

# 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 Big Data Preparation, Statistics, and Visual Exploration

### Question 1

How do you access the Data Management Studio Options window?

- A. from the Tools menu
- B. from the Administration riser bar
- C. from the Information riser bar
- D. in the app.cfg file in the DataFlux Data Management Studio installation folder

**correct\_answer = "A"**

### Question 2

How are the Field name analysis and Sample data analysis methods similar?

- A. They both utilize a match definition from the Quality Knowledge Base.
- B. They both require the same identification analysis definition from the Quality Knowledge Base.
- C. They both utilize an identification analysis definition from the Quality Knowledge Base.
- D. They both require the same match definition from the Quality Knowledge Base.

**correct\_answer = "C"**

### Question 3

A sample of data has been clustered and found to contain many multi-row clusters. To construct a "best" record for each multi-row cluster, you need to select information from other records within a cluster. Which type of rule allows you to perform this task?

- A. Clustering rules
- B. Record rules
- C. Business rules
- D. Field rules

**correct\_answer = "D"**

### Question 4

Which option in the properties of a Clustering node allows you to identify which clustering condition was satisfied?

- A. Condition matched field prefix
- B. Cluster condition field matched

- C. Cluster condition field count
- D. Cluster condition met field

**correct\_answer = "A"**

### Question 5

A Data Quality Steward creates these items for the Supplier repository:

- A row-based business rule called Monitor for Nulls
- A set-based business rule called Percent of Verified Addresses
- A group-based rule called Low Product Count
- A task based on the row-based, set-based, and group-based rules called Monitor Supplier Data

Which one of these can the Data Quality Steward apply in an Execute Business Rule node in a data job?

- A. set-based business rule called Percent of Verified Addresses
- B. row-based business rule called Monitor for Nulls
- C. group-based rule called Low Product Count
- D. task based on the row-based, set-based, and group-based rules called Monitor Supplier Data

**correct\_answer = "B"**

### Question 6

A financial analyst wants to know whether assets in portfolio A are more risky (have higher variance) than those in portfolio B. The analyst computes the annual returns (or percent changes) for assets within each of the two groups and obtains the following output from the GLM procedure:

<i>Levene's Test for Homogeneity of Return Variance</i>					
<i>ANOVA of Squared Deviations from Group Means</i>					
<i>Source</i>	<i>DF</i>	<i>Sum of Squares</i>	<i>Mean Square</i>	<i>F Value</i>	<i>Pr &gt; F</i>
<i>Portfolio</i>	1	217.2	217.2	16.29	0.0005
<i>Error</i>	23	306.7	13.3352		

Which conclusion is supported by the output?

- A. Assets in portfolio A are significantly more risky than assets in portfolio B.
- B. Assets in portfolio B are significantly more risky than assets in portfolio A.
- C. The portfolios differ significantly with respect to risk.
- D. The portfolios do not differ significantly with respect to risk.

**correct\_answer = "C"**

### Question 7

A linear model has the following characteristics:

- a dependent variable (y)
- one continuous predictor variables (x1) including a quadratic term (x1<sup>2</sup>)
- one categorical predictor variable (c1 with 3 levels)
- one interaction term (c1 by x1)

Which SAS program fits this model?

- A. 

```
proc glm data=SASUSER.MLR;  
  class c1;  
  model y = c1 x1 x1sq c1byx1 /solution;  
run;
```
- B. 

```
proc reg data=SASUSER.MLR;  
  model y = c1 x1 x1sq c1byx1 /solution;  
run;
```
- C. 

```
proc glm data=SASUSER.MLR;  
  class c1;  
  model y = c1 x1 x1*x1 c1*x1 /solution;  
run;
```
- D. 

```
proc reg data=SASUSER.MLR;  
  model y = c1 x1 x1*x1 c1*x1;  
run;
```

**correct\_answer = "C"**

### Question 8

When selecting variables or effects using SELECTION=BACKWARD in the LOGISTIC procedure, the business analyst's model selection terminated at Step 3.

What happened between Step 1 and Step 2?

- A. DF increased.
- B. AIC increased.
- C. Pr > Chisq increased.
- D. - 2 Log L increased.

**correct\_answer = "D"**

### Question 9

In SAS Visual Analytics Explorer, when a date data item is dragged onto an Automatic Chart visualization either a bar chart or a line chart will be created. What determines the type of chart created?

- A. The format applied to the date data item determines the type of chart displayed.
- B. A bar chart is created if the Model property of the data item is set to Discrete, and a line chart is created if the Model property is set to Continuous.
- C. The properties associated with the automatic chart determines the type of chart displayed.
- D. A line chart is created if the Model property of the data item is set to Discrete, a bar chart is created if the Model property is set to Continuous.

**correct\_answer = "B"**

### Question 10

Using SAS Visual Analytics Explorer, a content developer would like to examine the relationship between two measures with high cardinality. Which visualization should the developer use?

- A. Scatter Plot
- B. Heat Map
- C. Scatter Plot Matrix
- D. Treemap

**correct\_answer = "B"**