Machine Learning Assignment 1

Name: Akhila Boddu

Student ID: 700742171

Github link: https://github.com/AkhilaBoddu/ML-Assignment1.git

Videolink:
https://drive.google.com/file/d/1q4CsBd8W1n9CEigrSwRutyDguXGaeAFY/view?usp=share_link

Question1

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)

Source Code:

```
# Importing library called statistics which helps in calculating mathematical data
import statistics
ages = [19, 22, 19, 24, 20, 25, 26, 24, 25, 24]
# Sorting the List of ages in order
ages.sort()
print(ages)      # Displays sorted values
# Minimum age
print("The minimum age in the list is %d" %min(ages))
# Displays min value as we used min() method
# Maximum age
print("The maximum age in the list is %d"%max(ages))
# Displays max value as we used max() method
# Adding min and max values again so we use append() method to insert values to the list
```

ages.append(min(ages))

ages.append(max(ages))

ages.sort()

print(ages) **# Displays min and max value again to the sorted list**

 # Average age

print("The average of the ages is %d"%(sum(ages)/len(ages)))

 # Range

print("The range of the ages in the list is %d"%(max(ages)-min(ages)))

 # Median (one middle item or two middle items divided by two, as we imported statistics library it calculates easily and provides the opt)

print ("The median of the ages is %d"%(statistics.median(ages)))

Description:

In the above source code firstly we imported a library called statistics which will be useful to find the median later. Then we have inserted a data with integer type in list, we performed finding minimum and maximum age using min() and max() functions on next we added them using append() and then we calculated median, average and range using formulas and function.

```
Question1.py > ...
1   import statistics
2   ages = [19, 22, 19, 24, 20, 25, 26, 24, 25, 24]
3   # Sorting the List of ages in order
4   ages.sort()
5   print(ages)
6   print("The minimum age in the list is %d" %min(ages))
7   print("The maximum age in the list is %d"%max(ages))
8   # Adding the min and max values of ages to the list
9   ages.append(min(ages))
10  ages.append(max(ages))
11  ages.sort()
12  print(ages)
13  print("The average of the ages is %d"%(sum(ages)/len(ages)))
14  print("The range of the ages in the list is %d"%(max(ages)-min(ages)))
15  print ("The median of the ages is %d"%(statistics.median(ages)))
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL

\Local\Microsoft\WindowsApps\python3.10.exe' 'c:\Users\akhia\.vscode\extensions\ms-python.python-2022.20.2\pyth
The maximum age in the list is 26
[19, 19, 19, 20, 22, 24, 24, 24, 25, 25, 26, 26]
The average of the ages is 22
The range of the ages in the list is 7
The median of the ages is 24
PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main>
```

