550.400: Mathematical Modeling and Consulting

Lecture Notes

Instructor:
Dr. N. H. Lee

JHU AMS 2012 FALL

Last Compiled on September 10, 2012

Outline

Syllabus

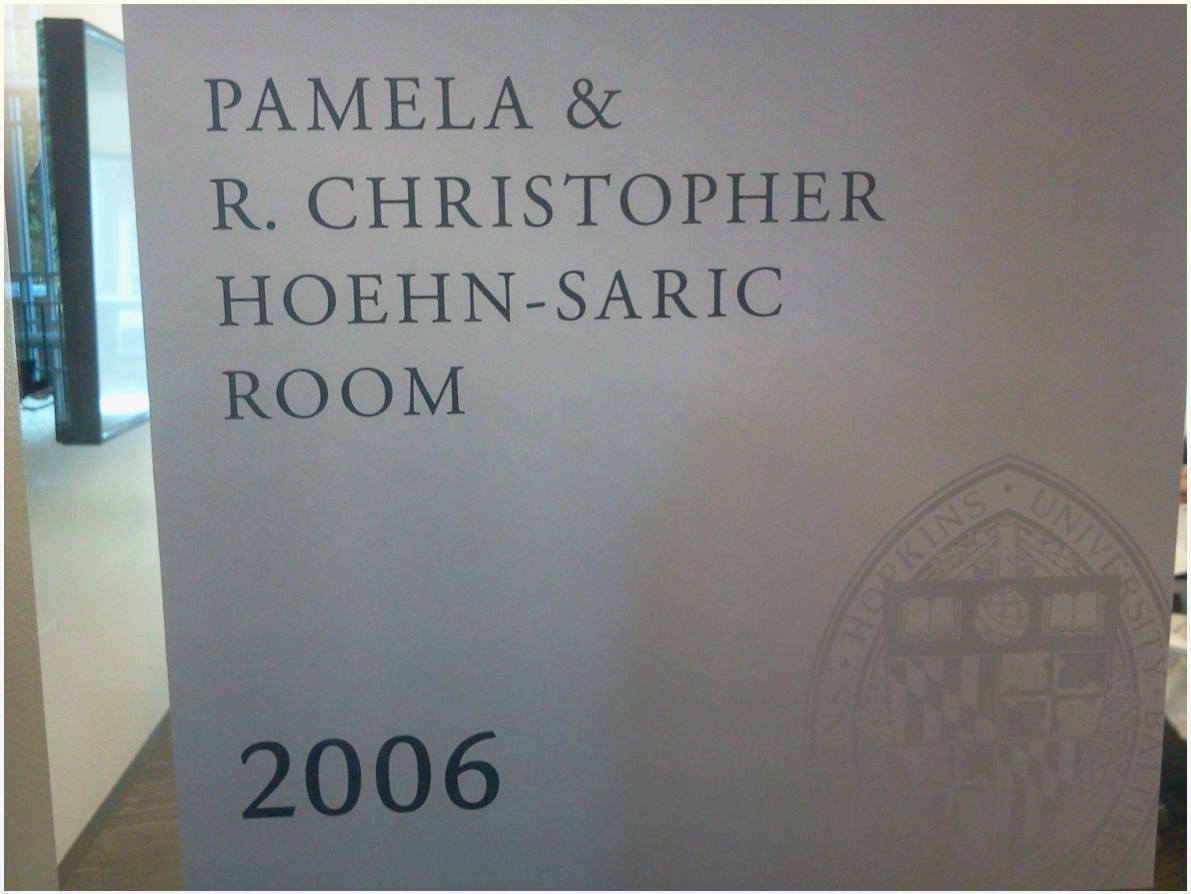
- Grade Policy
- Attendance
- *Tentative* Schedule
- Blackboard
- Misc.

Course Book Reserve

JHU Library Reserve Service

Presentations in this class

For your presentation recording needs

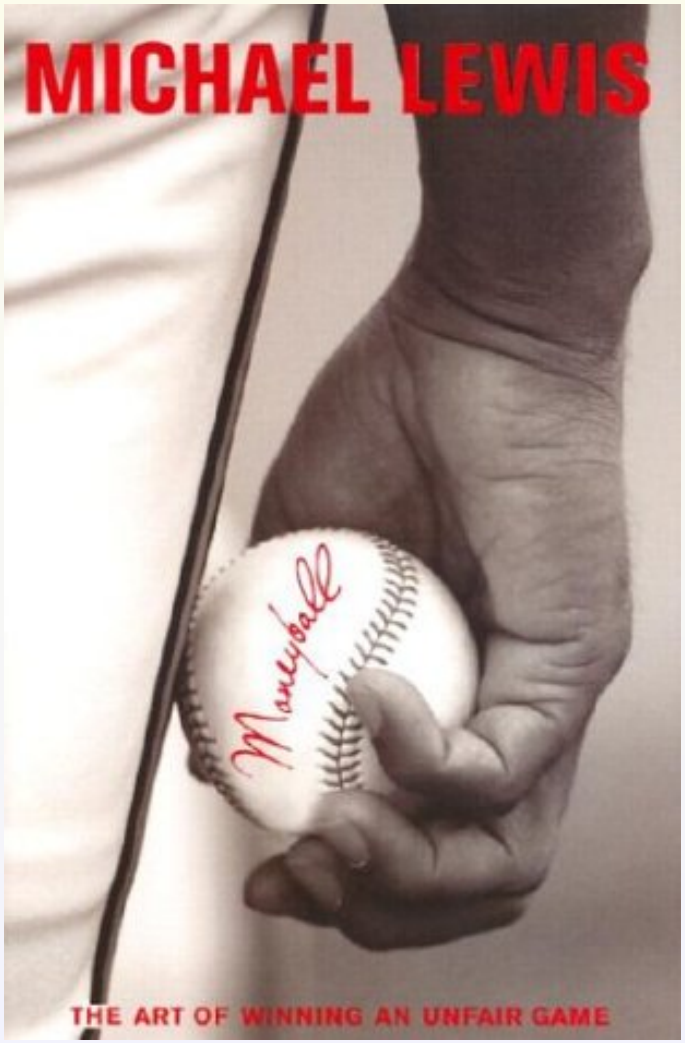


Unofficial Way to Access the Course Folder

`http://cis.jhu.edu/~nhlee/550400.html/`

What is Mathematical Modeling?

