1. Aim: what *substantive conclusions* about politics can be drawn from the summary statistics and histograms?

2. Explain the meaning of the statistics: “the distribution of Democrats’ support for Obama was strongly right-skewed. **This means that** most respondents approved of the President, and only a handful disapproved”
Describing Data

1. Aim: what substantive conclusions about politics can be drawn from the summary statistics and histograms?

2. Explain the meaning of the statistics: “the distribution of Democrats’ support for Obama was strongly right-skewed. This means that most respondents approved of the President, and only a handful disapproved”

3. Don’t just list the measures: “the mean was $x$, the standard deviation was $y$”

4. But don’t fail to mention them either: e.g. “On average, Republican support for Obama was very low” is better than “Republicans generally disapproved of Obama”
Good answer: “...in 1980 the average army was quite small compared to 1945 and 1915 - not much bigger than in 1900 or 2005. This is because there was no major ongoing world war. However, the largest army was much larger than in 1900 or 2005, and the standard deviation was also much greater. This indicates that the distribution features mostly small armies, with a small number of large armies. This was probably due to the cold war being at its height, so that the US and Russia had very large numbers of soldiers, with most other countries falling under their protection.”
All of the information needed to understand the figure should be conveyed within the plot itself

- Always have a title
- Label axes and scale points on axes
- Don’t leave the reader guessing what the scale means
- If you want the reader to compare between charts, plot them on the same scale
- Include a note acknowledging source
A bad figure
hist CC308a, discrete percent
Better figure

```stata
histogram CC308a, discrete percent
ytitle(Percent of Voters)
xtitle("Approval Rating" "(1='strongly approve' - 4='strongly disapprove')", margin(0 0 0 2))
xlabel(, notick)
title("Distribution of Voters’ Approval Ratings for President Obama, 2012", size(medium))
note("Source: 2012 Cooperative Congressional Election Study", size(vsmall))
```
Distribution of Voters' Approval Ratings for President Obama, 2012

Source: 2012 Cooperative Congressional Election Study
Republicans' Approval of Supreme Court in 2012, by Support for Tea Party Movement

(a) Republicans who Support Tea Party

(b) Republicans who Oppose Tea Party

Source: 2012 Cooperative Congressional Election Study
One way is with the `by()` option

- e.g., `hist CC308c, by(pid3)` plots supreme court approval by party supported
- but it’s hard to control how each chart looks

Better way is using `graph combine`

- code each chart separately, including the option `saving(“filename1”)`
- then `graph combine filename1 filename2, rows(1) cols(2)` plus other options (see solutions)