Question2

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

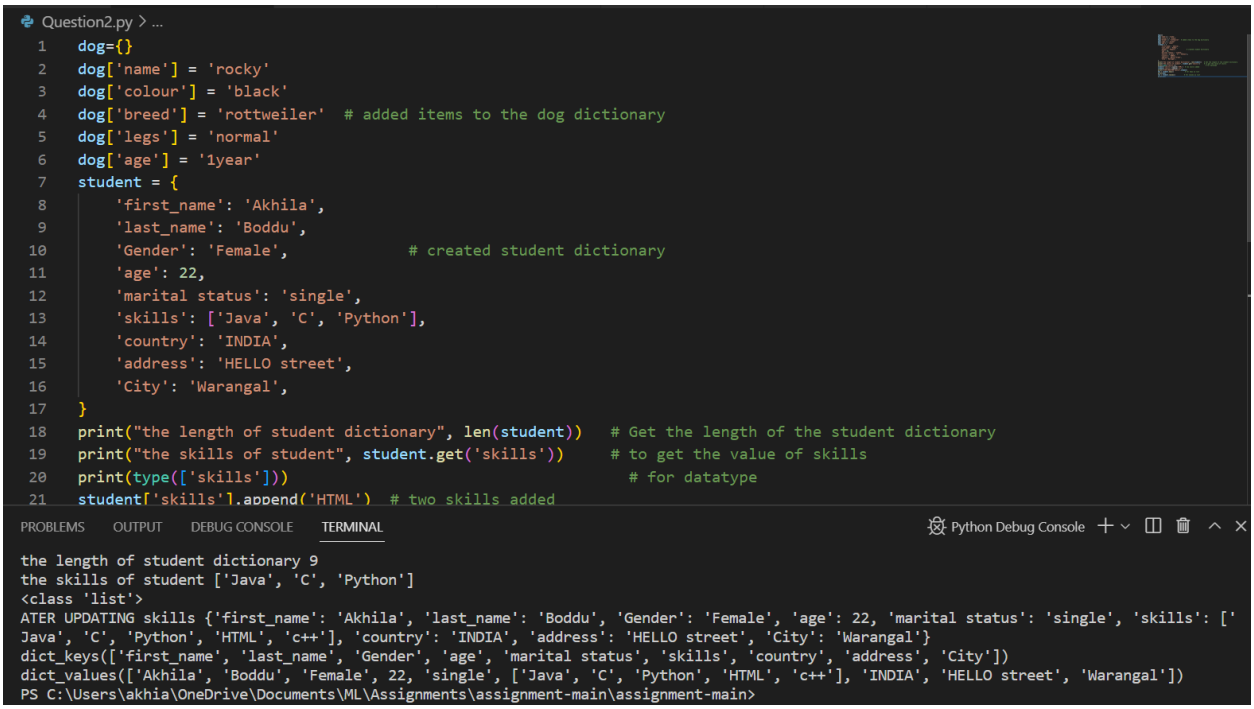
Source Code:

```
# Dog dictionary is created with given key and values
dog={}
dog['name'] = 'rocky'
dog['colour'] = 'black'
dog['breed'] = 'rottweiler'
dog['legs'] = 'normal'
dog['age'] = '1year'
# Added items to the dog dictionary
# Student dictionary is created with given key and values
student = {
   'first_name': 'Akhila',
   'last_name': 'Boddu',
   'Gender': 'Female',
```

```
    'age': 22,
    'marital status': 'single',
# Created student dictionary
# Create another dictionary for skills
    'skills': ['Java', 'C', 'Python'],
    'country': 'INDIA',
    'address': 'HELLO street',
    'City': 'Warangal',
}
# Find the length of student dictionary
print("the length of student dictionary", len(student))   # Get the length of the student dictionary
print("the skills of student", student.get('skills'))     # to get the value of skills
print(type(['skills']))                                   # for datatype
student['skills'].append('HTML')  # two skills added
student['skills'].append('c++')
print("ATER UPDATING skills",student)
key = student.keys()            # for keys as list
print(key)
val = student.values()          # for values as list
print(val)
```

Description:

In the above source code, we have created dictionary called dog added keys and values and printed them on screen, again we created student and skills dictionary and given values and printed them onscreen. Now we calculated length, datatype using len() and type().we added an item to skill then printed the dog keys and student values using print().



```
Question2.py > ...
1   dog={}
2   dog['name'] = 'rocky'
3   dog['colour'] = 'black'
4   dog['breed'] = 'rottweiler'  # added items to the dog dictionary
5   dog['legs'] = 'normal'
6   dog['age'] = '1year'
7   student = {
8       'first_name': 'Akhila',
9       'last_name': 'Boddu',
10      'Gender': 'Female',             # created student dictionary
11      'age': 22,
12      'marital status': 'single',
13      'skills': ['Java', 'C', 'Python'],
14      'country': 'INDIA',
15      'address': 'HELLO street',
16      'City': 'Warangal',
17  }
18  print("the length of student dictionary", len(student))   # Get the length of the student dictionary
19  print("the skills of student", student.get('skills'))     # to get the value of skills
20  print(type(['skills']))                                    # for datatype
21  student['skills'].append('HTML')  # two skills added

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL                    Python Debug Console

the length of student dictionary 9
the skills of student ['Java', 'C', 'Python']
<class 'list'>
ATER UPDATING skills {'first_name': 'Akhila', 'last_name': 'Boddu', 'Gender': 'Female', 'age': 22, 'marital status': 'single', 'skills': ['
Java', 'C', 'Python', 'HTML', 'c++'], 'country': 'INDIA', 'address': 'HELLO street', 'City': 'Warangal'}
dict_keys(['first_name', 'last_name', 'Gender', 'age', 'marital status', 'skills', 'country', 'address', 'City'])
dict_values(['Akhila', 'Boddu', 'Female', 22, 'single', ['Java', 'C', 'Python', 'HTML', 'c++'], 'INDIA', 'HELLO street', 'Warangal'])
PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main>
```

