# Consumer Wine Segmentations

Consumers can be grouped into three distinct segments when it comes to their wine preferences. Given that the average consumer is likely not aware of the different types of elements in wine, it is unlikely that they choose their wine by any of the metrics descried below. It is more likely that the consumer segments are divided based on the type of wine that these variables are describing. For example, judging by the distribution of the phenol content amongst the groups, it is very likely that group 1 is describing consumers who prefer red wine, while group 2 prefer roses and group 3 prefer white wine.

Description of the metrics:

```
* Alcohol - The percentage of alcohol in the wine.
* Ash - The amount (in grams per 100 ml) of the wine that is not decomposed to a mineral after burning (i.e. amount of inorganic matter).
* Proline - The amount (in miligrams per 1L) of this particular amino acid, most prevalent in wine.
* Magnesium - The amount (in mg per 1L) of this chemical element.
* Phenols - Bitterness from the skin of the fruit. Used to deter animals from eating them. Mostly present in dry wine.
* Flavanoids - Subset of Phenols. Wine's source of antioxidants.
* Nonflavanoid Phenols - Subset of Phenols, excluding flavanoids.
* Alcanity Ash - A measure of the basicity of the ash. (Alcanity is likely a typo for Alkalinity.)
* Color Intensity - The amount of color that is able to pass through the wine.
* Malic acidy - Give the milky aspects of wine. Produced by lactic acid bacteria and also found in yogurt and sauerkraut.
* Hue - The brightness of the color.
* Proanthocyanins - Condensed tanins that determine the bitterness of wine. (Proanthocyanins is likely a typo for Proanthocyanidins.)
* OD280 - Spectrophotometric analysis to determine the amount of total phenols in a wine based on the color.
```
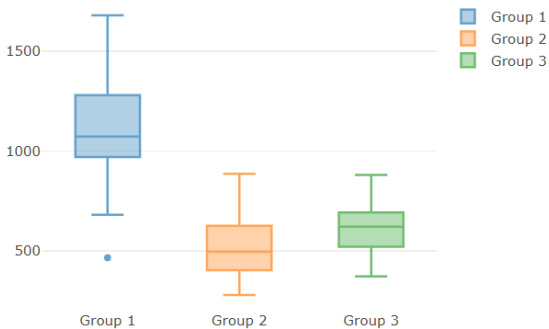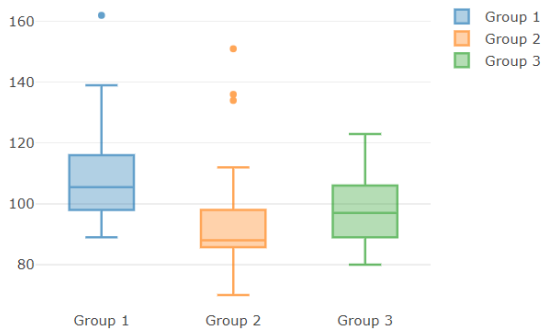

Preference of Alcohol Percentage
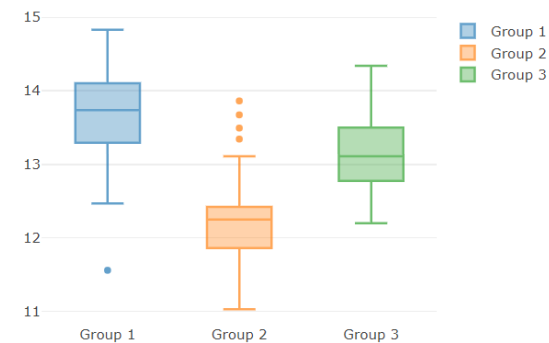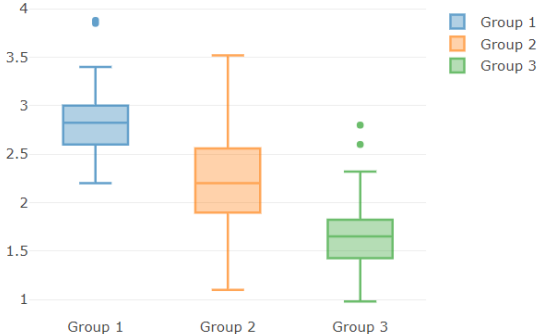

Preference of Ash (in g per 100 ml)
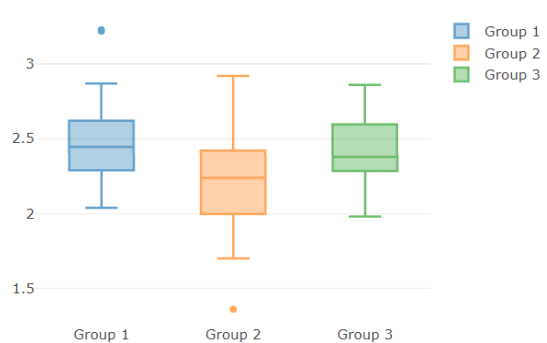
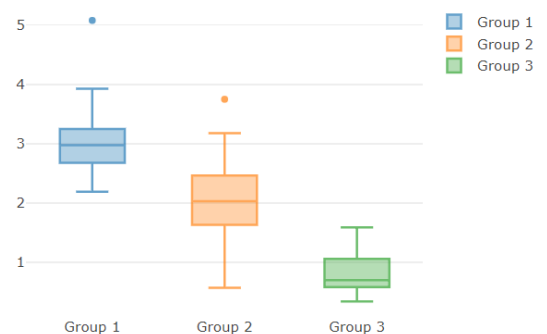Preference of Proline (in mg per L)

