

# **INTD 0318: Quantitative Finance**

## **Spring 2016**

**Professor David Kane**

*dave.kane at gmail.com*

### **Logistics**

Location: Wright Memorial Theatre SEM. 96 Chateau Road. Bring your laptop and headphones to class.

Times: Monday 1:30 -- 4:00

Be [prepared for class](#), including the first. Install the latest version of R ([3.2.3](#)), the latest version of Git ([2.7.0](#)), and the latest release of RStudio ([0.99.878 or later](#)). *R Packages* by Hadley Wickham ([available for free on-line](#)) provides useful background information on using R. Files for the course are [available via MiddFiles](#).

Final projects must be [submitted](#) for permanent public display at the [Digital Collections at Middlebury](#). Do not enroll if you are not comfortable with this requirement.

This is a rigorous and time-consuming course. The normal prerequisites should include knowledge of R and a previous course in finance. Students with such backgrounds can expect to devote 20 hours a week to the class. However, I allow students without these prerequisites to take the class, but only if they acknowledge that, for them, 30 to 40 hours per week is a better estimate.

### **Overview**

This class will introduce students to applied quantitative equity finance. First, we will develop the technical skills needed to do serious research, the most important of which is proficiency with R and RStudio. Second, we will briefly review the history and approach of academic research in equity pricing via selected readings. Students will work as teams to replicate the results of a published academic paper and then extend those results in a non-trivial manner. This course is designed for two types of students: first, those interested in applied financial research, and second, those curious about how that research is used and evaluated by finance professionals. See "[Publication, Publication](#)" for useful background on my approach.

### **Goal**

If you had tried to conduct a similar piece of financial research before taking this class, you would have done X well. Now that you have taken the class, you will do Y well, both with your actual paper and with any future financial research you choose to undertake. The success (or failure) of the class can be measured by comparing Y with X. A second goal of the class is to enable you to get a better internship/job (and to perform better when given that opportunity) then you would have if you had not taken the class.

# Logistics and Prerequisites

This class requires a lot of independent work, especially if you are not familiar with the R programming language. I will assist students and provide a number of learning resources, but most of the discussion in class will be devoted to the replication project.

There will be significant amount of preparation for every class, so that class meetings can be primarily working sessions. Each team/student should bring her laptop and be prepared to work on data analysis during class. I will do very little lecturing. Think about the class like an apprenticeship. If you wanted to be a blacksmith in 15th century France, there was no book to read nor class to take. Your best option was to become the apprentice of the local blacksmith. He taught you what you needed to know by working closely with you. Similarly, there is no (good) book to read about how to conduct applied quantitative equity analysis. The best way to learn is to work closely with someone who already knows how to do it. That is what we will do during class. Each student/team will work on their own project as I move around the room, answering questions and providing suggestions. If a particularly important point comes up, I will stop everyone and give a brief overview to the whole class about the topic.

## Outside Material

There is no reason for us to re-invent the wheel so, to the greatest extent possible, we will make use of teaching material created elsewhere. This is good for three reasons:

1. We can use the best-in-class material. As wonderful (?) as I am as a teacher, I am not the world's best lecturer on R. So, instead of having me lecture about R, we will use the [Johns Hopkins Data Science Specialization](#) from Coursera. You do not have to pay for a verified certificate in taking these classes (although you can if you want). Simply select the "Audit" option.
2. Students can learn at their own pace. If you are a statistics expert who already understands R very well, then it is stupid for you to listen to introductory lectures about R, whether by me or on-line. So, don't. Spend that time doing other stuff. Or, if a particular topic, like factors, is giving you trouble, you can re-watch the on-line lecture as many times as you like, without feeling bad that you are holding up the entire class.
3. We devote class time to things that can't be found or replicated on-line: student/professor interaction about the specific questions that you have, especially as you try to replicate your academic paper and extend its results.

I have been [urging Williams College to get rid of all lectures](#) for almost a decade, so far without success. In this class, at least, I practice what I rant.

