

Exam Questions A00-240

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

<https://www.2passeasy.com/dumps/A00-240/>



NEW QUESTION 1

Select the equivalent LOGISTIC procedure model statements. (Choose two.)

- A. Mode1 Purchase * Gender Age Region;
- B. Mode1 Purchase * Gender | Age | Region;
- C. Mode1 Purchase * Gender|Age|Region @1;
- D. Mode1 Purchase * Gender|Age|Region @2;

Answer: AC

NEW QUESTION 2

Identify the correct SAS program for fitting a multiple linear regression model with dependent variable (y) and four predictor variables (x1-x4).

- ☐ A.

```
proc reg data=SASUSER.MLR;
    var y x1 x2 x3 x4;
    model y = x1-x4;
run;
```
- ☐ B.

```
proc reg data=SASUSER.MLR;
    model y = x1-x4;
run;
```
- ☐ C.

```
proc reg data=SASUSER.MLR;
    model y = x1;
    model y = x2;
    model y = x3;
    model y = x4;
run;
```
- ☐ D.

```
proc reg data=SASUSER.MLR;
    model y = x1 x2 x3 x4 /solution;
run;
```

- A. Option A
- B. Option B
- C. Option C
- D. Option D

Answer: B

NEW QUESTION 3

Suppose training data are oversampled in the event group to make the number of events and non-events roughly equal. A logistic regression is run and the probabilities are output to a data set NEW and given the variable name PE. A decision rule considered is, "Classify data as an event if probability is greater than 0.5." Also the data set NEW contains a variable TG that indicates whether there is an event (1=Event, 0= No event).

The following SAS program was used.

```
data NEW;
    set NEW;
    Solicit = PE > .5;
run;
proc means data=NEW(where = (TG=1)) mean;
    var Solicit;
run;
```

What does this program calculate?

- A. Depth
- B. Sensitivity
- C. Specificity
- D. Positive predictive value

Answer: B

NEW QUESTION 4

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 AIC statistic, which model is the champion model?

- A. Age Weight RunTime RunPulse MaxPulse
- B. Age Weight RunTime RunPulse RestPulse MaxPulse
- C. RestPulse
- D. RunTime

Answer: A

NEW QUESTION 5

The question will ask you to provide a missing statement. Given the following SAS program:

```
proc logistic data = MYDIR.DEFAULT_DATA des;
  model Purchase = Money Acct_type Debt Employment;
  <insert statement here>
run;
```

Which SAS statement will complete the program to correctly score the data set NEW_DATA?

- A. Scoredata data=MYDIR.NEW_DATA out=scores;
- B. Scoredata data=MYDIR.NEW_DATA output=scores;
- C. Scoredata=HYDIR.NEU_DATA output=scores;
- D. Scoredata=MYDIR,NEW DATA out=scores;

Answer: D

NEW QUESTION 6

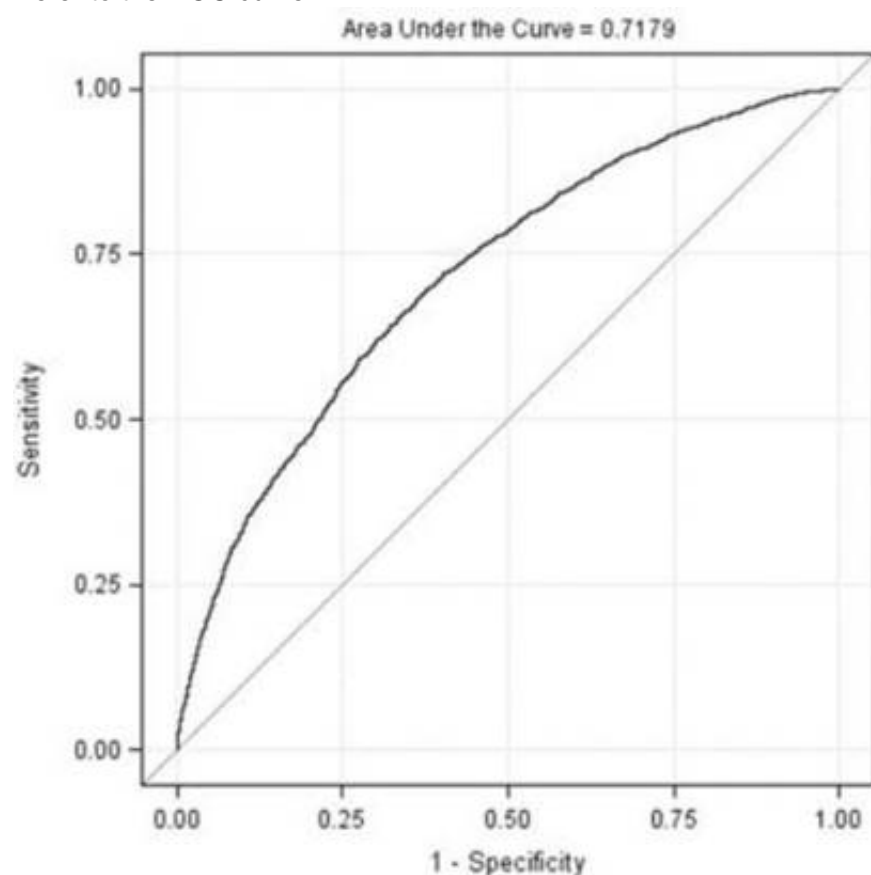
What is a drawback to performing data cleansing (imputation, transformations, etc.) on raw data prior to partitioning the data for honest assessment as opposed to performing the data cleansing after partitioning the data?

