Financial Econometrics can be broadly defined as the area of statistics and econometrics devoted to the analysis of financial data.

Three main topics will be covered in this course:

1. **Linear regression model** and its application to understand the factors driving stock returns and to measure their riskiness.

2. **Time series models** use the past of a variable to forecast its future; one of the applications that we will consider is to forecast quarterly revenue of a company.

3. **Volatility models** are time series model that are used to forecast the variance or standard deviation of a financial variable; we will consider the application of these models to financial risk management.

In addition to learn about methods and models, you will also learn to do data analysis in the R programming language.
Resources

- S. Manzan, *Introduction to Financial Econometrics* [link]
- Class slides
- **WRDS** Wharton Research Data Services
  - In Blackboard is posted a class username and password to access the database
  - NB: you agree to not share these credentials with any other individual outside of the class; the account can accommodate up to 15 simultaneous users
  - Datasets available in WRDS (among others):
    - CRSP
    - Compustat
    - TAQ
I created a group and sent out invitations (who did not get it?)
Weekly assignments to learn R
Currently assigned:

1. **DC1** (02/07): course *Introduction to R*
2. **DC2** (02/09):
   - chapter *Compiling Reports* from course *Reporting with R Markdown*
   - chapter *Embedding Code* from course *Reporting with R Markdown*
   - chapter *Authoring R Markdown Reports* from course *Reporting with R Markdown*
3. **DC3** (02/14): course *Data Visualization with ggplot2 (Part 1)*
4. **DC4** (02/16):
   - chapter *Importing data from flat files with utils* from course *Importing Data in R (Part 1)*
   - chapter *readr & data.table* from course *Importing Data in R (Part 1)*
   - chapter *Importing Excel data* from course *Importing Data in R (Part 1)*
5. **DC5** (02/21): course *Data Manipulation in R with dplyr*
6. **DC6** (02/28): course *Intro to Statistics with R: Multiple Regression*
7. More **DC?** assignments TBA
Supplemental material

- Tsay, *An Introduction to Analysis of Financial Data with R*, Wiley
- R Core Team, *An Introduction to R*
- Albert and Rizzo, *R by Example*, Springer
- Wickham, *ggplot2: Elegant Graphics for Data Analysis*, Springer
- Datacamp course on *Intro to Computational Finance with R* by Eric Zivot, UW
- Coursera course on *R Programming* (part of Data Specialization Specialization)
- **Swirl**: a package to learn R in R
  - Install the package: `install.packages("swirl")`
  - Install a course: `install_from_swirl("Course name")` (details)
  - The courses related to our course are: (S1) *R Programming*, (S2) *Data Analysis*, (S3) *Regression Models*, (S4)

- **CRAN task views**
- **Quick-R**
- **R-Bloggers**
The overall grade is determined as follows:

1. Datacamp assignments (between 5 and 8, 2% each)
2. WRDS assignments (between 4 and 7, 3% each)
3. Group Project (20%)
4. Midterm and Final (non-cumulative and both in computer lab 12-120; weight: 100 minus the sum of 1-4 above divided by 2, i.e. between 21.5% and 29% each)
Apply the concepts discussed in class to a relevant economic and financial problem using R

Group composed of at least 2 and max 4 students

10-page report due on Thursday 05/10/2018 by 10am
Presentations on Sat 05/12/2018 (last class)

Topics:

- I posted in the course webpage a series of articles that could be used as the base for your project on topics ranging from the housing market to momentum trading
- You can choose to develop your project based on one of these articles or find another article on a topic of your interest
<table>
<thead>
<tr>
<th>Date</th>
<th>Day</th>
<th>Topic</th>
<th>CH</th>
<th>ASS</th>
<th>DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/27</td>
<td>Sat</td>
<td>Introduction to the course, R, Rstudio, Rmarkdown, Datacamp, WRDS</td>
<td>CH1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/27</td>
<td>Sat</td>
<td>Getting started with R: Classes, types, structures</td>
<td>CH2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/3</td>
<td>Sat</td>
<td>Reading and writing data</td>
<td></td>
<td></td>
<td>DC1</td>
</tr>
<tr>
<td>2/3</td>
<td>Sat</td>
<td>Plotting data with the base and ggplot2 packages</td>
<td></td>
<td></td>
<td>DC2</td>
</tr>
<tr>
<td>2/10</td>
<td>Sat</td>
<td>Summary statistics</td>
<td></td>
<td></td>
<td>DC3</td>
</tr>
<tr>
<td>2/10</td>
<td>Sat</td>
<td>LAB 12-120 Loops, apply, functions</td>
<td></td>
<td></td>
<td>DC4</td>
</tr>
<tr>
<td>2/17</td>
<td>Sat</td>
<td>Linear regression model Review, estimator, assumptions</td>
<td>CH3</td>
<td>PS1</td>
<td>DC5</td>
</tr>
<tr>
<td>2/17</td>
<td>Sat</td>
<td>LAB 12-120 Nonlinear regression model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/24</td>
<td>Sat</td>
<td>CAPM and Fama-French 3 factor model</td>
<td></td>
<td></td>
<td>PS2</td>
</tr>
<tr>
<td>2/24</td>
<td>Sat</td>
<td>additional topics</td>
<td></td>
<td></td>
<td>DC6</td>
</tr>
<tr>
<td>3/3</td>
<td>Sat</td>
<td>Time series models</td>
<td></td>
<td>PS3</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Day</td>
<td>Topic</td>
<td>CH</td>
<td>ASS</td>
<td>DC</td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>----------------------------------------------------------------------</td>
<td>----</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>3/23</td>
<td>Thu</td>
<td>Random walks with and without drift, testing non-stationarity</td>
<td></td>
<td></td>
<td>WRDS3</td>
</tr>
<tr>
<td>3/28</td>
<td>Tue</td>
<td>Review session</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/30</td>
<td>Thu</td>
<td>MIDTERM (Chapter 1, 2, and 3)</td>
<td></td>
<td>WRDS3</td>
<td>LAB 12-120</td>
</tr>
<tr>
<td>4/4</td>
<td>Tue</td>
<td>Volatility models: intro, MA, EMA</td>
<td>CH4</td>
<td>WRDS4</td>
<td></td>
</tr>
<tr>
<td>4/6</td>
<td>Thu</td>
<td>ARCH and GARCH model, asymmetric models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/25</td>
<td>Tue</td>
<td>forecasting, estimation, R packages</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4/27</td>
<td>Thu</td>
<td>LAB 12-120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/2</td>
<td>Tue</td>
<td>High Frequency Data</td>
<td>CH5</td>
<td>WRDS5</td>
<td></td>
</tr>
<tr>
<td>5/4</td>
<td>Thu</td>
<td>Measuring financial risk</td>
<td>CH6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/9</td>
<td>Tue</td>
<td>Historical simulation, Monte Carlo, FHS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/11</td>
<td>Thu</td>
<td>LAB 12-120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/16</td>
<td>Tue</td>
<td>Portfolio VaR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/18</td>
<td>Thu</td>
<td>Project presentations (report due on 05/15/2017 by 10am)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5/25</td>
<td>Thu</td>
<td>FINAL (chapter 4, 5, and 6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LAB 12-120</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Checklist

1. Datacamp
2. WRDS
3. Rstudio:
   - Importing and viewing data
   - Console
   - Script
   - Rmarkdown

- Some examples of financial data that we will analyze in this course
A long-run look at the S&P 500

Figure 1: Annual Price-to-Earnings and Price-to-Dividends ratio for the Standard and Poors 500 Index starting in 1871.

- Why are market valuations fluctuating so much?
- Are these fluctuations driven by the tendency of economies to have cycles of expansions and recessions?
- What explains the extreme valuations in the late 1990s?
A short-run look at the S&P 500

Figure 2: Daily prices and returns for the Standard and Poor 500 Index.

- What forces determine the boom-bust dynamics of asset prices?
- Why do we have these clusters of calm and turbulent times rather than having “normal” times with returns fluctuating on a constant range?
- Are returns predictable?
Figure 3: Intra-day mid-point between bid and ask price for the USD/JPY FX rate at the 1 minute frequency.

1,320 1-minute quotes for one day from the 31,076 1-minute quotes for December 2016 obtained from 14,237,744 quotes

- Is the bid-ask spread constant over time or subject to fluctuations due to market events?
- Is it possible to use intra-day information to construct measures of volatility?
- Does the size of a trade have an impact on the price?
- Can we predict the direction of the next trade and the price change?
Figure 4: Logarithm of the market capitalization for aNYSE, AMEX, or NASDAQ stock against the percentage return in the following month.

- Do stocks of large (small) capitalization companies outperform in the following month stocks of small (large) caps? Is size a predictor of future returns?
- In addition to size, what other company characteristics can be used to predict future stock performance?
- Are small caps “riskier” relative to large caps?
- Small caps stocks provide (on average) higher returns relative to large cap stocks: what are the factors explaining it? why?
There are several sources of financial data that, in some cases, are publicly available, while in others are subscription-based. In this book we will use publicly available data when possible, and commercial datasets otherwise. A short-list of data providers is:

- [http://www.truefx.com](http://www.truefx.com)
- [https://fred.stlouisfed.org/](https://fred.stlouisfed.org/)
- [http://www.quandl.com](http://www.quandl.com)
- WRDS:
  - CRSP
  - Compustat
  - TAQ