Week 3 Quiz

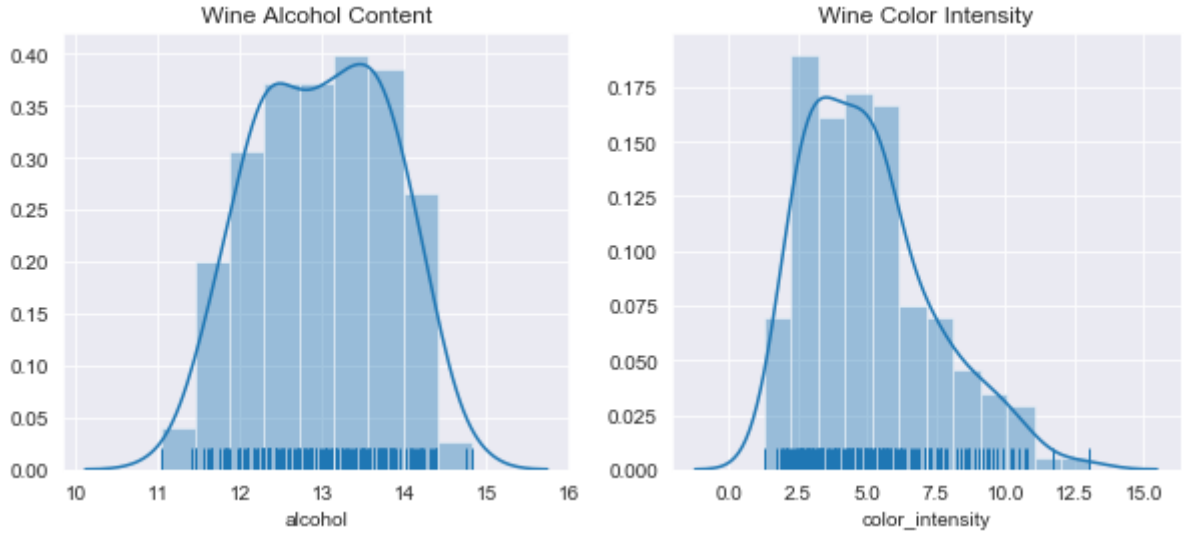
Perrin Anto - paj2117

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

sns.set_style('darkgrid')

%matplotlib inline
```

In [2]:
```
# Use pandas to read in 'wine_dataset.csv'
# This is a dataset of various wines with a target of categorical variab
le 'class'
df = pd.read_csv('../data/wine_dataset.csv')
```

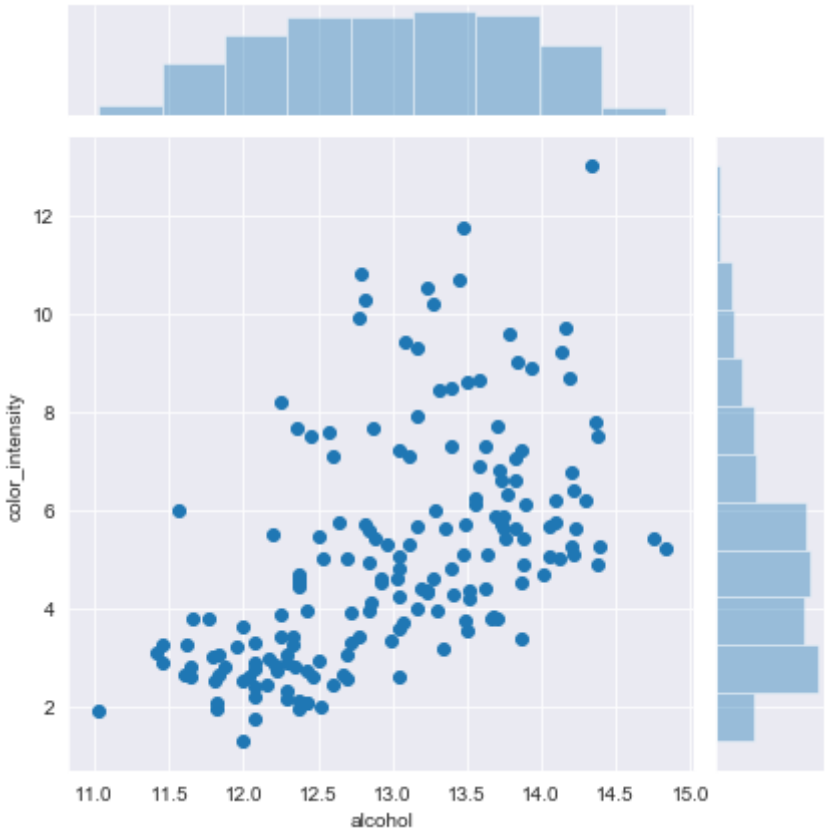
In [3]:
```
# 1. Create two axes using plt.subplots with 1 row , 2 columns, figsize=(10,4)
fig, ax = plt.subplots(1, 2, figsize=(10,4))

