Lecture 1: Introduction to Econometrics

ECON 480 - Econometrics - Fall 2018

Ryan Safner

August 27, 2018

HOOD COLLEGE

What is Econometrics?

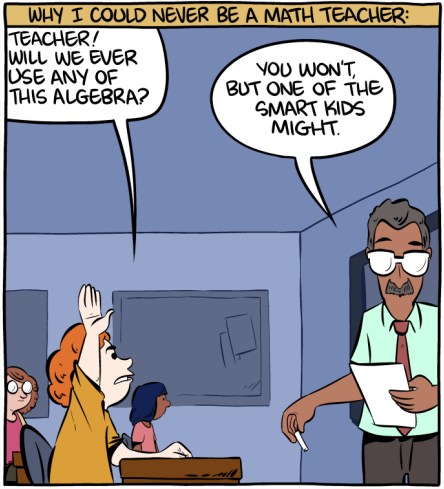
The Basic Empirical Research Process

About the Course

HOOD COLLEGE

What is Econometrics?

Why Everyone, Yes *Everyone*, Should Learn Statistics



SMBC

Why Everyone, Yes *Everyone*, Should Learn Statistics

SMBC

HOOD COLLEGE

We're Not so Good at Statistics: Votes

- Votes in the U.S. House of Representatives in favor of passing the Civil Rights Act of 1964:

Democrat	Republican
61%	80%

HOOD COLLEGE

We're Not so Good at Statistics: Votes

- Votes in the U.S. House of Representatives in favor of passing the Civil Rights Act of 1964:

Democrat	Republican
61%	80%

- Simple enough: "on average, Republicans tended to vote for passage more than Democrats"

HOOD COLLEGE

We're Not so Good at Statistics: Votes II

- Suppose we break it down further between Dems vs. Reps from Northern vs. Southern states:

	Democrat	Republican
North	94% (145/154)	85% (138/162)
South	7% (7/94)	0% (0/10)
Overall	61% (152/248)	80% (138/172)

HOOD COLLEGE

We're Not so Good at Statistics: Votes II

- Suppose we break it down further between Dems vs. Reps from Northern vs. Southern states:

	Democrat	Republican
North	94% (145/154)	85% (138/162)
South	7% (7/94)	0% (0/10)
Overall	61% (152/248)	80% (138/172)

- In *both* North and South, Democrats are more likely to vote for passage than Republicans, despite a higher proportion of Republicans *overall* voted for the act than Democrats!

HOOD COLLEGE

We're Not so Good at Statistics: Votes III

	Democrat	Republican
North	94% (145/154)	85% (138/162)
South	7% (7/94)	0% (0/10)
Overall	61% (152/248)	80% (138/172)

- A far greater proportion of Democrats at the time were from the South (94/248, 38%) vs. Republicans (10/172, 6%)

HOOD COLLEGE

We're Not so Good at Statistics: Votes III

	Democrat	Republican
North	94% (145/154)	85% (138/162)
South	7% (7/94)	0% (0/10)
Overall	61% (152/248)	80% (138/172)

- A far greater proportion of Democrats at the time were from the South (94/248, 38%) vs. Republicans (10/172, 6%)
- The 7% of southern Democrats voting for the act dragged down the Democrats' overall percentage much more than the 0% of southern Republicans

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones

- Suppose you suffer from kidney stones, your doctor offers you *treatment A* vs. *treatment B*

We're Not so Good at Statistics: Kidney Stones

- Suppose you suffer from kidney stones, your doctor offers you *treatment A* vs. *treatment B*
- In clinical trials, treatment A was effective for a higher overall percentage of patients than treatment B

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones

- Suppose you suffer from kidney stones, your doctor offers you *treatment A* vs. *treatment B*
- In clinical trials, treatment A was effective for a higher overall percentage of patients than treatment B
- But, treatment B was effective for a higher percentage of patients in *both* groups (large and small stones)

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones

- Suppose you suffer from kidney stones, your doctor offers you *treatment A* vs. *treatment B*
- In clinical trials, treatment A was effective for a higher overall percentage of patients than treatment B
- But, treatment B was effective for a higher percentage of patients in *both* groups (large and small stones)
- How can this be?

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones II

- From a real medical study:

	Treatment A	Treatment B
Small Stones	93% (81/87)	87% (234/270)
Large Stones	73% (192/263)	69% (55/80)
Overall	78% (273/350)	83% (289/350)

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones II

- From a real medical study:

	Treatment A	Treatment B
Small Stones	93% (81/87)	87% (234/270)
Large Stones	73% (192/263)	69% (55/80)
Overall	78% (273/350)	83% (289/350)

- The sizes of the two groups (i.e. who gets A vs B) are *very* different

HOOD COLLEGE

We're Not so Good at Statistics: Kidney Stones II

- From a real medical study:

	Treatment A	Treatment B
Small Stones	93% (81/87)	87% (234/270)
Large Stones	73% (192/263)	69% (55/80)
Overall	78% (273/350)	83% (289/350)

- The sizes of the two groups (i.e. who gets A vs B) are *very* different
- A **lurking variable** in the study is the *severity* of the case: doctors tended to give *B* for less severe cases

