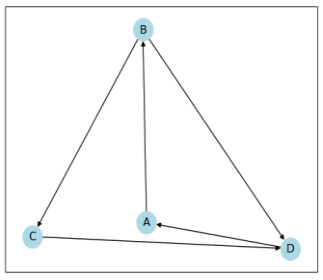
L.J. Institute of Engineering & Technology, Ahmedabad										
APP Practice Book_2024										
Note: The Practice Book is for reference only, LJU Test paper may not be compulsory set from this										
Sr. No.	**unit_number**	**keyword**	**question_text**	**answer_text**	**marks**	**previous_year**	**option1 (A)**	**option2 (B)**	**option3 (C)**	**option4 (D)**
1	1	Cleaning Data	Which of the following functions can be used to fill all null values in a data frame?	A	1		fillna()	filled()	fillnull()	filler()
2	1	Cleaning Data	Which attribute of dropna() can be used to select the columns from which null values are to be considered for removing rows?	C	1		thresh	how	subset	superset
3	1	Two-way cross tabulation	Which of the following pandas functions is used to generate cross tabulation?	D	1		crosstabulation	cross_tabulation	cross_tab	crosstab
4	1	DataFrame	Which of the following DataFrame attributes is used to return one or more specified row(s)?	C	1		locate	location	loc	find
5	1	DataFrame	Which of the following attributes can be used to show the number of rows and columns in a Pandas dataframe?	D	1		size	info	describe	shape
6	1	DataFrame	Which of the following is not displayed by the Pandas DataFrame info function?	D	1		column names	non-null count	data types	column average
7	1	DataFrame	Which of the following is not displayed by the Pandas DataFrame describe function?	C	1		count	mean	correlation	std
8	1	Statistical Analysis	What does it indicate if the corr() function shows correlation as 1 between any two columns of the DataFrame?	A	1		perfect correlation	good correlation	bad correlation	none of these
9	1	Qualitative Data	Which of the following is an example of qualitative data?	B	1		mean	gender	median	mode
10	1	Quantitative Data	Which of the following is an example of quantitative data?	B	1		eye colour	weight	skin colour	names
11	1	Cleaning Data	What does DataFrame.dropna(how='all') do?	A	1		drops those rows from the DataFrame which contain all null values	drops all rows from the DataFrame	drops even numbered rows from the DataFrame	drops odd numbered rows from the DataFrame
12	1	Outliers	A data point that differs significantly from other observations is known as	D	1		mean	median	mode	outlier
13	1	DataFrame	Which of following pandas functions can be used to display the specified number of rows from the beginning of the dataset?	A	1		head()	tail()	begin()	end()
14	1	DataFrame	Which of following pandas functions can be used to display the specified number of rows from the end of the dataset?	B	1		head()	tail()	begin()	end()
15	1	Statistical Analysis	Which of the following represents each data sample as polyline connecting parallel lines where each parallel line represents an attribute of that data sample?	A	1		parallel coordinates	parallelogram	straight lines	long lines
16	1	Cleaning Data	What is the output of the code shown below? import pandas as pd import numpy as np df=pd.DataFrame([[0,1.0,2.0,np.nan,5],[2.0,0,1.0,5.0,np.nan],[5.0,0,1.0,np.nan,5.0]]) df.dropna() print(df.loc[1,3])	B	1		0	5.0	1.0	2.0
17	1	DataFrame	What is the output of the code shown below? import pandas as pd import numpy as np df=pd.DataFrame([[0,1,2,np.nan,5],[2,0,1,5,np.nan],[5,0,1,np.nan,5],[2,0,1,np.nan,np.nan]]) df=df.drop_duplicates(subset=[1,2]) df=df.drop_duplicates(subset=[4]) df.dropna(thresh=2,axis=1) print(df.shape)	C	1		(5, 2)	(2, 3)	(2, 5)	(3, 2)
18	1	Series	What type of Error the following code produces? import pandas as pnd pnd.Series([1,2], index= ['a','b','c'])	D	1		Syntax	Index	Key	Value
19	1	DataFrame	To remove multiple values from the Pandas dataframe and to keep only the first occurrence values, what will be the correct syntax?	C	1		df.drop_duplicate()	df.drop()	df.drop_duplicates()	df.dropduplicates()
20	1	Series	From a Pandas series 's', if we need to extract indices (1,5,7,12) what will be the syntax used?	B	1		s (1,5,7,12)	s ([1,5,7,12])	s.index([1,5,7,12])	s.index(1,5,7,12)
21	1	DataFrame	What is the output for following program? import pandas as pd import numpy as np df=pd.DataFrame([[0,1,2,np.nan,5],[2,0,1,5,np.nan],[5,0,1,np.nan,5]]) print(df.iloc[1,4])	A	1		nan	5	1	None
22	1	DataFrame	What is the method used to calculate the mean of a numeric column in a DataFrame ?	D	1		average()	calculate_mean()	get_mean()	mean()
23	1	DataFrame	What is the output of the below code? import pandas as pd import numpy as np df=pd.DataFrame({"a":[1,2,np.nan,3,4],"b":[1,5,np.nan,2,1]}) df=df.drop_duplicates(subset="b") df.dropna() df.fillna(20,inplace=True) print(df.shape[0])	A	1	LJU 2023	4	3	2	1
24	1	DataFrame	What is the output of the below code? import pandas as pd import numpy as np df=pd.DataFrame([[1,2,3,4,5],[2,1,3,4,5],[np.nan,np.nan,np.nan,np.nan,np.nan]]) df.dropna(thresh=3,axis=1,inplace=True) print(df.shape[1])	A	1	LJU 2023	0	1	2	3
25	1	DataFrame	What is the output of the below code? import pandas as pd import numpy as np df=pd.DataFrame([[1,2,3,4,5],[2,1,3,4,5],[np.nan,np.nan,np.nan,np.nan,np.nan]]) df.drop(1,inplace=True) df=df.dropna() print(df.shape[0])	A	1	LJU 2023	1	2	3	0
26	1	Cleaning Data	Create a Pandas DataFrame from the following table and write code to remove all rows from this table containing at least one NaN value name region sales expenses 0 William NaN 50000.0 42000.0 1 Emma North 52000.0 43000.0 2 Sofia East NaN NaN 3 Markus NaN NaN NaN 4 Edward West 42000.0 38000.0 5 Thomas West 72000.0 39000.0 6 Ethan South 49000.0 42000.0 7 NaN NaN NaN NaN 8 Arun West 67000.0 39000.0 9 Anika East 65000.0 50000.0 10 Paulo South 67000.0 45000.0		3					
27	1	Cleaning Data	Create a Pandas DataFrame from the following table and write code to remove all rows from this table only if all of their values are NaN name region sales expenses 0 William NaN 50000.0 42000.0 1 Emma North 52000.0 43000.0 2 Sofia East NaN NaN 3 Markus NaN NaN NaN 4 Edward West 42000.0 38000.0 5 Thomas West 72000.0 39000.0 6 Ethan South 49000.0 42000.0 7 NaN NaN NaN NaN 8 Arun West 67000.0 39000.0 9 Anika East 65000.0 50000.0 10 Paulo South 67000.0 45000.0		3					
28	1	Cleaning Data	Create a Pandas DataFrame from the following table and write code to drop all columns containing NaN name region sales expenses 0 William NaN 50000.0 42000.0 1 Emma North 52000.0 43000.0 2 Sofia East NaN NaN 3 Markus NaN NaN NaN 4 Edward West 42000.0 38000.0 5 Thomas West 72000.0 39000.0 6 Ethan South 49000.0 42000.0 7 NaN NaN NaN NaN 8 Arun West 67000.0 39000.0 9 Anika East 65000.0 50000.0 10 Paulo South 67000.0 45000.0		3					
29	1	Outliers	Write Python code to remove outliers from any given DataFrame.		4					

30	1	Cleaning Data	Consider the following data:
			data = {
			"A": ["TeamA", "TeamB", "TeamB", "TeamC", "TeamA"],
			"B": [50, 40, 40, 30, 50],
			"C": [True, False, False, False, True]
			}
			
			Convert this to a Pandas DataFrame and remove duplicate rows from it. Reset index values.
					4					

