Clarity**: Clearly define the problem statement, research objectives, and policy goals related to your topic. Include concise explanations of the techniques used, key findings, and recommendations.
- **Organization and Structure**: Structure the content logically, starting with an introduction, followed by methodology, analysis, results, and conclusion. Ensure that the content flows smoothly and maintains coherence.
- **Visualizations**: Include relevant visualizations like graphs, charts, and maps that help illustrate your findings. Ensure that the visualizations are clearly labeled and support the narrative.
- **Readability**: Maintain a consistent font size and style throughout the poster. The main text should be readable from a distance of at least 3 feet. Use bullet points, headers, and subheaders to enhance readability.
- **Design Aesthetics**: Utilize a professional and visually pleasing design. Maintain a balance between text and visuals and avoid cluttering. Use colors that complement each other and enhance readability.
- **Citations and References**: Include appropriate citations for data sources, methodologies, and key literature that informs your research. Follow a standard citation style consistently.
- **Engagement and Communication**: Prepare to engage with the audience by explaining your poster, answering questions, and discussing implications. Practice communicating complex information in an accessible manner.
- **Size and Printing**: Follow the specified size guidelines for printing the poster, and ensure that it is of high quality with clear and sharp images and text. **Be aware that this may involve additional costs and time. Alternatively, you may choose to write a report and give a standard presentation.**

General Guidelines for the Report

1. **Introduction**
   - Outline the purpose and scope of the report.
   - Provide a brief summary of the project idea you're addressing.
   - State the hypotheses or questions you intend to answer.
2. **Methodology**
   - Describe the methods used for data collection (e.g., case studies, industry data, surveys).
   - Explain the analytical techniques used (e.g., econometric analysis).
3. **Literature Review**
   - Summarize relevant theories and previous research related to your topic.
   - Identify gaps in existing research to justify your study.
4. **Findings**
   - Organize the findings logically, aligning them with the project's purpose.
   - Use graphs, tables, and charts to present data clearly.
   - Clearly state what the data reveals about the topic.
5. **Discussion**
   - Interpret the findings in the context of the research questions or hypotheses.
   - Compare your findings with existing literature.
6. **Conclusions**
   - Summarize the key findings.
   - Discuss the real-world implications of the results.
   - Suggest areas for further research.
7. **References**
   - Cite all sources and references used in the report according to a recognized citation style like APA, MLA, or Chicago.
8. **Appendix (if necessary)**
   - Include supplementary material like detailed statistical analyses, additional graphs, etc.

Sample Topic: News Sentiment and Stock Price: A Deep Learning Approach

Introduction

- **Issue Explanation:** Explain the potential relationship between news sentiment and stock price fluctuations, highlighting the relevance of this analysis in modern finance and investment strategies.
- **Objective:** Investigate how dispersion in news sentiment affects stock price, using data from the CRSP (Center for Research in Security Prices), COMPUSTAT, and ProQuest databases.

Guideline for Poster Presentation

Problem Statement

- **Description:** Define the problem of predicting stock price movements and the role of news sentiment in influencing these movements.
- **Objective:** Detail the purpose of the study, which is to apply Deep Learning and Sentiment Analysis techniques to understand and possibly predict stock price changes based on news sentiment.

Methodology

- **Modeling Technique:** Apply LSTM (Long Short-Term Memory) networks, specialized in recognizing patterns in sequences of data.
- **Sentiment Analysis:** Analyze the sentiment of news articles using Natural Language Processing techniques, including data drawn from ProQuest for media analysis.
- **Data Preprocessing:** Describe the process of data cleaning and preparation from CRSP, COMPUSTAT, and ProQuest, including feature extraction, normalization, etc.

Analysis

- **Stock Price Correlation:** Visualization showing the relationship between news sentiment and stock price over time.
- **Dispersion in Sentiment:** Graph illustrating how the dispersion in sentiment is correlated with stock price fluctuations.
- **Model Performance:** Table or chart detailing the performance metrics of the LSTM model, including accuracy, precision, recall, etc.

Results

- **Findings:** Summarize key findings, such as the effectiveness of news sentiment as an indicator for stock price movement.
- **Model Insights:** Highlight insights gained from the LSTM model, such as the ability to capture long-term dependencies between news sentiment and stock price.

Conclusion & Implications

- **Summary:** Conclude the analysis with the potential utility of news sentiment dispersion in predicting stock price and its broader applications in finance.
- **Future Research:** Suggest areas for further exploration, such as integrating additional data sources, refining LSTM models, or extending the analysis to different sectors.

