

**Microsoft: DAT203x Data Science and Machine Learning Essentials**

- ▶ Before You Start
- ▶ Module 1: Introduction and Data Science Theory
- ▶ Module 2: Working with Data
- ▶ Module 3: Visualization, and Building and Evaluating Models
- ▼ **Module 4: Regression, Classification, and Unsupervised Learning**
- Chapter 16: Regression Modeling
- Lab 4A: Working with Regression Models
- Chapter 17: Classification Modeling
- Lab 4B: Working with Classification Models
- Chapter 18: Unsupervised Learning Models
- Lab 4C: Working with Unsupervised Learning Models

**QUESTION 1** (1 point possible)

While evaluating the performance of a regression model you discover that the residuals are randomly distributed and exhibit no significant structure with respect to the values of the label or the features.

This indicates which two of the following conditions are true?

☐ The model is a good fit to the data. ✓

☐ The data are exhibiting nonlinear behavior not accounted for by the model.

☒ The information in the features is being exploited for the most part. ✓

☒ The residuals do not vary with the value of the label as one would expect.




Note: Make sure you select all of the correct options—there may be more than one!

**EXPLANATION**

The residuals of a good regression model should exhibit a random distribution with no particular structure with respect to the values of the label or the features. The randomness and lack of structure indicates that model fits the data well and that the information in the features has been exploited by the model.

*You have used 2 of 2 submissions*

**Module 4 Review**

Homework due Oct 30,  
2015 at 00:00 UTC 

- ▶ Module 5:  
Recommenders  
and Publishing  
Your Work
- ▶ Final Exam

© All Rights Reserved



© 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**

