Georgia State University
J. Mack Robinson College of Business
Master of Science in Finance

FI 8090 – Financial Data Analytics
CRN: 90736

Course Syllabus
Fall Semester 2021

Instructor: Rasha Ashraf

Contact Information:

Professor: Rasha Ashraf

Office: 1201 RCB

Email: rashraf@gsu.edu

Tel: 404-413-7348

Class Meetings:

RCB Buckhead Executive Education Center, Room

Time: Wednesday (weekly), 5:30 – 9:45 pm

Dates: 08/18, 08/25, 09/01, 09/08, 09/15, 09/22, 09/29, 10/06

Office Hours: By appointment

Course Description:

This course provides the foundation for financial data analytics. Students will develop knowledge of data analytics modeling aimed at addressing a range of financial questions in practice. The objective of this course is to gain experience in analyzing financial data using modern machine learning techniques, statistical methods, and prediction models. Students will perform data analysis using programming environment Python.

The topics discussed include an introduction to Python, visualization of financial data, simple and multiple linear regression, classification models, high dimension data analysis using Lasso, tree based regression, model assessment and selection using cross validation and textual analysis. Students will have hands-on experience on the development of data analytics applications to analyze real world financial problems.

Course Co-requisite:
1. MBA 8135

Required Materials:

1. An Introduction to Statistical Learning with Applications on R, by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, *Springer Texts in Statistics*
   URL: http://www-bcf.usc.edu/~gareth/ISL/

2. Introduction to Python

Download Python - Anaconda

URL: https://www.anaconda.com/products/individual

Some useful links:

Learn Python: https://realpython.com/

https://www.coursera.org/specializations/data-science-python

Machine Learning:
https://developers.google.com/machine-learning/crash-course

Course Objectives:
The course has three broad objectives: to develop an understanding of modern machine learning and data analytics methods; to introduce the Python programming language; and to apply machine learning techniques for financial data analysis using Python programming language.

Students will gain hands-on experience using the data analytics methods discussed in the course for analyzing real world financial data. Upon completion of the course students will have the ability to perform the following tasks:

1. Understand and use the Python language: Commands and Syntax, Packages, Data types, Objects, Graphics and Visualization
2. Download financial data directly from some open sources, such as Yahoo Finance.
3. Visualize financial data: Construct histograms, bar chart.
4. Compute statistical measures including Mean, Median, Variance, Standard Deviation.
5. Understand the difference between prediction and inference, parametric methods vs. non-parametric methods.
6. Understand difference between regression versus classification methods
7. Construct simple and multiple linear regression models: estimate coefficients, model fit, predictions, assessing the accuracy of model, potential problems.
8. Perform classification models: Logistic Regression.
9. Model Fit and Testing: Training and Test Data
10. Use Shrinkage Method LASSO to deal with high dimensionality of data.
11. Use Tree Based Methods: Classification Trees and Regression Trees.
12. Model assessment and selection using cross validation.
13. Textual Analysis

Additional Class Materials:
Class lectures, reading assignments, instructor's notes, data related to projects and other related materials will be available at iCollege. Students are responsible for reading the materials that are required for each class.

The course uses a building block design. Each day we will learn new techniques that will be useful for future classes as well as the current class. Make sure that you have studied all designated readings prior to class. Success in the course depends on being fully prepared for each class.

Communication:

There will be times during the term that I will want to communicate with you either individually or as a class. I will do this via your GSU email. I will be using the list of e-mail addresses as indicated on the Gosolar class roster.

Grading:
Your grade will be determined by the following.

Assignments	20 %
Quizzes	30%
Midterm Exam	25%
Final Project	25 %

Department of Finance Grading Policy
The Department of Finance provides its instructional faculty with the following guidance.

For masters level courses (MBA and FI prefixed), it is normally expected that no more than 35 percent of the students in a given class will receive a grade of A (including A-). The majority of the remaining students are expected to receive grades of B. Those students demonstrating significantly lagging performance shall earn grades of C or lower as appropriate. The above referenced grade levels (i.e., A, B, and C) are intended to include the +/− variations. Deviations from the above guidelines may occur based on the instructor's professional assessment of overall class performance. The grade of A+ will be assigned only in rare cases of exceptional performance. No A+ grades are expected in most semesters. In fairness to all students, there will be no extra credit assignments.

