Lab Assignment 4

For this assignment, you will continue using the wheat seeds data set. Start by opening up the starter code located in Module3/**assignment4.py**, and write code that...

1. Loads up the seeds dataset, located at Module3/Datasets/**wheat.data** into a dataframe
2. Drop the **id**, **area**, and **perimeter** features from your dataset
3. Plot a parallel coordinates chart, grouped by the **wheat_type** feature. Be sure to set the optional display parameter **alpha** to 0.4

Once you're done, answer the following questions about your work:

Lab Questions

(2/2 points)
Which class of wheat do the two outliers you found previously belong to?

Canadian Wheat ✔ **Answer:** Canadian Wheat

Which feature has the largest spread of values across all three types of wheat?

Asymmetry ✔ **Answer:** Asymmetry

EXPLANATION

When you drop columns from a dataframe, be sure to specify **axis=1**.

Recall that parallel coordinates group your polylines by a feature. The feature you were supposed to use was the **wheat_type** feature. The legend of your parallel coordinates plot should give you the colors that correspond to the wheat types, so just look for the two samples that don't "play well" with the curve.

You have used 1 of 2 submissions


Download on the
App Store
ANDROID APP ON
Google play