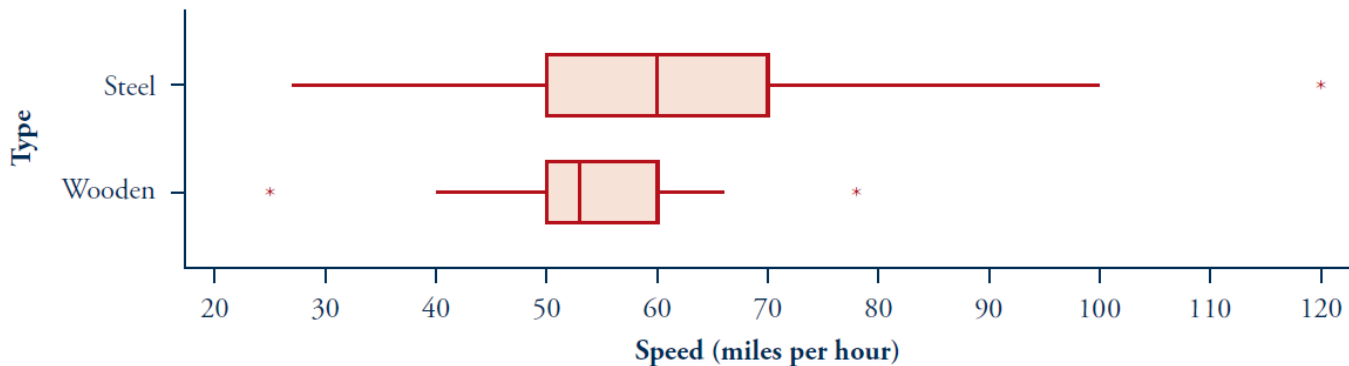


## Rollercoaster Boxplots

The following boxplots display the distributions of top speeds (in miles per hour) for roller coasters in the United States, classified by whether the coaster is wooden or steel:



Observational unit: rollercoasters in the United States.

Explanatory variable: Material made of      Type: Categorical - Binary

Response variable: Top Speed                      Type: Quantitative

- a. What proportion of the steel coasters have a top speed of 60 miles per hour (mph) or greater? Explain how you can tell from the boxplots.

**50 percent of the steel coasters have a top speed of 60 miles per hour, or greater. We see this since the median is at 60, and therefore half of the data points are greater than 60.**

- b. What proportion of the wooden coasters have a top speed of 60 mph or greater? Explain how you can tell from the boxplots.

**25 percent of wooden coasters have a top speed of 60 miles per hour or greater. We can see this since the upper quartile is at 60, and therefore one quarter of the data points are greater than 60.**

- c. Which type of coaster (steel or wooden) has a higher proportion of coasters with a top speed greater than 50 miles per hour, or are the 2 types of coasters the same, or is it possible to determine from the boxplots? Explain how you can tell (or why you cannot tell) from the boxplots.

**The proportion of coasters with top speed over 50 miles per hour are the same for both types of coasters. We can see this on the box plots since 50 miles per hour occurs at the lower quartile for both types of coasters. Meaning that 75 percent of all coasters have Speeds over 50 miles per hour, regardless of whether they are wooden or steel.**

- d. Which type of coaster has a higher proportion of coasters w/top speed greater than 45 miles per hour, or are the 2 types of coasters the same, or is it impossible to determine from the boxplots? Explain how you can tell (or why you cannot tell) from the boxplots.

**We cannot tell. The reason is that 45 mph occurs between the lower quartile and the minimum for both coasters. And between these numbers, we don't have information about how the data is distributed. So we do not know if more of the wooden coaster data is above 45mph, or more of the steel coaster data is above 45 miles per hour within this range of data.**

- e. Which type of coaster (steel or wooden) has more variability in its speeds, or is the variability in the 2 types of coasters the same, or is it impossible to determine from the box plots? Explain how you can tell (or why you cannot tell) from the box plots.

**This steel coaster has more variability. We see this by looking at the interquartile range. For the wooden coasters, this is 10. For the steel coasters, we see it is 20.**

- f. Which type of coaster is more prevalent, or are there equal numbers of both types, or is it impossible to tell from the boxplots? Explain how you can tell (or cannot) from the boxplots.

**Box plots don't give us information about the # of data pts, only the % of those pts that are in particular ranges. So, we don't have information about the prevalence of these coasters.**

Build a boxplot for these 5-number summaries on the chart below (use a straight edge for your range):

	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
Advanced Compact	63	69	70	73	78
Compact	62	65	71	73.5	76
Subcompact	53	61.75	65.5	69.25	75
Super-Zoom	66	72.25	75.5	79	81