Question3

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

Source Code:

```
brother=("Nithin","Akhil","Samanth")
siblings=("Nikhil","Dayakar")
#Adding brothers to siblings list
siblings=brother+siblings
# Displays siblings' output and length of siblings
print("No of siblings are ",len(siblings))
print(siblings)
```

parents=("Sudhakar","Shobha")

#Adding parents to siblings list

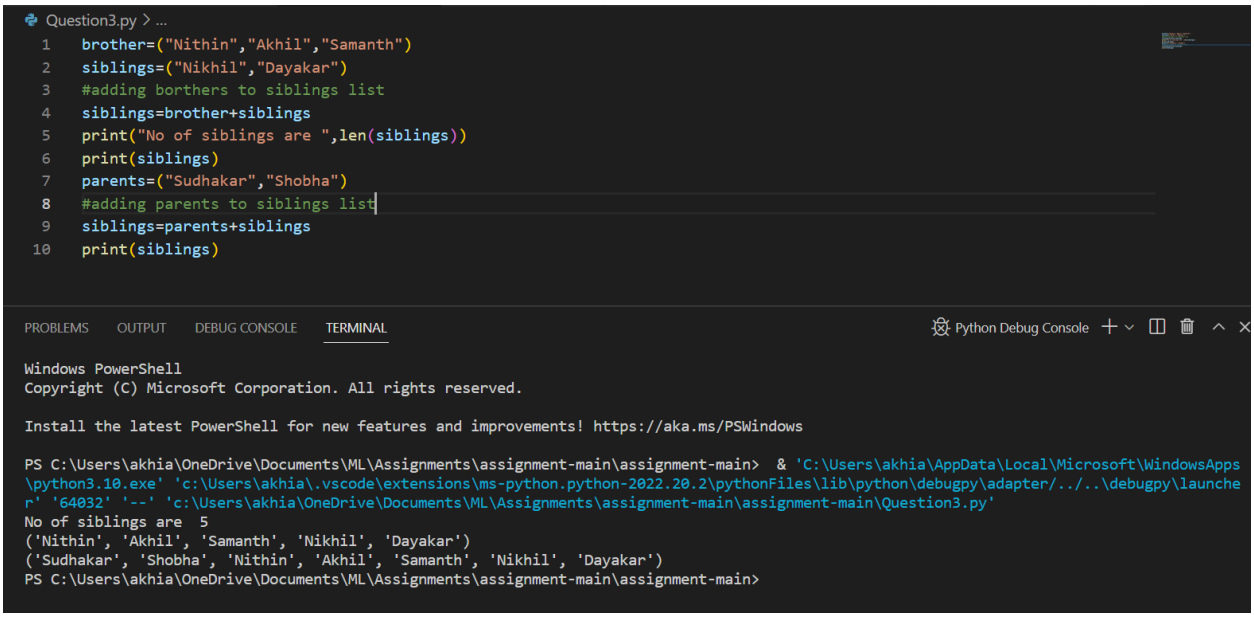
siblings=parents+siblings

 # Displays family_members output

print(siblings)

Description:

In the above source code we have created a tuple sisters and brothers and the created another tuple siblings and added siblings and brother tuple, displayed by using print(). Now we created parents tuple and printed them.



Question4

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.

