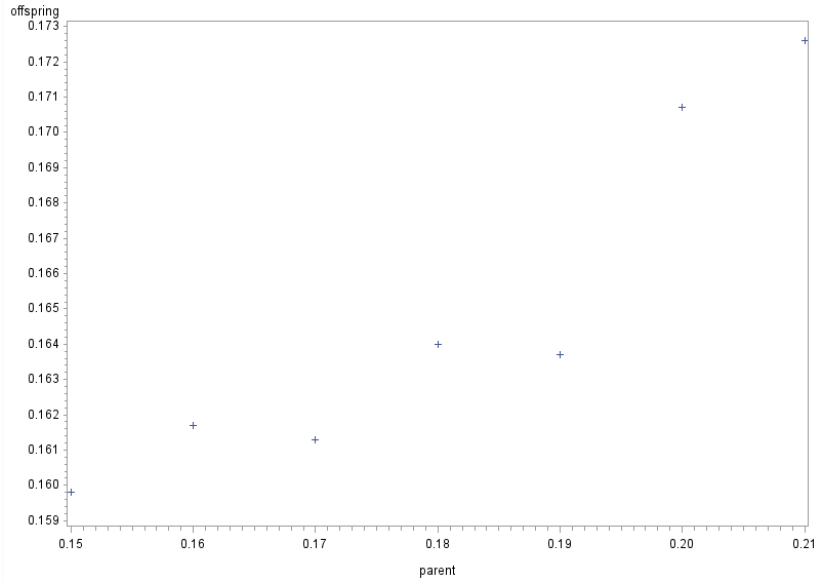
Galton's Peas Write Up

Introduction:

In 1877 Galton presented data from an experiment performed on sweet peas. For different diameters of sweet pea parents, he measured the diameter of the offspring sweet peas. He observed seven pairs of parent and offspring peas, trying to predict what offspring size will be based on the parent's size.

We calculated the correlation coefficient to be 0.92488, showing a strong correlation. This can be easily visualized by creating a scatterplot.



Analysis:

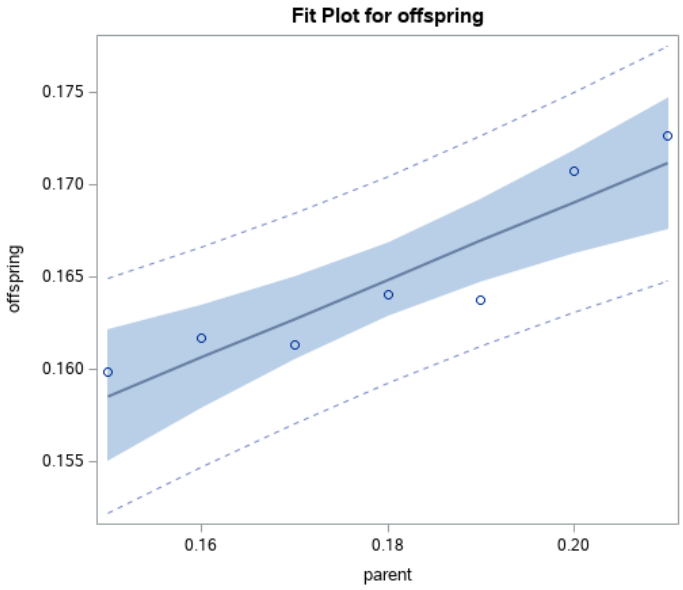
Our response variable is the diameter (in) of the Offspring pea, and the explanatory variable is the diameter of the Parent pea. The model we created is as follows:

$$Offspring = \beta_0 + \beta_1 * Parent + \dot{\varepsilon}, \dot{\varepsilon} \sim N(0, \sigma^2)$$

Upon creating the model, we find the parameter estimates and standard error as shown below:

Variable	DF	Parameter Estimate	Standard Error
Intercept	1	0.12703	0.00699
parent	1	0.21000	0.03861

The value of $\hat{\beta}_1$ is 0.21, which means that for a 1 inch change in parent diameter, offspring diameter will increase 0.21 inches on average.



Paying attention to only the solid black line and blue shaded area in the diagram to the left shows our level of uncertainty. The blue area represents our 95% confidence interval, where we are 95% confident that future values of pea diameters will fall within this range.

To test our model and prediction, I was asked to make a prediction for a parent pea with size 0.2 inches, and the offspring's predicted size comes out to be 0.1690. I also made a prediction with parent diameter 0.18 in. This can be represented in the above graph, where instead of focusing on the blue area, focus on the dotted lines for the prediction interval. To see how well the prediction performed I calculated the R^2 value, which came out to be 0.8554, proving a good prediction. All in all, there is a statistically significant inheritance effect.