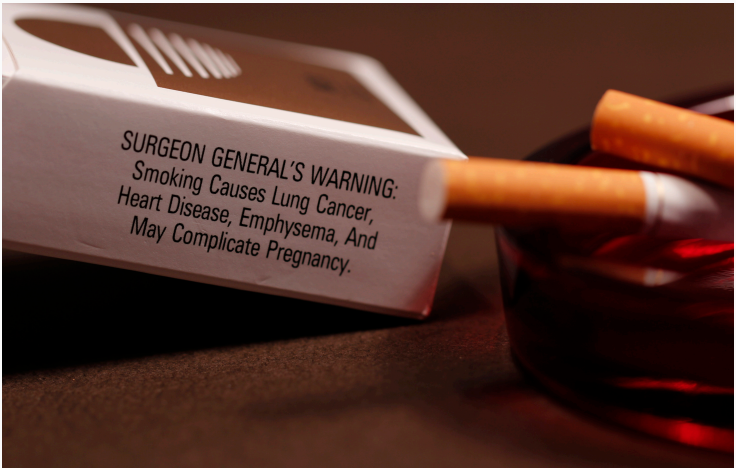
HOOD COLLEGE

Simpson's Paradox

The correlation between two variables can change (even reverse!) when additional variables are considered

We're Not so Good at Statistics: Smoking

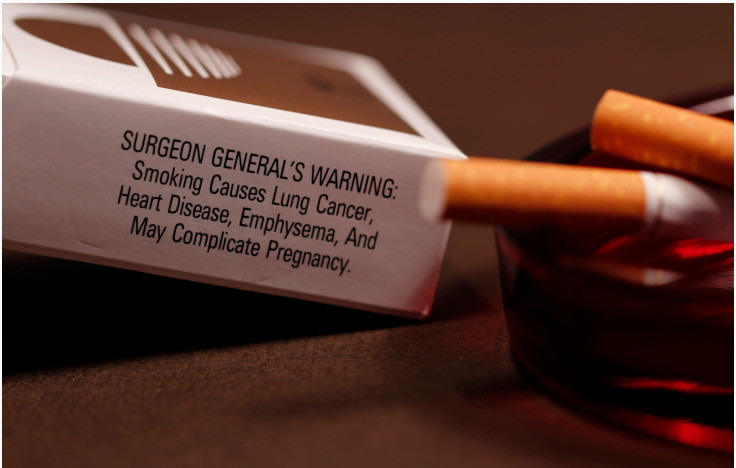
- 1964: U.S. Surgeon General issued a report claiming that cigarette smoking causes lung cancer



HOOD COLLEGE

We're Not so Good at Statistics: Smoking

- 1964: U.S. Surgeon General issued a report claiming that cigarette smoking causes lung cancer
- Evidence based primarily on *correlations* between cigarette smoking and lung cancer



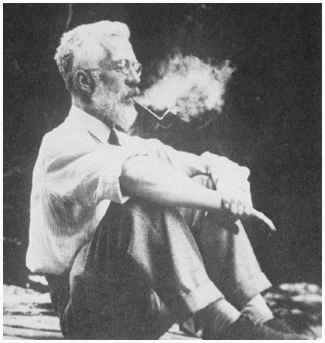
HOOD COLLEGE

We're Not so Good at Statistics: Smoking II

- Tobacco companies attacked the report, naturally

We're Not so Good at Statistics: Smoking III

- But so did Ronald Fisher, the "father of modern statistics"

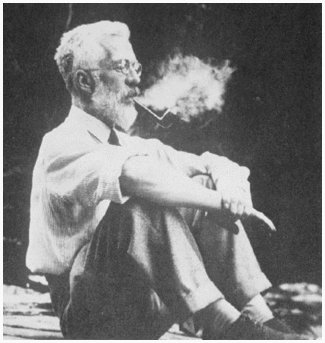


Ronald A. Fisher

1890–1924

HOOD COLLEGE

- But so did Ronald Fisher, the "father of modern statistics"
- There could be a lurking variable (genetics?) that causes *both* lung cancer *and* people wanting to smoke (i.e. nicotine craving)



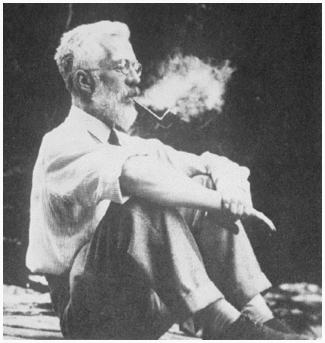
Ronald A. Fisher

1890–1924

HOOD COLLEGE

We're Not so Good at Statistics: Smoking III

- But so did Ronald Fisher, the "father of modern statistics"
- There could be a lurking variable (genetics?) that causes *both* lung cancer *and* people wanting to smoke (i.e. nicotine craving)
- If true: despite correlation between smoking and lung cancner, decision to smoke or not would have *no impact* on whether or not you get lung cancer!



