TOPSTECHNOLOGIES
Training | Outsourcing | Placement | Study Abroad

Module 2) Introduction to Excel

Assignment

https://github.com/TopsCode/Data_Analysis_2024/tree/main/Module%203%20DA%20-%20Introduction%20to%20Excel%20/Assignment%20Data

1) Use the average function and calculate the average of all the three category of weight. (for this question use excel file named **average 1**).

2) The excel file named **Average 3,** the table below contains precipitation measurement as measured in the Rochester NY area last year and we sampled 3 days in each of the first three months of 2018. Complete all the question in the file given.

3) In excel file named **Count 1**, The table below shows survey responses; the respondents could use any value for their answers.

4) Answer all the questions using COUNT and COUNTA function.

5) In excel file named **COUNT 2,** The following table represents a bank statement of ExcelMaster company. Column E shows the total dollar value amount of each of the accounts. Answer all the questions using COUNT and COUNTA function.

6) In excel file named **COUNT 3,** Solve all the question by using formulas COUNT, COUNTA and COUNTBLANK:

7) In excel file named **HLOOKUP**, solve all the question using HLOOKUP only.

8) In excel file named **IF 1,** Table A contains names and their respective grades for Excel 101 Course. Complete column C using only IF formula.

9) In excel file named **IF 2**, The following table is an extract from an accounting system that contains four journal entries. Check if column A's cells match column B's cell. if they match - return "match", otherwise return "no match".

10) In excel file named **IF 3**, The table below contains details of high school student's names and ages, use IF formula to complete columns D and E.If the student's age is 16 or above, he/she is eligible for a driver's license. Check if they are eligible or not. Answer in column D.

11) If the student is younger than 18 years old, he/she is a minor. Check whether the student is a minor or not. for Minor return "Minor" and non-minor = "Adult" answer in column E.

12) In excel file named **IF** 4, An A+ student gets 100% scholarship and non A+ gets 50% scholarship, the following table contains the names of students from 2024 class. Use IF function to calculate the scholarships' amounts each of them will get.

13) In excel file named **Math 1,** Use the following guidelines to calculate the statements given the file.

14) In excel file named **MAX MIN 1,** Use max, min and average formulas to answer all the following questions given in the file.

15) In the file named **MAX MIN 2**, The following table contains details about the scores of 4 students in a driving theory test. If a student fails at least one test - she or he needs to retake the course. Use IF and MAX/MIN to check if a student passed the test.

16) In the file named **MAX MIN 3**, IF at least one student got 99 points or more in a test - the test considered easy, Use MAX and IF to create a logic that checks if the test was "Easy" or not.

17) In the file named **Nested IF 1**, The school decided to use the following grade system:
    a. Grade higher or equal to 80 - Excellent
    b. Grade higher or equal to 60 but lower than 80 – Good
    c. Grade lower than 60 - Failed Complete all the task given in the file.

18) In the file named **SUM 1**, The following table includes ABC company's revenue by month. The company's CFO asked you to use SUM formula to calculate the total revenue for the year.

19) In the file named **SUM 2**, The following table represents daily costs by

20) day for the first quarter of 2015. Calculate the total costs at the bottom of the table. Hint: to save time, use sum shortcuts.

21) In the file named **SUM 3,** Find the number of residents for each of the following groups from the table below, complete all the question in the file.

22) In the file named **SUMIF 1**, answer all the question given in the file.

23) In the file named **SUMIF 2,** answer all the question given in the file based on table.

24) In the file named **VLOOKUP APPROXIMATE MATCH**, Retrieve the GBP: USD exchange rate for the following dates using VLOOKUP function, from the table in columns G-H. In case there is no exchange rate for a certain date entry, return the the last known rate for that day.

25) In the file named **VLOOKUP 1,** Below is a list of the employees who

work in your company: Answer all the question given in the file using vlookup function.

26)In the file named **VLOOKUP 2**a, according to the table, answer all the question given in the file using vlookup.

27)In Excel file **first exercise,** for a table of populations, change data types and make other changes in Power Query. Do the following things to make this table easier to read:

a. Tell Power Query to use the first row as column headings.
b. Delete the Source column (we don't need it).
c. Change the data type of the Date column to Date.
d. Change the data type of the Population column to Whole Number.
e. Shorten the name of the Country column.
f. And make other changes if needed to read data properly.

