

30-Apr-2023

What is the satiating price for a Constant-elasticity price response function?

Options :

6406531885611. ✘ Zero

6406531885612. ✓ Does not exist

6406531885613. ✘ Equal to the Market size

6406531885614. ✘ Can be anything

What does the term "Multicollinearity" refer to? (Select all that are applicable)

Options :

6406531885607. ✘ The dependent and independent variables are not-related

6406531885608. ✘ The dependent and independent variables are linearly related

6406531885609. ✘ The dependent variable is linearly related to another dependent variable

6406531885610. ✓ None of these

independent variables are correlated = multi-collinearity

What is the productive efficiency frontier?

Options :

6406531885656. ✘ It is an aspect of economic efficiency focussing on maximizing the output under given constraints

6406531885657. ✘ Productive efficiency frontier does not worry about optimal allocation, or choice of products

6406531885658. ✘ Effective usage of technology for allocating resources optimally

6406531885659. ✓ Consists of all combinations of outputs such that the production of one product cannot be increased without sacrificing the output of the other (without any change in technology)

Question Label : Multiple Choice Question

Let the objective function of the DEA problem for the DMU  $k$  be

Maximize:  $y_{1k}O_{1k} + y_{2k}O_{2k} + y_{3k}O_{3k} + \dots y_{Mk}O_{Mk}$

Which is not a constraint for the DEA problem

Options :

6406531885660. ✘ Normalizing constraint for the denominator

6406531885661. ✓ DMU  $k$  cannot have an efficiency of less than 1

6406531885662. ✘ Non-negativity constraint for the decision variables

6406531885663. ✘ None of these

In DEA, we make inefficient DMUs efficient by:

**Options :**

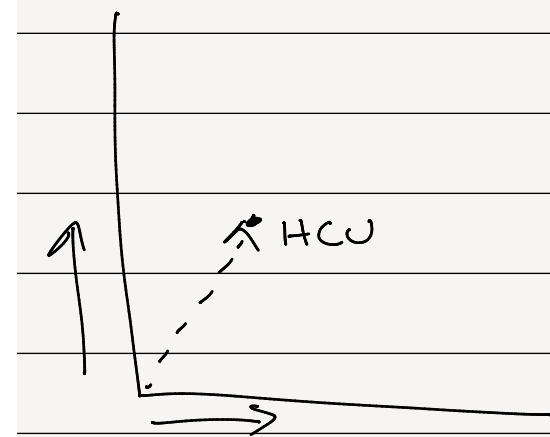
6406531885668. \* Moving horizontally towards the efficiency frontier

6406531885669. \* Moving vertically towards the efficiency frontier

6406531885670. \* By moving towards Hypothetical Composite Unit (HCU)

6406531885671. ✓ Includes Moving horizontally towards the efficiency frontier, Moving vertically towards the efficiency frontier & By moving towards Hypothetical Composite Unit (HCU)

6406531885672. \* Only Moving horizontally towards the efficiency frontier and Moving vertically towards the efficiency frontier



What kind of data is unsuitable for performing the conjoint analysis using the Statistical or Linear Regression Approach?

**Options :**

6406531885652. \* Consumer Choice Data is Ratings

6406531885653. ✓ Consumer Choice Data is Pairwise Comparison

6406531885654. ✓ Value of the attributes are continuous

6406531885655. \* Value of the product attributes are categorical

There are 7 business units and you are using the DEA to compare them. You solve the LP for business unit 2. You find from the constraint expression that business unit 3 has obtained an efficiency of 1 and business unit 6 has obtained an efficiency of 1 with the optimal weights of business unit 2. Which of the following statements is correct?

**Options :**

6406531885664. \* Business unit 3 may be inefficient

If efficiency = 1  
=> The DMU is efficient.

6406531885665. ✓ Business unit 3 will be efficient

6406531885666. \* Business unit 7 may be inefficient

6406531885667. ✓ Business unit 6 will be efficient

What kind of data is unsuitable for performing the conjoint analysis using the Linear Programming Approach?

**Options :**

6406531885674. ✓ Consumer Choice Data is Ratings

6406531885675. \* Consumer Choice Data is Pairwise Comparison

6406531885676. \* Values of the attributes are continuous

6406531885677. ✓ Values of the product attributes are categorical

A multiple linear regression model that uses Price, quantity and rating score is built to predict the sales in a company. In the multiple linear regression, it is found that the direct effect of price on sales is 0.3, the direct effect of quantity on sales is 0.2 and the direct effect of rating score on sales is 0.4. Moreover, it is known that the effect of price on quantity is 0.2, and the effect of price on rating score is 0.1. Then what is the total effect of price on sales?

**Response Type :** Numeric

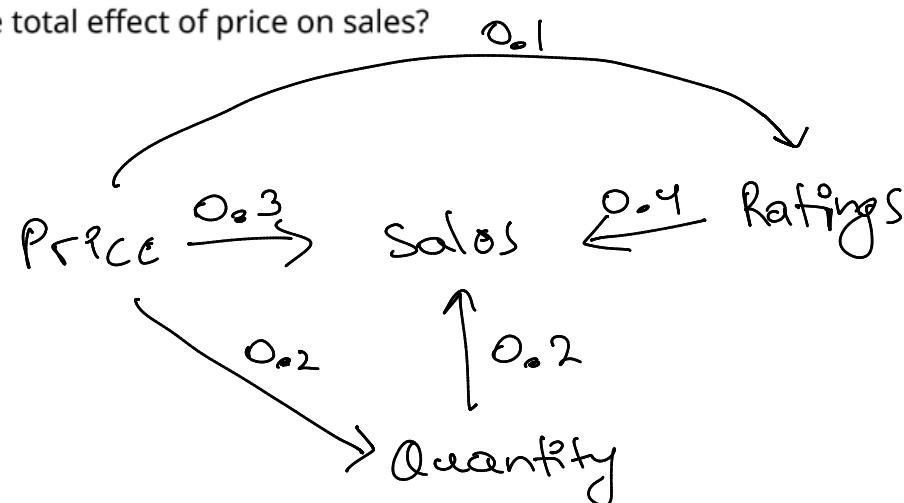
**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**