Ronald A. Fisher

1890–1924

We're Not so Good at Statistics: Correlation vs. Causation

- The goal of statistics professors everywhere is for you to learn one thing above all else:

Correlation does not imply causation

We're Not so Good at Statistics: Correlation vs. Causation

- The goal of statistics professors everywhere is for you to learn one thing above all else:

Correlation does not imply causation

- Can NOT conclude *from correlation alone* that *X causes Y*

We're Not so Good at Statistics: Correlation vs. Causation

- The goal of statistics professors everywhere is for you to learn one thing above all else:

Correlation does not imply causation

- Can NOT conclude *from correlation alone* that *X causes Y*
- May be other confounding variables (e.g. *Z*, etc) that affect *X* and/or *Y*

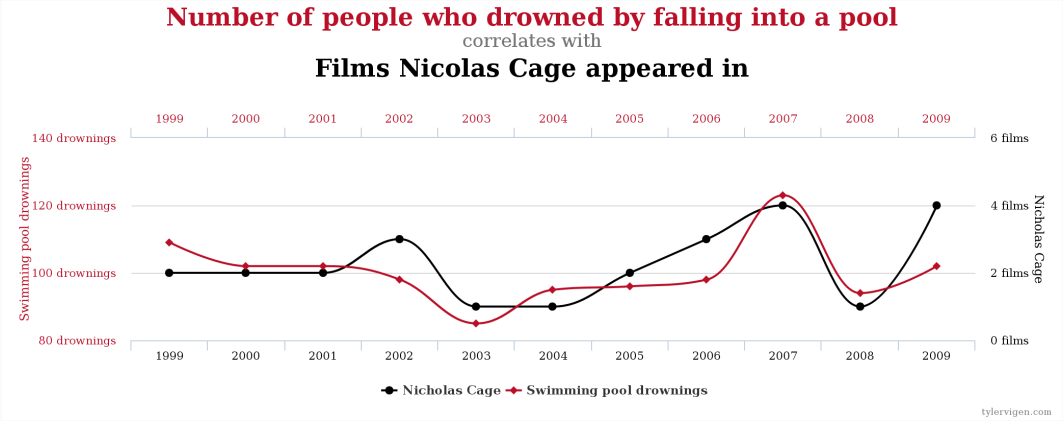
We're Not so Good at Statistics: Correlation vs. Causation

- The goal of statistics professors everywhere is for you to learn one thing above all else:

Correlation does not imply causation

- Can NOT conclude *from correlation alone* that *X causes Y*
- May be other confounding variables (e.g. *Z*, etc) that affect *X* and/or *Y*
- May be reverse causation (maybe *Y* causes *X*, or both cause each other)

We're Not so Good at Statistics: Correlation vs. Causation I



- Now you might think that we can never *really* say what causes something...

- Now you might think that we can never *really* say what causes something...
- It's always good to be skeptical, *but*

We're Not so Good at Statistics

- Now you might think that we can never *really* say what causes something...
- It's always good to be skeptical, *but*
- This is where econometrics comes in

HOOD COLLEGE

- Econometrics is the application of statistical and other empirical tools to quantify economic relationships in the real world.

Econometrics

- Econometrics is the application of statistical and other empirical tools to quantify economic relationships in the real world.
- Uses real data to:

HOOD COLLEGE

Econometrics

- Econometrics is the application of statistical and other empirical tools to quantify economic relationships in the real world.
- Uses real data to:
    - test economic hypotheses

HOOD COLLEGE

Econometrics

- Econometrics is the application of statistical and other empirical tools to quantify economic relationships in the real world.
- Uses real data to:
  - test economic hypotheses
  - estimate quantitatively the causal effects between economic variables

HOOD COLLEGE

Econometrics

- Econometrics is the application of statistical and other empirical tools to quantify economic relationships in the real world.
- Uses real data to:
  - test economic hypotheses
  - estimate quantitatively the causal effects between economic variables
  - make forecasts of future events

HOOD COLLEGE

- What we *really* use econometrics for is demonstrating evidence for causal relationships between real world entities/forces of interest, and to attempt to *quantify* them, e.g.

- What we *really* use econometrics for is demonstrating evidence for causal relationships between real world entities/forces of interest, and to attempt to *quantify* them, e.g.
  - "For every additional police officer in a metropolitan district, crime rates decrease by 0.12%, on average"

- What we *really* use econometrics for is demonstrating evidence for causal relationships between real world entities/forces of interest, and to attempt to *quantify* them, e.g.
  - "For every additional police officer in a metropolitan district, crime rates decrease by 0.12%, on average"
  - "Each additional year of education tends to increase a person's yearly wages by $2,500, on average"

HOOD COLLEGE

- What we *really* use econometrics for is demonstrating evidence for causal relationships between real world entities/forces of interest, and to attempt to *quantify* them, e.g.
  - "For every additional police officer in a metropolitan district, crime rates decrease by 0.12%, on average"
  - "Each additional year of education tends to increase a person's yearly wages by $2,500, on average"
  - "For every 1% tariffs are lowered, a country's GDP growth rate rises by 0.025%, on average"

HOOD COLLEGE

Example
Does reducing class size improve elementary school education?

- Claim: with smaller class sizes, students get more one-on-one attention, learn better

HOOD COLLEGE

Example
Does reducing class size improve elementary school education?

- What is the precise effect of class size on performance?

Example
Does reducing class size improve elementary school education?

- What is the precise effect of class size on performance?
- A policy-relevant tradeoff with a budget constraint

Example
Does reducing class size improve elementary school education?

- What is the precise effect of class size on performance?
- A policy-relevant tradeoff with a budget constraint
- Biases due to wealthy school districts?

HOOD COLLEGE

Example
Is there racial discrimination in home mortgage lending?

- Boston Fed: 28% of African-Americans are denied mortgages compared to only 9% of White Americans

HOOD COLLEGE

Example
Is there racial discrimination in home mortgage lending?

- What factors determine whether a person is denied a mortgage?

