

The goal of this course is to help students to handle financial data using R computer language. More specifically, the course will help students to write R programs to conduct financial analytics, while the purpose of financial analytics is to transform financial data to insights (through visualization in particular) that support, improve, and automate financial decisions. The majority of class time will be used to explain various R language skills so that students can eventually use R to do a real-world financial analysis. In addition to the general knowledge on R programming language, a special type of R commands from the so-called ‘tidyverse’ system will also be taught in depth. An important area of R usage is time series analysis which will be covered extensively in the class. This knowledge will then be applied to financial prices and this hopefully will be useful for students who are interested in quantitative finance.

The class will be taught in Chinese. Students will receive teaching materials in the format of the **R Markdown file** each week and they will be followed closely in the class. All the registered students must bring a personal computer to the very first class and, preferably, have **R** and **RStudio** installed and running well. Go to google these two terms first and find the web sites for downloading. The installment process is simple and straightforward. You should search for solution on the internet if any problem occurs. The first class should be attended. Otherwise, you will have a lot of problems with the fairly complicated class setup (for example, you will have no idea about what **R Markdown files** mean and it is devastating). Students will need to download and upload quite a few files during the class time. So you should always be able to connect your computer to the internet.

The course grade will be based on a number of in-class quizzes/homeworks (40%) and a final report (60%). The subject of the final report however is not limited to finance areas and can be about any social issues. A proposal for your final report should be turned in three weeks before the end of this semester. You will then be advised whether your final report is feasible and acceptable.

The main teaching materials will be provided by the teacher through a computer software called PresTree, which runs only on Windows PC. It will be a problem for those students who only have Apple or Linux computers. These students should either try to install a Windows subsystem in their computers to run PresTree or to buy/borrow a Windows PC. However, there are a great many reference materials on the internet which can be used to supplement PresTree.

The main topics to be covered in the course are as follows. But how they are taught in the class will not follow the order of this list. Due to time constraint all the topics will be taught only briefly in the class and once they are mentioned in the class they will be tested either in homework or in quiz so long as they are included in the PresTree software.

- R language basics
 1. Presentation skills: R markdown
 2. Using packages and getting datasets
 3. Operation units: vectors, lists, factors, matrices, data-frames,
 4. Statistical operations: Statistical functions, Handling missing values, Simulation, Matrix algebra
 5. Basic plots: scatter plots, line plots, bar plots and tables, histogram, box plots

6. Control structures: conditional statements, iteration statements, writing functions

- Tidyverse ecosystem
 1. Data-visualization: **ggplot2** packages
 2. Data wrangling: **dyplr** packages
 3. Pivot Table technique: **tidyr** package
 4. Factor manipulation: **forcats** package
 5. String manipulation: **stringr** package
- Time-series analysis:
 1. Time-series classes: ts and xts series
 2. Date classes: formats, series, and **lubridate** package
- Financial data analysis
 1. Downloading and plotting financial data: **quantmod** package
 2. Measuring the performance of financial returns: **PerformanceAnalytics** and **tidyquant** packages
 - Returns calculations and backtesting
 - Sharpe ratios
 - CAPM
 - Drawdowns
 - Downside risks
 - VaR, ES

Many of you might know Python and love it. R is in many ways similar to Python and with almost identical functionality. There are many people who know both languages well and continue use both regularly. Knowing only one and bad mouth the other is childish. R language is started off by statisticians so that it is quite capable of statistical analysis. However, there are tens of thousand of R packages from hundreds of different fields other than statistics cumulated over the past 40 years. Furthermore, R (particularly with ggplot2) is extremely powerful in plotting. Almost no one who knows how to make plots in R have negative feeling about it. For those of you who only know Excel plotting, you probably will never go back to excel for plotting any more after this class.