Big Data for Finance
Business 9334
Spring 2023

Overview
Instructors
- Simona Abis
- Kirsten Burr
- Harry Mamaysky (coordinator)
- Ciamac Moallemi
- Tomek Piskorski
- Razvan Popescu

Room	Lecture hours
Kravis 680	Fridays, 8:30 – 11:45am

TAs	Office hours
Daheng Yang	TBD
Meha Sadasivam	TBD

Course Description
This course provides an introduction to financial data, data analysis tools and approaches, and analyzing statistical models using output from these datasets. The course schedule below lists the data sets that will be covered.

Course Requirements
The course will have extensive weekly homework assignments and a take-home final exam.

Course Prerequisites
This course requires three PhD courses: B9320, Econometrics and Statistical Inference (our first term PhD course); B9122: Computing for Business Research; and B9302 Finance Theory I. We assume students are familiar with the finance theories and econometric techniques used in these courses.

In terms of programming background, we expect proficiency in Python as well as knowledge of SQL (both covered in Fall Term "Computing").

Assignment and Grading

Problem Sets:

i.   There are weekly problem sets during the semester, each will be graded on a scale of 0-10 points. The lowest one will be dropped from your grade.

ii. Students are allowed/encouraged to work in groups, but <u>each student must do their own analysis and</u> <u>submit their own original work, which includes computer code</u>. Every student is required to do their own coding (you won't learn to code unless you actually code). It is, however, useful to discuss issues with others so we encourage students to work in groups with each other. Turning in another student's code (even if you change all the variable and function names) is a violation of the honor code.

iii. Homework should be turned in at the beginning of class on the day it is due. Late homework is not be accepted under *any* circumstances. Since your lowest homework is dropped, you have a cushion to work with.

Exams: There is a take home final, details TBA.

Course Grade: The grade is 40% final and 60% homework.

Course Schedule

Date	Instructor	Topic
1/27	Razvan Popescu	Grid and parallel computing
2/03	''	''
2/10	Harry Mamaysky	Thomson-Reuters and natural language processing
2/17	''	''
2/24	Simona Abis	Mutual fund data
3/03	''	''
3/10	No class	Exam period (this class has no midterm)
3/17	No class	Spring break
3/24	Kirsten Burr	CRSP and Compustat and Fama-French
3/31	''	''
4/07	''	''
4/14	Tomek Piskorski	Real estate data: Residential real estate
4/21	''	Real estate data: Commercial real estate
4/28	Ciamac Moallemi	Cryptocurrency data