Machine Learning (Assignment 1)

700743830 – CHEVUTUKUR SAI KRISHNA REDDY

Source code link: https://github.com/csk17/ML-Assignment-1

Question 1

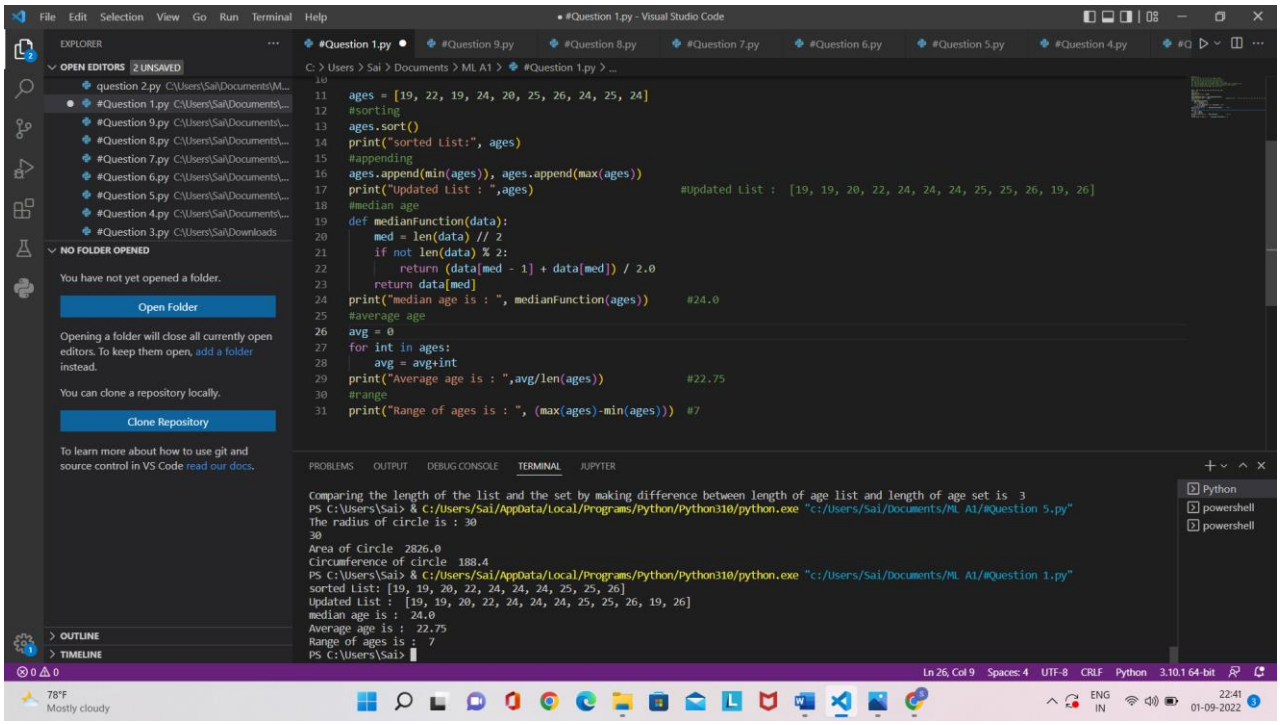
The following is a list of 10 students ages:

ages = [19, 22, 19, 24, 20, 25, 26, 24, 25, 24]

- Sort the list and find the min and max age
- Add the min age and the max age again to the list
- Find the median age (one middle item or two middle items divided by two)
- Find the average age (sum of all items divided by their number)
- Find the range of the ages (max minus min)

Solution:

- From the given list of age(), the list is sorted by using 'sort()' function. The minimum and maximum age is obtained by using min() and max() function respectively.
- The list was updated by adding min and max age again to the list by using 'append()' function.
- The median of age is obtained by using 'median()' function and here the 'if' statement is used as to perform for two cases i.e., in even and odd case.
- The average age is obtained by using division operation for sum and total number of observations.
- The range of ages is obtained by using difference of max() and min() functions of ages.

