

Research Demo

Undergraduate Field Seminar: American Politics

Examining the 2016 Election:
Inequality, Populism and Voting

Jan Zilinsky
NYU
Fall 2017

You will be a researcher for the next 6-8 weeks

- You will ask one interesting (and original) question
- You will dig into datasets on voting and attitudes
- You will report your results

Types of topics people you can study (be creative)

- Was Trump's opposition to trade particularly popular in rural areas? (Impact of globalization)
- Impact of ethnic diversity on voting
- Why did Trump receive a larger Hispanic vote share in 2016 than Romney in 2012?
- What are the demographic characteristics of Trump voters? Were low-income voters more likely to view Trump favorably?
- Is there a positive relationship between poor health and the Trump vote?

Methods

- You will mostly interpret correlations
 - Cross-tabulating data
 - OLS / probit / logit
 - Let's talk about interpretation
- Claims you can make: “the data is consistent with my hypothesis”. (Don't say you have “proven” XYZ.)

How do I find the voters who preferred Obama in 2012 and then preferred Trump in 2016?

- CCES data: a nationally representative sample of 64,600 American adults.
- Voter Study Group (VSG) data (smaller but rich dataset)

How do I find the voters who preferred Obama in 2012 and then preferred Trump in 2016? **CCES data**

* Vote choice in 2012:

```
tab CC16_326
```

* Vote choice in 2016:

```
tab CC16_410a
```

```
gen obama12_trump16 = (CC16_326 == 1 & CC16_410a == 1)
```

```
summarize obama12_trump16 [aw=commonweight]
```

* Why is the reported share incorrect?

* Look at the N. The denominator is not right!

How do I find the voters who preferred Obama in 2012 and then preferred Trump in 2016?

The point is that you **must** only label an observation as a voter/non-voters (1 vs. 0) if the respondent was actually asked whether he/she voted or not.

More than 7,000 people were not asked, *and they are not coded as missing*; see the next slide.

A person who was not asked

- Must be coded as a missing observation
- You cannot assume that these respondents “**did not vote for candidate X**”

How do I find the voters who preferred Obama in 2012 and then preferred Trump in 2016?

* Vote choice in 2016:

tab CC16_410a

```
. * Vote choice in 2016:
. tab CC16_410a
```

President vote	Freq.	Percent	Cum.
Donald Trump (Republican)	18,755	35.45	35.45
Hillary Clinton (Democrat)	22,136	41.85	77.30
Gary Johnson (Libertarian)	1,829	3.46	80.76
Jill Stein (Green)	913	1.73	82.48
Other	1,136	2.15	84.63
I didn't vote in this election	81	0.15	84.78
I'm not sure	229	0.43	85.22
Evan McMullin (Independent)	163	0.31	85.53
Skipped	27	0.05	85.58
Not Asked	7,630	14.42	100.00
Total	52,899	100.00	

Run also: tab CC16_410a, nol

How do I find the voters who preferred Obama in 2012 and then preferred Trump in 2016?

* Generate vote variables

```
gen clinton16 = (CC16_410a==2)
```

```
replace clinton = . if CC16_410a==.
```

```
replace clinton = . if CC16_410a==98
```

```
replace clinton = . if CC16_410a==99
```

Can attitudes help you predict the whether someone switched their vote?

2012/16 switch based on views on illegal immigrants:

Grant legal status to all illegal immigrants who have held jobs?

```
. tabstat obama12_trump16 [aw=commonweight], by(CC16_331_1)

Summary for variables: obama12_trump16
    by categories of: CC16_331_1 (Immigration - Grant legal
> status to all illegal immigrants who have held jobs an)
```

CC16_331_1	mean
Yes	.0329149
No	.0452237
Total	.0386632

Contextual variables

Is it the case that people who live in particular types of areas share certain behaviors?