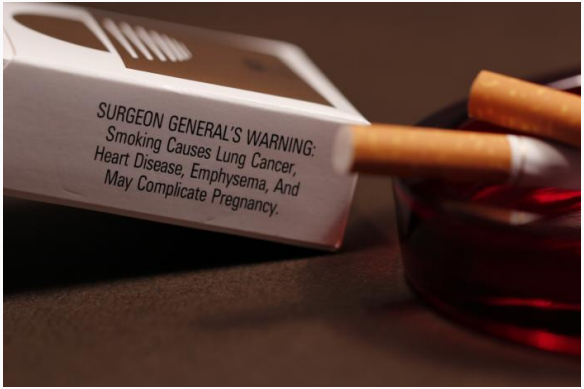
Example
Is there racial discrimination in home mortgage lending?

- What factors determine whether a person is denied a mortgage?
- How many denials are explained *purely* by race, vs. other factors?

HOOD COLLEGE

Example
How much do state cigarette taxes reduce smoking rates?

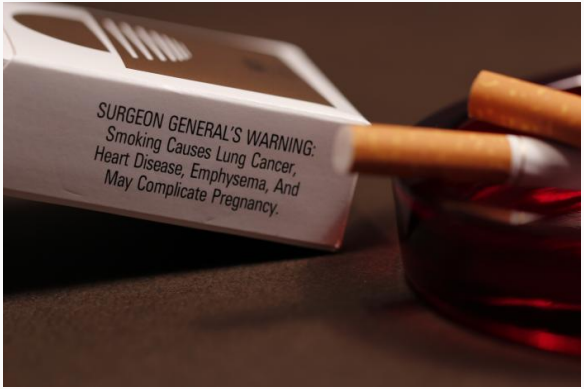
- Econ 101: raise price $\Longrightarrow$ lower quantity consumed



HOOD COLLEGE

Example
How much do state cigarette taxes reduce smoking rates?

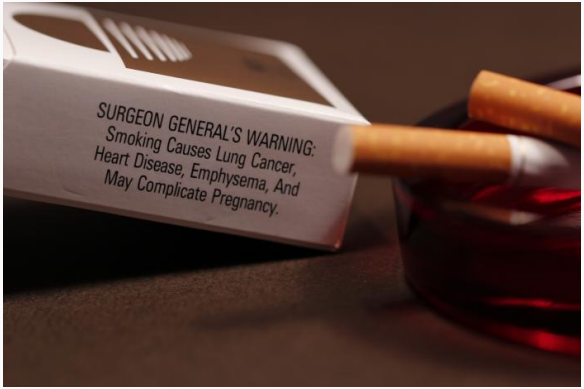
- Econ 101: raise price $\Longrightarrow$ lower quantity consumed
- What is the *price elasticity of demand* for smoking?



HOOD COLLEGE

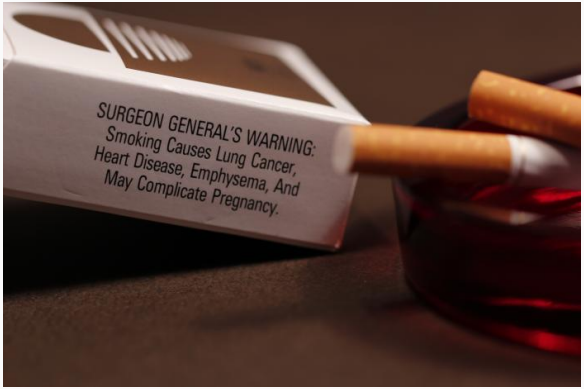
Example
How much do state cigarette taxes reduce smoking rates?

- Low cigarette taxes probably causes more smoking (& vice versa).



Example
How much do state cigarette taxes reduce smoking rates?

- Low cigarette taxes probably causes more smoking (& vice versa).
- Might more smokers also cause low cigarette taxes?



HOOD COLLEGE

Example
What will the inflation rate be next year?

- Banks, governments, businesses make forecasts about economic indicators to make future plans

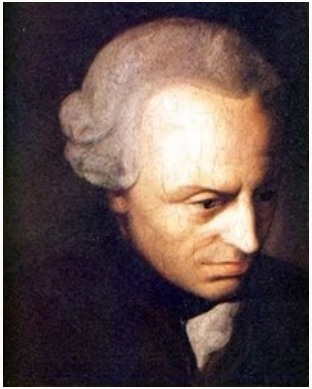
HOOD COLLEGE

Example
What will the inflation rate be next year?

- Banks, governments, businesses make forecasts about economic indicators to make future plans
- This course is less about forecasting

The Basic Empirical Research Process

"Experience without theory is blind, but theory without experience is mere intellectual play."



Immanuel Kant

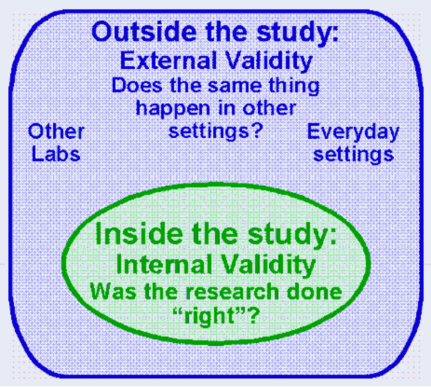
1724-1804

- First, we must formulate a question we are interested in answering (harder than you think!)

