LeRoy Rosenbohm

We will start with the heat map. It has some interesting details

* Alcohol is positively correlated with the quality of the red wine.
* pH negative correlation suggesting less pH equals higher quality wine
* Interesting how you can break down the quality of wine into the science of “how much or how little of this” in a basic sense.

**The squares with positive values show direct co-relationships between features. The higher the values, the stronger these relationships are. That means, if one feature increases, the other one also tends to increase, and vice-versa.**

**The squares that have negative values show an inverse co-relationship. The more negative these values get, the more inversely proportional they are. This means that if the value of one feature is higher, the value of the other one gets lower.**

**Finally, squares close to zero indicate almost no co-dependency between those sets of features.**

1. **This shows pH versus quality distribution before outliers were taken out. You’ll see that the median of pH looks to be around 3.3 for those quality levels.**
2. **The box plot maps the inverse relationship between quality and pH. As you can see the higher the quality, the lower the pH level tends to be.**
3. **Over the next 2 slides we have histograms for each category of the data set. Most of them are right-skewed. We applied a boxcox transformation which adjusts the data so that it more closely resembles a normal distribution.**
4. **This is the result is our dataset and the variables over a more normal distribution with outliers removed. Judging from this, sulfer dioxide might be the best way to test for wine quality as it spans a good range for testing and it’s positively correlated to the quality as opposed to the pH.**