Money Ball



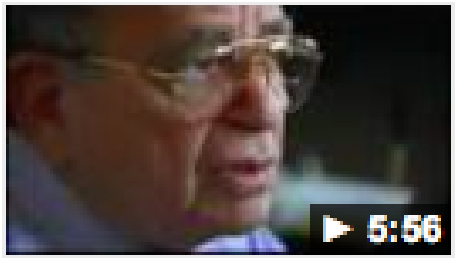
What is Mathematical Modeling?

Trillion Dollar Bet

NOVA Online | Trillion Dollar Bet
www.pbs.org/wgbh/nova/stockmarket/
Welcome to the companion Web site to "**Trillion Dollar Bet**," originally broadcast on February 8, 2000. The film tells the fascinating story of the invention of the ...
The Formula that Shook The ... - Transcripts - A Trader's Lexicon - Resources

Videos for trillion dollar bet - Report videos

Trillion Dollar Bet 1 - YouTube
youtube.com
Jan 8, 2009

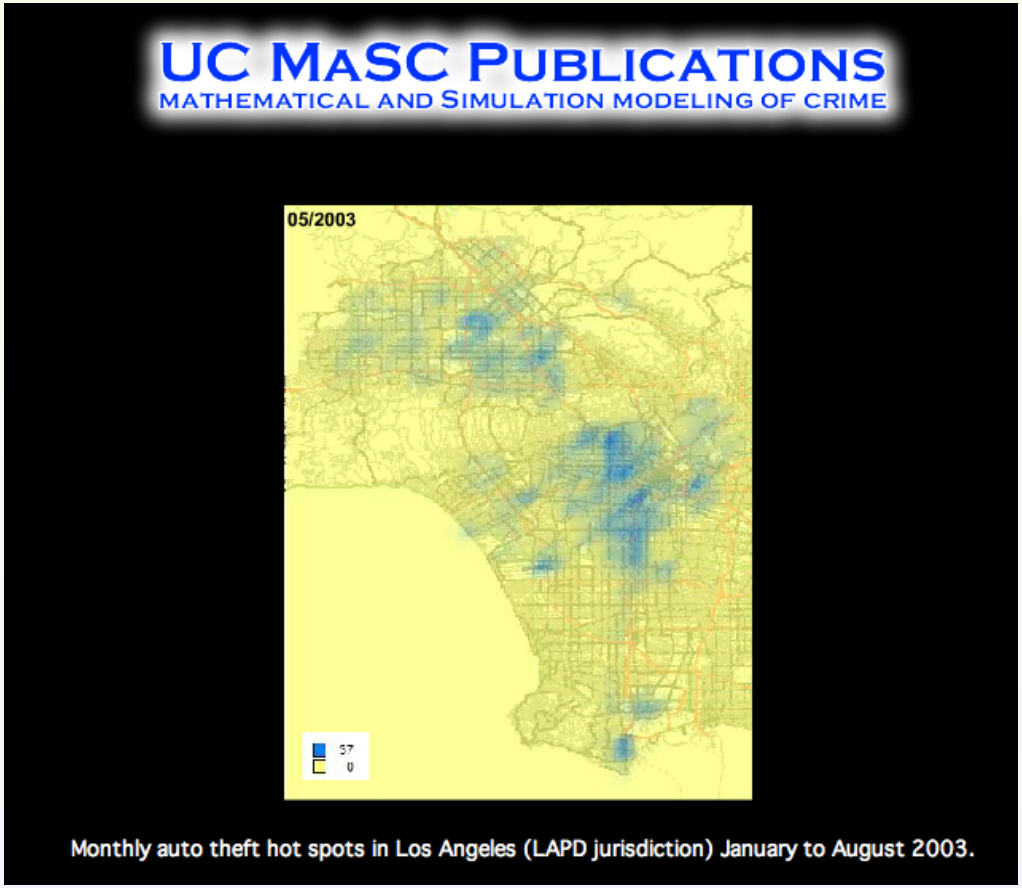


The **Trillion Dollar Bet** - YouTube
youtube.com
Sep 15, 2007

The Midas Formula Stockmarket ...
youtube.com
Aug 23, 2011

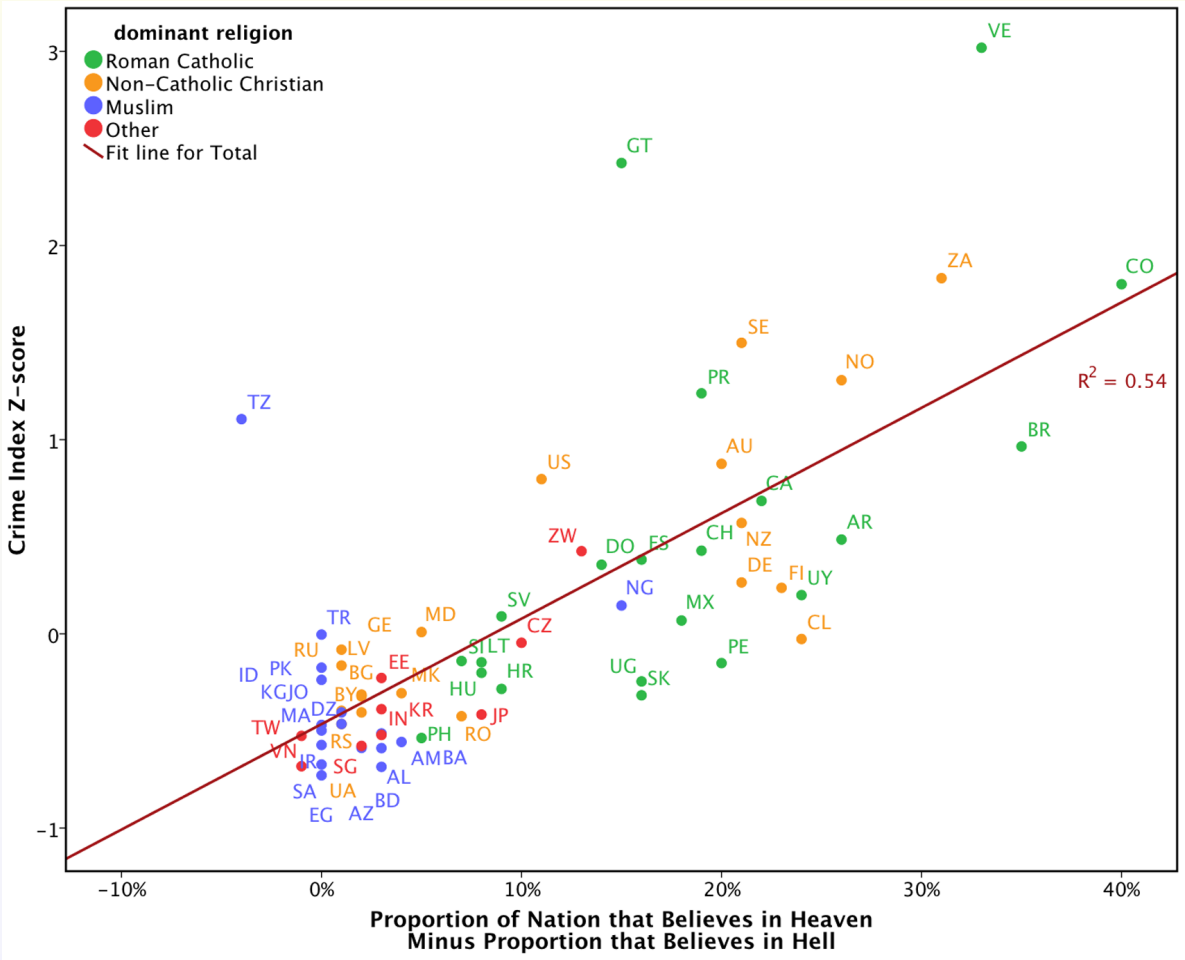
What is Mathematical Modeling?

LAPD Fighting Crime with Math



What is Mathematical Modeling?

Crime rates and religious beliefs



More Project Ideas

http://www.stat.berkeley.edu/

http://www.math.msu.edu/

http://www.mathgoespop.com/

http://www.math.hmc.edu/clinic/

Models and Reality

The ultimate test of a model is how well it performs when it is applied to the problem it was designed to handle.

A model is used, it may lead to incorrect predictions. The model is often modified, frequently discarded, and sometimes used anyway because it is better than nothing. This is the way science develops.

Models and Reality

What makes Mathematical models useful? We must/have/have/have:

- formulate our ideas precisely and so are less likely to let implicit assumptions slip by,
- concise "language" which encourages manipulation,
- a large number of potential theorems available,
- high speed computers available for carrying out calculations.

Properties of Models

As far as a model is concerned, the world can be divided into three parts:

- Things whose effects are neglected,
- Things that affect the model but whose behavior the model is not designed to study,
- Things the model is designed to study the behavior of.

A recurring theme

Frequently Recurring Elements of doing a Project in Industry:

1. Work Statement,
2. Midterm Presentation,
3. Progress Report,
4. Final Presentation,
5. Final Report.

What is Work Statement?

- The written proposal and definition of the project
- Your consulting team's "contract" with the sponsor
- It is ultimately given to the sponsor for review and signature

What is Work Statement?

It sets forth:

- the nature of the project,
- the specific objectives of the project,
- the result expected,
- the "deliverable" for the project.

What is Work Statement?

- The scope of the project must be within the time table for the course
- The deliverables are reasonable and appropriate

What is Work Statement?

- Given the nature of research, it should not include promises that your consulting team cannot be certain to achieve

- It may be necessary after discussion and agreements among various parties to modify and renegotiate the work statement as the project progresses