In CCES, look for geographic indicators:
`countyname` `countyfips` `inputstate`

How can I find the voters who preferred Obama in 2012 and then preferred Trump in 2016? **VOTER Survey**

```
svyset [pw=weight]
```

```
* Vote choice
```

```
gen obama08 = (presvote08_baseline==1)
```

```
gen mccain08 = (presvote08_baseline==2)
```

```
gen obama12 = (post_presvote12_2012==1)
```

```
gen romney12 = (post_presvote12_2012==2)
```

```
gen obamaTrumpVoter = (presvote08_baseline==1 &  
presvote16post_2016==2)
```

Who supported a particular candidate?

- Correlate multiple factors such as Hispanic, Republican party ID, and feelings about Trump's immigration positions ...
- [See the provided .do file]

Angle: rural vs. urban voters

- See `urbancity_baseline` variable in the VOTER Survey

Rising unemployment

- Go to <http://guides.nyu.edu/c.php?g=276966&p=1846659>
- Data-planet contains county-level data

Finding Data: Quick Statistics

enter your search here

Search

Home

Search Strategies

Quick Statistics

Sources by Topic

Sources by Type

Software & Analysis


How to Cite Data & Statistics

Classes & Workshops

Data Sources for Instructional purposes

General Information

Resources



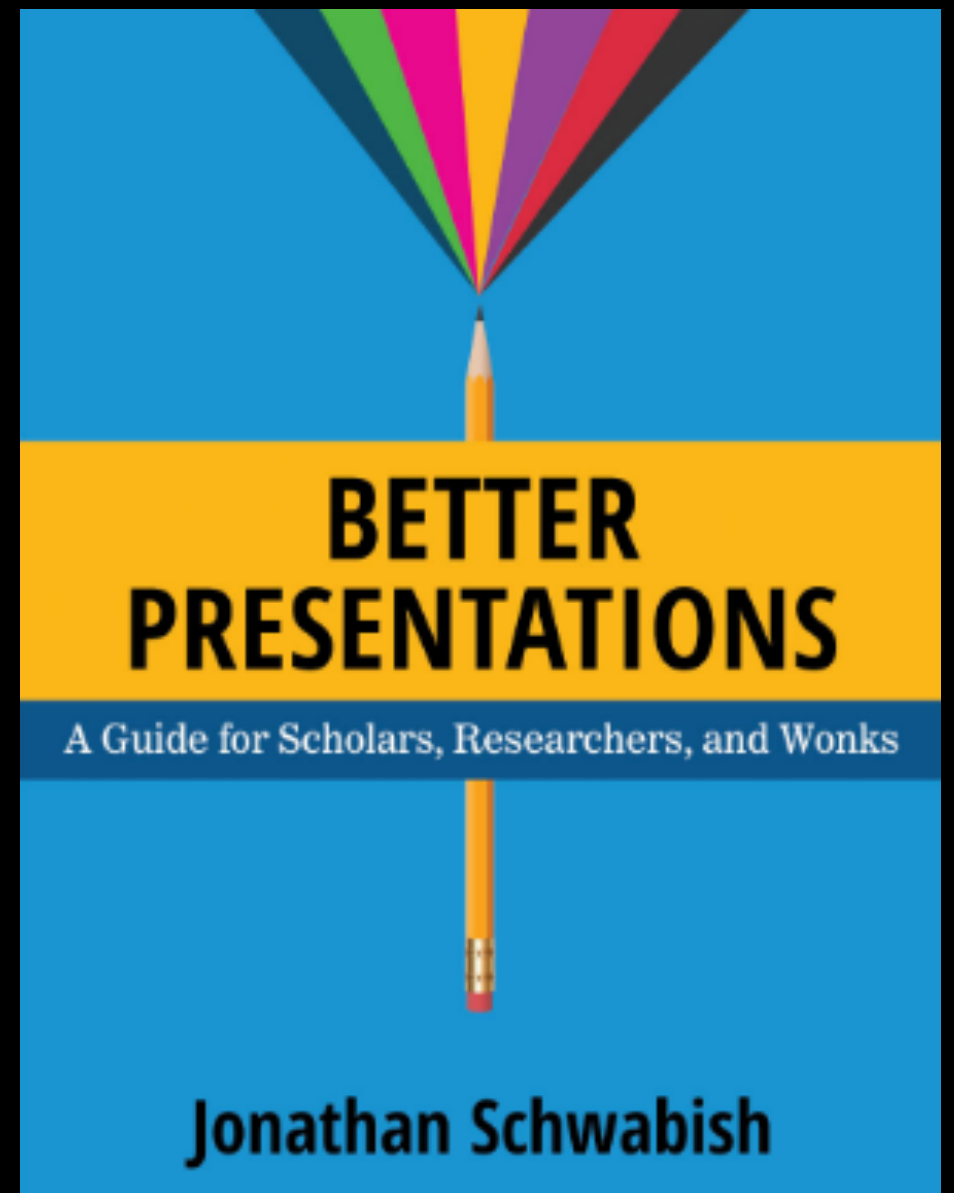
NYU LIBRARIES & IT

- [Data-Planet Statistical Datasets](#) **DS**

Data-Planet Statistical Datasets aggregates datasets sourced from reputable public and private organizations, such as OECD, the United Nations, and the U.S. Census. It covers topics across many subject areas, including education, population and income, industry, commerce, trade, housing and construction. Users can create custom extracts and download files in delimited, SAS, or shapefile formats.

