

S.NO - 5036

SET_A

Name of Course: CBCS B.A. (Prog.) SEC, Semester VI

Name of the Paper: Basic Computational Techniques for Data Analysis

Unique Paper Code: 62273601

Maximum Marks: 50 Marks

Duration: 3 hours

Maximum marks: 50

Instructions for Candidates:

- Write your Roll number on the top immediately on receipt of the question paper.
- Attempt any four questions.
- Each question carries 12.5 marks each.
- Answers may be written in Hindi or English, but same medium will be used throughout the paper.
- Use of simple calculator is allowed.

अभ्यर्थियों के लिए दिशा निर्देश:

- प्रश्न पत्र मिलते ही सबसे ऊपर तुरंत अपना अनुक्रमांक लिखें।
- किन्हीं 4 प्रश्नों के उत्तर दीजिए।
- सभी प्रश्न समान (12.5) अंक के हैं।
- उत्तर अंग्रेजी या हिंदी में लिखे जा सकते हैं, लेकिन पूरे पेपर में एक ही माध्यम का उपयोग किया जाना चाहिए।
- साधारण कैलकुलेटर के उपयोग की अनुमति है।

1. (a) How would you account for the predominant choice of arithmetic mean of statistical data as a measure of central tendency? Under what circumstances would it be appropriate to use mode or median?
(b) The linear correlation coefficient for the following data is -0.68. Interpret the relationship between these two variables X and Y on the basis of the following data.

X	Y
3	54
4	23
7	76
9	56
12	87
34	12
56	3

- (c). Is null hypothesis always rejected? Substantiate your answer.

1. a. केंद्रीय प्रवृत्ति के माप के रूप में आप सांख्यिकीय आंकड़ों के अंकगणितीय माध्य की प्रमुख पसंद के लिए कैसे जिम्मेदार होंगे? बहुलक या माधिका का प्रयोग किन परिस्थितियों में करना उचित होगा?

b. निम्न डेटा के लिए रैखिक सहसंबंध गुणांक -0.68 है। निम्नलिखित आँकड़ों के आधार पर इन दो चरों X और Y के बीच संबंध की व्याख्या कीजिए।

X	Y
3	54
4	23
7	76
9	56
12	87
34	12
56	3

c. क्या शून्य परिकल्पना हमेशा खारिज होती है? अपने उत्तर की पुष्टि कीजिए।

2. (a) What are the various steps in filtering data in MS Excel/R?

(b) We have a list of the marks of students during 2014 to 2019 and we want to sum all based on their genders. Write the syntax "SUMIF" to sum the marks of boys and girls separately. State which gender scored higher marks.

Year	Gender	Marks
2014	Boys	72
2015	Girls	75
2016	Girls	74
2017	Girls	71
2018	Boys	73
2019	Boys	71

(c) "Measures of dispersion and central tendency are complementary to each other in highlighting the characteristics of a frequency distribution." Explain the statement giving suitable examples.

2. a. MS Excel/R में डेटा फ़िल्टर करने के विभिन्न चरण क्या हैं?

b. हमारे पास 2014 से 2019 के दौरान छात्रों के अंकों की एक सूची है और हम सभी को उनके लिंग के आधार पर जोड़ना चाहते हैं। लड़कों और लड़कियों के अंकों को अलग-अलग जोड़ने के लिए सिटैक्स "SUMIF" लिखें। बताएं कि किस लिंग ने अधिक अंक प्राप्त किए।

Year	Gender	Marks
2014	Boys	72
2015	Girls	75
2016	Girls	74
2017	Girls	71
2018	Boys	73
2019	Boys	71

c. "विक्षेपण के उपाय और केंद्रीय प्रवृत्ति बारंबारता वितरण की विशेषताओं को उजागर करने में एक दूसरे के पूरक हैं।" उपयुक्त उदाहरण देते हुए इस कथन की व्याख्या कीजिए।

3 (a) "A coefficient estimate is statistically significant at the 5% significance level". Explain.

(b) Write the syntax for calculation of skewness in excel/R from the following data set and also interpret the result of skewness found if skewness is -0.027.

Class	A	B	C	D	E	F	G
Marks	54	23	76	56	87	12	3

(c) Imagine you invested \$10,000 in a portfolio with the returns outlined below:

- From Jan 1, 2014, to Jan 1, 2015, your portfolio grew to \$13,000 (or 30% in year one).
- On Jan 1, 2016, the portfolio was \$14,000 (or 7.69% from Jan 2015 to Jan 2016).
- On Jan 1, 2017, the portfolio ended with \$19,000 (or 35.71% from Jan 2016 to Jan 2017).