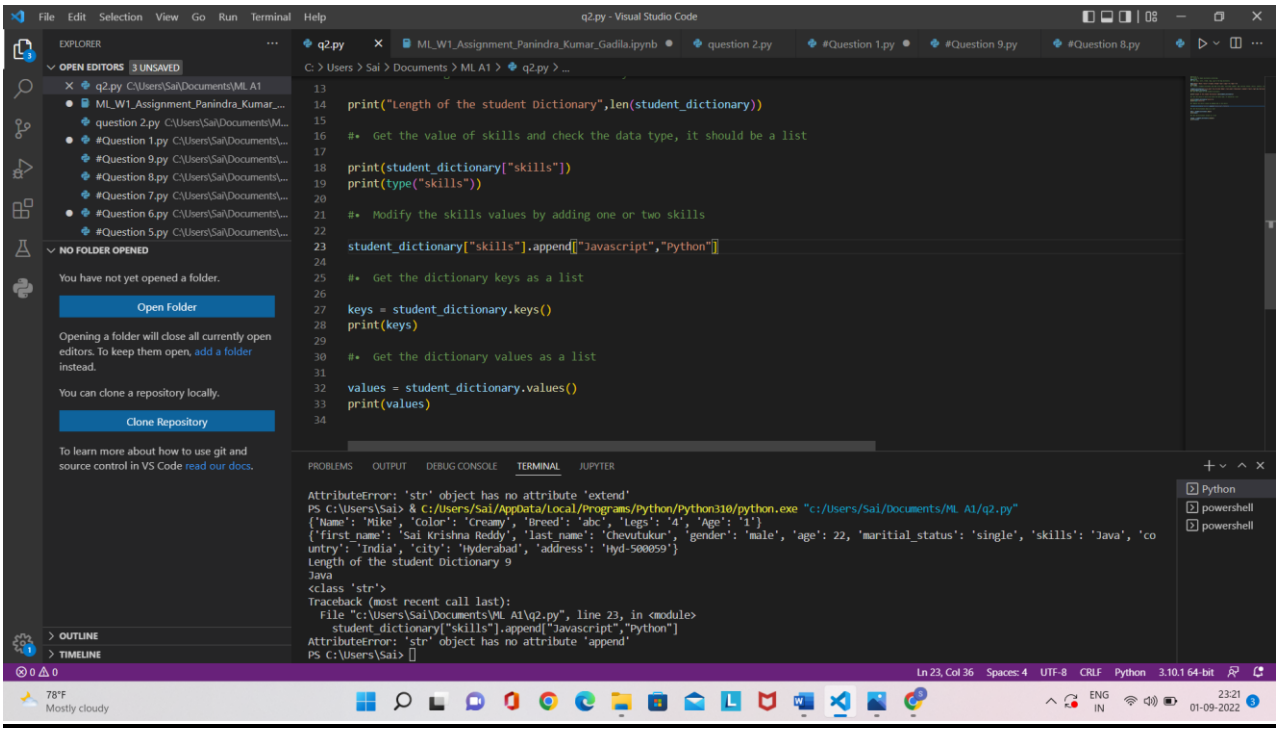


Question 2

- Create an empty dictionary called dog
- Add name, color, breed, legs, age to the dog dictionary
- Create a student dictionary and add first_name, last_name, gender, age, marital status, skills, country, city and address as keys for the dictionary
- Get the length of the student dictionary
- Get the value of skills and check the data type, it should be a list
- Modify the skills values by adding one or two skills
- Get the dictionary keys as a list
- Get the dictionary values as a list

Solution:

- An empty dictionary is created by naming ‘dog’.
- The name colour breed legs age were added to the dictionary.
- The student dictionary is also created and the details of first name, last name, gender, age, marital status, skills, country, city and address were mentioned in the dictionary.
- The length of student dictionary is obtained by using ‘len()’ function.
- The value of skills were obtained and data type is checked in a list.
- The skill values were modified by using ‘update()’ function.
- The dictionary Keys and Values are obtained as a list.



