

550.400: Mathematical Modeling and Consulting

Lecture Notes

Instructor:

Dr. N. H. Lee

JHU AMS 2012 FALL

Last Compiled on September 12, 2012

Outline

Class Info.

Theory

- Writing about numbers

- Math. Modeling

- Course Project

Example

- Random Bits

- Insurance Redlining

- Sherlock Holmes and the Bicycle Tracks

Tutorials

- L^AT_EX

- Git

- Vim

- R

Syllabus

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

OH Location

Clark Hall 320B

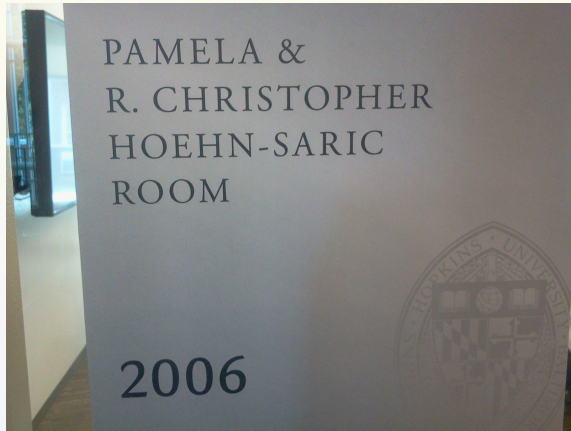


Course Book Reserve

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/`

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

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

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

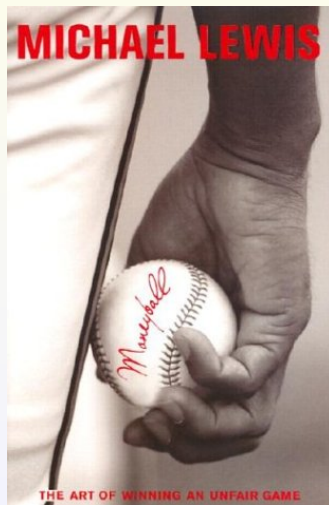
“When I decide the time needed for the project, I first approximate the time that I might actually need, and then, request the sponsor the double of the approximated time.”

From Team to Sponsor Presentations, Reports, Special Softwares.

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

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



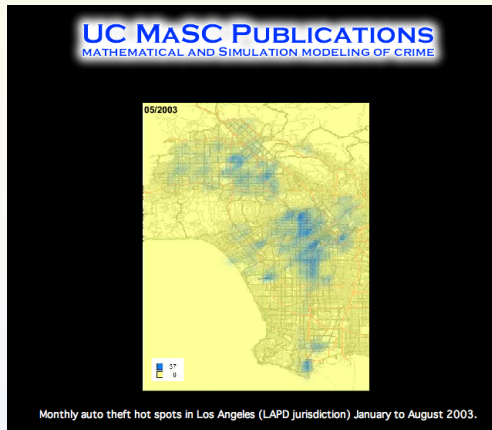
[The Midas Formula Stockmarket ...](#)

youtube.com

Aug 23, 2011

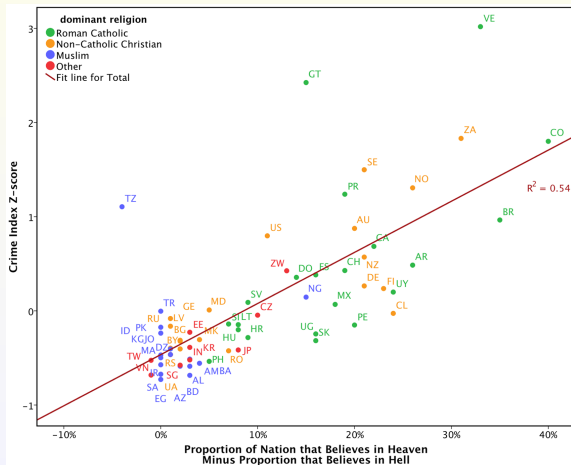
What is Mathematical Modeling?

