# Business Analytics using R (FIN456A)
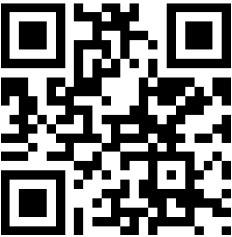
## *(CRN: 37935, Fall2016)*

| | |
|---|---|
| Instructor: | Paul Yan |
| Contact Information: | Email:  yany@canisius.edu <br> Phone:  (716) 888-2604    Office:   CT308 |
| Lecture: | TR 1:00pm-2:15pm                 @  OM111 (Financial Services Lab) |
| Office hours: | TR: 11:30-12:30pm, 4:00pm-4:30  @  CT308 |
| Prerequisites: | FIN311 (Corporate Finance) and Investments (FIN312) , or equivalents, or with permission from the instructor (about half  hour interview) |
| Textbook: | Financial Modeling using R by Yuxing Yan <br> ISBN: 978-1-68187-530-9, Tate Publishing, 2016 <br><br>     |
| Websites: | *An Introduction to R*        http://canisius.edu/~yany/doc/R-intro.pdf <br> *The R Language Definition* http://canisius.edu/~yany/doc/R-lang.pdf <br> *My related R web site:*        http://canisius.edu/~yany/R.shtml |
| QR codes for two more web sites | https://www.r-project.org/         http://www3.canisius.edu/~yany/R.shtml <br><br>     |
| One-line R codes | *> source("http://canisius.edu/~yany/fin456.R")* <br>  Note: I will explain this line in week 2 |
| Course Description: | This course introduces students to the open source (free) software language R which is used extensively on Wall Street to access, process and analyze data.  Students will learn how to download and process data from the many publicly available data sources such as Yahoo!Finance, Google Finance, SEC filings, and the Federal Reserve Bank's Data Library.  R will be applied to the analysis of financial ratios, investment portfolios, options, CAPM, Monte Carlo simulation, Value at Risk |

| | |
|---|---|
| | modeling and market microstructure. |
| Capacity of the class | 20 (ideally the number of students should be less than 15 since hands-on is critical ) |
| Three objectives: | 1) Learn/review basic financial concepts such as Ratio Analysis, Portfolio Theory, CAPM, Fama-French-Carhart Factor Model, Monte Carlo simulation, Options Theory, VaR (Value at Risk) and Market microstructure<br>2) Learn and apply R to finance<br>3) Focus on publicly available financial data such as Yahoo Finance, Google Finance, Prof. French's Data Library and Federal Reserve Economic Data Library (FRED). |
| R Software: | R is open source statistical and computational software, see www.r-project.org. For this course, it is optional. Students could use R as a financial calculator. I will discuss R in the first 4 lectures. |
| Academic Integrity: | Students are expected to know and understand college policies with regard to Academic Integrity Code.   Violations of academic integrity will be prosecuted fully.  Please note that you are responsible for reporting any instances where other students have violated these policies.  Failure to do so will result in penalties as well.  If you have any questions about this policy, please see the instructor. |
| Attendance Policy: | Attending classes regularly is required. Before-class preparation and in-class participation is an integral part of this course. Students are strongly encouraged to participate in class discussions and ask questions. Students are encouraged to discuss current events relevant to this course or their own experiences. Homework problems are regularly assigned. |
| Academic and Accessibility Support Services: | The GRIFF Center for Academic Engagement provides comprehensive programs, tutoring services, and resources to support student academic and career success. If you would like to learn more about academic support, please stop in Old Main 013 or call 716-888-2170.  Visit the GRIFF Center webpage at: http://www.canisius.edu/griff-center/ . Accessibility Support (716-888-2170), which is located in the Griff Center for Academic Engagement (OM 013), is responsible for arranging appropriate academic accommodations for students with documented disabilities.  If anyone in this course falls into this category, please contact Accessibility Support so that an appropriate course of action may be determined.  For additional information, see http://www.canisius.edu/dss/ |
| Course Level Learning Goals: | Learn basic finance theory; understand various decision rules, such as NPV, IRR and payback rules to evaluate different investment projects; understand the term-structure of interest rate, spread, and default risk; learn how to download data from public sources, such as Yahoo!Finance, Federal Reserve Bank's Data Library, Prof. French's Data Library, be able to use Excel to process data and calculate certain measures, .e.g., estimate the market risk for a given set of tickers by applying CAPM. |
| | This course is designed to help students achieve one or more College Core, Business Program and/or Major level learning goals and |