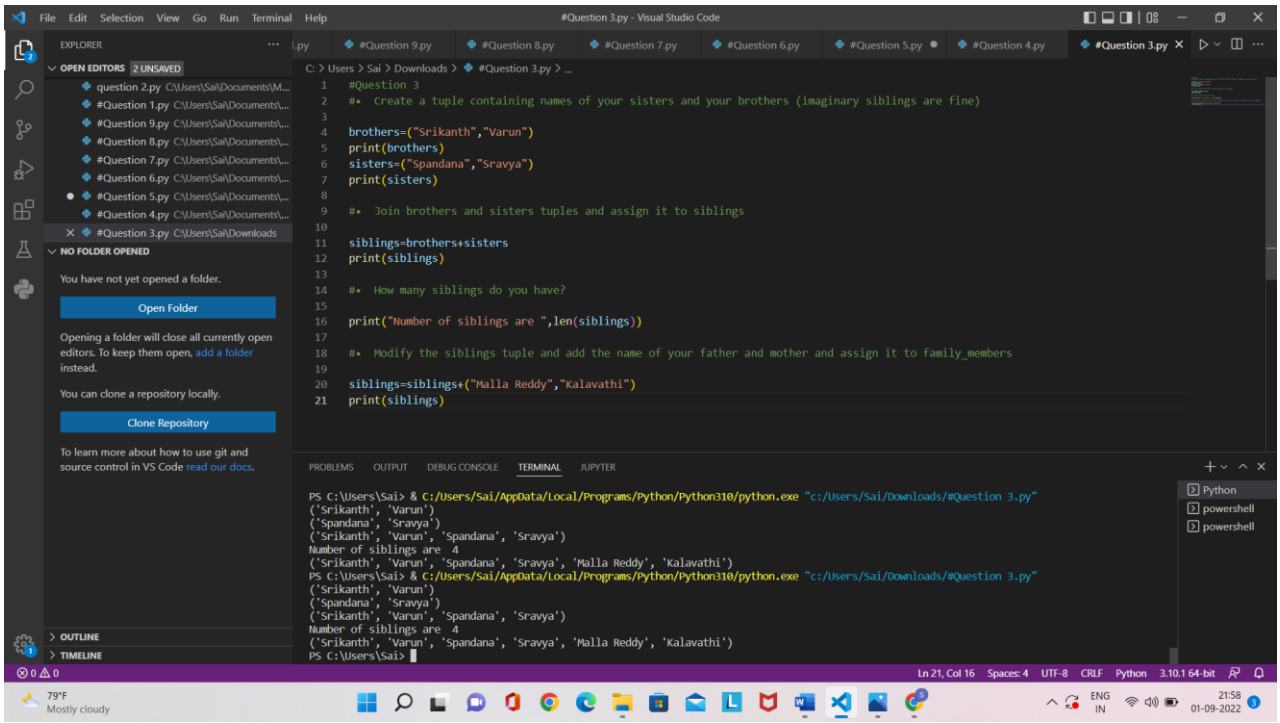
Question 3

- Create a tuple containing names of your sisters and your brothers (imaginary siblings are fine)
- Join brothers and sisters tuples and assign it to siblings
- How many siblings do you have?

- Modify the siblings tuple and add the name of your father and mother and assign it to family_members

Solution:

- The tuple was created with the names of sisters and brothers.
- Now, both the tuples were joined by using addition (+) operator and named as siblings.
- I have founded the total number of siblings by using ‘len’ function and printed the output.
- The tuple of siblings is modified by adding the name of my mother and father by using ‘+’ operator.



Question 4

it_companies = {'Facebook', 'Google', 'Microsoft', 'Apple', 'IBM', 'Oracle', 'Amazon'}

A = {19, 22, 24, 20, 25, 26}

B = {19, 22, 20, 25, 26, 24, 28, 27}

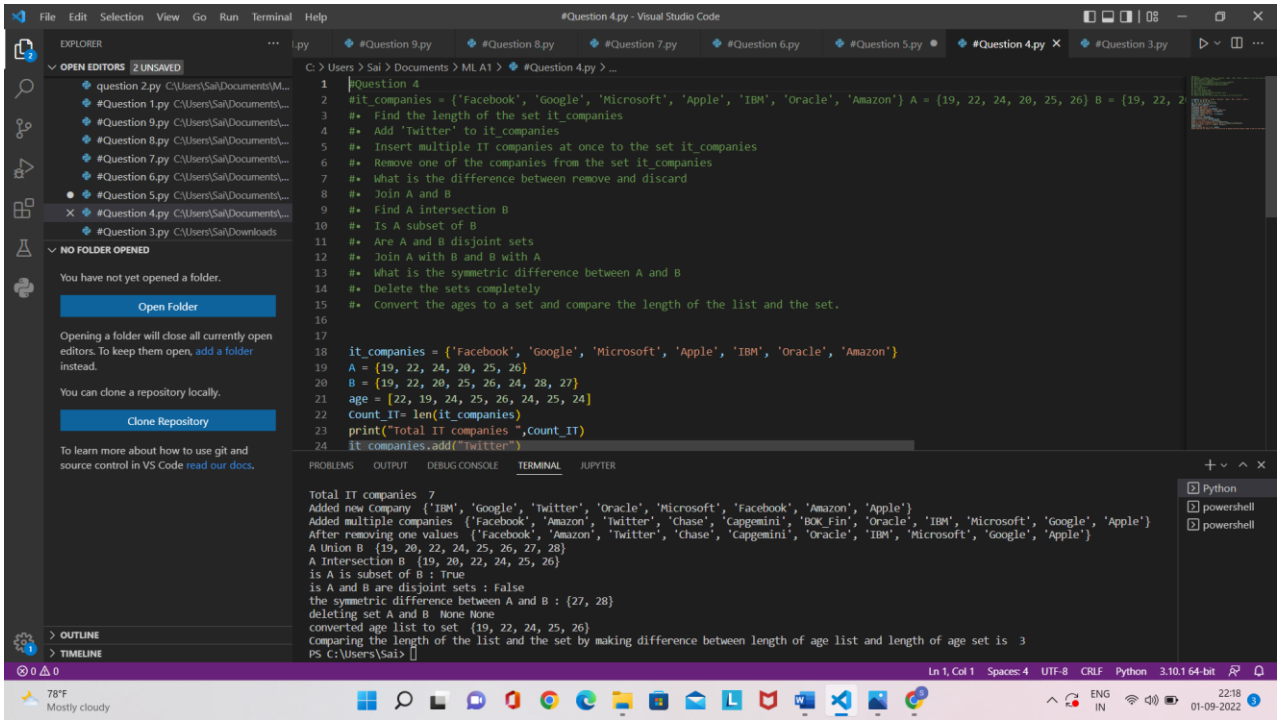
age = [22, 19, 24, 25, 26, 24, 25, 24]