Late Work Policy

Late work (e.g., submitting a deliverable after the established deadline) will earn a grade of zero (though feedback will still be provided) unless it is associated with an excused absence and/or is prearranged with the instructor ahead of time. Work submitted by a student as makeup work for an excused absence is not considered late work and is exempted from the late work policy (Student Rule 7).

Attendance Policy

The university views class attendance and participation as an individual student responsibility. Students are expected to attend class and to complete all assignments.

Please refer to Student Rule 7 in its entirety for information about excused absences, including definitions, and related documentation and timelines.

University Policies

Makeup Work Policy

Students will be excused from attending class on the day of a graded activity or when attendance contributes to a student's grade, for the reasons stated in Student Rule 7, or other reason deemed appropriate by the instructor.

Please refer to Student Rule 7 in its entirety for information about makeup work, including definitions, and related documentation and timelines.

Absences related to Title IX of the Education Amendments of 1972 may necessitate a period of more than 30 days for make-up work, and the timeframe for make-up work should be agreed upon by the student and instructor" (Student Rule 7, Section 7.4.1).

"The instructor is under no obligation to provide an opportunity for the student to make up work missed because of an unexcused absence" (Student Rule 7, Section 7.4.2).

Students who request an excused absence are expected to uphold the Aggie Honor Code and Student Conduct Code. (See Student Rule 24.)

Academic Integrity Statement and Policy

"An Aggie does not lie, cheat or steal, or tolerate those who do."

"Texas A&M University students are responsible for authenticating all work submitted to an instructor. If asked, students must be able to produce proof that the item submitted is indeed the work of that student. Students must keep appropriate records at all times. The inability to authenticate one's work, should the instructor request it, may be sufficient grounds to initiate an academic misconduct case" (Section 20.1.2.3, Student Rule 20).

Texas A&M at College Station
You can learn more about the Aggie Honor System Office Rules and Procedures, academic integrity, and your rights and responsibilities at aggiehonor.tamu.edu.

Americans with Disabilities Act (ADA) Policy

Texas A&M University is committed to providing equitable access to learning opportunities for all students. If you experience barriers to your education due to a disability or think you may have a disability, please contact the Disability Resources office on your campus (resources listed below) Disabilities may include, but are not limited to attentional, learning, mental health, sensory, physical, or chronic health conditions. All students are encouraged to discuss their disability related needs with Disability Resources and their instructors as soon as possible.

Texas A&M at College Station
Disability Resources is located in the Student Services Building or at (979) 845-1637 or visit disability.tamu.edu.

Title IX and Statement on Limits to Confidentiality

Texas A&M University is committed to fostering a learning environment that is safe and productive for all. University policies and federal and state laws prohibit gender-based discrimination and sexual harassment, including sexual assault, sexual exploitation, domestic violence, dating violence, and stalking.

With the exception of some medical and mental health providers, all university employees (including full and part-time faculty, staff, paid graduate assistants, student workers, etc.) are Mandatory Reporters and must report to the Title IX Office if the employee experiences, observes, or becomes aware of an incident that meets the following conditions (see University Rule 08.01.01.M1):

- The incident is reasonably believed to be discrimination or harassment.
- The incident is alleged to have been committed by or against a person who, at the time of the incident, was (1) a student enrolled at the University or (2) an employee of the University.

Mandatory Reporters must file a report regardless of how the information comes to their attention – including but not limited to face-to-face conversations, a written class assignment or paper, class discussion, email, text, or social media post. Although Mandatory Reporters must

file a report, in most instances, a person who is subjected to the alleged conduct will be able to control how the report is handled, including whether or not to pursue a formal investigation. The University's goal is to make sure you are aware of the range of options available to you and to ensure access to the resources you need.

Texas A&M at College Station
Students wishing to discuss concerns in a confidential setting are encouraged to make an appointment with Counseling and Psychological Services (CAPS).

Students can learn more about filing a report, accessing supportive resources, and navigating the Title IX investigation and resolution process on the University's Title IX webpage.

Statement on Mental Health and Wellness

Texas A&M University recognizes that mental health and wellness are critical factors that influence a student's academic success and overall wellbeing. Students are encouraged to engage in healthy self-care by utilizing available resources and services on your campus

Texas A&M College Station
Students who need someone to talk to can contact Counseling & Psychological Services (CAPS) or call the TAMU Helpline (979-845-2700) from 4:00 p.m. to 8:00 a.m. weekdays and 24 hours on weekends. 24-hour emergency help is also available through the 988 Suicide & Crisis Lifeline (988) or at 988lifeline.org Links to an external site..