0.4 to 0.44

$$0.1 \times 0.4 + 0.2 \times 0.2 + 0.3 = \underline{\underline{0.38}}$$

There are 6 business units. There are two outputs and one input under consideration. You are solving the optimization problem for business unit 4 and find that the efficiency is 0.8. You find that the dual variables corresponding to the constraints of business units 3 and 5 are non-zero and the dual variables corresponding to the constraints of other units are zero. The dual variables corresponding to the constraints of business units 3 and 5 are 0.3 and 0.5 respectively. You are given the following table where sales and number of leads are the two outputs. What is the No. of leads in HCU 4?

$$\Rightarrow \text{No. of leads} = \frac{0.3 \times 12}{0.8} + \frac{0.5 \times 10}{0.8}$$

	Sales	No. of leads
DMU 3	12000	12
DMU 5	8000	10

0.3  
0.5

$$\Rightarrow \frac{3.6 + 5}{0.8} = \frac{8.6}{0.8} \Rightarrow \underline{\underline{10.75}}$$

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

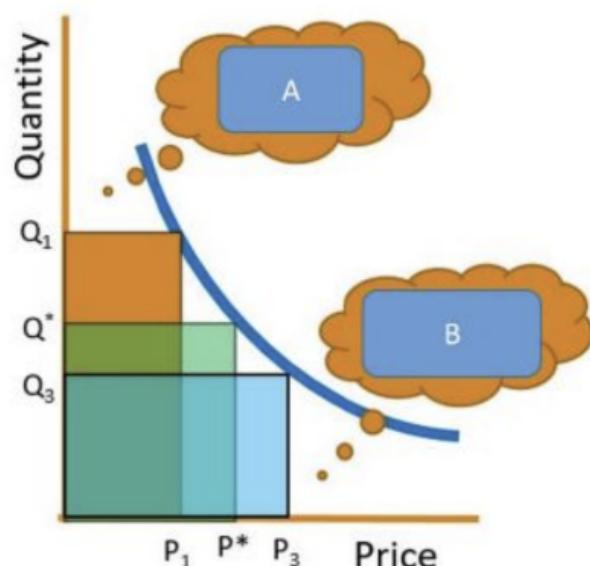
**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

13 to 14

For the figure given below, answer the subquestions.



What does region "A" represent?

**Options :**

6406531885615. \* Market Size

6406531885616. ✓ Latent demand

6406531885617. \* Consumer Surplus

6406531885618. \* None of these

What does region "B" represent?

**Options :**

6406531885619. \* Market Size

6406531885620. \* Latent demand

6406531885621. ✓ Consumer Surplus

6406531885622. \* None of these

Say the curve in the figure is modelled as a constant elasticity curve. If  $Q_1$  is 2400,  $Q_3$  is 1500,  $P_1$  is 100 and  $P_3$  is 200, then what is the elasticity of the curve?

$$\text{Elasticity} = \frac{\frac{1500 - 2400}{2400}}{\frac{200 - 100}{100}}$$

$$\Rightarrow 0.375$$

**NOTE:** If your answer is in decimal, enter it rounded to two decimal places. For example, if your answer is "10.256", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

2.28 to 2.36

A logistic regression model is fit to predict if a student will fail the BA course ( $Y=0$  pass,  $Y=1$  fail).

The results of the fitted model are given in the table below. Then answer the given subquestions.

(Note: For all questions in this comprehension, enter the answer as a numeric percentage value rounded to two decimal places without the % symbol. For example, if your answer is "10.256 %", enter it as "10.26")

Student ID	Actual "Pass" or "Fail"	Probability of Passing based on the logistic model
ABC101	Fail	0.79
ABC102	Pass	0.72
ABC103	Fail	0.45
ABC104	Fail	0.47
ABC105	Fail	0.73
ABC106	Pass	0.78

Threshold = 0.7

0  
0  
1  
1  
0  
0

0 = pass

1 = fail

Student Pass Course  
Predicted

Actual	Pass	2	0
Fail		2	2

Student Fail Course

Predicted	Fail	Pass
Actual	Fail	2
Pass	0	2

Considering a threshold of 0.7, how many "True Positives" are present in the confusion matrix if the aim is to predict if a student will pass the course?

**NOTE:** Enter the answer as a numeric percentage value rounded to the nearest integer without the % symbol.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

2

?

Considering a threshold of 0.7, how many "False Negatives" are present in the confusion matrix if the aim is to predict if a student will fail the course?

**NOTE:** Enter the answer as a numeric percentage value rounded to the nearest integer without the % symbol.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

2

?

TP = 2 (student pass course)

FN = 2 (student fail course)

FP = 2 (student fail course)

TN = 0 (student pass course)

Considering a threshold of 0.7, how many "False Positives" are present in the confusion matrix if the aim is to predict if a student will fail the course?

**NOTE:** Enter the answer as a numeric percentage value rounded to the nearest integer without the % symbol.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

2



Considering a threshold of 0.7, how many "True Negatives" are present in the confusion matrix if the aim is to predict if a student will pass the course?

**NOTE:** Enter the answer as a numeric percentage value rounded to the nearest integer without the % symbol.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

0



What is the accuracy of the built logistic model for predicting if a student will fail the course if a threshold of 0.7 is taken?

**NOTE:** Enter the answer as a numeric percentage value rounded to two decimal places without the % symbol. For example, if your answer is "10.256 %", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

33.00 to 34.00



What is the precision of the built logistic model for predicting if a student will pass the course if a threshold of 0.7 is taken?

**NOTE:** Enter the answer as a numeric percentage value rounded to the nearest integer without the % symbol.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

0



What is the recall of the built logistic model for predicting if a student will **fail** the course if a threshold of 0.7 is taken?

**NOTE:** Enter the answer as a numeric percentage value rounded to two decimal places without the % symbol. For example, if your answer is "10.256 %", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

50.00 to 50.50

?