Grades will be determined relative to class performance and according to the guidelines provided above.

Class Attendance
Students are expected to attend the class regularly (in-class and virtually whenever scheduled), be on time and stay in the class for the entire duration of the class. An attendance will be taken by the professor by your login to the class.

Any student who is absent for more than one class during the semester, including the first day of class, will be withdrawn from the course (this means you can miss only one class to prevent from being withdrawn). You will not receive any formal update on your attendance record so you should keep track of your absences. If you are not sure of your record, email me and ask for the number of absences. Students who leave class early will be considered absent even if they have signed in earlier in the class period. Students arriving late will be considered present only at the discretion of the professor. Normally, permission will not be granted.

Homework Assignments
Students will be assigned 4 homework assignments. The requirements of the assignments will be discussed during the class. The assignments pertain to the development of Python programs related to topics discussed in the class and will involve analyzing financial data using various data analytics methods taught in class. The assignment submission due date will be provided with the assignment. No late submission is accepted and failure to submit the assignments on time will result a score of zero.

Assignments will be performed by groups. For each group assignment, each group should submit the Python program electronically in iCollege.

You are not allowed to assist other student in providing explanations related to the conceptual development of models and/or providing clarifications related to the analysis or programs with other group members. Any student violating this policy will receive the appropriate University disciplinary action.

There are several sections in each assignments. Each team members must actively participate in writing part of the code of a particular section and provide the name of the actual contributor. Other group members can validate the code to make sure that part is working. Team members must decide how to distribute the code assignments across the team members. Group members must NOT provide any help to members of the team for quizzes related to the assignments or any other quizzes and exam.
Team assignments will be done by the professor. Be respectful to all the team members as you work. Maintain professionalism at all times.

Final Project

Students will complete a group final project at the end of the semester. Each group will identify a decision problem related to finance and perform an analysis using the Python programming environment. Each group must receive prior approval of the project topic from the professor. The last date to receive approval is indicated in the course outline. The final project is due at the beginning of the last class. The final project is group assignment and each group is expected to work on it completely independently without any help from others. Each group need to submit Python codes and program output related to the project electronically. Additionally, each group will present the final project in front of the class on the day of last class.

Working in groups:

Submit your group name (your start-up company) by the end of the first class. You are responsible to coordinate with your group members regarding how you will accomplish tasks of the group assignments and final project. You will be required to submit an evaluation of your group members, indicating the extent of their contribution along with your own involvement. You should maintain professionalism in working with others at all times.

Data files:

Many of the assignments require downloading data from online financial websites. For other assignments data are provided at the GSU iCollege.

Academic honesty:

It is your responsibility to read the University Policy on Academic Honesty. It can be found in "GSU Student Code of Conduct" (http://codeofconduct.gsu.edu/). This document establishes clear standards for academic honesty and it will be enforced in this course. My policy is completely consistent with that document.

Anyone violating the academic honesty standards will receive the appropriate University disciplinary action. If at any time you are not sure whether something is a violation of the academic honesty policy please contact me for clarification.

To give you an idea of the types of actions that I consider to be violations of academic honesty, here is a non-exhaustive list.

- You may not look at another student/group's project solution, copy from another student/group's R program or output, communicate in any way with others about the solution methods, or get help from people not in the class. A person who allows someone else to look, copy, or share information is equally guilty of a violation.
- You may not use other people's work that are available online and claim that as yours. If you do use any online resources for coding purposes, you have to provide list of references.
- You may not sign someone else's name to the attendance sheet.

You will be required to sign and submit a Statement of Academic Honesty for all submitted projects.

Disability Services:
Students who wish to request accommodation for a disability may do so by registering with the Office of Disability Services in Suite 230 or the Student Center. Students may only be accommodated upon issuance by the Office of Disability Services of a signed Accommodation Plan and are responsible for providing a copy of that plan to instructors of all classes in which an accommodation is sought.

