**Sample Questions**

*The following sample questions are not inclusive and do not necessarily represent all of the types of questions that comprise the exams. The questions are not designed to assess an individual's readiness to take a certification exam.*

**SAS Viya 3.4 Supervised Machine Learning Pipelines Exam**

**Question 1:**

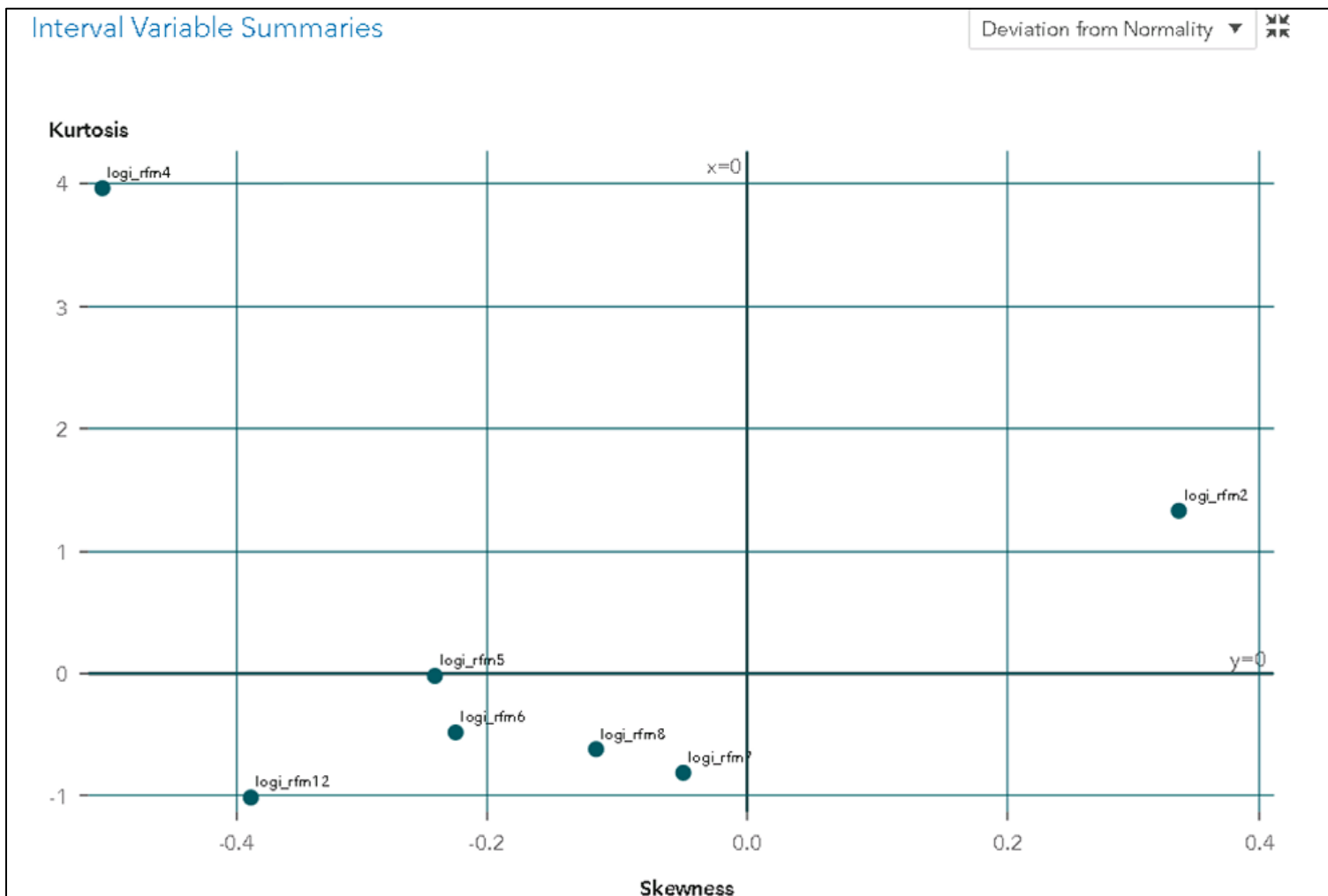A project has been created and a pipeline has been run in Model Studio.

Which project setting can you edit?

- A. **Advisor Options** for missing values
- B. **Rules** for model comparison statistic
- C. **Partition Data** percentages
- D. **Event-based Sampling** proportions

Answer: B

**Question 2:**

Refer to the exhibit below:

Based on the output from the **Data Exploration** node shown in the exhibit, which variable has the most thin tails (most platykurtic distribution)?

    A. Logi_rfm4
    B. Logi_rfm6
    C. Logi_rfm8
    D. Logi_rfm12

Answer: D


**Question 3:**

Which feature extraction method can take both interval variables and class variables as inputs?