- Find the length of the set it_companies
- Add 'Twitter' to it_companies
- Insert multiple IT companies at once to the set it_companies

- Remove one of the companies from the set it_companies
- What is the difference between remove and discard
- Join A and B
- Find A intersection B
- Is A subset of B
- Are A and B disjoint sets
- Join A with B and B with A
- What is the symmetric difference between A and B
- Delete the sets completely
- Convert the ages to a set and compare the length of the list and the set.

Solution:

- From the given string of it_companies the length was obtained by using 'len' function
- The Twitter added to the it_companies.by using 'add' function.
- Some of multiple IT companies are inserted to it_companies by using 'update' function.
- Removed one of it by using 'remove' function.
- The difference between remove and discard is the discard() method removes the specified item from the set where the remove() method raises an error if the specified item does not exist as the discard() method won't.
- Both A and B were joined by using 'union' function.
- The Intersection of A and B is obtained by using 'intersection' function.
- Yes, A is a subset of B as we obtained by using 'issubset()' function.
- No, A and B are not disjoint sets as it obtained by using 'isdisjoint()' function.
- Joined A with B and B with A.
- The symmetric difference between A and B is {27,28} as it obtained by using 'symmetric_difference()' function.
- The sets were deleted by using 'clear()' function.
- The ages to a set was converted and compared the length of list and set by using 'set()' function.



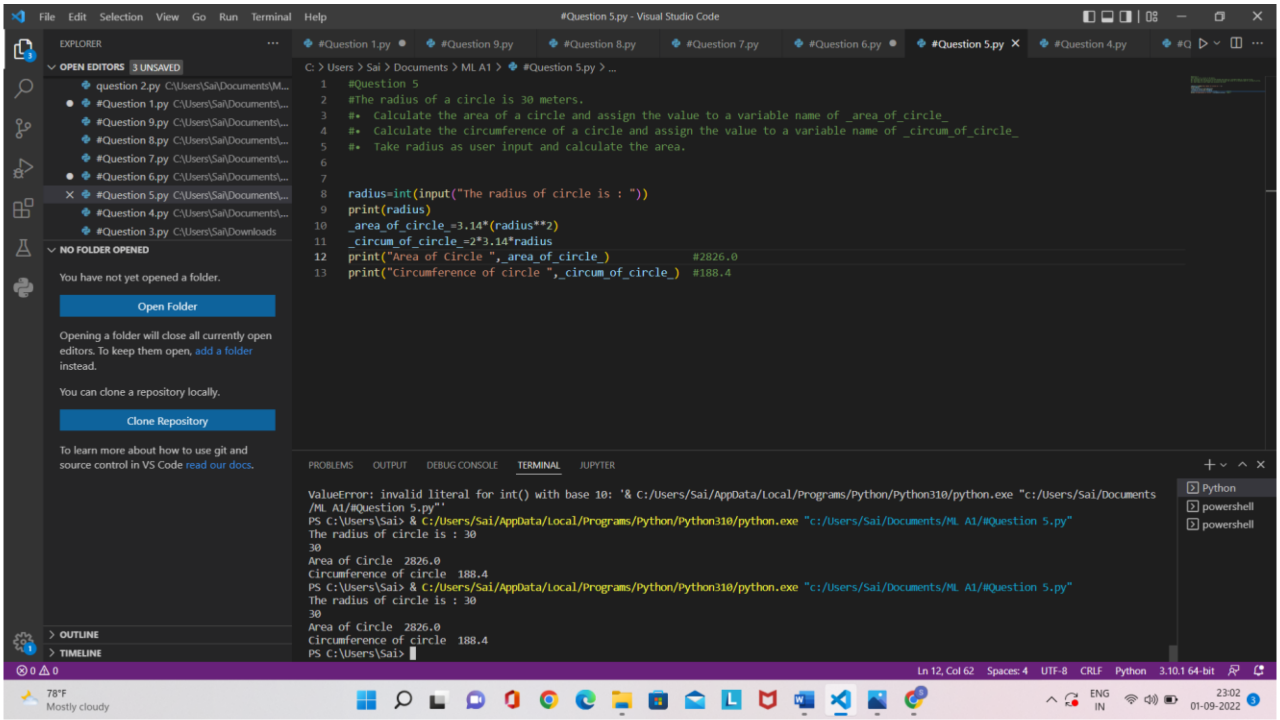
Question 5

The radius of a circle is 30 meters.