•

Regression Statistics	
Multiple R	0.156176746
R Square	0.024391176
Adjusted R Square	-0.090386333
Standard Error	6.212632834
Observations	20

#### ANOVA

	df	SS	MS	F	Significance F
Regression	A1	16.40428549	8.202143	A5	0.810669095
Residual	A2	656.1457145	A4		
Total	A3	A6			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	55.8459249	9.257705483	6.032372	1.35E-05	36.31387366	75.37797614	36.31387366	75.37797614
Price	0.198655711	0.407861203	0.487067	0.632432	-0.661856209	1.059167631	-0.661856209	1.059167631
Income	0.004272622	0.006852519	0.623511	0.541225	-0.010184929	0.018730174	-0.010184929	0.018730174

#### Sub questions

$P < 2 = \text{Reject Null}$

Null = Model is **insignificant**  
 $\alpha = 0.05$  (unless mentioned different)

p-value of model

calc A1, A2, A3, A4, A5,  
A6

$$A3 = n-1$$

$$\Rightarrow n = \text{observations} = 20$$

$$\Rightarrow n-1 = 19$$

$$\underline{\underline{A3 = 19}}$$

$$A1 = k = (\text{Price, Income}) \\ = 2$$

$$\underline{\underline{A1 = 2}}$$

$$A2 = A3 - A1 = n - k - 1 \\ = 19 - 2 \\ = \underline{\underline{17}}$$

$$A4 = \frac{656.145}{A2} = \frac{656.145}{17} = \underline{\underline{38.59}}$$

$$A5 = \frac{8.202}{A4} = \frac{8.202}{38.59} = \underline{\underline{0.212}}$$

$$A6 = 16.404 + 656.145$$

$$\Rightarrow \underline{\underline{672.549}}$$

Which of the following inferences is/are correct? (Choose all that are applicable)

**Options :**

6406531885630. ✘ The model is significant

6406531885631. ✓ The model is insignificant

6406531885632. ✘ The intercept is insignificant

6406531885633. ✓ The intercept is significant

6406531885634. ✘ Price, the dependent variable, is insignificant

6406531885635. ✘ Price, the dependent variable, is significant

6406531885636. ✘ None of these

Price, Income are  
independent and insignificant

The owner of a cycle shop is curious to know about his customer base. Hence, he collects data on the total sales made to different age groups in the year 2022. This data is given in the table below. He hires you as a data analyst and asks for your opinion. Then answer the given subquestions.

Age group	Actual number of cycles sold	
less than 10 years	30	37
11 to 15 years	50	37
16 to 20 years	90	37
21 to 24 years	40	37
25 to 30 years	12	37
31 to 40 years	25	37
greater than 41 years	12	37

$$\text{Expected} = \frac{\text{Total no. of cycles}}{\text{Total no. of obsv}} = \frac{259}{7} = 37$$

Based on the given table, you think the data is **normally distributed** such that the buckets split the distribution into areas that have equal probability. Then what is the probability of selling a cycle in any given bucket?

**NOTE:** If your answer is in decimal, enter it rounded to two decimal places. For example, if your answer is "10.256", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

$$P = \frac{37}{259} = 0.\underline{\underline{14}}$$

**Possible Answers :**

0.13 to 0.15

Based on the given table, you think the data is **normally distributed** such that the buckets split the distribution into areas that have equal probability. Then what is the expected number of cycles that will be sold to the age group of "25 to 30 years"?

**NOTE:** If your answer is in decimal, enter it rounded to two decimal places. For example, if your answer is "10.256", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

36 to 38

37

What is the value of the computed test statistic?

**NOTE:** If your answer is in decimal, enter it rounded to two decimal places. For example, if your answer is "10.256", enter it as "10.26"

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

144 to 148

$$(7)^2 + (13)^2 + (43)^2 + (3)^2 + (25)^2 + \\ (12)^2 + (25)^2 \\ 37$$

293.70

How many degrees of freedom does the test have?

**NOTE:** Enter your answer to the nearest integer.

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

7

?

Given that the "Null Hypothesis" is rejected at the significance level of 0.5, then what will be the conclusion at a significance level of 0.1?

**Options :**

6406531885642. \* Cannot say, insufficient data

6406531885643. ✓ Null Hypothesis is rejected

?

6406531885644. \* Null Hypothesis is not rejected

Dec-11-2022

Select which of the following are discrete data?

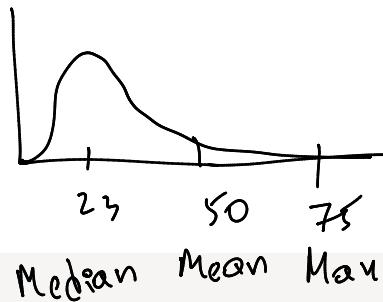
**Options :**

- 6406531502815. ✓ Number of courses you are crediting in this IIT Madras BSc Degree program
- 6406531502816. ✗ Average weight of your batch in this BSc program
- 6406531502817. ✓ Size of your batch in this BSc program
- 6406531502818. ✗ Average age of your batch in this BSc program

You are given a data of size 1000, which is centred at 50 and has a variance of 25. The median, mode, max and min values are 23, 22, 75 and 1 respectively. Then

**Options :**

- 6406531502819. ✓ The data is right-tailed
- 6406531502820. ✗ The data is left-tailed
- 6406531502821. ✗ The data is symmetric
- 6406531502822. ✗ Cannot say



Let  $Y_{jk}$  be the weight of the output  $O_{jk}$  and  $X_{jk}$  be the weight of the input  $I_{jk}$ . Which of the following is the correct objective function of the DEA if we solve it as LP?

**Options :**

- 6406531502859. ✗ Max  $Y_{jk} * I_{jk}$
- 6406531502860. ✓ Max  $Y_{jk} * O_{jk}$
- 6406531502861. ✗ Max  $(1/Y_{jk}) * I_{jk}$
- 6406531502862. ✗ Max  $- (Y_{jk} * I_{jk})$

In DEA, for calculating the weights of efficiency (weighted outputs/weighted inputs), when can the Linear Programming model can be used?

**Options :**

- 6406531502867. ✗ After converting the ratio into the linear objective function
- 6406531502868. ✗ After normalizing the denominator
- 6406531502869. ✗ By setting a constraint on the efficiency of all DMUs to be lesser than or equal to 1
- 6406531502870. ✓ All of these

Match the type of conjoint analysis to its method of administration

Conjoint Analysis Type	Method of administration
1. Choice-Based Conjoint Analysis	A. Questions are dynamically decided based on customer response
2. Adaptive Conjoint Analysis	B. All the options are presented to the customer, and their preference is sought
3. Full-profile Conjoint Analysis	C. Customer chooses their most preferred full-profile product among 3-4 options provided
4. Menu-based Conjoint Analysis	D. List of attributes and their prices are shown to customers, and they choose what they want in the ideal product.

