**Homework week 8 – Data Visualization Emmy Pahmer 11-NOV-2019**

**Data**: wine

**Description**:

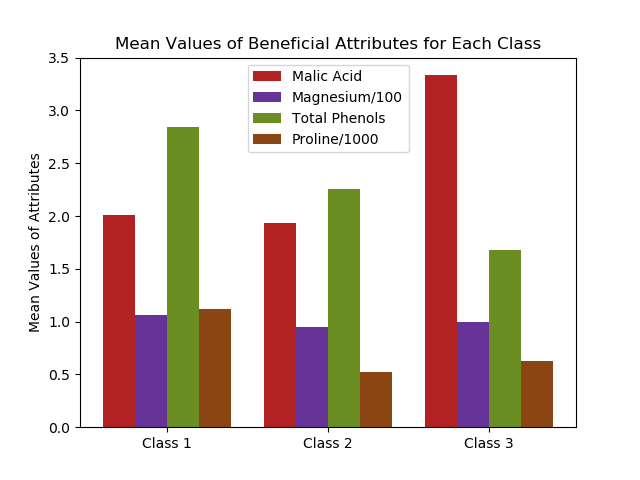
Chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. All attributes are continuous. 1st attribute is class identifier (1-3)

**Data:**

1. Alcohol
2. Malic acid – sour/acid flavour, used for treating fibromyalgia
3. Ash -  inorganic matter that remains after evaporation and incineration (eg. potassium, sodium, calcium, magnesium, iron, copper, lead, arsenic, etc.). Ash content is one of the important indicators in wine quality determination
4. Alcalinity of ash - defined as the sum of cations, other than the ammonium ion, combined with the organic acids in the wine.
5. Magnesium - Alcohol causes a rapid loss of magnesium, so beer and wine are not the recommended way of getting Mg; none-the-less, beverages containing more Mg are healthier than beverages containing less Mg.
6. Total phenols – antioxidant
7. Flavanoids – antioxidant
8. Nonflavanoid phenols – antioxidant. Known to enhance and stabilize the color of red wines by intra- and intermolecular reactions. They contribute to wine flavor (volatile phenolic acids) and some of them (e.g. resveratrol) exhibit potent biological activities.
9. Proanthocyanidins - the principal polyphenols in red wine that are under research to assess risk of [coronary heart disease](https://en.wikipedia.org/wiki/Coronary_heart_disease) and lower overall mortality. With [tannins](https://en.wikipedia.org/wiki/Tannins), they also influence the aroma, flavor, [mouth-feel](https://en.wikipedia.org/wiki/Mouth-feel) and [astringency](https://en.wikipedia.org/wiki/Astringency) of red wines. In red wines, total OPC content, including [flavan-3-ols](https://en.wikipedia.org/wiki/Flavan-3-ol) (catechins), was substantially higher (177  mg/L) than that in white wines (9  mg/L).
10. Color intensity
11. Hue
12. OD280/OD315 of diluted wines - ? to examine further. OD280 seems to have something to do with proteins and TPI total phenol index.
13. Proline - is typically the most abundant amino acid present in grape juice and wine. The amount present is influenced by viticultural and winemaking factors and can be of diagnostic importance. Used in the [biosynthesis](https://en.wikipedia.org/wiki/Protein_biosynthesis) of proteins. In brewing, proteins rich in proline combine with polyphenols to produce haze (turbidity).

**Analysis 1**

Based on the information above, after discerning which attributes could be considered beneficial to health, these attributes were compared among the 3 classes of wines.