- Calculate the area of a circle and assign the value to a variable name of _area_of_circle_
- Calculate the circumference of a circle and assign the value to a variable name of _circum_of_circle_
- Take radius as user input and calculate the area.

Solution:

- The radius of circle is a user input as it assigned using ‘input()’ function.
- The area and circumference of a circle is calculated and assigned it as _area_of_circle_ and _circum_of_circle_ respectively.



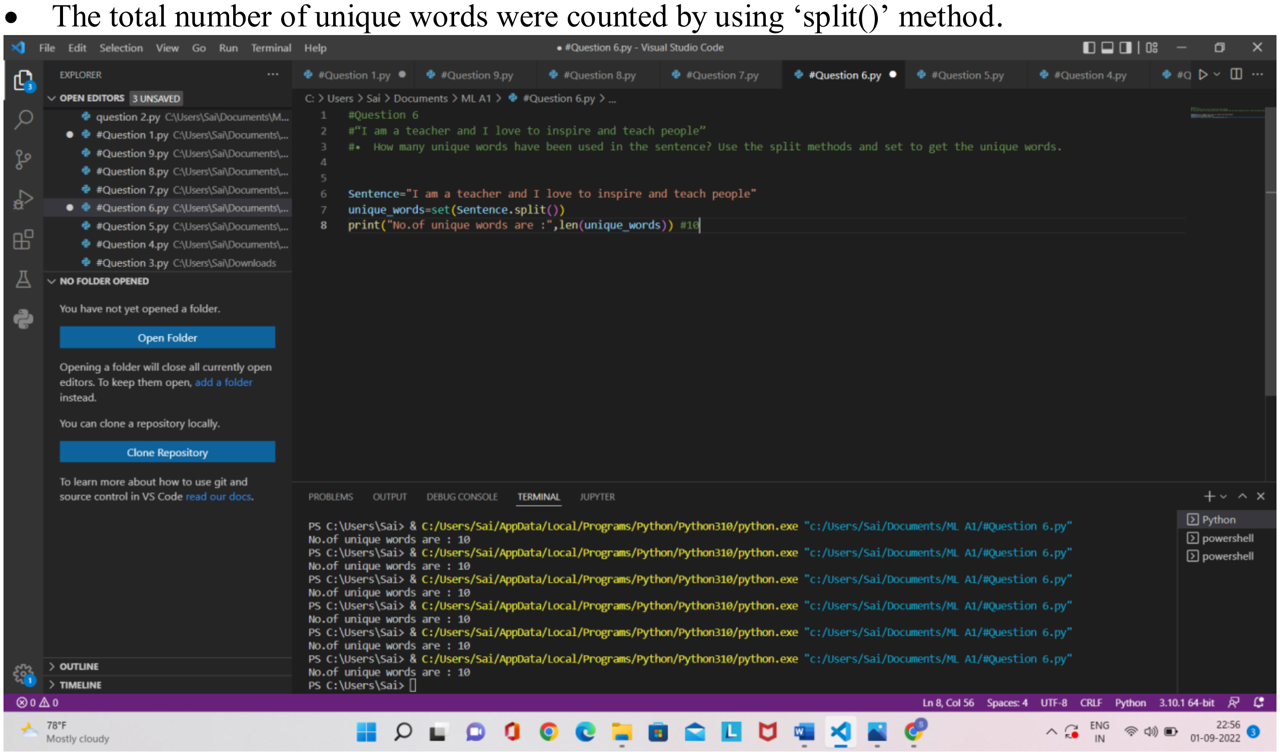
Question 6

"I am a teacher and I love to inspire and teach people"

- How many unique words have been used in the sentence? Use the split methods and set to get the unique words.

Solution:

- The total number of unique words were counted by using 'split()' method.



Question 7

Use a tab escape sequence to get the following lines.

