**MIDHUN KUMAR CHINTAPALLI 700739502**

CRN-23922

**Assignment 1**

**GitHub Link:** <https://github.com/midhun-ch/ML_Assignment-1>

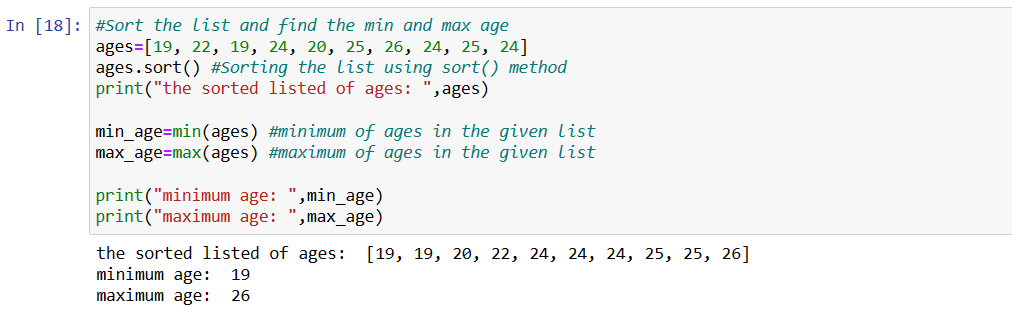
**Video Link:** [**https://drive.google.com/file/d/1wwlMA2cKvM7J9ICNPxEH301whPpkxD7K/view?usp=share\_link**](https://drive.google.com/file/d/1wwlMA2cKvM7J9ICNPxEH301whPpkxD7K/view?usp=share_link)

**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:**

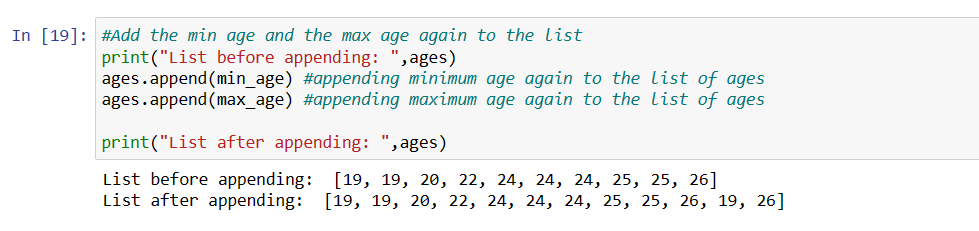
Output:



Description:

Sort() function is used to sort the elements in list in ascending order. So the first element in the list will be the minimum value and the last element in the list will be the maximum value.

* **Add the min age and the max age again to the list**

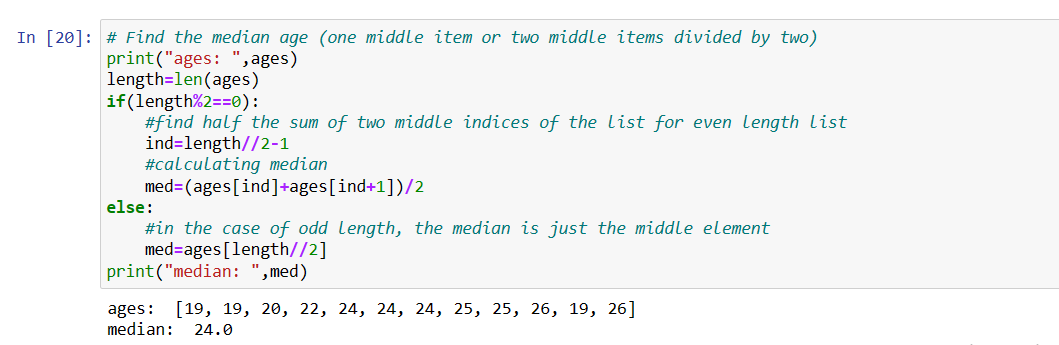
Output: 

Description:

Append() is used to insert the new values to the list. So in order to add minimum and maximum values to the list I have used append() function.

* **Find the median age (one middle item or two middle items divided by two):**

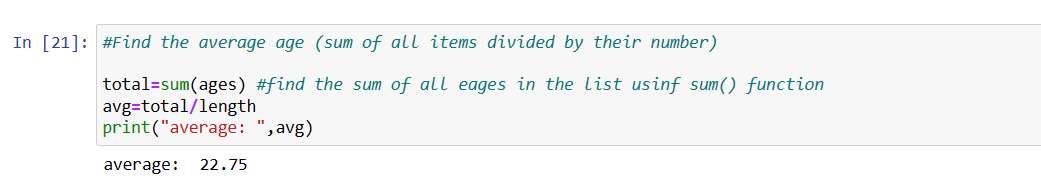
Output:



Description:

Median for odd length list will be the middle element and for even it will be the average of two middle elements. In order to do that I have checked the condition whether the number of elements in the list are even or odd by checking whether the number of elements count is divisible by 2.