## Coursera Data Science Specialization

All students in the class will be required to register for and complete six of the nine courses in the [Coursera Data Science Specialization](#) sequence. You can sign up for a certificate if you like, but doing so a) is not required, b) costs money and c) requires taking an additional three classes. Conveniently, these courses are offered (in parallel) starting in February. Note that, according to Coursera, the combined workload for the courses is around 20 to 30 hours a week. Having sampled the classes, this estimate is not unreasonable, especially for students with no prior exposure to statistics/programming. But, in order to accomplish the aggressive goals for this class, there is no alternative. If you aren't interested in working this hard, don't take my class.

# Reference Material

There is lots of interesting material related to R, data analysis and finance around the web.

- AQR, a major quantitative asset manager, has created a [data library](#) which provides public access to (some of) the data which underlies their academic publications.
- Ken French maintains a similar [data library](#).
- [Course in data visualization](#) by Kieran Healy.
- [Swirl](#) is an R package for learning R within R. (Full disclosure: I am contributor to the project.) It shares some personnel with the Coursera courses discussed above.
- STAT 545 at the University of British Columbia has the most impressive collection of [teaching materials](#) that I have ever seen. Many thanks to Professor Jenny Bryan for making the material open source.
- R materials [here](#).
- **Videos:** This [video](#), from the [How To R](#) series on YouTube, provides step-by-step instructions for installing R and R Studio. The whole series is well-done, if you like video lectures. Another Youtube series is by prolific author Stuar51XT. [This is the start](#) of a five part series on data frames. He has dozens of other R videos, all nicely organized. Other relevant videos include dplyr parts [one](#) and [two](#).
- **R and R Studio:** This [pdf](#) provides a simple introduction. This [EdX course](#) has useful homework assignments. [How to write an R package](#). How to [scrape data from the web using R](#).
- [Beautiful Graphics in R](#): Good class.
- [R Data Class](#): Not as comprehensive as the Coursera classes but still nicely done.
- [R Study Group](#): Somewhat dated (2010) notes on R. Clear and thorough.
- [R Fiddle](#):
- [Git/Github Videos](#):
- [Try R](#): An introductory lesson for R from Code School.
- [Things I wish I had been told](#): Words of wisdom which I try to incorporate into this class.
- [R Statistics.net](#): New to me but interesting.
- Other resources include [Introduction to R](#), [Data Analysis and Statistical Inference](#) and [Introduction to Computational Finance and Financial Econometrics](#) from Data Camp.
- [Rticles](#): For simple (?) formatting of journal articles within R Studio.
- [Abnormal Returns](#) is a daily blog with links to great finance-related articles. Pick five links each day (avoid anything having to do with individual trading or personal finance), read the underlying article, look up words/concepts you don't understand on Wikipedia and, after a year or two, you will know as much about the finance world as the typical first year analyst at a major investment bank.
- [Matt Levine](#) at Bloomberg.

## Books

*Where the Money Grows* by Garet Garrett. ([Available for free online](#).)

*Quantitative Value* by Gray and Carlisle.

*A Random Walk Down Wall Street* by Burton Malkiel. Feel free to use editions prior to the current one (the 11th). Material with advice about personal investing (chapters 12ff) is less relevant to this class.

R Packages by Hadley Wickham. ([Available for free on-line.](#))

## Articles

The articles that we read come in two categories: Those that work with daily/monthly aggregate data and those that use data for individual stocks on a daily basis.

### Daily/Monthly Aggregate Data

We will explore these articles during class.

- "[Betting against beta](#)" by Andrea Frazzini and Lasse Heje Pedersen, 2014. *Journal of Financial Economics*, 111, 1–25.
- "[Quality Minus Junk](#)" by Clifford S. Asness, Andrea Frazzini and Lasse H. Pedersen, 2013. AQR Working Paper.
- "[Value and Momentum Everywhere](#)" by Clifford S. Asness, Tobias Moskowitz and Lasse Heje Pedersen, 2013. *Journal of Finance*, 68, 929-985.
- "[Time Series Momentum](#)" by Tobias J. Moskowitz, Yao Hua Ooi and Lasse Heje Pedersen, 2012. *Journal of Financial Economics*, 104, 228–250.

