

Basic Python

1. Write a Python program which accepts the user's first and last name and print them in reverse order with a space between them.

2. Write a Python program which accepts a sequence of comma-separated numbers from user and generate a list and a tuple with those numbers.

Sample data : 3, 5, 7, 23

Output :

List : ['3', '5', '7', '23']

Tuple : ('3', '5', '7', '23')

3. Write a Python program to display the first and last colors from the following list.

color_list = ["Red", "Green", "White", "Black"]

4. Write a Python program to print the documents (syntax, description etc.) of Python built-in function(s).

Sample function : abs()

Expected Result : mat

abs(number) -> number

Return the absolute value of the argument.

5. Write a Python program to print the calendar of a given month and year.

Note : Use 'calendar' module.

6. Write a Python program to calculate number of days between two dates.

Sample dates : (2014, 7, 2), (2014, 7, 11)

Expected output : 9 days

7. Write a Python program to check whether a specified value is contained in a group of values.

Test Data :

3 -> [1, 5, 8, 3] : True

-1 -> [1, 5, 8, 3] : False

8. Write a Python program to create a histogram from a given list of integers. 9. Write a

Python program to concatenate all elements in a list into a string and return it.

10. Write a Python program to print out a set containing all the colors from color_list_1 which are not present in color_list_2.

Test Data :

color_list_1 = set(["White", "Black", "Red"])

```
color_list_2 = set(["Red", "Green"])
```

Expected Output :

```
{'Black', 'White'}
```

11. Write a Python program to check whether a file exists.
12. Write a python program to call an external command in Python.
13. Write a Python program to find out the number of CPUs using.
14. Write a Python program to list all files in a directory in Python.
15. Write a python program to access environment variables.
16. Write a Python program to get the current username
17. Write a program to get execution time for a Python method.
18. Write a Python program to get an absolute file path.
19. Write a Python program to get file creation and modification date/times.
20. Write a Python program to sort three integers without using conditional statements and loops.
21. Write a Python program to sort files by date.
22. Write a Python program to get the command-line arguments (name of the script, the number of arguments, arguments) passed to a script.
23. Write a Python program to find the available built-in modules.
24. Write a Python program to get the size of an object in bytes.
25. Write a Python program to get the current value of the recursion limit.
26. Write a Python program to count the number occurrence of a specific character in a string.
27. Write a Python program to get the system time.
28. Write a Python program to clear the screen or terminal.

29. Write a Python program to get the name of the host on which the routine is running.
30. Write a Python program to access and print a URL's content to the console.

31. Write a Python program to get system command output.

32. Write a Python program to get the effective group id, effective user id, real group id, a list of supplemental group ids associated with the current process.

Note: Availability: Unix.

33. Write a Python program to get the users environment.

34. Write a Python program to retrieve file properties.

35. Write a Python program to get numbers divisible by fifteen from a list using an anonymous function.

36. Write a Python program to determine if variable is defined or not.

37. Write a Python program to empty a variable without destroying it.

Sample data: n=20

d = {"x":200}

Expected Output : 0

{}

38. Write a Python program to add leading zeroes to a string.

39. Write a Python program to find files and skip directories of a given directory. 40.

Write a Python program to extract single key-value pair of a dictionary in variables.

41. Write a Python program to convert an integer to binary keep leading zeros.

Sample data : 50

Expected output : 00001100, 0000001100

42. Write a Python program to determine if the python shell is executing in 32bit or 64bit mode on operating system.

43. Write a Python function to find the maximum and minimum numbers from a sequence of numbers.

Note: Do not use built-in functions.

Python Data Structure

Array

1. Write a Python program to create an array of 5 integers and display the array items. Access individual element through indexes.
2. Write a Python program to reverse the order of the items in the array.
3. Write a Python program to get the number of occurrences of a specified element in an array.
4. Write a Python program to remove the first occurrence of a specified element from an array.

Dictionary

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

2. Write a Python script to add a key to a dictionary.

Sample Dictionary : {0: 10, 1: 20}
Expected Result : {0: 10, 1: 20, 2: 30}

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

Sample Dictionary :
dic1={1:10, 2:20}
dic2={3:30, 4:40}
dic3={5:50,6:60}
Expected Result : {1: 10, 2: 20, 3: 30, 4: 40, 5: 50, 6: 60}

4. Write a Python program to iterate over dictionaries using for loops.

5. Write a Python script to generate and print a dictionary that contains a number (between 1 and n) in the form (x, x*x).

Sample Dictionary (n = 5) :
Expected Output : {1: 1, 2: 4, 3: 9, 4: 16, 5: 25}

6. Write a Python program to remove a key from a dictionary.

7. Write a Python program to print all unique values in a dictionary.

Sample Data : [{"V": "S001"}, {"V": "S002"}, {"VI": "S001"}, {"VI": "S005"}, {"VII": "S005"}, {"V": "S009"}, {"VIII": "S007"}]

Expected Output : Unique Values: {'S005', 'S002', 'S007', 'S001', 'S009'}

8. Write a Python program to create a dictionary from a string.

Note: Track the count of the letters from the string.

Sample string : 'w3resource'

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

9. Write a Python program to print a dictionary in table format.

10. Write a Python program to count the values associated with key in a dictionary.

Sample data: = [{"id": 1, 'success': True, 'name': 'Lary'}, {"id": 2, 'success': False, 'name': 'Rabi'}, {"id": 3, 'success': True, 'name': 'Alex'}]

Expected result: Count of how many dictionaries have success as True 11.

Write a Python program to convert a list into a nested dictionary of keys.