Source Code:

```
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]

print("THE LENGTH OF IT COMPANIES IS ",len(it_companies))  # Length

it_companies.add('TCS')                  # Adding TCS to set it_companies

it_companies.update(['INFOSYS','COGNIZANT','WIPRO'])  # Adding multiple it companies to set

print("THE IT COMPANIES ARE", it_companies)

it_companies.pop()     # Remove one value from set

print("THE COMPANIES AFTER REMOVING RANDOM COMPANY",it_companies)

# The difference between remove and discard is that remove() method displays an error if the element is missing in a set , where as discard() method doesn't display anything if any element is missing

C=A.union(B)  # Join a and b sets

print("A UNION B",C)

print("A INTERSECTION B",A.intersection(B))  # a intersection b

print("A SUBSET B?",A.issubset(B))  # a subset b?
```

```
print("A IS DISJOINT OF B?",A.isdisjoint(B))  # a disjoint b?

A.update(B)  # joining a with b

B.update(A)  # joining b with a

print("UPDATED A after joining with b",A)

print("UPDATED B after joining with a",B)

print("SYMMETRIC DIFFERENCE BETWEEN A AND B is :",A.symmetric_difference(B))

# symmetric difference

del A  # deleting all sets

del B

del it_companies

a=len(age)  # converting age list into set

b=set(age)

print("THE AGE AFTER CONVERTING INTO SET",b)

c=len(b)

print("the length of age is " ,a,"the length of age after becoming set is", c)
```

Description:

In the above source code initially we assigned some values to the list. We have performed several operation like find length, adding values, remove, discard, using len(), add(), remove(), discard(). Next we performed basic operations like union, intersection, subset, disjoint, length, symmetric difference and deletion of sets.

Question4.py > ...

```python
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]
print("THE LENGTH OF IT COMPANIES IS ",len(it_companies))  # length
it_companies.add('TCS')                    # adding to set it_companies
it_companies.update(['INFOSYS','COGNIZANT','WIPRO'])  # adding multiple it companies to set
print("THE IT COMPANIES ARE", it_companies)
it_companies.pop()      # remove one value from set
print("THE COMPANIES AFTER REMOVING RANDOM COMPANY",it_companies)
# The difference between remove and discard is that remove() method displays an error if the element is missing in a
C=A.union(B)  # join a and b sets
print("A UNION B",C)
print("A INTERSECTION B",A.intersection(B))  # a intersection b
print("A SUBSET B?",A.issubset(B))  # a subset b?
print("A IS DISJOINT OF B?",A.isdisjoint(B))  # a disjoint b?
A.update(B)  # joining a with b
B.update(A)  # joining b with a
print("UPDATED A after joining with b",A)
print("UPDATED B after joining with a",B)
print("SYMMETRIC DIFFERENCE BETWEEN A AND B is :",A.symmetric_difference(B))  # symmetric difference
del A  # deleting all sets
del B
del it_companies
a=len(age)  # converting age list into set
b=set(age)
print("THE AGE AFTER CONVERTING INTO SET",b)
c=len(b)
print("the length of age is " ,a,"the length of age after becoming set is", c)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL Python Debug Console

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main>  & 'C:\Users\akhia\AppData\Local\Microsoft\WindowsApps
\python3.10.exe' 'c:\Users\akhia\.vscode\extensions\ms-python.python-2022.20.2\pythonFiles\lib\python\debugpy\adapter/../..\debugpy\launche
r' '65495' '--' 'c:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main\Question4.py'
THE LENGTH OF IT COMPANIES IS  7
THE IT COMPANIES ARE {'Facebook', 'Microsoft', 'WIPRO', 'INFOSYS', 'Amazon', 'Google', 'IBM', 'COGNIZANT', 'TCS', 'Oracle', 'Apple'}
THE COMPANIES AFTER REMOVING RANDOM COMPANY {'Microsoft', 'WIPRO', 'INFOSYS', 'Amazon', 'Google', 'IBM', 'COGNIZANT', 'TCS', 'Oracle', 'App
le'}
A UNION B {19, 20, 22, 24, 25, 26, 27, 28}
A INTERSECTION B {19, 20, 22, 24, 25, 26}
A SUBSET B? True
A IS DISJOINT OF B? False
UPDATED A after joining with b {19, 20, 22, 24, 25, 26, 27, 28}
UPDATED B after joining with a {19, 20, 22, 24, 25, 26, 27, 28}
SYMMETRIC DIFFERENCE BETWEEN A AND B is : set()
THE AGE AFTER CONVERTING INTO SET {19, 22, 24, 25, 26}
the length of age is  8 the length of age after becoming set is 5
PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main>
```