28)In Excel file **first exercise**, import a population table using Power Query, then tidy up the data: a. In the "**question_3_power_query_file**", all the steps are mentioned to do the above exercise, please look and follow the steps.

29)In Excel file first exercise, divide exchange rate and investment symbols into more parts by splitting and removing columns: a. In the "**question_4_power_query_file**", all the steps are mentioned to do the above exercise, please look and follow the steps.

Question 27: - Our Main goal is to create a macro that will convert this CSV File into nice and clean table as shown below. (Use Macros file from GitHub)

Aarav,Patel,50000,30
Aisha,Kumar,60000,35
Amit,Sharma,75000,40
Ananya,Choudhury,55000,28
Arjun,Reddy,80000,45
Avni,Gupta,65000,33
Dev,Verma,70000,38
Dia,Singh,60000,32
Ishaan,Saxena,85000,42
Jiya,Khan,70000,36
Kabir,Mishra,60000,31
Kriti,Joshi,55000,29
Mohan,Kumar,90000,48
Neha,Shah,65000,34
Pranav,Jain,75000,39
Riya,Das,70000,37
Rohan,Pandey,80000,41
Sneha,Chopra,60000,33
Vikram,Singhania,70000,38

Zoya,Mehta,65000,35

Step 1=	**Convert The CSV file into Table**
Step 2=	**With Text to Columns**
Step 3=	**Add Discount Column**
Step 4=	**Do formatting and colouring**

Sr. No.	First Name	Last Name	Salary	Age
1	Aarav	Patel	50000	30
2	Aisha	Kumar	60000	35
3	Amit	Sharma	75000	40
4	Ananya	Choudhury	55000	28
5	Arjun	Reddy	80000	45
6	Avni	Gupta	65000	33
7	Dev	Verma	70000	38
8	Dia	Singh	60000	32
9	Ishaan	Saxena	85000	42
10	Jiya	Khan	70000	36
11	Kabir	Mishra	60000	31
12	Kriti	Joshi	55000	29
13	Mohan	Kumar	90000	48
14	Neha	Shah	65000	34
15	Pranav	Jain	75000	39

16	Riya	Das	70000	37
17	Rohan	Pandey	80000	41
18	Sneha	Chopra	60000	33
19	Vikram	Singhania	70000	38
20	Zoya	Mehta	65000	35

Question 28: - Our Main goal is to create a Macros that will convert this CSV File into nice and clean table as shown below. (Use VBA file from GitHub)

OrderID,Name,Product,Quantity,Price,Total
11280,Bill Smith, Volkswagen Golf,15,30000,450000
11281,Kennedi Singh, Toyota Yaris,10,25000,250000
11282,Harley Fritz, Seat Panda,150,28000,4200000
11283,Nyla Novak, Ford Focus,12,30000,360000
11284,Ivan Hines, Vauxhall Corsa,20,30000,600000

11285,Jonah Higgins, Volkswagen Polo,20,26000,520000
11286,Jordan Boone,Kia Sportage,2,30000,60000
11287,Kylee Townsend,Ford Fiesta,5,35000,175000
11288,Nora Rollins,Seat Ibiza,80,30000,2400000
11289,Brendan Walls,Renault Clio,100,26000,2600000
11290,Steven Michael,Honda CR-V,20,35000,700000
11291,Lucia Mckay,GMC Sierra,25,40000,1000000
11292,Josue Roach,Seat Panda,10,28000,280000
11293,Franklin Wright,Ford Focus,10,30000,300000
11294,Denzel Flores,Vauxhall Corsa,5,30000,150000

11295,Bruno Cordova,Volkswagen

Polo,180,26000,4680000

11296,Jaylynn Knapp,Kia Sportage,5,30000,150000

11297,Bruce Rich,Ford Fiesta,250,35000,8750000

1=	**Convert The CSV file into Table**
2=	**With Text To Columns**
3=	**Add Discount Column**
4=	**Do formatting And colouring**
5=	**Create a Function to give Discount of 20 % if Buying Quantity more than 20**

Module 3) Applied Statistics in Excel

Assignment

https://github.com/TopsCode/Data_Analysis_2024/blob/main/ALL_CSV/airline_passenger_satisfaction.csv

First of all, I choice this "Airline Passenger Satisfaction" dataset because of airline statistics intrigued me. Also I choice the "Flight Distance" column because I think that can find there is interesting statistics relations between that passenger satisfaction and flight distance. This column shows us the flight distance of all passengers. According to the column information mean is 119k, standard deviation is 997 and quantiles are Q1 = 414, Q2(Median) = 844, Q3 = 1744. Also, we see the minimum value is 31 and the maximum value is 4983.