* **Find the average age (sum of all items divided by their number):**

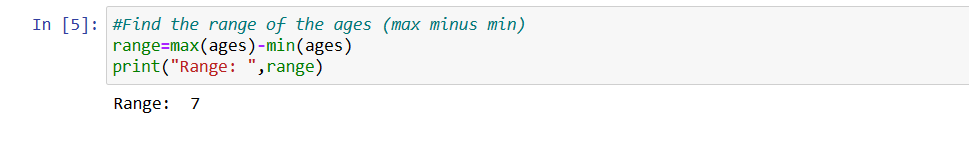
Output: 

Description:

Used **for** loop to sum up all the elements in the list and len() function to get the number of elements in the list (which is the length of the list).

* **Find the range of the ages (max minus min):**

Output:



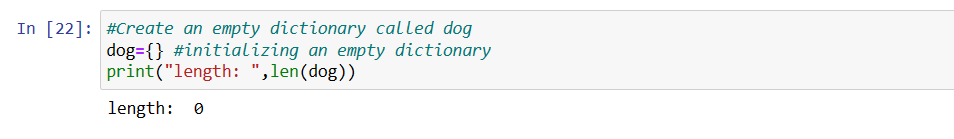
Description:

When we used sort() function the elements in the list get sorted in ascending order ,so the range can be calculated by getting the difference between the last element(which is the maximum value in the list) and first element(which is the minimum value in the list)

**Question 2:**

* **Create an empty dictionary called dog:**

Code:

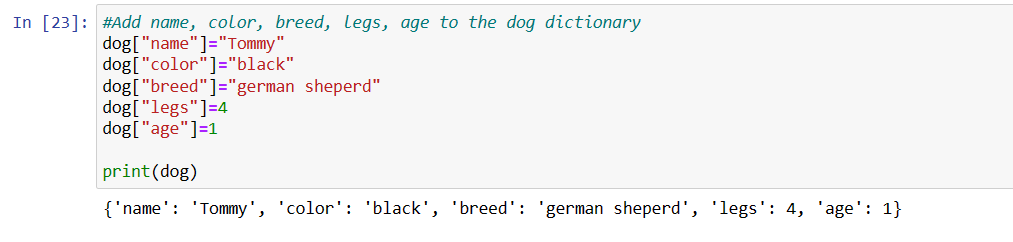


Description:

{} represents empty dictionary

* **Add name, color, breed, legs, age to the dog dictionary:**

Code:



Description:

Data is added to dictionary as key-value pairs. Dictionary[keys]=value is the format for data insertion.

* **Create a student dictionary and add first\_name, last\_name, gender, age, marital status, skills, country, city and address as keys for the dictionary:**

Code:

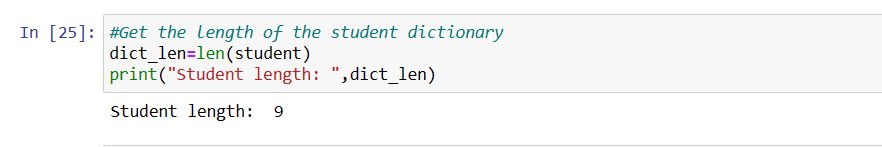


Description:

Created Student dictionary with key-value pairs.

* **Get the length of the student dictionary:**

Output:

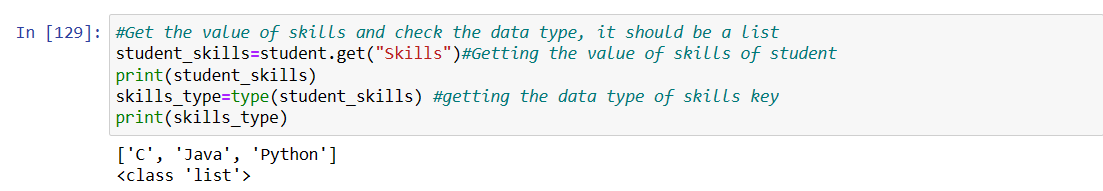


Description:

len() function is used to retrieve the number of key-value pairs in the dictionary.