31	1	Cleaning Data	Consider the following autompg dataset: https://raw.githubusercontent.com/Jovita7/Data-Analysis-and-Visualization/main/auto-mpg.csv Write Python code to convert it to a DataFrame and remove mpg and cylinders columns from it		3					
32	1	Statistical Analysis	Use the file heights_weights.csv (https://raw.githubusercontent.com/Jovita7/Data-Analysis-and-Visualization/main/heights_weights.csv) which contains 10000 non-null values for heights and weights. The Male column shows 1 if the person is a Male and 0 if the person is a Female. 1. Convert this file into a pandas Data Frame. (0.5 marks) 2. Display basic information like memory and data types for this data frame. (0.5 marks) 3. Display basic statistics like mean, std, quartiles, etc. for this data frame. (0.5 marks) 4. Create a correlation table for the data frame and comment about what kind of correlation is there between Height and Weight. (0.5 marks) 5. Do Height and Weight contain any outliers? (1 mark)		3					
33	1	Statistical Analysis	Use the file ipl-matches.csv which contains data of all the IPL matches from year 2008 to 2022. Read this csv file and display the basic information like memory and data types for this data frame. Write python code for the following cases: 1. List out all matches gone in superover. 2. How Many Matches won by Chennai Super Kings at Kolkata. 3. In How Many Matches MS Dhoni is Player of Match Vs Mumbai Indians. 4. Display list of all matches in which Gujarat Titans won the Toss and Elected to Bat and won the match. 5. Display list of all matches won by Gujarat Titans.		5					
34	1	Statistical Analysis	Use the file spotify.csv 1. Convert this file into a pandas Data Frame. (0.5 marks) 2. Display basic information like memory and data types for this data frame. (0.5 marks) 3. Display basic statistics like mean, std, quartiles, etc. for this data frame. (0.5 marks) 4. Create a correlation table for the data frame and comment about what kind of correlation is there between danceability and energy (0.5 marks) 5. Display first five rows for this data frame.(0.5 marks) 6. Display last five rows for this data frame. (0.5 marks) 7. Display the rows between 15 to 39 for this data frame. (0.5 marks) 8. Display the data only for last five rows and last five columns for this data frame. (0.5 marks) 9. Display the shape for this data frame. (0.5 marks) 10. Display the sum of NULL values for all the columns. (0.5 marks) 11. Remove first 3 columns from this Data Frame. (0.5 marks) 12. Remove first 10 rows from this Data Frame. (0.5 marks) 13. After removing first 3 columns and first 10 rows from this data frame find outliers for the column popularity. (1 marks) 14. After removing first 3 columns and first 10 rows from this data frame remove outliers for the column energy then display the data frame. (1 marks)		9					
35	1	Statistical Analysis	1. Load the dataset into a pandas DataFrame (data_result.csv) and answer the following questions. 2. View the first few rows of the dataset 3. Check the shape of the dataset 4. View the first last rows of the dataset 5. Get summary statistics of numerical columns 6. Get summary statistics of numerical columns with 0.58 and 0.87 percentiles 7. Get summary statistics of all types of columns 8. Information of all columns 9. Check for missing values 10. Removing duplicates if duplicates 11. List out female students who have greater than 7 spi in all semesters. 12. Find number of students those who have greater than 8 spi in all 5 semesters		9					
36	1	Statistical Analysis	Use the file movies.csv which contains 1629 rows and 18 columns. Read this csv file and display the basic information like memory and data types for this data frame. Write python code for the following cases: 1. List out Movies Released in Year 2019. 2. How Many Movies are having IMDB Rating > 7 (Display Number of Movies). 3. List out the Movies with 'title' and 'story' whose IMDB Votes > 20000. 4. List out Movies Released in Year 2018, Display only Movie Title with Release Date of Year 2018 Movies. 5. Display only Movie Title with its Wikipedia Link.		6					
37	2	Area Plots	Which of the following commands is used to create an area plot in Matplotlib?	C	1		plt.scatter()	plt.area()	plt.fill_between()	plt.plot()
38	2	Data Visualization	Which of the following is not a visualization under matplotlib?	D	1		Scatter Plot	Area Plot	Box Plot	Table Plot
39	2	Data Visualization	Which python package is used for data visualization?	A	1		matplotlib.pyplot	matplotlib.pip	matplotlib.numpy	matplotlib.plt
40	2	Data Visualization	Which of the following commands is used to show a Matplotlib plot in a Jupyter notebook?	C	1		plt.plot()	plt.display()	plt.show()	plt.draw()
41	2	Data Visualization	Plot which is used to give statistical summary is	B	1		Scatter Plot	Box Plots	Bar Plot	Area Plot
42	2	Data Visualization	Which of the following chart element is used to identify data series by its color patterns?	B	1		Data Series	Legend	Title	Markers
43	2	Scatter Plots	Which of the following is best suitable chart to show data correlation?	D	1		Histogram	Bar	Pie	Scatter
44	2	Area Plots	Which of the following parameters is used to specify the transparency of an area plot in Matplotlib?	A	1		alpha	linewidth	color	label
45	2	Area Plots	Which of the following commands is used to create a stacked area plot in Matplotlib?	B	1		plt.plot()	plt.stackplot()	plt.fill_between()	plt.area()
46	2	Box Plots	What type of data is best suited for box plots?	C	1		Categorical data	Binary data	Continuous numerical data	Time-series data
47	2	Box Plots	In a box plot, the bottom line of the box represents which quartile?	A	1	LJU 2023	First quartile	Second quartile	Third quartile	Fourth quartile
48	2	Box Plots	In a box plot, the top line of the box represents which quartile?	C	1		First quartile	Second quartile	Third quartile	Fourth quartile
49	2	Waffle Charts	What is a waffle chart in Python?	D	1		A type of pie chart	A type of stacked bar chart	A type of heatmap	A type of visualization that displays progress towards a goal
50	2	Word Clouds	Which of the statement is true for Word Clouds?	A	1		A graphical representation of the most frequently occurring words in a text corpus	A cloud computing service for analyzing text data	A machine learning algorithm for text classification	A programming language for natural language processing
51	2	Word Clouds	Which of the following types of data is best suited for creating a word cloud?	C	1		Categorical data	Numerical data	Text data	Image data
52	2	Word Clouds	Which of the following parameters in the WordCloud() function is used to set the maximum number of words in the cloud?	A	1		max_words	words	max	word_size
53	2	Word Clouds	Which of the following methods in the WordCloud() function is used to generate the word cloud image?	A	1		generate()	fit()	transform()	predict()
54	2	Word Clouds	What does STOPWORDS contain in wordcloud?	A	1		Words that are used very frequently in a language and have little meaning, such as "the", "is", and "and"	Words that are used very rarely in a language and have little meaning, such as "zephyr", "ebullient", and "myriad"	Words that are used in a specific domain, such as "computer", "internet", and "programming"	Words that are used in formal contexts, such as "therefore", "moreover", and "thus"
55	2	Word Clouds	What is the purpose of removing stopwords from a text before generating a word cloud?	C	1		To improve the readability of the word cloud	To reduce the number of words in the word cloud	To remove words that have little meaning and contribute to noise in the visualization	To highlight the most important words in the word cloud
56	2	Regression Plots	Which Python library is commonly used to create regression plots?	B	1		pandas	seaborn	Matplotlib	NumPy
57	2	Regression Plots	Which type of regression plot is used to visualize the relationship between two continuous variables?	B	1		lmplot	regplot	residplot	jointplot
58	2	Heatmaps	What is a heatmap used for?	B	1		To visualize categorical data	To visualize numerical data in a grid-like format	To fit a regression line to the data	To perform clustering on the data
59	2	Heatmaps	Which parameter in the sns.heatmap() function is used to show numerical values in heatmap?	A	1		annot	annotate	percent	show
60	2	Heatmaps	What is the purpose of the cbar parameter in the sns.heatmap() function?	C	1		To adjust the transparency of the colorbar	To adjust the size of the colorbar	To add a colorbar to the heatmap	To adjust the color scale of the heatmap
61	2	Geospatial Data with Folium	Which of the following methods is used to create a map in Folium?	B	1		folium.create_map()	folium.Map()	folium.make_map()	folium.new_map()
62	2	Geospatial Data with Folium	Which of the following methods is used to add a marker to a map in Folium?	D	1		add_marker()	add_point()	add_location()	add_child()
63	2	Geospatial Data with Folium	Which of the following statements is true about the CircleMarker class in Folium?	A	1		It is used to create a circle markers on a map	It is used to create a polygon markers on a map	It is used to add a single marker to a map	It is not a valid class in Folium
64	2	Choropleth Maps	Which of the following statements is true about the Choropleth class in Folium?	B	1		It is used to create a heatmap	It is used to create a choropleth map	It is used to group markers together	It is not a valid class in Folium
65	2	Choropleth Maps	Which of the following methods is used to add a Choropleth to a map in Folium?	B	1		map.add_choropleth()	Choropleth.add_to(map)	map.add_layer()	Choropleth.add_marker()
66	2	Choropleth Maps	Which of the following methods is used to create a Choropleth map in Folium?	B	1		folium.Map()	folium.Choropleth()	folium.Marker()	folium.Circle()
67	2	NetworkX	Which of the following types of graphs is not supported by NetworkX?	C	1		Directed graphs	Undirected graphs	Hypergraphs	None of the above
68	2	NetworkX	Which of the following methods is used to add nodes to a graph in NetworkX?	A	1		graph.add_node()	graph.add_nodes()	graph.nodes()	graph.node()
69	2	Waffle Charts	To plot a pywaffle chart, what will be the correct syntax used?	A	1		plt.figure(FigureClass=Waffle, rows=10, values=values, labels=labels)	plt.waffle(rows=10, values=values, labels=labels)	plt.pywaffle(rows=10, values=values, labels=labels)	plt.figure(figureclass=Waffle, rows=10, values=values, labels=labels)
70	2		Write a python program which creates following graph using networkx module in python		2					

		NetworkX									
71	2	Box Plots	Create a boxplot of the distribution of temperatures in different cities. Take data from 'temperatures.csv' from below: https://raw.githubusercontent.com/kavit88/Data-Sets/main/temperatures.csv		3						
72	2	NetworkX	The following dictionary shows how five people follow each other on Instagram: instagram = {'person1': [0,1,1,0,1], 'person2': [0,0,1,0,1], 'person3': [1,1,0,1,1],'person4': [1,1,1,0,0], 'person5': [1,1,0,0,0]} E.g., the list for person1 has the value on index 2 as 1 which means person1 followsperson3 and a directed edge should be added from person1 to person3. Using networkx library, create a directed graph.		4						
73	2	Scatter Plots	You have been given a dataset of car prices and their respective horsepower, mileage, and weight. You have been tasked to analyze the relationship between these variables and create a scatter plot to visualize the patterns. Dataset: The dataset, named "car_data.csv" : https://raw.githubusercontent.com/kavit88/Data-Sets/main/car_data.csv		5						
74	2	Scatter Plots	You have been given a dataset of house prices and their respective lot size and square footage. Your task is to create a scatter plot to determine if there is any correlation between these variables. Dataset: The dataset, named "house_data.csv": https://raw.githubusercontent.com/kavit88/Data-Sets/main/house_data.csv		5						
75	2	Data Visualization	Use the file heights_weights.csv which contains 10000 non-null values for heights and weights. The Male column shows 1 if the person is a Male and 0 if the person is a Female. Take file of dataset from: https://raw.githubusercontent.com/kavit88/Data-Sets/main/heights_weights.csv 1. Convert this file into a pandas Data Frame. 2. Display basic information like memory and data types for this data frame. 3. Display basic statistics like mean, std, quartiles, etc. for this data frame. 4. Create a correlation table for the data frame and comment about what kind ofcorrelation is there between Height and Weight. 5. Do Height and Weight contain any outliers? Answer by creating boxplots for both. 6. Finally, create a scatter plot of Weight v/s Height with the following specifications: (i) use + sign, colour green and size 50 for markers. (ii) Label X Axis as Weight and Y Axis as Height. (iii) Display title on top as Weight vs Height		6						
76	2	Area Plots	The file "sales.csv" contains the monthly sales data for a store over a year. Each row contains the month (in the format "yyyy-mm"), the total sales for that month, and the number of items sold. Create a pandas DataFrame from this data and plot the monthly sales using an area plot. Take the dataset from below: https://raw.githubusercontent.com/kavit88/Data-Sets/main/sales.csv		3						
77	2	Regression Plots	The file "survey.csv" contains the results of a survey that asks people how many hours they sleep per night, how much coffee they drink per day, and how many hours they spend exercising per week. Create a pandas DataFrame from this data and plot the relationships between these variables using regression plots. Specifically, create the following plots: 1. A regression plot of hours of sleep versus cups of coffee per day, with a regression line and confidence interval. 2. A regression plot of hours of sleep versus hours of exercise per week, with a regression line and confidence interval. 3. A regression plot of cups of coffee per day versus hours of exercise per week, with a regression line and confidence interval. Label each axis appropriately and give each plot a title. Take Dataset from below: https://raw.githubusercontent.com/kavit88/Data-Sets/main/survey.csv		5						
78	2	Geospatial Data with Folium	Use the California_Houses.csv file to create a map with the first 200 rows using the latitudes and longitudes given in the file with the following customizations: 1. Colour of circle markers should be green with red fill and the type of map should be stamen terrain 2. Add pop up labels using the population from the file. Take the dataset fom below: https://raw.githubusercontent.com/kavit88/Data-Sets/main/California_Houses.csv		4						
79	2	Heatmaps	The file "student_scores.csv" contains the marks scored by a group of students in three subjects: Maths, Science, and English. Each row contains the name of the student, their score in Maths, Science, and English. Create a pandas DataFrame from this data and create a heatmap to visualize the correlations between the scores in these three subjects. Take Dataset from below: https://raw.githubusercontent.com/kavit88/Data-Sets/main/student_scores.csv		3						
80	2	Choropleth Maps	You are given a dataset that contains the unemployment rate of different US states for the year 2021. You have to create a choropleth map of the US using the unemployment rate data. csv file: https://raw.githubusercontent.com/Jovita7/Data-Analysis-and-Visualization/main/US_Unemployment_Oct2012.csv json file: https://raw.githubusercontent.com/Jovita7/Data-Analysis-and-Visualization/main/us-states.json		3						
81	2	Word Clouds	You are given a text file named "speech.txt" which contains the transcript of a speech. You need to create a Word Cloud for the most frequent words used in the speech. https://raw.githubusercontent.com/kavit88/Data-Sets/main/speech.txt		3						
82	2	Word Clouds	You are given a dataset containing customer reviews of a restaurant. Your task is to create a wordcloud of the most frequent words used in the reviews after removing the stopwords. https://raw.githubusercontent.com/kavit88/Data-Sets/main/restaurant_reviews.csv		4						
83	2	Waffle Charts	Suppose you have data on the number of medals won by a country in the 2020 Tokyo Olympics. You want to visualize this data using a waffle chart to show the proportional representation of each country's medal count. Data={'USA': 113, 'China': 88, 'Japan': 58, 'Great Britain': 65, 'ROC': 71, 'Australia': 46, 'Netherlands': 36, 'France': 33, 'Germany': 37, 'Italy': 40}		3	LJU 2023					
84	2	NetworkX	You have been hired as a network analyst by a company to analyze the social network of their employees. The company has provided you with the following data: There are 5 employees in the company, each identified by a unique ID from 1 to 5. The following relationships exist between the employees: 1. Employee 1 is friends with Employee 2 and Employee 3. 2. Employee 2 is friends with Employee 4. 3. Employee 3 is friends with Employee 5. Your task is to create a NetworkX graph representing this social network and display it.		3	LJU 2023					
85	2	Area Plots	Consider the following numpy arrays: Time=np.arange(12) income=np.array([5,9,6,6,10,7,6,4,4,5,6,4]) expense=np.array([6,6,8,3,6,9,7,8,6,6,4,8]) Use Time array for X-axis and create two separate lines in the same graph with income & expense on Y-axis. Give Appropriate labels. Create an area fill graph between the two lines in such a way that where income is more than expense, are filled with Green and areas where expense is more than income are filled with red.		3						