Question5

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_

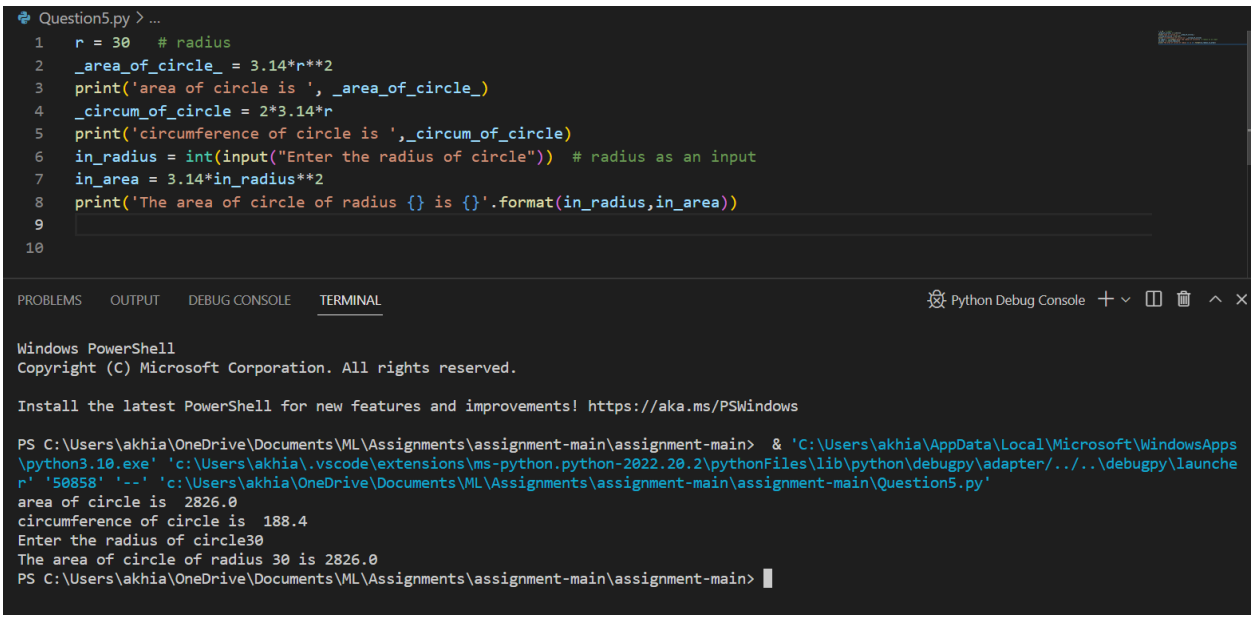
Take radius as user input and calculate the area.

Source Code:

```
# Initialize r where r value can be read from user input
r = 30   # radius
# Calculate area of circle and circumference of circle
_area_of_circle_ = 3.14*r**2
print('area of circle is ', _area_of_circle_)
_circum_of_circle = 2*3.14*r
# Display area of circle and circumference of circle
print('circumference of circle is ',_circum_of_circle)
in_radius = int(input("Enter the radius of circle"))  # radius as an input
in_area = 3.14*in_radius**2
print('The area of circle of radius {} is {}'.format(in_radius,in_area))
```

Description:

In the above code we have created an input variable R that takes input from the user, from the given input we calculated area of the circle, circumference of the circle using formulas (3.14*r*r, 2*3.14*r) and print them on the screen.



