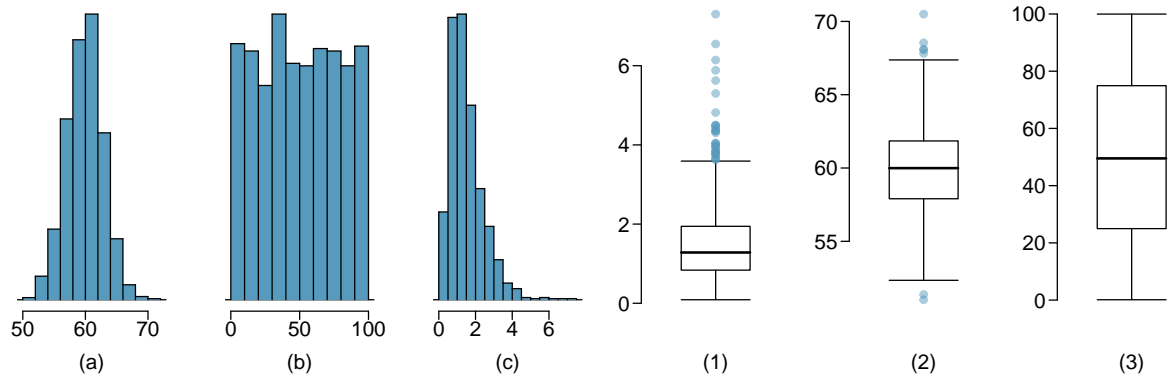


## STAT 310, HW 2

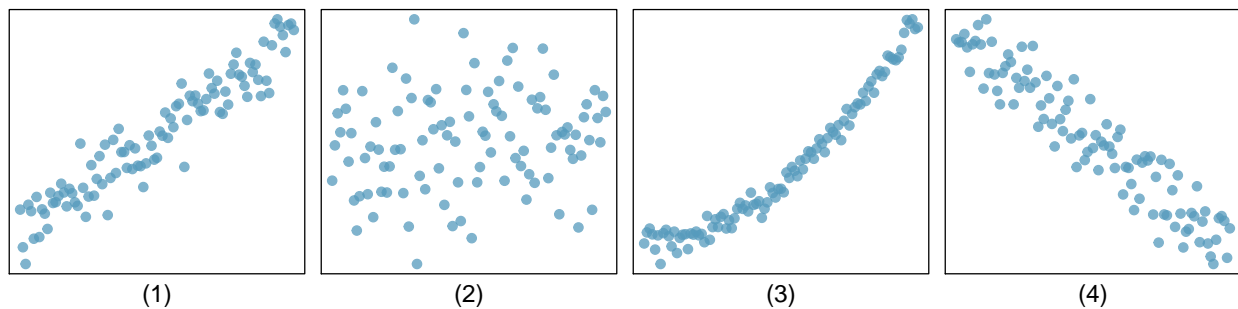
Due: Thursday, February 16

**Directions:** Please submit your completed assignment to Canvas in PDF format.

**Exercise 1.** Describe the distributions in the histograms below and match them to the box plots.



**Exercise 2.** For each scatter plot below indicate whether there is a positive association, negative association, or no association. Also determine if the positive or negative association is linear or nonlinear.



**Exercise 3.** Without doing an calculations, determine which set has the larger standard deviation. Briefly explain your reasoning, and then use R to verify.

Set A: 3,5,5,5,8,11,11,11,13

Set B: 3,5,5,5,8,11,11,11,20

**Exercise 4.** Without doing an calculations, determine which set has the larger standard deviation. Briefly explain your reasoning, and then use R to verify.

Set A: 100, 200, 300, 400, 500

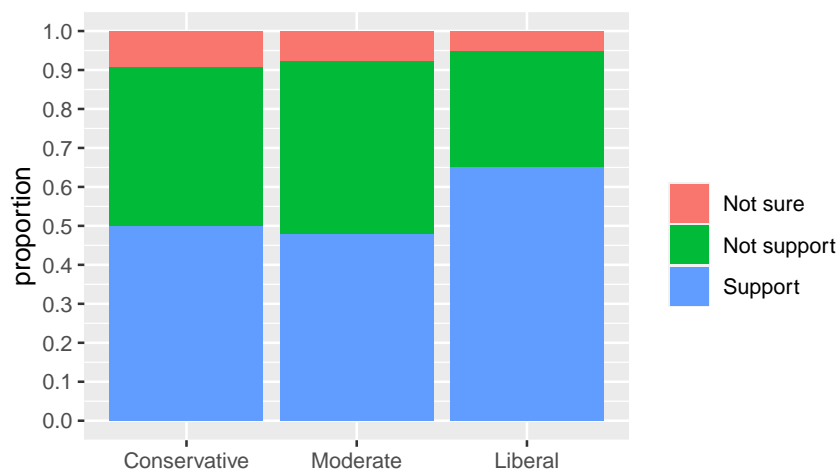
Set B: 0, 50, 300, 550, 600

**Exercise 5.** Hearing anecdotal reports that some patients undergoing treatment for the eating disorder anorexia seemed to be responding positively to the antidepressant Prozac, medical researchers conducted an experiment to investigate. They found 93 women being treated for anorexia who volunteered to participate. For one year, 49 randomly selected patients were treated with Prozac and the other 44 were given an inert substance called a placebo. At the end of the year, patients were diagnosed as healthy or relapsed, as summarized in the contingency table:

	Prozac	Placebo	Total
Healthy	35	32	67
Relapse	14	12	26
Total	49	44	93

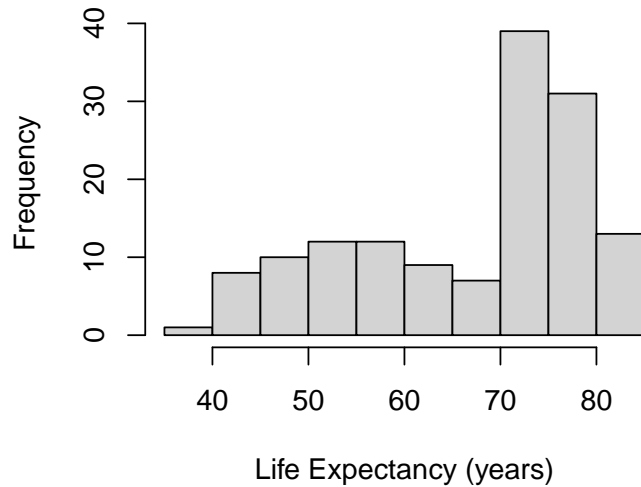
- Is this an experimental or observational study?
- What percent of patients who received Prozac relapsed?
- What percent of patients who received the Placebo relapsed?
- Do these results provide evidence that Prozac might be helpful in treating anorexia? Explain.

**Exercise 6.** A random sample of registered voters from Tampa, FL were asked if they support the DREAM Act, a proposed law which would provide a path to citizenship for people brought illegally to the US as children. The survey also collected information on the political ideology of the respondents.



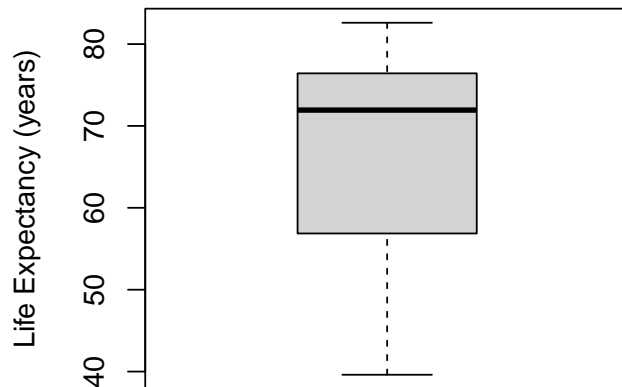
- Based on the standardized bar plot, approximately what proportion of conservative respondents supported the Dream Act?
- Approximately what proportion of liberal respondents supported the Dream Act?
- Does there appear to be an association between views on the Dream Act and political ideology, or are they independent? Explain.

**Exercise 7.** The histogram below shows the distribution of life expectancy in years for 142 countries in 2007. The data come from the website [www.gapminder.org](http://www.gapminder.org).



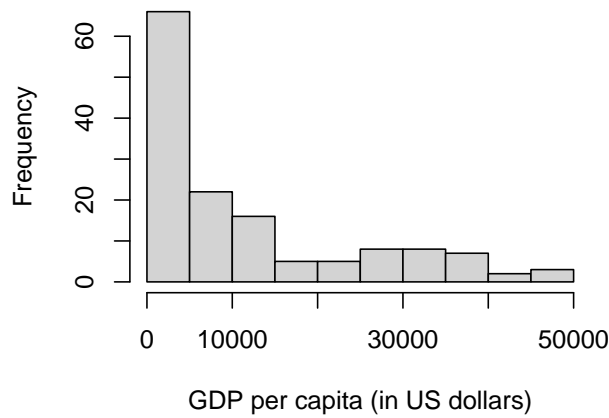
- Describe the shape of the distribution.
- Based on the histogram, would you expect the mean to be higher or lower than the median?
- Approximately what percent of countries in this data set have life expectancy between 70-80 years.

**Exercise 8.** Below is a box plot of life expectancy in years for 142 countries in 2007.



- Using the box plot, estimate  $Q_1$ , the median,  $Q_3$ , and the IQR. Also, are there any potential outliers?
- What characteristics of the distribution are apparent in the histogram, but not in the box plot?

**Exercise 9.** The histogram below shows the distribution of the GDP per capita (in US dollars) for 142 countries in 2007.



- Describe the shape of the distribution.
- Based on the histogram, would you expect the mean to be higher or lower than the median?

**Exercise 10.** Displayed below is a scatter plot of life expectancy (in years) versus GDP per capita (in US dollars) for 142 countries in 2007. The points are colored according to the continent where each country is located. Write a few sentences describing the relationship between the variables in the scatter plot.