Options :

6406531502871. ✓ 1-C, 2-A, 3-B, 4-D

6406531502872. ✗ 1-A, 2-C, 3-B, 4-D

6406531502873. ✗ 1-B, 2-C, 3-A, 4-D

6406531502874. ✗ 1-C, 2-B, 3-C, 4-D

6406531502875. ✗ None of these

A hypothesis was tested at 5% significance level and *p-value* turned out to be 0.054. What will be your correct decision?

$$p < \alpha = \text{Reject Null}$$

Options :

6406531502888. ✗ Reject the null hypothesis

6406531502889. ✓ Fail to reject the null hypothesis

6406531502890. ✗ Sufficient evidence to accept the null hypothesis

6406531502891. ✗ both Fail to reject the null hypothesis and Sufficient evidence to accept the null hypothesis are correct

In Multiple Linear Regression, if the explanatory variables are highly correlated, then that phenomenon is called

Options :

6406531502892. ✗ Normality

6406531502893. ✗ Singularity

6406531502894. ✓ Collinearity

6406531502895. ✗ Variation Inflation

In Regression, Marginal slope and partial slope coincide, if the explanatory variables are dependent. Is the statement True or False?

Options :

6406531502823. ✗ TRUE

6406531502824. ✓ FALSE

2

VIF value less than 1 indicates that

Options :

6406531502825. ✗ Multicollinearity is very high

6406531502826. ✗ Multicollinearity is very low

6406531502827. ✓ Calculation error

6406531502828. ✗ None of these

$$VIF = \frac{1}{1 - R^2}$$

$$R^2 \text{ max value} = 1$$

$$R^2 \text{ min value} = 0$$

Logistic Regression can create a non-linear decision boundary.

**Options :**

6406531502845. ✓ TRUE

6406531502846. ✗ FALSE

B

not B

A C

From the Table given below, what is the number of True Negatives for Class B?

		Predicted Class		
		A	B	C
Actual Class	A	100	0	10
	B	10	80	10
	C	30	0	70

2  
-

**Options :**

6406531502852. ✓ 0

6406531502853. ✗ 10

6406531502854. ✗ 20

6406531502855. ✗ 110

What is productive efficiency?

**Options :**

6406531502856. ✓ It is an aspect of economic efficiency focussing on maximizing the output under given constraints (without worrying about optimal allocation, or choice of products)

6406531502857. ✗ Effective usage of technology for maximizing the profitability

6406531502858. ✗ Consists of all combinations of outputs such that the production of one product cannot be increased without sacrificing the output of the other (without any change in technology)

The object of the conjoint analysis problem will be \_?

**Options :**

6406531502878. ✗ Minimize the distance between the attribute combination for the ideal product and the customer preferences

6406531502879. ✗ Minimize the distance between the attribute combination for the ideal product and the attribute combination for each possible variant

6406531502880. ✗ Minimize the distance between the possible customer preferences

6406531502881. ✓ None of these

You are solving a regression problem with 7 explanatory variables. The data has 50 observations and the R-square value was found to be 0.85. Then answer the given subquestions

**Sub questions**

**Question Number : 290 Question Id : 640653451578 Question Type : SA Calculator : None**

**Response Time : N.A Think Time : N.A Minimum Instruction Time : 0**

**Correct Marks : 2**

Question Label : Short Answer Question

What is the value of adjusted R-squared (Round off to three decimal values)?

**Response Type : Numeric**

**Evaluation Required For SA : Yes**

**Show Word Count : Yes**

**Answers Type : Range**

**Text Areas : PlainText**

**Possible Answers :**

0.825 to 0.830

$$n = 50$$

$$k = 7$$

$$R^2 = 0.85$$

$$R_{\text{Adj}} = 1 - \frac{(1-R^2)(n-1)}{n-k-1}$$

$$\Rightarrow 1 - \frac{(0.15)(49)}{42}$$

$$\Rightarrow \underline{\underline{0.825}}$$

You are adding one more explanatory variable to the dataset making a total of 8 variables. The

new R squared value is 0.862 adjusted R square value increases to 0.864. What does it signify?

**Options :**

6406531502830. \* The new variable does not improve the model

6406531502831. \* The new variable alone has high explanatory power

6406531502832. ✓ Adjusted R squared can never be greater than R squared. Calculation error ✓

6406531502833. \* None of these

You are given the following contingency table based on a sample data with people belonging to two cities (City A, City B) and their brand preferences. You perform a chi-squared test of independence to make inferences about the population from this sample. Then answer the given subquestions.

	Brand A	Brand B	Brand C	Brand D
City A	142	147	227	100
City B	75	114	103	85

From the given contingency table, find the expected frequency of people belonging to City A preferring brand B?

**Response Type : Numeric**

**Evaluation Required For SA : Yes**

**Show Word Count : Yes**

**Answers Type : Range**

**Text Areas : PlainText**

**Possible Answers :**

158 to 164

What is the calculated value of chi-squared?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

13 to 19

At the significance level 0.01, chi-squared tabular value is 11.34. What do you conclude?

**Options :**

6406531502836. ✓ Reject the null hypothesis and conclude that the categorical variables are not independent

6406531502837. ✗ Fail to reject the null hypothesis and conclude that the categorical variables are not independent

6406531502838. ✗ Fail to reject the null hypothesis and conclude that the categorical variables are independent

6406531502839. ✗ Reject the null hypothesis and conclude that the categorical variables are independent

Suppose a factory manufactures products on three machines A, B and C. Suppose 35% of total output comes from machine A, 40% of total output comes from machine B and 25% of total output comes from machine C. From the past data, it is known that 5% of products by machine A are defectives, 10% of products by machine B are defectives and 12% of products by machine C are defectives. What is the probability that the product has come from machine C given that it is a defective?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

0.31 to 0.37

Select the correct option from below:

**Options :**

6406531502841. ✗ For inelastic product demand ( $\epsilon < 1$ ) the revenue can be increased by setting price close to zero.

6406531502842. ✓ For elastic product demand ( $\epsilon > 1$ ) the revenue can only be increased by setting price close to zero.

6406531502843. ✓ For inelastic product demand ( $\epsilon < 1$ ) the revenue can be increased by simply increasing the prices

6406531502844. ✗ For elastic product demand ( $\epsilon > 1$ ) the revenue can only be increased by simply increasing the prices

There are 7 business units and you are using the DEA to compare them. You solve the LP for business unit 5. You find from the constraint expression that business unit 4 has obtained an efficiency of 1 and business unit 7 has obtained an efficiency of 1 with the optimal weights of business unit 5. Which of the following statements is correct?

