Ch. 6: Comparing Two Means

Navigation

By Date

- March 31: start - end
- April 2: start - end
- April 7: start - end

By Section

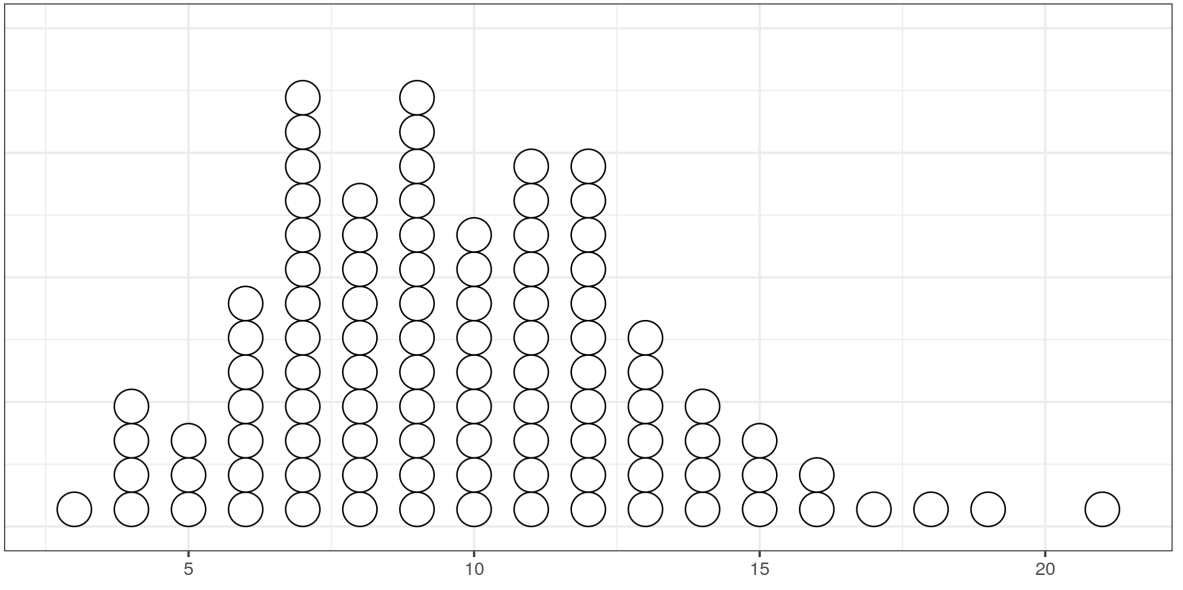
- 6.1: start - end
- 6.2: start - end
- 6.3: start - end

6.1: Comparing Two Groups

Quantitative Response

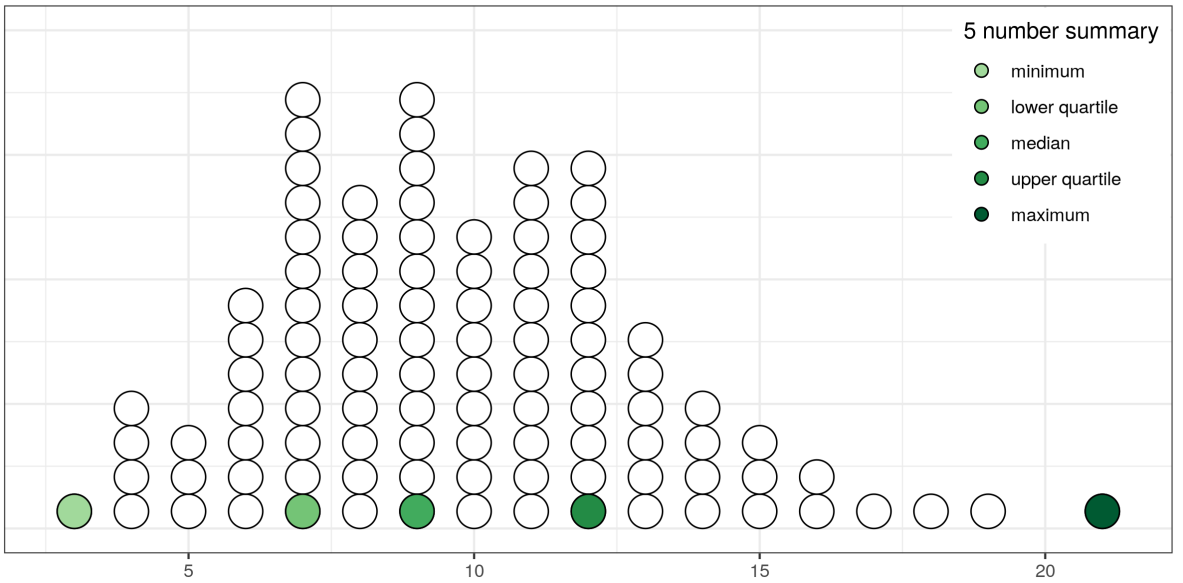
Describing Distributions of Quantitative Data

