Seat No. / Enrolment No.:

GUJARAT TECHNOLOGICAL UNIVERSITY

Diploma Engineering – SEMESTER – 6 – EXAMINATION – Summer-2026

Subject Code: 4360707 **Date: 20-05-2026**

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)**	Define Data analysis. Explain how data analysis is applied in real-world applications.	**03**
પ્રશ્ન.1	(અ)	ડેટા એનાલિસિસ ને વ્યાખ્યાયિત કરો. વાસ્તવિક વિશ્વના ઉપયોગમાં ડેટા એનાલિસિસ કેવી રીતે લાગુ થાય છે તે સમજાવો.	૦૩
	(b)	Explain the challenges faced in visualizing big data.	**04**
	(બ)	બિગ ડેટા વિઝ્યુઅલાઇઝેશનમાં આવતા પડકારો સમજાવો.	૦૪
	(c)	Explain the analytical process.	**07**
	(ક)	એનલિટિકલ પ્રક્રિયા સમજાવો.	૦૭
		OR	
	(c)	Explain the various phases of the data analytics lifecycle.	**07**
	(ક)	ડેટા એનાલિટિક્સ લાઈફ સાઇકલના વિવિધ તબક્કાઓ સમજાવો.	૦૭
Q.2	**(a)**	Explain how to generate random numbers using NumPy.	**03**
પ્રશ્ન.2	**(અ)**	NumPy નો ઉપયોગ કરીને રેન્ડમ નંબરો કેવી રીતે જનરેટ કરવાં તે સમજાવો.	**૦૩**
	(b)	Explain how to handle missing values in a DataFrame.	**04**
	(બ)	ડેટાફ્રેમમાં ખૂટતી કિંમતોને કેવી રીતે હેન્ડલ કરવી તે સમજાવો.	૦૪
	(c)	Explain dimensionality reduction in data analysis.	**07**
	(ક)	ડેટા એનાલીસીસમાં ડાઈમેનસ્નાલીટી રિડકશન સમજાવો.	૦૭
		OR	
Q.2	**(a)**	Explain how to find the nlargest() and nsmallest() values in a DataFrame.	**03**
પ્રશ્ન.2	(અ)	ડેટાફ્રેમમાં nlargest() અને nsmallest() મૂલ્યો શોધવાની પ્રક્રિયા સમજાવો.	૦૩
	(b)	Discuss the methods to handle outliers in a dataset with example.	**04**
	(બ)	ઉદાહરણ સાથે ડેટાસેટમાં આઉટલાયર્સને હેન્ડલ કરવાની પદ્ધતિઓની ચર્ચા કરો.	૦૪
	(c)	Discuss vertical and horizontal stacking in NumPy.	**07**
	(ક)	NumPy માં વર્ટિકલ અને હોરીઝન્ટલ સ્ટેકિંગ સમજાવો.	૦૭
Q. 3	**(a)**	List various probability distribution techniques and their applications.	**03**
પ્રશ્ન.3	(અ)	વિવિધ સંભાવના વિતરણ તકનીકો અને તેમની એપ્લિકેશનોની સૂચિ બનાવો.	૦૩
	(b)	Discuss the use of Bernoulli distribution in binary outcome experiments.	**04**
	(બ)	બાઇનરી આઉટકમ પ્રયોગોમાં બર્નૂલી વિતરણનો ઉપયોગ પર ચર્ચા કરો.	૦૪

	(c)	Describe the process of conducting multivariate analysis for examining multiple variables.	**07**
	(ક)	બહુવિધ ચલોની તપાસ કરવા માટે મલ્ટિવેરિયેટ વિશ્લેષણ કરવાની પ્રક્રિયાનું વર્ણન કરો.	૦૭