GSU Policy Prohibiting Students from Posting Instructor-Generated Materials on External Sites
The selling, sharing, publishing, presenting, or distributing of instructor-prepared course lecture notes, videos, audio recordings, or any other instructor-produced materials from any course for any commercial purpose is strictly prohibited unless explicit written permission is granted in advance by the course instructor. This includes posting any materials on websites such as Chegg, Course Hero, OneClass, Stuvia, StuDocu and other similar sites. Unauthorized sale or commercial distribution of such material is a violation of the instructor's intellectual property and the privacy rights of students attending the class, and is prohibited.

Courtesy Rules
1. Turn off your cell phones.
2. Be respectful to others
3. Do not engage in any discussions non-related to the course during the duration of the class.
4. Students in this class are encouraged to speak up and participate during class meetings and online class discussions. The students on our class represent a diversity of individual beliefs, backgrounds, and experiences, and therefore, every member of this class must show respect for every other member of this class.

Student Evaluation of Instructor:
Your constructive assessment of this course plays an indispensable role in shaping education at Georgia State. Upon completing the course, please take the time to fill out the online course evaluation.

Course Outline:
This course outline provides a general plan for the course; deviation may be necessary. Any schedule change will be announced in class or transmitted via email, and you are responsible for the announced change regardless of whether you are in class at the time of the announcement or have read your email. This includes announcements made the first night of class.

Class Day	Topics

Week 1
 	
(08/18)
	Introduction to Python and Financial Data
	
	• Installation of Python Anaconda
	• Python Basics
	• Data manipulation
	
	Financial Data Visualization
	
	Data Visualization
	    • Downloading financial data directly from some open sources.
	    • Time plots of financial data
	Distributional Properties of Returns
	    • Mean, Variance, Standard Deviation, One Sample t-test
	    • Histogram of daily simple returns
	
	Introduction to Data Analytics

	- What is statistical learning? - Difference between prediction and inference - Parametric methods vs. Non-parametric methods. - Supervised versus unsupervised learning **Read: :** JWHT Chapter 2, Instructors' notes and supplemental material **Homework 1 assigned: Group Submission**
Week 2 **(08/25)**	**Liner Regression Part 1: Simple Regression** - Simple Linear Regression: Hypothesis testing, Estimating the coefficients, assessing the accuracy of the model - Potential problems: non-linearity of the data, outliers, collinearity, and correlation of error terms. **Application:** The Marketing Plan: Predicting future sales. **Read:** JWHT Chapter 3 and Instructor's notes
Week 3 **(09/01)**	**Liner Regression Part 2: Multiple Regression** - Multiple Linear Regression: Deciding on important variables, model fit and predictions - Qualitative predictors **Application:** Analyze credit card balance based on customer history and attributes **Read:** JWHT Chapter 3 and Instructor's notes **Homework 2 assigned: Group Submission** **Submit initial proposal of the Company Sponsored Project**
Week 4 **(09/08)**	**Classification Model** - An Overview of Classification models. - Logistic Regression - Model Accuracy - Model Fit and Testing: Training vs. Test Data **Application:** Fit a logistic regression model to predict market direction (up or down) using lag return data. **Read:** JWHT Chapter 4 and Instructors' notes **Homework 3 assigned: Group Submission** **Meet with Professor/Industry PM to Finalize the Company Sponsored Project Scope** **Submit initial proposal for the final data analytics project topic.**
Week 5 **(09/15)**	**Exam: On topics discussed till week 4** **Class Lecture** **Models for High Dimension Data** Shrinkage methods - LASSO

	Model Assessment and Selection • Cross validation • K-fold Cross Validation Read: JWHT Chapter 5 and Instructor's notes **Application:** Prediction of baseball player salary using LASSO model **Read:** JWHT Chapter 6 and Instructor's notes **Homework 4 assigned : Group Submission**
Week 6 **(09/22)**	**Tree Based Methods** • The Basics of Decision Trees: Regression and Classification Trees **Application:** Prediction of baseball player salary using Decision Trees **Read:** JWHT Chapter 8 and Instructors' notes **Last date to get approval for the final project topic.**
Week 7 **(09/29)**	Web Scraping and Textual Analysis: Read: Scraping EDGAR with Python by Rasha Ashraf
Week 8 **(10/06)**	**Final Project Submission and Presentation: Group Assignment**