Calculate the compound annual growth rate from January 2014 to January 2017?.

3 a. "एक गुणांक अनुमान 5% महत्व स्तर पर सांख्यिकीय रूप से महत्वपूर्ण है"। समझाना।

b. निम्नलिखित डेटा सेट से एक्सेल/आर में तिरछापन की गणना के लिए वाक्य रचना लिखें और यदि तिरछापन - 0.027 है तो पाए गए तिरछापन के परिणाम की व्याख्या करें।

Class	A	B	C	D	E	F	G
Marks	54	23	76	56	87	12	3

c. कल्पना कीजिए कि आपने नीचे दिए गए रिटर्न के साथ एक पोर्टफोलियो में \$10,000 का निवेश किया है:

- 1 जनवरी 2014 से 1 जनवरी 2015 तक, आपका पोर्टफोलियो बढ़कर \$13,000 (या एक साल में 30%) हो गया।
- 1 जनवरी 2016 को पोर्टफोलियो \$14,000 (या जनवरी 2015 से जनवरी 2016 तक 7.69%) था।
- 1 जनवरी, 2017 को, पोर्टफोलियो \$19,000 (या जनवरी 2016 से जनवरी 2017 तक 35.71%) के साथ समाप्त हुआ। जनवरी 2014 से जनवरी 2017 तक चक्रवृद्धि वार्षिक वृद्धि दर की गणना कीजिए।

4. (a) How do you create pivot tables in MS Excel?

(b) Write the MS Excel method or syntax for covariance and correlation of the following data set.

X	8	2	4	7	9	6	20	12	16
Y	10	9	4	2	8	8	9	3	45

(c) Explain the meaning and significance of the concept of regression.

4.a. आप एमएस एक्सेल में पिवट टेबल कैसे बनाते हैं?

b. निम्नलिखित डेटा सेट के सहप्रसरण और सहसंबंध के लिए एमएस एक्सेल विधि या सिंटैक्स लिखें।

X	8	2	4	7	9	6	20	12	16
Y	10	9	4	2	8	8	9	3	45

c. प्रतिगमन की अवधारणा के अर्थ और महत्व की व्याख्या करें।

5. (a) Can we compare the standard deviation for both the following data sets? State how can you determine the variability of these data sets?

Student	Height (cm)	Weight(Kg)
1	178	65
2	180	70
3	175	71
4	185	72
5	177	73
6	190	79

(b) Consider a project which has the following cash stream:

Investment	100000
Saving in Year	Cash Flow
1	200000
2	200000
3	300000
4	300000
5	350000

The cost of capital, k , for the firm is 10 per cent. Find the net present value of the proposal?

(c) Explain the process of capital budgeting.

5.a. क्या हम निम्नलिखित दोनों डेटा सेटों के लिए मानक विचलन की तुलना कर सकते हैं? बताएं कि आप इन डेटा सेटों की परिवर्तनशीलता कैसे निर्धारित कर सकते हैं?

Student	Height (cm)	Weight(Kg)
1	178	65
2	180	70
3	175	71
4	185	72
5	177	73
6	190	79

b. एक ऐसी परियोजना पर विचार करें जिसमें निम्नलिखित नकदी प्रवाह हो:

Investment	100000
Saving in Year	Cash Flow
1	200000
2	200000
3	300000
4	300000
5	350000

फर्म के लिए पूंजी की लागत, k , 10 प्रतिशत है। प्रस्ताव का शुद्ध वर्तमान मूल्य ज्ञात कीजिए?

c. पूंजी बजटन की प्रक्रिया को समझाइए।

6. (a) Given the estimated linear consumption function. Interpret its coefficients and R^2 .

A. The Linear Consumption of India

$$\hat{C}_t = \hat{\beta}_1 + \hat{\beta}_2 Y_t$$

$$= 137.607 + 0.561 Y_t$$

(se) (10.029) (0.011)

(t-statistic) (13.721) (50.377)

$df = n - k = 8$ $F[1, 8] = 2537.86$ $R^2 = 0.997$

DW = 2.899

(b) How to construct a bar graph and pie chart in MS Excel/R?

(c) Explain the datasets available in Ministry of Statistics and Programme Implementation (MOSPI).