- **5-number summary** - the minimum, lower quartile, median, upper quartile, and maximum of a set of data
  - **lower quartile** - 25% of the data lie below this value
  - **median** - 50% of the data lie below this value
  - **upper quartile** - 75% of the data lie below this value

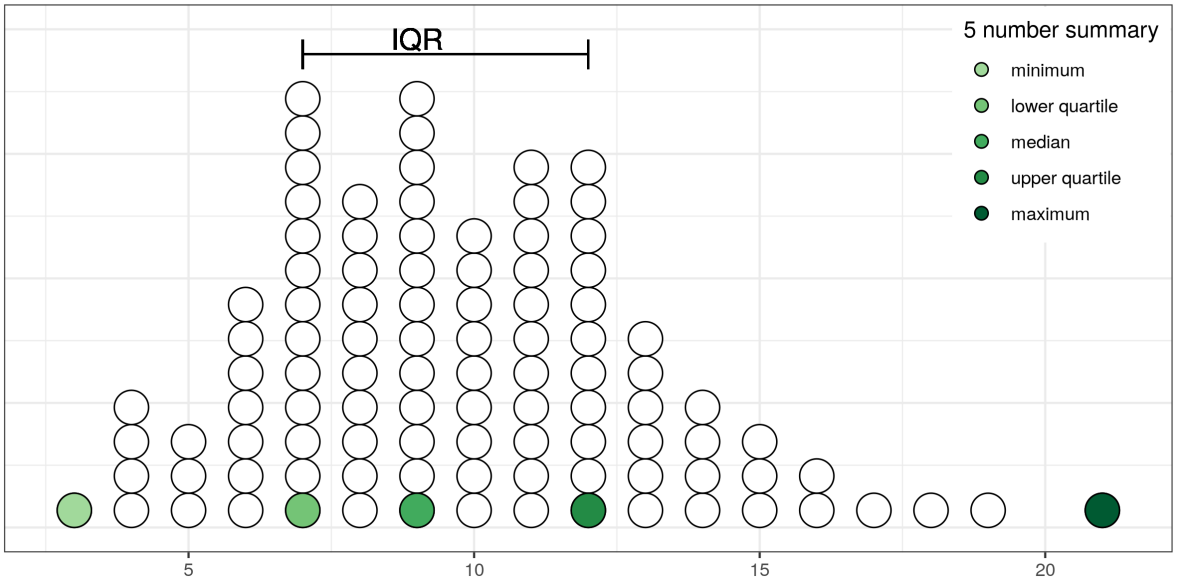


Describing Distributions of Quantitative Data

- **5-number summary** - the minimum, lower quartile, median, upper quartile, and maximum of a set of data
  - **lower quartile** - 25% of the data lie below this value
  - **median** - 50% of the data lie below this value
  - **upper quartile** - 75% of the data lie below this value



Describing Distributions of Quantitative Data



The distance between the two quartiles is called the **inter-quartile range**(IQR). The IQR is another measure of variability, along with the standard deviation.

