Question 1
Training a model using labeled data and using this model to predict the labels for new data is known as _____.
   a. Clustering
   b. Density estimation
   c. Unsupervised learning
   d. Supervised learning

Question 2
What are the two most common supervised tasks? (Check all that apply)
   a. Classification
   b. Reinforcement learning
   c. Clustering
   d. Regression

Question 3
Training a model using categorically labelled data to predict labels for new data is known _____.
   a. Clustering
   b. Regression
   c. Feature extraction
   d. Classification

Question 4
Training a model using labelled data where the labels are continuous quantities to predict labels for new data is known as _____.
   a. Regression
   b. Clustering
   c. Feature extraction
   d. Classification

Question 5
Modeling the features of an unlabeled dataset to find hidden structure is known as _____.
   a. Regression
   b. Unsupervised learning
   c. Supervised learning
   d. Classification

Question 6
You're running a company, and you want to develop learning algorithms to address each of two problems.
   ● **Problem 1:** You have a large inventory of identical items. You want to predict how many of these items will sell over the next 3 months.
   ● **Problem 2:** You'd like software to examine individual customer accounts, and for each account decide if it has been hacked/compromised.

Should you treat these as *classification* or as *regression* problems?
a.   Treat problem 1 as a regression problem, problem 2 as a classification problem.
b.   Treat both as regression problems.
c.   Treat both as classification problems.
d.   Treat problem 1 as a classification problem, problem 2 as a regression problem.

Question 7
Of the following examples, which would you address using an <u>unsupervised</u> learning algorithm? (Check all that apply)
a.   Given a database of customer data, automatically discover market segments and group customers into different market segments.
b.   Given a dataset of patients diagnosed as either having diabetes or not, learn to classify new patients as having diabetes or not.
c.   Given email labeled as spam/not spam, learn a spam filter.
d.   Given a set of news articles found on the web, group them into set of articles about the same story.

Question 8
In a simple regression model, if you increase the input value by 1 then you expect the output to change by:
a.   The value of the *slope* parameter
b.   The value of the *intercept* parameter
c.   Also 1
d.   Impossible to tell

Question 9
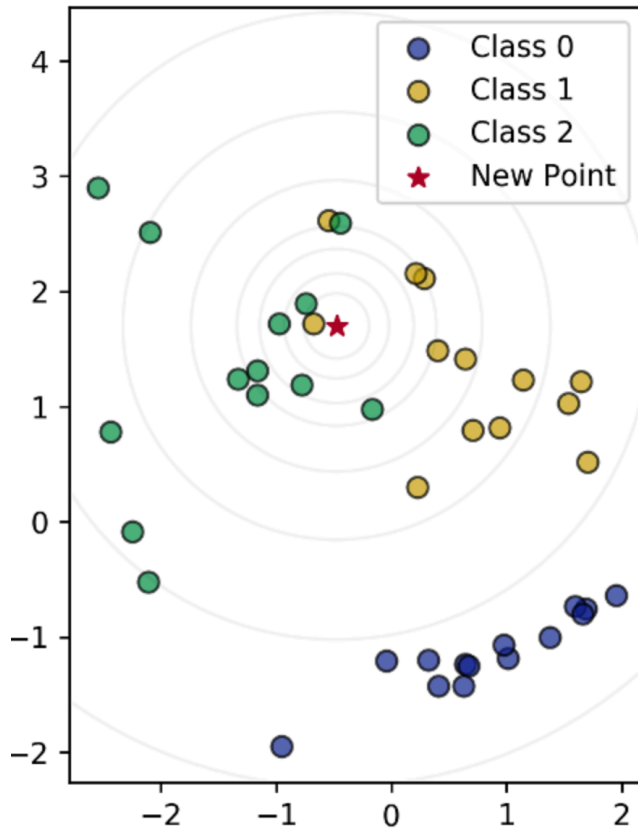Consider the following data set, and the regression line fitted on this data:
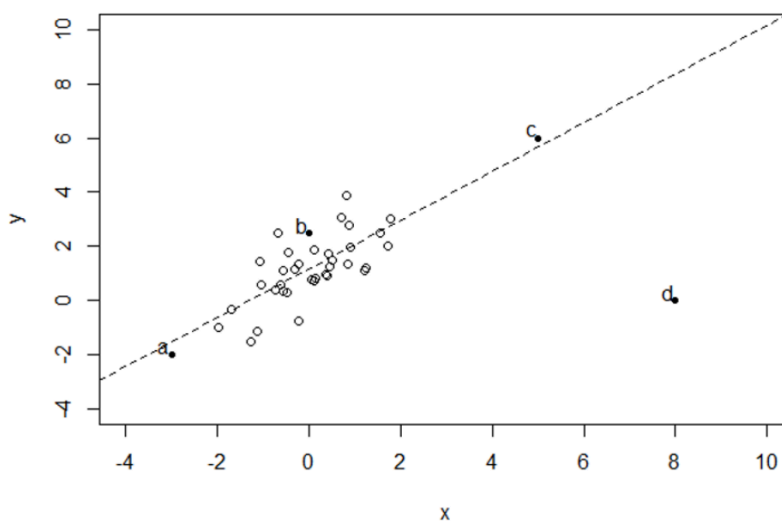
Which bold/labeled point, if removed, will have the largest effect on the fitted regression line (dashed)?

a. b
c. c
d. a

Question 10
Using the data for classes 0, 1, and 2 plotted below, what class would a KNeighborsClassifier classify the new point as for k = 1 and k = 3?

a. k = 1: Class 1; k = 3: Class 0
b. k = 1: Class 1; k = 3: Class 2
c. k = 1: Class 0; k = 3: Class 1
d. k = 1: Class 0; k = 3: Class 2
e. k = 1: Class 2; k = 3: Class 1