86	2	NetworkX	You have been hired by an Airlines company to analyze their routes. The company has provided you following data. Your task is to create a NetworkX directed graph representing the routes and display it. Figure size should be (15,15), node color should be green, take appropriate node size, edge color should be red. Data: Kolkata to Mumbai Mumbai to Pune Mumbai to Goa Kolkata to Delhi Kolkata to Bhubaneshwar Mumbai to Delhi Delhi to Chandigarh Delhi to Surat Kolkata to Hyderabad Hyderabad to Chennai Chennai to Thiruvananthapuram Thiruvananthapuram to Hyderabad Kolkata to Varanasi Delhi to Varanasi Mumbai to Bangalore Chennai to Bangalore Hyderabad to Bangalore Kolkata to Guwahati		4					
87	2	Statistical Analysis, Visualization	Using 'supermarket_sales.csv' file do the following operations and give required answer by using proper programming process. 1). Load the dataset into a pandas DataFrame and read first 8 rows. 2). Check for missing values and fill it by mean values of that particular column if any. 3). Find the number of orders which have 'Quantity' less than 3 and which have (either 'Rating' greater than 8.5 or 'Total' greater than 600). 4). Find the sum of 'Total' purchasing price spent by Member and Normal 'Customer type'. 5). Find the percentage of total of 'gross income' based on the different 'Payment' methods used by customers. (Ewallet, Cash and Credit card) 6). Analyze the purchasing behavior of male and female customers using 'Gender' column. Find their average purchase prices using 'Total' column. 7). Create a scatter plot that shows the relationship between total amount spent and rating. (keep '+' marker, with marker size 100 and green color). 8). Create a box plot that shows the distribution of 'Rating' and 'Quantity'. And comment about outliers in both columns.		9					
88	2	Statistical Analysis, Visualization	Use the file data.csv which contains 169 rows and 4 columns. 1. Convert this file into pandas Data Frame and Display basic statistics like mean, std, quartiles, etc. for this data frame. 2. Create a correlation table for the data frame and comment about what kind of correlation is there between Duration and Calories? 3. Find whether there any null or NA values, drop all such rows if found in the data frame and print the shape of the data frame after dropping. 4. Prepare a scatter matrix for the following data frame and prepare a parallel coordinates for Duration v/s Pulse, Maxpulse and Calories (all 3 other columns). 5. Do Maxpulse have any outliers? Find using function. 6. Show the outliers using box plot for Maxpulse, width of box plot should be 0.75 and notch should be True. 7. Create a scatter plot for Duration (x-axis) and then Pulse, Maxpulse and Calories (y-axis) with different colors. For each there should be different color and marker.		9					
89	2	Statistical Analysis, Visualization	The dataset provided in 'kc_house_data.csv' contains house sale prices for King County, which includes Seattle. It includes homes sold between May 2014 and May 2015. **Variable** — **Description** id — A notation for a house date — Date house was sold price — Price is prediction target bedrooms — Number of bedrooms bathrooms — Number of bathrooms sqft_living — Square footage of the home sqft_lot — Square footage of the lot floors — Total floors (levels) in house waterfront — House which has a view to a waterfront view — Has been viewed condition — How good the condition is overall grade — overall grade given to the housing unit, based on King County grading system sqft_above — Square footage of house apart from basement sqft_basement — Square footage of the basement yr_built — Built Year yr_renovated — Year when house was renovated zipcode — Zip code lat — Latitude coordinate long — Longitude coordinate sqft_living15 — Living room area in 2015(implies-- some renovations) This might or might not have affected the lotsize area sqft_lot15 — LotSize area in 2015(implies-- some renovations) Perform the following tasks : 1) Load the csv to a dataframe named 'house_survey'. 2) Display the first 5 rows of the dataframe. 3) Display the data types of each column. 4) Obtain a statistical summary of the dataframe. 5) Drop the columns "id" and "Unnamed: 0" 6) Check all the null values present in all the columns of the dataframe. 7) Replace the missing values of the column 'bedrooms' with the mean of the column. 8) Replace the missing values of the column 'bathrooms' with the mean of the column. 9) Count the number of houses with unique floor values. 10) Using boxplot determine whether houses with a waterfront view or without a waterfront view have more price outliers. (Mention your answer as comment in the next cell) 11) Use the function regplot in the seaborn library to determine if the feature sqft_above is negatively or positively correlated with price. (Mention your answer as comment in the next cell). 12) Find the feature other than price that is most correlated with price. (Mention your answer as comment in the next cell)		9					
90	2	Statistical Analysis, Visualization	For the given dataset – iris.csv, perform following exploratory data analysis using python - Use comment feature to answer appropriate questions – a) Load dataset into jupyter notebook using appropriate libraries. Check the datatypes of the dataset attributes. Does the data contain any missing /null values? b) Extract head and tail of the dataset using appropriate methods. c) Summarize statistical figures (i.e. mean, median, percentiles) in one table using appropriate method. d) Create correlation table of all variables. What can you infer about relation between petal length and sepal length? e) Create parallel coordinate plot of iris dataset. What can you infer about petal length and petal width? f) Create box plot of sepal width. Visualizing the plot, answer whether the sepal width data contains any outliers. g) Create cross tabulation of sepal length and petal width attributes. What does the table represent? h) Create scatter matrix of the dataset. i) Create a new column called 'SepalLengthSize' which contains "High" if sepal length ≥ 5 or "Low" if sepal length < 5.		9					
91	2	Pandas,Visulization	To upload the 'diabetes_unclean.csv' to your working folder First import the following libraries import pandas as pd import matplotlib.pyplot as plt import numpy as np 1. Make a data frame with the variable name df 2. To display the specific statistics or measures that are relevant for object-type columns 3. To display the specific statistics or measures that are relevant for numerical-type columns 4. How many rows and columns are in a given dataset 5. To check the missing values 6. To replace the missing values in the column "HbA1c" with their mean value 7. Dropping the missing values of other columns 8. Display the correlation between variables 9. Checking the outliers in the dataset for the following parameters: 'AGE', 'Urea', 'HbA1c', 'Chol', 'TG', 'HDL', 'LDL', 'VLDL', 'BMI' using box plot with labels and title 10. Visualized the "Urea", "HbA1c", "TG" and "BMI" parameters for different ages using parallel_coordinates with labels and title 11. Remove the rows whose gender column has an "f" value and give the frequency count of the "F" and "M" values in different CLASS values 12. Remove the outliers in the "HbA1c" columns and print the shape of the data frame Note: all task output with specific question numbers and follow the sequence Example: print("Ans 1")		9	LJU 2023				
92	3	Regular Expressions	Which module in Python supports regular expressions?	A	1		re	regex	pyregex	None of these
93	3	Python re - split()	What will be the output of the following Python code? re.split('\W+', 'Hello, hello, hello.')	D	1		['Hello', 'hello', 'hello.']	['Hello, 'hello', 'hello']	['Hello', 'hello', 'hello', '.']	['Hello', 'hello', 'hello', '']

