

QUESTIONS & ANSWERS

Kill your exam at first Attempt



A00-240 Dumps
A00-240 Braindumps
A00-240 Real Questions
A00-240 Practice Test
A00-240 dumps free



SASInstitute

A00-240

SAS Statistical Business Analysis SAS9: Regression and Model

<http://killexams.com/pass4sure/exam-detail/A00-240>



Question #87

What is a benefit to performing data cleansing (imputation, transformations, etc.) on data after partitioning the data for honest assessment as opposed to performing the data cleansing prior to partitioning the data?

- A. It makes inference on the model possible.
- B. It is computationally easier and requires less time.
- C. It omits the training (and test) data sets from the benefits of the cleansing methods.
- D. It allows for the determination of the effectiveness of the cleansing method.

Answer: D

Question #88

A researcher has several variables that could be possible predictors for the final model. There is interest in checking all 2-way interactions for possible entry to the model. The researcher has decided to use forward selection within PROC LOGISTIC. Fill in the missing code option that will ensure that all 2-way interactions will be considered for entry.

```
proc logistic data=finances;
  class gender branch;
  model default = gender branch income debt
                gender|branch|income|debt @2
  / selection=forward slentry=0.01 <insert option here> ;
run;
```

- A. start = 5
- B. include = 4
- C. include = 5
- D. start = 4

Answer: C

Question #89

FILL BLANK -

Refer to the confusion matrix:

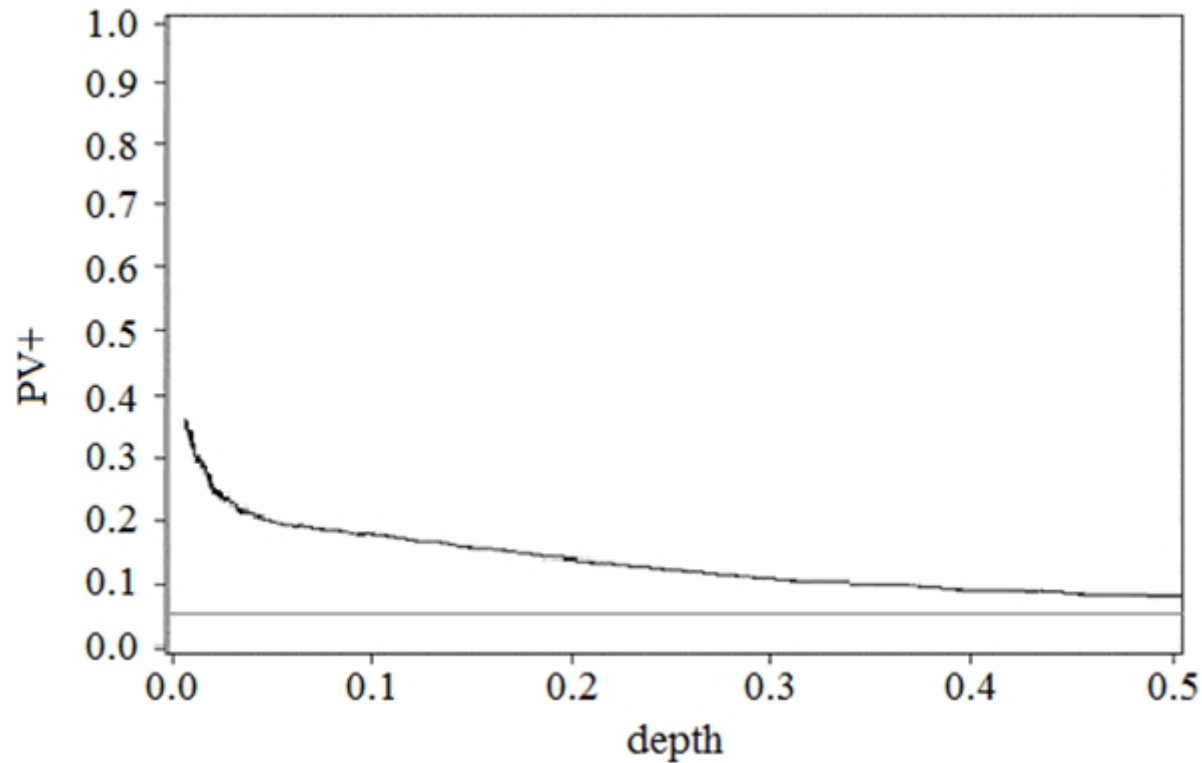
		Predicted Outcome	
		0	1
Actual Outcome	0	345	155
	1	188	312

An analyst determines that loan defaults occur at the rate of 3% in the overall population. The above confusion matrix is from an oversampled test set (1 = default). What is the sensitivity adjusted for the population event probability? Enter your answer in the space below. Round to three decimals (example: n.nnn).

Answer: 0.617

Question #90

Refer to the exhibit:



On the Gains Chart, what is the correct interpretation of the horizontal reference line?

- A. the proportion of cases that cannot be classified
- B. the probability of a false negative
- C. the probability of a false positive
- D. the prior event rate

Answer: B

Question #91

Refer to the confusion matrix:

		Predicted Outcome	
		0	1
Actual Outcome	0	58	44
	1	23	25

Calculate the accuracy and error rate (0 - negative outcome, 1 - positive outcome)

- A. Accuracy = 58/102, Error Rate = 23/48
- B. Accuracy = 83/102, Error Rate = 67/102
- C. Accuracy = 25/150, Error Rate = 44/150
- D. Accuracy = 83/150, Error Rate = 67/150

Answer: A

Question #92

Which statistic is based on the maximum vertical distance between the primary event EDF and the secondary event EDF?

