MSA 8200
Predictive Analytics
Sample Syllabus

1 Instructor Information

- **Instructor:** Dr. Mohammad "**Javad**" Feizollahi

- **Office:** Room No. 331, Buckhead Center

  Room No. 1626, 55 Park Place (Downtown)

- **Office Hours:** Mondays 5:00 PM - 6:00 PM

  Tuesdays 12:00 PM-1:00 PM

  Due to the COVID-19 situation, the classes and office hours will be held virtually via Webex. Students need to make appointments beforehand via email. You only need to send an email to the instructor and specify your availability during one of these office hours to get an appointment set up.

- **Webex:** https://gsumeetings.webex.com/meet/mfeizollahi

- **Email:** mfeizollahi@gsu.edu

2 Teaching Assistant Information

- **TA:** Minghan Sun

- **Email:** (msun6@student.gsu.edu)

- **Webex:** https://gsumeetings.webex.com/meet/msun6

- **Problem Solving Labs by TA:**

Option	Day	Time
1	Fridays	12:00-1:00 PM
2	Fridays	6:00-7:00 PM

3 Class Information

- **Location:** Both sections of the class will be online until the midterm exam (2/24).

- **Time:** Wednesdays 12:00-2:30PM (Section 1) 6:00-8:30PM (Section 2)

- **Website:** iCollege – `http://icollege.gsu.edu`

  - Please enable notifications for this course to receive announcements, updates, etc.
  - To connect to the online class sessions, in the iCollege page of the course, click on the "Webex" tab. Then, from the "virtual meetings" tab click on the "Join" button.

- **Prerequisite:** MSA 8190, or consent of the instructor.

- **Software:** You will need to be familiar with statistical programming environments. We will rely on R and RStudio for the coursework.

  You can download and install R for free from the following website:

  $$\texttt{http://www.cran.r-project.org}$$

4 Course Description

This course introduces students to different predictive models with a focus on real-world applications and datasets. The course covers three primary topics: 1) the analyses of time series data, including estimation and inference for ARIMA models as well as more recent developments; 2) a collection of other important models including survival analyses, models for panel data; 3) the set of skills required to analysis real world data, including data pre-processing, and data type identification. In addition, the students will also have hand on experience of working with real world data.

5 Course Outcomes

By the end of the semester you will be able to:

1. Understand different data types and identify the correct model for time series data analysis;

2. Model and forecast time series data;

3. Deal with messy data that has missing values or censored observations.

4. Analyze real world data sets and provide business insights.

6 Lecture Notes and Textbooks

Lecture notes will be posted on iCollege. Nevertheless, it is strongly advised that you take notes during lecture as there may be ideas presented in the class which are not included in the posted notes. Useful references will be given for certain concepts and further reading. Handouts as well as computer programs will also be posted there. As a general approach, I first discuss models and methods conceptually, and I then provide and discuss a variety of example problems/programs that illustrate the concepts.

6.1 Main Reference:

[S] Shumway, R.H., and Stoffer D.S., "Time Series Analysis and Its Applications." 4th Edition. Springer, 2017.

(The book and relevant material are freely available at `https://www.stat.pitt.edu/stoffer/tsa4/`)

[K] Kleiber, C., and Zeileis, A., "Applied Econometrics with R." Springer.

[D] Kleinbaum, D.G., and Klein M., "Survival Analysis: A Self-Learning Text." Third Edition. Springer.

6.2 Other Useful References:

[R] `https://www.econometrics-with-r.org/`

[E] The elements of statistical learning: `https://web.stanford.edu/~hastie/Papers/ESLII.pdf`

7 COVID-19 Updates and Guidelines

GSU updates and guidelines related to the COVID-19 (such as class format, face covering, social distancing, etc) can be found in `https://ahead.gsu.edu`. All students are required to read these guideline and abide them.