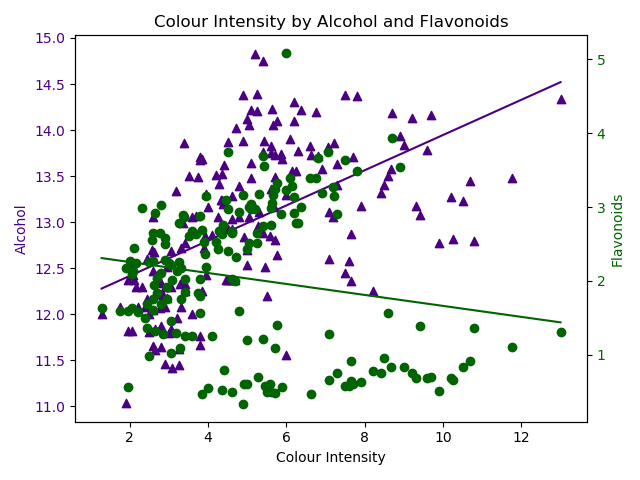


For each class, the mean values of the health attributes were presented. Magnesium ad Proline values were adjusted to fit within the same values on the Y axis. So while each attribute may have a different importance in health, and the proportions of the attributes compared to each other (eg Malic Acid to Magnesium) might not reflect actual amounts, we can still glean some information. We can compare one attribute over all groups and see that, for example, class 1 has the highest level of Total Phenols and class 3 has the highest level of Malic Acid. There wasn’t one class which was higher than the others for all health indicators. In proportion to each other, it’s quite surprising that, for example, there would be half as much Proline in class 2 than class 1. Magnesium was relatively consistent among the groups.

Further exploration: do these attributes play different size roles in terms of health benefits. How can they be compared to each other? Can we create a total health benefit index?

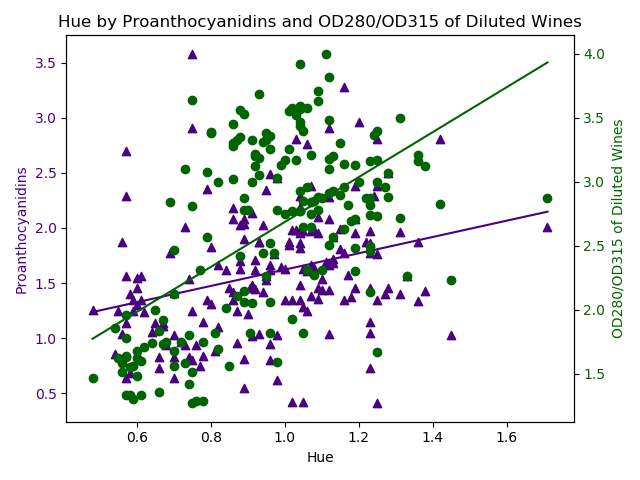
**Analysis 2**

Determine which attributes which may be correlated to colour intensity and hue.



During the initial exploratory analysis, there seemed to be a relationship between Colour Intensity and both Alcohol and Flavonoids. They are presented here, with Alcohol in purple (scale on left side), and Flavonoids in green (scale on right side). The relationship appears to be stronger between colour intensity and alcohol (purple triangles), than between colour intensity and flavonoids (green dots). The purple triangles are clustered in a fairly clear pattern indicating an increase in alcohol corresponding to an increase in color intensity. The green dots are not as concentrated and do not produce a clear pattern.

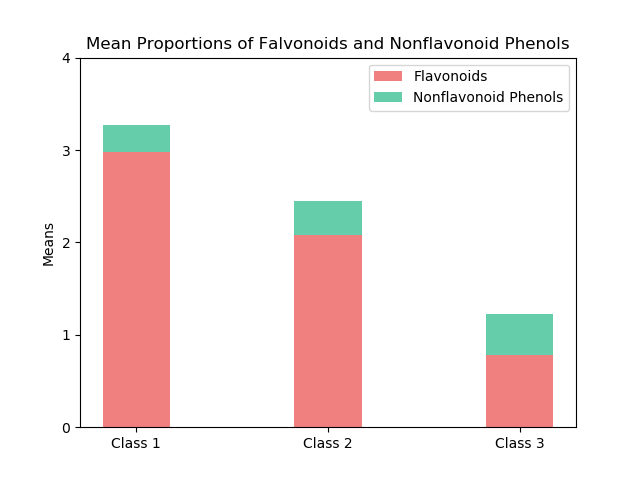
Further exploration: is there an actual relationship between colour intensity and alcohol? Why?



Plots for Proanthocyanidins and OD280/OD315 of Diluted Wines show that there may be a relationship between these two attributes and hue. Research indicates that Proanthocyanidins are substantially higher in red wines so this plot supports that finding. This is perhaps why we have always heard that red wine in particular has health benefits. I was unable to find any information on what the attribute OD280/OD315 of Diluted Wines is.

**Analysis 3**

There appear to be several types of phenols. Two types were presented here, stacked, since they each represent a fraction of the total.



Here we can see the mean values of flavonoids and non-flavonoid phenols (NFP) for each class. For class 1, the ratio of NFP to flavonoids is smaller than class 2, which is smaller than class 3. However, the overall quantity of the two combined is greatest in class 1, less in class 2, and again less in class 3. This also corresponds to the Total Phenol levels in the figure for Analysis 1.

Further exploration: would it be possible to include all or most types of phenols to see the proportions?