

GUJARAT TECHNOLOGICAL UNIVERSITY
Diploma Engineering – SEMESTER – 6 (NEW) – EXAMINATION – Summer-2025

Subject Code: 4360707

Date: 16-05-2025

Subject Name: Introduction to Data Analysis

Time: 10:30 AM TO 01:00 PM

Total Marks: 70

Instructions:

1. Attempt all questions.
2. Make Suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Use of simple calculators and non-programmable scientific calculators are permitted.
5. English version is authentic.

		Marks
Q.1	(a) Classify types of data.	03
પ્રશ્ન.1	(અ) ડેટાના પ્રકારોનું વર્ગીકરણ કરો.	૦૩
	(b) Describe vertical stacking and horizontal stacking with example.	04
	(બ) વર્ટિકલ સ્ટેકીંગ અને હોરિઝોન્ટલ સ્ટેકીંગ ઉદાહરણ આપી વર્ણવો.	૦૪
	(c) Discuss characteristics of data and need of data analysis.	07
	(ક) ડેટાની લાક્ષણિકતાઓની ચર્ચા કરો અને ડેટા એનાલીસીસની જરૂરિયાત વર્ણવો.	૦૭
OR		
	(c) Discuss various phases of data analytics life cycle.	07
	(ક) data analytics ની જુદા જુદા ફેઝીસની ચર્ચા કરો.	૦૭
Q.2	(a) List random module functions.	03
પ્રશ્ન.2	(અ) રેન્ડમ મોડ્યુલ ફંક્શનસની યાદી લખો.	૦૩
	(b) Describe merge operations with multiple data frames using pandas library.	04
	(બ) પાનાડસ લાઇબ્રેરીની મદદ થી મલ્ટીપલ ડેટા ફ્રેમ્સનો ઉપયોગ કરીને મર્જ ઓપરેશન વર્ણવો.	૦૪
	(c) Write a pandas program to implement following operations : (1) Use the loc function to display rows where 'Survived' is 1 (2) Display the top 3 passengers with the largest 'Age' (3) Show the 3 passengers with the smallest 'Age'	07
	(ક) નીચેની ક્રિયાઓ અમલમાં મૂકવા માટે pandas પ્રોગ્રામ લખો: (1) Use the loc function to display rows where 'Survived' is 1 (2) Display the top 3 passengers with the largest 'Age' (3) Show the 3 passengers with the smallest 'Age'	૦૭
OR		
Q.2	(a) List steps for handling missing values in dataframe.	03
પ્રશ્ન.2	(અ) ડેટા ફ્રેમમાં મિસિંગ વેલ્યુ હેન્ડલ કરવા માટેના સ્ટેપ્સ લખો.	૦૩
	(b) Describe data normalization and transformation.	04
	(બ) ડેટા નોર્મલાઇઝેશન અને ટ્રાન્સફોર્મેશન વર્ણવો.	૦૪
	(c) Write basics of sending GET/POST requests, accessing response content and authenticating with requests.	07
	(ક) GET/POST વિનંતીઓ મોકલવા, પ્રતિસાદ સામગ્રીને એક્સેસ કરવા અને વિનંતીઓ સાથે પ્રમાણિત કરવાની મૂળભૂત બાબતો લખો.	૦૭
Q. 3	(a) List different correlation techniques.	03