8 Attendance Policy

It is strongly suggested that students do not miss the classes as most students will have difficulties completing the assignments without attending the lectures. So, students need to attend the classes. In case of emergency, students need to inform the instructor before the class that he/she will not be able to attend that session.

9   Grading

Percentages of course works in students' final scores are as follow:

Course Works	Percentages
Homework (3 sets of problems, each 10%)	30%
Midterm Exam	35%
Final Project	35%

The final grade conversion is based on the following table:

A+	A	A-	B+	B	B-	C+	C	C-	D	F
≥ 97	≥ 93	≥ 90	≥ 87	≥ 83	≥ 80	≥ 77	≥ 73	≥ 70	≥ 60	<50

10   Homework, Exam, and Final Project

- You should start working on each homework early, that way you will have time to ask questions in class before the homework is due.

- Late homework will be accepted only in case of unavoidable occurrences, such as illness or death in the family.

- You are encouraged to discuss homework and learn from each other, but each person must submit his/her own work , unless the homework specifically indicates that you should work in groups.

- Students copying from their classmates or from previous years' assignments, or from any source will receive a zero score. In addition, the student who let others copy from her/his assignment will receive a zero score. There are no exceptions to this rule! Further consequences are possible.

- The exam is open book and open note. But, you are not allowed to get help from any source including the students in your class, other people, internet, etc. Any cheating and violation of the honor code will result a grade of F for the course.

- For HW, exam and the project, you will need to do all computations in R and submit your R file. Any explanation can be added as comments in R. Make sure to write the final answer as a comment by using # sign.

- When you are done with your solutions, you need to record a 10-minute video presenting your solution and post the link of recording in the comments your submission. The instructions for the recording will be posted as an announcement in iCollege.

- During the exam time, you are not allowed to communicate (in anyway) with anyone regarding the exam problems. Any relevant question should be posted on Piazza and only the instructor is allowed to respond to these questions (during the exam time).

- Any queries on homework, exam, or project grades must be submitted in writing via email to the instructor, together with the documents and evidences.

11 Academic Honor Code

All course participants (myself, teaching assistant, and students) are expected and required to abide by the Georgia State University Honor Code. See the University's policy on Academic Honesty (Section 409, http://www2.gsu.edu/ wwwfhb/sec409.html) for details. Please familiarize yourself with the code, and use it to guide your conduct. Specifically, you must do your own work in all homework (unless the homework is specifically designated as a group homework), quizzes and exams. Any form of academic dishonesty, such as plagiarism, can result in a serious deduction from your final grade or even a grade of F in the course.

12 Tentative Lectures Schedule

Date	Topic	References	Readings (links)	Due
01/13	Time series models; ACF; Stationary process	[S] Ch 1-3	A comprehensive beginner's guide to create a Time Series Forecast (with Codes in Python and R)	
01/20	AR models	[S] Ch 3	A Complete Tutorial on Time Series Modeling in R	
01/27	MA model; PACF	[S] Ch 3	Time Series Analysis with R	HW1
02/03	ARIMA model; Alternative ways?	[S] Ch 3	Stock Prices Prediction Using Machine Learning and Deep Learning Techniques (with Python codes)	HW2
02/10	Regression with time series data; Case Study: Sales prediction; Predicting arrivals: Survival analysis	[K] Ch 3.5 [D] Ch 2-3	Walmart Store Sales Forecast; Introduction to Survival Analysis ; Survival Analysis in R	
02/17	Bayesian analysis; Panel data; Fixed effect model; Random effect model	[R] Ch 6,10 [E] Ch 8.3		HW3
02/24	Midterm Exam			
03/03	Data Cleaning and Data exploration		An introduction to data cleaning with R	
03/10	Feature selection and dimension reduction			
03/17	Spring break (No class)			
03/24	Incorporating text analytics into insights		Text Analysis	
03/31	Incorporating image analytics into insights			
04/07	Model selection			
04/14	Effective communication of business insights			
04/21	Final presentation			

Note that this syllabus provides a general plan for the course; deviations may be necessary.