A00-240 Q&As

SAS Certified Statistical Business Analyst Using SAS 9: Regression and Modeling Credential

Pass SASInstitute A00-240 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.leads4pass.com/a00-240.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by SASInstitute Official Exam Center

- **Instant Download** After Purchase
- **100% Money Back** Guarantee
- **365 Days** Free Update
- **800,000+** Satisfied Customers

QUESTION 1

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

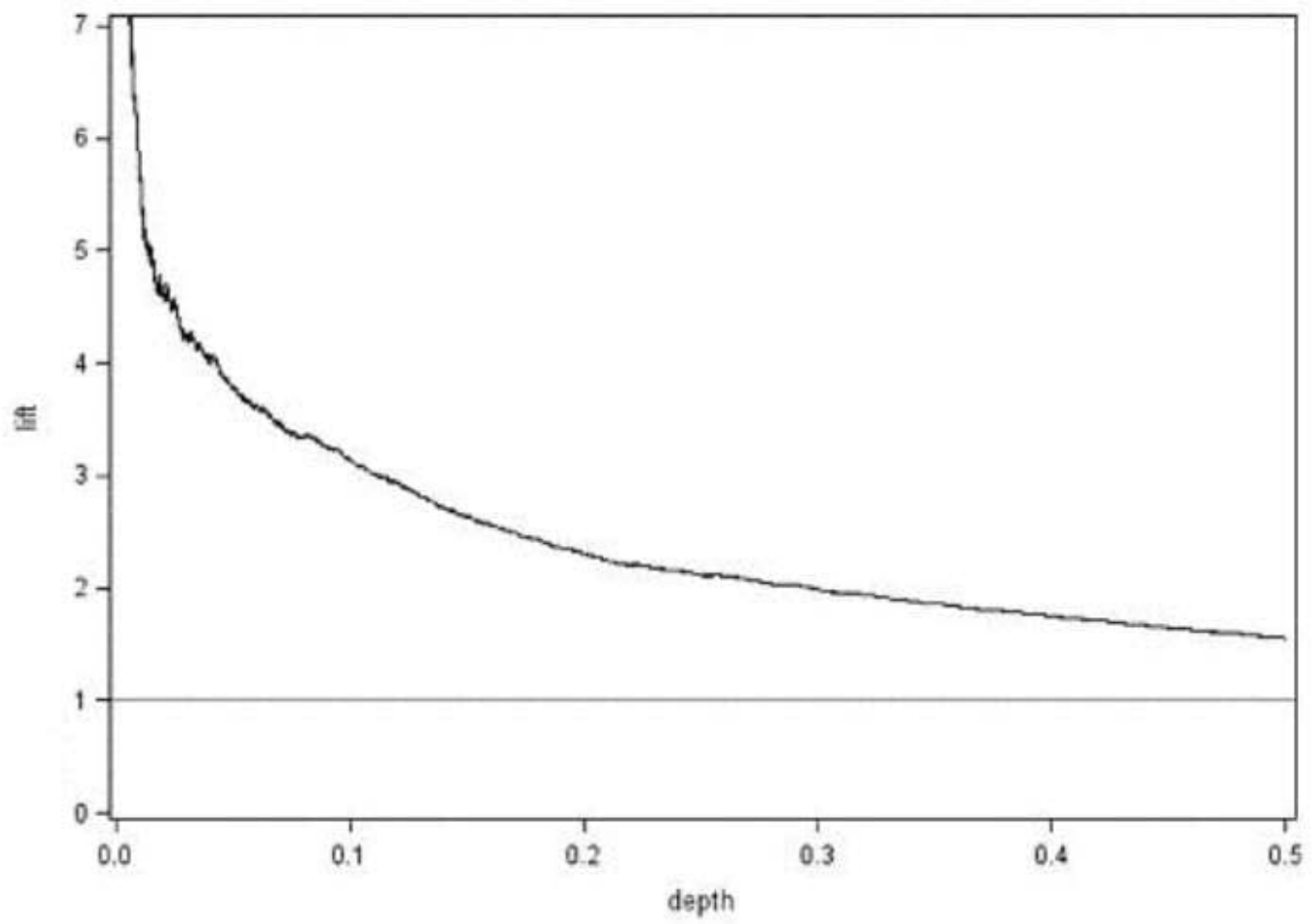
C. Accuracy = 25/150, Error Rate = 44/150

