



# Descriptive Statistics

March 1, 2022

POLS 095

Drake University

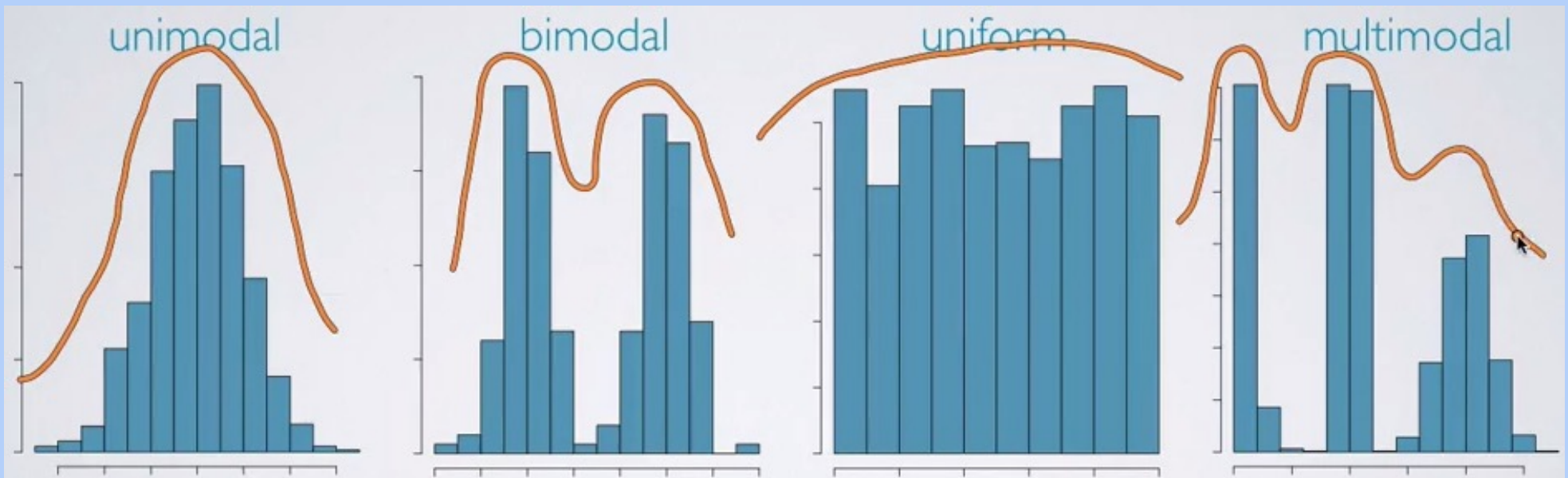
Describing the  
distribution of  
a variable

1. Central tendency or typical value
2. Dispersion (spread)



# Describing the distribution of a variable

- Other ways to describe distribution
  - Modality (unimodal, bimodal, multimodal)

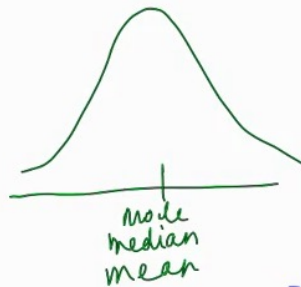


# Describing the distribution of a variable

- Other ways to describe distribution
- Skew (negative or positive)

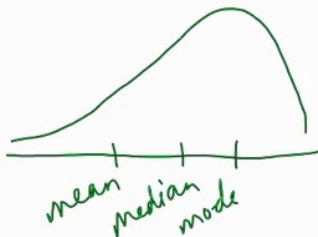
The relative location of the mode, median, and mean in a **unimodal** distribution:

Symmetric



For a symmetric distribution, the mean, median, and mode are all approximately the same.

Left-skewed



For a left-skewed distribution, the mode is larger than the median which is larger than the mean.