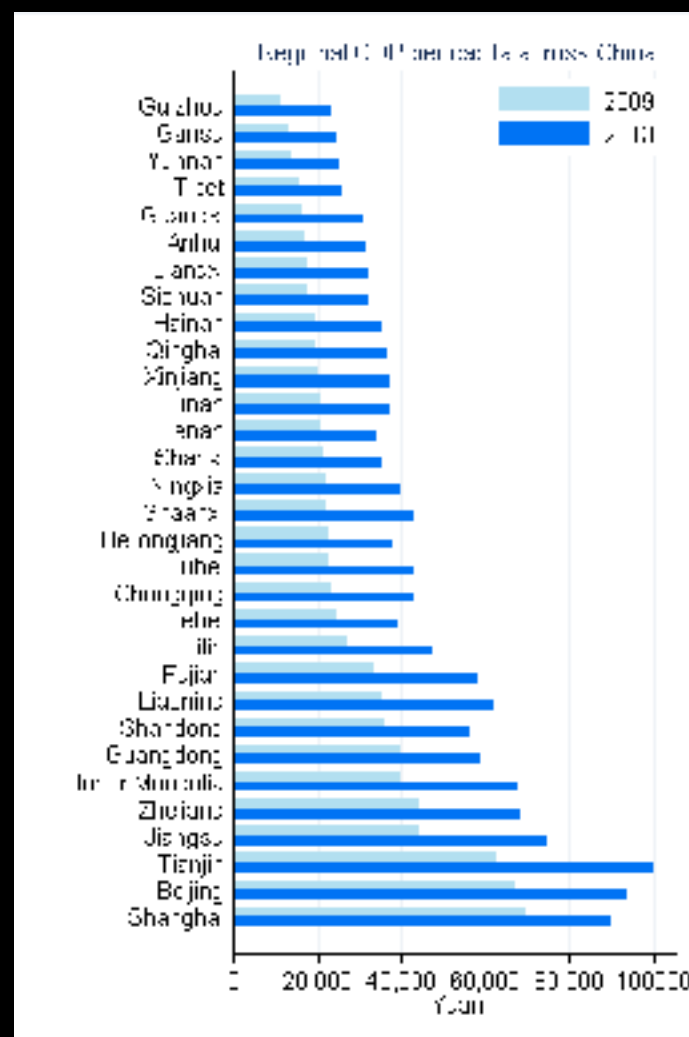
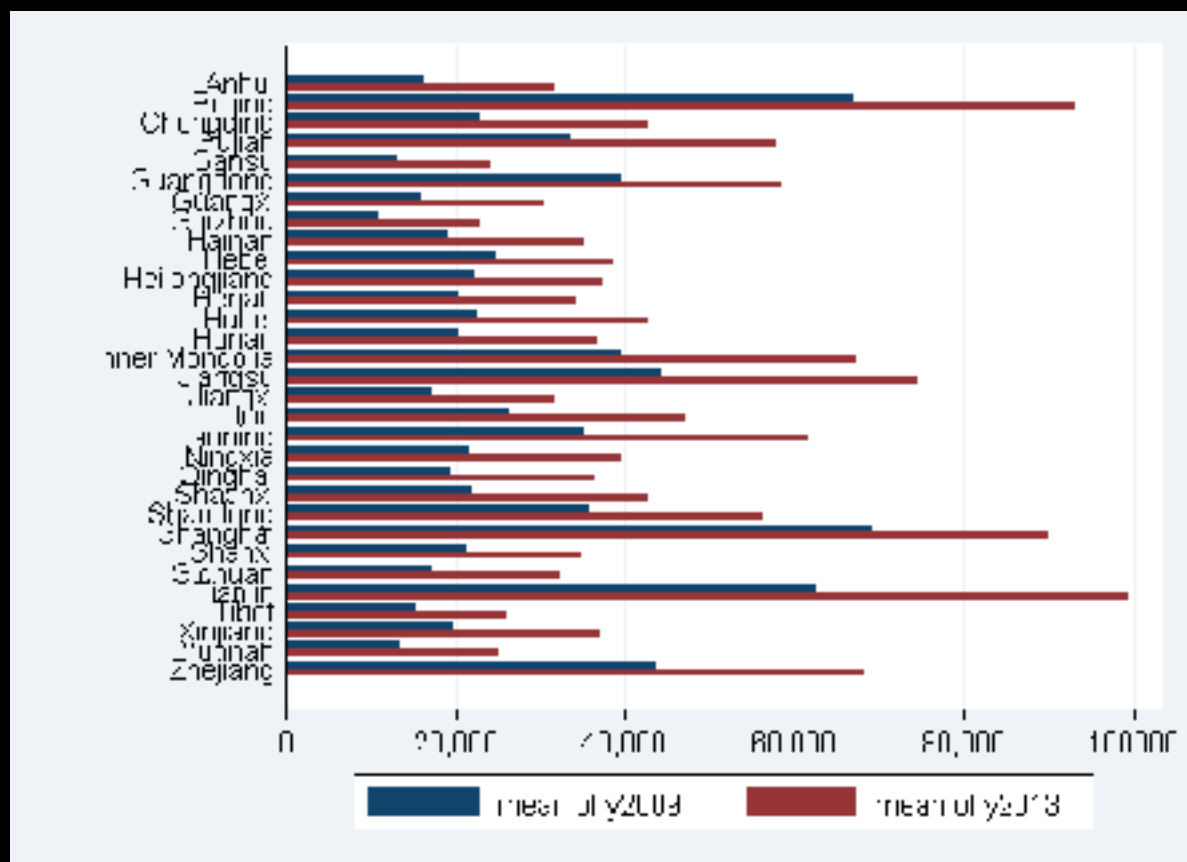
Communicating what you found

- I brought my copy today, you can borrow it.
- And consider getting the book.



Things to remember

1. Explore the data
2. Decide what you want to convey
3. Treat early figures as "drafts"
4. Eliminate chartjunk from the final version



Don't do what you see on the left

Clean your code following the advice at <http://janzilinsky.com/stata-tip-bar-charts-chinese-regions/>

Resources

- See the visualizations tips at:
 - <https://guides.library.duke.edu/datavis/topten>
 - <https://www.vis4.net/blog/2012/06/doing-the-line-charts-right/>
 - <https://policyviz.com/blog/>
 - <http://junkcharts.typepad.com/>
- Read The Upshot and 538.