12. Write a Python program to check multiple keys exists in a dictionary.

13. Write a Python program to count number of items in a dictionary value that is a list.

Sets

1. Write a Python program to create a set.

2. Write a Python program to iteration over sets.

3. Write a Python program to add member(s) in a set.

4. Write a Python program to remove item(s) from set

5. Write a Python program to remove an item from a set if it is present in the set.

6. Write a Python program to create an intersection of sets.

7. Write a Python program to create a union of sets.

8. Write a Python program to create set difference.

9. Write a Python program to create a symmetric difference.

10. Write a Python program to clear a set.

11. Write a Python program to use of frozensets.

12. Write a Python program to find maximum and the minimum value in a set.

List

1. Write a Python program to sum all the items in a list.

2. Write a Python program to multiplies all the items in a list.

3. Write a Python program to get the smallest number from a list.

4. 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.

Sample List : ['abc', 'xyz', 'aba', '1221']

Expected Result : 2

5. Write a Python program to get a list, sorted in increasing order by the last element in each tuple from a given list of non-empty tuples.

Sample List : [(2, 5), (1, 2), (4, 4), (2, 3), (2, 1)]

Expected Result : [(2, 1), (1, 2), (2, 3), (4, 4), (2, 5)]

6. Write a Python program to remove duplicates from a list.

7. Write a Python program to clone or copy a list.

8. Write a Python program to find the list of words that are longer than n from a given list of words.

9. Write a Python function that takes two lists and returns True if they have at least one common member.

10. Write a Python program to print a specified list after removing the 0th, 4th and 5th elements.

Sample List : ['Red', 'Green', 'White', 'Black', 'Pink', 'Yellow']

Expected Output : ['Green', 'White', 'Black']

11. Write a Python program to generate all permutations of a list in Python.

12. Write a Python program to get the difference between the two lists. 13.

Write a Python program to append a list to the second list.

14. Write a python program to check whether two lists are circularly identical. 15. Write a Python program to find common items from two lists.
16. Write a Python program to split a list based on first character of word.
17. Write a Python program to remove duplicates from a list of lists.
Sample list : `[[10, 20], [40], [30, 56, 25], [10, 20], [33], [40]]`
New List : `[[10, 20], [30, 56, 25], [33], [40]]`

Tuple

1. Write a Python program to create a tuple.
2. Write a Python program to create a tuple with different data types.
3. Write a Python program to unpack a tuple in several variables.
4. Write a Python program to create the colon of a tuple.
5. Write a Python program to find the repeated items of a tuple.
6. Write a Python program to check whether an element exists within a tuple.
7. Write a Python program to convert a list to a tuple.
8. Write a Python program to remove an item from a tuple.
9. Write a Python program to slice a tuple.
10. Write a Python program to reverse a tuple.

Strings

1. Write a Python program to calculate the length of a string.
2. Write a Python program to count the number of characters (character frequency) in a string.

Sample String : `google.com`

Expected Result : `{'o': 3, 'g': 2, '.': 1, 'e': 1, 'l': 1, 'm': 1, 'c': 1}`

3. Write a Python program to get a string from a given string where all occurrences of its

first char have been changed to '\$', except the first char itself.

Sample String : 'restart'

Expected Result : 'resta\$t'

4. Write a Python program to add 'ing' 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.

Sample String : 'abc'

Expected Result : 'abcing'

Sample String : 'string'

Expected Result : 'stringly'

5. Write a Python function that takes a list of words and returns the length of the longest one.

6. Write a Python script that takes input from the user and displays that input back in upper and lower cases.

7. Write a Python program that accepts a comma separated sequence of words as input and prints the unique words in sorted form (alphanumerically).

Sample Words : red, white, black, red, green, black

Expected Result : black, green, red, white,red

8. Write a Python program to get the last part of a string before a specified character.

<https://www.w3resource.com/python-exercises>

<https://www.w3resource.com/python>

9. Write a Python program to display formatted text (width=50) as output.

10. Write a Python program to count occurrences of a substring in a

string. 11. Write a Python program to reverse a string.

12. Write a Python program to lowercase first n characters in a string.

Linear Algebra

1. Write a python program to add below matrices

X = [[12,7,3],

[4 ,5,6],

[7 ,8,9]]

Y = [[5,8,1],

[6,7,3],

[4,5,9]]

2. Write a program to perform scalar multiplication of matrix and a number

X = [[12,7,3],

[4 ,5,6],

[7 ,8,9]]

Y = 9

3. Write a program to perform multiplication of given matrix and vector

X = [[5, 1 ,3], [1, 1 ,1], [1, 2 ,1]], Y = [1, 2, 3]

4. Write a program to multiply matrices in problem 1

5. Write a program to find inverse matrix of matrix X in problem 1 . 6. Write

a program to find transpose matrix of matrix Y in problem 1

Probability and Statistics

1. Write a program to find probability of drawing an ace from pack of cards 2. Write a program to find the probability of drawing an ace after drawing a king on the first draw

3. Write a program to find the probability of drawing an ace after drawing an ace on the first draw

4. You toss a fair coin three times write a program to find following:

a. What is the probability of three heads, HHH?

b. What is the probability that you observe exactly one heads?

c. Given that you have observed at least one heads, what is the probability that you observe at least two heads?

5. In my town, it's rainy one third of the days. Given that it is rainy, there will be heavy traffic with probability 12, and given that it is not rainy, there will be heavy traffic with probability 14. If it's rainy and there is heavy traffic, I arrive late for work with probability 12. On the other hand, the probability of being late is reduced to 18 if it is not rainy and there is no heavy traffic. In other situations (rainy and no traffic, not rainy and traffic) the probability of being late is 0.25. You pick a random day.

