Quantitative Data
Graphical Summary: Boxplots

Dr. Brenda Gunderson

Lecturer IV in Statistics and Research Fellow, Statistics,
College of Literature, Science, and the Arts

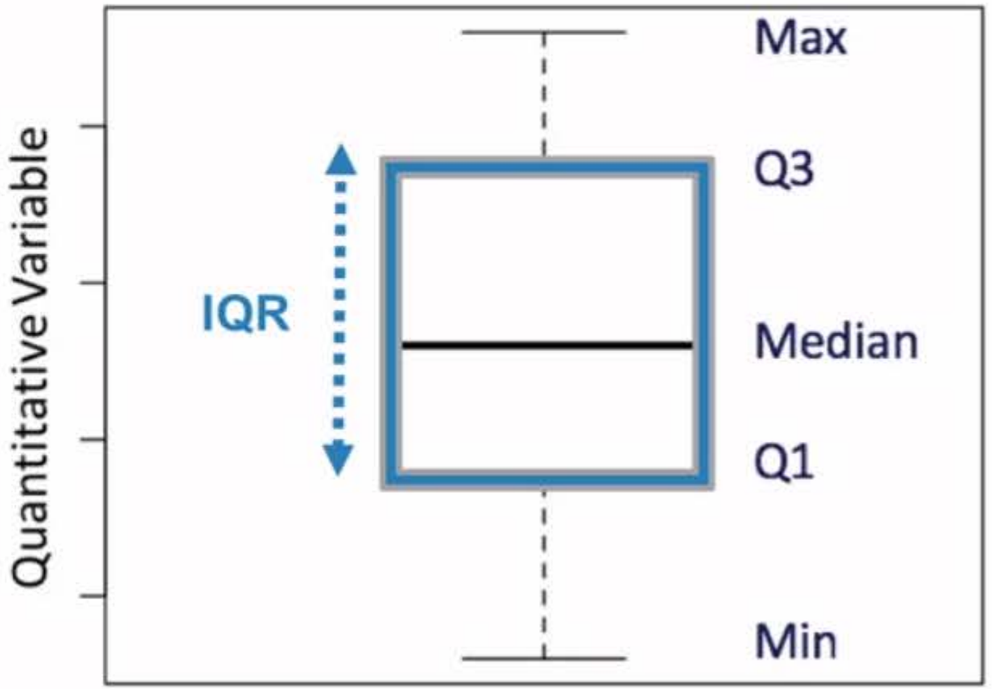
What is a Boxplot?

Five Number Summary

Min	Q1	Median	Q3	Max
		Center		

IQR

Range



So, visually the length of the box is our interquartile range,

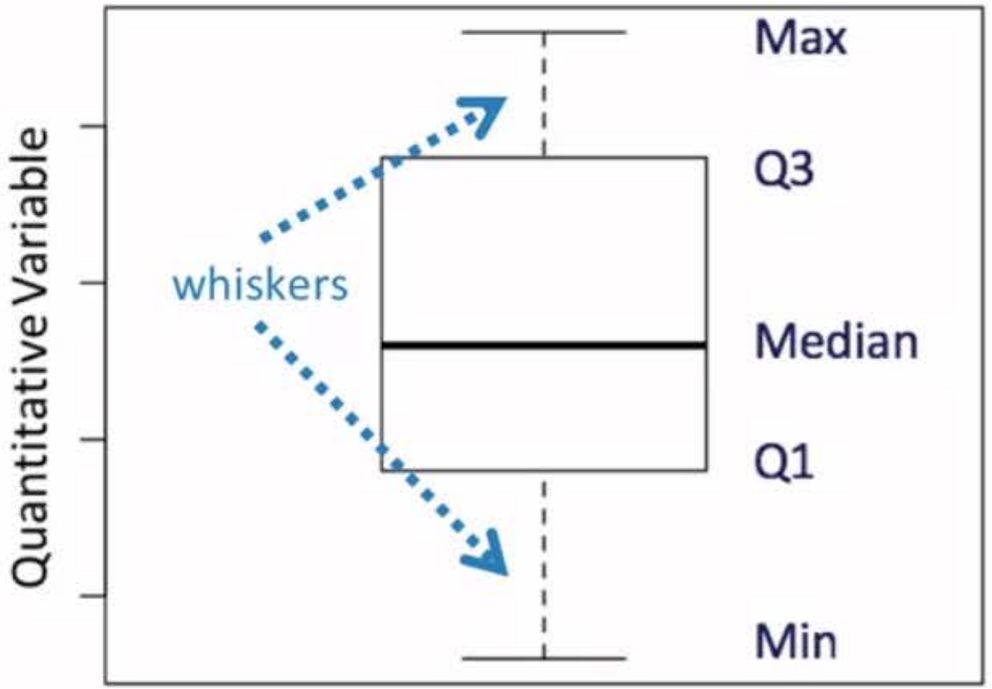
What is a Boxplot?

