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.certbus.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

Which SAS program will correctly use backward elimination selection criterion within the REG procedure?

A.
```
proc reg data=SASUSER.MLR;
   model y = x1-x10 /selection=backward sls=aic;
run;
```

B.
```
proc reg data=SASUSER.MLR;
   model y = x1-x10 /selection=backward sls=0.15;
run;
```

C.
```
proc reg data=SASUSER.MLR;
   model y = x1-x10 /selection=backward sle=cp;
run;
```

D.
```
proc reg data=SASUSER.MLR;
   model y = x1-x10 /selection=backward sle=all;
run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: B

QUESTION 2

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;

Correct Answer: A

Reference: https://rpubs.com/JsoLab/Stat01_L02

QUESTION 3

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 Pearson statistics?

A. OUTPUT=estimates

B. OUTP=estimates

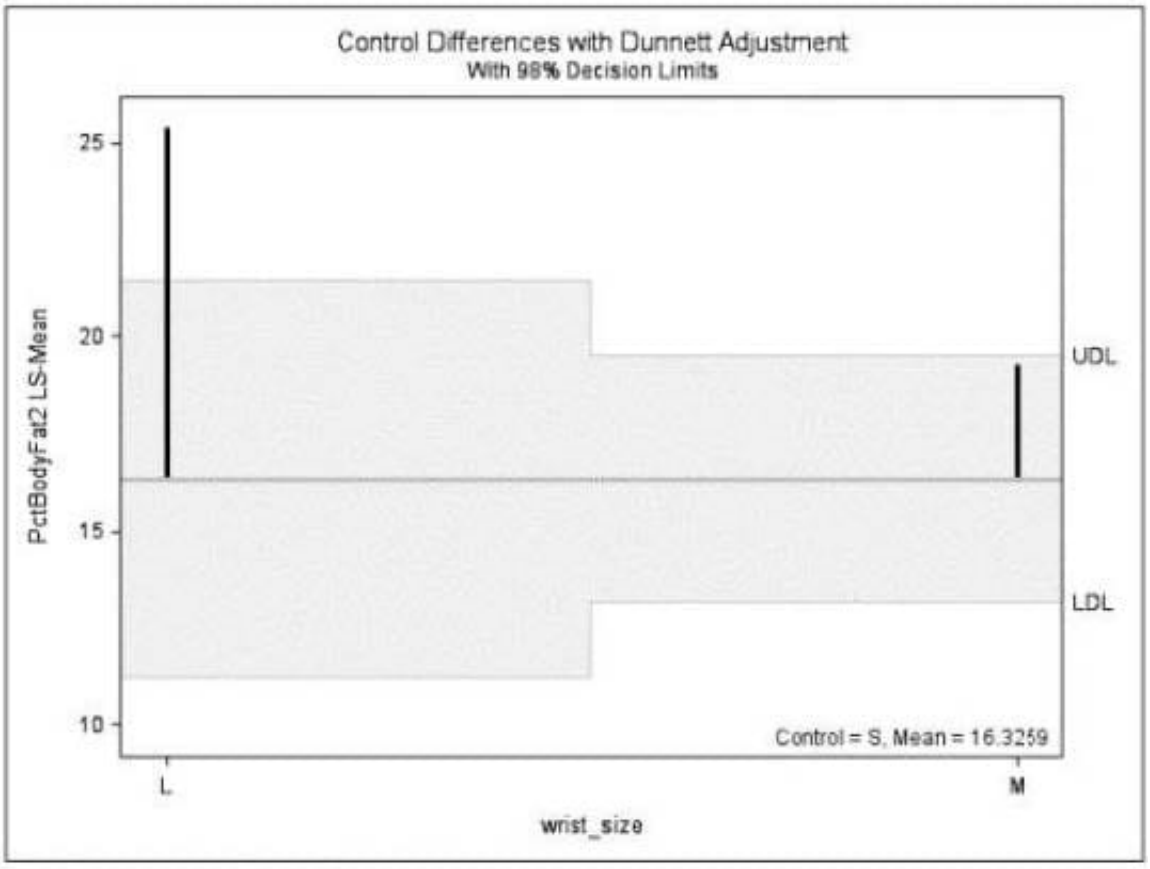
C. OUTSTAT=estimates

D. OUTCORR=estimates

Correct Answer: B

QUESTION 4

Refer to the exhibit.



Given alpha=0.02, which conclusion is justified regarding percentage of body fat, comparing small (S), medium (M), and large (L) wrist sizes?

A. Medium wrist size is significantly different than small wrist size.

B. Large wrist size is significantly different than medium wrist size.

C. Large wrist size is significantly different than small wrist size.

D. There is no significant difference due to wrist size.

Correct Answer: C

QUESTION 5

In order to perform honest assessment on a predictive model, what is an acceptable division between training, validation, and testing data?

A. Training: 50% Validation: 0% Testing: 50%

B. Training: 100% Validation: 0% Testing: 0%

C. Training: 0% Validation: 100% Testing: 0%

D. Training: 50% Validation: 50% Testing: 0%

Correct Answer: D

QUESTION 6

Which of the following describes a concordant pair of observations in the LOGISTIC procedure?