Example: Insurance Redlining

Insurance Redlining

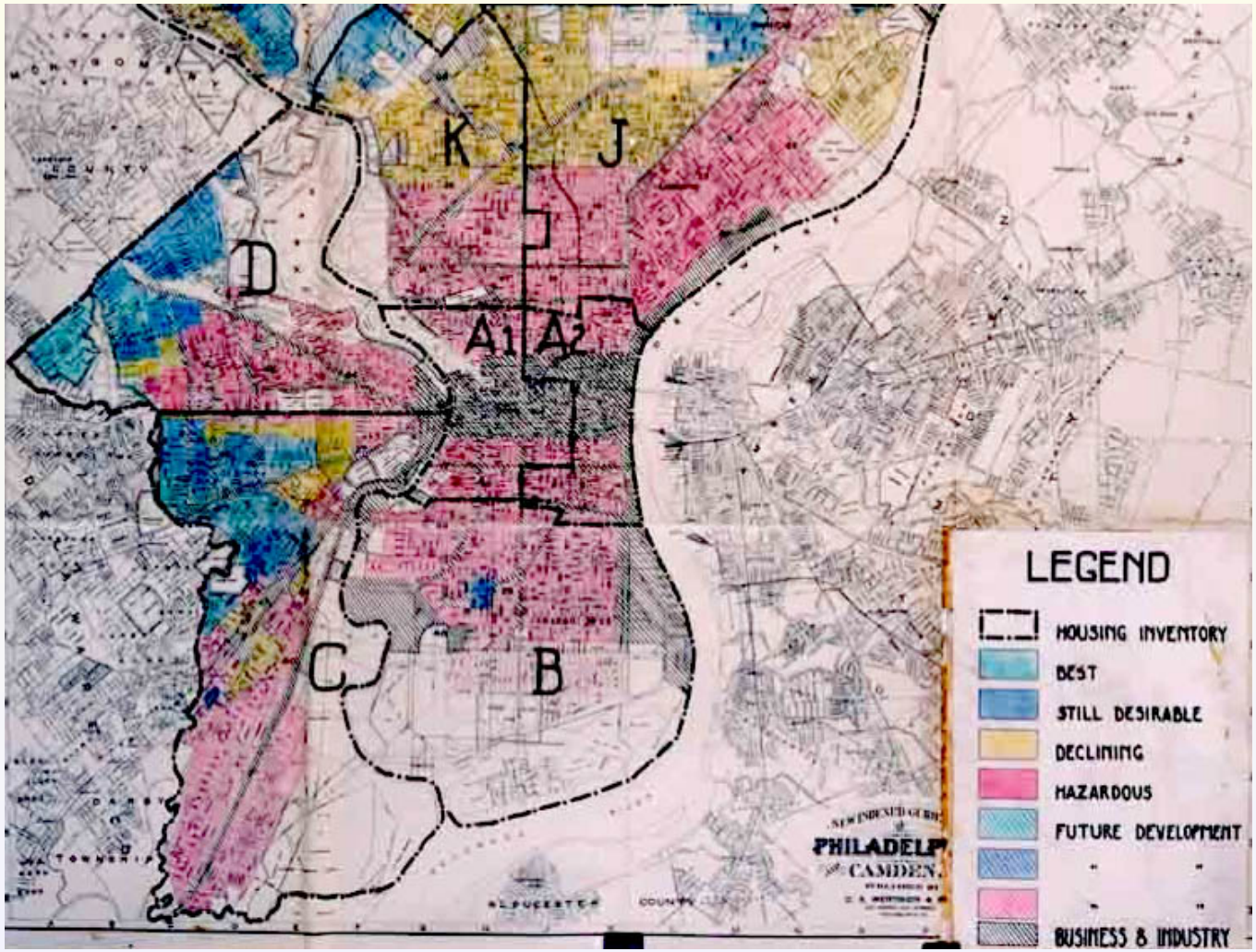
Insurance redlining refers to the practice of refusing to issue insurance to certain types of people or within some geographic area.

FAIR

The FAIR plan was offered by the city of Chicago as a default policy to homeowner who had been rejected by the voluntary market.

Example: Insurance Redlining

Insurance Redlining



Example: Insurance Redlining

Sponsor

The U.S. Commission on Civil Rights examined charges by several Chicago community organizations that insurance companies were redlining their neighborhoods.

Data

The number of FAIR plan policies written and renewed in Chicago by zip code for the number of months of December 1977 through May 1978.

Example: Insurance Redlining

Variables to consider:

`race`	Racial composition in percentage of minority,
`fire`	Fire per 100 housing units,
`theft`	Theft per 100 housing units,
`age`	Theft per 1000 population,
`involact`	New FAIR plan policies and renewal per 100 housing units,
`income`	Median family income in thousands of dollars,
`side`	North or South side of Chicago.

Work Statement: Introduction

Describe:

- the purpose of the project,
- a brief introduction of the sponsoring organization,
- a suitably condensed statement of the problem,
- some discussion of the relevance of the project to the sponsor.

Work Statement: Introduction

The work statement should contain a short description of your sponsor.

For the insurance redlining example, *U.S. Commision on Civil Rights* would be the sponsor.

Boilerplating from the sponsor's webpage is often acceptable.

http://www.usccr.gov

Work Statement: Problem Statement

Can the insurance companies claim that the discrepancy is due to greater risks in some zip codes?

The insurance companies could claim that they were denying insurance in neighborhoods where they had sustained large fire-related losses and any discriminatory effect was a by-product of legitimate business practice.

Work Statement: Timeline & Deliverable

"When I decide the time needed for the project, I first approximate the time that I might actually need, and then, request the sponsor the double of the approximated time."

From Team to Sponsor Presentations, Reports, Special Softwares.

From Sponsor to Team Regular meetings, Data & Contingences, Code & Code Documentation.