Five Number Summary

Min	Q1	Median	Q3	Max
		Center		

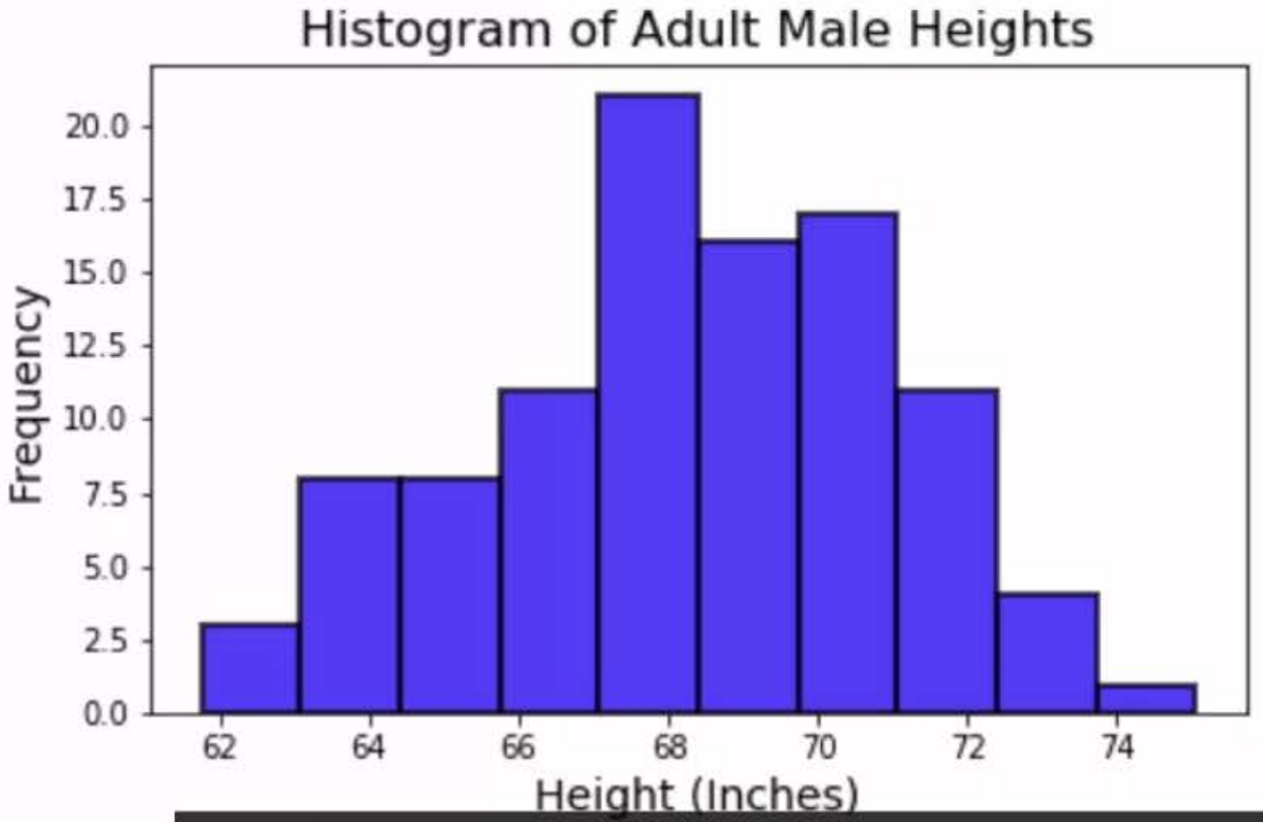
IQR

Range



so these are often referred to as box and whisker plots.