Content:

1. Mean of Column Data
2. Median of Column Data
3. Variance, Standard Deviation and Standard Error
4. Decide the Shape of Distribution
5. Find Outliers
6. Graph the Column Data and Comment
7. Boxplot

Mean of Column Data

To find a mean of column data, we have to sum all elements of the column data and divided by the column length.

Median of Column Data

To find the median value of the column, we have to sort the dataset and check if the dataset length is even or odd. If the dataset length is an odd number, the median

value will be the middle value of the dataset. If the dataset length is an even number, the median value will be the average of the two middle values.

Variance:

Variance tells you the degree of spread in your data set. The more spread the data, the larger the variance is in relation to the mean. A large variance means that the values have a large deviation from the arithmetic mean. To calculate variance, first of all we will subtract the mean from each value and square the results obtained. After that, we have to sum all squares, then finally divided the sum of squares by n (when you work with population).

The standard deviation is the average amount of variability in your dataset. The standard deviation shows us on average, how far each value lies from the mean. If the value of the standard deviation is high that means values are generally far from the mean. It is calculated by squaring the variance.

Standard Error:

The standard error tells us; how different the population mean is likely to be from a sample mean. It is calculated by dividing the standard deviation by the square root of the number of elements.

Decide the Shape of Distribution

Shape of distribution type is determined by the values of the average and the median. If the mean bigger than median, the shape of distribution will be Right-Skewed (Positively Skewed) distribution, if median bigger than mean, the shape of distribution will be Left-Skewed (Negatively Skewed) distribution. If mean and median are equal than the shape of distribution will be Normal(Symmetric)

distribution. Our case we have a Right-Skewed (Positively Skewed) distribution for "Flight Distance" column, because the mean bigger than the median.

