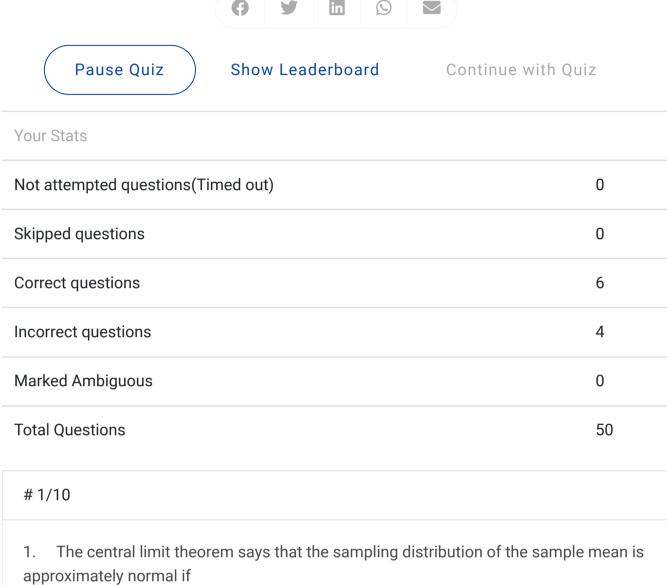
8/6/2019 Datamin | Review

Review your last 10 questions

5 days, 5 hours remaining until quiz ends.



- - All possible samples are selected
 - The standard error of the sampling distribution is small
 - The sample size is large
 - All of the above

Incorrect

```
# 2/10
```

```
What does the output of the following code signify?
import spacy
nlp = spacy.load('en_core_web_md')
sent1 = 'Data science is love'
sent2 = 'Love data and science'
doc1 = nlp(sent1)
doc2 = nlp(sent2)
print(doc1.similarity(doc2))
```

- The code outputs a similarity index between -1 and 1, with 1 being most similar and -1 least similar
- The code outputs a similarity index between 0 and 1, with 1 being most similar and 0 being least similar
- The code outputs a similarity index between 0 and 1, with 0 being most similar and Online Chat ^

1 being least similar

8/6/2019

• The code outputs a similarity index between -1 and 1, with -1 being most similar and 1 least similar Incorrect #3/10 Two data scientists want to make a non-linear regression model from a similar dataset. The first one tries to find the P-values of the variables to check the significance of the variables in the model. While the second used the confidence intervals to select the parameters. Which model should be trusted? First Second both · Cannot be determined Incorrect # 4/10 Which of the following options will group the dataframe df by multiple columns(column1 and column2)? df.groupby(`column1`,`column2`) df.groupby([`column1`, `column2`]) df.groupby((`column1`, `column2`)) None of the above Correct # 5/10 In which of the following applications can we use deep learning to solve the problem? Protein structure prediction · Prediction of chemical reactions • Detection of exotic particles All of these Correct

6/10

Select the correct output: l=['abcd','zx','def'] I.sort() print(l) ['abcd','zx','def'] ['abcd','def','zx'] ['zx','def','abcd'] error(not a numpy array,sort() function available for np array only) Correct # 7/10 What is the run time complexity of a decision tree of a depth 'k' used on a dataset with 'n' datapoints and 'd' dimensions? • BigOh(ndk) BigOh(k) • BigOh(d) • BigOh(nd) • None of the above Incorrect #8/10 If the cosine distance between two vectors, corresponding to two datapoints, is 2 then what is the L2 norm value for a vector defined by the difference between the two given vectors? 2 • 1.41 • 1 • 4 Correct # 9/10

Which of the following is used as a benchmark to compare other models for a text classification probem?

3/4

KNN Decision trees SVM None of the above Correct # 10/10 _ is the analysis of variance and _____ is the analysis of covariance. ANCOVA, ANOVA ANOVA, ANCOVA MANOVA, ANOVA ANCOVA, MANOVA Correct i Suggested reading 9 Challenges on Data Merging & Subsetting in R & Python (beginners) Pause Quiz **Show Leaderboard** Continue with Quiz **Analytics Vidhya Data Scientists** Companies About Us <u>Blog</u> Post Jobs <u>Hackathon</u> Our Team <u>Trainings</u> <u>Careers</u> **Discussions** Hiring Hackathons Contact us Apply Jobs <u>Advertising</u> Visit us **Download App** GET IT ON Google Play Sign up to our newsletter your email address

https://datamin.analyticsvidhya.com/quiz/review