Question 1 (2 points)

OK, now it's time to do some predictive analytics. What is a business question that points to predictive analytics (vs. descriptive or diagnostic analytics)?

A) What were the sales of the xPhone 8 in July?

B) What caused the decline in sales of the xPhone 8 in July?

C) What should I do to increase the sales of xPhone 8?

D) What factors determine the sales of xPhone 8?

Question 2 (2 points)

My CEO comes to me and asks about predictive analytics. "If you're so good at predictive analytics, why can't you tell me what the stock market is going to do tomorrow?" What do you say?

A) It's good for estimating trends but there's not enough data for an accurate specific prediction.

B) It's good for estimating an unknown value in the past or present, but not in the future.

C) Computers aren't powerful enough to build a good model of the stock market.

D) An accurate prediction of the stock market requires prescriptive analytics, not predictive analytics.

Question 3 (2 points)

Rats. I forgot what a model is. Can you remind me?

A) It's a set of attributes collected for a purpose and stored in a database. B) It's a simplified representation of reality.

C) It's a mathematical formula such as f = ma.

D) It's a "blueprint" that describes the information infrastructure of an organization.

Question 4 (2 points)

I hate definition questions but sometimes you need to know the right term for something. I have a set of data and I want to create a model from it. What's that process called?

A) Deduction B) Induction C) Reduction D) Production

Question 5 (2 points)

My data scientist came to me and says that a new software product can
reduce the actual data analysis time by 20%. Is it worth it?

A) Yes! That reduces the overall analytics effort by over 80%.

B) Yes! This reduces the most difficult part of the analytics modeling process.

C) No. Data analysis represents at most only 20% of the analysis process. It's probably not worth it.

D) No! Data analysis can be made faster by buying faster computers.

Question 6 (2 points)

Your analytics person comes to you with a model. "This model has an entropy of 0.90!" What do you think about that?

A) It's really good. That means there's only 0.10 uncertainty in the model.

B) You can't tell. Entropy is not a good measure of model goodness.

C) Entropy is the opposite of disorder. This means you have a disorder of 0.90.

D) It's not so good. The model is only a little better than random chance.

Question 7 (2 points)

We have a pretty nifty model. It's entropy is 0.40. We collect another set of data and add it to the model, reducing entropy to 0.25. What is the information gain of the new data?

A) 0.25 B) 0.65 C) 0.15 D) Can't be determined from the information given.

Question 8 (2 points)

What's the biggest barrier to organizations adopting predictive analytics?

A) Lack of good data scientists

B) Lack of good models

C) Lack of good managers

D) Lack of good data.

Question 9 (2 points)

I've been digging around and I found a model on the shelf that was able to accurately predict customer purchasing behavior. The model is about 10 years old. Can I still use it?

A) Use it with caution. Its predictions may not be accurate any more. B) Sure. Customer behavior patterns don't change much over time.

C) No. The data that it was based on is old and is no longer valid.

D) Maybe. If the product you're predicting is the same, then the model is valid.

Question 10 (2 points)

I've built a model that can cluster my customers very well into three different groups. But there's a few customers that don't fit any groups. What should I do with them?

A) Ignore them. They are outliers.

B) Investigate them further. See why they don't fit.

C) They are anomalies that are probably due to data errors. Discard them.

D) Adjust the model until the outliers merge with one of the three groups.