94	3	Python re - findall()	What will be the output of the following Python function? re.findall("hello world", "hello", 1)	B	1	LJU 2023	[“hello”]	[]	hello	hello world
95	3	Python re - sub()	What will be the output of the following Python code? re.sub('morning', 'evening', 'good morning')	A	1		‘good evening’	‘good’	‘morning’	‘evening’
96	3	Python re - split()	What will be the output of the following Python code? re.split('mum', 'mumbai*', 1)	B	1		Error	[‘’, ‘bai*’]	[‘’, ‘bai’]	[‘bai*’]
97	3	Python re - split()	What will be the output of the following Python code? re.split(r'(n\d)=', 'n1=3.1, n2=5, n3=4.565')	B	1		Error	[‘’, ‘n1’, ‘3.1, ‘, ‘n2’, ‘5, ‘, ‘n3’, ‘4.565’]	[‘n1’, ‘3.1, ‘, ‘n2’, ‘5, ‘, ‘n3’, ‘4.565’]	[‘3.1, ‘, ‘5, ‘, ‘4.565’]
98	3	Python re - split()	What will be the output of the following Python code? re.split(r'(a)(t)', 'Maths is a difficult subject')	D	1		[‘M a t h s i s a d i f f i c u l t s u b j e c t’]	[‘Maths’, ‘is’, ‘a’, ‘difficult’, ‘subject’]	‘Maths is a difficult subject’	[‘M’, ‘a’, ‘t’, ‘hs is a difficult subject’]
99	3	Python re - split()	What will be the output of the following Python code? re.split(r'\s+', 'Chrome is better than explorer', maxsplit=3)	B	1		[‘Chrome’, ‘is’, ‘better’, ‘than’, ‘explorer’]	[‘Chrome’, ‘is’, ‘better’, ‘than explorer’]	(‘Chrome is’, ‘better’, ‘than explorer’)	‘Chrome is better’ ‘than explorer’
100	3	Python re - sub()	What will be the output of the following Python code? re.sub('Y', 'X', 'AAAAAA', count=2)	D	1		‘YXAAAA’	(‘YXAAAA’)	(‘AAAAAA’)	‘AAAAAA’
101	3	Python re-functions	Which function returns a list containing all matches?	A	1		findall	search	split	find
102	3	Special Sequences	Which character stand for Starts with in regex?	B	1		&	^	#	$
103	3	Metacharacters	Which character stand for Zero or more occurrences in regex?	A	1		*	#	@	|
104	3	Special Sequences	In Regex, s stands for?	C	1		Returns a match where the string DOES NOT contain digits	Returns a match where the string DOES NOT contain a white space character	Returns a match where the string contains a white space character	Returns a match if the specified characters are at the end of the string
105	3	Regular Expressions	Which of the following options is the correct way to import the regex library?	B	1		import regex	import re	import Regex	import Re
106	3	Metacharacters	_____ matches the start of the string. _____ matches the end of the string.	A	1		‘^’, ‘$’	‘$’, ‘^’	‘$’, ‘?’	‘?’, ‘^’
107	3	Metacharacters	What does the command ab+c search for?	C	1		ac,abc,abbc, and so on	ab,abc,abcc and so on	abc,abbc,abbbc and so on	None of the above
108	3	Python re - search()	Which of the following command is used to search a match for 1,2,3,4?	D	1		[1-4]	(1-3)	[1234]	Both A and C
109	3	Python re - split()	What is the output of the code shown below? print(re.split(‘\d’,’abc123xyz’,maxsplit=1))	A	1		['abc', '23xyz']	['abc', '123xyz']	['abc123xyz']	['abc1', '23xyz']
110	3	Python re - sub()	What is the output of the below code? re.sub(‘a’,’u’,’aeiou!’)	A	1		ueiou!'	eiou!'	eio!'	None of these
111	3	Python re - sub()	What is the output of the code shown below? import re text = "Is this Python?" pattern = r'\w{2}\W+[^\W]' result = re.sub(pattern, "**", text) print(result)	C	1		****ython?	Is this Py**	**h**ython?	Is th** ython?
112	3	Python re - sub()	What is the output of the code shown below? import re text = " This is some text with bold and italic text. " pattern = r'<[^>]?>' result = re.sub(pattern, "", text) print(result)	A	1		This is some text with bold and italic text.	This is some text with bold and italic text.	This is some text with bold and italic text.	This is some text with
113	3	Python re - findall()	What will be the output of the following Python code? import re text = "My phone number is 123-456-7890 and my friend's number is 987-654-3210." pattern = r'\d{10}' result = re.findall(pattern, text) print(result)	A	1		[ ]	['123-456-7890', '987-654-3210']	['123-456-7890']	['123-456-7890-987-654-3210']
114	3	Python re - findall()	What will be the output of the following Python code? import re text = "The code is AAA333BB and PQR365RRR." pattern = r'[A-Z]{3}\d{3}[A-Z]{3}' result = re.findall(pattern, text) print(result[0])	D	1		A	PQR	AAA333BB	PQR365RRR
115	3	Python re - search()	What is the output of the following code? import re txt='The rain in Spain' x=re.search('\s',txt) print(x.start())	C	1		1	2	3	4
116	3	Python re - search()	What is the output of the code shown below? import re txt="Today is 31st December 2022" x=re.search('\W+',txt) print(x.start())	D	1		0	4	1	5
117	3	Python re - search()	What is the output of the code shown below? import re txt="8 times before 04:00 PM" x=re.search('\D+',txt) print(x.end())	B	1		6	15	14	8
118	3	Python re - findall()	What is the output of the code shown below? import re txt="That will be 59 dollars till 2000" x=re.findall('\d+',txt) print(x)	B	1		59,2000	[‘59’,’2000’]	[‘59’,’20’,’00’]	59,20,00
119	3	Python re - split()	What is the output for following program? import re text = "The quick brown @fox*jumps#over$the^ lazy&dog." pattern = r'[a-z]+' result = re.split(pattern, text) len(result[0])	B	1		5	1	2	3
120	3	Python re - findall()	What is the output for following program? import re text = "The quick brown fox jumps over the lazy dog." result = re.findall(r'\w{3}', text) result	A	1		['The', 'qui', 'bro', 'fox', 'jum', 'ove', 'the', 'laz', 'dog']	['qui', 'bro', 'fox', 'jum', 'ove', 'the', 'laz',]	['The']	[ ]
121	3	Python re - search()	What is the output for following program? import re text = "the password is p@ssword." pattern = r'[A-Z0-9]+' result = re.search(pattern, text) print(result)	A	1		None	0	error	[ ]
122	3	Python re - findall()	What is the output of the code shown below? import re txt = "08 times before 11:45 AM" x = re.findall("[1-5][0-9]", txt) print(x)	A	1		['11', '45']	[‘08’, '11', '45']	[‘8’, '11', '45']	['11:45']
123	3	Python re - search()	What is the output of the code shown below? import re txt="Hello Good Morning" x=re.search("\s",txt) print(x.start(),x.end())	A	1		5 6	6 7	5 10	‘Hello Morning’
124	3	re.sub()	What is the output of the below code? import re text = "Hello, how are you?" pattern = r'\w{3}\W+' result = re.sub(pattern, "###", text) (result)	A	1	LJU 2023	'He############'	'He###ow ###ou?'	'He###how are you?'	None of these
125	3	re.findall()	What is the output of the below code? import re text = "The quick brown fox jumps over the lazy dog." result = re.findall(r'\w{4}\s', text) print(len(result))	A	1	LJU 2023	5	4	3	2
126	3	re.findall()	What is the output of the below code? import re s = "black, blue and brown" pattern = r'bl\w+\W' matches = re.findall(pattern,s) print(len(matches[0]))	B	1	LJU 2023	5	6	4	3
127	3	re.findall()	What is the output of the below code? import re text = "The code is ABC123XYZ and XYZ789." pattern = r'[A-Z]{3}\d{3}[^\s]{3}' result = re.findall(pattern, text) print(result)	A	1	LJU 2023	['ABC123XYZ']	ABC123XYZ']	[]	[ABC123XYZ]

128	3	re.findall()	What is the output of the below code? import re pattern = r'\d{3}' string = 'The price of the product is 1234 dollars.' match = re.findall(pattern, string) print(match[0])	A	1	LJU 2023	123	1234	12	12
129	3	Regular Expressions	Write a python program to print Phone number from given string using regular expressions.		3					
130	3	Regular Expressions	Write a Python program to check that a string contains only a certain set of characters (in this case a-z, A-Z and 0-9) using regular expressions.		3					
131	3	Regular Expressions	Write a Python program using regular expressions that matches a string that has an a followed by zero or more b's.		4					
132	3	Regular Expressions	Write a Python program that matches a string that has an 'a' followed by one or more b's using regular expressions.		4					
133	3	Regular Expressions	Write a Python program that matches a string that has an 'a' followed by zero or one 'b' using regular expressions.		4					
134	3	Regular Expressions	Write a Python program that matches a string that has an 'a' followed by three 'b' using regular expressions.		4					
135	3	Regular Expressions	Write a Python program to find sequences of lowercase letters joined by an underscore using regular expressions.		4					
136	3	Regular Expressions	Write a Python program to find the sequences of one upper case letter followed by lower case letters using regular expressions.		4					
137	3	Regular Expressions	Write a Python program that matches a word at the end of a string, with optional punctuation using regular expressions.		4					
138	3	Regular Expressions	Write a Python program that matches a word containing 'z' using regular expressions.		4					
139	3	Regular Expressions	Write a Python program to match a string that contains only upper and lowercase letters, numbers, and underscores using regular expressions.		4					
140	3	Regular Expressions	Write a Python program that starts each string with a specific number using regular expressions.		4					
141	3	Regular Expressions	Write a Python program to remove leading zeros from an IP address using regular expressions.		4					
142	3	Regular Expressions	Write a Python program to check for a number at the end of a string using regular expressions.		4					
143	3	Regular Expressions	Write a Python program to search for literal strings within a string using regular expressions.		4					
144	3	Regular Expressions	Write a Python program to extract year, month and date from an URL using regular expressions.		4					
145	3	Regular Expressions	Write a Python program to convert a date of yyyy-mm-dd format to dd-mm-yyyy format using regular expressions.		4					
146	3	Regular Expressions	Write a Python program to find all words starting with 'a' or 'e' in a given string using regular expressions.		4					
147	3	Regular Expressions	Write a Python program to abbreviate 'Road' as 'Rd.' in a given string using regular expressions.		3					
148	3	Regular Expressions	Write a Python program to replace all occurrences of a space, comma, or dot with a colon using regular expressions.		3					
149	3	Regular Expressions	Write a Python program to replace maximum 2 occurrences of space, comma, or dot with a colon using regular expressions.		3					
150	3	Regular Expressions	Write a Python program to convert a camel-case string to a snake-case string using regular expressions.		4					
151	3	Regular Expressions	Write a Python program to remove multiple spaces from a string and store the output in list using regular expressions.		3					
152	3	Regular Expressions	Write a Python program to split a string into uppercase letters using regular expressions.		3					
153	3	Regular Expressions	Write a Python program to remove the parenthesis area in a string.		3					
154	3	Regular Expressions	Write a Python program to insert spaces between words starting with capital letters.		4					
155	3	Regular Expressions	Write a Python program that reads a given expression and evaluates it.		7					
156	3	Regular Expressions	Write a Python program to remove lowercase substrings from a given string.		4					
157	3	Regular Expressions	Write a Python program that checks whether a word starts and ends with a vowel in a given string. Return true if a word matches the condition; otherwise, return false. Sample Data: ("Red Orange White") -> True ("Red White Black") -> False ("abcd dkise eosksu") -> True		4					
158	3	Regular Expressions	Write a Python program that takes a string with some words. For two consecutive words in the said string, check whether the first word ends with a vowel and the next word begins with a vowel. If the program meets the condition, return true, otherwise false. Only one space is allowed between the words. Sample Data: ("These exercises can be used for practice.") -> True ("Following exercises should be removed for practice.") -> False ("I use these stories in my classroom.") -> True		4					
159	3	Regular Expressions	Write a Python Program to find all five-character words in a string. For example: Input : text = 'The quick brown fox jumps over the lazy dog.' Output : ['quick', 'brown', 'jumps']		2					
160	3	Regular Expressions	Write a python program that executes following tasks (strictly using regex module) Given text – " hello welcome to the python exam my email is alice@google.com, world this is bob@meta.com appearing for python exam " a) Remove leading and trailing spaces of the given text. b) Replace space between words of the given text by '$' symbol c) Extract username and host name (i.e. alice,bob,google, meta) in a list		4					
161	3	Regular Expressions	Write a Python Program to find all URLs from a given text. Consider URLs to be of only this format. http://github.com https://github.com Can Start with either http or https followed by :// domain name dot com Example: Text="Hello all Students must visit at my website https://www.pandasrockstar.com for more information. Also, check out http://www.google.com" Output: Found URLs: https://www.pandasrockstar.com		3					
162	4	Basic EDA	Which of the following pandas functions is used to convert categorical data into numeric data?	A	1		get_dummies()	numeric()	get_categorical()	get_dumps()
163	4	Basic EDA	How do you handle missing or corrupted data in a dataset?	D	1		Drop missing rows or columns	Replace missing values with mean/median/mode	Assign a unique category to missing values	All of these
164	4	Basic EDA	What is Scikit-learn?	A	1		A machine learning library in Python	A data visualization library in Python	A natural language processing library in Python	A web development framework in Python
165	4	Basic EDA	Which of the following is an example of a regression algorithm in Scikit-learn?	C	1		K-means clustering	Decision tree	Linear regression	Support vector machines (SVM)
166	4	Basic EDA	How would you access the column "symboling" from the dataframe df?	A	1		df["symboling"]	df=="symboling"	df[:"symboling"]	df[{"symboling"}]
167	4	Basic EDA	What is the correct symbol for missing data?	B	1		na	nan	none	non
168	4	Basic EDA	Why do we convert values of Categorical Variables into numerical values?	A	1		Most statistical models cannot take in objects or strings as inputs	To save memory	To save time	None of these
169	4	Regression	What is the main difference between regression and classification in supervised learning?	A	1		Regression predicts continuous outcomes, while classification predicts categorical outcomes	Regression predicts categorical outcomes, while classification predicts continuous outcomes	Regression uses labeled data, while classification uses unlabeled data	Regression is unsupervised, while classification is supervised
170	4	Regression	What evaluation metric is commonly used for regression tasks?	C	1		Accuracy	Precision	Mean Squared Error (MSE)	Recall
171	4	Regression	What type of target variable is typically used in a regression problem?	C	1		Discrete	Categorical	Continuous	Binary
172	4	Feature Engineering	What is feature selection in supervised learning?	B	1		It is the process of creating new features from existing ones.	It is the process of removing irrelevant or redundant features from the dataset.	It is the process of selecting the target variable for prediction.	It is the process of transforming categorical features into numerical features.
173	4	Feature Engineering	What is feature transformation in machine learning?	D	1		It is the process of creating new features from existing ones.	It is the process of removing irrelevant or redundant features from the dataset.	It is the process of selecting the target variable for prediction.	It is the process of transforming categorical features into numerical features.
174	4	Feature Engineering	You've been given a dataset with apartment area and price information. There's a noticeable non-linear relationship between area and price. To address this you intend to categorize them into 'High', 'Medium', and 'Low' groups. Prices above $3,000,000 are 'High', below $2,000,000 are 'Low', and between $2,000,000 and $3,000,000 are 'Medium'. Write a code to achieve this assuming that dataset has two columns named area and price.		3					