- A. KS
- B. SBC
- C. Max EDF
- D. Brier Score

Answer: A

Reference:

<https://support.sas.com/documentation/onlinedoc/ets/132/severity.pdf>

Question #93

DRAG DROP -

Drag the adjustment formulas for oversampling from the left and place them into the correct location in the confusion matrix shown on the right.

Select and Place:

$\pi_0 \times \text{specificity}$
$\pi_0 \times (1 - \text{specificity})$
$\pi_1 \times \text{sensitivity}$
$\pi_1 \times (1 - \text{sensitivity})$

		Predicted Class	
		0	1
Actual Class	0	formula	formula
	1	formula	formula

$\pi_0 \times \text{specificity}$		Predicted Class	
$\pi_0 \times (1 - \text{specificity})$		0	1
$\pi_1 \times \text{sensitivity}$	Actual Class	0	$\pi_1 \times \text{sensitivity}$
$\pi_1 \times (1 - \text{sensitivity})$		1	$\pi_0 \times (1 - \text{specificity})$
			$\pi_1 \times (1 - \text{sensitivity})$

Answer:

Question #94

An analyst knows that the categorical predictor, zip_code, is an important predictor of a binary target. However, zip_code has too many levels to be a feasible predictor in a model. The analyst uses PROC CLUSTER to implement Greenacre's method to reduce the number of categorical levels. What is the correct application of Greenacre's method in this situation?

- A. Clustering the levels using the target proportion for each zip_code as input.
- B. Clustering the levels using the zip_code values as input.
- C. Clustering the levels using the number of cases in each zip_code as input.
- D. Clustering the levels using dummy coded zip_code levels as inputs.

Answer: A

Reference:

<https://support.sas.com/resources/papers/proceedings/proceedings/sugi31/079-31.pdf>

Question #95

What does the Pearson product moment correlation coefficient measure?

- A. nonlinear and nonmonotonic association between two variables
- B. linear and monotonic association between two variables
- C. linear and nonmonotonic association between two variables
- D. nonlinear and monotonic association between two variables

Answer: B

Reference:

http://d-scholarship.pitt.edu/8056/1/Chokns_etd2010.pdf

Question #96

This question will ask you to provide a segment of missing code.

The following code is used to create missing value indicator variables for input variables, fred1 to fred7.

```
data work.train_mi;

    set work.train;

    array mi{*} MI_fred1-MI_fred7;
    array x{*} fred1-fred7;
```

<insert code here>

```
run;
```

Which segment of code would complete the task?

A.

```
do i=1 to dim(mi);  
    mi{i}=(x{i}=" ");  
end;
```

B.

```
do i=1 to dim(mi);  
    mi{i}=(x{i}=.);  
end;
```

C.

```
do 1 to 7;  
    if missing(array(x{*})) then array(mi{*})=1;  
end;
```

D.

```
do 1 to 7;  
    array(mi)=missing(array(x));  
end;
```

Answer: C

Question #97

This question will ask you to provide a missing option.

Given the following SAS program:

```
proc corr data = MYDATA <insert option here> ;  
    var x1 x2 x3 x4 x5;  
    with Target;  
run;
```

What option must be added to the program to obtain a data set containing Spearman statistics?

- A. OUTCORR=estimates
- B. OUTS=estimates
- C. OUT=estimates
- D. OUTPUT=estimates

Answer: D

Question #98

This question will ask you to provide a missing option.

A business analyst is investigating the differences in sales figures across 8 sales regions. The analyst is interested in viewing the regression equation parameter estimates for each of the design variables.

Which option completes the program to produce the regression equation parameter estimates?

```
proc glm data=sales_spring;  
    class region;  
    model finalsales = region / <insert option here> ;  
run;  
quit;
```

- A. Solve
- B. Estimate
- C. Solution
- D. Est

Answer: C

Reference:

https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_ods_examples06.htm&docsetVersion=14.3&locale=en

Question #99

After performing an ANOVA test, an analyst has determined that a significant effect exists due to income. The analyst wants to compare each Income to all others and wants to control for experimentwise error.

Which GLM procedure statement would provide the most appropriate output?

- A. lsmeans Income / pdiff=control adjust=dunnett;
- B. lsmeans Income / pdiff=control adjust=t;
- C. lsmeans Income / pdiff=all adjust=tukey;
- D. lsmeans Income / pdiff=all adjust=t;

Answer: A

Reference:

https://rpubs.com/JsoLab/Stat01_L02

Question #100

SIMULATION -

A linear model has the following characteristics:

*A dependent variable (y)

*One continuous variable (x1), including a quadratic term (x1²)

*One categorical (d with 3 levels) predictor variable and an interaction term (d by x1)

How many parameters, including the intercept, are associated with this model?

Enter your numeric answer in the space below. Do not add leading or trailing spaces to your answer.

Answer: 7

For More exams visit <https://killexams.com/vendors-exam-list>



Kill your exam at First Attempt....Guaranteed!