- 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.

Answer: D

NEW QUESTION 7

Refer to the ROC curve:



As you move along the curve, what changes?

- A. The priors in the population
- B. The true negative rate in the population
- C. The proportion of events in the training data
- D. The probability cutoff for scoring

Answer: D

NEW QUESTION 8

Spearman statistics in the CORR procedure are useful for screening for irrelevant variables by investigating the association between which function of the input variables?

- A. Concordant and discordant pairs of ranked observations
- B. Logit link ($\log(p/1-p)$)
- C. Rank-ordered values of the variables
- D. Weighted sum of chi-square statistics for 2x2 tables

Answer: C

NEW QUESTION 9

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	2.82667	0.71581	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.

Answer: C

NEW QUESTION 10

Screening for non-linearity in binary logistic regression can be achieved by visualizing:

- A. A scatter plot of binary response versus a predictor variable.
- B. A trend plot of empirical logit versus a predictor variable.
- C. A logistic regression plot of predicted probability values versus a predictor variable.
- D. A box plot of the odds ratio values versus a predictor variable.

Answer: B

NEW QUESTION 10

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

Answer: A

NEW QUESTION 11

This question will ask you to provide a missing option.

Complete the following syntax to test the homogeneity of variance assumption in the GLM procedure:

Means Region / <insert option here> =levene;

- A. test
- B. adjust
- C. var
- D. hovtest

Answer: D

NEW QUESTION 15

Refer to the following odds ratio table:

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Effect	Unit	Estimate	95% Confidence Limits	
salary	1.0000	1.142	1.083	1.220

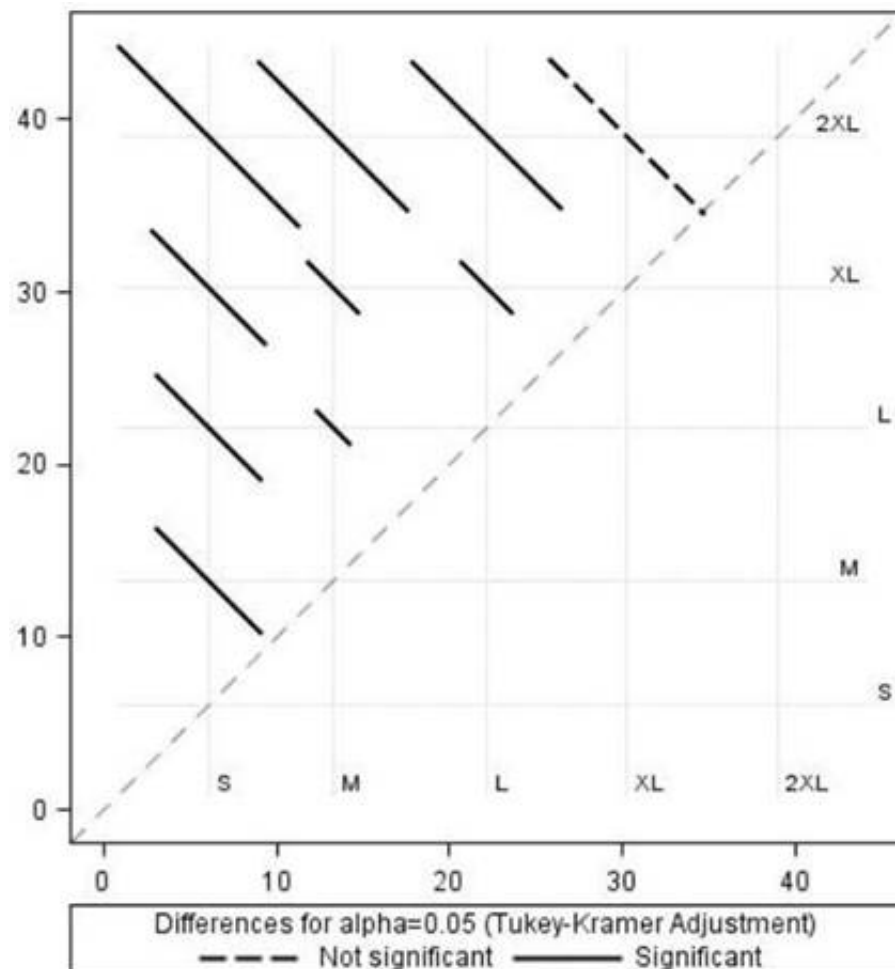
What is a correct interpretation of the estimate?