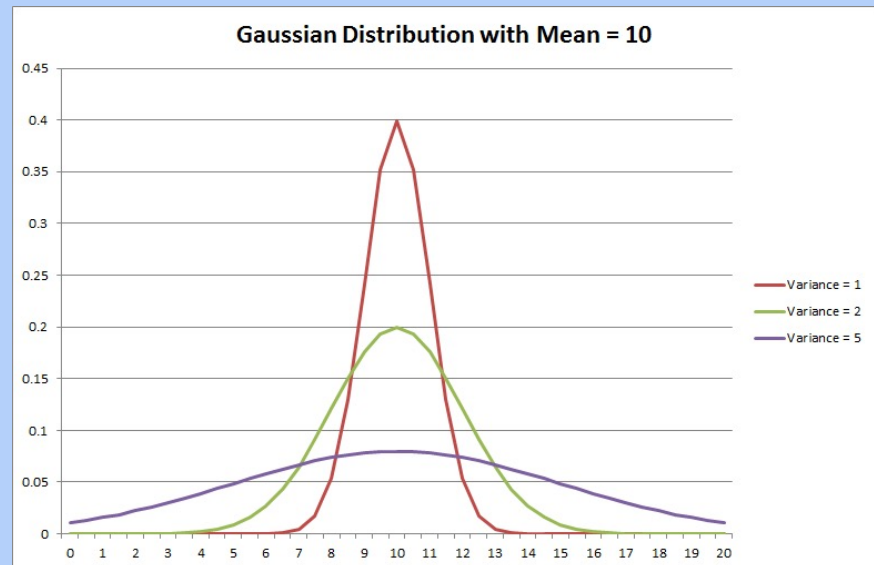
Right-skewed



For a right-skewed distribution, the mode is less than the median, which is less than the mean.

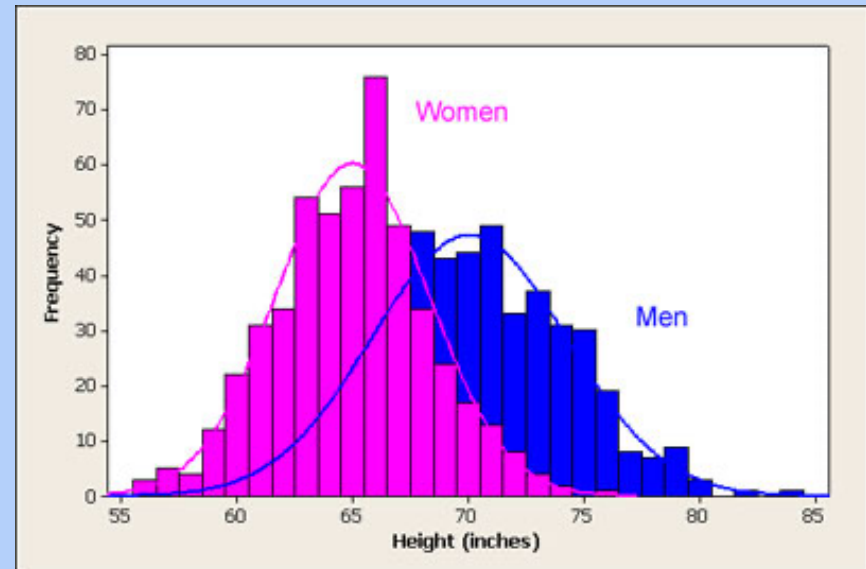
# Describing the distribution of a variable

- Frequency and relative frequency tables
- Bar charts and histograms
- Quantiles (e.g., quartiles, percentiles)
- How stretched the distribution is
- What sorts of description are possible for nominal, ordinal, and continuous levels?