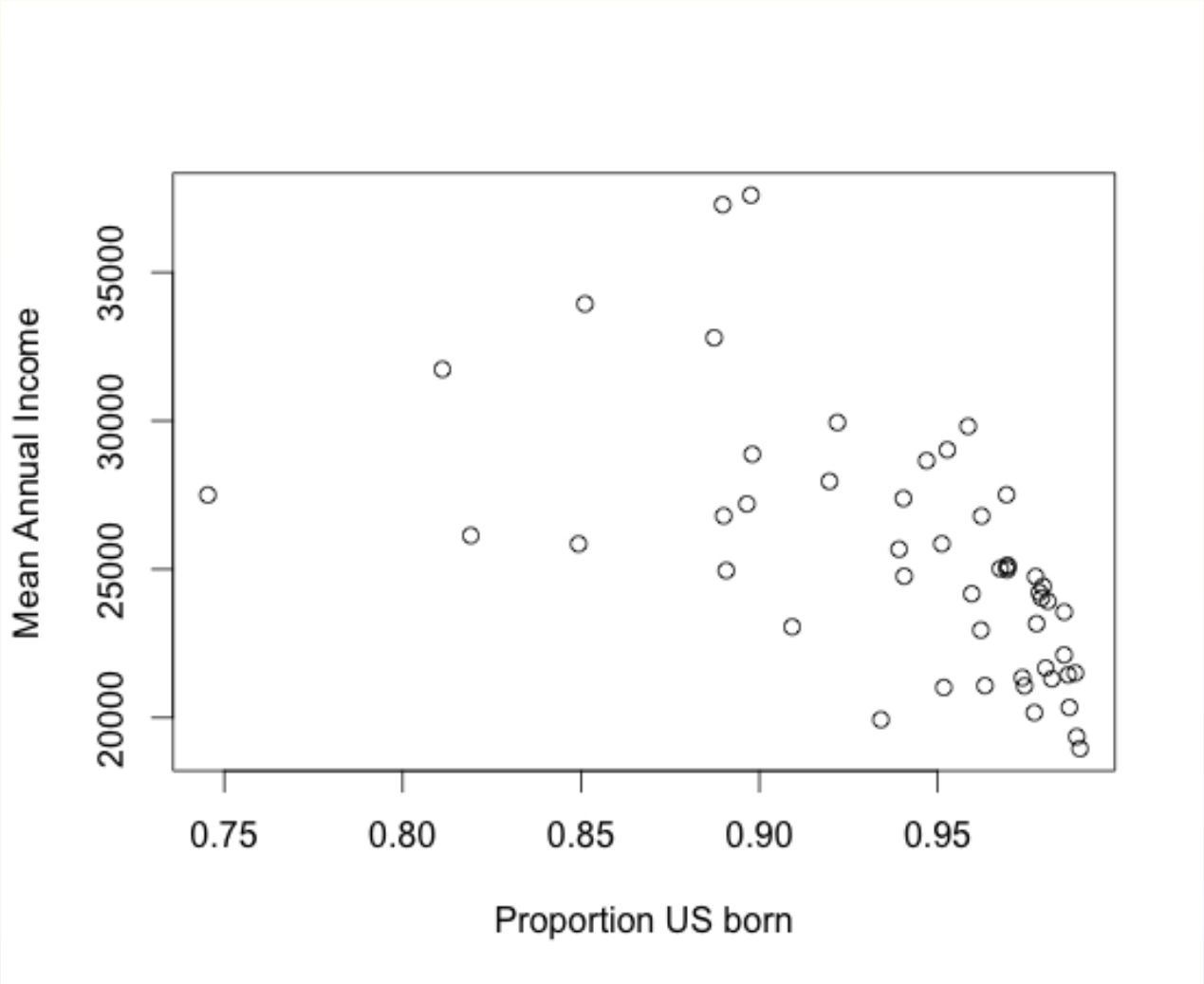
Example: Ecological Fallacy

Ecological Fallacy

When data are collected at the group level, we may observe a correlation between two variables. The ecological fallacy is concluding that the same correlation holds at the individual level.

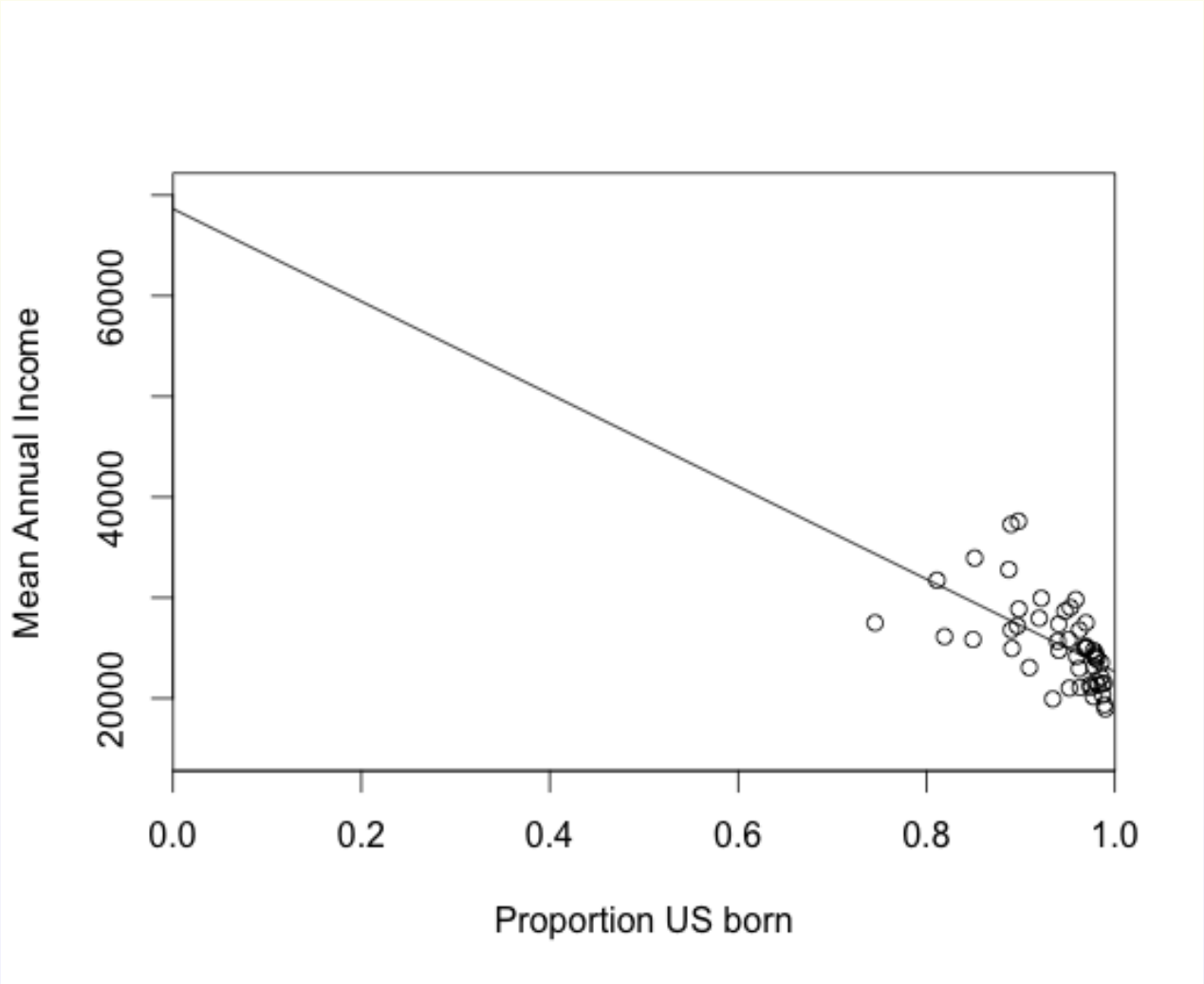
Example: Ecological Fallacy

1998 annual per capita income and proportion U.S. born for 50 states plus D.C.



Example: Ecological Fallacy

1998 annual per capita income and proportion U.S. born for 50 states plus D.C.