### Daily Stock Data

All 6 of these articles are important papers in the quantitative equity literature. Although we do not have the time and data resources to *exactly* replicate *all* of their results, we should be able to use the same basic approaches, albeit over a different universe and time period. These articles will be used for the main replication project.

- "[Does the stock market overreact?](#)" by Werner DeBondt and Richard Thaler, 1985. *Journal of Finance*, 40, 793–805.
- "[Evidence of predictable behavior of security returns](#)" by Narasimhan Jegadeesh, 1990. *Journal of Finance*, 45, 881-898.
- "[Returns to buying winners and selling losers: Implications for market efficiency](#)" by Narasimhan Jegadeesh and Sheridan Titman, 1993. *Journal of Finance*, 48, 65–91.
- "[Do industries explain momentum?](#)" by Tobias Moskowitz and Mark Grinblatt, 1999. *Journal of Finance*, 54, 1249–1290.
- "[Price momentum and trading volume](#)" by Charles M.C. Lee and Bhaskaran Swaminathan, 2000. *Journal of Finance*, 55, 2017–2069.
- "[The 52-Week High and Momentum Investing](#)" by Thomas J. George and Chuan-Yang Hwang, 2004. *Journal of Finance*, 59, 2145-2176

If there is a different article that you would prefer to replicate, please let me know.

## Special Projects

Especially adventurous students (who want to devote twice as much time to the project as the typical member of the class) are invited to discuss special projects. In general, these would involve creating an R package for installation on CRAN and writing an associated article for *The R Journal*. Two recent examples that I have been involved with are [here](#) and [here](#). If you are interested in doing a project outside of finance, I am open minded,

especially if it includes data related to Middlebury.

## Schedule

### Week 1: February 15

- Read *Where the Money Grows* by Garet Garrett. (PDF provided via MiddFiles. Also available for free [on-line](#) and via [Amazon](#).)
- *Quantitative Value*, chapters 1 -- 3. (PDFs provided.)
- Read "[Publication, Publication](#)" by Gary King.
- Sign up for Coursera and GitHub with a username that is recognizable to someone who knows your real name. I am davidkane9.
- Complete this [15 minute overview of Git](#).
- Begin the "The Data Scientist's Toolbox," "R Programming" and "Reproducible Research" on Coursera. You should sign up for these at the beginning of February and keep up with the work even though our first class is not until the middle of the month.
- In class, look at Middlebury course data and Bloomberg index data.

### Week 2: February 22

- Read *A Random Walk Down Wall Street*, chapters 1 -- 11.
- Begin the Wharton "Introduction to Financial Accounting" on Coursera.
- "[A Quick Introduction to Version Control with Git and GitHub](#)" by Blischak et al.
- Introductions to [R Markdown](#), [dplyr](#), and [ggplot2](#). [dplyr](#) and [ggplot2](#) are covered in more detail in the Coursera courses assigned for next month, but we need to start using them now.
- This completes the first part of the course. My goal has been to provide an overview of the quantitative approach and to start you on the path of learning the necessary tools.
- In class, explore Ken French's [Data Library](#).

### Week 3: February 29

- Begin working with monthly data. This session, and the next two, are the second part of the course. My goal is to introduce you to the academic literature on quantitative equities and to start working with financial data, albeit at a monthly frequency.
- *Quantitative Value*, chapters 4 - 12. (You need to get the book.)
- Introductions to [knitr](#), [readr](#) and [readxl](#),
- "[Quality Minus Junk](#)" by Clifford S. Asness, Andrea Frazzini and Lasse H. Pedersen, 2013. AQR Working Paper.
- "[Value and Momentum Everywhere](#)" by Clifford S. Asness, Tobias Moskowitz and Lasse Heje Pedersen, 2013. *Journal of Finance*, 68, 929-985.
- In class, use [AQR Data Library](#).