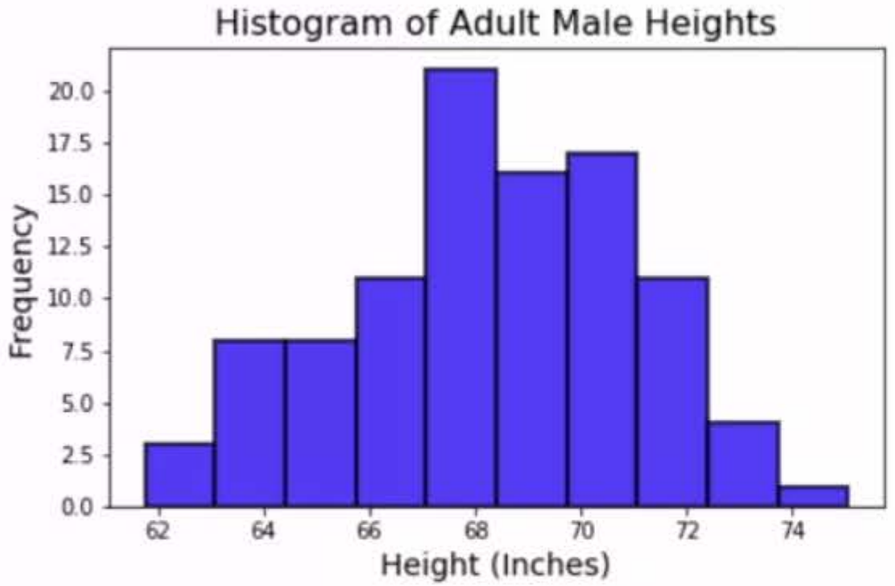
Example 1: Heights of Adult Males



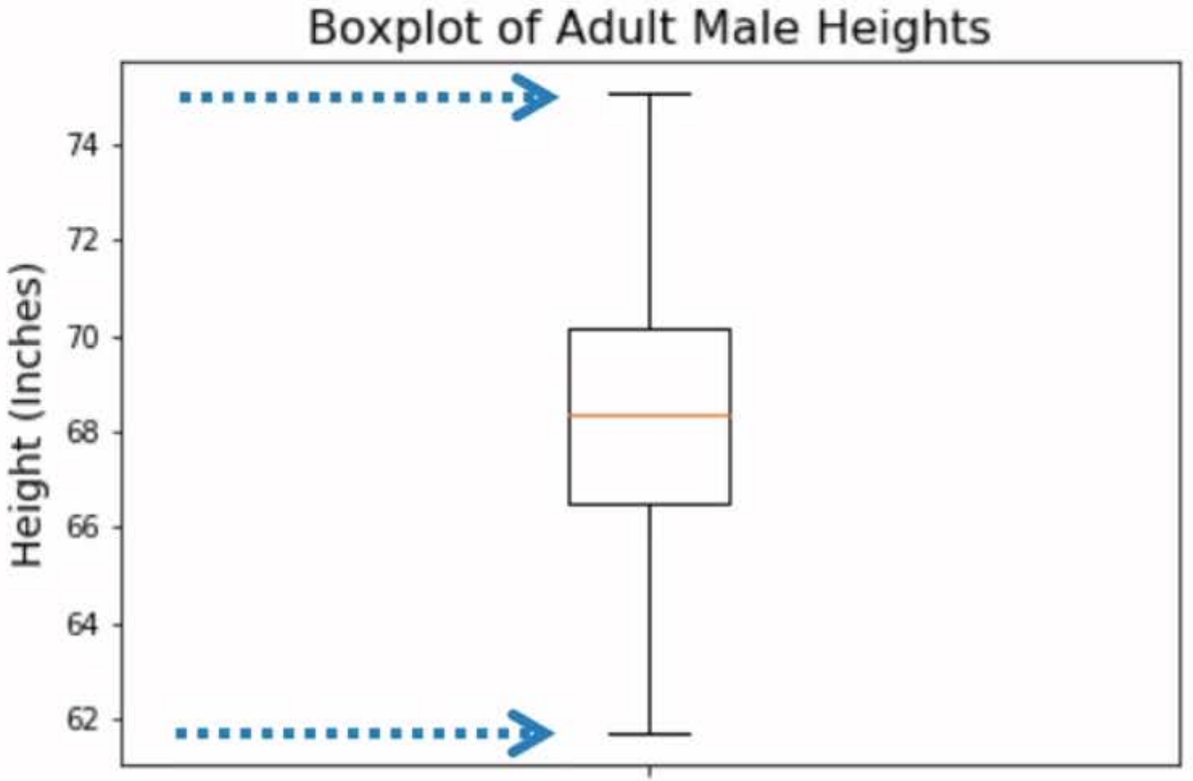

```
     Height
Min.   :61.7
1st Qu.:66.5
Median :68.3
Mean   :68.3
3rd Qu.:70.1
Max.   :75.1
```

We'll take our five-number summary and create the boxplot.