Example: Insurance Redlining

For the ecological fallacy example, the assumption would be that the incomes of the native born do not depend on the proportion of native born within the state (and similarly for naturalized citizens).

For the insurance redlining example, we only have aggregate data. We must inform the sponsor that unless more detailed data becomes available, the results for the aggregated data may not hold true at the individual level.

Programmings in this class

- LaTeX :
  - `moderncv`
  - `beamer`
  - `report`
  - `pgf/TikZ`
- Git
  - `git gui`
- R:
  - `lm`
  - `ggplot2`
  - `tikzDevice`
  - `R CMD build`

Where to get some help for LaTeX

http://en.wikibooks.org/wiki/LaTeX/

Tutorial: LaTeX

LaTeX is a computer language for writing a scholarly paper:

Table: HTML vs LaTeX

	HTML	LaTeX

Code	`<html>`
	`. . .`
	`</html>`
		`\begin{document}`
		`. . .`
		`\end{document}`

Compiler	Firefox and etc.	pdflatex and etc.
Output	Web-page	PDF file

Tutorial: LaTeX

TeXworks is:

- an editing tool that is separate from LaTeX,
- available in Linux, OSX and Windows,
- avaiable in:

http://code.google.com/p/texworks/

Tutorial: LaTeX

- Demo on preparing a resume using LaTeX `moderncv` package:
  - Install LaTeX (MikTeX in Windows and MacTeX in OSX),
  - Download `moderncv` package files from the course folder,
  - Change file names to reflect you,
  - Edit the TeX file,
  - Compile using your favorite LaTeX editor,
  - Look at the resulting PDF file.

Demo: LaTeX

Obama's resume

Cautions: LaTeX

There are numerous quirky LaTeX rules:

- opening quotation is not the same as the closing quotation,
- period yields *two* blank spaces,
- for %, need to type \%,
- for \, need to type `\textbackslash`,
- for /, need to type /,
- for {, need to type \{,
- for $, need to type \$,
- ~ yields a single blank space,
- and etc.

The place to get some Git helps

http://git-scm.com/doc/

Demo: LaTeX + Git

The Blind Men and the Elephant

- Start up a git folder
- Create and edit the `.gitignore` file
- Download the template for a beamer file
- Look up the poem from the book
- One slide per stanza
- Compile after each stanza
- Commit after creating each stanza
- Repeat until done.