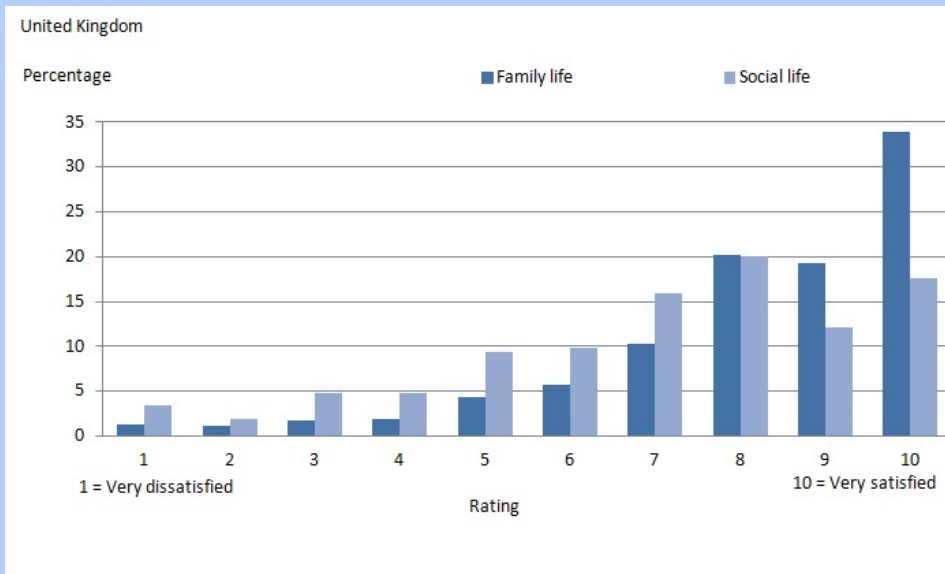
# Mean

- For interval level data that has a symmetric distribution
  - Can also be used for dichotomous data where the mean will yield the % for each category (ex. % red apples vs. % not red apples)

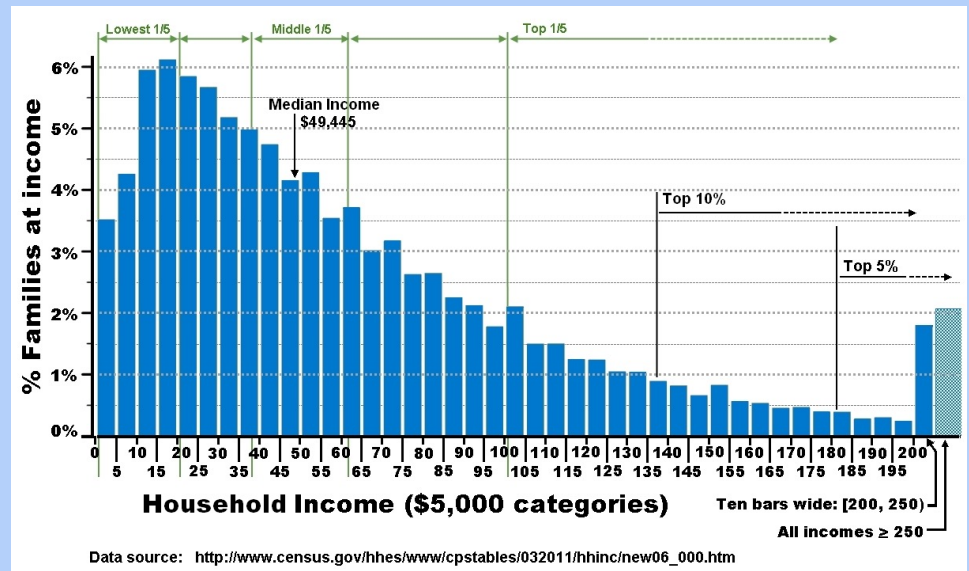


# Median

- For:
  - Ordinal level data  
(Satisfaction with life)