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/>

Example: Insurance Redlining

Insurance Redlining

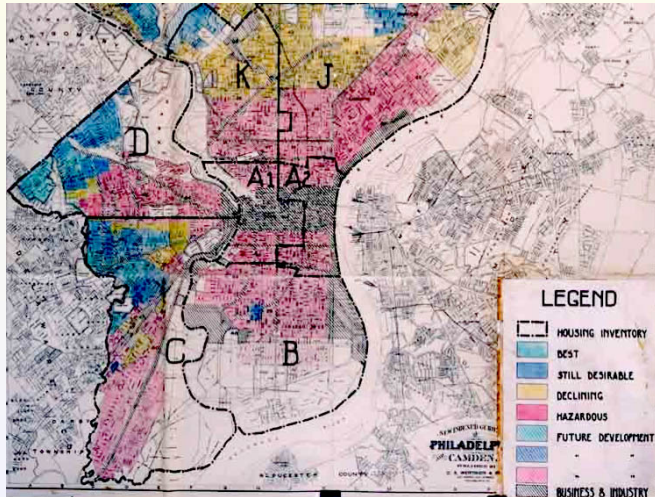
Insurance redlining refers to the practice of refusing to issue insurance to certain types of people or within some geographic area.

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 1000 population,
- `age` percent of housing unit built before 1939,
- `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.

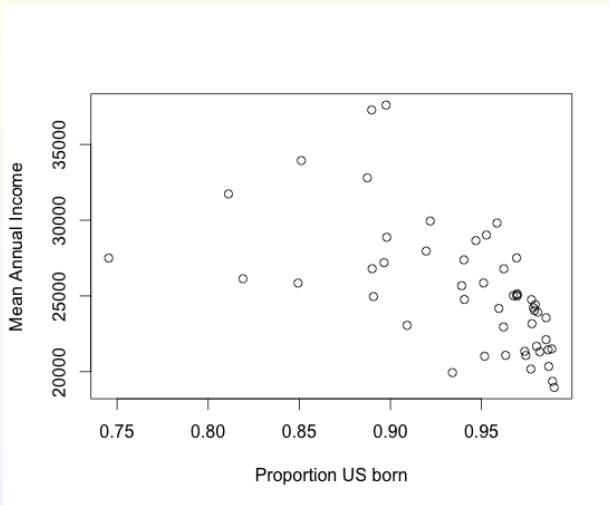
Example: Ecological Fallacy

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.*

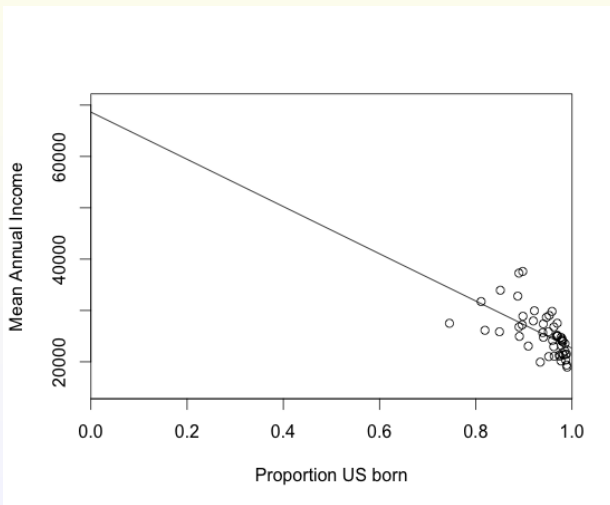
Example: Ecological Fallacy

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.

Example: Sherlock Holmes and the Bicycle Tracks

“This track, as you perceive, was made by a rider who was going from the direction of the school.”

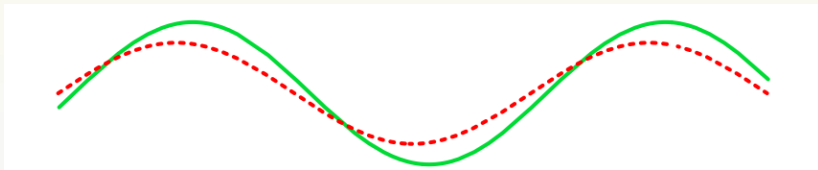
“Or Toward it?”