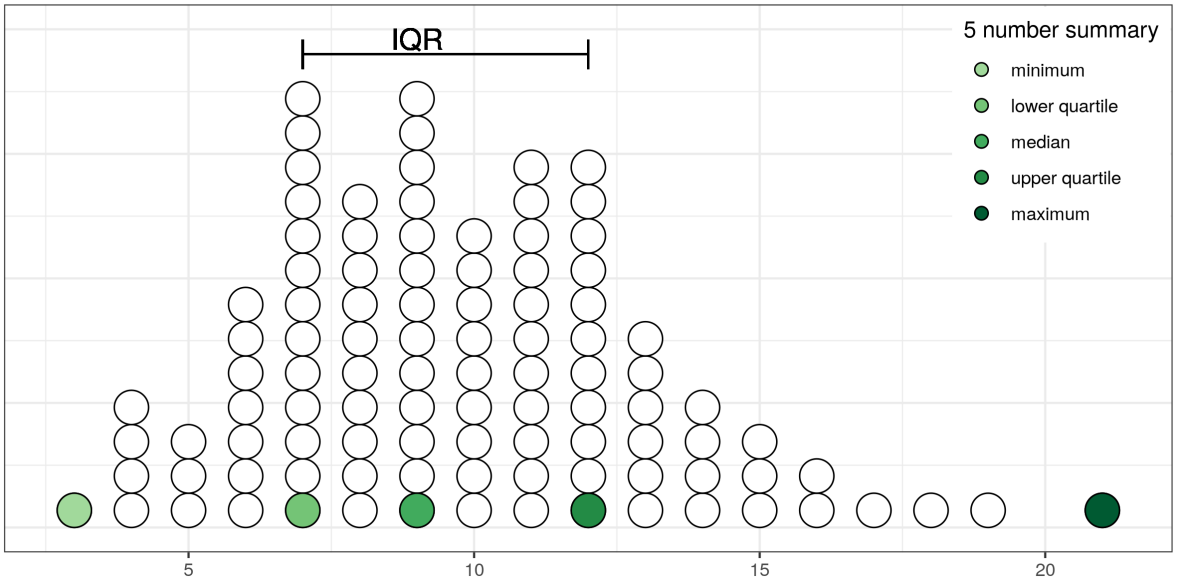
The IQR is resistant (or *robust*) to extreme values and skewness, unlike the standard deviation.

Describing Distributions of Quantitative Data

A **boxplot** (or box-and-whisker plot) is a visual display of the 5-number summary.

- The box displays the middle 50% of the distribution and its width (the IQR) shows the spread of the bulk of the distribution.
- The 'whiskers' extend to the
  - smallest and largest values in the dataset
  - OR the values in the dataset that are within 1.5*IQR away from the edges of the box.
    - Observations outside of the 1.5*IQR range may be shown as dots and are outliers