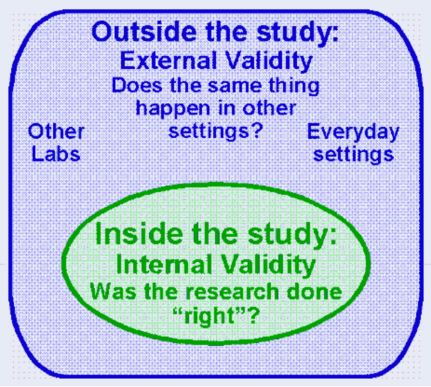
- First, we must formulate a question we are interested in answering (harder than you think!)
- Ex: How does education affect wages? Does going to school longer raise your wages?

- First, we must formulate a question we are interested in answering (harder than you think!)
- Ex: How does education affect wages? Does going to school longer raise your wages?
- We need a theory to tell us how and why phenomena cause each other! Otherwise all we have are correlations!

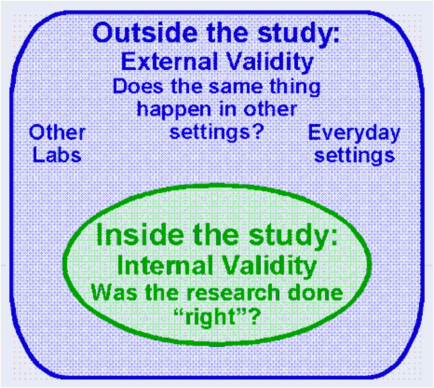
- We need a good theory for external validity of a study



- We need a good theory for external validity of a study
- Is what is true of one sample/population true of others?



- We need a good theory for external validity of a study
- Is what is true of one sample/population true of others?
  - Lucas (1976) critique: can't base policy off of historical econometric relationships that, when changed, will cause people to change their behavior!



- Economic theorizing often involves building a formal model to relate economic phenomena and build intuitions

Example
Becker (1968) famously models crime as a rational choice:

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

y:	hours spent on criminal activities	x_1:	hourly "wage" of criminal activity
x_2:	hourly wage for legal employment	x_3:	other income
x_4:	probability of getting caught	x_5:	probability of conviction if caught
x_6:	expected sentence if convicted	x_7:	age

HOOD COLLEGE

- Economic theorizing often involves building a formal model to relate economic phenomena and build intuitions

Example
Becker (1968) famously models crime as a rational choice:

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

y:	hours spent on criminal activities	x_1:	hourly "wage" of criminal activity
x_2:	hourly wage for legal employment	x_3:	other income
x_4:	probability of getting caught	x_5:	probability of conviction if caught
x_6:	expected sentence if convicted	x_7:	age

- We can at least predict the "sign" of each relationship between y and each x_i (then estimate the quantitative impact)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- y is the dependent variable of interest
  - AKA "response variable," "regressand," "Left-hand side (LHS) variable"

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- y is the dependent variable of interest
  - AKA "response variable," "regressand," "Left-hand side (LHS) variable"
- x_1 and x_2 are independent variables
  - AKA "explanatory variables," "regressors," "Right-hand side (RHS) variables," "covariates," "control variables"

HOOD COLLEGE

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- y is the dependent variable of interest
  - AKA "response variable," "regressand," "Left-hand side (LHS) variable"
- x_1 and x_2 are independent variables
  - AKA "explanatory variables," "regressors," "Right-hand side (RHS) variables," "covariates," "control variables"
- We have observed values of y, x_1, and x_2 & "regress y on x_1 and x_2"

HOOD COLLEGE

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- y is the dependent variable of interest
  - AKA "response variable," "regressand," "Left-hand side (LHS) variable"
- x_1 and x_2 are independent variables
  - AKA "explanatory variables," "regressors," "Right-hand side (RHS) variables," "covariates," "control variables"
- We have observed values of y, x_1, and x_2 & "regress y on x_1 and x_2"
- β_0, β_1, and β_2 are unknown parameters to *estimate*

HOOD COLLEGE

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

- y is the dependent variable of interest
  - AKA "response variable," "regressand," "Left-hand side (LHS) variable"
- x_1 and x_2 are independent variables
  - AKA "explanatory variables," "regressors," "Right-hand side (RHS) variables," "covariates," "control variables"
- We have observed values of y, x_1, and x_2 & "regress y on x_1 and x_2"
- β_0, β_1, and β_2 are unknown parameters to *estimate*
- ϵ is the error term
  - It is **stochastic** (random)
  - We can never measure the error term

HOOD COLLEGE

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters (“the betas”)

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters ("the betas")
- Suppose we (hypothetically) find that:

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters ("the betas")
- Suppose we (hypothetically) find that:
  - $\hat{\beta}_0 = 6.2$

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters ("the betas")
- Suppose we (hypothetically) find that:
  - $\hat{\beta}_0 = 6.2$
  - $\hat{\beta}_1 = 4.1$

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters ("the betas")
- Suppose we (hypothetically) find that:
  - $\hat{\beta}_0 = 6.2$
  - $\hat{\beta}_1 = 4.1$
  - $\hat{\beta}_2 = 7.1$

HOOD COLLEGE

Example

$$wage = \beta_0 + \beta_1 educ + \beta_2 expr + \epsilon$$

- Most of econometrics is about trying to estimate the parameters ("the betas")
- Suppose we (hypothetically) find that:
  - $\hat{\beta}_0 = 6.2$
  - $\hat{\beta}_1 = 4.1$
  - $\hat{\beta}_2 = 7.1$
- Then we can make predictions based on our empirical model of the causal relationship between X's and Y

