PARVATHANENI BRAHMAYYA
SIDDHARTHA COLLEGE OF ARTS & SCIENCE
Autonomous
Siddhartha Nagar, Vijayawada–520010
Re-accredited at 'A+' by the NAAC

Course Code				23STMDL201			
Title of the Course				Applied Statistical Analysis			
Offered to: (Programme/s)				B.Com.Hons(GEN,TPP,BPM,BANKING,FINANCE,CA)			
L	2	T	0	P	0	C	2
Year of Introduction:		2024-25		Semester:			3
Course Category:		MDC		Course Relates to:		Local, Regional, National, Global	
Year of Introduction:		2024 - 25		Percentage:		NA	
Type of the Course:				Multi Disciplinary Course			
Crosscutting Issues of the Course:				NA			
Pre-requisites, if any				NA			

Course Description:

This course provides an in-depth understanding of correlation and regression analysis, which are vital techniques in statistical analysis. It introduces the concepts, methods, and applications of different types of correlation, such as Karl Pearson's and Spearman's rank correlation, as well as regression analysis, both linear and non-linear. Additionally, the course covers index numbers, their calculation, and their application, along with tests of adequacy for index number formulae. Students will learn to apply these statistical tools in real-world business and economic situations.

Course Objectives:

S. No	COURSE OBJECTIVES
1	introduce students to fundamental concepts of correlation and regression analysis.
2	equip students with knowledge of various methods for studying correlation and regression.
3	enable students to apply statistical methods like linear regression and correlation coefficients in business scenarios.
4	understand and apply the concept of index numbers and test their adequacy in economic indices.
5	develop the ability to analyze and interpret statistical data using advanced correlation and regression techniques.

Course Outcomes

At the end of the course, the student will be able to…

NO	COURSE OUTCOME	BTL	PO	PSO
CO1	apply statistical methods like Pearson's coefficient and regression techniques to analyze data and draw conclusions about relationships between variables in business.	K2	1	1
CO2	critically assess the strengths and weaknesses of methods such as the scatter diagram, Karl Pearson's coefficient, and Spearman's rank correlation.	K5	1	1
CO3	evaluate the use of index numbers in various economic scenarios.	K5	1	1
CO4	delve into regression analysis, focusing on explained and unexplained variations.	K4	1	1
CO5	understand advanced topics like multiple and partial correlations and logistic regression	K2	1	1

For BTL: K1: Remember; K2: Understand; K3: Apply; K4: Analyze; K5: Evaluate; K6: Create

CO-PO-PSO MATRIX									
CO NO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PSO1	PSO2
CO1	3							3	
CO2	3							3	
CO3	3							3	
CO4	3							3	
CO5	2							2	

Use the codes 3, 2, 1 for High, Moderate and Low correlation Between CO-PO-PSO respectively

Course Structure:

Unit– 1: Correlation   (10 hours)

Introduction, Types of correlation, methods of studying correlation-scatter diagram, Karl Pearson's co-efficient of correlation and Spearman's Rank correlation co-efficient –merits, demerits properties and Applications. Concurrent deviation method. Concept of partial and multiple correlations.

Examples:

1. Portfolio diversification: Identifying assets that move in opposite directions canhelp create diversified portfolios, reducing risk.
2. Risk management: Understanding the correlation between market movements and investment performance can aid in risk management strategies.

Case Study

1. **Relationship between Education and Income**
2. **Objective:** To examine the correlation between education level and income.
3. **Data:** Data on education level (e.g., years of schooling) and income for a large sample of individuals.
4. **Analysis:** A scatter plot can be used to visualize the relationship between education and income. The correlation coefficient can be calculated to quantify the strength and direction of the relationship.
   VISULIZATION
5. Visualization plays a crucial role in understanding and interpreting correlation between variables. It provides a visual representation of the relationship, making it easier to identify patterns, trends, and outliers.
6. **Scatter Plot:** This is the most common method for visualizing correlation.

7. Each data point is represented by a dot on the graph, with the x-axis representing one variable and the y-axis representing the other.
8. The pattern of the dots reveals the correlation:
9. **Positive correlation:** The dots form an upward-sloping line.
10. **Negative correlation:** The dots form a downward-sloping line.
11. **No correlation:** The dots are scattered randomly.

Exercises

1.Data Collection:
i. Gather data on two variables of interest (e.g., height and weight, temperature and ice cream sales, stock prices of two companies).
ii. Create a scatter plot to visualize the relationship between the variables.
iii. Calculate the correlation coefficient using a statistical software or calculator.
iv. Interpret the correlation coefficient and describe the relationship between the variables.

Correlation and Causation:
i. Explain the difference between correlation and causation.
ii. Provide examples of situations where a strong correlation exists but causation is not clear.
iii. Discuss how additional evidence or experiments can help establish causation.
iv. Projects

Specific Resources (web): file:///C:/Users/user25/Downloads/Unit-6.pdf

Unit – 2: Regression   **(10 hours)**
Concept of Regression, Linear and Non Linear regression. Linear Regression – Regression lines, Regression coefficients and it properties. Correlation vs Regression. Explained and Unexplained variations. Coefficient of determination. Concept of Multiple and Partial correlation coefficient. Concept of Logistic regression

Examples
1. Sales forecasting: Predicting future sales based on factors like advertising expenditure, price, and competition.
2. Market research: Understanding customer behavior and preferences.
3. Financial modeling: Predicting stock prices, interest rates, and economic indicators.