Describing Distributions of Quantitative Data

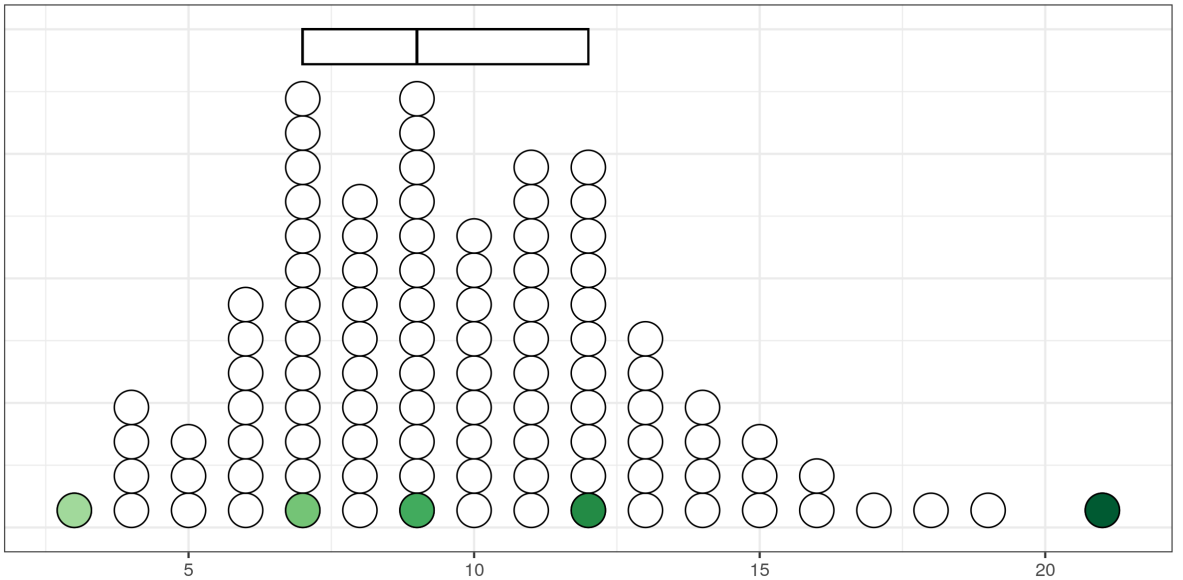


What are the following values?

- bottom whisker:
- lower quartile:
- median:
- upper quartile:
- top whisker:

Are there any outliers?

Describing Distributions of Quantitative Data

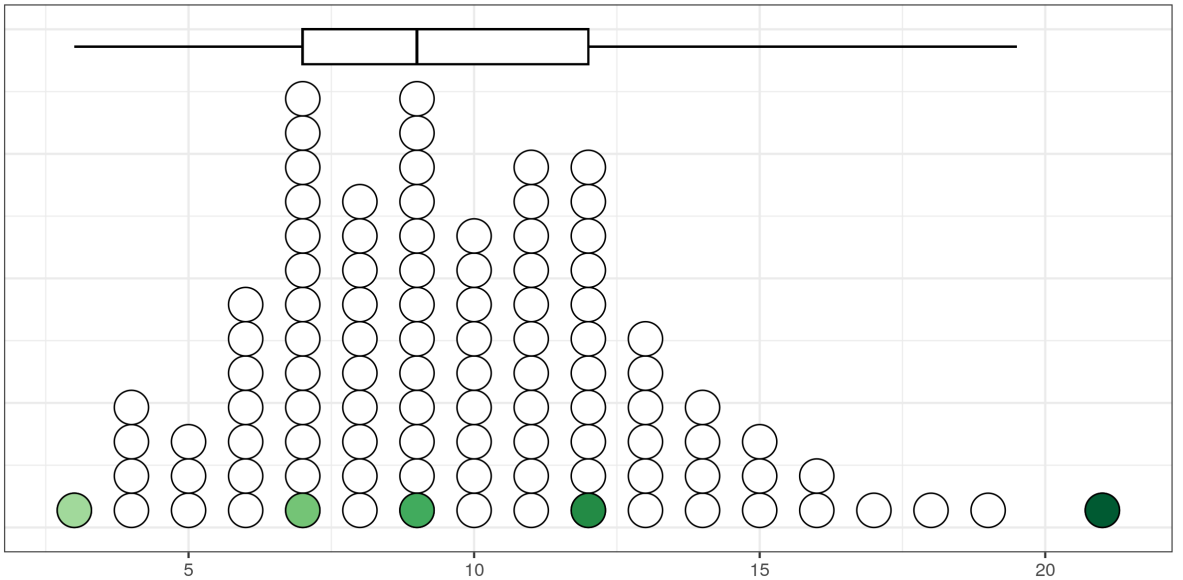


What are the following values?

- bottom whisker:
- lower quartile: 7
- median: 9
- upper quartile: 12
- top whisker:

Are there any outliers?