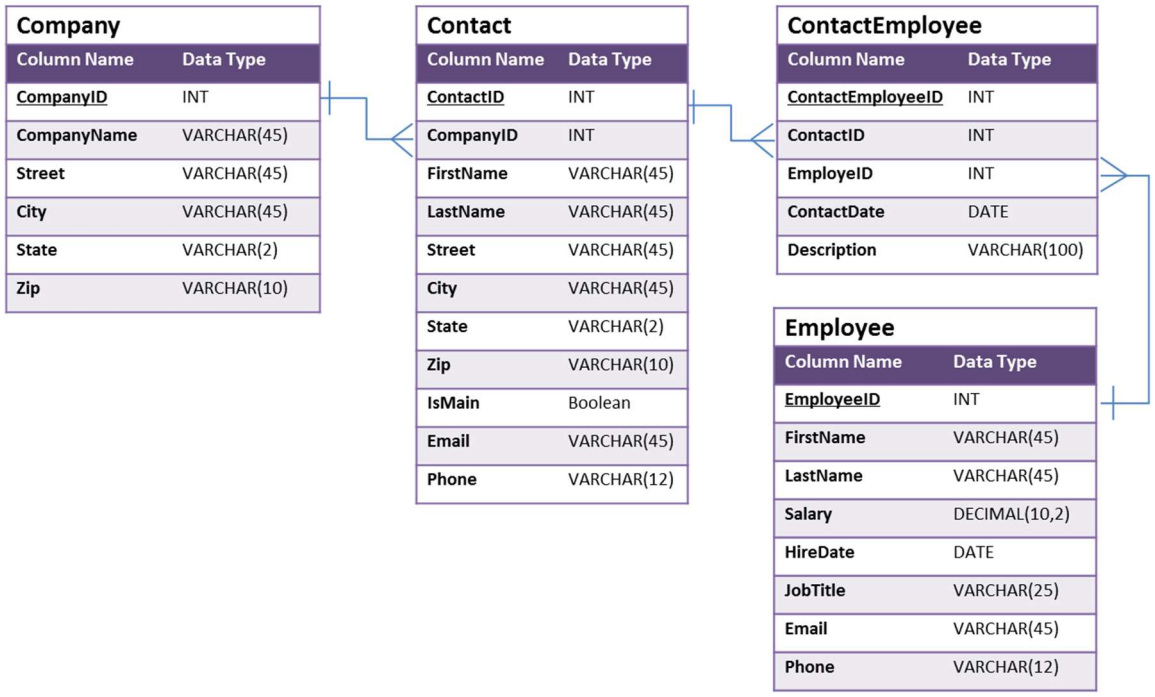
Find Outliers

The outliers are as known as the extreme values of the dataset. This means extreme values are different from all other values on the dataset. If we want a consistent statistical result, we should clear them, because they can cause a huge effect on your statistic. To find outliers we have to calculate lower fence and upper fence, first of all, we will subtract q1 from q3 to find IQR. After that, to find the lower fence we have to will subtract (1.5 * IQR) from q1 and to find the upper fence we have to sum Q3 with (1.5 * IQR). After we find fences, values which bigger than the upper fence and smaller than the lower fence are called outliers.

Module 4) Working with Database using SQL

Assignment

For this assignment, you will finish building the contact management database for MarketCo



Company

Column Name	Data Type
CompanyID	INT
CompanyName	VARCHAR(45)
Street	VARCHAR(45)
City	VARCHAR(45)
State	VARCHAR(2)
Zip	VARCHAR(10)

Contact

Column Name	Data Type
ContactID	INT
CompanyID	INT
FirstName	VARCHAR(45)
LastName	VARCHAR(45)
Street	VARCHAR(45)
City	VARCHAR(45)
State	VARCHAR(2)
Zip	VARCHAR(10)
IsMain	Boolean
Email	VARCHAR(45)
Phone	VARCHAR(12)

ContactEmployee

Column Name	Data Type
ContactEmployeeID	INT
ContactID	INT
EmployeID	INT
ContactDate	DATE
Description	VARCHAR(100)

Employee

Column Name	Data Type
EmployeeID	INT
FirstName	VARCHAR(45)
LastName	VARCHAR(45)
Salary	DECIMAL(10,2)
HireDate	DATE
JobTitle	VARCHAR(25)
Email	VARCHAR(45)
Phone	VARCHAR(12)

1) Statement to create the Contact table

2) Statement to create the Employee table

3) Statement to create the ContactEmployee table
HINT: Use DATE as the datatype for ContactDate. It allows you to store the date in this format: YYYY-MM-DD (i.e., '2014-03-12' for March 12, 2014).

4) In the Employee table, the statement that changes Lesley Bland's phone number to 215-555-8800

5) In the Company table, the statement that changes the name of "Urban Outfitters, Inc." to "Urban Outfitters" .

6) In ContactEmployee table, the statement that removes Dianne Connor's contact event with Jack Lee (one statement).
HINT: Use the primary key of the ContactEmployee table to specify the correct record to remove.

7) Write the SQL SELECT query that displays the names of the employees that have contacted Toll Brothers (one statement). Run the SQL SELECT query in MySQL Workbench. Copy the results below as well.

8) What is the significance of "%" and "_" operators in the LIKE statement?

9) Explain normalization in the context of databases.

10) What does a join in MySQL mean?

11) 19.What do you understand about DDL, DCL, and DML in MySQL?

12) What is the role of the MySQL JOIN clause in a query, and what are some common types of joins?

Module 5) Creating Dashboard with Visualization Tool

Assignment

1) What is Power BI and how does it differ from Excel?
2) Explain the concept of data modeling in Power BI.
3) What are the different types of connections available in Power BI?
4) How do you handle data transformation in Power BI?
5) What is DAX (Data Analysis Expressions) and why is it important in Power BI?
6) Can you explain the difference between calculated columns and measures in Power BI?
7) How do you handle relationships between tables in Power BI?
8) What is the purpose of a Power BI Gateway?
9) How can you schedule data refresh in Power BI Service?
10) Explain the concept of row-level security in Power BI.
11) What is the Power BI Desktop and how does it differ from Power BI Service?

12) Explain the concept of Direct Query in Power BI.

13) What are Power BI templates and how are they useful?

14) How do you handle incremental data refresh in Power BI?

15) What is the role of Power Query in Power BI?