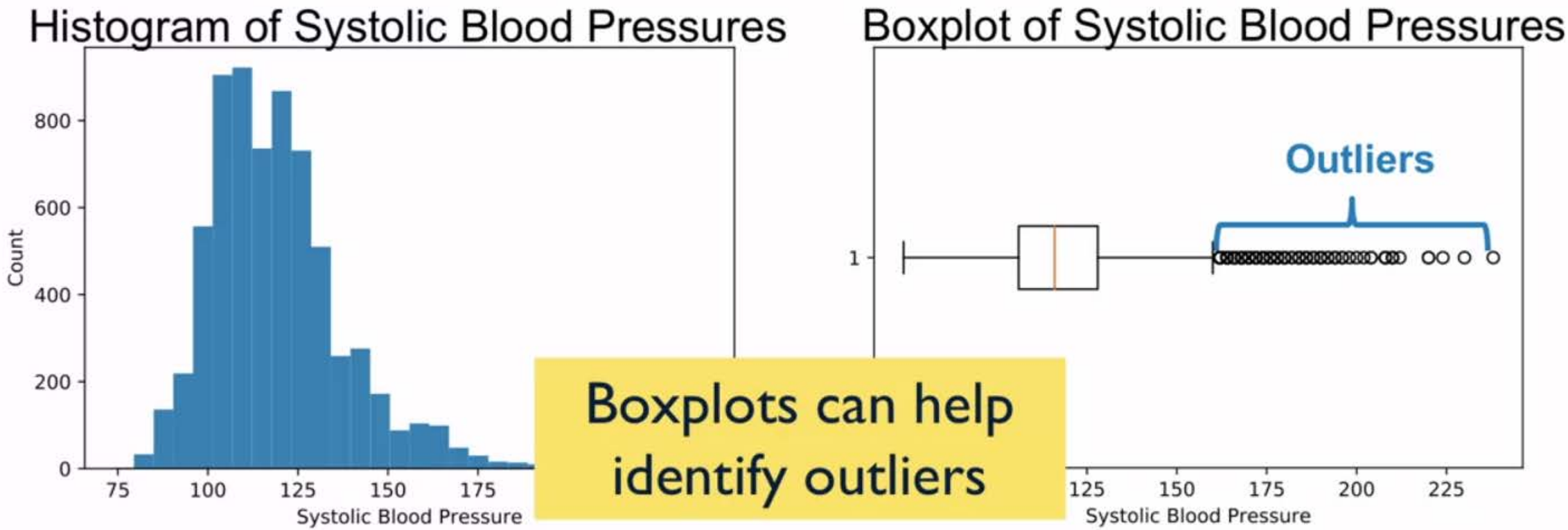
Example 1: Heights of Adult Males




```
     Height
Min.   :61.7
1st Qu.:66.5
Median :68.3
Mean   :68.3
3rd Qu.:70.1
Max.   :75.1
```