Tutorial: Git

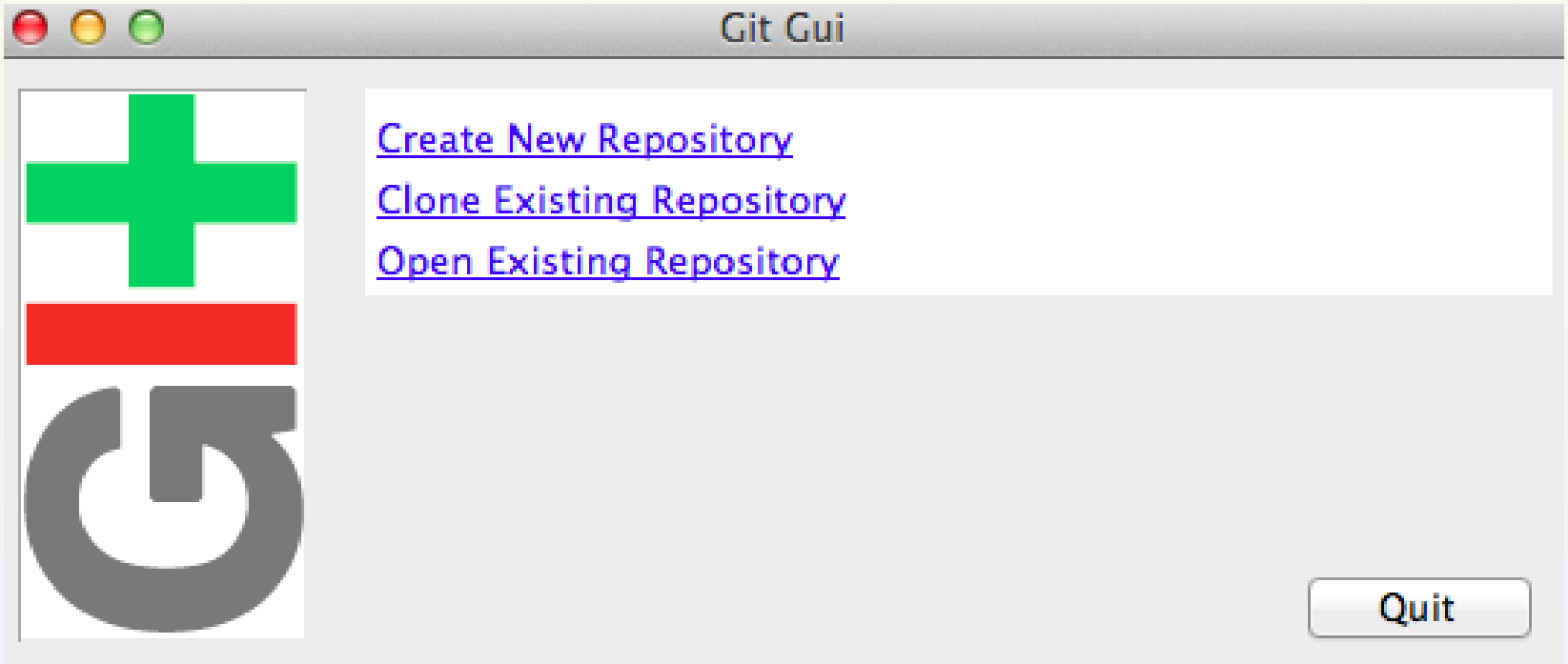
```
sudo apt-get install git
```

An alternative: `git gui`

Tutorial: Git

```
cd ~/
git clone http://cis.jhu.edu/~nhlee/550400.git
```

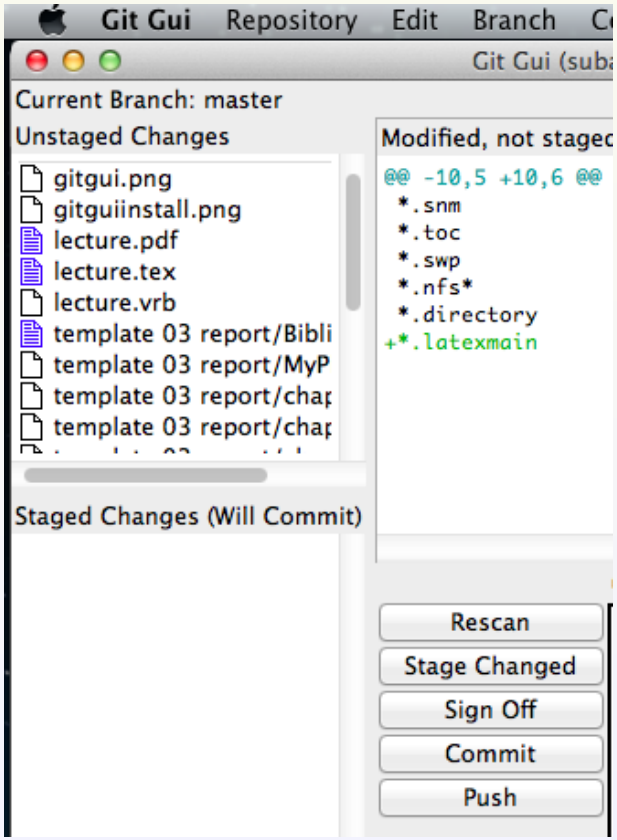
An alternative: `git gui`



Tutorial: Git

```
cd ~/550400.git