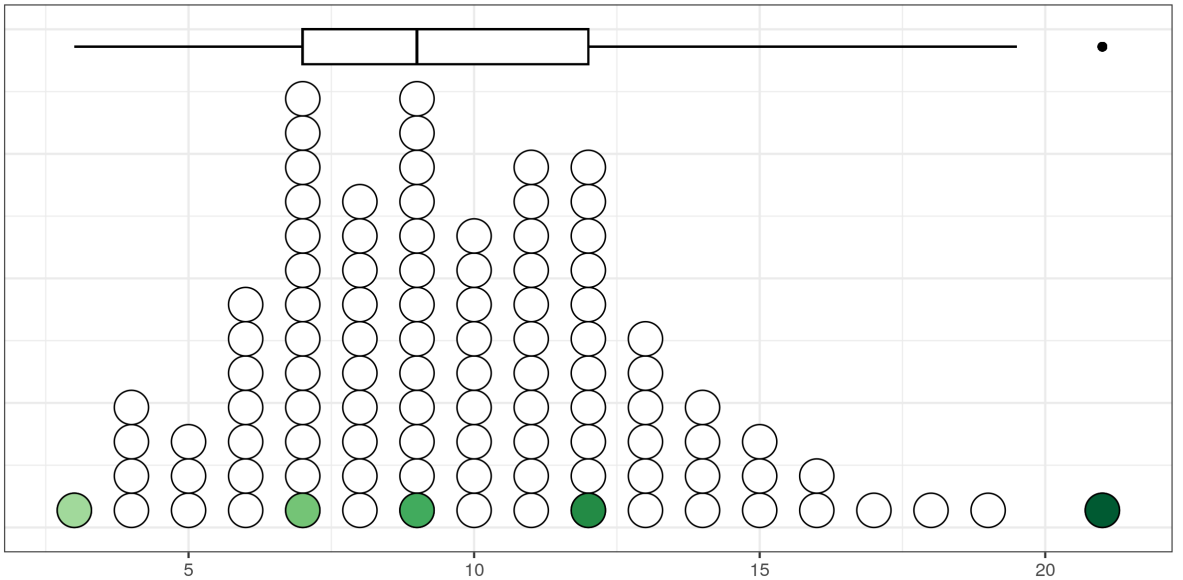
Describing Distributions of Quantitative Data



- bottom whisker: 1.5 x IQR below lower quartile, or smallest observation
  7 - 1.5(12-7) = -0.5 or 3
  bottom whisker: 3
- top whisker: 1.5 x IQR above upper quartile, or largest observation
  12 + 1.5(12-7) = 19.5 or 21
  top whisker: 19.5

Are there any outliers?

Describing Distributions of Quantitative Data



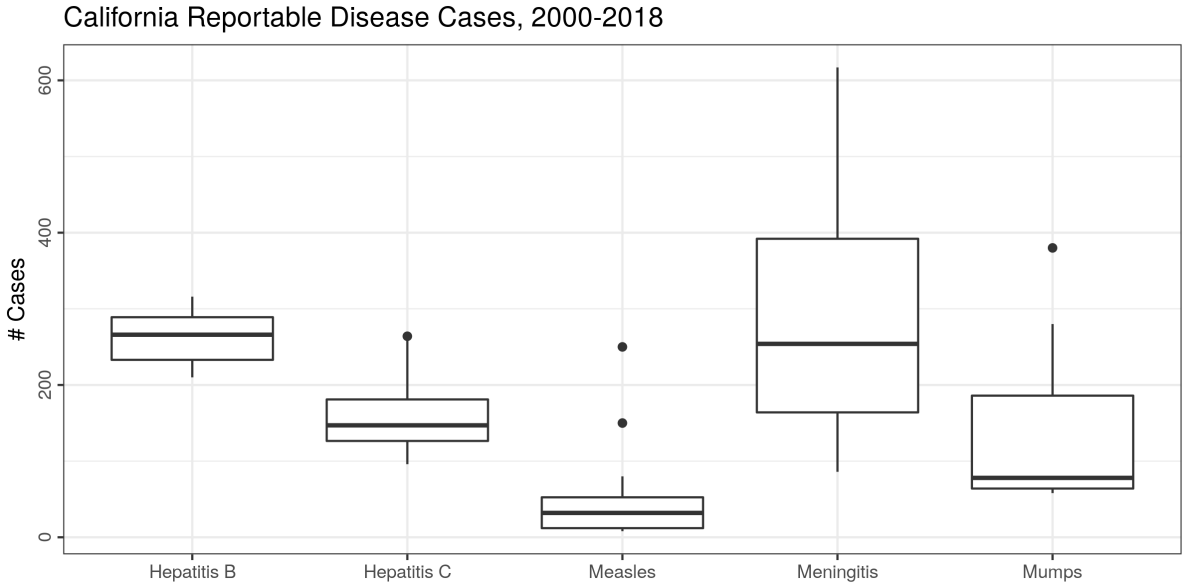
What are the following values?

