

1. The GPAs of the 345 students in the *StudentSurvey* dataset had an average of $\bar{x} = 3.16$ and a standard deviation of $s = 0.4$.

- a. Compute the z -score for a 2.5 GPA.

Using the definition, $z = \frac{2.5 - 3.16}{0.4} = -1.65$.

- b. Approximate the percentile rank for a 3.6 GPA.

The z -score for a 3.6 GPA is $z = (3.6 - 3.16) / 0.4 = 1.1 \approx 1$, and we know that roughly 68% of GPAs have z -scores between -1 and 1. Since about 16% of z -scores are less than -1, roughly $68\% + 16\% = 84\%$ of GPAs are less than 3.6.

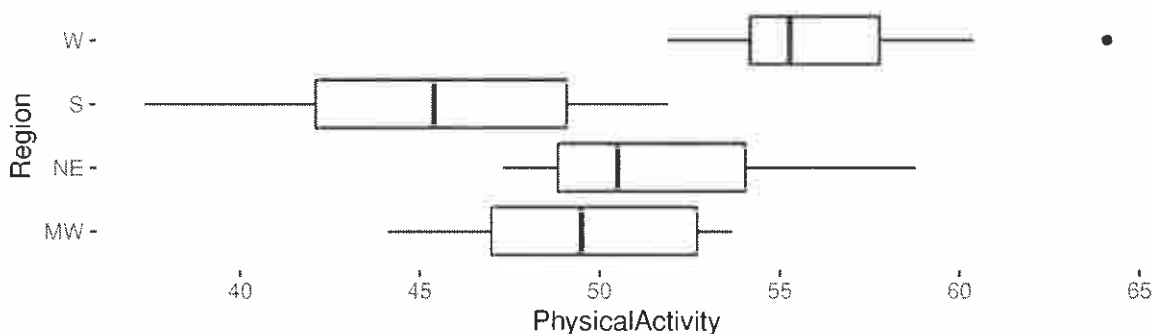
- c. About what percentage of the GPAs in this dataset are likely between 2.8 and 3.6?

The z -score for a 2.8 GPA is $z = (2.8 - 3.16) / 0.4 = -0.9 \approx -1$, and the z -score for a 3.6 GPA is close to 1. It follows that about 68% of GPAs are between 2.8 and 3.6.

- d. Determine an interval of GPAs that likely contains about 95% of the GPAs in this dataset.

About 95% of the GPAs will be between $3.16 - (2 \times 0.4) = 2.36$ and $3.16 + (2 \times 0.4) = 3.96$.

2. The *USStates* dataset includes the columns *Activity*, the percentage of people in each state who engage in at least 150 minutes of physical activity per week, and *Region*, the region of the country in which the state is located (West, South, Northeast, Midwest).



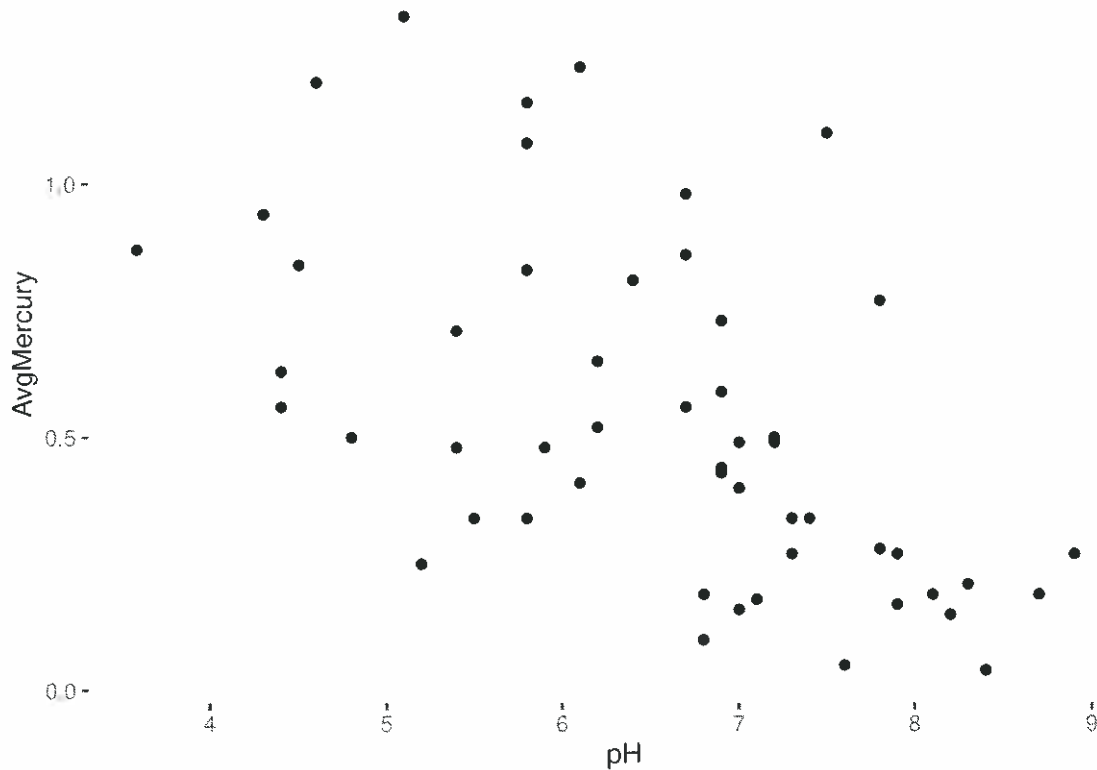
- a. Which region shows the lowest level of physical activity? Estimate the median activity level for this region.

The South is the least active region; the median for this region is about 45%.

- b. Which region shows the highest level of physical activity? Estimate the median activity level for this region.

The West is the most active region; the median for this region is about 55%.

3. Here is a scatterplot of the average mercury and pH levels of 53 lakes in Florida (taken from the aptly named *FloridaLakes* dataset):



- a. Which of the following is the correlation between these two variables? Circle your choice.

-0.925 -0.575 -0.115 0.115 0.575 0.925

- b. Explain your choice and what it says about the relationship between these two variables.

There is clearly a negative relationship between pH and AvgMercury, which means that AvgMercury level tends to decrease as pH increases. This negative relationship means that the correlation is negative; we can eliminate -0.925 for being too close to -1 (a perfect linear relationship), and we can eliminate -0.115 for being too close to 0 (no relationship). The remaining negative choice, -0.575, is the correlation between these two variables.