A. An observation with the event has an equal probability as another observation with the event.

B. An observation with the event has a lower predicted probability than the observation without the event.

C. An observation with the event has an equal predicted probability as the observation without the event.

D. An observation with the event has a higher predicted probability than the observation without the event

Correct Answer: D

QUESTION 7

A company has branch offices in eight regions. Customers within each region are classified as either "High Value" or "Medium Value" and are coded using the variable name VALUE. In the last year, the total amount of purchases per customer is used as the response variable.

Suppose there is a significant interaction between REGION and VALUE. What can you conclude?

A. More high value customers are found in some regions than others.

B. The difference between average purchases for medium and high value customers depends on the region.

C. Regions with higher average purchases have more high value customers.

D. Regions with higher average purchases have more medium value customers.

Correct Answer: B

QUESTION 8

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

A. RSQ

B. MSE

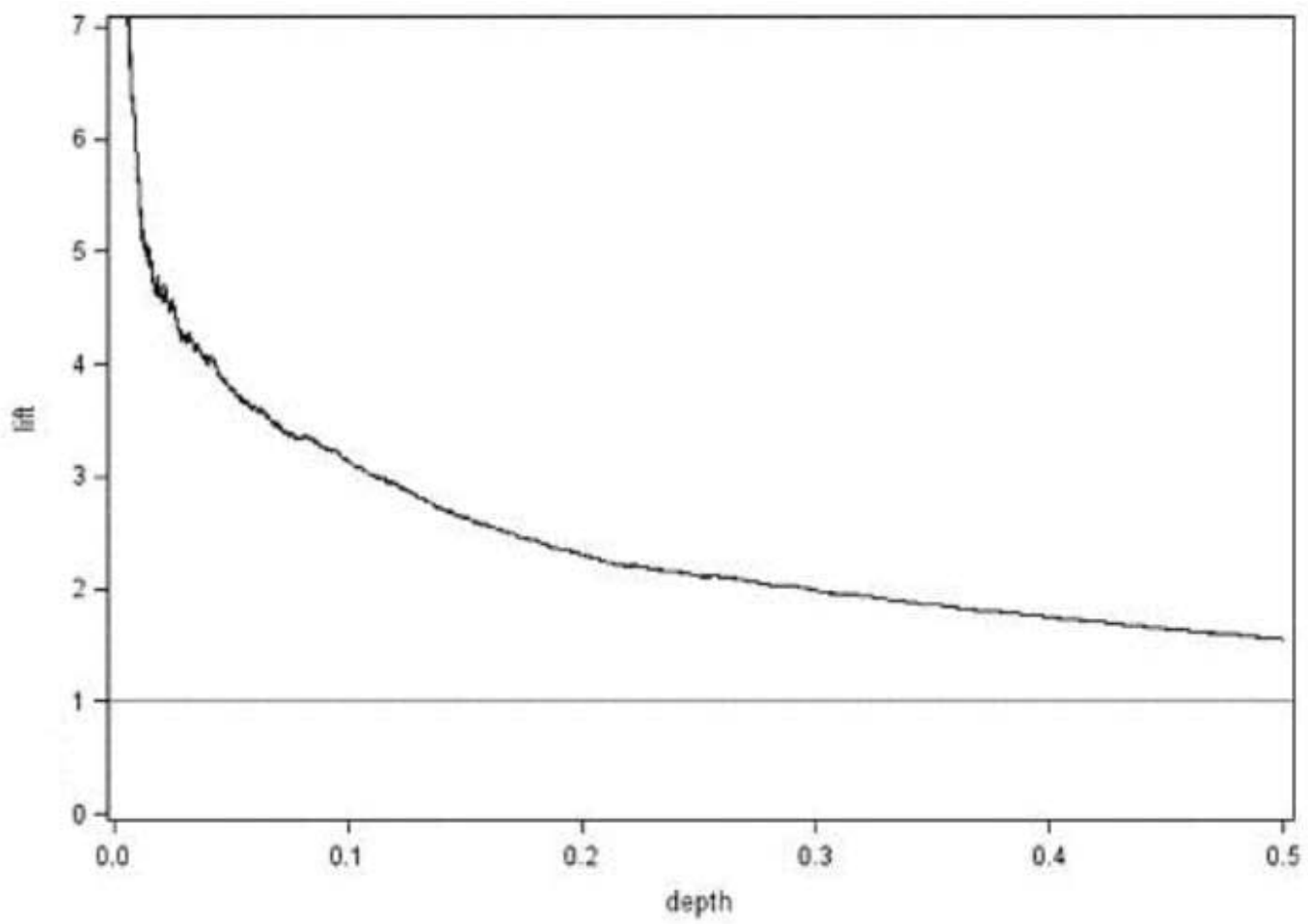
C. R-squared

D. AIC

Correct Answer: D

QUESTION 9

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 10

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

Correct Answer: C

Reference: https://documentation.sas.com/?docsetId=statuganddocsetTarget=statug_ods_examples06.htmanddocsetVersion=14.3andlocale=en

QUESTION 11

Refer to the REG procedure output:

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Standardized Estimate
Intercept	1	618.44051	40.03665	15.45	<.0001	0
overhead	1	4.99845	0.00157	3181.24	<.0001	0.99993
scrap	1	2.02667	0.71501	3.95	<.0001	0.00124
training	1	-50.95436	2.82069	-18.06	<.0001	-0.00568