Example 2: Systolic Blood Pressures

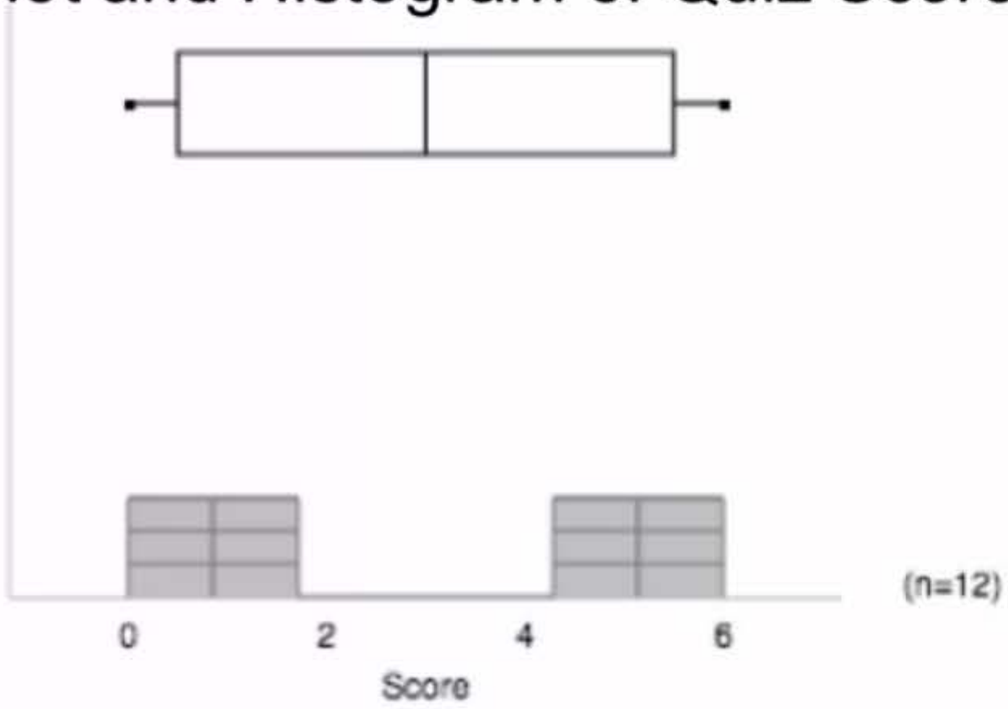


Boxplots can help identify outliers

But, certainly look, we want to take a look at them.

Example 3: Quiz Scores

Boxplot and Histogram of Quiz Scores

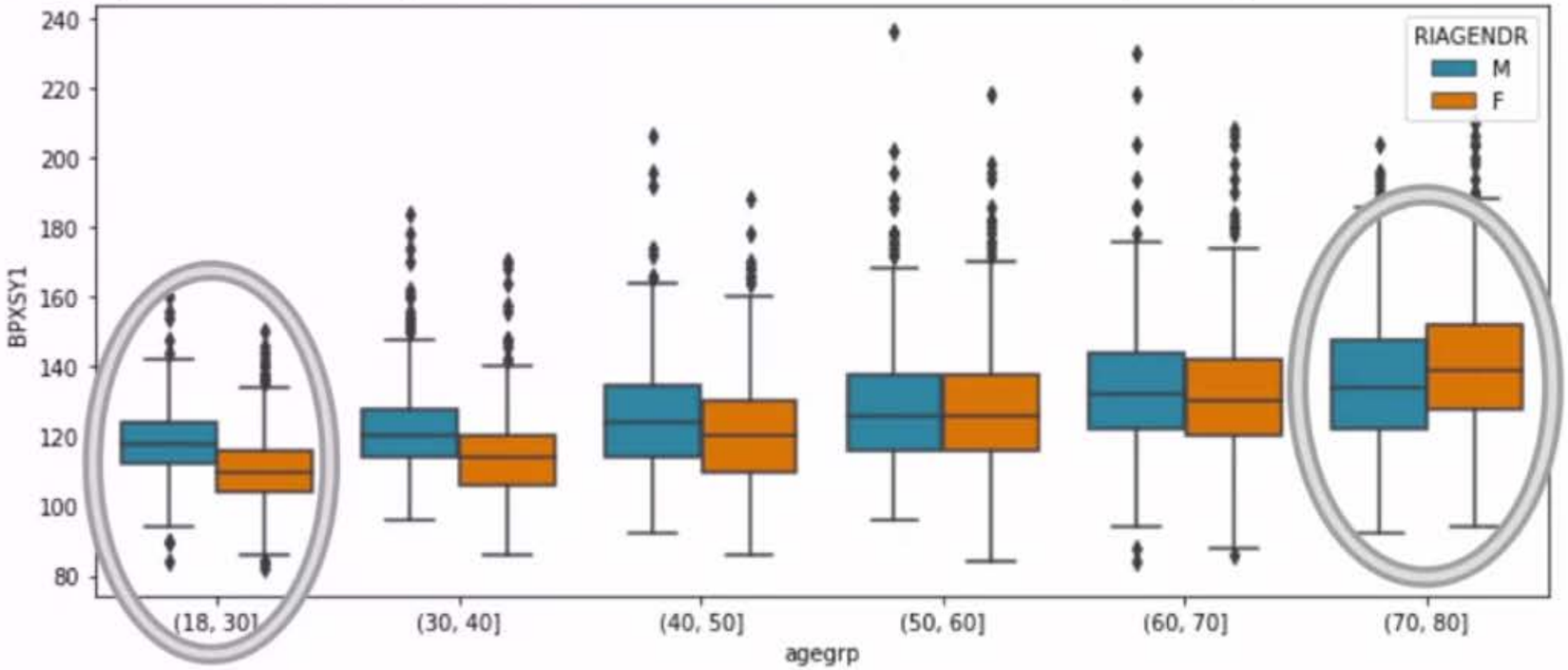


Boxplots can hide gaps and clusters

Our final example, we have a lot of boxplots.

