

← Back

Graded Quiz: Module 3 - Exploratory Data Analysis and Feature Engineering

Graded Assignment • 10 min

Due Oct 3, 11:59 PM PDT

Your grade: 100%

Your latest: 100% • Your highest: 100%

To pass you need at least 70%. We keep your highest score.

Next item →

1. Which scaling approach converts features to standard normal variables?

☐ Nearest neighbor scaling

☒ Standard scaling

☐ Robust scaling

☐ MinMax scaling

☒ Correct

Correct. Standard scaling converts variables to standard normal variables.

2. Which variable transformation should you use for ordinal data?

☐ One-hot encoding

☐ Min-max scaling

☐ Standard scaling

☒ Ordinal encoding

☒ Correct

Correct. Use ordinal encoding if there is some order to the categorical features.

3. What are polynomial features?

☐ They are represented by linear relationships in the data.

☐ They are lower order relationships in the data.

☐ They are logistic regression coefficients.

☒ They are higher order relationships in the data.

☒ Correct

Correct. Polynomial features are estimated by higher order polynomials in a linear model, like squared, cubed, etc.

4. What does boxcox transformation do?

☐ It transforms categorical variables into numerical variables.

☐ It makes the data more left skewed

☐ It makes the data more right skewed.

☒ It transforms the data distribution into more symmetrical bell curve

☒ Correct

Correct. Boxcox is one of the ways we can transform our skewed dataset to be more normally distributed.

5. Select three important reasons why EDA is useful.

☐ To examine correlations, to sample from dataframes, and to train models on random samples of data

☐ To analyze data sets, to determine the main characteristics of data sets, and to use sampling to examine data

☒ To determine if the data makes sense, to determine whether further data cleaning is needed,, and to help identify patterns and trends in the data

☐ To utilize summary statistics, to create visualizations, and to identify outliers

☒ Correct

Correct. EDA helps us analyze data to summarize its main characteristics.

6. What assumption does the linear regression model make about data?

☐ This model assumes a transformation of each parameter to a linear relationship.

☐ This model assumes an addition of each one of the model parameters multiplied by a coefficient.

☐ This model assumes that raw data in data sets is on the same scale.

☒ This model assumes a linear relationship between predictor variables and outcome variables.

☒ Correct

Correct. The linear regression model assumes a linear relationship between predictor and outcome variables.

7. What is skewed data?

☐ Raw data that may not have a linear relationship.

☒ Data that is distorted away from normal distribution; may be positively or negatively skewed.

☐ Raw data that has undergone log transformation.

☐ Data that has a normal distribution.

☒ Correct

Correct. Often raw data, both the features and the outcome variable, can be negatively or positively skewed.

8. Select the two primary types of categorical feature encoding.

☐ Frequency encoding and label encoding

☐ Log and polynomial transformation

☒ One-hot encoding and ordinal encoding

☐ Encoding and scaling

☒ Correct

Correct. Encoding that transforms non-numeric values to numeric values is often applied to categorical features.

9. Which scaling approach puts values between zero and one?

☐ Robust scaling

☒ Min-max scaling

☐ Standard scaling

☐ Nearest neighbor scaling

☒ Correct

Correct. Min-max scaling converts variables to continuous variables in the (0, 1) interval by mapping minimum values to 0 and maximum values to 1.

10. Which variable transformation should you use for nominal data with multiple different values within the feature?

☐ Min-max scaling

☐ Ordinal encoding

☒ One-hot encoding

☐ Standard scaling

☒ Correct

Correct. Use one-hot encoding if there are multiple different values within a feature.

https://www.coursera.org/learn/ibm-exploratory-data-analysis-for-machine-learning/assignment-submission/UNr7l/graded-quiz-module-3-exploratory-d...1/1