Question6

“I am a teacher and I love to inspire and teach people”

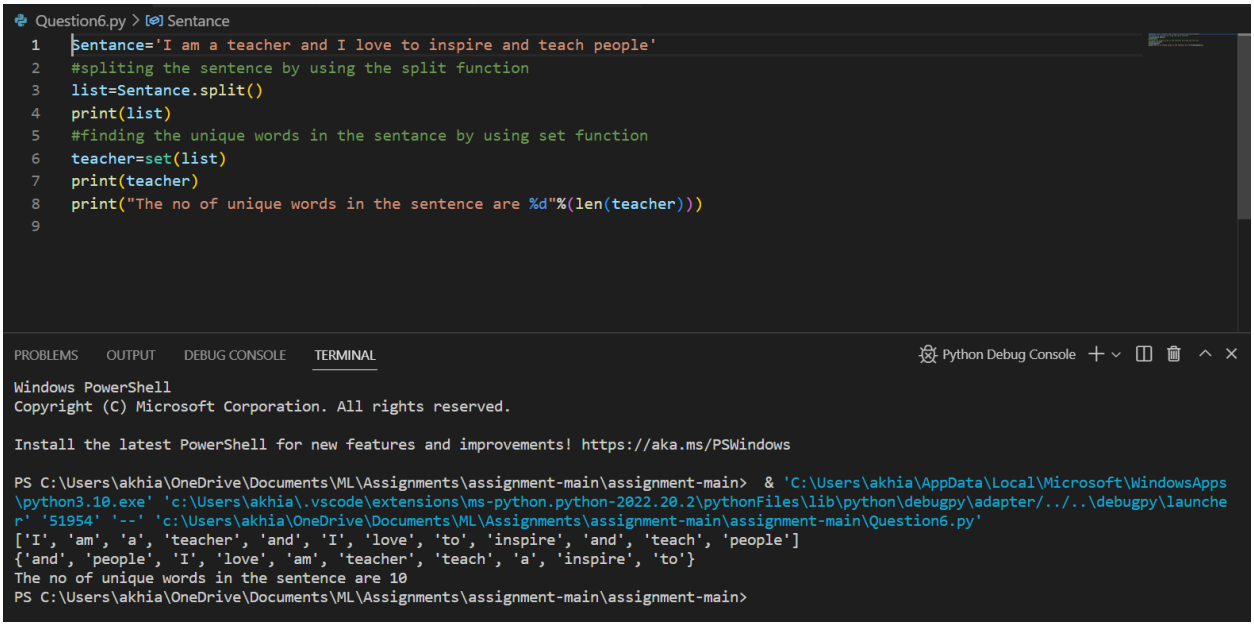
- How many unique words have been used in the sentence? Use the split methods and set to get the unique words.

Source Code:

```
Sentance='I am a teacher and I love to inspire and teach people'

#spliting the sentence by using the split function

list=Sentance.split()

print(list)

#finding the unique words in the sentance by using set function

teacher=set(list)

print(teacher)

print("The no of unique words in the sentence are %d"%(len(teacher)))
```

Description:

In the above code we assigned a string "I am a teacher and I love to inspire and teach people" to sentance by using the split () we found the unique words in the given string and the print the unique values in the display. Now we find the length using len() and print it on the screen.



Question7

Use a tab escape sequence to get the following lines.