HOOD COLLEGE

Example

$$\widehat{wage} = 6.2 + 4.1educ + 7.1expr$$

- Predictions based on our empirical model of the causal relationship between X's and Y

$$\widehat{wage} = 6.2 + 4.1(10) + 7.1(3) = \$68.5 \text{ thousand}$$

HOOD COLLEGE

Example

$$\widehat{wage} = 6.2 + 4.1educ + 7.1expr$$

- Predictions based on our empirical model of the causal relationship between X's and Y
  - A year of education (holding experience constant) increases wages by \$4.1 thousand

$$\widehat{wage} = 6.2 + 4.1(10) + 7.1(3) = \$68.5 \text{ thousand}$$

Example

$$\widehat{wage} = 6.2 + 4.1educ + 7.1expr$$

- Predictions based on our empirical model of the causal relationship between X's and Y
  - A year of education (holding experience constant) increases wages by \$4.1 thousand
  - A year of experience (holding education constant) increases wages by \$7.1 thousand

$$\widehat{wage} = 6.2 + 4.1(10) + 7.1(3) = \$68.5 \text{ thousand}$$

HOOD COLLEGE

Example

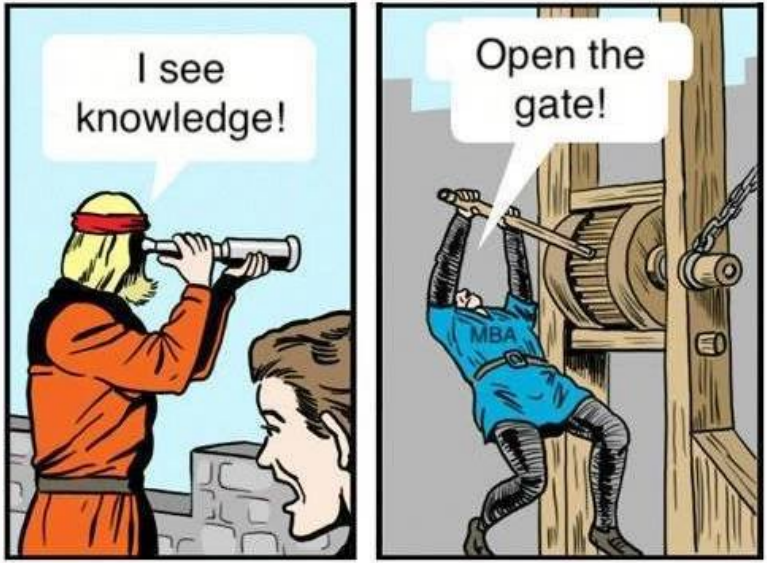
$$\widehat{wage} = 6.2 + 4.1educ + 7.1expr$$

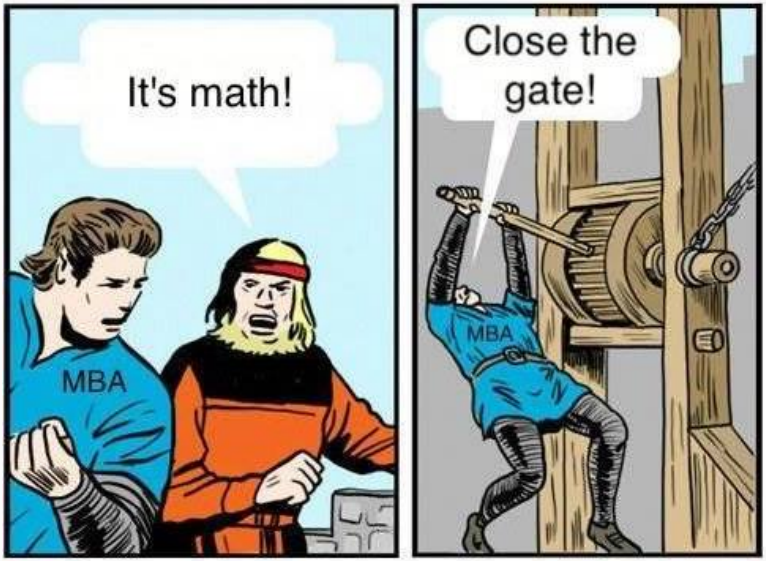
- Predictions based on our empirical model of the causal relationship between X's and Y
  - A year of education (holding experience constant) increases wages by \$4.1 thousand
  - A year of experience (holding education constant) increases wages by \$7.1 thousand
  - For someone with 10 years of education and 3 years of experience, their predicted wage is

