Welcome to Module 3's Labs!

In order to complete the labs in this module, please make sure you download and unarchive this .zip file with all the datasets and files necessary.

Lab Assignment 1

For this assignment, you'll be using the seeds data set, generated by recording X-Ray measurements of various wheat kernels. Start by opening up the starter code located in Module3/**assignment1.py**, and reading through it. Then, write code that...

1. Loads the seeds dataset, located at Module3/Datasets/**wheat.data** into a dataframe
2. Creates a slice of your dataframe that only includes the **area** and **perimeter** features
3. Creates another slice that only includes the **groove** and **asymmetry** features
4. Creates a histogram for the 'area and perimeter' slice, and another histogram for the 'groove and asymmetry' slice. Set the optional display parameter: **alpha**=0.75

Once you're done, run your code and then answer the following questions about your work:

Lab Questions

(2/2 points)

Looking at your first plot, the histograms of area and perimeter, which feature do you believe more closely resembles a Gaussian / normal distribution?

Perimeter ▾ ✔ **Answer:** Perimeter

In your second plot, does the groove or asymmetry feature have more variance?

Asymmetry ▾ ✔ **Answer:** Asymmetry

EXPLANATION

Start by doing some basic array slicing of your data using the **your_slice = df[['column1', 'column2']]** syntax.

Then, plot by using the code provided in the lesson, **your_slice.plot.hist(alpha=0.75)**

All features (columns) selected in the slice will be plotted.

You have used 1 of 2 submissions