The Intercept estimate is interpreted as:

A. The predicted value of the response when all the predictors are at their current values.

B. The predicted value of the response when all predictors are at their means.

C. The predicted value of the response when all predictors = 0.

D. The predicted value of the response when all predictors are at their minimum values.

Correct Answer: C

QUESTION 12

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 13

An analyst compares the mean salaries of men and women working at a company. The SAS data set SALARY contains variables:

1.

Gender (M or F)

2.

Pay (dollars per year)

Which SAS programs can be used to find the p-value for comparing men\\'s salaries with women\\'s salaries? (Choose two.)

A.
```
proc glm data = SALARY;
   class Gender;
   model Pay = Gender;
run;
```

B.
```
proc ttest data = SALARY;
   class Gender;
   var Pay;
run;
```

C.
```
proc glm data = SALARY;
   class Pay;
   model Pay = Gender;
run;
```

D.
```
proc ttest data = SALARY;
   class Gender;
   model Pay = Gender;
run;
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: AB

QUESTION 14

The total modeling data has been split into training, validation, and test data. What is the best data to use for model assessment?

A. Training data

B. Total data

C. Test data

D. Validation data

Correct Answer: D

QUESTION 15

When mean imputation is performed on data after the data is partitioned for honest assessment, what is the most appropriate method for handling the mean imputation?

A. The sample means from the validation data set are applied to the training and test data sets.

B. The sample means from the training data set are applied to the validation and test data sets.

C. The sample means from the test data set are applied to the training and validation data sets.

D. The sample means from each partition of the data are applied to their own partition.

Correct Answer: B