git reset --hard HEAD
git pull origin master
```

An alternative: `git gui`

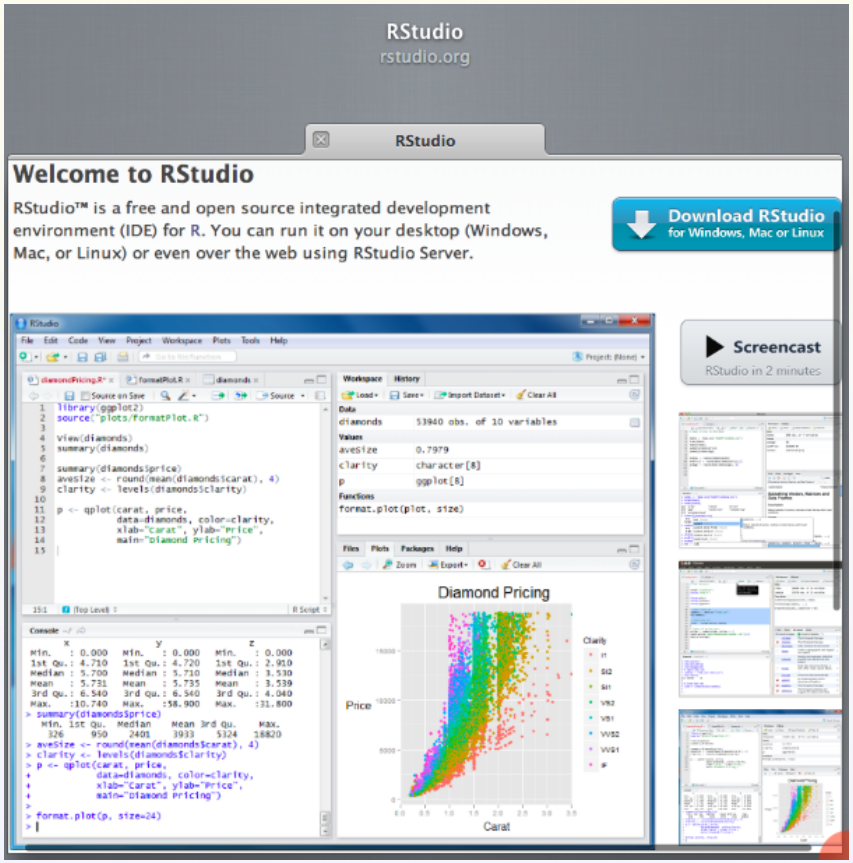


Tutorial: Git

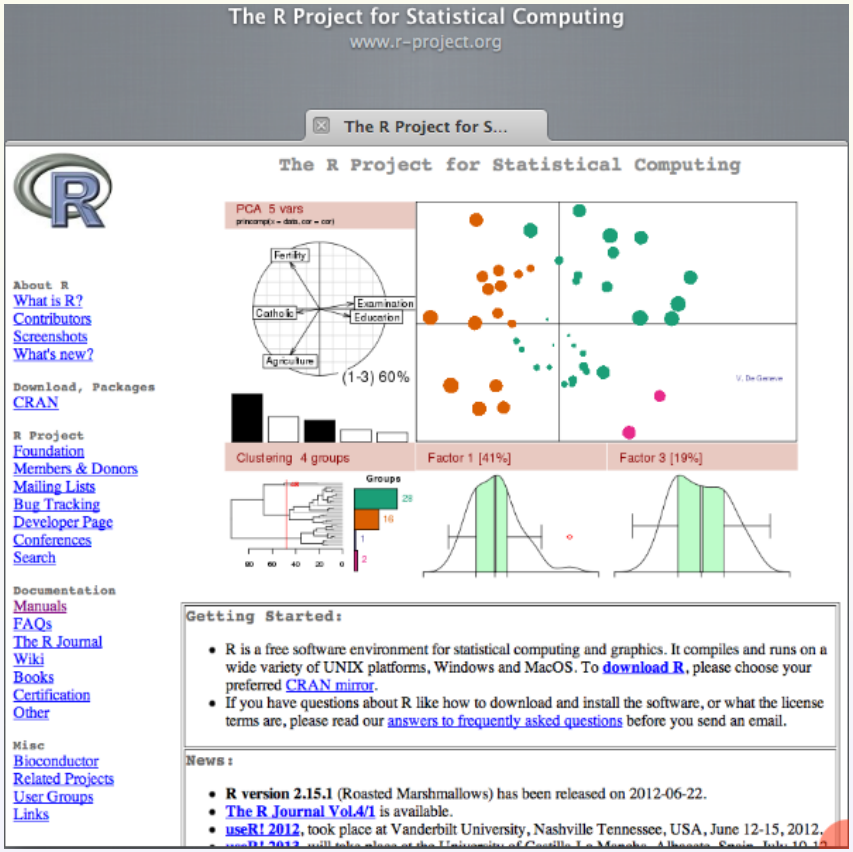
- Demo I: build a *personal* Git folder
  - Create some files
  - Stage the files
  - Commit the files
- Demo II: build the *course* Git folder