**Options :**

6406531502863. ✗ Business unit 4 may be inefficient

6406531502864. ✓ Business unit 4 will be efficient

6406531502865. ✗ Business unit 7 may be inefficient

6406531502866. ✓ Business unit 7 will be efficient

Using the data given in Table 1, answer the given subquestions

Table 1

Classification confusion matrix		
	Predicted Class	
Actual Class	1	0
1	187	19
0	63	31

What is the Precision for predicting 1's? (enter your value in percentage without the percentage symbol, rounded to two decimal places. Eg: If your answer is 0.12345, enter the answer as "12.3")

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

$$\frac{TP}{TP + FP} = \text{Precision}$$

**Text Areas :** PlainText

$$\frac{187}{187+63} = \frac{187}{250} = 0.748$$

**Possible Answers :**

74.60 to 75.00

What is the Recall for predicting 1's? (enter your value in percentage without the percentage symbol, rounded to two decimal places. Eg: If your answer is 0.12345, enter the answer as "12.3")

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

90.50 to 90.90

$$\frac{TP}{TP+FN} = \frac{187}{187+19} = 0.9077$$

What is the accuracy of the predictor? (enter your value in percentage without the percentage symbol, rounded to two decimal places. Eg: If your answer is 0.12345, enter the answer as "12.3")

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

0.71 to 0.74

How many False Negatives is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

19

How many True Negatives is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

31

A manufacturer is going to make a product, and he is exploring possible variants (V1, V2, V3) for it.

The variants of the products are determined by 5 attributes (A1, A2, A3, A4, A5). Pair-wise

preference data will be collected from customers to perform the analysis. Using this information,

answer the given subquestions.

How many pairs (for comparisons) will be generated in this problem?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

15

5 x 3

Then what is the size of the set of options on which the preference judgements are made?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

3

?

✓

The BSc Team wanted to see if the live sessions get a uniform response across different terms.

Accordingly, the BA course was taken as a pilot, and the total number of students who participated in the live sessions over the past three terms was obtained (data in the table below). Given this information, answer the given subquestions.

Observed

Term	Total Number of Participants across Live Sessions
Jan-2022	12
May-2022	14
Sep-2022	12

$$\text{Expected} = \frac{12+14+12}{3}$$

12.67

12.67

12.67

What statistical test needs to be carried out to see if the attendance is uniformly distributed across the terms?

**Options :**

6406531502882. ✓ Chi-squared Goodness of Fit Test

6406531502883. \* Chi-Squared Test of Independence

6406531502884. \* Chi-Square Test of variance

6406531502885. \* Test of Means (Z-Test or T-test)

What is the value of the computed test statistic (Round your answer to one decimal place)?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

12.4 to 12.6

J

$$(0.67)^2 + (1.33)^2 + (0.67)^2$$

12.67

=>

0.21

If the attendance is indeed uniformly distributed, then how many participants will you expect to see across live sessions in any given term? (round your answer to two decimal places)

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

$$\text{Expected} = 12.67$$

**Answers Type :** Set

**Text Areas :** PlainText

**Possible Answers :**

12.66

12.67

Aug 7 2022

A leading European two-wheeler manufacturer is trying to build an ideal scooter for the Indian market. Which of the following tools will help him understand: how important the attributes such as Engine Capacity, LED Lights, Fuel Efficiency, and Smart Connect Technology are to the customers?

**Options :**

6406531183921. ✘ Data Envelopment Analysis

6406531183922. ✓ Conjoint Analysis

6406531183923. ✘ Chi-Square test

6406531183924. ✘ Both Data Envelopment Analysis & Chi-Square test

The objective function of the linear programming model using pair-wise judgments:

**Options :**

6406531183938. ✘ Maximize the distance from the ideal point

6406531183939. ✘ Minimize the distance from the ideal point

6406531183940. ✓ Minimize the poorness of fit

6406531183941. ✘ Both Maximize the distance from the ideal point & Minimize the poorness of fit

6406531183942. ✘ Both Minimize the distance from the ideal point & Minimize the poorness of fit

What is the format of data needed for performing the conjoint analysis using the Statistical or Linear Regression Approach?

**Options :**

6406531183943. ✓ Consumer Choice Data is Ratings

6406531183944. ✘ Consumer Choice Data is Pairwise Comparison

6406531183945. ✘ Value of the attributes are continuous

6406531183946. ✓ Value of the product attributes are categorical

Suppose you conduct a chi-squared test of independence on the categorical variables cities and brand preferences at the significance level 0.1. You obtain a p-value of 0.12. What will you conclude?

**Options :**

6406531183900. ✘ Reject the null hypothesis and conclude that the categorical variables are independent

6406531183901. ✘ Reject the null hypothesis and conclude that the categorical variables are not independent

6406531183902. ✓ Fail to reject the null hypothesis and conclude that the categorical variables are independent

6406531183903. ✘ Fail to reject the null hypothesis and conclude that the categorical variables are not independent

$P < \alpha = \text{Reject Null}$

Which of the following is not a form of conjoint analysis?

**Options :**

6406531183929. ✘ Choice based

6406531183930. ✘ Adaptive based

6406531183931. ✘ Full profile

6406531183932. ✓ Selective attribute type

6406531183933. ✘ Menu-based

In a conjoint problem with 4 products and 2 attributes, how many pair-wise preferences are

possible?

**Options :**

6406531183934. ✘ 8

6406531183935. ✓ 6

6406531183936. ✘ 16

6406531183937. ✘ 12

$\frac{n(n-1)}{2}$  only applicable for 2 attributes.

$$\frac{4(3)}{2} = \underline{\underline{6}}$$

Say a data is distributed as "Normal" with a right tail. If it is compared with an "Symmetric Normal" distribution, then which of the following states is/ are true (choose all that is applicable)

**Options :**

6406531183888. ✘ In a "P-P plot", the data will fall on a line which is indicate at 45-degree to the X-axis

6406531183889. ✓ In a "Q-Q plot", the data will fall on a line which is indicate at 45-degree to the

X-axis

6406531183890. ✓ In a "P-P plot", the data will not entirely fall on a line which is indicate at 45-degree to the X-axis

6406531183891. ✘ In a "Q-Q plot", the data will not entirely fall on a line which is indicate at 45-degree to the X-axis

6406531183892. ✘ Cannot use P-P plot or Q-Q plot as the assumed distribution is discrete

There are 7 business units and you are using the DEA to compare them. You solve the LP for business unit 5. You find from the constraint expression that the business unit 3 has obtained an efficiency of 1 and the business unit 7 has obtained an efficiency of 1 with the optimal weights of business unit 5. Which of the following statements are correct?