Write a program to find following

- a. What is the probability that it's not raining and there is heavy traffic and I am not late?
- b. What is the probability that I am late?
- c. Given that I arrived late at work, what is the probability that it rained that day?

6. Given the following statistics, write a program to find the probability that a woman has cancer if she has a positive mammogram result?

- a. One percent of women over 50 have breast cancer.
- b. Ninety percent of women who have breast cancer test positive on mammograms.
- c. Eight percent of women will have false positives.

7. A bank teller serves customers standing in the queue one by one. Suppose that the service time X_i for customer i has mean $E(X_i)=2$ (minutes) and $Var(X_i)=1$. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers. Write a program to find $P(90 < Y < 110)$

8. In a communication system each data packet consists of 1000 bits. Due to the noise, each bit may be received in error with probability 0.1. It is assumed bit errors occur independently. Find the probability that there are more than 120 errors in a certain data packet.

9. In a particular pain clinic, 10% of patients are prescribed narcotic pain killers. Overall, five percent of the clinic's patients are addicted to narcotics (including pain killers and illegal substances). Out of all the people prescribed pain pills, 8% are addicts. If a patient is an addict, write a program to find the probability that they will be prescribed pain pills?

10. X is a normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Write a program to find

- a. $P(x < 40)$
- b. $P(x > 21)$
- c. $P(30 < x < 35)$

11. A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. Write a program to find the probability that a car picked at random is travelling at more than 100 km/hr?

12. Write a program to find the probability of getting a random number from the interval [2, 7]

13. The table below shows the height, x, in inches and the pulse rate, y, per minute, for 9 people. Write a program to find the correlation coefficient and interpret your result.

x \Rightarrow 68 72 65 70 62 75 78 64 68

y \Rightarrow 90 85 88 100 105 98 70 65 72

Python Libraries

NumPy -

1. Write a Python program to convert a list of numeric value into a one-dimensional NumPy array.

Expected Output:

Original List: [12.23, 13.32, 100, 36.32]

One-dimensional numpy array: [12.23 13.32 100. 36.32]

2. Create a 3x3 matrix with values ranging from 2 to 10.

Expected Output:

[[2 3 4]

[5 6 7]

[8 9 10]]

3. Write a Python program to create a null vector of size 10 and update sixth value to 11.

[0. 0. 0. 0. 0. 0. 0. 0. 0. 0.]

Update sixth value to 11

[0. 0. 0. 0. 0. 0. 11. 0. 0. 0.]

4. Write a Python program to reverse an array (first element becomes last).

Original array:

[12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37]

Reverse array:

[37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12]

5. Write a Python program to create a 2d array with 1 on the border and 0 inside.

Expected Output:

Original array:

```
[[ 1.  1.  1.  1.  1.]
```

```
 [ 1.  1.  1.  1.  1.]
```

```
 [ 1.  1.  1.  1.  1.]
```

```
 [ 1.  1.  1.  1.  1.]
```

```
 [ 1.  1.  1.  1.  1.]]
```

1 on the border and 0 inside in the array

```
[[ 1.  1.  1.  1.  1.]
```

```
 [ 1.  0.  0.  0.  1.]
```

```
 [ 1.  0.  0.  0.  1.]
```

```
 [ 1.  0.  0.  0.  1.]
```

```
 [ 1.  1.  1.  1.  1.]]
```

6. Write a Python program to add a border (filled with 0's) around an existing array.

Expected Output:

Original array:

```
[[ 1.  1.  1.]
```

```
 [ 1.  1.  1.]
```

```
 [ 1.  1.  1.]]
```

1 on the border and 0 inside in the array

```
[[ 0.  0.  0.  0.  0.]
```

```
 [ 0.  1.  1.  1.  0.]
```

```
 [ 0.  1.  1.  1.  0.]
```

```
 [ 0.  1.  1.  1.  0.]
```

```
 [ 0.  0.  0.  0.  0.]]
```

7. Write a Python program to create a 8x8 matrix and fill it with a checkerboard pattern.

Checkerboard pattern:

```
[[0 1 0 1 0 1 0 1]
```

```
 [1 0 1 0 1 0 1 0]
```

```
 [0 1 0 1 0 1 0 1]
```

```
 [1 0 1 0 1 0 1 0]
```

```
 [0 1 0 1 0 1 0 1]
```

```
 [1 0 1 0 1 0 1 0]
```

```
 [0 1 0 1 0 1 0 1]
```

```
 [1 0 1 0 1 0 1 0]]
```

8. Write a Python program to convert a list and tuple into arrays.

List to array:

```
[1 2 3 4 5 6 7 8]
```

Tuple to array:

```
[[8 4 6]
```

```
[1 2 3]]
```

9. Write a Python program to append values to the end of an array.

Expected Output:

Original array:

```
[10, 20, 30]
```

After append values to the end of the array:

```
[10 20 30 40 50 60 70 80 90]
```

10. Write a Python program to find the real and imaginary parts of an array of complex numbers.

Expected Output:

Original array [1.00000000+0.j 0.70710678+0.70710678j]

Real part of the array:

```
[ 1. 0.70710678]
```

Imaginary part of the array:

```
[ 0. 0.70710678]
```

11. Write a Python program to find the number of elements of an array, length of one array element in bytes and total bytes consumed by the elements. Expected Output:

Size of the array: 3

Length of one array element in bytes: 8

Total bytes consumed by the elements of the array: 24

12. Write a Python program to find common values between two arrays. Expected Output:

Array1: [0 10 20 40 60]

