Introduction to Financial Data Science

Instructor: Poojan Patel
Email: Poojan.patel001@umb.edu

Pre-requisites: No Prerequisites

Course Description:
This course, "Introduction to Data Science for Finance," is designed to give students a comprehensive overview of how data science is applied within the financial sector. Students will learn to handle and analyze financial data, explore market instruments, acquire and clean data, and apply statistical and machine learning methods for insights. The course covers core topics like time series analysis and sentiment analysis, blending theory with hands-on applications.

Course Objective:

By the end of this course, students will be able to:
1. **Understand Financial Data Basics**: Learn the types and importance of financial data, including key sources like APIs, databases, and web scraping.
2. **Learn Data Preparation**: Acquire and preprocess financial data, focusing on cleaning techniques and handling missing values to ensure accuracy in analysis.
3. **Conduct Exploratory Data Analysis (EDA)**: Develop skills to summarize and visualize data, uncovering patterns and insights relevant to finance.
4. **Apply Statistical and Time Series Analysis**: Use statistical methods and time series analysis to interpret financial data, supporting data-driven decision-making.
5. **Explore Machine Learning Applications**: Gain an understanding of machine learning tools for finance, emphasizing practical applications.
6. **Leverage Alternative Data and Sentiment Analysis**: Integrate non-traditional data sources, including sentiment analysis, to enhance financial insights.
7. **Demonstrate Proficiency Through Projects**: Apply course concepts in a mid-term presentation and a final project, building real-world financial data science skills.

Grading:

Class Presentation	25%
Idea Presentation	10%
Midterm Presentation	20%
Final Project Report and Presentation	40%
Attendance	5%
	100%

Class Presentation - There will be group presentations throughout the semester, with groups assigned by the instructor. Each group will be given a topic covered in class and will be tasked with creating a presentation on the application of the techniques taught in class in the field of Finance.

• Objective: Students should demonstrate an understanding of core concepts and showcase applications relevant to financial data science.

• Requirements:

- **Content**:
    - Overview of the topic, real-world applications, and any relevant techniques or tools.
    - Include a case study, recent news, or a dataset example to illustrate the concept

Introduction of the concept	**30**
Explanation of the technique or tool	**30**
Practical example or demo	**40**
	100

Final Project – Students will develop a comprehensive data analysis project that showcases their ability to apply data science techniques within a financial context. The project can focus on exploring a dataset, modeling financial trends, or conducting a detailed analysis of a specific financial topic. Students will collect and prepare data, apply various analytical methods, and present their findings. **Progress will be tracked through a midterm presentation update**, with a final presentation and report at the end of the semester.
Here are some guiding questions to help students develop and present their project:
1. **What financial area or topic are you analyzing?**
   Describe the focus of your project, whether it's an exploration of a financial market, an asset class, or a type of financial data. Explain why this area is interesting or important for analysis.
2. **What data have you chosen, and how did you prepare it?**
   Outline your data sources (e.g., APIs, financial databases) and describe any preprocessing steps, including cleaning, handling missing values, or formatting.
3. **What analytical methods are you using?**
   Discuss the data science techniques you're applying, such as exploratory data analysis, statistical methods, time series analysis, or machine learning. Explain why these methods are appropriate for your topic.
4. **What patterns, trends, or insights did you uncover?**
   Summarize the key findings from your analysis, such as trends in the data, relationships

between variables, or unexpected patterns. Highlight how these insights could add value or enhance understanding of your topic.

5. **What challenges did you face, and how did you address them?**
   Describe any obstacles encountered during data collection, processing, or analysis. Share how you overcame these issues and any limitations that may impact your findings.

6. **What are the practical applications or implications of your project?**
   Reflect on how your analysis could be used in the real world, such as informing investment strategies, identifying market trends, or enhancing risk assessment.

This project allows students to apply data science in a meaningful financial context, demonstrating their ability to handle data, apply analytics, and draw valuable insights.

Presentation & Report Grading Rubrics (Midterm & Final)

Depth of Research	**20**
Potential outlook	**10**
Applications	**30**
Insights	**20**
Presentation Skills & Formatting	**10**
Q&A	**10**
	100

Use these questions to frame your pitch and ensure that you present a clear, scalable, and cost-effective solution with a strong revenue potential.

Reference Books

1. Charu C. Aggarwal, Data Classification: Algorithms and Applications. CRC Press, 2015. (ISBN: 978-1-4665-8674-1)
2. Charu C. Aggarwal, Data Mining. Springer, 2015. (ISBN: 978-3-319-14141-8)
3. Deborah Nolan and Duncan T. Lang, Data Science in R: A Case Studies Approach to Computational Reasoning and Problem Solving, CRC Press, 2015. (ISBN: 978-1-4822-3481-7)
4. Norman Matloff, The Art of R Programming, No Starch Press, 2011. (ISBN: 978-1-59327-384-2)

Class Schedule

Weeks		Presentation Schedule
	Introduction to Data Science for Finance	
1	**Introduction to Financial Data Science** Overview Importance of Data Types of Data	
2	**Financial Markets and Instrumetns** Basics of Financial Markets Types of financial instruments (stocks, bonds, derivatives) Market participants and their roles	
3	**Data Acquisition and Preprocessing** Data sources: APIs, web scraping, financial databases Data cleaning techniques Handling missing values and outliers	Group 1
4	**Data Acquisition and Preprocessing**	Group 2
5	**Exploratory Data Analysis**	**Idea Presentation**
6	**Statistical Methods in Finance**	Group 3
7	**Statistical Methods in finance**	Group 4

8	**Mid Term Presentation**	**Mid term Presentation**
9	**Time Series Analysis**	Group 5
10	**Machine Learning Basics**	Group 6
11	**Machine Learning Basics**	Group 7
12	**Alternative Data and Sentiment Analysis**	Group 8
13	**Art of Story Telling**	
14	**Final Project**	**Final Report**