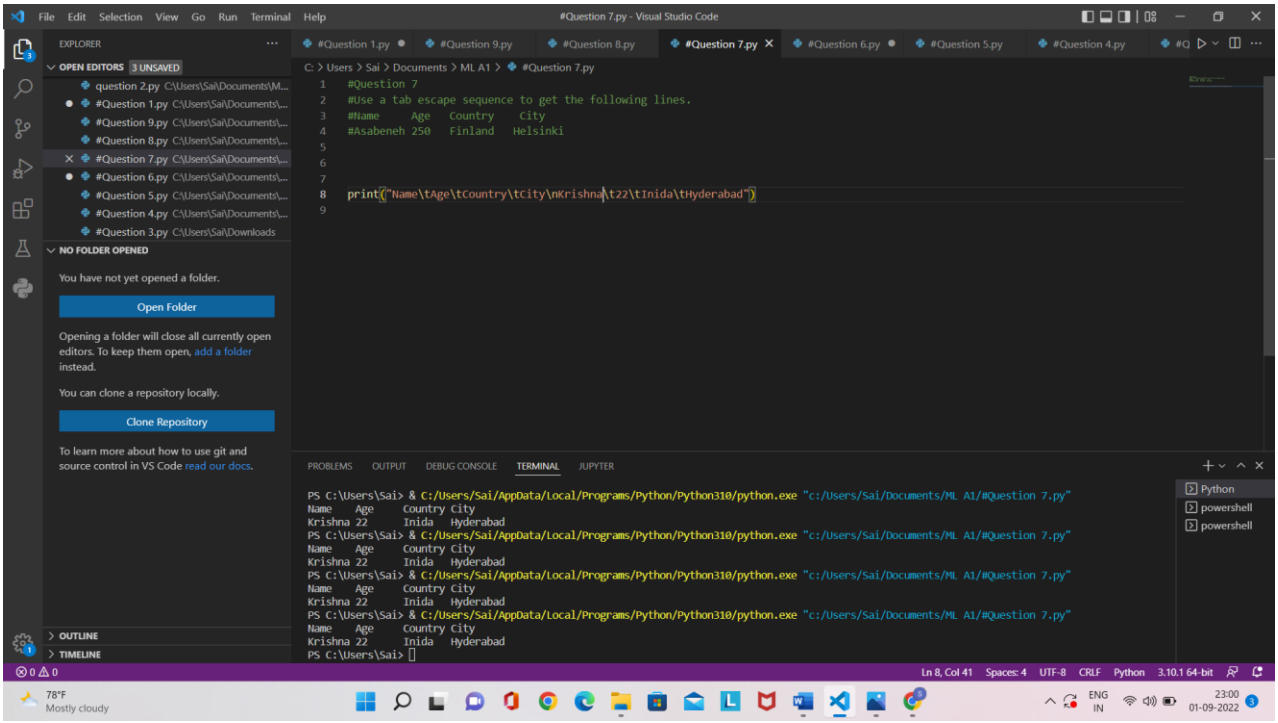
Name Age Country City

Asabeneh 250 Finland Helsinki

Solution:

- By using tab escape the above

  sentence is obtained.



Question 8

Use the string formatting method to display the following:

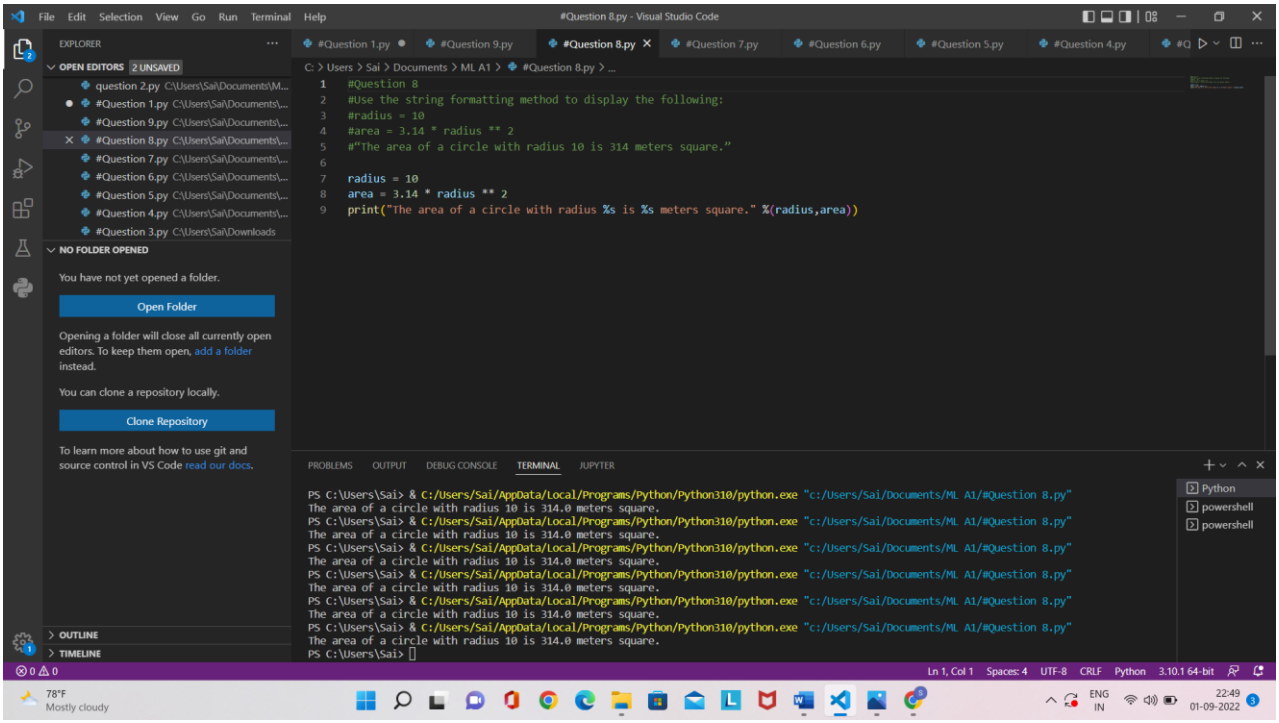
radius = 10

area = 3.14 * radius ** 2

“The area of a circle with radius 10 is 314 meters square.”