- bottom whisker: 3
- lower quartile: 7
- median: 9
- upper quartile: 12
- top whisker: 19.5

Are there any outliers? Yes, at 21

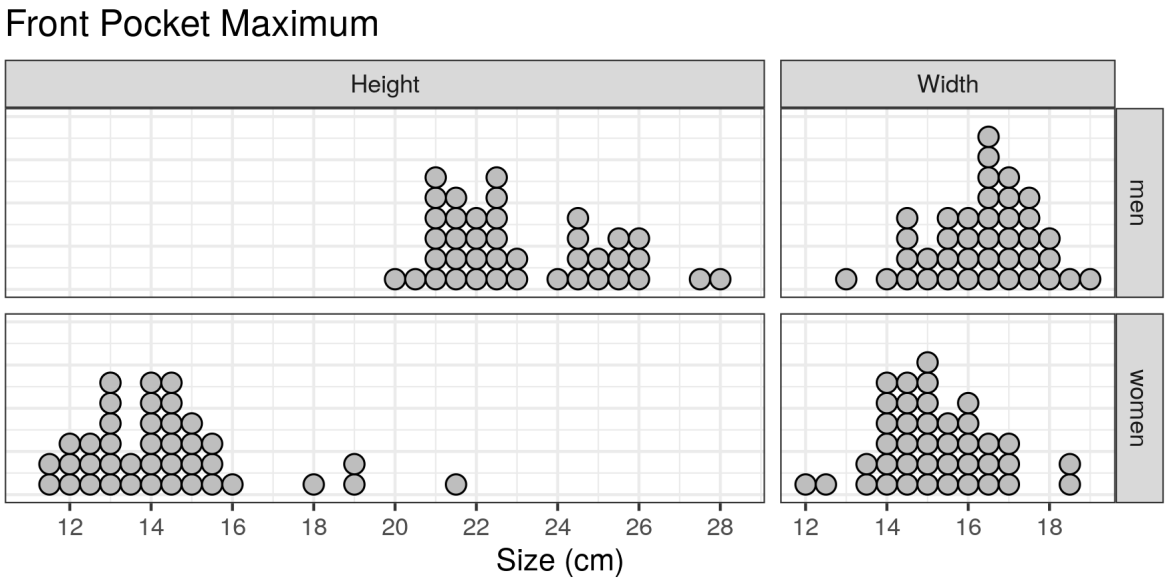
Describing Distributions of Quantitative Data

Boxplots make it easy to compare distributions of different groups.



