Exam Questions

Different levels of difficulty

easy

Big Data is best described as:

- a. Data that requires new architectures, techniques, and analytics due to its scale, diversity, and complexity
- b. A term used only to describe social media data
- c. A collection of large-sized Excel spreadsheets
- d. Data stored in relational databases

Big Data is best described as:

- a. Data that requires new architectures, techniques, and analytics due to its scale, diversity, and complexity
- b. A term used only to describe social media data
- c. A collection of large-sized Excel spreadsheets
- d. Data stored in relational databases

medium

In a linear regression, if a coefficient estimate is not statistically significant and has the expected sign, what does the teaching material suggest regarding its inclusion in the model for prediction?

- a. It must be transformed to achieve significance.
- b. It indicates a strong need for more data.
- c. It should always be excluded to simplify the model.
- d. It is generally fine to keep it in, as it likely isn't hurting predictions.

In a linear regression, if a coefficient estimate is not statistically significant and has the expected sign, what does the teaching material suggest regarding its inclusion in the model for prediction?

- a. It must be transformed to achieve significance.
- b. It indicates a strong need for more data.
- c. It should always be excluded to simplify the model.
- d. It is generally fine to keep it in, as it likely isn't hurting predictions.

challenging

In a simple linear regression model where kid.score is predicted by mom.hs (mother graduated high school: 1=yes, 0=no), if the fitted model is kid.score = 78 + 12 * mom.hs, what does the coefficient 12 represent?

- a. The average kid.score for children whose mothers did not complete high school.
- b. The average kid.score for children whose mothers completed high school.
- c. The predicted kid.score for a child with a mother who completed high school.
- d. The average difference in kid.score between children whose mothers completed high school and those who did not.

In a simple linear regression model where kid.score is predicted by mom.hs (mother graduated high school: 1=yes, 0=no), if the fitted model is kid.score = 78 + 12 * mom.hs, what does the coefficient 12 represent?

- a. The average kid.score for children whose mothers did not complete high school.
- b. The average kid.score for children whose mothers completed high school.
- c. The predicted kid.score for a child with a mother who completed high school.
- d. The average difference in kid.score between children whose mothers completed high school and those who did not.