- A. The odds of the event are 1.142 greater for each one dollar increase in salary.
- B. The odds of the event are 1.142 greater for each one thousand dollar increase in salary.
- C. The probability of the event is 1.142 greater for each one dollar increase in salary.
- D. The probability of the event is 1.142 greater for each one thousand dollar increase in salary.

Answer: B

NEW QUESTION 16

Refer to the exhibit.



Based on the control plot, which conclusion is justified regarding the means of the response?

- A. All groups are significantly different from each other.
- B. 2XL is significantly different from all other groups.
- C. Only XL and 2XL are not significantly different from each other.
- D. No groups are significantly different from each other.

Answer: C

NEW QUESTION 19

Given the following GLM procedure output:

Source	DF	Type III SS	Mean Square	F Value	Pr > F
School	3	17905.24929	5968.41643	4.14	0.0073
Gender	1	1578.63006	1578.63006	1.09	0.2971
School*Gender	3	17205.36689	5735.12230	3.97	0.0091

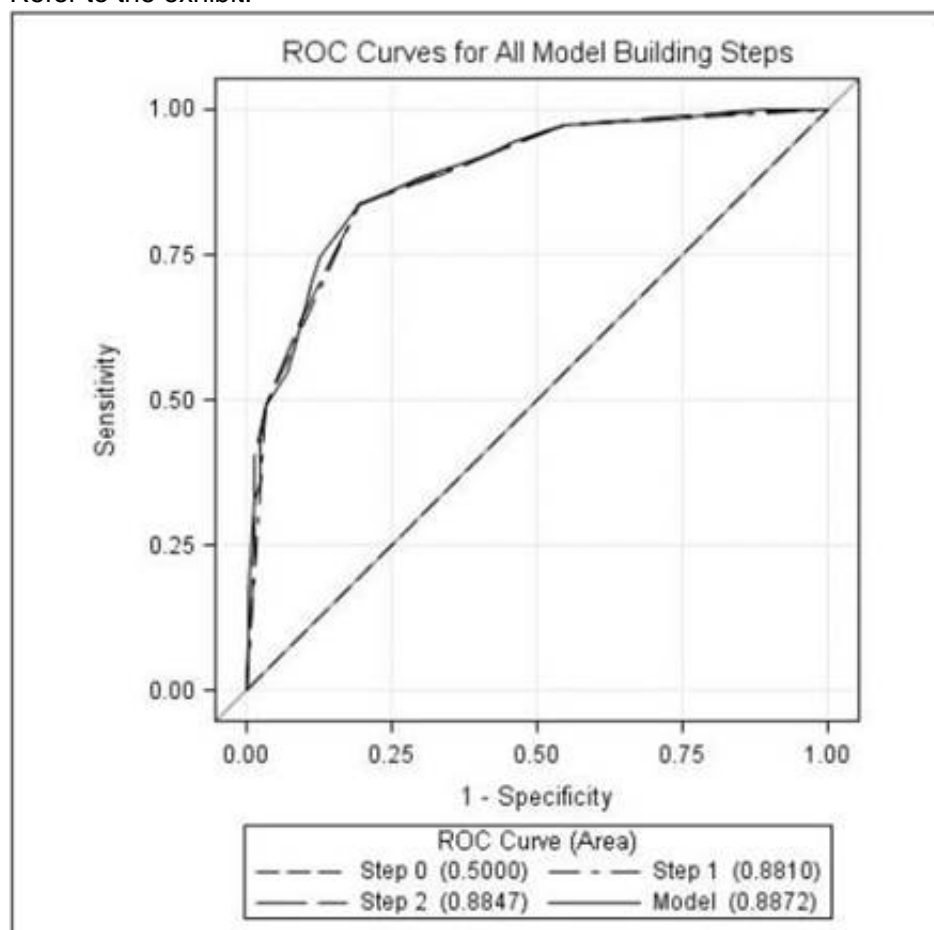
Which statement is correct at an alpha level of 0.05?