OR

Q. 3	**(a)**	List and explain various measures of central tendency with example.	**03**
પ્રશ્ન.3	(અ)	સેન્ટ્રલ ટેંડેંસી ના વિવિધ માપદંડો ઉદાહરણ સાથે સમજાવો.	૦૩
	(b)	Explain Normal distribution in statistical analysis.	**04**
	(બ)	આંકડાકીય વિશ્લેષણમાં સામાન્ય વિતરણને સમજાવો.	૦૪
	(c)	Explain regression modeling techniques used to analyze relationships between variables.	**07**
	(ક)	ચલો વચ્ચેના સંબંધોનું વિશ્લેષણ કરવા માટે ઉપયોગમાં લેવાતી રીગ્રેશન મોડેલિંગ તકનીકને સમજાવો.	૦૭
Q. 4	**(a)**	Describe how effective data visualization can enhance decision-making processes.	**03**
પ્રશ્ન.4	(અ)	ડેટા વિઝ્યુલાઇઝેશન કેવી રીતે અસરકારક રીતે નિર્ણય લેવાની પ્રક્રિયામાં વધારો કરી શકે છે તેનું વર્ણન કરો.	૦૩
	(b)	Discuss the significance of Exponential distribution in statistical analysis.	**04**
	(બ)	આંકડાકીય વિશ્લેષણમાં એક્સ્પોનેન્શિયલ વિતરણના મહત્ત્વની ચર્ચા કરો.	૦૪
	(c)	List features of plotly and explain process of creating a simple line plot using plotly.	**07**
	(ક)	પ્લોટલીની વિશેષતાઓની ચર્ચા કરો અને પ્લોટલીનો ઉપયોગ કરીને સરળ લાઇન પ્લોટ બનાવવાની પ્રક્રિયા સમજાવો.	૦૭

OR

Q. 4	**(a)**	Explain the characteristics and use cases of violin plots in data analysis.	**03**
પ્રશ્ન.4	(અ)	ડેટા વિશ્લેષણમાં વાયોલિન પ્લોટના લક્ષણો અને ઉપયોગના કેસને સમજાવો.	૦૩
	(b)	Describe how scatterplots and other graphical techniques are used to identify correlations between variables.	**04**
	(બ)	ચલો વચ્ચેના સહસંબંધને ઓળખવા માટે સ્કેટરપ્લોટ અને અન્ય ગ્રાફિકલ તકનીકોનો કેવી રીતે ઉપયોગ થાય છે તેનું વર્ણન કરો.	૦૪
	(c)	List features of Seaborn. Describe advantages of using Seaborn compared to Matplotlib.	**07**
	(ક)	સીબોર્નની સુવિધાઓની સૂચિ બનાવો. Matplotlib ની તુલનામાં Seaborn નો ઉપયોગ કરવાના ફાયદાઓનું વર્ણન કરો.	૦૭
Q.5	**(a)**	Define web scraping. Discuss applications of web scraping.	**03**
પ્રશ્ન.5	(અ)	વેબ સ્ક્રેપિંગ વ્યાખ્યાયિત કરો. વેબ સ્ક્રેપિંગના ઉપયોગોની ચર્ચા કરો.	૦૩
	(b)	Explain recent trends in data collection and analysis techniques.	**04**
	(બ)	ડેટા સંગ્રહણ અને એનાલિસિસ પદ્ધતિઓમાં તાજેતરના વલણો સમજાવો.	૦૪
	(c)	Discuss features of matplotlib and explain process of creating a simple line plot using Matplotlib.	**07**
	(ક)	મેટપ્લોટલીબ ની વિશેષતાઓની ચર્ચા કરો અને મેટપ્લોટલીબ નો ઉપયોગ કરીને એક સરળ લાઇન પ્લોટ બનાવવાની પ્રક્રિયા સમજાવો.	૦૭

OR

Q.5	**(a)**	Define Data Transformation and Discuss Its Importance in Data Preprocessing	**03**
પ્રશ્ન.5	(અ)	ડેટા ટ્રાન્સફોર્મેશનને વ્યાખ્યાયિત કરો અને ડેટા પ્રીપ્રોસેસિંગમાં તેના મહત્વની ચર્ચા કરો	૦૩
	(b)	List various data visualization tools. Describe the key features of data visualization tools.	**04**

(બ)	વિવિધ ડેટા વિઝ્યુલાઇઝેશન ટૂલ્સની યાદી બનાવો. ડેટા વિઝ્યુલાઇઝેશન ટૂલ્સની મુખ્ય લાક્ષણિકતાઓનું વર્ણન કરો.	૦૪
(c)	Explain how to create a bar chart and customize its labels using Matplotlib.	**07**
(ક)	મેટપ્લોટલીબનો ઉપયોગ કરીને બાર ચાર્ટ કેવી રીતે બનાવવો અને તેના લેબલ્સને કસ્ટમાઇઝ કેવી રીતે કરવું તે સમજાવો.	૦૭

****