16) Explain the difference between calculated columns and calculated tables in Power BI.

17) How do you create custom visuals in Power BI?

18) What are the best practices for optimizing performance in Power BI?

19) How can you integrate Power BI with other Microsoft products like Azure and Office 365?

20) Explain the concept of aggregations in Power BI.

21) How do you handle error handling and data quality in Power BI?

22) What is the purpose of Power BI Embedded and when would you use it?

Module 7) DA - Introduction to Python

Assignment

1) What are the types of Applications?
2) What is programing?
3) What is Python?
4) Write a Python program to check if a number is positive, negative or zero.
5) Write a Python program to get the Factorial number of given numbers.
6) Write a Python program to get the Fibonacci series of given range.
7) How memory is managed in Python?
8) What is the purpose continuing statement in python?
9) Write python program that swap two number with temp variable and without temp variable.
10) Write a Python program to find whether a given number is even or odd, print out an appropriate message to the user.
11) Write a Python program to test whether a passed letter is a vowel or not.
12) Write a Python program to sum of three given integers. However, if two values are equal sum will be zero.
13) Write a Python program that will return true if the two given integer values are equal or their sum or difference is 5.
14) Write a python program to sum of the first n positive integers.
15) Write a Python program to calculate the length of a string.
16) Write a Python program to count the number of characters (character frequency) in a string
17) What are negative indexes and why are they used?
18) Write a Python program to count occurrences of a substring in a string.
19) Write a Python program to count the occurrences of each word in a

given sentence

20) Write a Python program to get a single string from two given strings, separated by a space and swap the first two characters of each string.

21) Write a Python program to add 'in' at the end of a given string (length should be at least 3). If the given string already ends with 'ing' then add 'ly' instead if the string length of the given string is less than 3, leave it unchanged.

22) Write a Python function to reverses a string if its length is a multiple of 4.

23) Write a Python program to get a string made of the first 2 and the last 2 chars from a given a string. If the string length is less than 2, return instead of the empty string.

24) Write a Python function to insert a string in the middle of a string.

25) What is List? How will you reverse a list?

26) How will you remove last object from a list?

27) Suppose list1 is [2, 33, 222, 14, and 25], what is list1 [-1]?

28) Differentiate between append () and extend () methods?

29) Write a Python function to get the largest number, smallest num and sum of all from a list.

30) How will you compare two lists?

31) Write a Python program to count the number of strings where the string length is 2 or more and the first and last character are same from a given list of strings.

32) Write a Python program to remove duplicates from a list.

33) Write a Python program to check a list is empty or not.

34) Write a Python function that takes two lists and returns true if they have at least one common member.

35) Write a Python program to generate and print a list of first and last 5 elements where the values are square of numbers between 1 and 30.

36) Write a Python function that takes a list and returns a new list with unique elements of the first list.

37) Write a Python program to convert a list of characters into a string.

38) Write a Python program to select an item randomly from a list.

39) Write a Python program to find the second smallest number in a list.

40) Write a Python program to get unique values from a list

41) Write a Python program to check whether a list contains a sub list

42) Write a Python program to split a list into different variables.

43) What is tuple? Difference between list and tuple.

44) Write a Python program to create a tuple with different data types.

45) Write a Python program to unzip a list of tuples into individual lists.

46) Write a Python program to convert a list of tuples into a dictionary.

47) How will you create a dictionary using tuples in python?

48) Write a Python script to sort (ascending and descending) a dictionary by value.

49) Write a Python script to concatenate following dictionaries to create a new one.

50) Write a Python script to check if a given key already exists in a dictionary.

51) How Do You Traverse Through a Dictionary Object in Python?

52) How Do You Check the Presence of a Key in A Dictionary?

53) Write a Python script to print a dictionary where the keys are numbers between 1 and 15.

54) Write a Python program to check multiple keys exists in a dictionary

55) Write a Python script to merge two Python dictionaries

56) Write a Python program to map two lists into a dictionary

Sample output: Counter ({'a': 400, 'b': 400,'d': 400, 'c': 300}).

57) Write a Python program to find the highest 3 values in a dictionary

58) Write a Python program to combine values in python list of dictionaries.
Sample data: [{'item': 'item1', 'amount': 400}, {'item': 'item2', 'amount': 300}, o {'item': 'item1', 'amount': 750}]
Expected Output:

- Counter ({'item1': 1150, 'item2': 300})

59) Write a Python program to create a dictionary from a string.
Note: Track the count of the letters from the string.