**Options :**

6406531183905. ✘ Business unit 3 may be inefficient

6406531183906. ✓ Business unit 3 will be efficient

6406531183907. ✘ Business unit 7 may be inefficient

6406531183908. ✓ Business unit 7 will be efficient

There are 4 business units. Using the DEA, you solve the LP for all the four business units and find the efficiencies for these units. The efficiency is denoted by E. For these units,  $E_1 = 1$ ,  $E_2 = 1$ ,  $E_3 = 1$ ,  $E_4 = 0.91$ . Which of these units are efficient?

**Options :**

6406531183912. ✓ 1

6406531183913. ✓ 2

6406531183914. ✓ 3

6406531183915. ✗ 4

What is the format of data needed for performing the conjoint analysis using the Linear Programming Approach?

**Options :**

6406531183925. ✗ Consumer Choice Data is Ratings

6406531183926. ✓ Consumer Choice Data is Pairwise Comparison

6406531183927. ✓ Values of the attributes are continuous

6406531183928. ✗ Values of the product attributes are categorical

You have estimated the demand to follow the following relationship:  $D(p) = 100 - 10p$ . Now, you intend to maximize the revenue  $R(p) = D(p)p$ . You find the first derivative of  $R(p)$  with respect to  $p$ , equate it to 0 and find  $p^*$ . What is the value of  $p^*$ ?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

$$\frac{D_0}{2m} = \frac{100}{2(10)} = \underline{\underline{5}}$$

Direct formula.

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

4.9 to 5.1

In a multiple linear regression with 3 explanatory variables, you find that R-squared value is 0.6. The number of observations is 25. What is the value of adjusted R-squared?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

0.5 to 0.6

$$R_{Adj} = 1 - \frac{(1-0.6)(24)}{21}$$

$$\Rightarrow 1 - \frac{(0.4)(24)}{21}$$

$$\Rightarrow \underline{\underline{0.54}}$$

$$n = 25$$

$$k = 3$$

$$R^2 = 0.6$$

You are conducting a multiple linear regression with sales as the dependent variable. Price, quantity and rating score are the independent variables. In the multiple linear regression, you find that the direct effect of price on sales is 0.4, the direct effect of quantity on sales is 0.2 and the direct effect of rating score on sales is 0.4. And, you also know that the effect of price on quantity

is 0.2 and the effect of price on rating score is 0.1. What is the total effect of price on sales?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

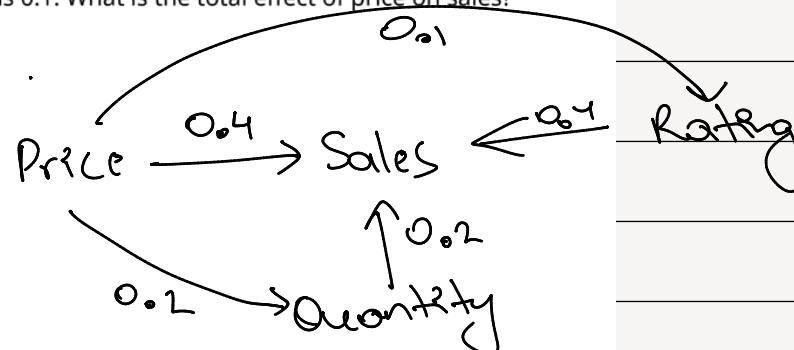
**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

0.46 to 0.50



$$\text{Total Effect} = 0.4 + 0.2 \times 0.2 + 0.1 \times 0.4 \\ \Rightarrow \underline{\underline{0.48}}$$

You solve the primal of a linear program with maximization objective, three decision variables and two constraints of the less than or equal to type. Non-negativity restrictions apply on the decision variables. After solving the linear program, you find that the first constraint is binding ( $\text{lhs} = \text{rhs}$ ) and the second constraint is binding ( $\text{lhs} = \text{rhs}$ ). Which of the following statements are correct?

**Options :**

6406531183916. ✘ There are three decision variables in the dual

6406531183917. ✓ The dual variable corresponding to the first constraint is non-zero

6406531183918. ✘ There are four decision variables in the dual

6406531183919. ✘ The dual variable corresponding to the second constraint is zero

There are 6 business units. There are two outputs and one input under consideration. You are solving the optimization problem for business unit 3 and find that the efficiency is 0.7. You find that the dual variables corresponding to the constraints of business units 4 and 5 are non-zero and the dual variables corresponding to the constraints of other units are zero. The dual variables corresponding to the constraints of business units 4 and 5 are 0.3 and 0.4 respectively. You are given the following table where sales and number of leads are the two outputs.

	Sales	Number of leads
Business unit 4	8500	10
Business unit 5	8000	12

$$\text{Sales HCU} =$$

$$\frac{8500 \times 0.3}{0.7} + \frac{8000 \times 0.4}{0.7}$$

$$\Rightarrow \underline{\underline{8214.29}}$$

What is the sales in HCU 3?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

8210 to 8220

You are conducting a multiple linear regression with sales as the dependent variable. Price, quantity and rating score are the independent variables. In order to calculate the VIF for the variable quantity, you implement a linear regression with quantity as the dependent variable and other variables as independent variables and obtain R-squared of 0.4. What is the VIF for the

variable quantity?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

1.6 to 1.7

$$VIF = \frac{1}{1 - R^2}$$

$$\Rightarrow \frac{1}{1 - 0.4}$$

$$\Rightarrow \underline{\underline{1.66667}}$$

Ms. X is working is with the data given in Table-1 below. Using this information answer the given subquestions.

Product	Sales of a Product in a City for a given Year								
	City-1 (1990)	City-2 (1990)	City-3 (1990)	City-1 (1991)	City-2 (1991)	City-3 (1991)	City-1 (1992)	City-2 (1992)	City-3 (1992)
A	100	90	250	120	50	120	140	20	500
B	145	300	500	175	250	250	195	230	1000
C	90	180	30	100	110	15	110	95	58
D	130	220	132	140	200	61	150	180	270

Table-1

If the message that X wants to convey is focused on the market share of each product in the year 1992 in each city, then which visualization will be most appropriate?