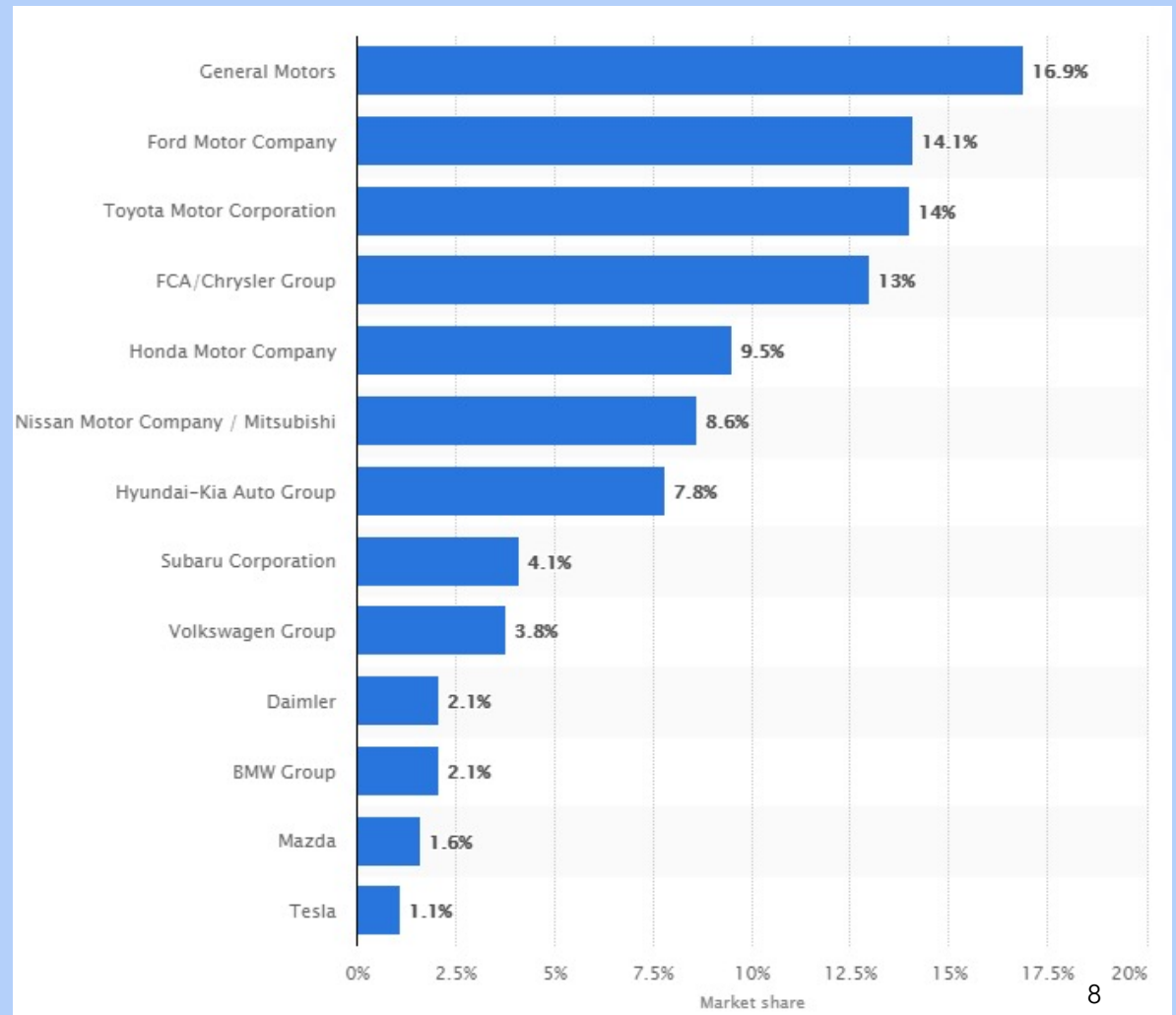
Interval level data that has a skewed distribution  
(Income)