“No, no, my dear Watson. The more deeply sunk impression is, of course, the hind wheel, upon which the weight rests. You perceive several places where it has passed across and obliterated the more shallow mark of the front one. It was undoubtedly heading away from the school.”

– *The Adventure of the Priory School* by Arthur Conan Doyle

Example: Sherlock Holmes and the Bicycle Tracks

Which one is the front wheel?



Example: Sherlock Holmes and the Bicycle Tracks

$$f_x(t) = r_x(t) + \frac{L}{\sqrt{1 + (r'_y(t)/r'_x(t))^2}}$$

$$f_y(t) = r_y(t) + \frac{Lr'_y(t)/r'_x(t)}{\sqrt{1 + (r'_y(t)/r'_x(t))^2}}$$

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: L^AT_EX

L^AT_EX is a computer language for writing a scholarly paper:

Table: HTML vs L^AT_EX

	HTML	L ^A T _E X
Code	<pre><html> . . . </html></pre>	<pre>\begin{document} . . . \end{document}</pre>
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,
- available in:

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

Tutorial: L^AT_EX

- Demo on preparing a resume using L^AT_EX moderncv package:
 - Install L^AT_EX (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 L^AT_EX editor,
 - Look at the resulting PDF file.

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: L^AT_EX + Git

The Blind Men and the Elephant

Exercise :

- 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,
- Use `verse` environment,
- Compile after each stanza,
- Commit after creating each stanza,
- Repeat until done.

Tutorial: Git

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

An alternative: `git gui`

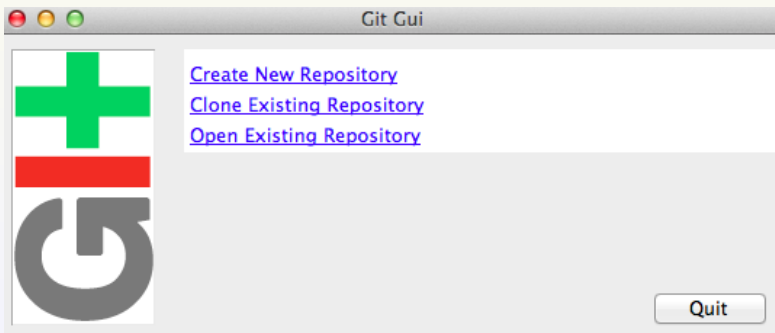


The screenshot shows the Git website homepage. At the top, there is a red diamond logo with a white branching diagram, followed by the word "git" in a large, bold, black font. To the right of the logo, the tagline "--distributed-is-the-new-centralized" is written in a smaller, grey font. Below the logo and tagline, there is a navigation menu with the following items: "About", "Documentation", "Downloads" (in red), "GUI Clients" (in red), "Logos", and "Community". To the right of the navigation menu, there is a white box with a grey border. Inside this box, the heading "GUI Clients" is displayed in a large, bold, black font. Below the heading, the text "Git comes with built-in GUI tools for comm" and "third-party tools for users looking for plat" is visible. At the bottom of the box, there is a button labeled "Show GUIs for all OSes" and a red text element that says "7 Mac GUIs a".

Tutorial: Git

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

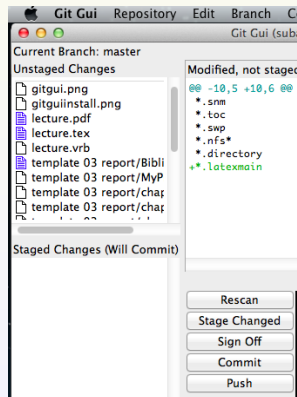
An alternative: git gui



Tutorial: Git

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

An alternative: git gui



Tutorial: Git

After years of using git, you might find this funny:



```
git pull origin master
```


Tutorial: Git

After years of using git, you might find this funny:



```
git push origin master
```

Tutorial: Git

For \$19.99, you can also have your own:



```
cd ~/
mkdir hub.git
mkdir computerA.git
mkdir computerB.git
```

```
git init --bare hub.git
```

```
cd hub.git
cd hooks
cp post-update.sample post-update
```