| | |
|---|---|
| College, Program and Major Learning Goals: | objectives. You can see the specific College, Program or Major level learning goals and objectives associated with the course from this page on the College website: http://bit.ly/bcoreLG |
| Grade Evaluation: | Homework 30%<br>Midterm 20%<br>Final exam 25%<br>Group project 10%<br>Group presentation 5%<br>Class participation 10%<br>-----------------------------------------------<br>Total 100% |
| Teaching Methods: | Each class will be consist of two parts: lecture (including discussion of homework) and hands-on. |
| Group project | Each group can have up to three members. A topic should be closely associated with this course. The maximum number of pages of your report is 15 with 12-point font. Please discuss with me your topic before you start to work on it. Some basic criterions are listed below.<br>Real world topics are especially encouraged. Three parts are essential:<br>1) theory and background of the topic,<br>2) R programs with a short explanation of the codes,<br>3) final data set (plus the codes to process the data, the source of raw data)<br> Note: please do not send me your raw data.<br>The second type of projects is to study one of R packages. Three parts are critical:<br>1)why this specific package is useful in finance<br>2)a summary of all or most important functions offered by the package<br>3)examples to use them<br>Note: see a list of potential topics, at the end of the syllabus, for the group projects. |
| Makeup Exams: | To be eligible for a makeup examination, a student must submit to the instructor written documentation of the reason for missing a scheduled examination due to medical problems or death of an immediate family member. This has to be done before the scheduled exam. The instructor (not the student) determines whether and when a makeup is to be given. If a makeup examination is to be given, the instructor will determine the type of makeup examination. If the student misses (for any reason) the scheduled makeup examination, additional makeups are not permissible. |
| Laptop policy: | Students are encouraged to bring a laptop to the class. However, it should be used for class related activities only. |
| Course Schedule: | For the detailed schedule, see below. I reserve the right to change the course schedule throughout the semester. Changes to the schedule will be announced in class or via email. |

**Tentative schedule**

| # | Date | Topics | Description   (F for Finance) |
|---|---|---|---|
| 1 | | Syllabus discussion, introduction to R | A short survey, self-intro, syllabus, course structure,  mid-term and final<br>F: Review of basic financial concepts, financial formulae,  risk vs. return, how to measure risk,  time value of money, present value (PV), future value (FV), PV(annuity), etc.<br><br>R:  Installation, assignment, basic math functions: mean(), min(),max(), median(), sd() and use R as a scientific calculator |
| 2 | | Review of basic finance concepts and formulae (functions) | F: How to estimate returns? PV(bond), simple and compound interest, conversion of returns for different frequencies<br><br>R: How to write an R function? double_f(), pv_f(), fv_f(), IRR(), pv_annuity(), fv_annuity(), pv(perpetuity), pv(perpetuity_due) How to call your functions?  several ways to input data, matrix, differential operator  and use R as a financial calculator |
| 3 | | Options | F: Black-Scholes-Merton option model, trading strategies with options, implied volatility, Black's approximation for an American option, Greeks, put-call parity, hedging strategy, risk-neutral evaluation<br><br>R: pnorm() [cumulative standard normal distribution], bs_f.R, Implied_vol.R, greeks.R |
| 4 | | Financial Statement Analysis | F:  Financial statement analysis, ratio analysis, probability ratio: operating margin, net profit margin, ROA, ROE, current ratio, book debt-equity  ratio<br><br>R: Retrieve BS, IS, CF,  current ratio, debt equity ratio, |
| 5 | | Open data [Yahoo finance etc.] | F: daily vs. monthly returns, Yahoo finance, Rf, French's Data Library<br>R: How to input data from a text file? simple programs to download historical price data from Yahoo finance, download data from Prof. French's data library , download risk-free rate |
| 6 | | CAPM | F: CAPM, β estimation, rolling/portfolio β, hedging portfolio  market risk<br>R Several functions: as.Date, week_day_effect.R , data.frame, ,beta.R, rolling_beta.R, replicate |
| 7 | | Multi-factor models, Sharpe Treynor ratios | F: Fama-French 3-factor model, momentum strategy, Sharpe ratio, Treynor's ratio, 52-weeks high, Jensen's α,<br>R:  ff3factor, ff4,  weekday/January effect. sharpe.R, Treynor.R |
| 8 | | | **Midterm** |
| 9 | | T-test, F-test, Autocorrelation, Causality | F: T-test for significance, equality of means, F-test for difference of volatility, Granger causality test, Durbin-Watson autocorrelation test<br>R: t.test(), var.test(), dwtest(), Wilcoxon.test(), granger_test().<br>**Note: post about a dozen topics in detail** |
| 10 | | Monte Carlo Simulation | F: simulation and assumptions, normality test, estimate variance-covariance matrix, conversion variances between different  frequencies, path dependent options, sensitivity analysis,  scenario analysis<br>R: rnorm(x), random number from normal, uniform distribution     Several functions: as.Date, week_day_effect.R,In class-exercise: find 1 C among 500 Os and Monte Carlo Simulation to price European and Asian options |

Continued

| Lecture | Date | Topic | F (finance) and R |
|---|---|---|---|
| 11 | | CRSP for teaching using R | F: What is CRSP? CRSP monthly, daily time series data, event data (more topics for term projects)<br><br>R: stockMonthly, indexMonthly, indexDaily, stockD1925 to stockD2014, various R program to retrieve/process data efficiently |
| 12 | | Portfolio theory | F: variance, standard deviation, correlation, return matrix, portfolio return, portfolio volatility of 2-stock (n-stock) portfolio, variance-covariance matrix, portfolio optimization<br><br>R: package "fPortfolio" |
| 13 | | Value at Risk | F: introduction to VaR, (standard) normal distribution, thick tail distribution<br><br>R: standard normal distribution, VaR_01.R, VaR_02.R, introduction to packages in R. |
| 14 | | Group Presentation | 3 to 4 groups |
| 15 | | Group Presentation | All other groups |
| | Extra | Text Analysis | Word frequency table, how to generate a word frequency table keywords<br><br>R: searchKeyword.R, wordPicture.R |
| | Extra | Machine learning | Library: tm |
| | TBA | **Final** | **Final-exam** |