175	4	Feature Engineering	In a survey dataset, you have a column representing participants' ages. You want to categorize ages into 'Young', 'Middle-aged', and 'Elderly' groups. Ages below 30 are 'Young', ages between 30 and 60 are 'Middle-aged', and ages above 60 are 'Elderly'. Write a code to achieve this assuming the dataset has a column named 'age'.		3					
176	4	Feature Engineering	In a customer dataset, you have a column representing customer incomes. You want to categorize incomes into 'Low', 'Medium', and 'High' groups. Incomes below 30000 are 'Low', incomes between 30000 and 70000 are 'Medium', and incomes above 70000 are 'High'. Write a code to achieve this assuming the dataset has a column named 'income'.		3					
177	5	Linear Regression	From where you can import LinearRegression?	C	1		sklearn.metrics	sklearn.linearmodel	sklearn.linear_model	sklearn. model_selection
178	5	Linear Regression	From where you can import train_test_split?	D	1		sklearn.metrics	sklearn.linearmodel	sklearn.linear_model	sklearn. model_selection
179	5	Linear Regression	What is the purpose of the predict() method in sklearn?	B	1		To train a model using a given dataset	To make predictions using a trained model	To evaluate the performance of a model	To split the data in train and test data
180	5	Linear Regression	What is the purpose of the fit() method in sklearn?	A	1		To train a model using a given dataset	To evaluate the performance of a model	To create a plot of predicted values	All of these
181	5	Linear Regression	If we pass x and y to a function train_test_split(), we will get output in which order?	B	1		x_train, y_train, x_test, y_test	x_train, x_test, y_train, y_test	x_train, y_test, x_test, y_train	y_train, y_test, x_train, x_test
182	5	Linear Regression	Consider the following lines of code, what is the name of the column that contains the target values: from sklearn.linear_model import LinearRegression lm=LinearRegression() X = df[['highway-mpg']] Y = df['price'] lm.fit(X, Y) Yhat=lm.predict(X)	A	1		price	highway-mpg	Both A and B	None of these
183	5	Linear Regression	If X is a dataframe with 100 rows and 5 columns, and y is the target with 100 samples, and assuming all the relevant libraries and data have been imported, and the following line of code has been executed: LR = LinearRegression() LR.fit(X, y) yhat = LR.predict(X)	C	1		50	500	100	5
184	5	Linear Regression	What will be the size of training data if data is split like below? train_test_split(x,y,test_size=0.25,random_state=2)	A	1		75%	25%	80%	20%

185	5	Linear Regression	Consider the following code snippet that implements linear regression in Python, what will be printed as the output of the code snippet?
			import numpy as np
			from sklearn.linear_model import LinearRegression
			
			# Training data
			X_train = np.array([[1], [2], [3], [4]])
			y_train = np.array([2, 4, 6, 8])
			
			# Test data
			X_test = np.array([[5]])
			
			# Linear regression model
			model = LinearRegression( )
			model.fit(X_train, y_train)
			predicted_value = model.predict(X_test)
			
			print(predicted_value)
				B	1		10	[10]	5	[5]

186	5	Linear Regression	If the data contains 100 rows and 2 columns and if test_size=0.2 then how many rows will go into training and how many will undergo in testing?	A	1		80,20	70,30	50,60	30,70
187	5	Linear Regression	Consider the following lines of code having 200 non-null data in both x and y. what is the output of following code : import pandas as pd import numpy as np dataset=pd.read_csv("advertising.csv") x=dataset[["TV","Radio","Newspaper"]] y=dataset["Sales"] print(x.shape)	A	1		(200,3)	(200,)	(200,1)	(160,3)
188	5	Linear Regression	Consider the following lines of code having 300 non-null data in both x and y. what is the output of following code : import pandas as pd import numpy as np dataset=pd.read_csv("Book1.csv") x=dataset[["cgpa"]] y=dataset["package"] from sklearn.model_selection import train_test_split x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.2, random_state=1) print(x_train.shape)	C	1		(240,3)	(240,)	(240,1)	(60,)
189	5	Linear Regression	If a dataframe with 400 rows and 5 columns, from the following code how many number of rows will go for x_test? from sklearn.model_selection import train_test_split x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.2, random_state=1)	A	1		80	100	10	200
190	5	Linear Regression	In scikit-learn's linear regression, what is the purpose of the "coef_" attribute?	B	1		It returns the intercept of the linear regression model.	It provides the coefficients of the features in the linear regression model.	It predicts the target variable values for new input data.	It computes the mean squared error (MSE) of the model.
191	5	Regression	What is the purpose of the LinearRegression() function in scikit-learn?	B	1		To perform classification tasks	To fit a linear model to the	To preprocess text data	To plot scatter plots
192	5	Regression	In linear regression, what does the coefficient of determination (R-squared) measure?	D	1		The strength of the relationship between independent and dependent variables	The slope of the regression	The accuracy of the model predictions	The variance explained by the regression model
193	5	Regression	When should you use linear regression for modeling data?	D	1		When the relationship between variables is nonlinear	When the dataset contains categorical variables	When the dependent variable is binary	When there is a linear relationship between independent and dependent variables
194	5	Regression	What does the coefficient of the independent variable (slope) in a simple linear regression model represent?	B	1		The y-intercept of the regression line.	The change in the dependent variable for a unit change in the independent variable.	The standard deviation of the residuals.	The correlation between independent and dependent variables
195	5	Regression	In polynomial regression, what does the degree of the polynomial represent?	B	1		The number of independent variables in the model.	The order of the polynomial curve fitted to the data.	The correlation between independent and dependent variables	The y-intercept of the regression curve.
196	5	Regression	What distinguishes polynomial regression from linear regression?	C	1		Polynomial regression can handle categorical variables.	Polynomial regression only works with two variables.	Polynomial regression fits a curve to the data instead of a straight line.	Polynomial regression always has a higher R-squared value than linear regression.
197	5	Regression	When would you choose polynomial regression over linear regression?	B	1		When there is a linear relationship between variables.	When the data points exhibit a non-linear pattern.	When dealing with categorical variables.	When the dataset contains missing values.
198	5	Regression	Which of the following statements about simple linear regression is true?	D	1		Simple linear regression is used to model the relationship between two categorical variables.	The equation for simple linear regression is $y=\beta0+\beta1x$, where $\beta0$ represents the slope and $\beta1$ represents the y-intercept.	The goal of simple linear regression is to minimize the sum of squared differences between the observed and predicted values of the dependent variable y.	Simple linear regression assumes a curvilinear relationship between the independent and dependent variables.
199	5	Linear Regression	For x = np.array([5, 15, 25, 35, 45, 55]) and y = np.array([5, 20, 14, 32, 22, 38]), apply simple linear regression using scikit learn library and calculate calculate R squared, coeficient and intercept. Predict the y values for x = np.arange(5). (Don't split data for training/testing)		5					
200	5	Linear Regression	Given a dataset with 'SAT' scores as independent variables and 'GPA' as the dependent variable, calculate R squared, coeficient and intercept using linear regression and scikitlearn library. (Don't split data for training/testing)		5					
201	5	Multiple linear Regression	Given a real estate price size year dataset, implement multiple linear regression using scikitlearn library. Using the model, make a prediction about an apartment price with size 750 sq.ft. for 2009.Also Calculate R squared, coeficient and intercept. (Don't split data for training/testing)		5					
202	5	Polynomial Regression	Predict salary based on job position of 6.5 using polynomial regression with a degree of 3 and scikit learn library for the given 'Position_Salaries.csv' dataset. (Don't split data for training/testing)		5					