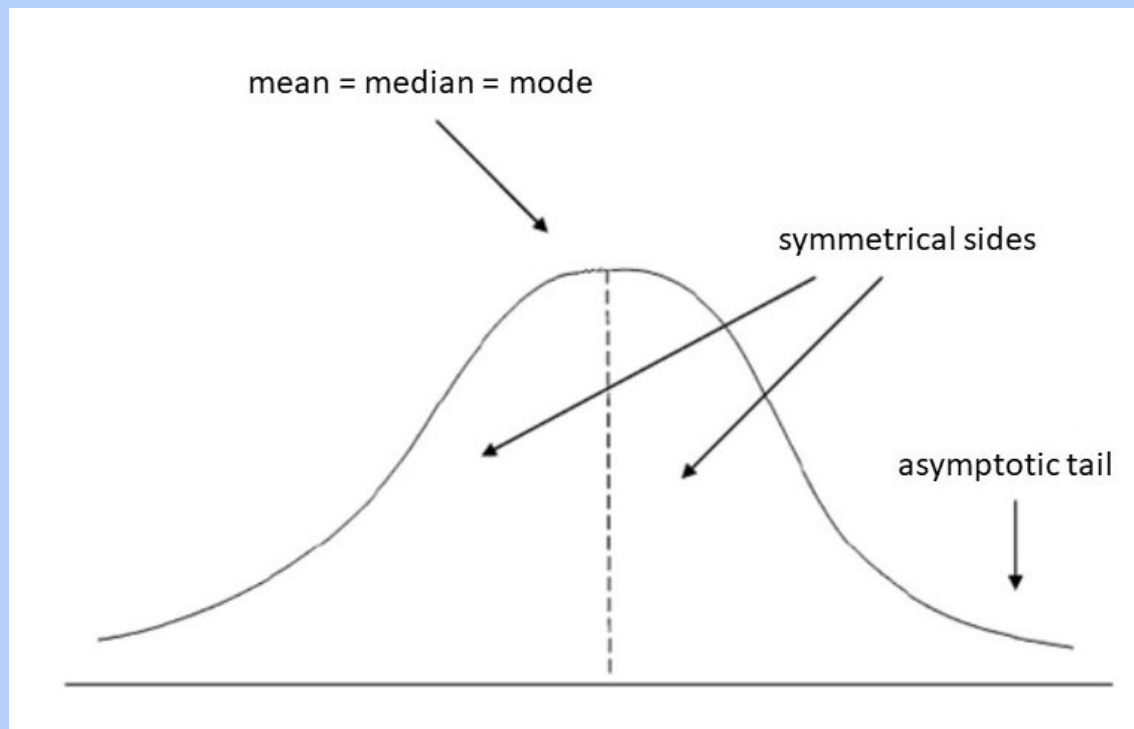
# Mode

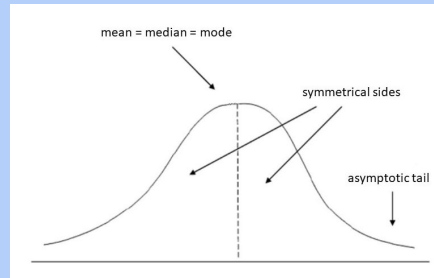
- For categorical level data
- “What is the typical car brand on the road?”





# THE NORMAL DISTRIBUTION





It's a SUPER  
TOOL because it  
has all sorts of  
good  
characteristics  
that help us

---

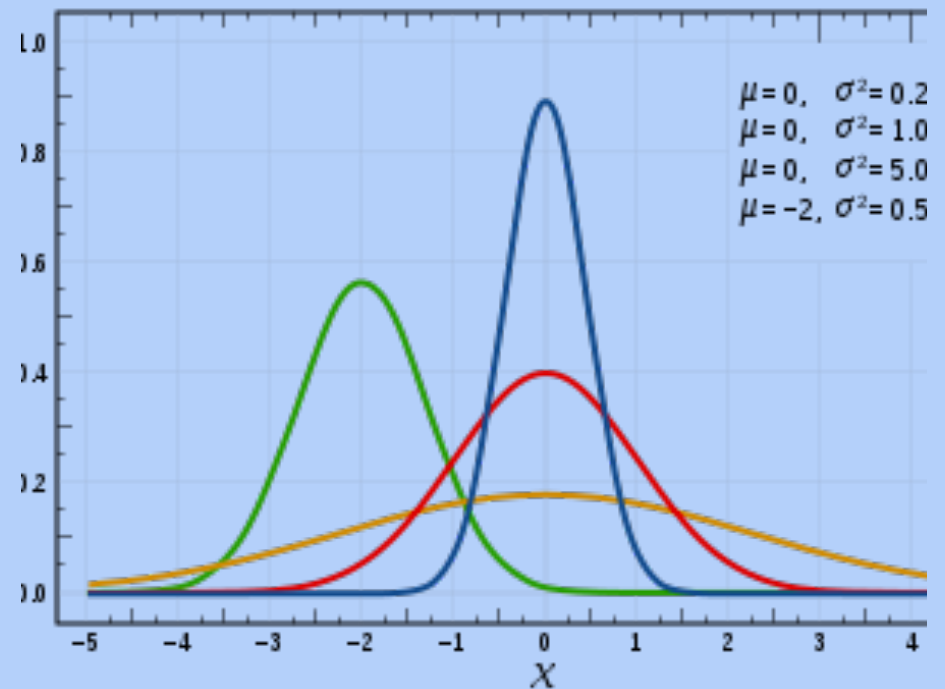
It is symmetrical.