### Week 4: March 7

- Complete "The Data Scientist's Toolbox," "R Programming" and "Reproducible Research." Submit course completion certificates via the class shared document.
- Begin three more Coursera classes: "Getting and Cleaning Data," "Exploratory Data Analysis" and "Developing Data Products." These must be completed by mid-April.
- Read [an introduction to tidyr](#) and "[Tidy Data](#)."

- "[Betting against beta](#)" by Andrea Frazzini and Lasse Heje Pedersen, 2014. *Journal of Financial Economics*, 111, 1–25.
- "[Time Series Momentum](#)" by Tobias J. Moskowitz, Yao Hua Ooi and Lasse Heje Pedersen, 2012. *Journal of Financial Economics*, 104, 228–250.
- [Overview](#) of Middlebury endowment and performance [pdf](#).
- In class, use [AQR Data Library](#).

## Week 5: March 14

- Guest lecture by Churchill Franklin '71, CEO of Acadian Asset Management and former chair of the Middlebury board of trustees. (All items available via [MiddFiles](#).)
- Acadian: The Case for Active Management
- Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler. 2011. "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly." *Financial Analysts Journal*, vol. 67, no. 1 (January/February):40–54.
- Baker, Malcolm, Brendan Bradley, and Ryan Taliaferro. 2014. "The Low Risk Anomaly: A Decomposition into Micro and Macro Effects." *Financial Analysts Journal*, vol. 70, no. 2 (March/April):43–58.
- Acadian Managed Volatility Profile
- Acadian Q4 2015 Report
- Acadian Smart Beta
- Acadian Diversified Alpha
- Acadian Managed Volatility Pitch Book

## Week 6: March 21

- Begin working with daily stock prices.
- Advanced uses of dplyr for [window functions](#).
- "[Does the stock market overreact?](#)" by Werner DeBondt and Richard Thaler, 1985. *Journal of Finance*, 40, 793–805.
- "[Evidence of predictable behavior of security returns](#)" by Narasimhan Jegadeesh, 1990. *Journal of Finance*, 45, 881-898.
- Complete "Introduction to Financial Accounting." Submit course completion certificates via the class shared document.

## Week 7: April 4

- Start work on replication project. This should be based in one of the articles that we have not examined together, or an article of your choosing.

## Week 8: April 11

- Read [R Packages](#) by Hadley Wickham.
- [Resources](#) about R packages from Jennifer Bryan.

## Week 9: April 18

- Finish "Getting and Cleaning Data," "Exploratory Data Analysis" and "Developing Data Products." Submit course completion certificates via the class shared document.

## Week 10: April 25

- Continue work on replication project. All students should have replicated the major results of their chosen article. Replication work must be stored in an R package on Github.

## **Week 11: May 2**

- Visiting lecture from Professor David Colander on *Capital Ideas* and *Capital Ideas Evolving* by Peter Bernstein. Both books are excellent but they are not required reading.
- Each student must evaluate a classmate's initial draft. Can you download her package? Can you compile his vignette? Do you get the same answers?
- Extend your own replication in some interesting non-trivial fashion. Much of class discussion will be devoted to the topic of what it means to be "interesting" in the context of quantitative equity research.

## **Week 12: May 9**

- Students will publicly present their work, either in class or in an open-to-the-public meeting. Details to be discussed.
- Final projects must be [submitted](#) for permanent public display at the [Digital Collections at Middlebury](#). You will have worked very hard on your paper and you should be proud of the results. You must select the "My work shall be available to anyone worldwide with Internet access" option. Make sure to include the strings "INTD 318" and "Quantitative Finance" in your list of keywords when submitting.
- Final papers must be submitted for publication to an academic journal. Student-run journals ([here](#)) are a good option.