203	5	Polynomial Regression	For x = np.arange(0, 30) and y = np.array([3, 4, 5, 7, 10, 8, 9, 10, 10, 23, 27, 44, 50, 63, 67, 60, 62, 70, 75, 88, 81, 87, 95, 100, 108, 135, 151, 160, 169, 179]), apply polynomial regression using scikit learn library and calculate R squared, coeficient and intercept. Predict the y values for x = np.arange(5). (Don't split data for training/testing)		5					
204	5	Linear Regression	Write a program to make a model based on linear regression for the following dataframe created from a csv file named "Package.csv" of x and y which follows equation y = a + bx. Write a program which can predict value of y based on any value of x, also write code to find value of a and b in above equation. Given Data in csv file: cgpa: 6.89, 5.12, 7.82, 7.42, 6.94, 7.89, 6.73, 6.75, 6.09, 8.31, 5.32, 6.61, 8.94, 6.93, 7.73 package: 3.26, 1.98, 3.25, 3.67, 3.57, 2.99, 2.60, 2.48, 2.31, 3.51, 1.86, 2.60, 3.65, 2.89, 3.42		3					
205	5	Linear Regression	Write a program to make a model based on linear regression for the following dataframe created from a csv file named "data.csv" of x1 and y which follows equation y = a+bx1. Write a program which can predict value of y based on any value of x, also write code to find value of a and b in above equation. Given Data in csv file: y: 140, 155, 159, 179, 192, 200, 212, 215 X_1: 60, 62, 67, 70, 71, 72, 75, 78		3					
206	5	Linear Regression	Write a program to create a Model using linear regression to predict the price of house using the csv file provided named "Housing.csv". Do the required process in the data before making a model. Find predicted values, co-efficients, intercept and mean squared error.		4					
207	5	Linear Regression	Write a program to create a Model using linear regression to predict the student scores using the csv file provided named "student_scores.csv". Do the required process in the data before making a model. Find predicted values, co-efficients, intercept and mean squared error.		4					
208	5	Linear Regression	Write a program to create a Model using linear regression to predict the gas consumption using the csv file provided named "petrol_consumption.csv". Do the required process in the data before making a model. Find predicted values, co-efficients, intercept and mean squared error.		4					
209	5	Linear Regression	Write a program to create a Model using linear regression to predict the gas consumption using the csv file provided named "FuelConsumptionCo2.csv". Do the required process in the data before making a model. Find predicted values, co-efficients, intercept and mean squared error. (Wherever required remove null values, convert categorical data into numeric data) (Print Output wherever required)		5					
210	5	Linear Regression	For the given RealEstate csv, write a python program satisfying following tasks to demonstrate application of machine learning through multiple linear regression as follows – Given: - Dataset RealEstate.csv ML Library to be used scikit-learn Dependent variable 'Y house price of unit area' Independent variables 'X1 transaction date', 'X2 house age', 'X3 distance to the nearest MRT station', 'X4 number of convenience stores', 'X5 latitude' and 'X6 longitude' 1. Import required libraries. 2. Load RealEstate dataset, create a dataframe and check datatypes of its attributes using appropriate method. 3. Remove 'No' column from the dataframe. 4. Check for any null values in features using appropriate method. 5. Create feature variables x and y as given above. 6. Create training and testing sets of feature variables with 70% of data for training and with random state of 110. 7. Create and fit regression model using appropriate method. 8. Use testing set created in step 6 to find and print the prediction of the outcome. 9. Find and print coefficient and mean squared error of the regression model.		5					
211	5	Linear Regression	Write a program to create a Model using linear regression to predict the charges of insurance using the csv file provided named "insurance.csv". Do the required process in the data before making a model. Find predicted values, co-efficients, intercept and mean squared error.		5					
212	5	Linear Regression	Write a program to create a Model using linear regression to predict the wine quality using the csv file provided named "winequality.csv". Do the required process in the data before making a model. If you find any null value in "winequality.csv" then replace null value with mean value of respected columns. Find co-efficient, intercept and mean squared error. also Predict the quality of red wine for the following data: fixed acidity: 8 volatile acidity: 0.4 citric acid: 0.40 residual sugar: 15 chlorides: 0.048 free sulfur dioxide: 40 total sulfur dioxide: 150 density: 0.99 pH: 3 sulphates: 0.45 alcohol: 10.5		5					
213	5	Linear Regression	Consider variables x and y created from a pandas dataframe "car.csv" . Create new column named "Age_car" (Age_car=2023-year) For multiple linear regression problem, x contains the independent variables (Age_car , Driven_kms , Fuel_Type , Selling_type , Transmission) and y contains the dependent (Selling_Price) variable which is to be predicted. Write a Python program to spilt x and y into training and testing datasets with a 20% split. Then create a multiple linear regression model using the training data and print its coefficients ,intercept and mean squared error.		4					
214	6	kNN	What does kNN stand for?	C	1		K-Neural Networks	K-Means Neighbours	k Nearest Neighbours	K-Cluster Neighbours
215	6	kNN	In the context of kNN, what does 'distance' refer to?	B	1		Geographical distance	Difference in attribute values	Time difference	None of these
216	6	kNN	What is the main disadvantage of a high 'k' value in kNN?	B	1		Overfitting	Underfitting	High bias	None of these
217	6	kNN	What is the main advantage of kNN?	A	1		No assumptions about data	Makes assumptions about data	Prone to overfitting	None of these
218	6	kNN	What is the main disadvantage of kNN?	B	1		No assumptions about data	Sensitive to irrelevant features and the scale of the data	Not prone to overfitting	None of these
219	6	kNN	What does kNN use to make decisions?	B	1		Splitting criteria like entropy or gini index	Distance measures like Euclidean or Manhattan	Similarity measures like cosine similarity	None of these
220	6	kNN	What does the following code snippet represent? from sklearn.neighbors import KNeighborsClassifier knn = KNeighborsClassifier(n_neighbors=5)	B	1		Training a decision tree classifier	Initializing a kNN classifier	Implementing logistic regression	Initializing a random forest classifier

221	6	kNN	What does the following code snippet accomplish? from sklearn.model_selection import train_test_split from sklearn.neighbors import KNeighborsClassifier X_train, X_test, y_train, y_test = train_test_split(features, labels, test_size=0.3, random_state=42) knn = KNeighborsClassifier(n_neighbors=5) knn.fit(X_train, y_train)	B	1		Initializes a kNN classifier with 5 neighbors and trains it on the entire dataset	Splits the dataset into training and testing sets, initializes a kNN classifier with 5 neighbors, and trains it on the training set	Trains a decision tree classifier with Gini index as the criterion	Initializes a logistic regression classifier and evaluates its performance on the test set
222	6	kNN	What class from scikit-learn is used to create a KNN classifier?	B	1		KNeighborsRegressor()	KNeighborsClassifier()	knn_classify()	nearest_neighbors()
223	6	kNN	What metric is used by default in KNeighborsClassifier() to calculate distance between data points?	C	1		Manhattan distance	Chebyshev distance	Euclidean distance	Jaccard similarity
224	6	kNN	What is the primary task of the k-Nearest Neighbors algorithm?	A	1		Classification	Regression	Clustering	Dimensionality reduction
225	6	kNN	In kNN, the value of k represents:	C	1		The number of features in the dataset	The number of clusters in the dataset	The number of nearest neighbors to consider	The number of classes in the dataset
226	6	kNN	What does the "fit" method in scikit-learn's KNeighborsClassifier class do?	A	1		Trains the model	Evaluates the model	Preprocesses the data	Visualizes the data
227	6	kNN	Which parameter of the KNeighborsClassifier determines the number of neighbors to consider?	D	1		neighbors	k_value	k_neighbors	n_neighbors
228	6	kNN	What happens if the value of k in kNN is too small?	A	1		Overfitting	Underfitting	Both overfitting and underfitting	No impact on model performance
229	6	kNN	Which of the following scenarios is an example where K-Nearest Neighbors (KNN) algorithm is not suitable?	A	1		Image classification with high-resolution images.	Fraud detection in credit card transactions.	Sentiment analysis of text data.	Speech recognition for voice commands.
230	6	DecisionTree	What is the primary criterion for a decision tree using entropy?	A	1		Information Gain	Gini Index	Chi-Square	Reduction in Variance
231	6	DecisionTree	What is entropy in the context of a decision tree?	D	1		A measure of impurity or disorder	A measure of similarity	A measure of distance	A measure of impurity or disorder
232	6	DecisionTree	What does a decision tree do?	C	1		It makes decisions	It predicts continuous outcomes	It classifies data into different classes	None of these
233	6	DecisionTree	What is the disadvantage of a decision tree?	A	1		Prone to overfitting	Prone to underfitting	Not sensitive to outliers	None of these
234	6	DecisionTree	What does a decision tree use to make decisions?	A	1		Splitting criteria like entropy or gini index	Distance measures like Euclidean or Manhattan	Similarity measures like cosine similarity	None of these
235	6	Decision Tree	What is the primary task of the Decision Tree algorithm?	A	1		Classification	Regression	Clustering	Dimensionality reduction
236	6	DecisionTree	What is a leaf node in a decision treeWhat is a leaf node in a decision tree?	A	1		A node with no children that contains the class label	A node with children that contains the class label	A node with no children that contains the splitting criterion	None of these
237	6	DecisionTree	What is a root node in a decision tree?	B	1		A node with no children that contains the class label	A node with children that contains the class label	A node with no children that contains the splitting criterion	None of these
238	6	DecisionTree	Which of the following measures is used to quantify the randomness in a decision tree?	B	1		Variance	Entropy	Standard Deviation	Mean Absolute Error
239	6	DecisionTree	What is the purpose of the following code snippet? from sklearn.tree import DecisionTreeClassifier dt_classifier = DecisionTreeClassifier(criterion='entropy')	C	1		Initializing a decision tree classifier with Gini index	Initializing a kNN classifier with Euclidean distance	Initializing a decision tree classifier with entropy as the criterion	Initializing a logistic regression classifier
240	6	ConfusionMatrix	What does the confusion matrix evaluate in classification models?	C	1		Precision and Recall	Accuracy and Error Rate	Sensitivity and Specificity	F1 Score and ROC Curve
241	6	ConfusionMatrix	How is accuracy calculated in the context of a confusion matrix?	A	1		(True Positives + True Negatives) / Total Predictions	True Positives / (True Positives + False Positives)	(True Positives + True Negatives) / Total Actual Positives	True Negatives / (True Negatives + False Negatives)
242	6	ConfusionMatrix	Which metric from the confusion matrix reflects the proportion of correctly classified negative instances?	C	1		Accuracy	Sensitivity	Specificity	Error Rate
243	6	ConfusionMatrix	In a confusion matrix, what does the false positive rate represent?	B	1		Proportion of correctly classified negative instances	Proportion of incorrectly classified positive instances	Proportion of correctly classified positive instances	Proportion of incorrectly classified negative instances
244	6	ConfusionMatrix	Which metric from the confusion matrix focuses on the ability of the model to correctly identify positive instances?	B	1		Accuracy	Sensitivity	Specificity	Error Rate
245	6	ConfusionMatrix	Which of the following best describes sensitivity?	A	1		Proportion of correctly classified positive instances	Proportion of correctly classified negative instances	Proportion of incorrectly classified positive instances	Proportion of incorrectly classified negative instances
246	6	ConfusionMatrix	How is specificity calculated in the context of a confusion matrix?	A	1		True Negatives / (True Negatives + False Positives)	True Positives / (True Positives + False Negatives)	(True Positives + True Negatives) / Total Actual Positives	(True Positives + True Negatives) / Total Predictions
247	6	ConfusionMatrix	In a confusion matrix, what does the false negative rate represent?	B	1		Proportion of correctly classified negative instances	Proportion of incorrectly classified positive instances	Proportion of correctly classified positive instances	Proportion of incorrectly classified negative instances
248	6	ConfusionMatrix	How is the error rate calculated from a confusion matrix?	D	1		True Negatives / (True Negatives + False Positives)	(True Positives + True Negatives) / Total Actual Positives	True Positives / (True Positives + False Negatives)	(False Positives + False Negatives) / Total Predictions
249	6	ConfusionMatrix	Which metric from the confusion matrix focuses on the ability of the model to correctly identify negative instances?	C	1		Accuracy	Sensitivity	Specificity	Error Rate
250	6	ConfusionMatrix	How is specificity calculated in the context of a confusion matrix?	A	1		True Negatives / (True Negatives + False Positives)	True Positives / (True Positives + False Negatives)	(True Positives + True Negatives) / Total Actual Positives	(True Positives + True Negatives) / Total Predictions
251	6	ConfusionMatrix	What is the purpose of the following code snippet? from sklearn.metrics import confusion_matrix conf_matrix = confusion_matrix(true_labels, predicted_labels)	A	1		Evaluating the confusion matrix	Training a decision tree classifier	Implementing kNN algorithm	Tuning hyperparameters for a random forest classifier
252	6	ConfusionMatrix	What does the following code snippet accomplish? print(conf_matrix[0, 0] / (conf_matrix[0, 0] + conf_matrix[0, 1]))	B	1		Calculating the sensitivity of the classifier	Calculating the specificity of the classifier	Evaluating the F1 score	Printing the accuracy score of the model
253	6	ConfusionMatrix	What is the purpose of the following code snippet? from sklearn.tree import DecisionTreeClassifier from sklearn.metrics import confusion_matrix dt_classifier = DecisionTreeClassifier(criterion='entropy') dt_classifier.fit(X_train, y_train) predicted_labels = dt_classifier.predict(X_test) conf_matrix = confusion_matrix(y_test, predicted_labels)	A	1		Initializes a decision tree classifier with entropy as the criterion and evaluates its performance using a confusion matrix	Initializes a kNN classifier with 3 neighbors and evaluates its performance using a confusion matrix	Initializes a logistic regression classifier and evaluates its performance using a confusion matrix	Trains a random forest classifier and evaluates its performance using a confusion matrix
254	6	ConfusionMatrix	Write Python code to train a kNN classifier using the following steps: Split the dataset X into training and testing sets with a test size of 0.3 and a random state of 42. Initialize a kNN classifier with 5 neighbors. Train the classifier on the training set. Make predictions on the test set. Calculate and print the accuracy score of the classifier.		4					
255	6	ConfusionMatrix	Write Python code to train a decision tree classifier with entropy as the criterion using the following steps: Initialize a Decision Tree classifier with entropy as the criterion. Train the classifier on the training set. Make predictions on the test set. Calculate and print the confusion matrix for the classifier.		3					
256	6	ConfusionMatrix	Write Python code to evaluate the performance of a classification model using the following steps: Import the necessary functions from sklearn.metrics. Calculate and print the classification report for the true labels and predicted labels. Calculate and print the accuracy score of the classifier.		4					
257	6	ConfusionMatrix	Using the Iris dataset (Iris.csv), write Python code to perform the following tasks: Split the dataset into features (X) and labels (y). Split the features and labels into training and testing sets with a test size of 0.2 and a random state of 42. Initialize a kNN classifier with 3 neighbors. Train the classifier on the training set. Make predictions on the test set. Calculate and print the accuracy score of the classifier.		4					
258	6	kNN	You are tasked with using the k-Nearest Neighbors (kNN) algorithm to classify whether patients have diabetes or not based on certain diagnostic measurements. You have been provided with diabetes.csv file. The datasets consist of several medical predictor (independent) variables and one target (dependent) variable, Outcome. Independent variables include the number of pregnancies the patient has had, their BMI, insulin level, age, and so on. Also perform Model Performance Analysis using confusion matrix.		7					

