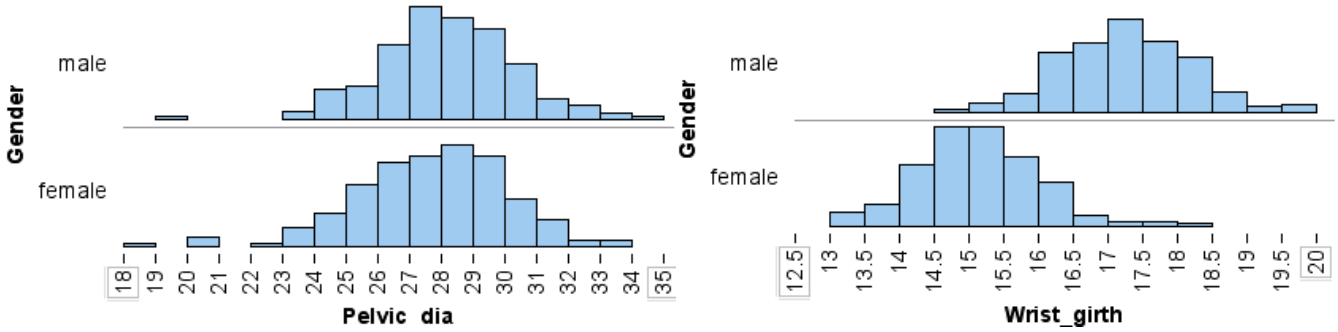
Learning Goal: Compare distributions from two or more groups

Specific Learning Objective: Create and interpret boxplots to compare distributions.

Overview: In this activity we return to comparing and describing distributions. We will again use the ideas of shape, center and spread (as well as noting outliers.). But this time we will use the 5-number summary (Min, Q1, Q2, Q3, Max) to make these ideas more precise.

1) Here is data from 247 men and 260 women who exercise regularly. Pelvic diameter is a measurement from hipbone to hipbone in centimeters. Wrist girth is a measurement around the wrist, also in centimeters.



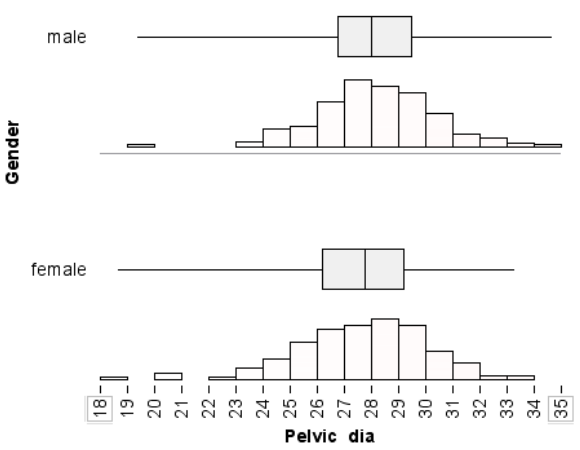
a) Fill in the blanks with either *pelvic diameter* or *wrist girth*.

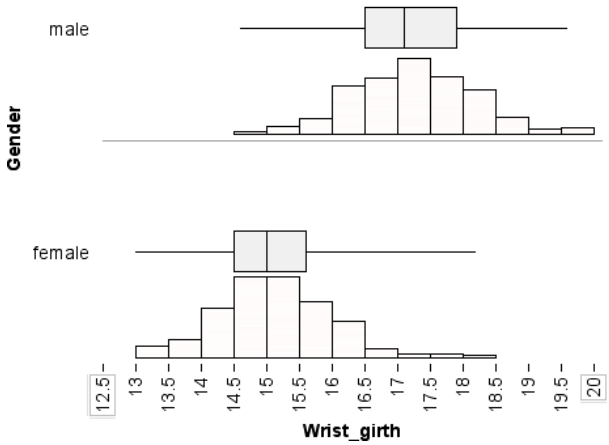
For the variable _____, men and women have similar measurements.
For the variable _____, men are substantially larger than women.

Previously, when we compared distributions of a quantitative variable, we described the shape and gave eyeball estimates of the center (a single typical measurement) and of the spread (both overall range and an interval of typical measurements.) Now we can make these estimates more precise using the 5-number summary as follows:

- The median can represent a typical measurement.
- The interval Q1 to Q3 can represent an interval of typical measurements.
- The interquartile range (IQR), which is Q3 minus Q1, describes the variability in the middle half of the distribution.

Here are the boxplots and the 5-number summaries for these distributions.



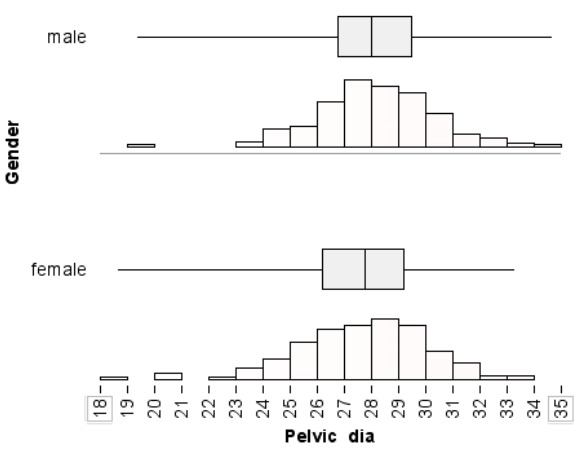


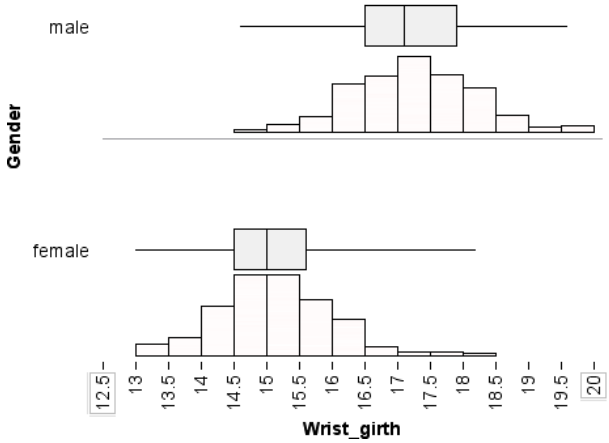
Pelvic Diameter (cm)						
Gender	N	Min	Q1	Q2	Q3	Max
Female	260	18.7	26.2	27.8	29.2	33.3
Male	247	19.4	26.8	28	29.5	34.7

Wrist Girth (cm)						
Gender	N	Min	Q1	Q2	Q3	Max
Female	260	13	14.5	15	15.6	18.2
Male	247	14.6	16.5	17.1	17.9	19.6

b) For the distributions that you identified as similar, compare the shapes, centers (medians and intervals of typical measurements) and spread (IQR and overall range). Make other observations using quartiles that support your conclusion. Then write a paragraph (or paragraphs) summarizing your observations in a way that supports your conclusion.

Here are the boxplots and the 5-number summaries for these distributions again.





Pelvic Diameter (cm)						
Gender	N	Min	Q1	Q2	Q3	Max
Female	260	18.7	26.2	27.8	29.2	33.3
Male	247	19.4	26.8	28	29.5	34.7

Wrist Girth (cm)						
Gender	N	Min	Q1	Q2	Q3	Max
Female	260	13	14.5	15	15.6	18.2
Male	247	14.6	16.5	17.1	17.9	19.6

c) For the distributions that you identified as substantially different, compare the shapes, centers (medians and intervals of typical measurements) and spread (IQR and overall range). Make other observations using quartiles that support your conclusion. Then write a paragraph summarizing your observations in a way that supports your conclusion.