**Shoes: What is Your Size & How Many Pairs do You Own?**

By: Taylor and Brooke

 Our survey topic was “ What is Your Shoe Size and How Many Pairs of Shoes do you Own?” Before conducting our survey, we decided to do a systematic sampling of all Fowler High School students by asking every 5th student in each grade level. We began by asking students in the library, then sent out a Google Form to the students we weren’t able to find. Our explanatory variable was the person’s shoe size, and our response variable was how many pairs of shoes the person owned.



The data set of shoe size and the number of pairs of shoes a person owns has a weak positive correlation. This was unexpected, because we predicted that as a person’s shoe size increased, the fewer pairs of shoes they would have owned. Therefore, instead of a predicted strong negative correlation, the data set followed a weak positive correlation.

The correlation coefficient *r* is .285. This means, in terms of our data set, that a person’s shoe size has a weak positive correlation compared to the number of pairs of shoes that a person owns. The $\overline{x}$, or the mean of the shoe size, is 9.41. The $\overline{y}$, or the mean of the number of pairs of shoes owned, is 12.66. The significance of the point ($\overline{x}$,$\overline{y}$) is that it is the mean for the data set of a person’s shoe size and the number of pairs of shoes they own, so the point is guaranteed to fall on the line of best fit. The marginal change for our data set is 1.567, which means that each time the shoe size changes by one unit, the number of pairs of shoes owned changes by 1.567. An influential point is (18, 69) because the x-value is an extreme high compared to other data values. The formula for the least squares regression line is 1.567x - 2.083. The coefficient of variation $r^{2}$ for our data set is .081. The percentage of explained variation in our data set is 8.1%, while the percent of unexplained variation in our data set is 91.9%. What this says about our data set is that, of the data set of a person's shoes size and the number of pairs of shoes they own, 91.9% has the potential to be affected by lurking variables.

 One possible lurking variable that could have affected our study on the correlation between a person’s shoe size and the number of pairs of shoes they own is if they have different shoe sizes according to brand. Another possible lurking variable could be if they did or did not include pairs of shoes shared with other people. When predicting what the number of pairs of shoes a person would own using interpolation, we used the linear regression formula. We chose a shoe size of 6.5 and plugged it into the formula, which said that a person should own 8.1 pairs of shoes, which would be rounded down to 8. When predicting what the number of pairs of shoes a person would own using extrapolation, we used the linear regression formula. We chose a shoe size of 15 and plugged it into the formula, which said that a person should own 21.4 pairs of shoes, which would be rounded down to 21.

 Shoe size and the number of pairs of shoes owned are not closely related. They share a weak positive correlation. We predicted that when shoe size increases, a person will have more difficulty finding shoes in their size, resulting in a moderate to strong negative correlation. This was not the case. We learned that shoe size cannot easily determine the number of pairs of shoes that a person would own.