



Bookmarks

- ▶ Start Here
- ▶ 1. The Big Picture
- ▶ 2. Data And Features

▼ 3. Exploring Data

Lecture: Visualizations

Lecture: Basic Plots

Quiz



Lecture: Higher Dimensionality

Quiz

**Lab: Visualizations**

Lab



Dive Deeper

- ▶ 4. Transforming Data
- ▶ 5. Data Modeling

3. Exploring Data > Lab: Visualizations > Assignment 1

Assignment 1

Bookmark this page

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 points possible (graded)

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

▶ 6. Data Modeling II

▶ 7. Evaluating Data

▶ Course Wrap-up

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

Select an option ▼

Submit

You have used 0 of 2 attempts

© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY
OPENedX®

