Name: Hagay Ringel

Report- WineQT

In the next report, I will demonstrate how each plot shows a different relationship between the components of the wine, as taken from the "WineQT" database and I will do this by using charts and graphs of different types so that the information presented is clear and visually legible.

In plot 1, I used a basic graph called "line" in order to determine whether or not there is a pattern between chlorides and alcohol. According to this plot, the less chlorides there is in the wine, the more alcohol content there is.

In plot 2, focusing on acidity, we can see that the major amount of the wine has less than a rate of 10 in the fixed acidity and less than a rate of 0.6 in the citric acid. It means that the overall acidity amount in the wine is low.

Plot 3 shows the relationship between the fixed acidity to the total sulfur dioxide, and we can see that wines with a low rate of fixed acidity are more likely to have a high total sulfur dioxide rate.

In plot 4 I wanted to examine what is the pattern of alcohol and acidity: as we clearly see in the graph, that the volatile acidity trend is stable and low and getting higher only when the alcohol curve is between 0.0 – 1.0.

In plot 5 we can see that there is no relationship between the fixed acidity rate and the sulphates rate as these two variables have a stable and fixed rate.

Plot 6 presents the relation of quality and alcohol. Here, the goal is to predict the quality of the wine upon the alcohol rate. As we can see at this plot, wines with an

alcohol rate of more than 11, will be more likely to have a high quality rate (7 and above).

In plot 7 and the next plots I took the first 30 values and as shown in the plot, I was trying to find a connection between the pH and total sulfur dioxide rate. As we see the dots in the graph- there is a wide scattering, therefore, there is no connection between pH and total sulfur dioxide.

Plot 8 is about whether a residual sugar rate influence the quality of the wine or not. According to this plot, wines with a residual sugar rate of less than 2.1 will be more likely to have a high quality rate of 7 and above than these with a residual sugar value of 2.1 and above.

By taking the first 30 values in plot 9 as well as the previous one, we can see an interesting trend as only when the citric acid value is 0.51, the value of the chlorides is more than 0.1. In all the other case, it's less.

In plot 10, we examine the relationship between chlorides and citric acid as well as plot 9, but in a different visualization way- and we can see the same trend but in clearer pattern as most of the values are concentrated in the circle of 0.00-0.15 for the chlorides and -0.2-0.6 for the citric acid. The scattered value in the small circle is probably fallen behind the standard deviation.

Part 2:

As part of my work in procurement for a governmental office, my main job is communicating with suppliers for the benefit of the organization for which I work. Procurement work requires recognition of the main market of the organization for which I am purchasing for, communicating in other languages, heightened interpersonal communication and the ability to analyze market trends, using several of computerized systems and financial analysis.

SAP is the main system in which I work with, and in my daily routine I work with other systems and applications for analyzing purchase history, consuming statistics, pricing analyzing and knowing the market trends.

I work with people who worked in my office in the 1980s and according to what they say the work was completely different, from analyzing paperwork to making connections with YELLOW PAGES and days of conversations to reach relevant professionals, just like LINKEDIN and the social employment revolution created, that's how my office transformed By using DATA and smart software.

As mentioned in the article about the Red Sox, statistics is highly important (same as in sport) and by using statistics and data analyzing, I make a statistical assessment of resource utilization for future procurement, I know today what the status of my resources will be in three years from now and I will prepare the ground for me and my future employees, through which we build a procurement strategy, we even have a department designed for this.

The main negative consequence of the Big Data is the over data and the challenge to sort and organize this data. There is a huge amount of data that needs to be stored,

sorted and organized in order to be used as mentioned above, therefore, a lot of energy, human resources and time are spent for that and the current challenge is to do that in the most effective pattern and in a minimum time range.