* **Get the value of skills and check the data type, it should be a list:**

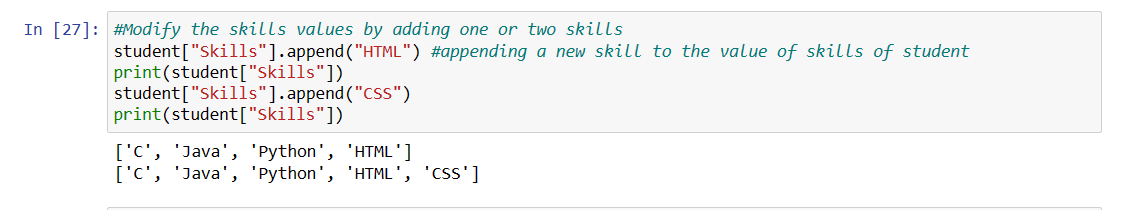
Output:



Description:

* By using get() function we can retrieve the value for the key skills and to define the datatype I have used the type() function.
* **Modify the skills values by adding one or two skills:**

Output:

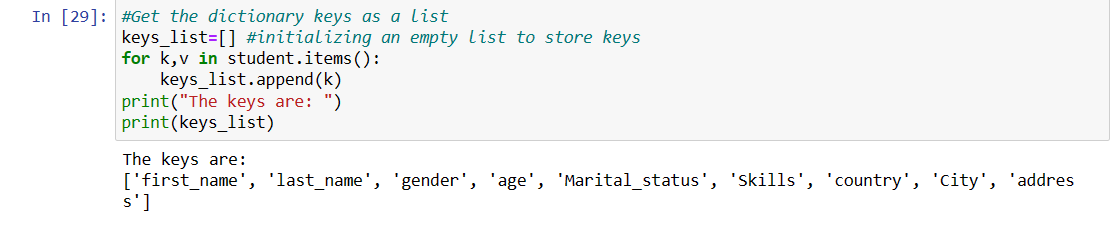


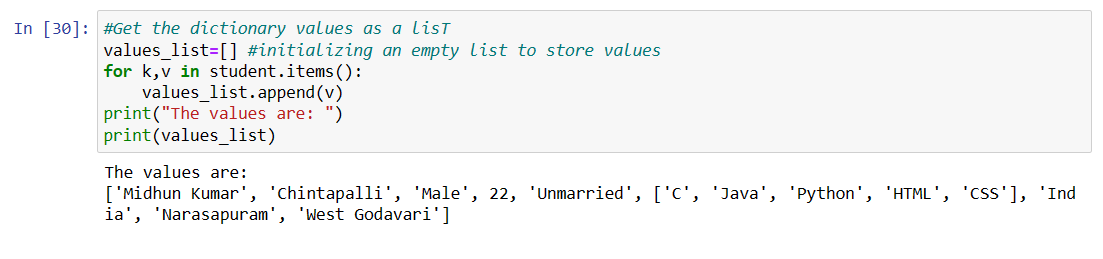
Description:

We can modify the values in dictionary by using the append() function to add values.

* **Get the dictionary keys as a list :**
* **Get the dictionary values as a list:**

Output:





Description:

In dictionary keys are unique identifiers and the values are assigned to them. We can display both keys and values separately with the function keys() and values()

**Question 3:**

* **Create a tuple containing names of your sisters and your brothers (imaginary siblings are fine):**

Code:

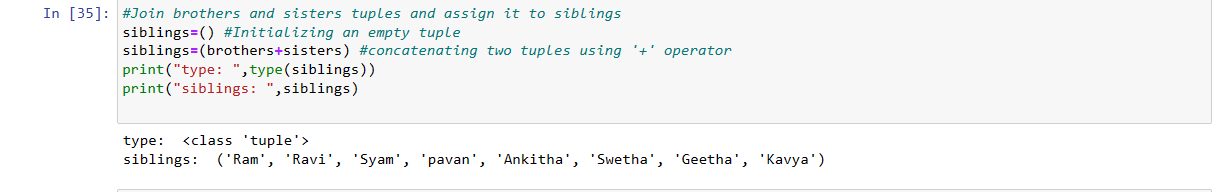


Description: Created two tuples sisters and brothers

Here in this code we created two tuples and added the values to the tuple. Then we added elements into the other tuple and re-assigned it to the same tuple. As tuples are immutable we added values into the tuples using the compound assignment operator.

* **Join brothers and sisters tuples and assign it to siblings:**

Output:

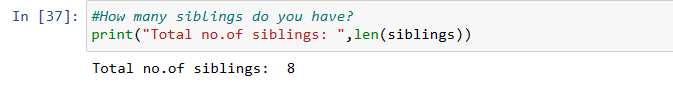


Description:

‘+’ operator is used to join tuples.

