Back

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

Graded Quiz. • 30 min

English

Due Nov 6, 1:29 PM IST

Congratulations! You passed!

Grade received 100%

Latest Submission Grade 100%

To pass 70% or higher

Go to next item

1.

Question 1

Which scaling approach converts features to standard normal variables?

1 / 1 point

Standard scaling

Nearest neighbor scaling

Robust scaling

MinMax scaling

Correct

Correct. Standard scaling converts variables to standard normal variables.

2.

Question 2

Which variable transformation should you use for ordinal data?

1 / 1 point

Min-max scaling

One-hot encoding

Ordinal encoding

Standard scaling

Correct

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

3.

Question 3

What are polynomial features?

1 / 1 point

They are logistic regression coefficients.

They are higher order relationships in the data.

They are represented by linear relationships in the data.

They are lower order relationships in the data.

Correct

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

4.

Question 4

What does Boxcox transformation do?

1 / 1 point

It makes the data more right skewed.

It transforms categorical variables into numerical variables.

It transforms the data distribution into more symmetrical bell curve

It makes the data more left skewed

Correct

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

5.

Question 5

Select three important reasons why EDA is useful.

1 / 1 point

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

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

Correct

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

6.

Question 6

What assumption does the linear regression model make about data?

1 / 1 point

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

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

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.

Correct

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

7.

Question 7

What is skewed data?

1 / 1 point

Raw data that has undergone log transformation.

Data that has a normal distribution.

Raw data that may not have a linear relationship.

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

Correct

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

8.

Question 8

Select the two primary types of categorical feature encoding.

1 / 1 point

Encoding and scaling

Log and polynomial transformation

One-hot encoding and ordinal encoding

Nominal encoding and ordinal encoding

Correct

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

9.

Question 9

Which scaling approach puts values between zero and one?

1 / 1 point

Nearest neighbor scaling

Standard scaling

Min-max scaling

Robust 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.

Question 10

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

1 / 1 point

Ordinal encoding

Min-max scaling

Standard scaling

One-hot encoding

Correct

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