Array2: [10, 30, 40]

Common values between two arrays:

```
[10 40]
```

13. Write a Python program to find the set difference of two arrays. The set difference will return the sorted, unique values in array1 that are not in array2. Expected Output:

Array1: [0 10 20 40 60 80]

Array2: [10, 30, 40, 50, 70, 90]

Set difference between two arrays:

[0 20 60 80]

14. Write a Python program to find the set exclusive-or of two arrays. Set exclusive-or will return the sorted, unique values that are in only one (not both) of the input arrays. Array1: [0 10 20 40 60 80]

Array2: [10, 30, 40, 50, 70]

Unique values that are in only one (not both) of the input arrays:

[0 20 30 50 60 70 80]

15. Write a Python program compare two arrays using numpy.

Array a: [1 2]

Array b: [4 5]

a > b

[False False]

a >= b

[False False]

a < b

[True True]

a <= b

[True True]

15. Write a Python program to save a NumPy array to a text file.

16. Write a Python program to create a contiguous flattened array.

Original array:

[[10 20 30]

[20 40 50]]

New flattened array:

[10 20 30 20 40 50]

17. Write a Python program to change the data type of an array.

Expected Output:

[[2 4 6]

[6 8 10]]

Data type of the array x is: int32

New Type: float64

[[2. 4. 6.]

[6. 8. 10.]]

18. Write a Python program to create a 3-D array with ones on a diagonal and zeros elsewhere.

Expected Output:

[[1. 0. 0.]

```
[ 0. 1. 0.]  
[ 0. 0. 1.]
```

19. Write a Python program to create an array which looks like below array.

Expected Output:

```
[[ 0. 0. 0.]  
 [ 1. 0. 0.]  
 [ 1. 1. 0.]  
 [ 1. 1. 1.]]
```

20. Write a Python program to concatenate two 2-dimensional arrays.

Expected Output:

Sample arrays: ([[0, 1, 3], [5, 7, 9]], [[0, 2, 4], [6, 8, 10]])

Expected Output:

```
[[ 0 1 3 0 2 4]  
 [ 5 7 9 6 8 10]]
```

21. Write a Python program to make an array immutable (read-only).

Expected Output:

Test the array is read-only or not:

Try to change the value of the first element:

Traceback (most recent call last):

File "19236bd0-0bd9-11e7-a232-c706d0968eb6.py", line 6, in

x[0] = 1

ValueError: assignment destination is read-only

22. Write a Python program to create an array of (3, 4) shape, multiply every element value by 3 and display the new array.

Expected Output:

Original array elements:

```
[[ 0 1 2 3]  
 [ 4 5 6 7]  
 [ 8 9 10 11]]
```

New array elements:

```
[[ 0 3 6 9]  
 [12 15 18 21]  
 [24 27 30 33]]
```

23. Write a Python program to convert a NumPy array into Python list structure.

Expected Output:

Original array elements:

```
[[0 1]
```

```
[2 3]
```

```
[4 5]]
```

Array to list:

```
[[0, 1], [2, 3], [4, 5]]
```

24. Write a Python program to convert a NumPy array into Python list structure.

Expected Output:

Original array elements:

```
[ 0.26153123 0.52760141 0.5718299 0.5927067 0.7831874 0.69746349  
0.35399976 0.99469633 0.0694458 0.54711478]
```

Print array values with precision 3:

```
[ 0.262 0.528 0.572 0.593 0.783 0.697 0.354 0.995 0.069 0.547]
```

25. Write a Python program to suppresses the use of scientific notation for small numbers in numpy array.

Expected Output:

Original array elements:

```
[ 1.60000000e-10 1.60000000e+00 1.20000000e+03 2.35000000e-01]
```

Print array values with precision 3:

```
[ 0. 1.6 1200. 0.235]
```

26. Write a Python program to how to add an extra column to an numpy array.

Expected Output:

```
[[ 10 20 30 100]
```

```
[ 40 50 60 200]]
```

27. Write a Python program to remove specific elements in a numpy array.

Expected Output:

Original array:

```
[ 10 20 30 40 50 60 70 80 90 100]
```

Delete first, fourth and fifth elements:

```
[ 20 30 60 70 80 90 100]
```

Pandas

1. Write a Python program to create and display a one-dimensional array-like

object containing an array of data using Pandas module.

2. Write a Python program to convert a Panda module Series to Python list and its type.
3. Write a Python program to add, subtract, multiple and divide two Pandas Series. Sample Series: [2, 4, 6, 8, 10], [1, 3, 5, 7, 9]

4. Write a Python program to get the powers of an array values element-wise.

Note: First array elements raised to powers from second array

Expected Output:

Original array

```
[0 1 2 3 4 5 6]
```

First array elements raised to powers from second array, element-wise:

```
[ 0 1 8 27 64 125 216]
```

5. Write a Python program to create and display a DataFrame from a specified dictionary data which has the index labels.

Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily',  
'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

6. Write a Python program to display a summary of the basic information about a specified Data Frame and its data.

Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',  
'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}  
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

7. Write a Python program to get the first 3 rows of a given DataFrame.

Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',  
'Laura', 'Kevin', 'Jonas'],  
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
```

```
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

8. Write a Python program to select the 'name' and 'score' columns from the following DataFrame.

Sample Python dictionary data and list labels:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

9. Write a Python program to select the specified columns and rows from a given data frame.

Select 'name' and 'score' columns in rows 1, 3, 5, 6 from the following data frame.

