Avocados

Cameron Hartling

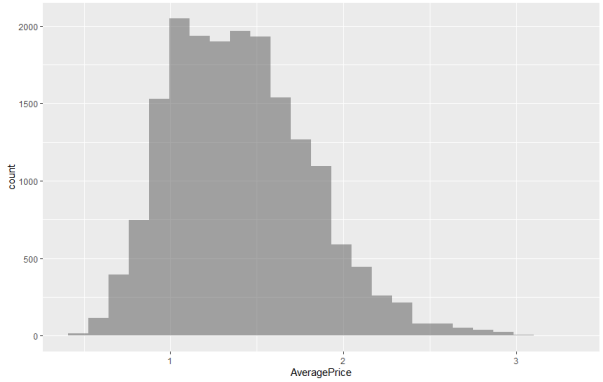
Research Question

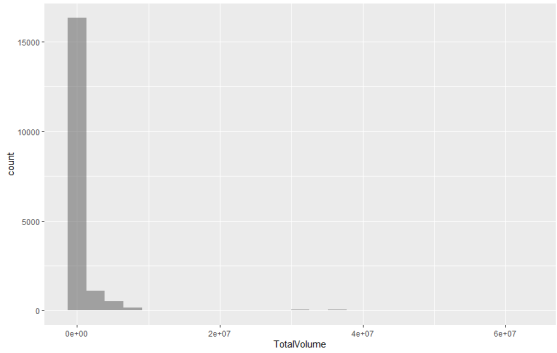
Does the average price of an avocado affect the number of avocados sold?

Methods

I went on Kaggle and found a dataset owned by the user JUSTIN KIGGINS on avocado sales from January 2015 to December 2018. There are 156 instances in the sample size. One instance for each week recorded. I used regression analysis to find how the average price of avocados affected the number of avocados sold. The null hypothesis is that the slope of the regression line is zero and the alternative hypothesis is that the slope of the regression line is not zero.

Results

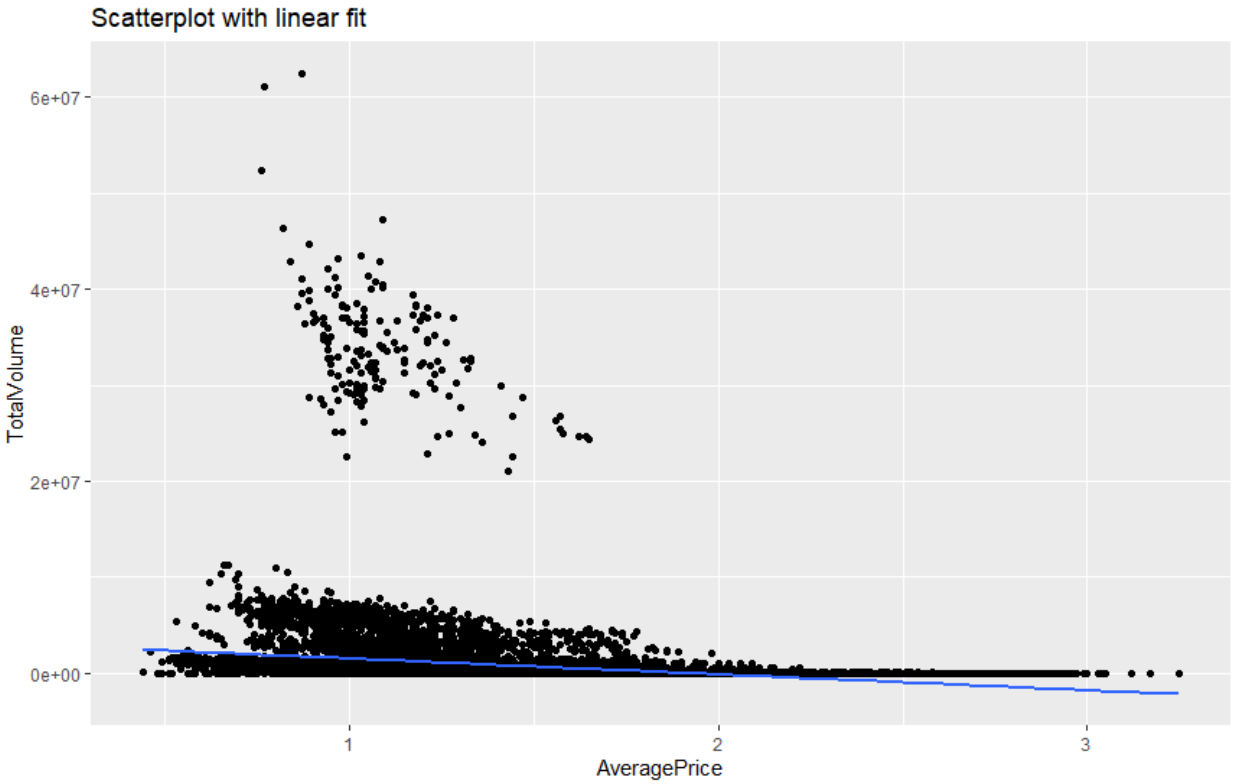




The average price has a normal distribution while the total volume is skewed to the right. This shows that the total number of avocados sold does not change very much while the average price of avocados can vary. There were a few outliers in the total volume graph.

The linear regression returned the correlation coefficient of r = -0.1928 and a regression line Total Volume = 3174918 + -1653136 * Average Price. The p-value for the t-test on the slope coefficient is $2.2*10^{-16}$. Since the p-value is so small this indicates that the

slope is non zero. R^2 = 0.0376 which indicates a high variance and a bad linear regression fit.



The average price of avocados does affect the total number of avocados sold.

Discussion

There are several issues of the research. I had no control of the data collection methods so I cannot confirm their accuracy or if any ethical issues arose. To prevent these issues I could collect the data on my own. I believe that a larger data set would have proved useful for getting a better regression line fit.