Tutorial: Git

```
cd computerA.git
git init
git remote add origin ~/hub.git
cat Hello >> commonfile.txt
git add commonfile.txt
git pull origin master
git commit -am 'from A'
git push origin master
```

```
cd computerB.git
git init
git remote add origin ~/hub.git
cat World >> fileB.txt
git add commonfile.txt
git pull origin master
git commit -am 'from B'
git push origin master
```

Tutorial: Git

```
cd ~/550400
```

```
git gui
```

```
git reset --hard HEAD
```

```
git branch personal
```

```
git branch
```

```
git checkout personal
```

```
edit some file
```

```
git status
```

```
git add .
```

```
git commit -am 'personal edit'
```

```
git checkout master
```

```
git branch -D personal
```

- checks if there has been any change to the folder
- build and update the master git branch
- create and update a personal git branch

Tutorial: Git

.gitignore?

- N.B. the course folder already has one
- Use it to let *git* know the files to *ignore* while version controlling
- one particular usage: create `.gitignore` at the root of your git folder
- files already been list under the git watch list will not be ignored even after creation of `.gitignore`

Vim

Vim

Vim is a highly customizable text editor

1. \LaTeX , R, C/C++, Java, Python, Git and etc.
2. Regular expression, syntax coloring, autocompletion
3. Try Firefox + Wasavi/Vimperator/Vimium
4. <ESC>-mode
 - :-mode, aka., the last line mode
 - i-mode, aka., the insert mode

Demo: R + \LaTeX

R Studio

R

RStudio
rstudio.org

Welcome to RStudio

RStudio™ is a free and open source integrated development environment (IDE) for R. You can run it on your desktop (Windows, Mac, or Linux) or even over the web using RStudio Server.

[Download RStudio for Windows, Mac or Linux](#)

[Screencast RStudio in 2 minutes](#)

```
R
# Welcome to R!
# Learn more about RStudio at rstudio.org
# Learn more about R at r-project.org

# Create a new R script
newScript("NewScript.R")

# View the console
viewConsole()

# Run the script
runScript("NewScript.R")

# Split the editor into panes
splitScreen()

# Change the background color
setTheme("rstudio")
```