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

10. Write a Python program to select the rows where the number of attempts in the examination is greater than 2.

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

11. Write a Python program to count the number of rows and columns of a DataFrame.
Sample data:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

12. Write a Python program to select the rows where the score is missing, i.e. is NaN. exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

13. Write a Python program to select the rows where number of attempts in the examination is less than 2 and score greater than 15.

14. Write a Python program to change the score in row 'd' to 11.5.
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

15. Write a Python program to calculate the sum of the examination attempts by the students.
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

16. Write a Python program to calculate the mean score for each different student in DataFrame.
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

17. Write a Python program to append a new row 'k' to data frame with given values for each column. Now delete the new row and return the original DataFrame. exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin',

```

'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
Values for each column will be:
name : "Suresh", score: 15.5, attempts: 1, qualify: "yes", label: "k"

```

18. Write a Python program to sort the DataFrame first by 'name' in descending order, then by 'score' in ascending order.

```

exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
Values for each column will be:
name : "Suresh", score: 15.5, attempts: 1, qualify: "yes", label: "k"

```

19. Write a Python program to replace the 'qualify' column contains the values 'yes' and 'no' with True and False.

```

exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

```

20. Write a Python program to delete the 'attempts' column from the DataFrame.

```

exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']

```

21. Write a Python program to insert a new column in existing DataFrame.

```

exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',

```

```
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

22. Write a Python program to iterate over rows in a DataFrame.

23. Write a Python program to get list from DataFrame column headers.

Sample data:

```
exam_data = {'name': ['Anastasia', 'Dima', 'Katherine', 'James', 'Emily', 'Michael', 'Matthew',
'Laura', 'Kevin', 'Jonas'],
'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],
'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],
'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
labels = ['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j']
```

Python Matplotlib

1. Write a Python program to draw a line with suitable label in the x axis, y axis and a title
2. Write a Python program to draw a line using given axis values with suitable label in the x axis , y axis and a title
3. Write a Python program to draw a line using given axis values taken from a text file, with suitable label in the x axis, y axis and a title.

Test Data:

test.txt

```
1 2
2 4
3 1
```

4. Write a Python program to draw line charts of the financial data of Alphabet Inc. between October 3, 2016 to October 7, 2016.

Sample Financial data (fdata.csv):

Date,Open,High,Low,Close

