INSTRUCTIONS

Answer **Question number 1** and **Any other THREE questions**
 Total possible mark is **100.**

Start **each** question on a new page in your answer
Booklet.

The marks allocated to **each** question are shown at the end of the section.

Show all your workings.

Credit will be awarded for logical, systematic and neat presentations.

1. Make a clear distinction of the following concepts, by highlighting the differences and giving concrete examples
    (a) Sample and population   **[5 Marks]**
    (b) Sample regression function and Population Regression Function
                                                                          [5 Marks]
    (c) R squared and adjusted R squared   **[5 Marks]**
    (d) Small Sample properties and large sample properties of a good estimator
                                                                          [5 Marks]
    (e) Cross sectional data and Time series data   **[5 Marks]**
    (f) Null Hypothesis and Alternative hypothesis   **[5 Marks]**
    (g) Significant and Insignificant parameter   **[5 Marks]**
    (h) Degrees of freedom and sample size   **[5 Marks]**

2. A research student has been investigating the relationship between office rents and vacancy rates in various US cities. She was able to obtain data from 30 cities on average monthly office rents in $ per square foot (Y) and vacancy rates in per cent (X). Running a simple linear regression using Excel she obtained the following results.
SUMMARY OUTPUT
Regression Statistics
Multiple R 0.53957932
R Square 0.29114585

	coeficients	*Std Error*	*t Stat*
Intercept	20.639	1.142	18.060
X Variable	-0.303	0.089	-0.48730158

a) Interpret the values obtained for the regression coefficients and explain whether they are in line with what you would have anticipated *a priori*.
                                                                          [4 marks]
b) Can you conclude at the 5% significance level that higher vacancy rates result in lower rents? Explain, giving a full interpretation of the calculated values for the t Stat and P-value for the X Variable.   **[4 marks]**
c) The Excel results include information that can be used to provide a 95% confidence interval for the slope parameter. Discuss the values for the Lower and Upper limits shown in the table above, explaining how they relate to the estimated Standard Error.   **[4 marks]**
d) Fully explain the meaning and interpretation of the R square value shown in the table.   **[4 marks]**
e) The student's supervisor comments that the regression results are based on a rather small sample, which means that they are associated with a low number of degrees of freedom. Explain what this means and note its implications for hypothesis testing of the regression parameters. **[4 marks]**

3.Given the volume of sales and the advertising expenditures follows

Advertising	8	12	15	18	18	20	20	23	24	24
Sales	10	20	30	40	50	60	70	80	90	100

(i)   Give apriori expectations of the signs of the parameters and justify your prediction.   **[4 marks]**

(ii)   Draw a scatter diagram and fit in the regression equation.   **[4 marks]**

(iii) Calculate and interpret the goodness of fit.   **[4 marks]**

(iv)   If advertising expenditure rises to 30 what would be the forecast for the sale   **[4 marks]**

(v)   Compute standard errors of estimated parameters and test for significance.   **[4 marks]**

4. (a) Two data sets were rejected by an econometrician, Christopher from Malawi. Please investigate the following data sets and establish the problem. **[5 marks]**

Data set A

Consumption	1000	1300	1500	1600	1700	1800	2000
Price	90	85	80	75	70	65	60
Income	180	170	160	150	140	130	120

Data set B

Consumption	1000	1300	1500	1600	1700	1800	2000	
Price	90	85	80	75	70	65	60	
Income	180	170	160	150	140	130	112	

Consumption is the explained variable

Income and Price are explanatory variable

(b) Demonstrate the effects of such a problem   **[5 marks]**

(c) What are the causes of the above problem?   **[5 marks]**

(d) What are the solutions to the problem?   **[5 marks]**

5. For a sample of 27 students, they have the examination mark, M, total hours spent studying, H, hours on primary study, P, and hours spent on revision, R

$$M = 45.6 + 0.15\ P + 0.21\ R - 0.5\ H$$
$$\quad (2.8) \quad (0.03) \quad (0.14) \quad (3.486)$$
$$R^2 = 0.99 \qquad\qquad D.W = 2.00$$

(a) What conclusions can you draw from this model about the relationship between final mark (M), and the explanatory variables **(H)** Hours spent studying hours, (P) Hours of primary study, (R) hours spent on revision? (justify your response quantitatively)   **[5 marks]**
(b) How would you use the goodness of fit to support of refute your conclusions in item (a)?   **[5 marks]**
(c) Test the significance of the whole model   **[5 marks]**
(d) Using as many examples as is possible demonstrate how the estimated equation can be used for forecasting   **[5 marks]**

5. Given the following data on Consumption (C) and Income (M)

Consumption	5	5.5	5.5	6	7.5	8	8.5	9	9	10
Income	12	13	150	15	18	18	20	21	25	26

(a) Plot a scatter diagram and ascertain apriori expectations.   **[2 marks]**

(b) Estimate the regression function and interpret the findings.   **[4 marks]**

(c) Fit the estimated regression function in the scatter diagram.   **[3 marks]**

(d) Compute the coefficient of determination and interpret it.   **[3 marks]**

(e) Panashe argued that Income does not affect Consumption. Demonstrate how you could prove him right or wrong.   **[4 marks]**

(f) Test the significance of the whole model   **[4 marks]**

6. (a) What is the difference between the coefficient of determination and the adjusted coefficient of determination?   **[4 Marks]**
(b) What is the difference between a null hypothesis and an alternative hypothesis.   **[4 Marks]**

(c) Two consultant companies were contracted to estimate the demand for product **Z** for IM private Limited. In order to minimize costs both companies restricted the sample size to 27 customers.

Coblar Consutants

$Q_Z = 100.325 + 1.5\text{Income} - 3.6\text{ Price}$

      (22.3)   (0.19)   (0.21)

$R^2 = 0.966$

Panashe Consultants

$Q_Z = 30,2 - 0.22\text{Income} - 1.9\text{ Price}$

      (15.9)   (0.71)   (0.33   )

$R^2 = 0.972$ DW

Compare these two models and determine the best one **[12 Marks]**

End of Paper