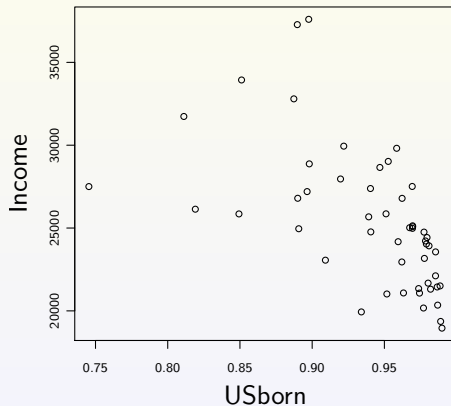
Diamond Pricing

Price vs Clarity

Clarity	Price
I1	~1000
I2	~1500
I3	~2000
I4	~2500
I5	~3000
I6	~3500
I7	~4000
I8	~4500
I9	~5000
I10	~5500
I11	~6000
I12	~6500
I13	~7000
I14	~7500
I15	~8000
I16	~8500
I17	~9000
I18	~9500
I19	~10000
I20	~10500
I21	~11000
I22	~11500
I23	~12000
I24	~12500
I25	~13000
I26	~13500
I27	~14000
I28	~14500
I29	~15000
I30	~15500
I31	~16000
I32	~16500
I33	~17000
I34	~17500
I35	~18000
I36	~18500
I37	~19000
I38	~19500
I39	~20000
I40	~20500
I41	~21000
I42	~21500
I43	~22000
I44	~22500
I45	~23000
I46	~23500
I47	~24000
I48	~24500
I49	~25000
I50	~25500
I51	~26000
I52	~26500
I53	~27000
I54	~27500
I55	~28000
I56	~28500
I57	~29000
I58	~29500
I59	~30000
I60	~30500
I61	~31000
I62	~31500
I63	~32000
I64	~32500
I65	~33000
I66	~33500
I67	~34000
I68	~34500
I69	~35000
I70	~35500
I71	~36000
I72	~36500
I73	~37000
I74	~37500
I75	~38000
I76	~38500
I77	~39000
I78	~39500
I79	~40000
I80	~40500
I81	~41000
I82	~41500
I83	~42000
I84	~42500
I85	~43000
I86	~43500
I87	~44000
I88	~44500
I89	~45000
I90	~45500
I91	~46000
I92	~46500
I93	~47000
I94	~47500
I95	~48000
I96	~48500
I97	~49000
I98	~49500
I99	~50000
I100	~50500
I101	~51000
I102	~51500
I103	~52000
I104	~52500
I105	~53000
I106	~53500
I107	~54000
I108	~54500
I109	~55000
I110	~55500
I111	~56000
I112	~56500
I113	~57000
I114	~57500
I115	~58000
I116	~58500
I117	~59000
I118	~59500
I119	~60000
I120	~60500
I121	~61000
I122	~61500
I123	~62000
I124	~62500
I125	~63000
I126	~63500
I127	~64000
I128	~64500
I129	~65000
I130	~65500
I131	~66000
I132	~66500
I133	~67000
I134	~67500
I135	~68000
I136	~68500
I137	~69000
I138	~69500
I139	~70000
I140	~70500
I141	~71000
I142	~71500
I143	~72000
I144	~72500
I145	~73000
I146	~73500
I147	~74000
I148	~74500
I149	~75000
I150	~75500
I151	~76000
I152	~76500
I153	~77000
I154	~77500
I155	~78000
I156	~78500
I157	~79000
I158	~79500
I159	~80000
I160	~80500
I161	~81000
I162	~81500
I163	~82000
I164	~82500
I165	~83000
I166	~83500
I167	~84000
I168	~84500
I169	~85000
I170	~85500
I171	~86000
I172	~86500
I173	~87000
I174	~87500
I175	~88000
I176	~88500
I177	~89000
I178	~89500
I179	~90000
I180	~90500
I181	~91000
I182	~91500
I183	~92000
I184	~92500
I185	~93000
I186	~93500
I187	~94000
I188	~94500
I189	~95000
I190	~95500
I191	~96000
I192	~96500
I193	~97000
I194	~97500
I195	~98000
I196	~98500
I197	~99000
I198	~99500
I199	~100000
I200	~100500
I201	~101000
I202	~101500
I203	~102000
I204	~102500
I205	~103000
I206	~103500
I207	~104000
I208	~104500
I209	~105000
I210	~105500
I211	~106000
I212	~106500
I213	~107000
I214	~107500
I215	~108000
I216	~108500
I217	~109000
I218	~109500
I219	~110000
I220	~110500
I221	~111000
I222	~111500
I223	~112000
I224	~112500
I225	~113000
I226	~113500
I227	~114000
I228	~114500
I229	~115000
I230	~115500
I231	~116000
I232	~116500
I233	~117000
I234	~117500
I235	~118000
I236	~118500
I237	~119000
I238	~119500
I239	~120000
I240	~120500
I241	~121000
I242	~121500
I243	~122000
I244	~122500
I245	~123000
I246	~123500
I247	~124000
I248	~124500
I249	~125000
I250	~125500
I251	~126000
I252	~126500
I253	~127000
I254	~127500
I255	~128000
I256	~128500
I257	~129000
I258	~129500
I259	~130000
I260	~130500
I261	~131000
I262	~131500
I263	~132000
I264	~132500
I265	~133000
I266	~133500
I267	~134000
I268	~134500
I269	~135000
I270	~135500
I271	~136000
I272	~136500
I273	~137000
I274	~137500
I275	~138000
I276	~138500
I277	~139000
I278	~139500
I279	~140000
I280	~140500
I281	~141000
I282	~141500
I283	~142000
I284	~142500
I285	~143000
I286	~143500
I287	~144000
I288	~144500
I289	~145000
I290	~145500
I291	~146000
I292	~146500
I293	~147000
I294	~147500
I295	~148000
I296	~148500
I297	~149000
I298	~149500
I299	~150000
I300	~150500
I301	~151000
I302	~151500
I303	~152000
I304	~152500
I305	~153000
I306	~153500
I307	~154000
I308	~154500
I309	~155000
I310	~155500
I311	~156000
I312	~156500
I313	~157000
I314	~157500
I315	~158000
I316	~158500
I317	~159000
I318	~159500
I319	~160000
I320	~160500
I321	~161000
I322	~161500
I323	~162000
I324	~162500
I325	~163000
I326	~163500
I327	~164000
I328	~164500
I329	~165000
I330	~165500
I331	~166000
I332	~166500
I333	~167000
I334	~167500
I335	~168000
I336	~168500
I337	~169000
I338	~169500
I339	~170000
I340	~170500
I341	~171000
I342	~171500
I343	~172000
I344	~172500
I345	~173000
I346	~173500
I347	~174000
I348	~174500
I349	~175000
I350	~175500
I351	~176000
I352	~176500
I353	~177000
I354	~177500
I355	~178000
I356	~178500
I357	~179000
I358	~179500
I359	~180000
I360	~180500
I361	~181000
I362	~181500
I363	~182000
I364	~182500
I365	~183000
I366	~183500
I367	~184000
I368	~184500
I369	~185000
I370	~185500
I371	~186000
I372	~186500
I373	~187000
I374	~187500
I375	~188000
I376	~188500
I377	~189000
I378	~189500
I379	~190000
I380	~190500
I381	~191000
I382	~191500
I383	~192000
I384	~192500
I385	~193000
I386	~193500
I387	~194000
I388	~194500
I389	~195000
I390	~195500
I391	~196000
I392	~196500
I393	~197000
I394	~197500
I395	~198000
I396	~198500
I397	~199000
I398	~199500
I399	~200000
I400	~200500
I401	~201000
I402	~201500
I403	~202000
I404	~202500
I405	~203000
I406	~203500
I407	~204000
I408	~204500
I409	~205000
I410	~205500
I411	~206000
I412	~206500
I413	~207000
I414	~207500