    A. Principal component analysis
    B. Autoencoder
    C. Singular value decomposition
    D. Robust PCA

Answer: B


**Question 4:**

Which statement is true regarding decision trees and models based on ensembles of trees?

    A. In the gradient boosting algorithm, for all but the first iteration, the target is the residual from the previous decision tree model.
    B. For a Forest model, the out-of-bag sample is simply the original validation data set from when the raw data partitioning took place.
    C. In the Forest algorithm, each individual tree is pruned based on using minimum Average Squared Error.
    D. A single decision tree will always be outperformed by a model based on an ensemble of trees.

Answer: A

**Question 5:**

Refer to the treemap shown in the exhibit below:
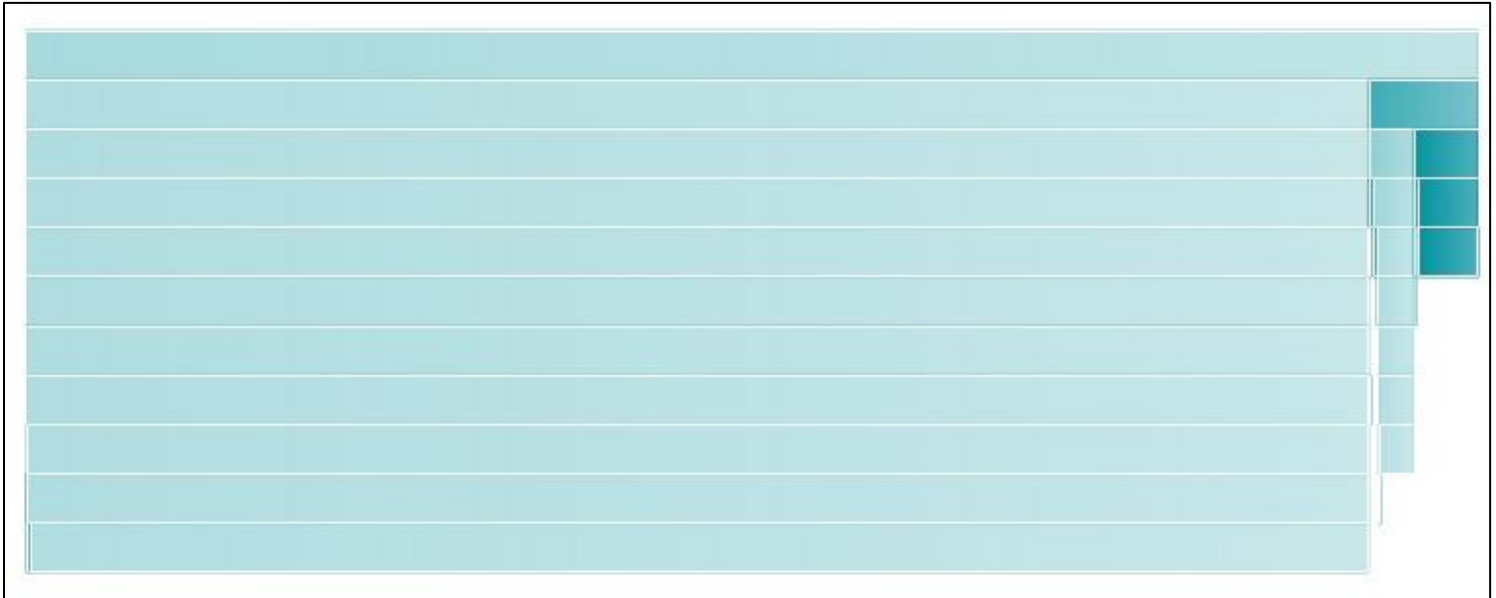


Which statement is true about the tree map for a decision tree with a binary target?

    A.   The top bar represents the node with the highest probability of event.
    B.   The darker bars represent nodes with a lower probability of event.
    C.   The top bar represents the node with the highest count.
    D.   The wider bars represent nodes with a higher probability of event.

Answer: C

**Question 6:**

Given the following properties for a neural network model, which statement is true regrading hidden units in the model? The following SAS program is submitted:

| Properties | |
|---|---|
| **Property name** | **Property value** |
| missAsLevl | false |
| inputStd | STD |
| nHidden | 1 |
| hiddenAll | false |
| hiddenAllNum | 50 |
| actFuncAll | TANH |
| hidden1 | 26 |
| actFunc1 | TANH |

    A. There are no hidden units in the model.
    B. The number of hidden units is 1.
    C. The number of hidden units is 50.
    D. The number of hidden units is 26.

Answer: D

**Question 7:**

Which statements are true for the F1 score? (Choose 2.)
    A. F1 score is calculated based on a depth value.
    B. F1 score is calculated based on a cut off value.
    C. F1 score is applicable to a model with a binary target.
    D. F1 score is applicable to a model with an interval target.

Answer: B and C