Quiz

1 ✓ What is the main purpose of standardizing data?

Your Answer

✓ **A)** It lets us view data in terms of standard deviations from the average case (i.e. mean)

○ **B)** It lets us use much smaller data values, which makes computation much quicker

○ **C)** It makes it easier to calculate the mean of each column in the data

○ **D)** It removes outliers from the dataset

2 ✓ What is data imputation?

● **A)** Combining multiple datasets into one large dataset

○ **B)** Removing certain features from a dataset

C) Subtracting the column-wise means from the values in a dataset

✓ **D)** Replacing missing data in the dataset with substituted values

---

3 ✓ Why do we use PCA?

○ **A)** To increase the number of features in a dataset

○ **B)** To group together correlated features

Your Answer

✓ **C)** To remove correlation from dataset features and perform dimensionality reduction

○ **D)** To remove repeated observations from the dataset

---

4 ✓ What is a good practice for plotting data separated by class label?

○ **A)** Plot a percentage of points from each class based on the class distribution

Your Answer

B) Use different colors and include a legend to distinguish data by class label

C) Put each class' points on separate plots, never on the same plot
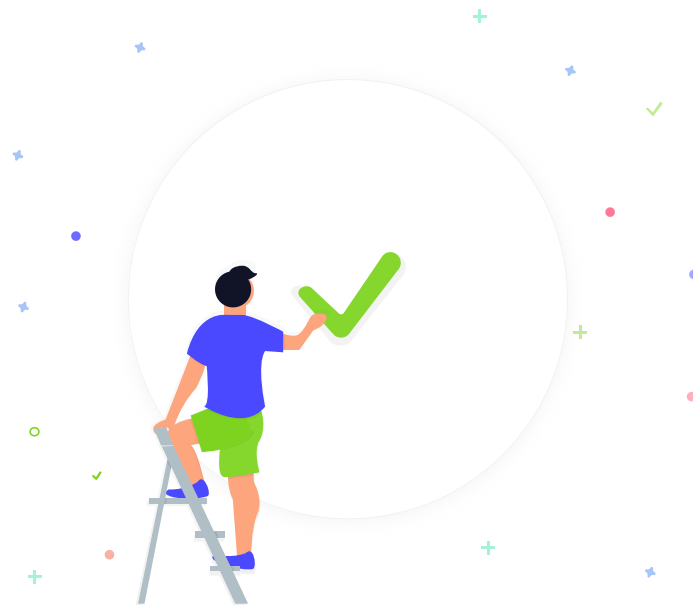
D) None of the above

SUMMARY

Correct	4
Incorrect	0

Awesome! you got all 4 correct!

Retake Quiz

✅ **Mark as Completed**

39% completed, meet the criteria and claim your course certificate!

Buy Certificate

⚠️ Report an Issue

❓ Ask a Question (https://discuss.educative.io/tag/quiz__data-preprocessing-with-scikit-learn__machine-learning-for-software-engineers)