Name Age Country City

Asabeneh 250 Finland Helsinki

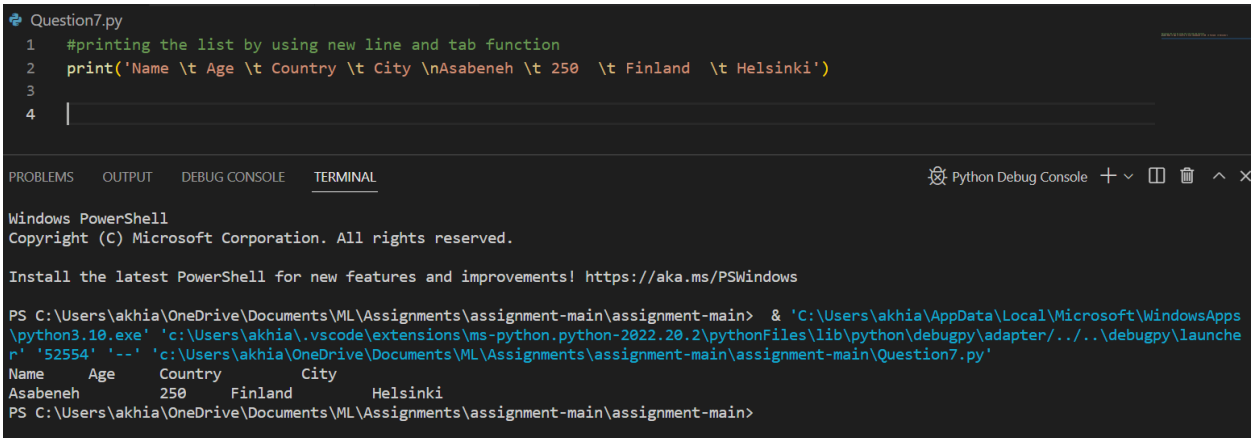
Source Code:

#printing the list by using new line and tab function

print('Name \t Age \t Country \t City \nAsabeneh \t 250 \t Finland \t Helsinki')

Description:

The above code uses tab escape sequence to print the data in particular format. So we used \t which gives a tab space in between the words which we can see below. It also provides a space at the precise location where the escape sequence is added.



Question8

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.”

Source Code:

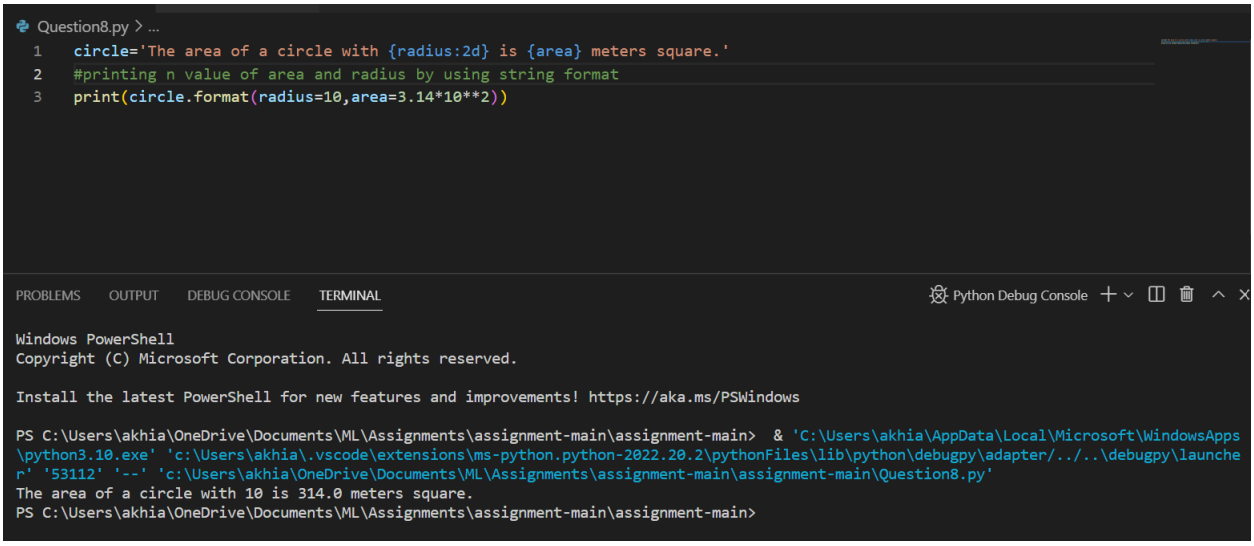
circle='The area of a circle with {radius:2d} is {area} meters square.'

