

# Research Methods in Political Science I

## 1. Introduction

Yuki Yanai

School of Law & Graduate School of Law

October 7, 2015



**KOBE UNIVERSITY**

# Today's Lecture



- 1 Introduction
  - Research Methods
  - Course Overview
  
- 2 Computing
  - R
  - RStudio

# Research in Political Science



## How should we conduct political-science research?

- Research Design
- How to solve puzzles
  - Formal Theory (analytic approach)
  - Qualitative Methods (case studies?)
  - **Quantitative Methods** (statistical approach)

## What are quantitative methods?



- Analyze quantified data
- Use statistics: both descriptive and inferential
  - Exploratory data analysis (EDA)
  - Statistical tests of hypotheses
  - Examine if the data conform to some theoretical expectations (observable implications)



## What will you learn in this course?

### Essentials of quantitative methods in political science

- How to use R (and some related languages/packages)
- How to collect data: web scraping
- How to analyze data given some specific research questions
- How to present the results of your research
- How to keep records of your research: **reproducible political-science research**

## Why should you take this course?



- ① Students who are interested in quantitative methods will be able to:
  - implement quantitative analyses for your research
  - conduct your research more efficiently
- ② Students who are *not* interested in quantitative methods will be:
  - interested in quantitative methods
  - a better reader in political science
  - able to criticize quantitative analyses in political science

## Class Format



- Lecture and lab work in each class meeting
- You may bring your own laptop computer
- You should always have your computer on tap
- PLEASE interrupt me to ask questions!

## Requirements and Grade



- Complete the readings prior to each lesson
- Your final grade is based on:
  - Class participation: 20%
  - Weekly assignments: 50%
  - Final project (a 10-page paper): 30%
- *F* (0 point) will be given if you do not submit the final paper

## Course website



- Visit my website: <http://yukiyanai.com>
- Go to “classes”, then “Research Methods in Political Science I”
- Syllabus is available at the website
- Check “Class Materials” regularly

# Syllabus

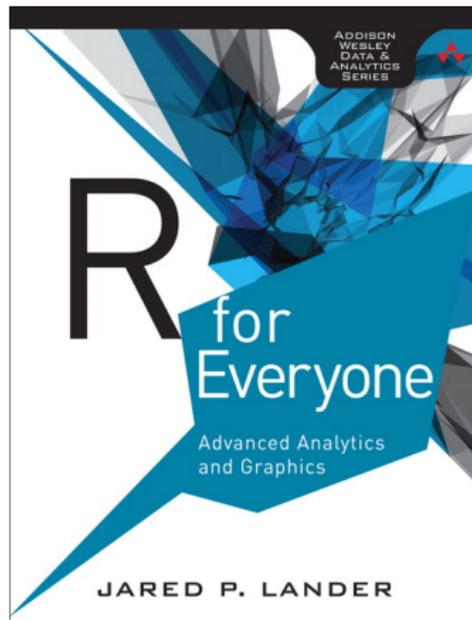


- PDF file of the syllabus is available at  
<http://www2.kobe-u.ac.jp/~yyanai/classes/rm1/docs/syllabus-rm1-2015fall.pdf>
- Use hyperlinks in the syllabus to get papers: you need to connect to the university's network to obtain most of the papers
- **Read the syllabus carefully**

## Required Book



- Jared P. Lander. 2014. *R for Everyone: Advanced Analytics and Graphics*. Addison-Wesley.
- Data used in this book is available at <http://www.jaredlander.com/r-for-everyone/>



## Required Book Translated into Japanese



- 高柳慎一, 秋山幸史, 蓑田高志 訳.  
2015. 『みんなのR：データ分析と  
統計解析の新しい教科書』マイ  
ナビ



## Office Hours



- Mon & Wed, 12–1pm
- and by appointment
- No.404, Building No.4 (Dai-yon Gakusha)
- Ask questions by email
- Ask questions **in class**