60) Sample string:
'w3resource' Expected output:

- {'3': 1,'s': 1, 'r': 2, 'u': 1, 'w': 1, 'c': 1, 'e': 2, 'o': 1}

61) Write a Python function to calculate the factorial of a number (a nonnegative integer)

62) Write a Python function to check whether a number is in a given range

63) Write a Python function to check whether a number is perfect or not.

64) Write a Python function that checks whether a passed string is palindrome or not

65) How Many Basic Types of Functions Are Available in Python?

66) How can you pick a random item from a list or tuple?

67) How can you pick a random item from a range?

68) How can you get a random number in python?

69) How will you set the starting value in generating random numbers?

70) How will you randomize the items of a list in place?

71) What is File function in python? What are keywords to create and write file.
72) Write a Python program to read an entire text file.
73) Write a Python program to append text to a file and display the text.
74) Write a Python program to read first n lines of a file.
75) Write a Python program to read last n lines of a file.
76) Write a Python program to read a file line by line and store it into a list
77) Write a Python program to read a file line by line store it into a variable.
78) Write a python program to find the longest words.
79) Write a Python program to count the number of lines in a text file.
80) Write a Python program to count the frequency of words in a file.
81) Write a Python program to write a list to a file.
82) Write a Python program to copy the contents of a file to another file.
83) Explain Exception handling? What is an Error in Python?
84) How many except statements can a try-except block have? Name Some built-in exception classes:
85) When will the else part of try-except-else be executed?
86) Can one block of except statements handle multiple exception?
87) When is the finally block executed?
88) What happens when „1“== 1 is executed?
89) How Do You Handle Exceptions with Try/Except/Finally in Python? Explain with coding snippets.
90) Write python program that user to enter only odd numbers, else will raise an exception.

Module 8 DA- Working with NumPy (python)

Assignment

1. How to Creating a 3x3 Identity Matrix with Float Data Type?
2. Create a 1D Array with Random Values between 0 and 1.
3. Create a 2D Array with Random Integer Values.
4. Creating an Array Using a Custom Function.
5. Reshaping a 1D Array into a 2D Array
6. How to Creating a 3x3 Array of Ones?

7. How to get the common items between two pythons
   NumPy? Input:
   a = np. array ([1,2,3,2,3,4,3,4,5,6])
   b = np. array ([7,2,10,2,7,4,9,4,9,8])
   Expected Output:
   array ([2, 4])

8. From array a remove all items present in
   array b Input:
   a = np. array ([1,2,3,4,5])
   b = np. array ([5,6,7,8,9])
   Expected Output:
   array ([1,2,3,4]

9. Limit the number of items printed in python NumPy array a to a maximum of 6 elements.

a = np. arrange (15)

Expected Output:

array ([0, 1, 2, ..., 12, 13, 14]

10. Drop all nan values from a 1D NumPy

array Input:

np. array ([1,2,3, np.nan,5,6,7, np.nan])

Desired Output:

array ([1., 2., 3., 5., 6., 7.])

11. Question: Create a 1D NumPy array of the first 20 natural numbers and a 2D NumPy array of shape (4, 5) with values ranging from 1 to 20.

12. Question: Given a 3D NumPy array of shape (2, 3, 4), find its shape, size, number of dimensions, and data type. Change its data type to float64 and verify the change.

13. Question: Reshape a 1D array of 12 elements into a 3x4 2D array and then flatten it back into a 1D array using ravel (). Verify that the flattened array matches the original.

14. Question: Given two arrays, a = np. array ([1, 2, 3]) and b = np. array ([4, 5, 6]), perform element-wise addition, subtraction, multiplication, and division. Explain the behavior when dividing by zero.

15. Question: Create a 2D array of shape (3, 1) and a 1D array of length 3. Perform element-wise addition using broadcasting. Explain how broadcasting rules apply in this scenario.

16. Question: Generate a random 2D array of integers between 0 and 10. Use

conditional operators to create a Boolean mask identifying elements greater than 5. Replace all elements greater than 5 with the value 5.