Example 4: Side by Side Boxplots

Boxplots of Systolic Blood Pressure by Age and Gender



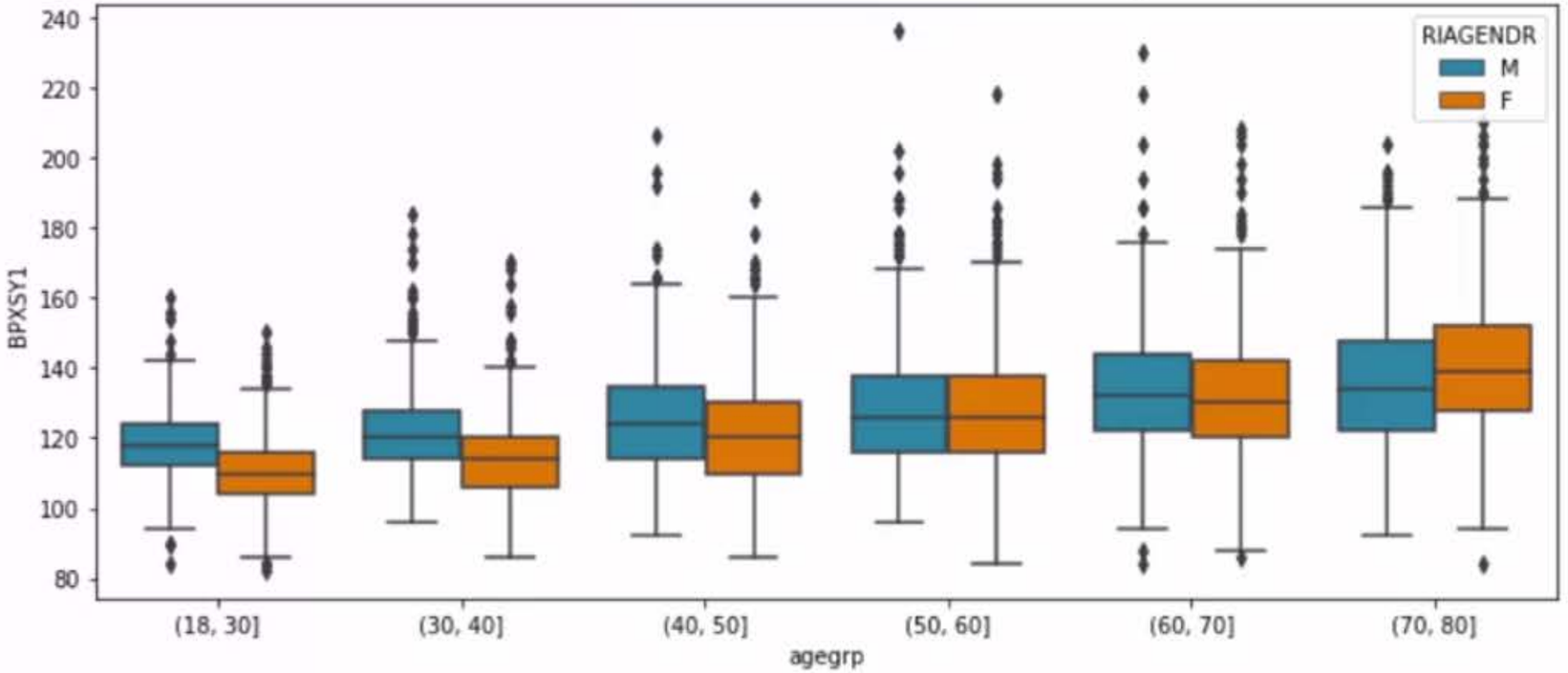
What do you see?

- Old vs Young: BP ***higher, more disperse***
- BP ***higher*** for **young men** vs young women
- BP ***lower*** for **old men** vs old women

are being used or compared in this final age group.

Example 4: Side by Side Boxplots

Boxplots of Systolic Blood Pressure by Age and Gender



Boxplots are useful for comparing sets of observations

useful for making comparisons of sets of observations.

Notes about Boxplots

- Boxplots provide a graphical picture of the five-number summary: showing center (median), spread (IQR and range), and identifies potential outliers.
- Boxplots can hide some shape aspects *(histograms do better job at displaying shape)*
- Side-by-Side Boxplots are useful for comparing two or more sets of observations.