259	6	kNN	The objective is to perform classification on the Iris dataset using the k-Nearest Neighbors (kNN) algorithm. The Iris dataset contains measurements of various iris flowers, including features such as sepal length, sepal width, petal length, and petal width, along with the corresponding species label. The problem involves two main tasks: Build a kNN classification model to predict the species of iris flowers based on their feature measurements. Train the model on a portion of the dataset and evaluate its performance on another portion to assess its accuracy. Experiment with different values of k and choose the optimal value that maximizes the model's performance. Use appropriate evaluation confusion matrix to evaluate the model's performance. Also calcualte accuracy, sensitivity and specificity. Use iris.csv file for dataset.		9					
260	6	kNN	Given the Breast Cancer Wisconsin (Diagnostic) dataset, the objective is to build a kNN classification model that accurately predicts whether a tumor is benign or malignant based on the diagnostic features provided. The model should be trained on a portion of the dataset and evaluated on another portion to assess its performance. The ultimate goal is to create a reliable classifier that can assist healthcare professionals in diagnosing breast cancer accurately and early. Use cancer.csv file for dataset.		5					
261	6	kNN	Given the credit card transaction dataset, the objective is to build a kNN classification model that accurately predicts whether a transaction is fraudulent or non-fraudulent based on the transaction features provided. The model should be trained on historical transaction data and evaluated on another portion of the dataset to assess its performance. The ultimate goal is to create a reliable classifier that can automatically detect fraudulent transactions and prevent financial losses for credit card companies and cardholders. Use card_transdata.csv for dataset.		5					
262	6	kNN	The task involves building a k-Nearest Neighbors (kNN) regression model to predict the Air Quality Index (AQI) based on the latitude and longitude coordinates of various countries. The dataset used for this task contains information about the AQI levels and geographic locations (latitude and longitude) of different countries. The AQI serves as an indicator of air quality, with higher values indicating poorer air quality and vice versa. Use AQI and Lat Long of Countries.csv for dataset.		5					
263	6	Decision Tree	The task involves building a Decision Tree classifier to predict whether to play tennis based on weather conditions. The dataset used for this task is the PlayTennis dataset, which contains information about various weather attributes such as outlook, temperature, humidity, and wind, along with the corresponding decision to play tennis or not. Use PlayTennis.csv for dataset.		6					
264	6	Decision Tree	Imagine that you are a medical researcher compiling data for a study. You have collected data about a set of patients, all of whom suffered from the same illness. During their course of treatment, each patient responded to one of 5 medications, Drug A, Drug B, Drug c, Drug x and y. Part of your job is to build a model to find out which drug might be appropriate for a future patient with the same illness. The feature sets of this dataset are Age, Sex, Blood Pressure, and Cholesterol of patients, and the target is the drug that each patient responded to. It is a sample of multiclass classifier, and you can use the training part of the dataset to build a decision tree, and then use it to predict the class of a unknown patient, or to prescribe it to a new patient. Use drug200.csv for dataset.		7					
265	7	Tensorflow	What is TensorFlow?	A	1		A machine learning library	A programming language	A deep learning framework	A database management system
266	7	Tensorflow	What is a tensor in TensorFlow?	A	1		A type of data structure	A machine learning model	A database management system	A programming language
267	7	Tensorflow	What is the default data type of TensorFlow tensors?	B	1		int64	float32	double	int32
268	7	Tensorflow	Which of the following is NOT a valid TensorFlow data type?	D	1		int32	bool	float16	char
269	7	Tensorflow	What is a placeholder in TensorFlow?	B	1		A variable that holds the output of a neural network	A variable that holds the input data for a neural network	A variable that holds the weights of a neural network	A variable that holds the bias of a neural network
270	7	Tensorflow	What is a variable in TensorFlow?	C	1		A fixed value that is used in a neural network	A data structure that holds the input data for a neural network	A data structure that holds the weights and biases of a neural network	A fixed value that is used to compute the output of a neural network
271	7	Tensorflow	What is a computation graph in TensorFlow?	B	1		A graph that represents the structure of a neural network	A graph that represents the mathematical operations performed by a neural network	A graph that represents the input and output of a neural network	A graph that represents the weights and biases of a neural network
272	7	Tensorflow	What is a tensor in TensorFlow?	C	1		A function that maps one tensor to another	A mathematical operation that can be applied to tensors	A multi-dimensional array	A unit of computation in a neural network
273	7	Tensorflow	Can we have multidimentional tensors	C	1		No tensor can have maximum two dimentions	Possible only in image data	Yes possible	Possible only in geo tagged data
274	7	Tensorflow	Why Tensorflow uses computational graphs?	D	1		Tensors are nothing but computational graphs	Gaphs are easy to plot	There is no such concept of computational graphs in TensorFlow	Calculations can be done in parallel
275	7	Tensorflow	Which of the following statements is not true about TensorFlow?	C	1		TensorFlow is an open-source library for machine learning and deep learning developed by Google.	TensorFlow provides support for both CPUs and GPUs to accelerate training and inference.	TensorFlow is limited to Python and does not support other programming languages for development.	TensorFlow can be used for building and training neural networks for a variety of tasks, including image classification, natural language processing, and reinforcement learning.
276	7	Basic TensorFlow operations	Which TensorFlow function is used to create a tensor with constant values?	B	1		tf.Variable()	tf.constant()	tf.placeholder()	tf.zeros()
277	7	Tensorflow	Which of the following methods is used to create a TensorFlow constant tensor?	B	1		tf.Variable()	tf.constant()	tf.placeholder()	tf.convert_to_tensor()
278	7	Tensorflow	What is the primary purpose of TensorFlow?	B	1		To perform simple arithmetic operations.	To facilitate high-performance numerical computation and machine learning.	To create graphical user interfaces.	To manage databases and perform SQL queries.
279	7	Tensorflow	From the following which instruction is used to print the version number of the TensorFlow library that is currently installed?	C	1		tf.version()	tensorflow.version.VERSION	tf.__version__	tf.get_version()
280	7	Basic TensorFlow operations	How can you convert a NumPy array to a TensorFlow tensor?	A	1		tf.convert_to_tensor()	tf.tensor()	tf.from_numpy()	tf.array()
281	7	Basic TensorFlow operations	What does the tf.zeros() function create?	C	1		A tensor with all elements set to one.	A tensor with all elements set to a specified value.	A tensor with all elements initialized to zero.	A tensor with random values.
282	7	Basic TensorFlow operations	Which TensorFlow method is used to create a tensor of ones?	C	1		tf.ones_like()	tf.zeros()	tf.ones()	tf.fill()
283	7	Basic TensorFlow	Which function is used to create a TensorFlow variable?	A	1		tf.Variable()	tf.constant()	tf.placeholder()	tf.random()
284	7	Basic TensorFlow operations	How can you initialize a tensor with random values in TensorFlow?	D	1		tf.random.normal()	tf.random.uniform()	tf.zeros()	Both A and B
285	7	Basic TensorFlow operations	Write a Python program that performs the following tasks: 1. Generates random positions for 10 nodes within a 2D space using TensorFlow. 2. Converts the TensorFlow tensor to a NumPy array for plotting. 3. Defines a set of edges to connect these nodes in a graph. 4.Uses Matplotlib to create a plot that visualizes the nodes and edges, where: Nodes are represented by red dots with black edges. Edges are plotted as blue lines. Each node is annotated with its index. 5. Customizes the plot with a title and removes the axis.		5					

