Special Topics: Data Analytics in Finance  
FIN 579 – Spring 2020  
Prof. Thomas Gilbert

The goal for this class is to learn new technical data analytics skills in three areas of business: finance, microeconomics, and macroeconomics. The focus will be on time-series models, causal experiments, and textual analysis. Time permitting, we will also touch on other topics such as clustering, principal component analysis, classification, and nonparametrics.

Class time: Tuesdays and Thursdays, 1.30pm to 3.20pm
Zoom: https://washington.zoom.us/j/262028014 (meeting ID 262-028-014)
Office hours (by appointment): gilbertt@uw.edu, 206-616-7184, or Zoom
TA: Anna Kruglova, kruglova@uw.edu
UW Foster Research Consultant: Jonathan Whittinghill, jwhitt@uw.edu
Class website: Canvas. All materials, submissions, and class discussions should be done electronically on Canvas.
Required software: We will be using R in this class and I will be coding in R via R Studio. Both are free and must be downloaded in installed prior to the first class. https://www.r-project.org/ and https://rstudio.com/products/rstudio/download/ (desktop version).
Individual homework: There will be eight weekly homework assignments, to be completed on your own. You must hand in your write-up and code electronically via Canvas on the due date.

Group project: The final assignment of this course will consist of a group project (nine teams of five students, to be entered on Canvas). Your goal is to apply the techniques learned in class to a new dataset with a novel question. There are multiple deliverables for this project:
1) Three progress reports, due 4/21, 5/5, 5/19
2) A 15-20 minute presentation to the class on 6/2 or 6/4 (randomly assigned)
3) A 5-10 page report, due 6/2 (with code and dataset)
4) A peer evaluation of your group members, due 6/4.

Grading: 50% on individual homework (equally weighted, lowest one dropped) and 50% on the group project.

Foster code of conduct: By being a student in this course, you acknowledge that you are a part of a learning community at the Foster School of Business that is committed to the highest academic standards. As a part of this community, you pledge to uphold the fundamental standards of honesty, respect, and integrity, and accept the responsibility to encourage others to adhere to these standards.

Religious accommodations: Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW’s policy, including more information about how to request an accommodation, is available at Religious Accommodations Policy. Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request form.
Class schedule and topics (subject to revisions)

Class 1: Tuesday 31 March
- Topics: uncertainty, prices and returns, plots, histograms, random walk versus stationarity, average versus expected returns, bootstrapping standard errors
- Dataset: monthly S&P500 Index

Class 2: Thursday 2 April
- Topics: portfolio diversification (2, 3, N stocks), exact mean-variance frontier, regression, CAPM, beta
- Dataset: monthly stock returns

Class 3: Tuesday 7 April
- Topics: performance of mutual funds, regression, multi-factor models, alpha, persistence
- Dataset: monthly mutual fund returns

Class 4: Thursday 9 April
- Topics: trading strategy
- Dataset: monthly stock returns and annual accounting statements
- Due: individual homework #1

Class 5: Tuesday 14 April
- Topics: forecasting recessions
- Dataset: term structure of interest rates and GDP

Class 6: Thursday 16 April
- Topics: experiments, randomized controlled trials, AB testing
- Dataset: Medicaid eligibility experiment in Oregon
- Due: individual homework #2

Class 7: Tuesday 21 April
- Topics: experiments, difference-in-difference
- Dataset: tbd
- Due: group project progress report #1

Class 8: Thursday 23 April
- Topics: experiments, regression discontinuity
- Dataset: tbd
  - Due: individual homework #3

**Class 9: Tuesday 28 April**
- Topics: experiments, endogeneity, instrumental variables
- Dataset: tbd

**Class 10: Thursday 30 April**
- Guest lecture: Kyle Kretschman, Senior Manager and Economist at Amazon Consumer Finance
  - Due: individual homework #4

**Class 11: Tuesday 5 May**
- Topics: text as data, web scraping (by Jonathan Whittinghill, UW Foster Research Consultant)
- Dataset: SEC Edgar, FOMC
  - Due: group project progress report #2

**Class 12: Thursday 7 May**
- Topics: text as data, pre-processing, textual analysis (by Jonathan Whittinghill, UW Foster Research Consultant)
- Dataset: SEC Edgar, FOMC
  - Due: individual homework #5

**Class 13: Tuesday 12 May**
- Topics: text as data, sentiment indicator
- Dataset: Dow Jones news

**Class 14: Thursday 14 May**
- Topics: classification, logistic regression, credit risk
- Dataset: loans and credit from a set of local lenders in Germany
  - Due: individual homework #6

**Class 15: Tuesday 19 May**
- Topics: predicting accounting fraud
- Dataset: annual accounting statements
- **Due**: group project progress report #3

**Class 16: Thursday 21 May**
- Topics: nonparametrics, decision trees, random forest
- Dataset: house prices in California
- **Due**: individual homework #7

**Class 17: Tuesday 26 May**
- Topics: factorization, k-means clustering
- Dataset: monthly stock returns

**Class 18: Thursday 28 May**
- Topics: factorization, principal component analysis
- Dataset: term structure of interest rates
- **Due**: individual homework #8

**Class 19: Tuesday 2 June**
- Group project presentations
- **Due**: group project reports, code, dataset

**Class 20: Thursday 4 June**
- Group project presentations