Demo: R + L^AT_EX

```
install.packages(faraway)
install.packages(tikzDevice)
require(faraway)
require(tikzDevice)
data(eco)
tikz('embeddedfig1.tex',
     standAlone=F,
     width=5,height=5)
plot(income ~ usborn,
     data=eco,
     xlab='Proportion US born',
     ylab='Mean Annual Income',
     )
dev.off()
```

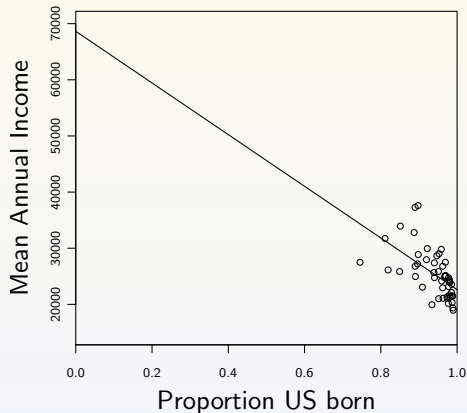


Demo: R + \LaTeX

```

tikz('embeddedfig2.tex',
     standAlone=F,
     width=5,height=5)
plot(income ~ usborn,
     data = eco,
     xlab='Proportion US born',
     ylab='Mean Annual Income',
     xlim=c(0,1),
     ylim=c(15000,70000),
     xaxs='i')
g<-lm(income~usborn,eco)
abline(coef(g))
dev.off()

```



Today's Lesson Plan

1. So, what else is wrong with “Crime rate and religious beliefs”?
2. Sherlock Homes and Bicycle
 - What is wrong with this?
 - What can we do to fix?
 - “Mathify”?
3. How do you update my git course git folder?
 - Doing status-add-commit-pull routine
 - Managing merge-conflicts
 - Branching out and checking in & out
4. Problem Set 2.1, 2.3, 2.5, 2.7, 2.9