6. a. अनुमानित रैखिक खपत समारोह को देखते हुए। इसके गुणांकों और R^2 की व्याख्या कीजिए।

A. The Linear Consumption of India

$$\hat{C}_i = \hat{\beta}_1 + \hat{\beta}_2 Y_i$$

$$= 137.607 + 0.561 Y_i$$

(se) (10.029) (0.011)

(t-statistic) (13.721) (50.377)

df = n - k = 8 F[1, 8] = 2537.86 R² = 0.997

DW = 2.899

- b. एमएस एक्सेल/आर में बार ग्राफ और पाई चार्ट कैसे बनाएं?
- c. सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय (MOSPI) में उपलब्ध डेटासेट की व्याख्या करें।

Year	Consumption (C _i)	Income (Y _i)
1980	137.607	0
1981	137.607 + 0.561 * Y ₁	Y ₁
1982	137.607 + 0.561 * Y ₂	Y ₂
1983	137.607 + 0.561 * Y ₃	Y ₃
1984	137.607 + 0.561 * Y ₄	Y ₄
1985	137.607 + 0.561 * Y ₅	Y ₅
1986	137.607 + 0.561 * Y ₆	Y ₆
1987	137.607 + 0.561 * Y ₇	Y ₇
1988	137.607 + 0.561 * Y ₈	Y ₈

Year	Consumption (C _i)	Income (Y _i)
1980	137.607	0
1981	137.607 + 0.561 * Y ₁	Y ₁
1982	137.607 + 0.561 * Y ₂	Y ₂
1983	137.607 + 0.561 * Y ₃	Y ₃
1984	137.607 + 0.561 * Y ₄	Y ₄
1985	137.607 + 0.561 * Y ₅	Y ₅
1986	137.607 + 0.561 * Y ₆	Y ₆
1987	137.607 + 0.561 * Y ₇	Y ₇
1988	137.607 + 0.561 * Y ₈	Y ₈

Name of Course: CBCS B.A. (Prog.) SEC, Semester VI

Name of the Paper: Basic Computational Techniques for Data Analysis

Unique Paper Code: 62273601_OC

Maximum Marks: 50 Marks

Time Limit: 3 Hours

1. (a) How to construct a line graph and histogram in MS Excel?
 (b) Arithmetic mean is a superior average as compared to other averages. Explain with examples.
 (c) Explain pivot tables along with their features.
2. (a) What do you understand by dispersion? Explain briefly the various methods used for measuring dispersion?
 (b) Consider the cash flows of two projects A and B.

Investment	100000	100000
Savings in Year	Cash Flow of A	Cash Flow of B
1	50000	20000
2	30000	20000
3	20000	20000
4	10000	40000
5	10000	50000
6	-	60000

Briefly explain which project is preferred according to payback criteria.

- (c) Given the data table with order numbers and the bill amount. Write the syntax "SUMIF" to sum the bill amount for a particular brand order where the order number starts with 'Amzcd' and 'ABVF'.

S No.	Order No.	Bill Amount (Rs)
1	Amzcd 008	450
2	Amzcd 009	500
3	AFJK 010	1102
4	Amzcd 011	850
5	Amzcd 012	564
6	ABVF 013	345
7	Amzcd 014	10.50
8	Amzcd 015	835.5
9	ABVF 016	1236
10	AFJK 017	478.20
11	Amzcd 018	100.25

3. (a) Explain the meaning and significance of the concept of regression.
 (b) Write the MS Excel method or syntax for Skewness, Kurtosis and standard deviation of the following data set. Also calculate the value of Skewness, Kurtosis and standard deviation of the given data.

X	4	2	3	7	9	10	23	12	16
---	---	---	---	---	---	----	----	----	----

- (c) What are the various steps in filtering data in MS Excel?
4. (a) Consider a project which has the following cash stream:

Investment	100000
Saving in Year	Cash Flow
1	200000
2	200000
3	300000
4	300000
5	350000

The cost of capital, k , for the firm is 10 per cent. Find the net present value of the proposal?

(b)) Explain the process of capital budgeting.

(c) Explain the term "Internal rate of return". Give suitable examples.

5. (a) Given the estimated linear equation. Interpret its coefficients and R^2 .

$$\hat{Y}_i = -0.0144 + 0.7241 X_i$$

$$\text{var}(\hat{\beta}_1) = 0.7649 \quad \text{se}(\hat{\beta}_1) = 0.8746$$

$$\text{var}(\hat{\beta}_2) = 0.00483 \quad \text{se}(\hat{\beta}_2) = 0.0695$$

$$r^2 = 0.9077 \quad \hat{\sigma}^2 = 0.8816$$

(b) Explain the method of sorting data in MS Excel.

(c) How would you interpret the sign and magnitude of a calculated 'r'? Consider in particular the values $r=0$, $r=-1$ and $r=+1$.

6. Write short notes on any two:

(a) R square

(b) Mode

(c) Coefficient of Variation