Solution:

- The string formatting method is used to display the radius and area as defined above.



Question 9

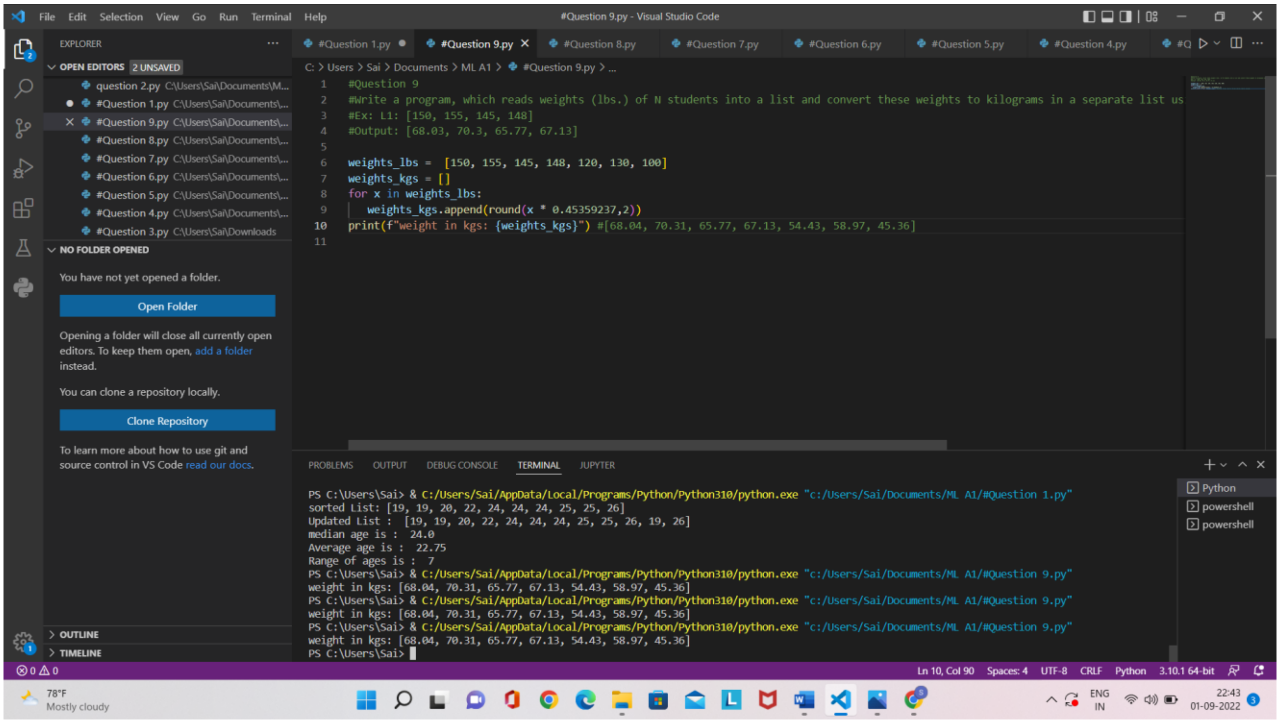
Write a program, which reads weights (lbs.) of N students into a list and convert these weights to kilograms in a separate list using Loop. N: No of students (Read input from user)

Ex: L1: [150, 155, 145, 148]

Output: [68.03, 70.3, 65.77, 67.13]

Solution:

- The weights in lbs is converted to kilograms by using ‘weights_kgs.append(round(x * 0.45359237,2))’ function.