## What Is R?



- Statistical computing software
- *The* standard software in statistics
- Run on Linux, Mac, and Windows
- Object oriented
- Open Source: **free** customization
- Available at CRAN for **free**

# Screenshot of R



```

R version 3.1.1 (2014-07-10) -- "Sock it to Me"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.1.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.65 (6784) x86_64-apple-darwin13.1.0]

[History restored from /Users/yuki/.Rhistory]

> 1 + 2
[1] 3
> 2 * 5
[1] 10
> 3^3
[1] 27
> a <- 1:100
> sum(a)
[1] 5050
> |

```

## Strengths of R



- Free
  - How much is Stata? How much is SPSS?
- Flexible
  - Lots of packages are available
  - You can write your own function
- Widely used
  - Enhance communication
  - Make it easy to conduct reproducible research
- Beautiful graphics

## Quick Overview



Watch the video

<http://www.revolutionanalytics.com/what-r>

## Weakness of R



- Not right out of the box: especially cumbersome for Japanese (Chinese, Korean, etc) users
- Need to program
- Difficult for beginners to use

## Run R in RStudio



RStudio:

- Integrated development environment (IDE) for R
- Free!
- Make your R life easier
  - 4-pane view to display relevant info
  - Easy to manage projects
  - Can edit not only R scripts but also R Markdown files
  - Keep all the figures you create within a session





# Screenshot of RStudio



The screenshot displays the RStudio environment with the following components:

- Script Editor:** Contains R code for generating data and fitting a linear model.
 

```

20 myd1$x5 <- rnorm(500, mean = 10, sd = 0.01)
21
22 ## set parameter values
23 b0 <- 1 #y intercept
24 b1 <- 4
25 b2 <- -0.1
26 b3 <- 0.8
27 b4 <- 10^(-6)
28 b5 <- 8
29
30 myd1$y <- b0 + b1*myd1$x1 + b2*myd1$x2 + b3*myd1$x3 + b4*myd1$x4 + b5*myd1$x5 + rnorm(500)
31
32
33 OLS
34 ... (r)
35 fit1 <- lm(y ~ x1, data = myd1)
36 summary(fit1)
37 ...
38
39 Compare the fit to the true model.
40 ... (r)
41 true.model <- lm(y ~ x1 + x2 + x3 + x4 + x5, data = myd1)
42 summary(true.model)
43 ...
      
```
- Environment Pane:** Shows the 'myd1' data frame with 500 observations and 6 variables. The 'Values' section lists:
 

Variable	Value
b0	1
b1	4
b2	-0.1
b3	0.8
b4	1e-06
b5	8
fit1	List of 12
model1	List of 12
true.model	List of 12
- Console:** Displays the output of the summary functions.
 

```

RStudio 1.1.1
R Markdown 1.1
27.14 Chunk 1.1

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 41.3259847 43.2662835  0.955  0.348
x1           3.9906856  0.0144468 276.232 <2e-16
x2          -0.0997436  0.0007321 -136.249 <2e-16
x3           0.8010602  0.0036837 217.462 <2e-16
x4           0.0013757  0.0014239  0.966  0.334
x5           3.9696815  4.3263289  0.918  0.359

Residual standard error: 0.9423 on 494 degrees of freedom
Multiple R-squared:  0.9966, Adjusted R-squared:  0.9965
F-statistic: 2.856e+04 on 5 and 494 DF, p-value: < 2.2e-16
      
```

# Let's Play with R



- Go to the course website
- Install R and RStudio
- Open RStudio and use R



## When You Are Stuck

- Google
- Search in Stack Overflow  
`http://stackoverflow.com/questions/tagged/r`
  - Read the questions and answers already there
  - Post a question if necessary
- Find a geek friend
- Become a geek

## Homework



- Install R and RStudio on your computer
- Enjoy R (at least for an hour)
- Complete Assignment 1 posted on the course website
- Due: 9am on Wednesday 14th.
- Read the syllabus

## Next Class



- Review elementary statistics
- Compute statistics with R