D. Accuracy = 83/150, Error Rate = 67/150

Correct Answer: A

QUESTION 2

Refer to the lift chart:



At a depth of 0.1, Lift = 3.14. What does this mean?

A. Selecting the top 10% of the population scored by the model should result in 3.14 times more events than a random draw of 10%.

B. Selecting the observations with a response probability of at least 10% should result in 3.14 times more events than a random draw of 10%.

C. Selecting the top 10% of the population scored by the model should result in 3.14 times greater accuracy than a random draw of 10%.

D. Selecting the observations with a response probability of at least 10% should result in 3.14 times greater accuracy than a random draw of 10%.

Correct Answer: A

QUESTION 3

The Model SS in a multiple linear regression model is equal to:

A. the total SS- MSE

B. the sum of Type I SS of all model terms

C. the sum of Type II SS of all model terms

D. the sum of SSE and MSE

Correct Answer: B

Reference: http://core.ecu.edu/psyc/wuenschk/SAS/SS1234.pdf

QUESTION 4

Which characteristic of Studentized residuals indicate potential outliers?

A. Only studentized residuals greater than negative two

B. Only studentized residuals less than negative two and greater than two

C. Only studentized residuals greater than two

D. Only studentized residuals less than two and greater than negative two

Correct Answer: C

QUESTION 5

The selection criterion used in the forward selection method in the REG procedure is:

A. Adjusted R-Square

B. SLE

C. Mallows\\' Cp

D. AIC

Correct Answer: B

QUESTION 6

While building a predictive model, median imputations are performed while preparing the training data. How should the imputations be addressed in the validation data?

A. The imputed values are irrelevant to the validation data, and are not used.

B. The imputed values must be applied directly to the validation data without recalculation.

C. The imputed values must be recalculated using the validation data.

D. The imputed values must be recalculated using both the training and the validation data.

Correct Answer: C

QUESTION 7

The standard form of a linear regression model is:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Which statement best summarizes the assumptions placed on the errors?

A. The errors are correlated, normally distributed with constant mean and zero variance.

B. The errors are correlated, normally distributed with zero mean and constant variance.

C. The errors are independent, normally distributed with constant mean and zero variance.

D. The errors are independent, normally distributed with zero mean and constant variance.

Correct Answer: D

QUESTION 8

Assume a $10 cost for soliciting a non-responder and a $200 profit for soliciting a responder. The logistic regression model gives a probability score named P_R on a SAS data set called VALID. The VALID data set contains the responder variable Pinch, a 1/0 variable coded as 1 for responder. Customers will be solicited when their probability score is more than 0.05.

Which SAS program computes the profit for each customer in the data set VALID?

A.
```
data VALID;
   set VALID;
   Profit = (P_R > .05)*Purch*200  - (P_R > .05)*(1 - Purch)*10;
run;
```

B.
```
data VALID;
   set VALID;
   Profit = (P_R <= .05)*Purch*200  - (P_R > .05)*(1 - Purch)*10;
run;
```

C.
```
data VALID;
   set VALID;
   if P_R > .05;
   Profit = (P_R > .05)*Purch*200  - (P_R > .05)*(1 - Purch)*10;
run;
```