$$\widehat{wage} = 6.2 + 4.1(10) + 7.1(3) = \$68.5 \text{ thousand}$$

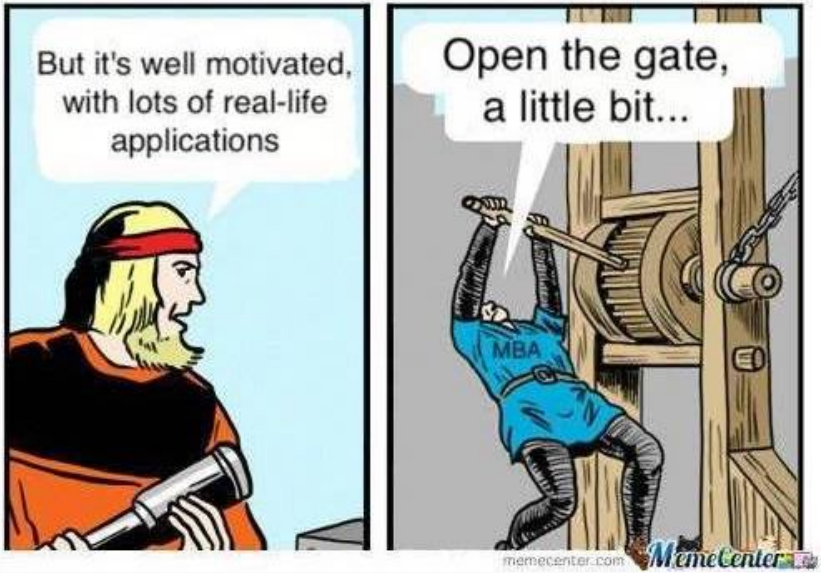
HOOD COLLEGE

About the Course


I see knowledge!
Open the gate!
MBA


It's math!
MBA
Close the gate!
MBA

Real Talk III



Real Talk IV

- This will be one of the hardest courses you take at Hood

HOOD COLLEGE

Real Talk IV

- This will be one of the hardest courses you take at Hood
- There will be moments where you have no idea WTF is going on (*this is normal*)

HOOD COLLEGE

Real Talk IV

- This will be one of the hardest courses you take at Hood
- There will be moments where you have no idea WTF is going on (*this is normal*)
- Yes, you can still get an A

HOOD COLLEGE

1. To understand and evaluate statistical & empirical claims

R Studio

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills

R Studio

What You'll Get out of this Course

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills
2. To understand research design, causation, and hypothesis testing

R Studio

HOOD COLLEGE

What You'll Get out of this Course

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills
2. To understand research design, causation, and hypothesis testing
3. To gain experience analyzing, interpreting, and communicating real data

R Studio

HOOD COLLEGE

What You'll Get out of this Course

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills
2. To understand research design, causation, and hypothesis testing
3. To gain experience analyzing, interpreting, and communicating real data
    3.1 Research project

R Studio

HOOD COLLEGE

What You'll Get out of this Course

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills
2. To understand research design, causation, and hypothesis testing
3. To gain experience analyzing, interpreting, and communicating real data
    3.1 Research project
    3.2 Use Tools of the Trade: data analysis with *R*

R Studio

HOOD COLLEGE

What You'll Get out of this Course

1. To understand and evaluate statistical & empirical claims
    1.1 Skepticism & critical thinking skills
2. To understand research design, causation, and hypothesis testing
3. To gain experience analyzing, interpreting, and communicating real data
    3.1 Research project
    3.2 Use Tools of the Trade: data analysis with *R*
    3.3 Ambitious: managing workflow and reproducibility with *R Markdown*

R Studio

HOOD COLLEGE

You are in Demand

DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

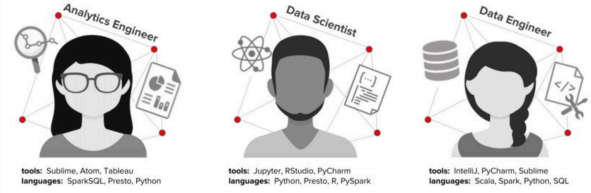
Harvard Business Review

HOOD COLLEGE

You are in Demand II

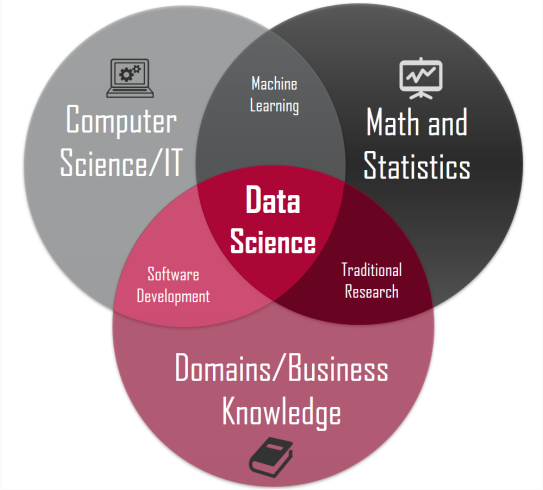
Data roles at Netflix

- business analyst
- data analyst
- quantitative analyst
- algorithm engineer
- analytics engineer
- data engineer
- data scientist
- machine learning scientist
- research scientist