```
03-10-16,774.25,776.065002,769.5,772.559998
04-10-16,776.030029,778.710022,772.890015,776.429993
05-10-16,779.309998,782.070007,775.650024,776.469971
06-10-16,779,780.47998,775.539978,776.859985
```

07-10-16,779.659973,779.659973,770.75,775.080017

5. Write a Python program to plot two or more lines on same plot with suitable legends of each line.

6. Write a Python program to plot two or more lines with legends, different widths and colors.

7. Write a Python program to plot two or more lines with different styles

8. Write a Python program to plot two or more lines and set the line markers. **9.** Write a

Python program to display the current axis limits values and set new axis values. **10.** Write

a Python program to plot quantities which have an x and y position.

11. Write a Python program to plot several lines with different format styles in one command using arrays.

12. Write a Python program to create multiple types of charts

13. Write a Python program to display the grid and draw line charts of the closing value of Alphabet Inc. between October 3, 2016 to October 7, 2016. Customized the grid lines with linestyle -, width .5. and color blue.

Date,Close

03-10-16,772.559998

04-10-16,776.429993

05-10-16,776.469971

06-10-16,776.859985

07-10-16,775.080017

14. Write a Python program to display the grid and draw line charts of the closing value of Alphabet Inc. between October 3, 2016 to October 7, 2016. Customized the grid lines with rendering with a larger grid (major grid) and a smaller grid (minor grid). Turn on the grid but turn off ticks.

15. Write a Python program to create multiple plots.

Matplotlib Barchart

1. Write a Python programming to display a bar chart of the popularity of programming Languages.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

2. Write a Python programming to display a horizontal bar chart of the popularity of programming Languages.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

3. Write a Python programming to display a bar chart of the popularity of programming Languages. Use uniform color.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

4. Write a Python programming to display a bar chart of the popularity of programming Languages. Use different color for each bar.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

5. Write a Python programming to display a bar chart of the popularity of programming Languages. Attach a text label above each bar displaying its popularity (float value).

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

6. Write a Python programming to display a bar chart of the popularity of programming Languages. Make blue border to each bar.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

7. Write a Python programming to display a bar chart of the popularity of programming Languages. Specify the position of each bar plot.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

8. Write a Python programming to display a bar chart of the popularity of programming Languages. Select the width of each bar and their positions.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

9. Write a Python programming to display a bar chart of the popularity of programming Languages. Increase bottom margin.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

10. Write a Python program to create bar plot of scores by group and gender. Use multiple X values on the same chart for men and women.

Sample Data:

Means (men) = (22, 30, 35, 35, 26)

Means (women) = (25, 32, 30, 35, 29)

11. Write a Python program to create bar plot from a DataFrame.

Sample Data Frame:

a b c d e

2 4,8,5,7,6

4 2,3,4,2,6

6 4,7,4,7,8

8 2,6,4,8,6

10 2,4,3,3,2

12. Write a Python program to create bar plots with error bars on the same figure. Sample Date

Mean velocity: 0.2474, 0.1235, 0.1737, 0.1824

Standard deviation of velocity: 0.3314, 0.2278, 0.2836, 0.2645

13. Write a Python program to create bar plots with errorbars on the same figure. Attach a text label above each bar displaying men means (integer value).

Sample Data

Mean velocity: 0.2474, 0.1235, 0.1737, 0.1824

Standard deviation of velocity: 0.3314, 0.2278, 0.2836, 0.2645

Matplotlib Piechart

1. Write a Python programming to create a pie chart of the popularity of programming Languages.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

2. Write a Python programming to create a pie chart with a title of the popularity of programming Languages.

Sample data:

Programming languages: Java, Python, PHP, JavaScript, C#, C++

Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7

3. Write a Python programming to create a pie chart of gold medal achievements of five most successful countries in 2016 Summer Olympics. Read the data from a csv file. Sample data:

medal.csv

country,gold_medal

United States,46

Great Britain,27

China,26

Russia,19

Germany,17

Matplotlib Scatterplot

1. Write a Python program to draw a scatter graph taking a random distribution in X and Y and plotted against each other.

2. Write a Python program to draw a scatter plot with empty circles taking a random distribution in X and Y and plotted against each other.

3. Write a Python program to draw a scatter plot using random distributions to generate balls of different sizes.

4. Write a Python program to draw a scatter plot comparing two subject marks of Mathematics and Science. Use marks of 10 students.

Test Data:

math_marks = [88, 92, 80, 89, 100, 80, 60, 100, 80, 34]

science_marks = [35, 79, 79, 48, 100, 88, 32, 45, 20, 30]

marks_range = [10, 20, 30, 40, 50, 60, 70, 80, 90, 100]

5. Write a Python program to draw a scatter plot for three different groups comparing weights and heights.

Seaborn Bar plot

1. Write a program to draw bar plot of sex against survived for a dataset given in the url

<https://github.com/mwaskom/seaborn-data/blob/master/titanic.csv>

Seaborn pointplot

1. Write a program to draw a point plot for sex against survived for a dataset given in url

<https://github.com/mwaskom/seaborn-data/blob/master/titanic.csv>

Seaborn Scatterplot

1. Write a program to draw a scatter plot of “day” against “total bill” for a dataset given in a url

<https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv>

Seaborn Violin Plot

1. Write a program to draw a violin plot of sex against total_bill for a given dataset

<https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv>

2. Write a program to draw a violin plot of “species” against “sepal length” for a dataset given in a url

<https://github.com/mwaskom/seaborn-data/blob/master/iris.csv>

Seaborn BoxPlot

1. Write a program to draw box plot of life expectancy by continent for a data set given in a url

<https://raw.githubusercontent.com/resbaz/r-novice-gapminder-files/master/data/gapminder-FiveYearData.csv>

2. Write a program to draw a box plot of day by tips for a dataset given in a url

<https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv>

Seaborn Swarm Plot

1. Write a program to draw a swarm plot of total bill against size for a given dataset

<https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv> 2. Write a

program to draw swarm plot of “total bill” against day for a dataset given in url

<https://raw.githubusercontent.com/mwaskom/seaborn-data/master/tips.csv>

Plotly Scatterplot

1. Write a program to draw a scatter plot for random 1000 x and y coordinates 2. Write