# 2.1 In the first axis, plot the distribution of df.alcohol using sns.distplot with rug=True
sns.distplot(df.alcohol, rug=True, ax=ax[0])

# 2.2 Add the title 'Wine Alcohol Content' to ax[0] using set_title
ax[0].set_title('Wine Alcohol Content')

# 3.1 In the second axis, plot the distribution of df.color_intensity using sns.distplot with rug=True
sns.distplot(df.color_intensity, rug=True, ax=ax[1])

# 3.2 Add the title 'Wine Color Intensity' to ax[1] using set_title
ax[1].set_title('Wine Color Intensity')
```

Out[3]: Text(0.5, 1.0, 'Wine Color Intensity')



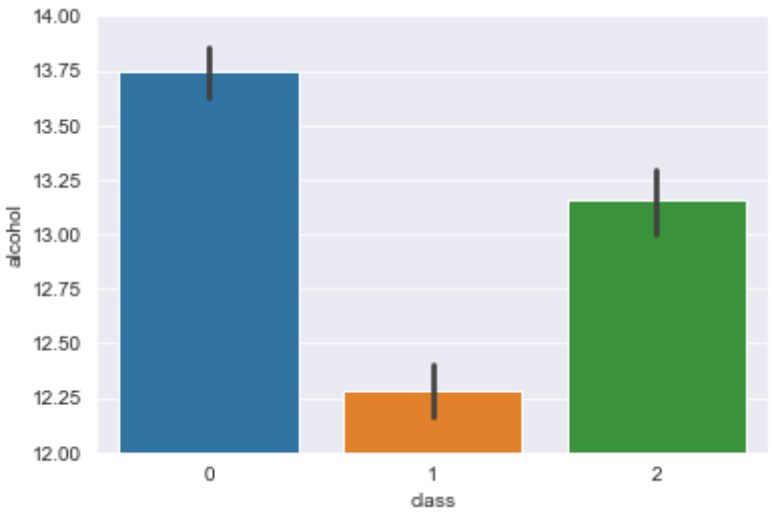
In [4]:
```
# Visualize correlation between alcohol and color_intensity using jointplot
sns.jointplot(x='alcohol',y='color_intensity',data=df)
```

Out[4]: <seaborn.axisgrid.JointGrid at 0x1a1a67f790>



In [5]:
```
# plot x=class vs y=alchohol using sns.barplot
ax = sns.barplot(x="class", y="alcohol", data=df)

# set the y-axis limits to (12,14) using set_ylim
ax.set_ylim(12,14)
```

Out[5]: (12, 14)



What's being plotted in the catplot?

The mean alcohol level for each class as the overall colored bars with additional black error bars to show the uncertainty around those means. The error bars are computed based on a normal distribution 95% confidence interval when the CI is not defined.

Does it look like there is a difference in alcohol level between class? Why?

Yes, the alcohol level normal distribitution range does not intersect for each class, showing there is a difference in alcohol level between classes with high confidence.

In [6]:
```
# when completed,
#   make sure you've replaced [Name] and [UNI] in the first cell and fil
ename
#   use Print Preview, Print-> Save to pdf
#   and post pdf to GradeScope
```