286	7	Basic TensorFlow operations	Write a Python script that performs the following tasks: 1. Generates random positions for 8 nodes in a 2D space using TensorFlow. 2. Converts the TensorFlow tensor of node positions into a NumPy array for plotting. 3. Defines a set of directed edges to connect these nodes in a graph. 4. Uses Matplotlib to create a plot that visualizes the directed graph where: Directed edges are represented by arrows. Nodes are shown as red dots with black edges. Each node is annotated with its index. 5. Customizes the plot with a title and removes the axis.		5					
287	7	Basic TensorFlow operations	Write a Python program that performs the following tasks to visualize a tree graph: 1. Generates random positions for nodes in a tree structure using TensorFlow. The tree should have 4 levels. 2. Converts the TensorFlow tensors of node positions into a NumPy array for plotting. 3. Defines edges for the tree, connecting nodes from each level to their parent nodes. 4. Uses Matplotlib to create a plot that visualizes the tree graph where: Edges are represented by blue lines. Nodes are shown as red dots with black edges. Each node is annotated with its index. 5. Customizes the plot with a title and removes the axis.		5					
288	7	Basic TensorFlow operations	Write a python code to generate 3D variable using tensorflow		2					
289	7	Basic TensorFlow operations	Write a python code that Creates a 3D tensor with random values from a normal distribution.		2					
290	8	Working with tensors	What does the tf.shape() function return?	C	1		The data type of a tensor	The total number of elements in a tensor	The dimensions of a tensor	The mean value of a tensor
291	8	Working with tensors	Given the tensor t = tf.constant([[[1, 2, 3], [4, 5, 6]], [[7, 8, 9], [10, 11, 12]]]), what is the result of tf.slice(t, [1, 0, 1], [1, 2, 2])?	C	1		[[[7, 8], [10, 11]]]	[[[8, 9]], [[11, 12]]]	[[[8, 9], [11, 12]]]	[[[2, 3], [5, 6]]]
292	8	Working with tensors	Given the tensor t = tf.constant([[1, 2, 3, 4], [5, 6, 7, 8]]), what is the result of tf.slice(t, [0, 1], [1, 2])?	B	1		[[1, 2]]	[[2, 3]]	[[2, 3], [6, 7]]	[[2, 3, 4], [6, 7, 8]]
293	8	Working with tensors	What happens if the indices tensor contains indices that are out of bounds of the dimension specified by axis in tf.gather?	B	1		TensorFlow will automatically adjust the indices to be within bounds.	An error is raised indicating out-of-bounds indices.	TensorFlow will wrap around the indices, effectively implementing modulo indexing.	TensorFlow will ignore out-of-bounds indices and return only valid indices.
294	8	Working with tensors	Consider the following code: import tensorflow as tf t = tf.constant([[1, 2, 3], [4, 5, 6]]) indices = tf.constant([2, 0]) result = tf.gather(t, indices, axis=1) What is the result?	D	1		[[1, 2], [4, 5]]	[[2, 1, 3], [5, 4, 6]]	[[3, 2, 1], [6, 5, 4]]	[[3, 1], [6, 4]]
295	8	Working with tensors	If you use tf.gather with an axis of 1, what does it imply?	B	1		Gathering along the rows of the tensor.	Gathering along the columns of the tensor.	Gathering along the depth of the tensor.	Gathering along a specific time step in a sequence.
296	8	Working with tensors	What will be the output of the following code? import tensorflow as tf t = tf.constant([[10, 20, 30], [40, 50, 60]]) indices = tf.constant([1]) result = tf.gather(t, indices, axis=0)	B	1		[[10, 20, 30]]	[[40, 50, 60]]	[10, 20, 30]	[40, 50, 60]
297	8	Working with tensors	In the tf.gather function, what does the axis argument specify?	B	1		The axis along which to concatenate the tensors.	The axis along which to gather slices from the tensor.	The axis along which to transpose the tensor.	The axis along which to compute the sum of elements.
298	8	Working with tensors	Given the tensor t = tf.constant([[1, 2, 3], [4, 5, 6]]), what is the result of tf.gather(t, [0])?	A	1		[[1, 2, 3]]	[[4, 5, 6]]	[1, 2, 3]	[[1], [4]]
299	8	Working with tensors	What does the tf.gather function do in TensorFlow?	B	1		Combines tensors by concatenation along a specified axis.	Gathers slices of a tensor according to specified indices along a specified axis.	Computes the element-wise multiplication of two tensors.	ransposes the dimensions of a tensor.
300	8	TensorFlow's gradient descent	What is a loss function in TensorFlow?	A	1		A function that measures the difference between the predicted output and the actual output	A function that computes the gradient of a neural network	A function that updates the weights and biases of a neural network	A function that initializes the weights and biases of a neural network
301	8	TensorFlow's gradient descent	What is an optimizer in TensorFlow?	C	1		A function that measures the difference between the predicted output and the actual output	A function that computes the gradient of a neural network	A function that updates the weights and biases of a neural network	A function that initializes the weights and biases of a neural network
302	8	TensorFlow's gradient descent	What is backpropagation in TensorFlow?	A	1		A process of computing the gradient of a neural network	A process of initializing the weights and biases of a neural network	A process of updating the weights and biases of a neural network	A process of measuring the difference between the predicted output and the actual output
303	8	TensorFlow's gradient descent	What is data augmentation in TensorFlow?	A	1		A technique for generating new data from existing data	A technique for reducing the size of the input data	A technique for increasing the complexity of the model	A technique for regularizing the model
304	8	Mathematical operations using TensorFlow	The following instruction indicates tf.reduce_sum(tf.square(out-Y))	C	1		Linear Model equation	Maximum Entropy loss function	Squared Error loss function	Feed_dict process
305	8	TensorFlow's automatic differentiation	What does tf.GradientTape do in TensorFlow?	B	1		It saves the state of the model during training.	It records operations for automatic differentiation to compute gradients.	It visualizes the model's performance metrics.	It executes TensorFlow operations on a specific device.
306	8	TensorFlow's automatic differentiation	What is the use of tf.gradients() in TensorFlow?	A	1		To calculate the gradients of a tensor with respect to a set of inputs.	To perform data augmentation.	To initialize model weights.	To create a dataset from a tensor.
307	8	Mathematical operations using TensorFlow	What does tf.math.argmax() return?	A	1		The index of the maximum value along a specified axis.	The maximum value in the tensor.	The index of the minimum value along a specified axis.	The sum of all elements in the tensor.
308	8	Mathematical operations using TensorFlow	How can you set the random seed for reproducibility in TensorFlow?	A	1		tf.random.set_seed()	tf.seed()	tf.random.initialize()	tf.config.set_seed()
309	8	Mathematical operations using TensorFlow	In TensorFlow, how can you perform element-wise multiplication of two tensors?	A	1		tf.multiply()	tf.matmul()	tf.mul()	None of these
310	8	Mathematical operations using TensorFlow	Which function is used to concatenate tensors along a specified axis in TensorFlow?	A	1		tf.concat()	tf.stack()	tf.split()	tf.merge()
311	8	Mathematical operations using TensorFlow	What does the tf.transpose() function do?	A	1		Reorders the dimensions of a tensor.	Transposes the tensor to its inverse.	Converts a tensor to a NumPy array.	Normalizes the tensor values.
312	8	Mathematical operations using TensorFlow	How can you perform a tensor reduction operation along a specific axis in TensorFlow?	D	1		tf.reduce_sum()	tf.reduce_mean()	tf.reduce_max()	All of the above
313	8	Mathematical operations using TensorFlow	What is the purpose of the tf.math module in TensorFlow?	B	1		To create and manage TensorFlow models.	To perform mathematical operations on tensors.	To save and load TensorFlow models.	To visualize data and results.
314	8	Mathematical operations using TensorFlow	What does the tf.reduce_mean() function compute?	C	1		The maximum value in a tensor.	The minimum value in a tensor.	The mean (average) value of elements in a tensor.	The sum of elements in a tensor.
315	8	Mathematical operations using TensorFlow	Which TensorFlow function would you use to compute the dot product of two tensors?	A	1		tf.matmul()	tf.add()	tf.reduce_sum()	tf.multiply()
316	8	Mathematical operations using TensorFlow	Which function in TensorFlow is used to perform element-wise addition between two tensors?	A	1		tf.add()	tf.concat()	tf.multiply()	tf.reshape()
317	8	Mathematical operations using TensorFlow	Implement a python program that performs basic arithmetic operations using TensorFlow.		3					
318	8	Mathematical operations using TensorFlow	Implement a python program that performs matrix operations using TensorFlow.		3					
319	8	Mathematical operations using TensorFlow	Write python to find the number of zeros at the end of a factorial of a given positive number using tensorflow.		3					

320	8	Mathematical operations using TensorFlow	Using tensorflow Write a Python function to check whether a number is perfect or not. The first perfect number is 6, because 1, 2, and 3 are its proper positive divisors, and 1 + 2 + 3 = 6. Equivalently, the number 6 is equal to half the sum of all its positive divisors: (1 + 2 + 3 + 6) / 2 = 6. The next perfect number is 28 = 1 + 2 + 4 + 7 + 14. This is followed by the perfect numbers 496 and 8128.		5					
321	8	Mathematical operations using TensorFlow	Using tensorflow Write a python program to check the validity of password without using any built-in functions or modules. Password checker program basically checks if a password is valid or not based on the password policies mention below: • Password should contain at least one lowercase letter(a-z). • Password should contain at least one uppercase letter(A-Z). • Password should contain at least one special character (@, #, %, &, !, $, etc…). • Password should not contain any space. • Password should contain at least one digit (0-9). • Password length should be between 8 to 15 characters. • It should not contain the repeated combination of consecutive 3 characters. Also check whether the password is strong or not without using any built-in function. Given a string, find its strength. Let a strong password is one that satisfies all above conditions. A moderate password is one that satisfies first three conditions and has length at least 6. Otherwise password is weak.		5					
322	8	Mathematical operations using TensorFlow	Using tensorflow write a python code to Take that list and finds all pairs of integers that differ by three.		5					
323	8	Mathematical operations using TensorFlow	Using tensorflow write a python code to Return all pairs of integers in a list. Also do the sum of missing numbers of that list of integers.		5					
324	8	Mathematical operations using TensorFlow	Using tensorflow write a python code to add all elements of that list of integers except the number at index. Return the new string.		5					
325	8	Mathematical operations using TensorFlow	Using tensorflow write a python code to find an element that divides a given list of integers with the same sum value.		5					
326	8	Mathematical operations using TensorFlow	Write a python code to calculate Square and Cube Using TensorFlow		3					
327	8	Mathematical operations using TensorFlow	Write a python code for solving a Linear Equation ax+b=0 Using TensorFlow		3					
328	8	TensorFlow's automatic differentiation	Write a python code to compute the gradient of a function x^2+3x-4 using tensorflow.		5					
329	8	TensorFlow's gradient descent	Write a Python code that demonstrates the implementation of a basic linear regression model using TensorFlow, including the optimization of model parameters using gradient descent.		5					