a program to draw line and scatter plots for random 100 x and y coordinates 3. Write a

program to draw a scatter plot for random 500 x and y coordinates and style it

4. Write a program to draw a scatter plot for a given dataset and show datalabels on hover

https://raw.githubusercontent.com/plotly/datasets/master/2014_usa_states.csv

ML Types

Classify the problems as supervised, unsupervised, reinforcement or semi supervised

- a. Spam filtering: Is an email spam or not
- b. Given a list of customers and information about them, discover groups of similar users. This knowledge can then be used for targeted marketing
- c. Robotics: A robot is in a maze, and it needs to find a way out.
- d. Training an AI for a complex game such as Civilization or Dota
- e. Anomaly detection: Given measurements from sensors in a manufacturing facility, identify anomalies, i.e. that something is wrong
- f. Discover patterns in data such as whenever it rains, people tend to stay indoors. When it is hot, people buy more ice-cream.
- g. Given information about a house, predict its price
- h. Netflix: Given a user and a movie, predict the rating the user is going to give to the movie
- i. Given an image, output which objects are present in the image

Data Preprocessing Steps

1. Apply following steps to dataset given in a url

https://drive.google.com/open?id=1NKMy-zIT3tfpNLnA7G0EmPxgZe0OPXp_

- a. Handling missing data
- b. Handling categorical data
- c. Split the dataset into training set and test set
- d. Feature scaling

Regression

Simple Linear Regression

1. Build a predictive linear regression model for given dataset, train the model for training set and test it against test dataset, plot the model using any plotting library.

Dataset url -

<https://drive.google.com/open?id=17Z5YVgk4hSzPvguWkck6tRb6Z2JEWdgh>

2. Build a predictive linear regression model for given dataset, given humidity predict apparent temperature

<https://drive.google.com/open?id=1WsJxbsh51SL1UhT0xEvNynZTKy4nOwLM>

Multiple Linear Regression

1. Build a predictive linear regression model for given dataset, given temperature, humidity, wind speed, wind bearing, visibility, pressure predict apparent temperature

<https://drive.google.com/open?id=1WsJxbsh51SL1UhT0xEvNynZTKy4nOwLM>

Classification

Logistic Regression

1. The data given in the url is related with direct marketing campaigns of a banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. Build a model to predict whether client will subscribe to term deposit

<https://drive.google.com/open?id=1E0EURgsF3L9Bt5hnOalE0d4Tw9mIOgTn>

2. Build a machine learning model to predict whether a person makes over 50K a year or not

<https://drive.google.com/open?id=1XVi34snXnh6qW2u71jwc05oCRnpOXoWF>

K- Nearest Neighbor (KNN)

1. The data given in the url is related with direct marketing campaigns of a banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. Build a model to predict whether client will subscribe to term deposit

<https://drive.google.com/open?id=1E0EURgsF3L9Bt5hnOalE0d4Tw9mIOgTn>

2. Build a machine learning model to predict whether a person makes over 50K a year or not

<https://drive.google.com/open?id=1XVi34snXnh6qW2u71jwc05oCRnpOXoWF>

Neural Networks

Single Layer Neural Network

1. The data given in the url is related with direct marketing campaigns of a banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. Build a model to predict whether client will subscribe to term deposit

<https://drive.google.com/open?id=1E0EURgsF3L9Bt5hnOalE0d4Tw9mIOgTn>

2. For a given dataset predict whether customer will exit (Output variable "Exited") the bank or not using artificial neural network

<https://drive.google.com/open?id=1aYpqp7vGd1vy76fhP-2lQJywn0BLt3eM>

Multi Layer Neural Network

1. The data given in the url is related with direct marketing campaigns of a banking institution. The marketing campaigns were based on phone calls. Often, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be ('yes') or not ('no') subscribed. Build a model to predict whether client will subscribe to term deposit

<https://drive.google.com/open?id=1E0EURgsF3L9Bt5hnOalE0d4Tw9mIOgTn>

2. For a given dataset predict whether customer will exit (Output variable "Exited") the bank or not using artificial neural network

<https://drive.google.com/open?id=1aYpqp7vGd1vy76fhP-2lQJywn0BLt3eM>

ML Using Libraries

Regression

Simple Linear Regression

1. Build a machine learning model to predict salary based on experience for a given dataset

<https://drive.google.com/open?id=1UAbkc53H9fpJiEKCUIcw1kXsrCdc0uU1>

2. For a given dataset predict number of bikes getting shared based on temperature of the day

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Multiple Linear Regression

1. Build a machine learning model to predict profit of the company based on different expenses for a given dataset

https://drive.google.com/open?id=1dic5dbHugytxCO6i9pcD6RPod_sLkA

[UA](#)

2. For a given dataset predict number of bikes getting shared based on different parameters

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Polynomial Regression

1. Build a machine learning model to predict salary based on position for a given dataset

[ol3](https://drive.google.com/open?id=1jKfNxQkybZKprVpCkoL16mTjmSSnC)

2. For a given dataset predict number of bikes getting shared based on different parameters

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Support Vector Regression

1. Build a machine learning model to predict salary based on position for a given dataset

[ol3](https://drive.google.com/open?id=1jKfNxQkybZKprVpCkoL16mTjmSSnC)

2. For a given dataset predict number of bikes getting shared based on different parameters

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Decision Tree Regression

1. Build a machine learning model to predict salary based on position for a given dataset

[ol3](https://drive.google.com/open?id=1jKfNxQkybZKprVpCkoL16mTjmSSnC)

2. For a given dataset predict number of bikes getting shared based on different parameters

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Random Forest Regression

1. Build a machine learning model to predict salary based on position for a given dataset

<https://drive.google.com/open?id=1jKfNxQkybZKprVpCkoL16mTjmSSnC0l3>

2. For a given dataset predict number of bikes getting shared based on different parameters

<https://drive.google.com/open?id=1ohN2o3zSZ2Xuy4CldTWN-dTmciUXizst>

Classification

Logistic Regression

1. Build a machine learning model to predict user will click the ad or not based on his experience and estimated salary for a given dataset.

<https://drive.google.com/open?id=1l8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

2. The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). Build a machine learning model for the dataset, please refer document inside zip file for additional information

https://drive.google.com/open?id=1Vm_zruT2djYympL0IG6_I57kAEhF5hXH

K- Nearest Neighbor

1. Build a machine learning model to predict user will click the ad or not based on his experience and estimated salary for a given dataset.

<https://drive.google.com/open?id=1l8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

2. The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). Build a machine learning model for the dataset, please refer document inside zip file for additional information