#printing n value of area and radius by using string format

print(circle.format(radius=10,area=3.14*10**2))

Description:

From the above code we used string formatting method to display “The area of a circle with radius 10 is 314 meters square.”. the code first prints radius = 10 then prints area = 3.14*radius**2. Now it prints The area of circle with radius 10 is 314 meters square.



Question9

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]

Source Code:

```
N = int(input("enter number of students"))  # input from user
lb = []
print("enter the weights of {} students in lbs".format(N))
for i in range(N):
 lb.append(int(input()))
 kg = []
#Using for loop to iterate the values and appending the list
for i in range(N):
```

value = "{:.2f}".format(lb[i]*0.45359237) **# converting lbs to kgs**

kg.append(float(value))

print(kg)

Description:

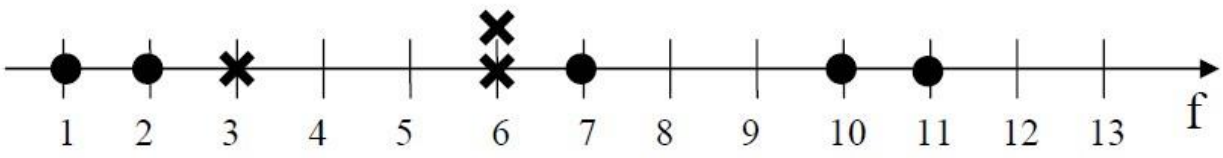
From the above code firstly we created a list N that takes input from the user and kg. initializing for loop to convert the weights we used a formula and append() then printed the converted weights.

```
assignment-main > Question9.py > ...
1   N = int(input("enter number of students"))  # input from user
2   lb = []
3   print("enter the weights of {} students in lbs".format(N))
4   for i in range(N):
5    lb.append(int(input()))
6    kg = []
7   for i in range(N):
8    value = "{:.2f}".format(lb[i]*0.45359237)  # converting lbs to kgs
9    kg.append(float(value))
10  print(kg)
```

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL        Python Debug Console

Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main>  & 'C:\Users\akhia\AppData\Local\Microsoft\WindowsApps\python3.10.exe'
 'c:\Users\akhia\.vscode\extensions\ms-python.python-2022.20.2\pythonFiles\lib\python\debugpy\adapter/../..\debugpy\launcher' '55389' '--'
'c:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main\assignment-main\Question9.py'
enter number of students2
enter the weights of 2 students in lbs
23
54
[10.43, 24.49]
PS C:\Users\akhia\OneDrive\Documents\ML\Assignments\assignment-main>
```

Question10

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. Provide stepwise mathematical solution, do not write code for it.



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.

10)

1) Dividing the whole data into 2 parts.

1, 2, 3, 10 as training

6, 6, 7, 11 as testing

* for KNN classfier :- for $K = 3$

We Know that $d = \sqrt{(w-w_1)^2 + (h-h_1)^2}$

→ The distance from 6 to 1, 2, 3, 10 is

$d_1 = \sqrt{(6-1)^2} = 5$

$d_2 = \sqrt{(6-2)^2} = 4$

$d_3 = \sqrt{(6-3)^2} = 3$ } X

$d_4 = \sqrt{(6-10)^2} = 4$

As there are maximum number of (.)s, 6 is changed from X to .

→ Distance from 7 to 1,2,3,10 is

$d_1 = \sqrt{(7-1)^2} = 6$

$d_2 = \sqrt{(7-2)^2} = 5$

$d_3 = \sqrt{(7-3)^2} = 4$

$d_4 = \sqrt{(7-10)^2} = 3$

} X

As there are maximum no.of(·), 7 is not changed from (·) to (·).

→ Distance from 11 to 1,2,3,10 is

$d_1 = \sqrt{(11-1)^2} = 10$

$d_2 = \sqrt{(11-2)^2} = 9$

$d_3 = \sqrt{(11-3)^2} = 8$

$d_4 = \sqrt{(11-10)^2} = 1$

} X

As there are maximum no.of(·)'s, there is no change, (·) is same as (·)