Demo: R

R Studio

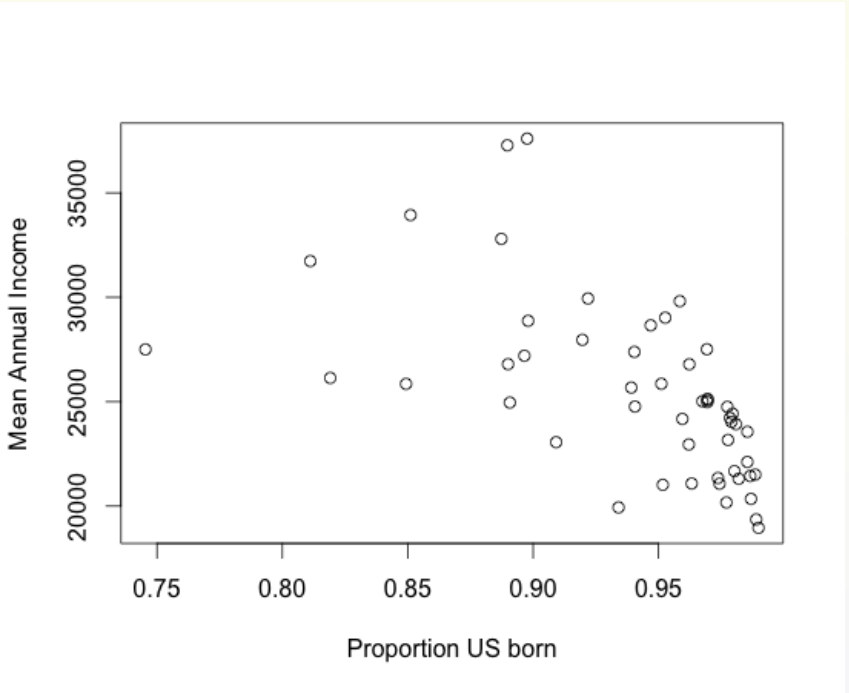


R



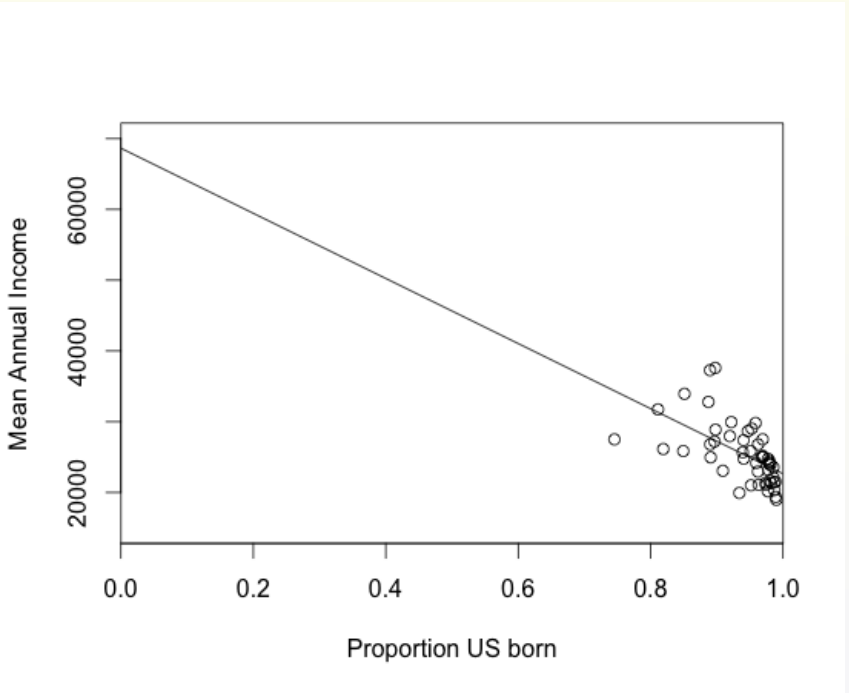
Demo: R

```
install.packages(faraway)
require(faraway)
data(eco)
plot(income ~ usborn,
     data=eco,
     xlab='Proportion US born'
     ylab='Mean Annual Income'
     )
```



Demo: R

```
g <- lm(income ~ usborn, eco)
summary(g)
plot(income ~ usborn,
     data = eco,
     xlab='Proportion US born',
     ylab='Mean Annual Income',
     xlim=c(0,1),
     ylim=c(15000,70000),
     xaxs='i')
abline(coef(g))
```



Example: Insurance Redlining

```
data(chredlin);
chredlin;
head(chredlin);
summary(chredlin);
par(mfrow=c(1,1));
pairs(chredlin);
summary(lm(involact ~ race, chredlin));
plot(involact ~ race, chredlin);
abline(lm(involact ~ race, chredlin));
plot(fire ~ race, chredlin);
abline(lm(fire ~ race, chredlin));
```

Seven Basic Principles

1. Set the context
2. Choose effective examples and analogies
3. Choose vocabulary to suit your readers
4. Decide whether to present #s in text, tables, or figures
5. Report and interpret #s in the text
6. Specify the direction *and* size of an association between variables
7. For many #s, summarize overall pattern

Creating Effective Tables

Example: Cost of Packaging

Example: The Nuclear Mission Arms Race

Example: Maintaining Inventory