**Options :**

6406531183873. \* Line graph for each product's sales in the year 1992 for each city

6406531183874. \* Line graph for each city's sale in the year 1992 for each product

6406531183875. \* Bar graph for each product's sale in the year 1992 for each city

6406531183876. \* Bar graph for each city's sales in the year 1992 for each product

6406531183877. \* Pie chart for each product's sales in the year 1992 for each city

6406531183878. ✓ Pie chart for each city's sales in the year 1992

6406531183879. \* The table with only sales of the year 1992 for the three cities for each product

If Ms. X has created the below visualization (Figure-1) for the data of product-A in Table-1. **What is**

**the primary message** that is conveyed through this?

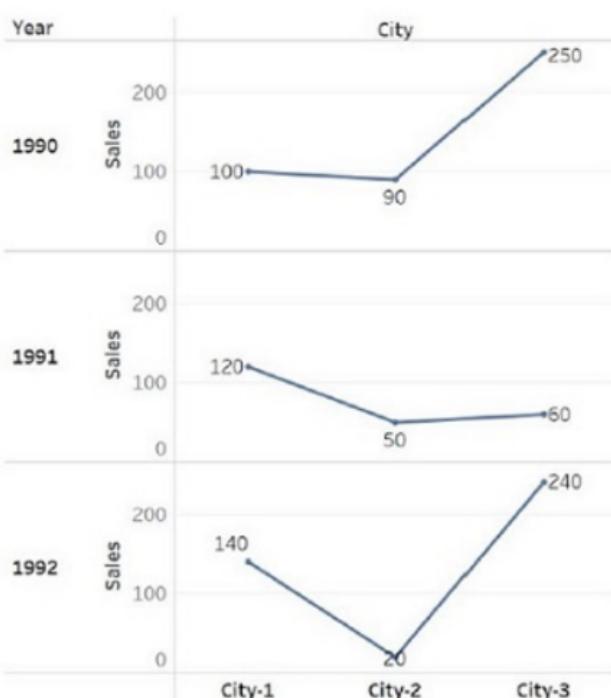


Figure-1

**Options :**

6406531183880. \* City-3 has the highest sales for the product-A in the current year

6406531183881. \* The sales trend of Product-A is constantly fluctuating over the years

6406531183882. ✓ The sales trend of Product-A is drastically different across cities in any given year

6406531183883. \* The sales trend of product-A in city-2 is lowest among all the cities in any given year

6406531183884. \* None of these

Say you want to see if the distribution of sales of Product-D in the Table-1 follows a uniform distribution within the range of 0 to 300 when split into bins as specified in Table-2. Then what is the expected frequency in any given bin (round your answer to two decimal places)?

D	130	220	132	140	200	61	150	180	270
---	-----	-----	-----	-----	-----	----	-----	-----	-----

Bin Number	Bin Range
Bin-1	Sales value less than or equal to 50
Bin-2	Sales value greater than 50 but less than or equal to 100
Bin-3	Sales value greater than 100 but less than or equal to 150
Bin-4	Sales value greater than 150 but less than or equal to 200
Bin-5	Sales value greater than 200 but less than or equal to 250
Bin-6	Sales value greater than 250 but less than or equal to 300

Table-2

obs	Exp
0	1.5
1	1.5
4	1.5
2	1.5
8	1.5
1	1.5

Expected value =  $\frac{9}{6} = 1.5$

1.5

If a Chi-Square Goodness-Of-Fit Test to check if the data for Product-D in Table-1 follows a Uniform Distribution with bins as specified in Table-2, then what is the value of the Test statistic (round your answer to two decimal places)?

$$\text{Hint: Chi-square} = \sum_k \frac{(observed_k - Expected_k)^2}{Expected_k}$$

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

6.20 to 6.40

$$(1.5)^2 + (0.5)^2 + (2.5)^2 + (-0.5)^2 + (0.5)^2 + (2.5)^2$$

1.5

$$7 \times 6.33$$

6.33

A Chi-squared Goodness-Of-Fit test with the bins as specified in Table-2 is going to be carried out to check if the data on sales (whole data in Table-1) is normal or not. Then what is the degrees of freedom for the test?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

3

$$n = 6$$
$$k = (\mu, \sigma^2) = 2$$
$$6 - 2 - 1 = \underline{\underline{3}}$$

A logistic model has been fit for a data set with the goal to predict the positive class (Y=1). The confusion matrix obtained for this model on the test dataset is provided in Table-3. Using this, answer the given subquestions.

		Predictions	
		1	0
Actual	1	23	77
	0	45	11

Table-3

How many "True Positives" is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

23

How many "False Positives" is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

45

How many "True Negatives" is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

11

How many "False Negatives" is the model predicting?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

77

What is the accuracy for the model (enter only the numerical value in percentage, without the "%" symbol and round it to two decimal places)?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

21.00 to 22.00

$$\frac{TP + TN}{TP + TN + FP + FN}$$

2)  $\underline{\underline{21.79}}$

What is the recall for the model when predicting the positive class (enter only the numerical value in percentage, without the "%" symbol and round it to two decimal places)?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

22.00 to 23.00

$$\frac{TP}{TP + FN}$$

$\underline{\underline{23}}$

What is the precision for the model when predicting the negative class (enter only the numerical value in percentage, without the "%" symbol and round it to two decimal places)?

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

12.0 to 12.80

$$\frac{TN}{TN + FN}$$

$\underline{\underline{12.5}}$

May - 2023

In an ideal scenario, if a distribution has "Negative Skewness" then,

[Marks: 1]

