|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1. a) | Create two series as shown using pd.series() function.  Series\_A = [10,20,30,40,50] Series\_B = [40,50,60,70,80]. Get the items common to both. | | | | | |
| b) | Describe the different types of machine learning algorithms with examples. | | | | | |
| 2. a) | What are the goals of artificial intelligence? | | | | | |
| b) | Create a data frame with following data | | | | | |
| Ename | Type | Dname | exp | salary |  |
| Roshan | regular | cs | 10 | 50000 |
| Amar | adhoc | cs | 20 | 15000 |
| Ashwini | regular | ec | 5 | 30000 |
| Lohith | adhoc | ec | 14 | 15000 |
| Mohan | contract | cs | 9 | 10000 |
| Pramod | regular | ec | 8 | 40000 |
| 1. Make a pivot table that shows the average salary of each employee for each department. 2. Make a pivot table that shows the sum and mean of the salaries of each type of employee and the number of employees of each type. | | | | | |

|  |  |
| --- | --- |
| 3) a | Use 'Cars93' dataset to answer the above questions.  The information that the columns of this dataset contain is given below:    Create the following plots to visualize/summarize the data and customize it appropriately.   * Use a box plot to determine the price range of all different cars available in the market? And interpret the five-number summary * Histogram to check the frequency distribution of the variable ‘Mpg.city’ (Miles per gallon) and note down the interval having the highest frequency. * Use a scatter plot to determine whether a car with higher horsepower gives lower mileage? * Use a line chart to observe the variations in ‘Engine Size', against 'Horsepower'. * Create a git repository and push source code to the repo. |
| b | 1. find a list of squares of the first five odd numbers using lambda and map function. 2. find the odd numbers from a given list using a filter 3. compute a sum of the first five integers using reduce function. |

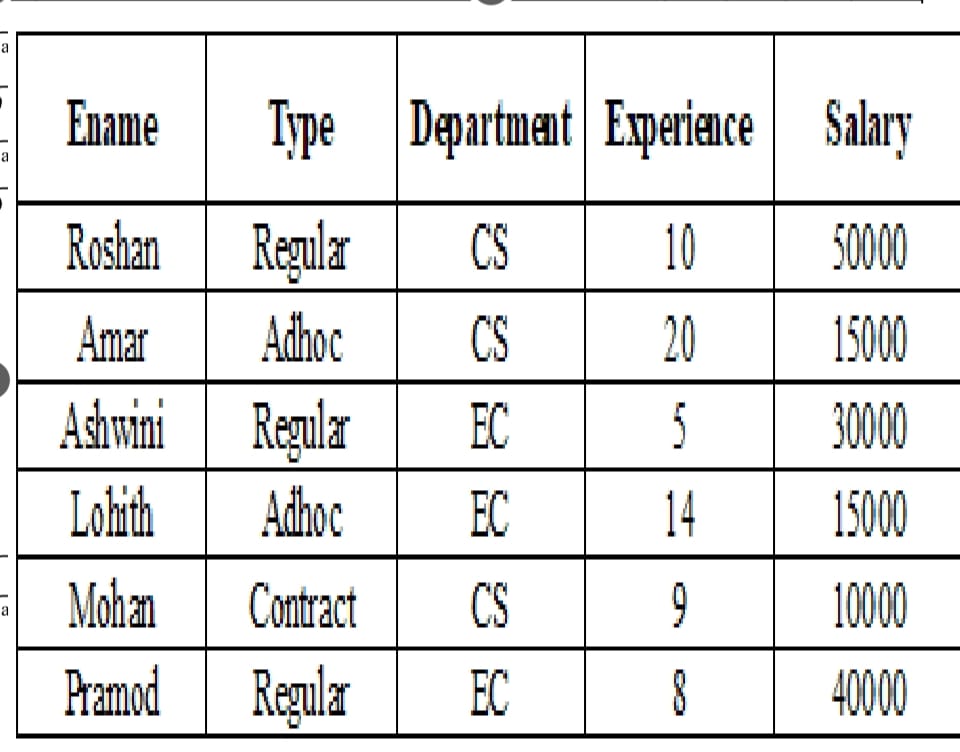
|  |  |
| --- | --- |
| 4) a) | Use the 'matcars.csv’ dataset to answer the above questions.  Create the following plots to visualize/summarize the data and customize it appropriately.   * histogram to check the frequency distribution of the variable ‘mpg’ (Miles per gallon)   and note down the interval having the highestfrequency.   * scatter plot to determine the relationship between the weight of the car and the mpg * bar plot to check the frequency distribution of transmission type of cars. * Box and Whisker plot of mpg and interpret the five-number summary. * Create a git repository and push source code to the repo. |
| b) | Write a Pandas program to split a given dataframe into groups and create a new column with count from GroupBy.  **Test Data:**  book\_name book\_type book\_id   1. Book1 Math 1 2. Book2 Physics 2 3. Book3 Computer 3 4. Book4 Science 4 5. Book1 Math 1 6. Book2 Physics 2 7. Book3 Computer 3 8. Book5 English 5 |

5.Create an array and apply all aggregation function to it

6.Find odd and even number from a given list using a filter Lambda function List[1,2,3,4,5,6,7,8,9,10]

7.create a dataframe as following using pandas

8. Create a data frame with following data

  
a. Make a pivot table that shows the average salary of each employee for each department.  
b. Make a pivot table that shows the sum and mean of the salaries of each type of employee and the number of employees of each type.

9. Consider the credit card dataset which contains the following columns:  
● Create a bivariate (scatter plot) plot to find if there is a correlation between credit card limit and average purchase made on the card.  
● Visualise the distribution of values for credit card limit and average purchase made on the card. Also, identify the outliers in the data, if any.  
● Provide a visual representation of the number of customers in each income group using a bar chart.  
● Plot the frequency distribution of the total transaction amount.  
● Graphically represent the percentage of customers retained and those attrited. Highlight the latter by slicing it apart from the main pie.

10. Find a list of squares of all the numbers in a given list using lambda and map function.

11. Find the odd numbers from a given list using a filter

12. Compute a sum of the first five integers using reduce function.

13. Create the following plots to visualize/summarize the data and customize it appropriately.  
● Histogram to check the frequency distribution of the variable ‘mpg’ (Miles per gallon) and note down the interval having the highest [0,3 2,5 2,3,4 15] frequency.  
● scatter plot to determine the relationship between the weight of the car and the mpg  
● bar plot to check the frequency distribution of transmission type of0 cars.  
● Box plot of mpg and interpret the five-number summary.  
Create a git repository and push source code to the repository

14.Consider the rainfall dataset. This data contains region(district) wise rainfall across India. Perform the following operations for the dataset

15.Find the district that gets the highest annual rainfall.

16. Drop the columns 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'.

17. Display the state-wise mean rainfall for all the months using a pivot table.

18.In recent Your most of the Companies are turning to text based chatbox for resolving Consumer queries. What is the reason for it & how is it impacting the business.

19.chess game / online trading /cyber attack / not entertainment Shopping/Medical Services (health care).

20.Create a df apply aggregation , grouping, Pivot and melt function ,Map, filter ,Reduce

21.create a df using Numpy modules aggregation function ,Vector, Map, filter, reduce and Lambda

22.Demonstrate create, Insert ,Select, operation for DB Connectivity using python.

23.Assume proper Dataset and apply different kinds plot using matplotlib.

24.Assume Stock Dataset and visualize Time series data using pandas.