- A. School*Gender should be removed because it is non-significant.
- B. Gender should be removed because it is non-significant.
- C. School should be removed because it is significant.
- D. Gender should not be removed due to its involvement in the significant interaction.

Answer: D

NEW QUESTION 20

Refer to the exhibit:



An analyst examined logistic regression models for predicting whether a customer would make a purchase. The ROC curve displayed summarizes the models. Using the selected model and the analyst's decision rule, 25% of the customers who did not make a purchase are incorrectly classified as purchasers. What can be concluded from the graph?

- A. About 25% of the customers who did make a purchase are correctly classified as making a purchase.
- B. About 50% of the customers who did make a purchase are correctly classified as making a purchase.
- C. About 85% of the customers who did make a purchase are correctly classified as making a purchase.
- D. About 95% of the customers who did make a purchase are correctly classified as making a purchase.

Answer: C

NEW QUESTION 21

An analyst has a sufficient volume of data to perform a 3-way partition of the data into training, validation, and test sets to perform honest assessment during the model building process.

What is the purpose of the test data set?

- A. To provide a unbiased measure of assessment for the final model.
- B. To compare models and select and fine-tune the final model.
- C. To reduce total sample size to make computations more efficient.
- D. To build the predictive models.

Answer: A

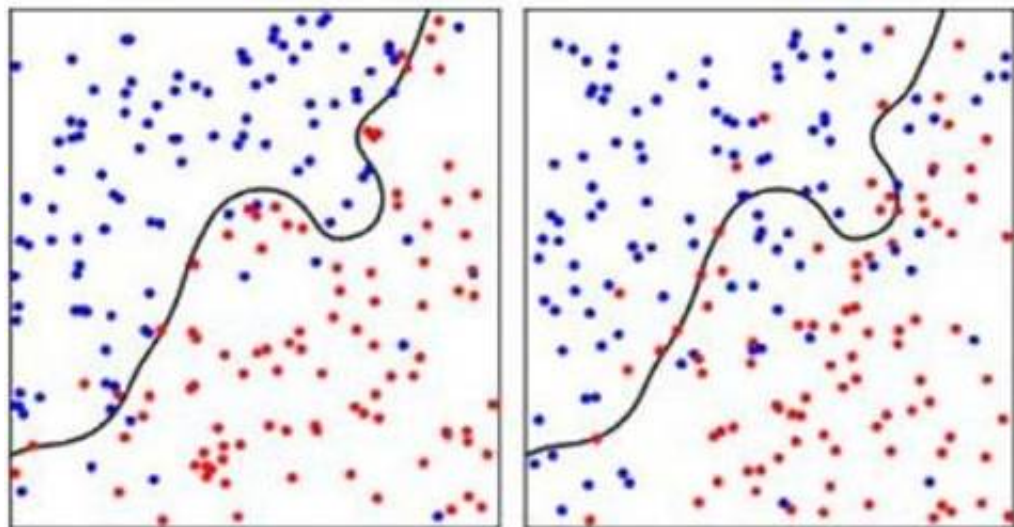
NEW QUESTION 25

Refer to the exhibit:

Model A

training data

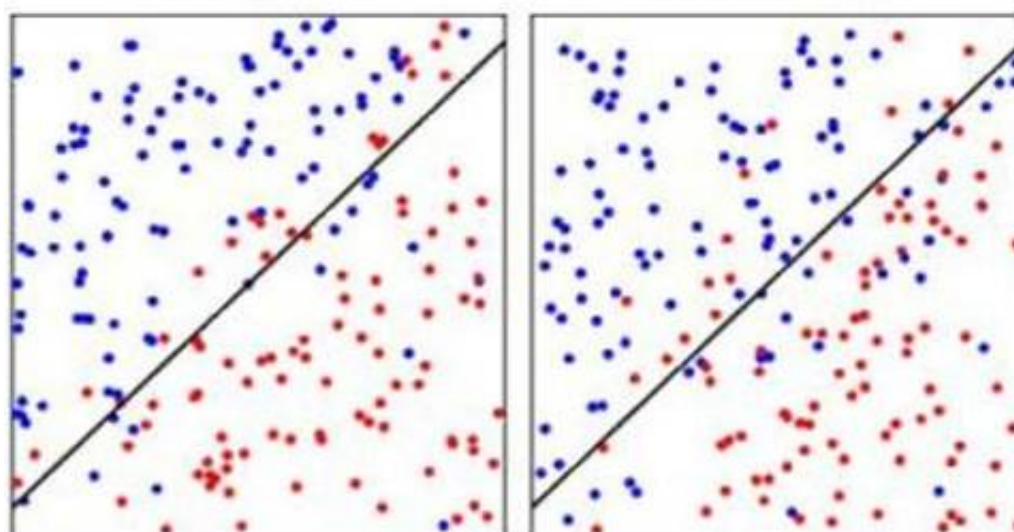
validation data



Model B

training data

validation data



The plots represent two models, A and B, being fit to the same two data sets, training and validation.

Model A is 90.5% accurate at distinguishing blue from red on the training data and 75.5% accurate at doing the same on validation data. Model B is 83% accurate at distinguishing blue from red on the training data and 78.3% accurate at doing the same on the validation data.

Which of the two models should be selected and why?

- A. Model
- B. It is more complex with a higher accuracy than model B on training data.
- C. Model
- D. It performs better on the boundary for the training data.
- E. Model
- F. It is more complex with a higher accuracy than model A on validation data.
- G. Model
- H. It is simpler with a higher accuracy than model A on validation data.

Answer: D

NEW QUESTION 29

One common approach for predicting rare events in the LOGISTIC procedure is to build a model that disproportionately over-represents those cases with an event occurring (e.g. a 50-50 event/non-event split).

What problem does this present?

- A. All parameter estimates are biased.
- B. Only the intercept estimate is biased.
- C. Only the non-intercept parameter estimates are biased.
- D. Sensitivity estimates are biased.

Answer: B

NEW QUESTION 31

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

Answer: C

NEW QUESTION 32

A marketing manager attempts to determine those customers most likely to purchase additional products as the result of a nation-wide marketing campaign. The manager possesses a historical dataset (CAMPAIGN) of a similar campaign from last year.

It has the following characteristics:

? Target variable Respond (0,1)

? Continuous predictor Income

? Categorical predictor Homeowner(Y,N)

Which SAS program performs this analysis?

- ☐ A.

```
proc logistic data=MYDIR.CAMPAIGN descending;
  class Homeowner;
  model Respond = Income Homeowner;
run;
```
- ☐ B.

```
proc logistic data = MYDIR.CAMPAIGN descending;
  by Homeowner;
  model Respond = Income Homeowner;
run;
```
- ☐ C.

```
proc logistic data = MYDIR.CAMPAIGN descending;
  model Respond = Income Homeowner;
run;
```
- ☐ D.

```
proc logistic data = MYDIR.CAMPAIGN descending;
  class Income Homeowner;
  model Respond = Income Homeowner;
run;
```

- A. Option A
 B. Option B
 C. Option C
 D. Option D

Answer: A

NEW QUESTION 36

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual A00-240 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the A00-240 Product From:

<https://www.2passeasy.com/dumps/A00-240/>

Money Back Guarantee

A00-240 Practice Exam Features:

- * A00-240 Questions and Answers Updated Frequently
- * A00-240 Practice Questions Verified by Expert Senior Certified Staff
- * A00-240 Most Realistic Questions that Guarantee you a Pass on Your FirstTry
- * A00-240 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year