Exercises/Project:
Real-world Data Regression Project
1. Divide students into groups.
2. Assign each group a real-world dataset (e.g., housing prices, student grades, climate data).
3. Guide students in selecting appropriate dependent and independent variables.
4. Have students create scatter plots to visualize the relationship between variables.
5. Use statistical software to fit a regression model and interpret the coefficients.
6. Discuss the findings and potential applications of the model.

Specific Resources (web): https://www.khanacademy.org/math/senior-high-school-statistics-probability/xc50280f9c512251f:2nd-quarter

Unit–3: Index numbers   **(10hours)**
Introduction, Uses and limitations of index numbers, types of Index Numbers, Simple or Unweighted and Weighted Index numbers, Base Shafting, Splicing, Test of Adequacy an index number formulae. Cost of Living Index number. Simple problems.

Project : Analyzing the Consumer Price Index (CPI) Over Time
Objective: Analyze the changes in the Consumer Price Index (CPI) over the last 10 years and examine how it reflects inflation trends in your country.
Tasks:
1. Collect data on the CPI from a reliable source, such as a government statistical agency or central

bank, for the last 10 years.
2. Calculate the percentage changes in CPI year-over-year and analyze the trends.
3. Compare the changes in the CPI with key economic events (e.g., financial crises, policy changes, global pandemics).
4. Assess how accurately the CPI reflects the actual cost of living and identify any limitations in the index.

Outcome: This project will help students understand how the CPI, a weighted index, is calculated and applied, and they will critically assess its limitations as an economic indicator.

Case Study: Base Shifting in GDP Deflator Calculations
Scenario: A country updates the base year for its GDP deflator from 2010 to 2020. This shift significantly changes the reported inflation and real GDP growth rates for the last decade. The government defends the base shifting decision as necessary to reflect the current structure of the economy.

Analysis:
1. Investigate how base shifting impacts the interpretation of inflation and growth trends over time.
2. Discuss the technical process of base shifting in index numbers and the potential challenges it poses for policy analysts and economists.
3. Evaluate the benefits and limitations of base shifting, particularly in developing countries where economic structures change rapidly.

Outcome: This case study will help students understand the complexities of base shifting, splicing, and the implications for long-term economic analysis.

Specific Resources (web): https://egyankosh.ac.in/bitstream/123456789/12281/1/Unit-12.pdf

Textbook:
S. C. Gupta, Fundamentals of Statistics, 8th Edition, 2023, Himalaya Publishing House Pvt. Ltd 'Ramdoot', Dr. Bhalerao Marg, Girgaon, Mumbai – 400 004, Maharashtra, India

References Books:
1. **Business Statistics A First Course, 8e Paperback – 30 October 2022, David.** Levine (Author)
2. Business Statistics: Problems & Solutions by J.K. Sharma (Author), Vikas Publishing House Pvt Ltd. Noida, UP, India

***

23STMDL201: **Applied Statistical Analysis**
Offered to: B.Com.Hons(GEN,TPP,BPM,BANKING,FINANCE,CA)
Max. Marks: 35   **Semester III**   **Max. Time: 2Hrs**

Section A (Short Answer Questions)
(Answer any Three questions; each question carries 5 marks)
(Total 15 Marks)

Q1. Explain the concept of correlation and describe the differences between **positive**, **negative**, and **zero correlation**. (CO1-K2)

Q2. Using the data given below, calculate **Karl Pearson's coefficient of correlation** between the scores of two students in a test.(CO1:K3)
Student A: 20,   25,   28,   35,   40
Student B: 18,   22,   24,   30,   36

Q3. Describe the key differences between **linear** and **non-linear regression**. What are the applications of linear regression in business analysis? (CO4-K4)

Q4. Calculate the **Cost-of-Living Index Number** using the following data: (CO3-K3)

Commodity	Base Year Price	Current Year Price	Weights
Food	100	120	50
Rent	80	100	20
Clothing	60	75	10
Misc.	40	60	20

Q5. Explain the **tests of adequacy** for index number formulae. How do these tests ensure the accuracy of an index number? (CO3-K2)

Section B (Long Answer/Problem-Solving Questions)
(Answer any 2 questions; each question carries 10 marks)
(Total 20 Marks)

Q6. Given the following data, calculate the **regression equations** of Y on X and X on Y, and find the values of **regression coefficients**:(CO4-K3)

X	10	12	14	16	18
Y	15	14	17	18	20

Q7. Define and explain the concepts of **partial correlation** and **multiple correlation**. In what type of business scenarios would these be applied, and how are they calculated? (CO5-K2)

Q8. Calculate the **Weighted Index Number** using **Paasche's Method** for the following data: (CO3-K3, K5)

Commodity	Price in Base Year	Price in Current Year	Quantity in Current Year
A	20	25	30
B	30	40	20
C	10	15	40

Also, explain the limitations of **Paasche's Index** compared to **Laspeyres Index**.

***

Distribution of Marks and Cognitive Levels:
- **Numerical Problems (60%)**: Q2, Q4, Q6, Q8
- **Theory Questions (40%)**: Q1, Q3, Q5, Q7
- **Cognitive Levels:**
  o K2:   Understanding: Q1, Q3, Q5, Q7
  o K3:   Applying: Q2, Q4, Q6, Q8
  o K4:   Analyzing: Q3, Q6
  o K5:   Evaluating: Q8