https://drive.google.com/open?id=1Vm_zruT2djYympL0IG6_I57kAEhF5hxH

Support Vector Machine

1. Build a machine learning model to predict user will click the ad or not based on his experience and estimated salary for a given dataset.

<https://drive.google.com/open?id=1I8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

2. The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). Build a machine learning model for the dataset, please refer document inside zip file for additional information

https://drive.google.com/open?id=1Vm_zruT2djYympL0IG6_I57kAEhF5hxH

Decision Tree Classification

1. Build a machine learning model to predict user will click the ad or not based on his experience and estimated salary for a given dataset.

<https://drive.google.com/open?id=1I8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

2. The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). Build a machine learning model for the dataset, please refer document inside zip file for additional information

https://drive.google.com/open?id=1Vm_zruT2djYympL0IG6_I57kAEhF5hxH

Random Forest Classification

1. Build a machine learning model to predict user will click the ad or not

based on his experience and estimated salary for a given dataset.

<https://drive.google.com/open?id=1l8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

2. The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). Build a machine learning model for the dataset, please refer document inside zip file for additional information

https://drive.google.com/open?id=1Vm_zruT2djYympL0IG6_I57kAEhF5hxH

Clustering

K-Means Clustering

1. Build a machine learning model to create group of mall customers based on their annual income and spending score for a given dataset

https://drive.google.com/open?id=1Lp3yU_ggl-OQSm-t4qbLs-8QzWmxRoc0

2. Apply K-Means clustering on below dataset

<https://archive.ics.uci.edu/ml/datasets/US+Census+Data+%281990%29>

Hierarchical Clustering

1. Build a machine learning model to create group of mall customers based on their annual income and spending score for a given dataset

https://drive.google.com/open?id=1Lp3yU_ggl-OQSm-t4qbLs-8QzWmxRoc0

2. Apply Hierarchical clustering on below dataset

<https://archive.ics.uci.edu/ml/datasets/US+Census+Data+%281990%29>

Dimensionality Reduction Techniques

Principal Component Analysis

1. Apply PCA for a dataset given in the url, build a classification

model and plot the graph

<https://drive.google.com/open?id=1a7zQj3ipKcp6jJikPv-exaTbrbbWy36x>

Linear Discriminant Analysis

1. Apply LDA for a dataset given in the url, build a classification model and plot the graph

<https://drive.google.com/open?id=1a7zQj3ipKcp6jJikPv-exaTbrbbWy36x>

Kernel PCA

1. Apply Kernel PCA for a dataset given in the url, build a classification model and plot the graph

<https://drive.google.com/open?id=1l8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

Reinforcement Learning

Upper Confidence Bound

1. Apply UCB to find ad which will perform better among other ads to improve ad CTR for a dataset in the given url

<https://drive.google.com/open?id=1Ntjt3oAXXafimA0vmUndY7X0SFqLdsb>

Thompson Sampling

1. Apply UCB to find ad which will perform better among other ads to improve ad CTR for a dataset in the given url

<https://drive.google.com/open?id=1Ntjt3oAXXafimA0vmUndY7X0SFqLdsb>

Association Rule Learning

Apriori

1. Generate association rules for dataset given in the url

<https://drive.google.com/open?id=1ke8crGAhKwSx1xbanwr8n5gNwpqOd>

[DXU](#)

Model Selection

K-Fold cross validation

1. Fit the model using SVM and apply k-fold cross validation technique to check accuracies for a dataset given in the url

<https://drive.google.com/open?id=1I8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

Grid Search

1. Fit the model using SVM and apply Grid search technique to find best model and best parameters for a dataset given in the url

<https://drive.google.com/open?id=1I8KsCufEa47XvzrkxhntEWSy1Su0E0NY>

Natural Language Processing

1. Using NLP predict whether the review is positive or negative for a given dataset

<https://drive.google.com/open?id=1-TJWzdxapGhp2aElncd6RH6zOpSAf69X>

Deep Learning

Artificial Neural Network

1. For a given dataset predict whether customer will exit (Output variable "Exited") the bank or not using artificial neural network

<https://drive.google.com/open?id=1aYqp7vGd1vy76fhP-2IQJywn0BLt3eM>

Convolutional Neural Network

1. Apply CNN to predict cat and dog from given image dataset

https://drive.google.com/open?id=1uWqERYD4wO4s-9U8-j6nKc_dxNtDbMT6

2. We need to identify the digit in given images. We have total 70,000

images, out of which 49,000 are part of train images with the label of digit and rest 21,000 images are unlabeled (known as test images). Now, We need to identify the digit for test images. Public and Private split for test images are 40:60 and evaluation metric of this challenge is accuracy

https://drive.google.com/open?id=1-TavdjU2ohg5T6ZN1KUiWE9BNWu9_vhi

XGBOOST

Linux Commands

Web scrapping

Additional Problems

1. The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class

URL - <https://drive.google.com/open?id=1-jZsOVcuBFB3PyTXWTo3n8qZaQ-skkxg>

2. Fashion-MNIST consists of 60,000 training images and 10,000 test images. It is a MNIST-like fashion product database. The developers believe MNIST has been overused so they created this as a direct replacement for that dataset. Each image is in greyscale and associated with a label from 10 classes.

URL - https://drive.google.com/open?id=1qr-EqIDV3FCj6_7iR1xx8XSd9n1yPdfB

3. Large Movie Review Dataset

We provide a set of 25,000 highly polar movie reviews for training, and 25,000 for testing. There is additional unlabeled data for use as well. Raw text and already processed bag of words formats are provided

URL - <https://drive.google.com/open?id=1UuFBGN8qEe01uldwRHF4-X66fz6dC0Gb>

4. Solve the task of identifying spoken digits in audio samples

<https://drive.google.com/open?id=1Rz807Ysl6A3wYFkorz6RGn9v-KYQPJuz>

5. VoxCeleb is a large-scale speaker identification dataset. It contains around 100,000 utterances by 1,251 celebrities, extracted from YouTube videos. The data is mostly gender balanced (males comprise of 55%). The celebrities span a diverse range of accents, professions and age. There is no overlap between the development and test sets. It's an intriguing use case for isolating and identifying which superstar the voice belongs to

<http://www.robots.ox.ac.uk/~vgg/data/voxceleb/vox1.html>

6. Twitter Sentiment Analysis

Hate Speech in the form of racism and sexism has become a nuisance on twitter and it is important to segregate these sort of tweets from the rest. We provide Twitter data that has both normal and hate tweets. Your task as a Data Scientist is to identify the tweets which are hate tweets and which are not.

<https://drive.google.com/open?id=1FbD-yB1E2tQ7SLtHSaYjlj851c2oPUEB>

7. Age Detection of Indian Actors

The dataset contains thousands of images of Indian actors and your task is to identify their age. All the images are manually selected and cropped from the video frames resulting in a high degree of variability in terms of scale, pose, expression, illumination, age, resolution, occlusion, and makeup.

<https://drive.google.com/open?id=1eOMF9c-wGb5oRemmRflqz2p8vWJ7r4qS>