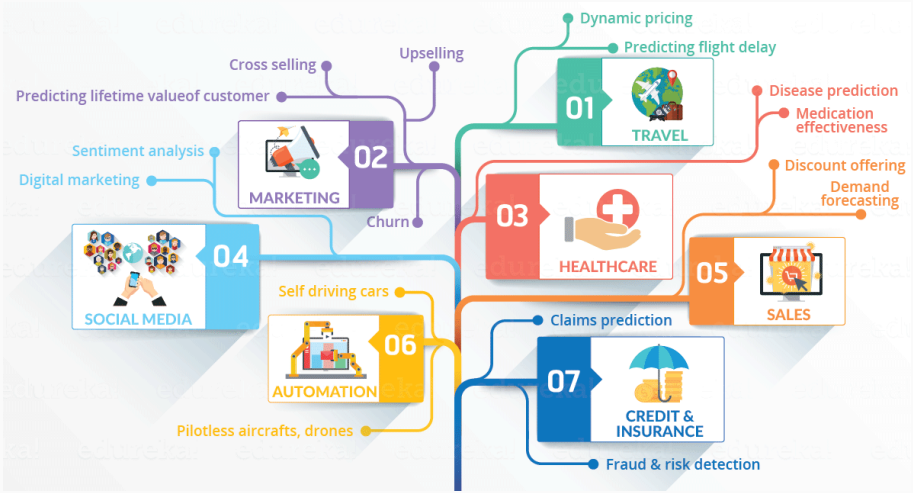
The Netflix Tech Blog

You are budding Data-Scientists



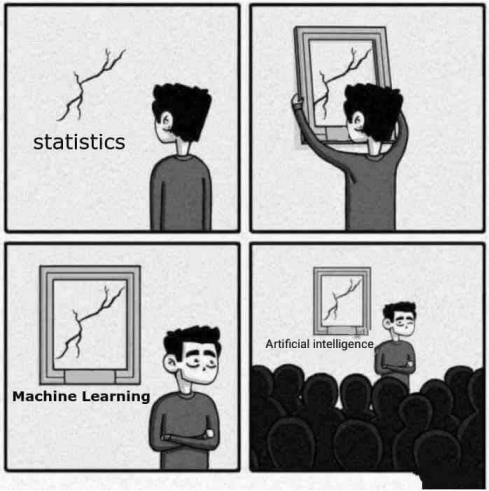
HOOD COLLEGE

You are budding Data-Scientists II



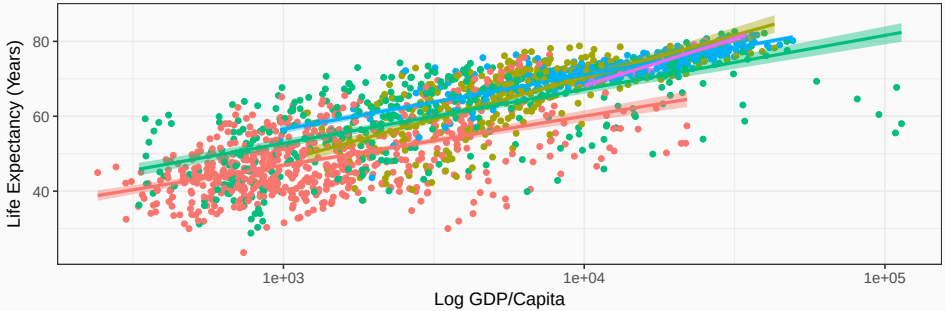
HOOD COLLEGE

The Sexiest Industry Terms are Basically Just Statistics

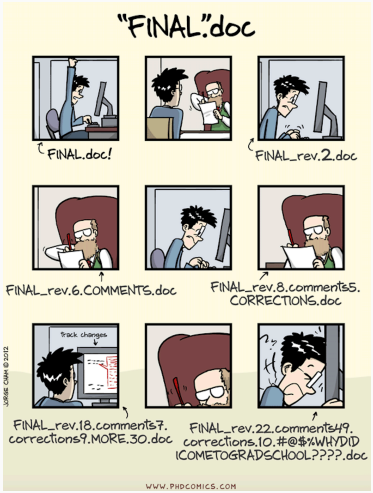


R is Very Powerful: I Literally Ran the Code in This Slide

```
ggplot(data = gapminder, aes(x = gdpPercap,
    y = lifeExp, color = continent, fill= continent))+
    geom_point()+geom_smooth(method = "lm") +
    scale_x_log10()+ylab("Life Expectancy (Years)")+
    xlab("Log GDP/Capita")
```



End This Nightmare!



PhD Comics

HOOD COLLEGE

.Rmd files

An R Markdown (.Rmd) file is a record of your research. It contains the code that a scientist needs to reproduce your work along with the narration that a reader needs to understand your work.

Reproducible Research

At the click of a button, or the type of a command, you can rerun the code in an R Markdown file to reproduce your work and export the results as a finished report.

Dynamic Documents

You can choose to export the finished report as a html, pdf, MS Word, ODT, RTF, or markdown document; or as a html or pdf based slide show.

From R Markdown Cheatsheet

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.

HOOD COLLEGE

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.
    - No more re-run everything and re-copy/paste when you change your data

HOOD COLLEGE

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.
    - No more re-run everything and re-copy/paste when you change your data
    - I can change the axes in the same file as I wrote the slide with my graph, and it will remake the graph in the slide automatically

HOOD COLLEGE

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.
    - No more re-run everything and re-copy/paste when you change your data
    - I can change the axes in the same file as I wrote the slide with my graph, and it will remake the graph in the slide automatically
- *R Markdown* is a simple human- and machine-readable language: do *all* your work in one file, export to PDF, Word, HTML, Slides

HOOD COLLEGE

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.
  - No more re-run everything and re-copy/paste when you change your data
  - I can change the axes in the same file as I wrote the slide with my graph, and it will remake the graph in the slide automatically
- *R Markdown* is a simple human- and machine-readable language: do *all* your work in one file, export to PDF, Word, HTML, Slides
- Using *R* is required, *R Markdown* is optional

HOOD COLLEGE

Ambitious: Use *R Markdown* to Think About Your Workflow

- Reproducible research: you run my code (e.g. the last slide) and you will get the same graph as me, *every time*.
- Manage your workflow: No more copy/paste pictures across 3 programs, 10 files, etc.
    - No more re-run everything and re-copy/paste when you change your data
    - I can change the axes in the same file as I wrote the slide with my graph, and it will remake the graph in the slide automatically
- *R Markdown* is a simple human- and machine-readable language: do *all* your work in one file, export to PDF, Word, HTML, Slides
- Using *R* is required, *R Markdown* is optional
    - Advanced: version control with GitHub

HOOD COLLEGE