પ્રશ્ન.3	(અ) જુદા જુદા પ્રકારની કોરીલેશન ટેકનીક્સની યાદી લખો.	૦૩
	(b) Describe regression modeling to analyze relationships between variables.	04
	(બ) વેરાએબલ્સ વચ્ચેની રીલેશનશીપને એનાલાઇઝ કરીને રેગ્રેશન મોડેલીંગ વર્ણવો.	૦૪
	(c) Apply basics of descriptive statistics including measures of central tendency. Include mean, median and mode.	07
	(ક) કેન્દ્રીય વલણના પગલાં સહિત વર્ણનાત્મક આંકડાઓની મૂળભૂત બાબતો લાગુ કરો. સરેરાશ, મધ્ય અને મોડ શામેલ કરો.	૦૭
	OR	
Q. 3	(a) List different probability distributions.	03
પ્રશ્ન.3	(અ) જુદાજુદા પ્રકારની સંભાવના વિતરણની યાદી લખો.	૦૩
	(b) Describe methods of least squares.	04
	(બ) લીસ્ટ સ્કેવરમેથડ વર્ણવો.	૦૪
	(c) Calculate and interpret person's correlation coefficient for examining the relationship between fare and passenger class on the titanic dataset.	07
	(ક) ટાઇટેનિક ડેટાસેટ પર ભાડું અને પેસેન્જર વર્ગ વચ્ચેના સંબંધની તપાસ કરવા માટે વ્યક્તિના સહસંબંધ ગુણાંકની ગણતરી કરો અને તેનું અર્થઘટન કરો.	૦૭
Q. 4	(a) Describe fundamentals of data visualization and its role in effective communication of insights.	03
પ્રશ્ન.4	(અ) ડેટા વિઝ્યુલાઇઝેશનને મૂળભૂત રીતે વર્ણવો અને આંતરદૃષ્ટિના અસરકારક સંચારમાં તેની ભૂમિકા પણ વર્ણવો.	૦૩
	(b) Describe multivariate analysis for a comprehensive examination of multiple variables simultaneously.	04
	(બ) એકસાથે બહુવિધ ચલોની વ્યાપક તપાસ માટે મલ્ટિવેરિયેટ વિશ્લેષણનું વર્ણન કરો.	૦૪
	(c) Create a python script that uses matplotlib to generate bar charts from sample data. Customize the appearance of plot including labels, colors and annotations.	07
	(ક) એક પાયથોન સ્ક્રિપ્ટ બનાવો જે સેમ્પલ ડેટામાંથી બાર ચાર્ટ બનાવવા માટે matplotlib નો ઉપયોગ કરે છે. લેબલ્સ, રંગો અને ટીકાઓ સહિત પ્લોટના દેખાવને કસ્ટમાઇઝ કરો.	૦૭
	OR	
Q. 4	(a) Discuss importance of data visualization.	03
પ્રશ્ન.4	(અ) ડેટા વિઝ્યુલાઇઝેશનના મહત્વની ચર્ચા કરો.	૦૩
	(b) Describe different probability distribution analysis techniques.	04
	(બ) વિવિધ સંભાવના વિતરણ વિશ્લેષણ તકનીકોનું વર્ણન કરો.	૦૪
	(c) Generate a Seaborn strip plot to visualize the distribution of total bill amounts within different days of the week in the "tips" dataset.	07
	(ક) "ટિપ્સ" ડેટાસેટમાં અઠવાડિયાના જુદા જુદા દિવસોમાં કુલ બિલની રકમના વિતરણને વિઝ્યુલાઇઝ કરવા માટે સીબોર્ન સ્ટ્રીપ પ્લોટ બનાવો.	૦૭
Q.5	(a) List data visualization techniques for big data.	03
પ્રશ્ન.5	(અ) બીગ ડેટા માટે ડેટા વિઝ્યુલાઇઝેશન તકનીકોની સૂચિ બનાવો.	૦૩
	(b) Describe data visualization tools.	04
	(બ) ડેટા વિઝ્યુલાઇઝેશન ટૂલ્સનું વર્ણન કરો.	૦૪
	(c) Describe basic data visualization with matplotlib.	07
	(ક) matplotlib સાથે મૂળભૂત ડેટા વિઝ્યુલાઇઝેશનનું વર્ણન કરો.	૦૭
	OR	
Q.5	(a) List various data visualization tools for handling large data sets.	03
પ્રશ્ન.5	(અ) બીગ ડેટા સેટને હેન્ડલ કરવા માટે વિવિધ ડેટા વિઝ્યુલાઇઝેશન ટૂલ્સની યાદી બનાવો.	૦૩

(b)	Discuss challenges of big data visualization.	04
(બ)	બીગ ડેટા વિઝ્યુલાઇઝેશનના પડકારોની ચર્ચા કરો.	૦૪
(c)	Discuss interactive data visualization with ploty.	07
(ક)	ploty સાથે ઇન્ટરેક્ટિવ ડેટા વિઝ્યુલાઇઝેશનની ચર્ચા કરો	૦૭