D.
```
data VALID;
   set VALID;
    if P_R >.05;
   Profit = (P_R > .05)*Purch*200  + (P_R <= .05)*(1 - Purch)*10;
run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: A

QUESTION 9

Refer to the REG procedure output:

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	31848	15924	13.42	<.0001
Error	97	115082	1186.40833		
Corrected Total	99	146930			

Root MSE	34.44428	R-Square	0.2168
Dependent Mean	606.38715	Adj R-Sq	0.2006
Coeff Var	5.68025		

An analyst has selected this model as a champion because it shows better model fit than a competing model with more predictors. Which statistic justifies this rationale?

A. R-Square

B. Coeff Var

C. Adj R-Sq

D. Error DF

Correct Answer: C

QUESTION 10

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

A. It violates assumptions of the model.

B. It requires extra computational effort and time.

C. It omits the training (and test) data sets from the benefits of the cleansing methods.

D. There is no ability to compare the effectiveness of different cleansing methods.

Correct Answer: D

QUESTION 11

A marketing campaign will send brochures describing an expensive product to a set of customers. The cost for mailing and production per customer is $50. The company makes $500 revenue for each sale. What is the profit matrix for a typical person in the population?

A.

	Purchase	
Solicit	No	Yes
No	-50	0
Yes	0	450

B.

	Purchase	
Solicit	No	Yes
No	0	0
Yes	-50	500

C.

	Purchase	
Solicit	No	Yes
No	0	0
Yes	-50	450

D.

	Purchase	
Solicit	No	Yes
No	-50	0
Yes	0	500

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: C

QUESTION 12

Refer to the exhibit:

Number in Model	R-Square	Adjusted R-Square	C(p)	AIC	Root MSE	SBC	Variables in Model
1	0.7434	0.7345	13.6988	64.5341	2.74478	67.40210	RunTime
1	0.1595	0.1305	106.3021	101.3131	4.96748	104.18108	RestPulse
2	0.7642	0.7474	12.3894	63.9050	2.67739	68.20695	Age RunTime
2	0.7614	0.7444	12.8372	64.2740	2.69337	68.57597	RunTime RunPulse
3	0.8111	0.7901	6.9596	59.0373	2.44063	64.77326	Age RunTime RunPulse
3	0.8100	0.7889	7.1350	59.2183	2.44777	64.95424	RunTime RunPulse MaxPulse
4	0.8368	0.8117	4.8800	56.4995	2.31159	63.66941	Age RunTime RunPulse MaxPulse
4	0.8165	0.7883	8.1035	60.1386	2.45133	67.30850	Age Weight RunTime RunPulse
5	0.8480	0.8176	5.1063	56.2986	2.27516	64.90250	Age Weight RunTime RunPulse MaxPulse
5	0.8370	0.8044	6.8461	58.4590	2.35583	67.06288	Age RunTime RunPulse RestPulse MaxPulse
6	0.8487	0.8108	7.0000	58.1616	2.31695	68.19952	Age Weight RunTime RunPulse RestPulse MaxPulse

SAS output from the RSQUARE selection method, within the REG procedure, is shown. The top two models in each subset are given. Based on the exhibit, which statement is true?

A. The AIC champion model is more parsimonious than the SBC champion.

B. The SBC champion model is more parsimonious than the AIC champion.

C. The R-Square champion model is the most parsimonious.

D. Adjusted R-Square and R-Square agree on the champion model.

Correct Answer: B

QUESTION 13

In partitioning data for model assessment, which sampling methods are acceptable? (Choose two.)

A. Simple random sampling without replacement

B. Simple random sampling with replacement

C. Stratified random sampling without replacement

D. Sequential random sampling with replacement

Correct Answer: AC

QUESTION 14

Refer to the confusion matrix:

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

Calculate the sensitivity. (0 - negative outcome, 1 - positive outcome)

Click the calculator button to display a calculator if needed.

A. 25/48

B. 58/102

C. 25/B9

D. 58/81

Correct Answer: A

QUESTION 15

What is the default method in the LOGISTIC procedure to handle observations with missing data?

A. Missing values are imputed.

B. Parameters are estimated accounting for the missing values.

C. Parameter estimates are made on all available data.

D. Only cases with variables that are fully populated are used.

Correct Answer: D