



Bookmarks

- ▶ [Module 1: The Basics of R and Introduction to the Course](#)
- ▶ [Entrance Survey](#)
- ▶ [Module 2: Fundamentals of Probability, Random Variables, Distributions, and Joint Distributions](#)
- ▶ [Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates](#)
- ▶ [Module 4: Joint, Marginal, and Conditional Distributions & Functions of Random Variable](#)

Module 11: Intro to Machine Learning and Data Visualization > Machine Learning I > Applying the Statistical Approach to Sentiment Analysis - Quiz

Applying the Statistical Approach to Sentiment Analysis - Quiz

🔖 Bookmark this page

Question 1

0 points possible (ungraded)

Consider a movie review with 250 words, all of which are unique. How many total unigrams and bigrams could be produced?

Note: Bigrams must appear consecutively and order does not matter (i.e. "the movie" is the same as "movie the")

✓ Answer: 499

499

Explanation

For a set of 250 words, one can produce a unigram out of each word, and for a set of 250 words you can produce 249 set of consecutive pairs, giving a total of 499 total unigrams and bigrams.

Submit

You have used 1 of 2 attempts

- ▶ [Module 5: Moments of a Random Variable, Applications to Auctions, & Intro to Regression](#)
- ▶ [Module 6: Special Distributions, the Sample Mean, the Central Limit Theorem, and Estimation](#)
- ▶ [Module 7: Assessing and Deriving Estimators - Confidence Intervals, and Hypothesis Testing](#)
- ▶ [Module 8: Causality, Analyzing Randomized Experiments, & Nonparametric Regression](#)
- ▶ [Module 9: Single and Multivariate Linear Models](#)
- ▶ [Module 10: Practical Issues in Running](#)

Question 2

1/1 point (graded)

Think back to the example provided in this lecture segment. Suppose you had 2,001 distinct words and 2,000 movie reviews. Should you try to run a regression in this scenario in which the distinct words are your x-variables?

☐ a. Yes

☒ b. No ✓

Explanation

No, you should never run a linear regression in which the number of independent variables (distinct words) in the regression is greater than the number of observations (movie reviews) you have.

Submit

You have used 1 of 1 attempt

✓ Correct (1/1 point)

Discussion

Topic: Module 11 / Applying the Statistical Approach to Sentiment Analysis
- Quiz

Show Discussion

Regressions, and
Omitted Variable Bias

▼ Module 11: Intro to
Machine Learning and
Data Visualization

Machine Learning I

due Dec 12, 2016 05:00 IST



Machine Learning II

due Dec 12, 2016 05:00 IST



Visualizing Data

due Dec 12, 2016 05:00 IST



► Module 12:
Endogeneity,
Instrumental Variables,
and Experimental
Design

© All Rights Reserved



© 2016 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.

