





**Business Analytics using R (FIN456A, MBA674A)**  
(CRN: 37935 & 37938, Fall2016)

Instructor:	Paul Yan
Contact Information:	Email: <a href="mailto:yany@canisius.edu">yany@canisius.edu</a> 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 (if pass a half- hour interview)
Textbook:	Financial Modeling using R by Yuxing Yan ISBN: 978-1-68187-530-9, Tate Publishing, 2016  
Websites:	<i>An Introduction to R</i> <a href="http://canisius.edu/~yany/R/181.pdf">http://canisius.edu/~yany/R/181.pdf</a> <i>The R Language Definition</i> <a href="http://canisius.edu/~yany/R/R-lang.pdf">http://canisius.edu/~yany/R/R-lang.pdf</a> <i>My related R web site:</i> <a href="http://canisius.edu/~yany/R.shtml">http://canisius.edu/~yany/R.shtml</a>
QR codes for two more web sites	<a href="https://www.r-project.org/">https://www.r-project.org/</a> <a href="http://www3.canisius.edu/~yany/R.shtml">http://www3.canisius.edu/~yany/R.shtml</a>   Fun question: how many lines of R codes to generate one image?
One-line R codes	<pre>&gt;source("http://canisius.edu/~yany/fin456.R")</pre> 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:	<ol style="list-style-type: none"> <li>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</li> <li>2) Learn and apply R to finance</li> <li>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).</li> </ol>
R Software:	R is open source statistical and computational software, see <a href="http://www.r-project.org">www.r-project.org</a> . 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 <a href="#">Academic Integrity Code</a> . 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: <a href="http://www.canisius.edu/griff-center/">http://www.canisius.edu/griff-center/</a> . 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 <a href="http://www.canisius.edu/dss/">http://www.canisius.edu/dss/</a>
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.
College, Program and	This course is designed to help students achieve one or more College Core, Business Program and/or Major level learning goals and objectives. You can see the specific College, Program or Major level

Major Learning Goals:	learning goals and objectives associated with the course from this page on the College website: <a href="http://bit.ly/bcoreLG">http://bit.ly/bcoreLG</a>																
Grade Evaluation:	<table> <tr> <td>Homework</td> <td>30%</td> </tr> <tr> <td>Midterm</td> <td>20%</td> </tr> <tr> <td>Final exam</td> <td>25%</td> </tr> <tr> <td>Group project</td> <td>10%</td> </tr> <tr> <td>Group presentation</td> <td>5%</td> </tr> <tr> <td>Class participation</td> <td>10%</td> </tr> <tr> <td colspan="2">-----</td> </tr> <tr> <td>Total</td> <td>100%</td> </tr> </table>	Homework	30%	Midterm	20%	Final exam	25%	Group project	10%	Group presentation	5%	Class participation	10%	-----		Total	100%
Homework	30%																
Midterm	20%																
Final exam	25%																
Group project	10%																
Group presentation	5%																
Class participation	10%																
-----																	
Total	100%																
Teaching Methods:	Each class will be consist of two parts: lecture (including discussion of homework) and hands-on.																
Group project <sup>1</sup>	<p>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. Real world topics are especially encouraged.</p> <p>Three parts are essential:</p> <ol style="list-style-type: none"> <li>1) theory and background of the topic,</li> <li>2) R programs with a short explanation of the codes,</li> <li>3) final data set (plus the codes to process the data, the source of raw data). Note: please do not send me your raw data.</li> </ol> <p>The second type of projects is to study one of R packages. Three parts are critical:</p> <ol style="list-style-type: none"> <li>1) why this specific package is useful in finance</li> <li>2) a summary of all or most important functions offered by the package</li> <li>3) examples to use them</li> </ol>																
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.																

<sup>1</sup> For group projects, see a list of potential topics, at the end of the syllabus. After mid-term, I will discuss a few topics in more detail.

## Tentative schedule

#	Date	Topics	Description (F for Finance)
1	8/30 9/1	Syllabus discussion, introduction to R	A short survey, self-intro, syllabus, course structure, mid-term and final 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.  R: Installation, assignment, basic math functions: mean(), min(),max(), median(), sd() and use R as a scientific calculator
2	9/6 9/8	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  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	9/13 9/15	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  R: pnorm() [cumulative standard normal distribution], bs_f.R, Implied_vol.R, greeks.R
4	9/20 9/22	Financial Statement Analysis	F: Financial statement analysis, ratio analysis, probability ratio: operating margin, net profit margin, ROA, ROE, current ratio, book debt-equity ratio  R: Retrieve BS, IS, CF, current ratio, debt equity ratio,
5	9/27 9/29	Open data [Yahoo finance etc.]	F: daily vs. monthly returns, Yahoo finance, Rf, French's Data Library 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	10/4 10/6	CAPM	F: CAPM, $\beta$ estimation, rolling/portfolio $\beta$ , hedging portfolio market risk R Several functions: as.Date, week_day_effect.R, data.frame, beta.R, rolling_beta.R, replicate
7	<b>10/11</b> 10/13	Multi-factor models, Sharpe Treynor ratios	<b>(10/11) no class</b>  F: Fama-French 3-factor model, momentum strategy, Sharpe ratio, Treynor's ratio, 52-weeks high, Jensen's $\alpha$ , R: ff3factor, ff4, weekday/January effect. sharpe.R, Treynor.R
8	<b>10/18</b> <b>10/20</b>		<b>Midterm</b>  <b>No Class (I will attend the FMA annual conference in Las Vegas)</b>
9	10/25 10/27	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 R: t.test(), var.test(), dwtest(), Wilcoxon.test(), granger_test(). <b>Note: post about a dozen topics in detail</b>
10	11/1 11/3	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 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	11/8 11/10	CRSP for teaching using R	F: What is CRSP? CRSP monthly, daily time series data, event data (more topics for term projects) R: stockMonthly, indexMonthly, indexDaily, stockD1925 to stockD2014, various R program to retrieve/process data efficiently
12	11/15 11/17	Portfolio theory	F: variance, standard deviation, correlation, return matrix, portfolio return, portfolio volatility of 2-stock (n-stock) portfolio, variance-covariance matrix, portfolio optimization R: package “fPortfolio”
13	11/22  <b>11/24</b>	Value at Risk	F: introduction to VaR, (standard) normal distribution, thick tail distribution R: standard normal distribution, VaR_01.R, VaR_02.R, introduction to packages in R.  <b>Thanks giving holiday (no Class)</b>
14	11/29 12/1	Group Presentation	(Value at Risk) 1 to 2 groups
15	12/6 12/8	Group Presentation	rest Backup day
	Extra	Text Analysis	Word frequency table, how to generate a word frequency table keywords R: R package called tm(), searchKeyword.R, wordPicture.R
	Extra	Machine learning	R package called rattle()
	TBA	<b>Final</b>	<b>Final-exam</b>

Canisius College Academic Calendar: <http://www.canisius.edu/academics/events/>

## 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.