The middle of the sampling  
distribution (the high point) will  
equal the true population mean,  
median, and mode.

---

The asymptotic tail means that VERY  
few cases will fall at the extremes  
(i.e. 0% or 100% yellow)

**J U S T  
L I K E  
A L L  
D A T A ...**

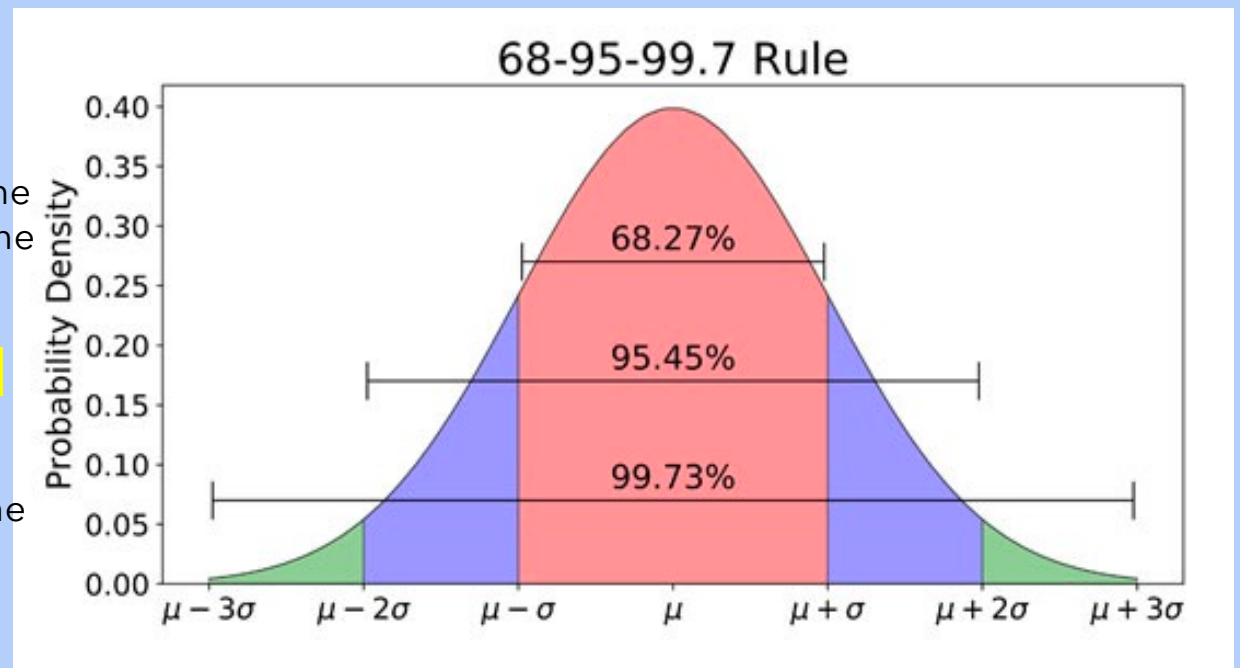


Just like all data...

68% of all samples will fall between one standard error below the mean and one standard error above the mean.

95% will fall between 2 S.E. below the mean and 2 S.E. above the mean.

99.7% will fall between 3 S.E. below the mean and 3 S.E. above the mean.



## Exercise

- 1) Find a categorical variable in the ANES data in which you are interested.
- 2) Find an ordinal variable in the ANES data in which you are interested.
- 3) Find a continuous variable in the ANES data in which you are interested.