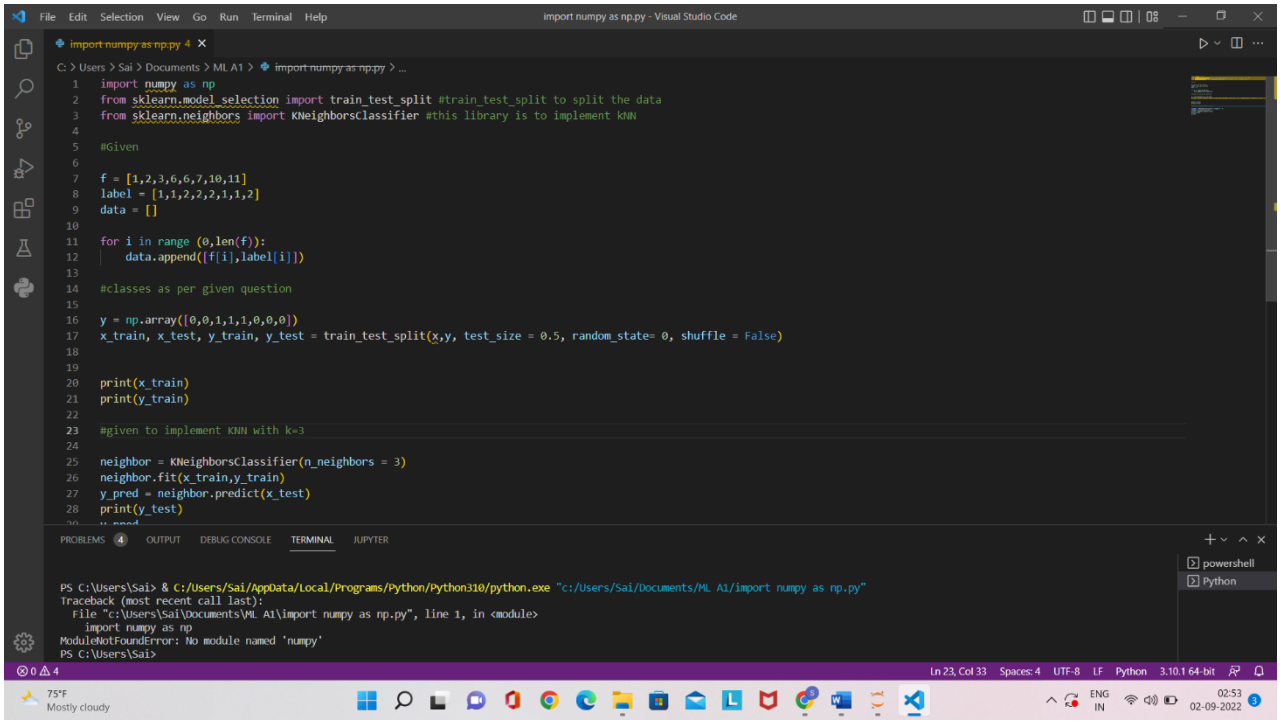
Question 10

The diagram below shows a dataset with 2 classes and 8 data points, each with only one feature value, labeled f. Note that there are two data points with the same feature value of 6. These are shown as two x's one above the other.

1. Divide this data equally into two parts. Use first part as training and second part as testing. Using KNN classifier, for K=3, what would be the predicted outputs for the test samples? Show how you arrived at your answer.
2. Compute the confusion matrix for this and calculate accuracy, sensitivity and specificity values.

Solution:

- Firstly, I have imported numpy, and from sklearn.model_selection imported the train_test_split which is used to split the data and from sklearn.neighbors imported KNeighborsClassifier which is used to implement KNN.
- Now, with the given data implemented the code by using 'for' statement and classes as per required data.
- Finally, implemented KNN with 3 and trained the model with x_train and its label y train. Then used x_test into neighbors.predict() method.
- The output is obtained by printing y_pred.


File Edit Selection View Go Run Terminal Help
import numpy as np.py - Visual Studio Code
import numpy as np.py 4
C: > Users > Sai > Documents > ML A1 > import numpy as np.py > ...
1 import numpy as np
2 from sklearn.model_selection import train_test_split #train_test_split to split the data
3 from sklearn.neighbors import KNeighborsClassifier #this library is to implement kNN
4
5 #Given
6
7 f = [1,2,3,6,6,7,10,11]
8 label = [1,1,2,2,2,1,1,2]
9 data = []
10
11 for i in range (0,len(f)):
12 data.append([f[i],label[i]])
13
14 #classes as per given question
15
16 y = np.array([0,0,1,1,1,0,0,0])
17 x_train, x_test, y_train, y_test = train_test_split(x,y, test_size = 0.5, random_state= 0, shuffle = False)
18
19
20 print(x_train)
21 print(y_train)
22
23 #given to implement KNN with k=3
24
25 neighbor = KNeighborsClassifier(n_neighbors = 3)
26 neighbor.fit(x_train,y_train)
27 y_pred = neighbor.predict(x_test)
28 print(y_test)
PROBLEMS 4 OUTPUT DEBUG CONSOLE TERMINAL JUPYTER
powershell
Python
PS C:\Users\Sai> & C:/Users/Sai/AppData/Local/Programs/Python/Python310/python.exe "c:/Users/Sai/Documents/ML A1/import numpy as np.py"
Traceback (most recent call last):
File "c:\Users\Sai\Documents\ML A1\import numpy as np.py", line 1, in <module>
import numpy as np
ModuleNotFoundError: No module named 'numpy'
PS C:\Users\Sai>
0 4
Ln 23, Col 33 Spaces: 4 UTF-8 LF Python 3.10.1 64-bit
75°F
Mostly cloudy
ENG
IN
02:53
02-09-2022