5

**Appendix A:  A list of potential topics for term projects**

| | | | |
|---|---|---|---|
| **Warm up** | | 1 | Financial statement analysis |
| | | 2 | An interned connected financial calculator (Yan, 2012) |
| | | 3 | Correlations among  stocks in US, UK, Canada, France, China, Japan and Australia |
| | | 4 | A Business cycle indicator (Yan and Zhang, 2015) |
| | | 5 | Use journal ranking data efficiently (SCImago Journal and Country Ranking) |
| | | 6 | Find an optimal portfolio |
| | | 7 | How much you need when you retire? Social Security Benefit calculator |
| | | 8 | Which party, Republican or Democratic, could manage the economy better? |
| | | 9 | Monte Carlo Simulation (standard normal distribution, one variables) VaR |
| | | 10 | PCA (Principal Component Analysis) |
| **Data** | **Public data** | 11 | Generate R data sets for 200 stocks, CPI, GDP, Unemployment rate etc. |
| | | 12 | Generate R data sets for Fama-French 3-factors, 5 factors etc. |
| | | 13 | Generate R data sets for all SEC 10Q and 10K index files from SEC (1993-2015) |
| | | 14 | Generate R data sets for TORQ (Trade, Order, Report and Quote) database |
| | | 15 | Generate R data sets for TDAQ (millisecond by millisecond transaction data) |
| | | 16 | Parse 10K data from SEC filings, generate related R data sets |
| | **CRSP/TAQ** | 17 | Generate R data sets for one month's TAQ data (MTAQ) |
| | | 18 | Generate R data sets for crspInfo, stockMonthly, indexMonthly for CRSP |
| | | 19 | Generate R data sets for stockDaily, indexDaily for CRSP |
| | | 20 | Generate R data sets for TDAQ for several months |
| **Using public data** | | 21 | Are annual beta mean reversion? |
| | | 22 | Test the January and weekday effects |
| | | 23 | Does size effect exist? |
| | | 24 | Tracking errors |
| | | 25 | Z-score (bankruptcy prediction, Altman, 1968) |
| | | 26 | 52-week High trading strategy using more than 200 stocks |
| | | 27 | estimate Roll (1984) spread from daily data (Roll, 1984) |
| | | 28 | Assessment of multiple choice questions using R |
| | | 29 | Monte Carlo Simulation (capital budgeting, replicate a Slot Machine) |
| | | 30 | Monte Carlo Simulation (one variables) VaR, n correlated stocks |
| **Using CRSP or TAQ** | | 31 | Replicate S&P500 EW (equal-weighted) and VW (value-weighted) returns |
| | | 32 | Is liquidity factor priced?  (Amihud, 2002) |
| | | 33 | What is the color of your firm, blue or red? (Yan, 2014) |
| | | 34 | Which model is the best, CAPM, FF3, FFC4, or FF5? |
| | | 35 | Estimate spread, relative spread, expected spread etc. by using TAQ |
| | | 36 | Process TAQ efficiently, how to process 30 year MTAQ data efficiently ? |
| | | 37 | Replicate momentum trading strategy (Jegadeesh and Titman, 1993) |
| | | 38 | Replicate industry momentum trading strategy (Moskowitz and Grinblatt, 1999) |
| | | 39 | Replicate 52-week high trading strategy (George and Huang, 2004) |
| | | 40 | Replicate max- trading strategy (Bali, Cakici and Whitelaw, 2011) |
| | | 41 | Impact of business cycle on the above four trading strategies, (Yan and Zhang , 2015) |

**References (most of them are for term projects)**

Altman, Edward, Altman, E. (1968), Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, 23(4), pp 598-608.

Amihud, Yakov, 2002, Illiquidity and Stock returns, *Journal of Financial Markets* 5, 31-56.

Bali, Turan G., Nusret Cakici, and Robert F. Whitelaw, 2011, Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns, *Journal of Financial Economics* 99 427-446.

George, Thomas J, and Chuan-Yang Huang, 2004, The 52-week High and Momentum Investing, *Journal of Finance* 54, 5, 2145-2176.

Jegadeesh, N., and S. Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* 48, 65-91.

Moskowitz, Tobias, and Mark Grinblatt, 1999, Do industries explain momentum? *Journal of Finance* 54, 2017-2069.

Pastor, L. & Stambaugh, R., 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642-685.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 39, 1127-1139.

Yan, Yuxing, An internet connected financial calculator, 2012, *Journal of Accounting and Finance* 12(5), 59-70.