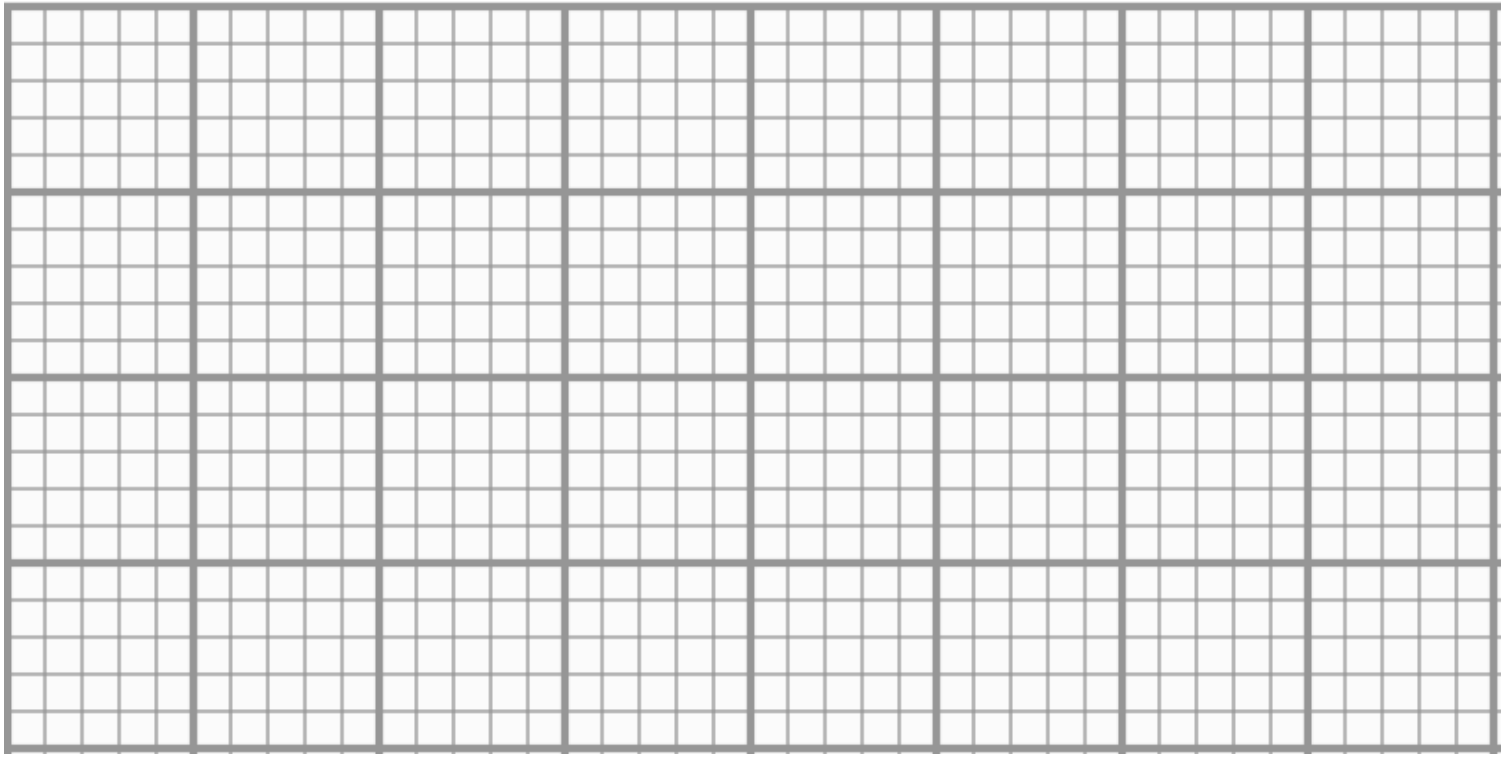
In-Class Practice: Jean Pocket Sizes

Using the graph paper on the next page, create boxplots for each of the dimensions in the chart below.



Data source: https://pudding.cool/2018/08/pockets/

In-Class Practice: Jean Pocket Sizes



What do you conclude about the distributions of front pocket maximum dimensions when comparing Men's jeans to Women's jeans?

Upload your graph and conclusions to Canvas.