Select an Option

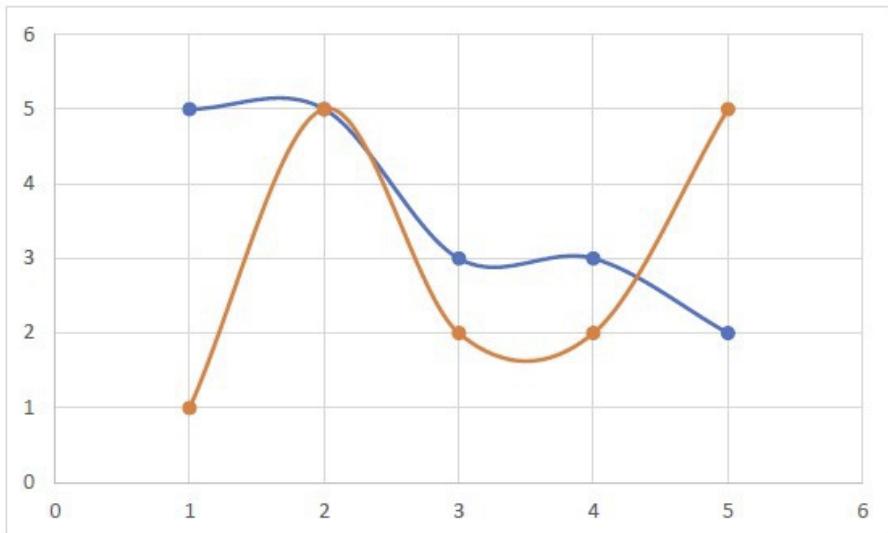
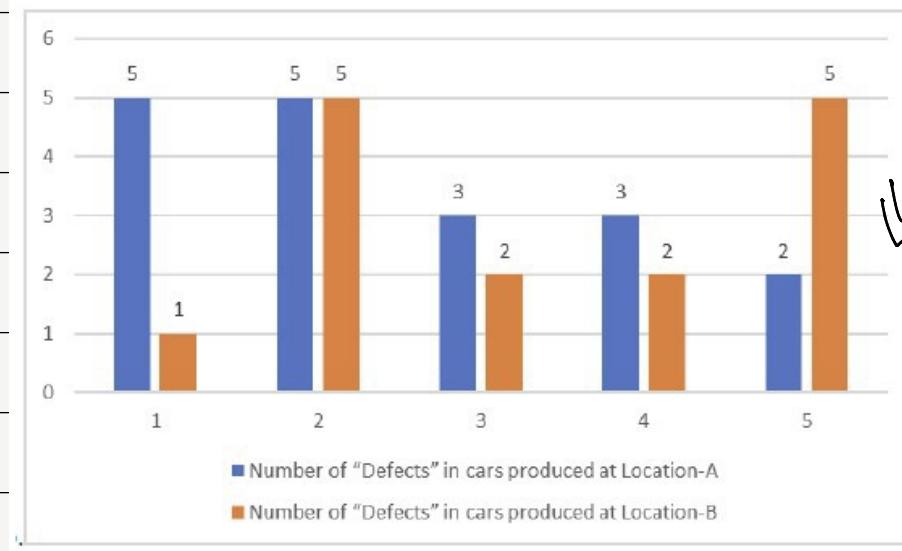
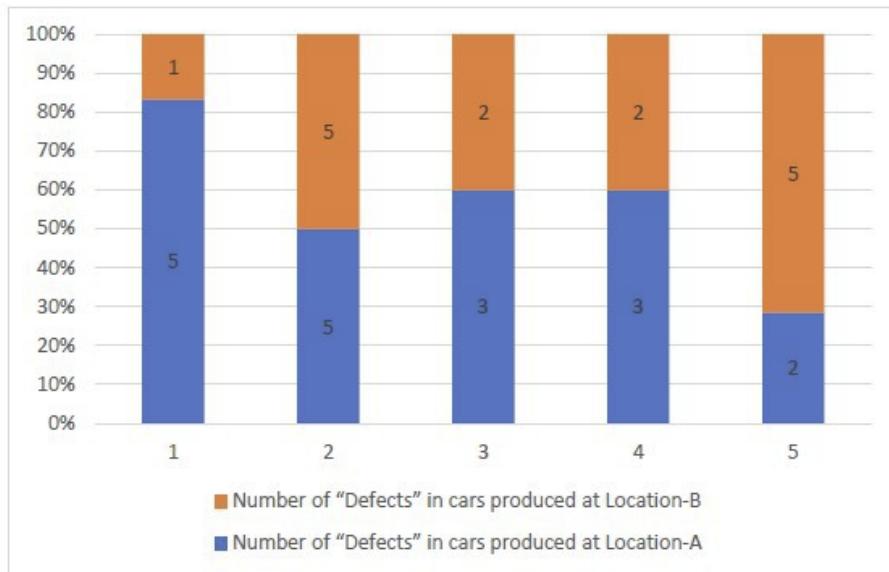
- Mode > Median > Mean ✓
- Mode < Median < Mean
- Mode = Median = Mean
- None of these

A company produces a car in two locations "A" and "B" using the same manufacturing process. A total of 20 cars in each location were taken and the number of defects in each car was computed. It has been established that the maximum number of defects per car is 5. Given this data in Table-1, answer the given sub-questions.

Number of Defects	Number of cars produced at Location-A with the specified number of defects	Number of cars produced at Location-B with the specified number of defects
1	5	1
2	5	5
3	3	2
4	3	2
5	2	5

Table- 1

If the focus is on seeing the distribution of defects (presented in Table-1) across the two locations, then which among the following graphs will be best suited? ( Note: While choosing an answer to this question, please do not worry about colour reproduction or other aesthetics. Make a choice only based on the concepts of visualization theory )



Any of these

A company produces a car in two locations "A" and "B" using the same manufacturing process. A total of 20 cars in each location were taken and the number of defects in each car was computed. It has been established that the maximum number of defects per car is 5. Given this data in Table-1, answer the given sub-questions.

Number of Defects	Number of cars produced at Location-A with the specified number of defects	Number of cars produced at Location-B with the specified number of defects
1	5	1
2	5	5
3	3	2
4	3	2
5	2	5

Table- 1

Aus Range  
2 to 3

	Defects	A	B	Total
	1	5	1	6
	2	5	5	10
	3	3	2	5
	4	3	2	5
	5	2	5	7
	Total	18	15	33

If the aim is to determine if the defect occurrence is independent of location, then how many cars would you expect to have "4" defects in Location A? (Round your answer to two decimal places. Eg: If your answer is 10.256, then round it to 10.26)

$$\text{Expectation of 4 defects in loc A} = \frac{5 \times 18}{33} = \underline{\underline{2.7272}}$$

If the aim is to determine if the defect occurrence is independent of location, then how many cars would you expect to have "1" defect in Location B? (Round your answer to two decimal places. Eg: If your answer is 10.256, then round it to 10.26)

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Equal

**Text Areas :** PlainText

**Possible Answers :**

3

$$\text{Expectation of 1 defect in loc B} = \frac{6 \times 15}{33} = \underline{\underline{2.7272}}$$

If the aim is to find if the number of defects is independent of location, then what is the value for the test statistic? (Round your answer to two decimal places. Eg: If your answer is 10.256, then round it to 10.26)

Defects	A	B
1		
2		
3		
4		
5		

**Response Type :** Numeric

**Evaluation Required For SA :** Yes

**Show Word Count :** Yes

**Answers Type :** Range

**Text Areas :** PlainText

**Possible Answers :**

5.40 to 5.80