DS-200^Q&As

Data Science Essentials

Pass Cloudera DS-200 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.geekcert.com/ds-200.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Cloudera Official Exam Center

Instant Download After Purchase

100% Money Back Guarantee

365 Days Free Update

800,000+ Satisfied Customers

QUESTION 1

When optimizing a function using stochastic gradient descent, how frequently should you update your estimate of the gradient?

A. Once after every pass through the data set

B. Once per observation

C. For each observation with a probability that you choose ahead of time

D. After a random number of observations

E. Once every N observations, where you decide N ahead of time

Correct Answer: AC

QUESTION 2

You have a large m x n data matrix M. You decide you want to perform dimension reduction/clustering on your data and have decide to use the singular value decomposition (SVD; also called principal components analysis PCA)

You performed singular value decomposition (SVD; also called principal components analysis or PCA) on you data matrix but you did not center your data first. What does your first singular component describe?

A. The mean of the data set

B. The variance of the data set

C. The standard deviation of the data set

D. The maximum of the data set

E. The median of the data set

Correct Answer: C

QUESTION 3

You have a large m x n data matrix M. You decide you want to perform dimension reduction/clustering on your data and have decide to use the singular value decomposition (SVD; also called principal components analysis PCA)

For the moment, assume that your data matrix M is 500 x 2. The figure below shows a plot of the data.
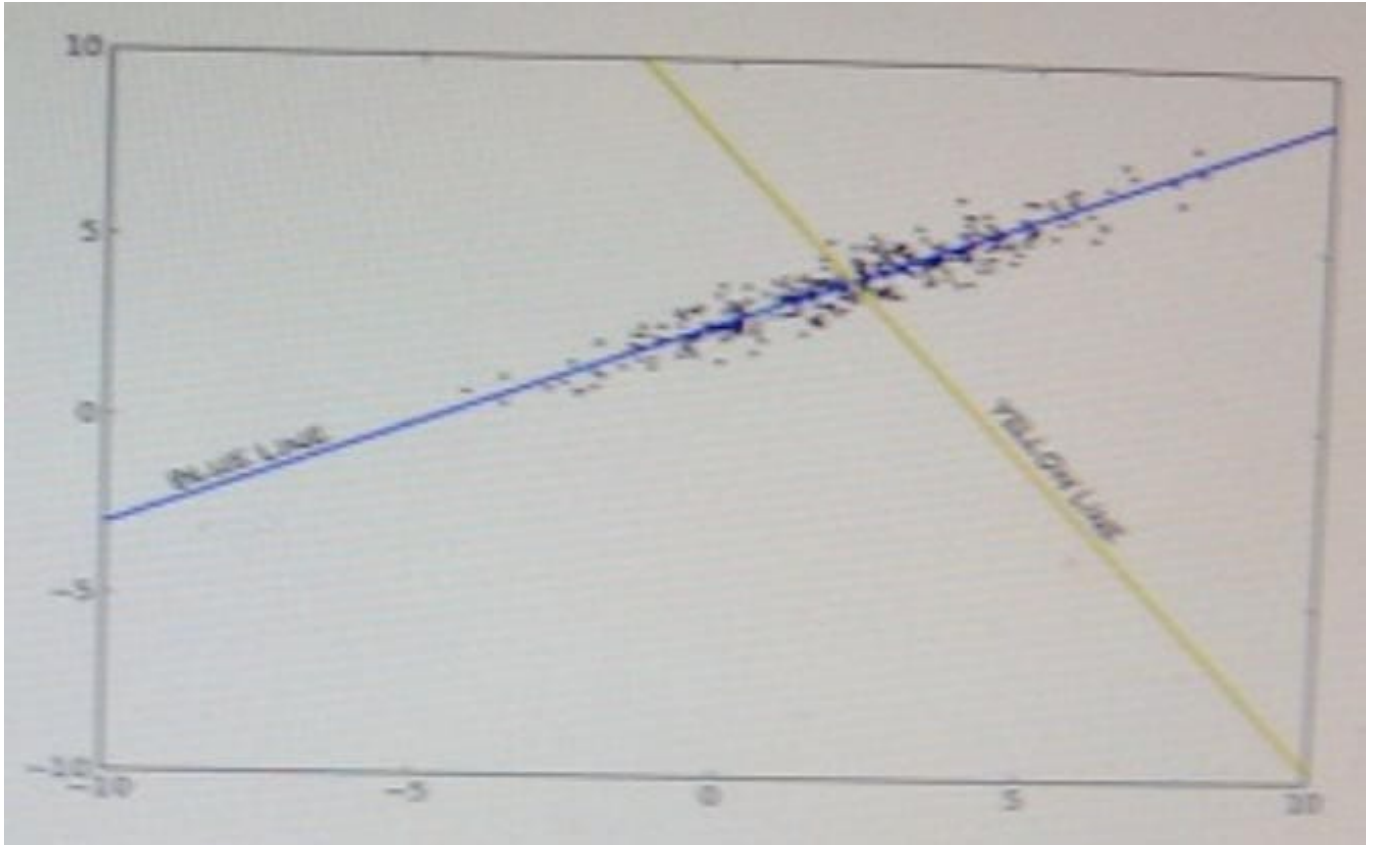
Which line represents the second principal component?

A. Blue

B. Yellow

Correct Answer: A

QUESTION 4

You have a large file of N records (one per line), and want to randomly sample 10% them. You have two

functions that are perfect random number generators (through they are a bit slow):

Random_uniform () generates a uniformly distributed number in the interval [0, 1] random_permotation (M)

generates a random permutation of the number O through M -1.

Below are three different functions that implement the sampling.

Method A

For line in file: If random_uniform ()

Method B

i = 0

for line in file:

if i % 10 = = 0;

print line

i += 1

Method C

idxs = random_permotation (N) [: (N/10)]

i = 0

for line in file:

if i in idxs:

print line

i +=1

Which method is least likely to give you exactly 10% of your data?
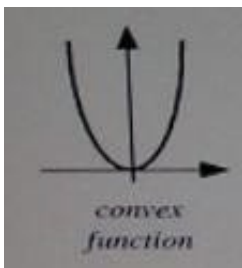
A. Method A

B. Method B

C. Method C

Correct Answer: B

QUESTION 5

Function is convex if the line segment between two points, a and b is greater than equal to the value of the a x b Which two functions are convex?

A. X1/2

B. Ex

C. 2x-1

D. 1-x2

Correct Answer: A

Latest DS-200 Dumps   DS-200 PDF Dumps   DS-200 Practice Test