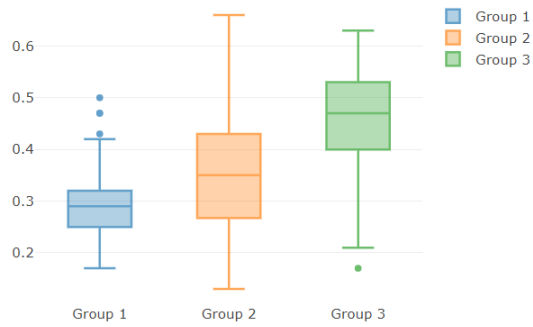Preference of Magnesium (in mg per L)
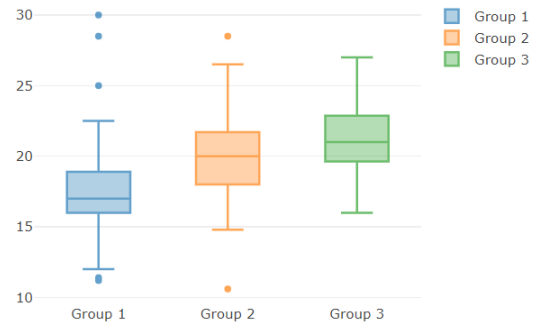

Preference of Phenols (in g per L)

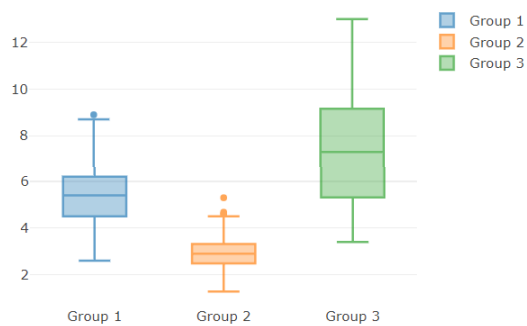Preference of Flavanoids (in g per L)

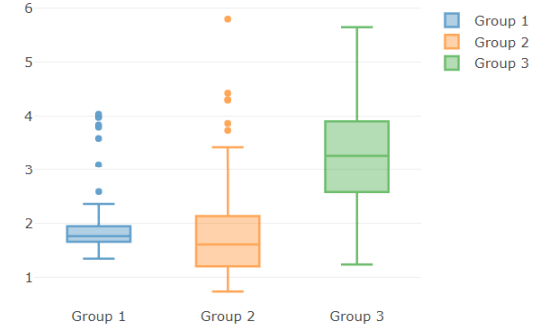**Preference of Non-flavanoid Phenols (in g per L)**

Preference of Ash Alcanity (in g per L)

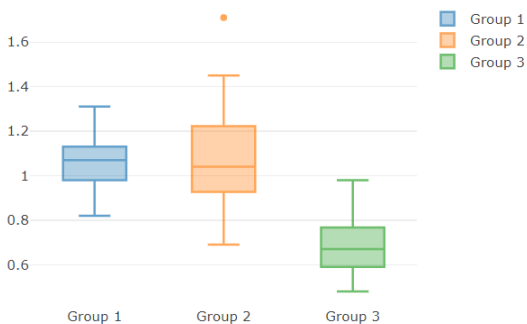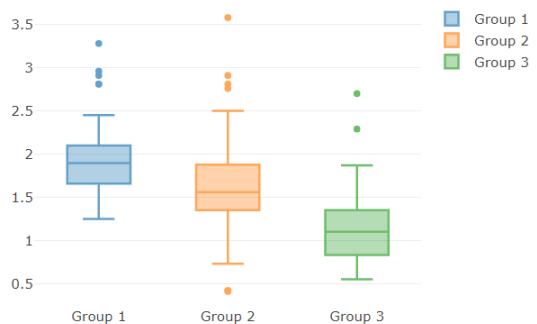**Preference of Color Intensity**

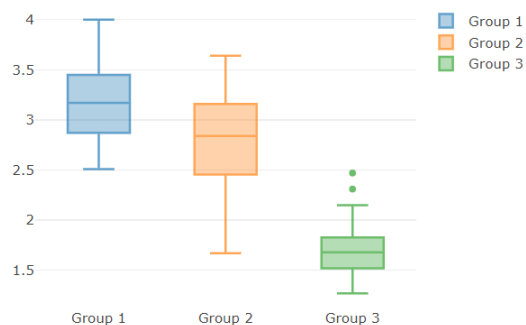**Preference of Malic Acid (in g per L)**

Preference of Hue

Preference of Proanthocyanins (in mg per L)

Preference of OD280

As you can see from the boxplot pairs that I have presented, there are quite a few metrics that seem to correlate. It would be interesting to explore further to see whether these correlations can be tied to the three factors that emerged from the factor analysis of my research.

Note: The metrics in the study were not explained in the Kaggle (https://www.kaggle.com/xvivancos/clustering-wines-with-k-means/data) documentation. All of the metrics descriptions above have instead been estimated from online research and may not be 100% accurate. Also, the markdown aspect of Dash by Plot.ly is current not working. As a result, links and bulleted lists can unfortunately not be made at this time.