17. Question: Given a 4x4 array of random integers, use indexing and slicing to extract:

- The entire second row
- The last column
- The subarray consisting of the first two rows and first two columns

18. Question: Describe a practical example where NumPy can be used in EDA, AI, ML, and DL, and implement a NumPy solution for a simple task in each area.

19. Question: Create a random 4x4 matrix and use NumPy to compute its eigenvalues and eigenvectors. Verify the eigenvalues by reconstructing the matrix.

.

20. Question: Create a 1D array of 27 elements and reshape it into a 3x3x3 3D array. Flatten it back into a 1D array and compare the flattened array with the original.

21. Question: Perform matrix multiplication of two 2D arrays using np.dot () and @. Compare the results and performance of both methods using a large dataset.

22. Question: Create a 3D array of shape (2, 1, 4) and a 2D array of shape (4, 1). Perform an element-wise operation using broadcasting and explain the result. Use np. new axis to achieve the same result without broadcasting.

23. Question: Generate a 2D array of random floats between 0 and 1. Use conditional operators to create a Boolean mask for values less than 0.5. Replace these values with their squares and leave the rest unchanged.

24. Question: Given a 5x5 array of sequential integers, use slicing to:

- Extract the diagonal elements
- Replace the elements of the middle row with zeros
- Flip the array vertically and horizontally

25. Question: Create a 4D array of shape (2, 3, 4, 5) with random integers. Use advanced slicing to extract a subarray and compute the mean along a specified axis.

26. Question: Given an array of shape (10, 20), reshape it to (20, 10) and (5, 40). Discuss the impact on the array's shape, size, and dimensionality.

27. Question: Generate a large 2D array and demonstrate the use of np. reshape () and unravel () to manipulate its shape for various linear algebra operations.

28. Question: Given a 6x6 matrix, use advanced indexing and slicing to extract the upper triangular part of the matrix and set the lower triangular part to zero. Verify the result.

Module 9) DA- Working with Pandas (python)

Assignment

1) Create a series of three different colors
2) View the series of different colors
3) Create a series of three different car types and view it
4) Combine the Series of cars and colors into a Data Frame
5) Find the different datatypes of the car data Data Frame
6) Describe your current car sales Data Frame using describe ()
7) Get information about your Data Frame using info ()
8) Create a Series of different numbers and find the mean of them
9) Create a Series of different numbers and find the sum of them
10) List out all the column names of the car sales Data Frame
11) Find the length of the car sales Data Frame
12) Show the first 5 rows of the car sales Data Frame
13) Show the first 7 rows of the car sales Data Frame
14) Show the bottom 5 rows of the car sales Data Frame
15) Use. loc to select the row at index 3 of the car sales Data Frame
16) Use. iloc to select the row at position 3 of the car sales Data Frame
17) Create a crosstab of the Make and Doors columns.

Module 10) DA- Visualization with Matplotlib and Seaborn

Assignment

Dataset contains information about some police deaths in US from 1984 to 2016. Create the following visualizations using this dataset.

1) Bar chart showing the total deaths per year for 1984-2016
2) Line chart comparing the yearly deaths in different states
3) A heat map overlaid on the map of the United States
4) A Choropleth comparing the number of police deaths in different states
5) A word cloud of the different causes of death (remove the string "Cause of Death:" for best results)
6) A marker cluster showing the shootings in the state of California (show person's name on hover)
7) Bar chart comparing the no. of deaths due to different causes, animated by year
8) A heatmap of total deaths per state per year i.e., showing "state" on one axis and "year" on the other axis.
9) A tree map with three levels: state, city (description), cause
10) A rug plot showing a timeline of canine deaths.

Download Dataset: https://github.com/tops121/TOPS_TECH

Module 11 DA- Working with Scrapping (python)

Assignment

Use Any Website as per Your Faculty Suggest and the requests library for Webpage

1) Inspect the website's HTML source and identify the right URLs to download.
2) Download and save web pages locally using the requests library.
3) Create a function to automate downloading for different topics/search queries.
4) Use Beautiful Soup to parse and extract information
5) Parse and explore the structure of downloaded web pages using beautiful soup.
6) Use the right properties and methods to extract the required information.
7) Create functions to extract from the page into lists and dictionaries.
8) Create functions for the end-to-end process of downloading, parsing, and saving CSVs.