* **How many siblings do you have?**

Output:

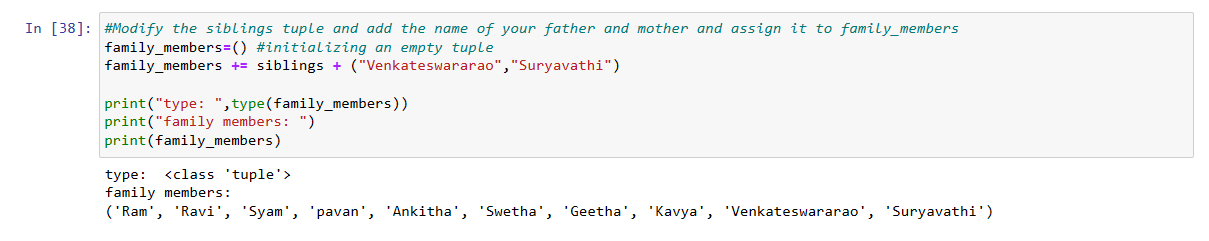


Description:

‘Len()’ function is used to find the length of the elements in the tuple.

* **Modify the siblings tuple and add the name of your father and mother and assign it to family\_members:**

Output:



Description:

In this code we created two more tuples and added them to the family\_members tuple finally and printed it.

**Question 4:**

it\_companies = {'Facebook', 'Google', 'Microsoft', 'Apple', 'IBM', 'Oracle', 'Amazon'}

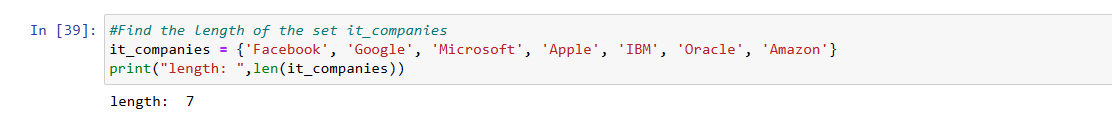
A = {19, 22, 24, 20, 25, 26}

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:**

Output:



Description:

Len() function is used to calculate the number of elements in the set

* **Add 'Twitter' to it\_companies:**

Output: 

Description:add() function is used to insert new element to the set

* **Insert multiple IT companies at once to the set it\_companies:**

Output:



Description:

Update() function is used to insert multiple elements in the set

* **Remove one of the companies from the set it\_companies:**

Output:



Description:

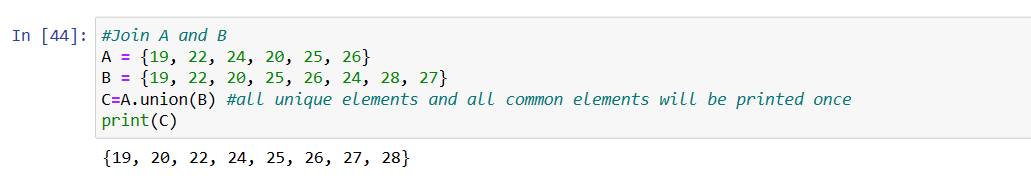
Remove() function is used to delete or remove the element from the list

* **What is the difference between remove and discard:**

Remove() and discard() both the functions are used to remove elements in the set. But whenever the element is not present in the set, discard() does not throw any exception but remove throws an exception.

* **Join A and B**

Output:

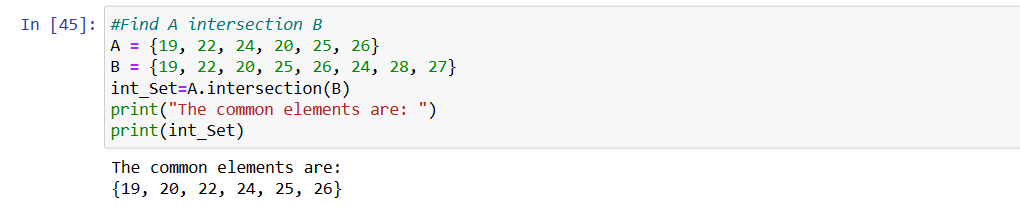


Description:

Union function is used to join two sets

* **Find A intersection B :**

Output:

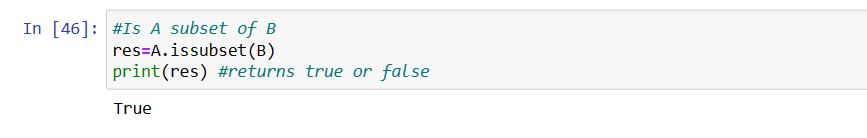


Description:

Intersection() function returns the common elements from both the sets

* **Is A subset of B:**

Output:



Description:

If all the elements in set A are present in set B then the function issubset return true else it returns false.

* **Are A and B disjoint sets**

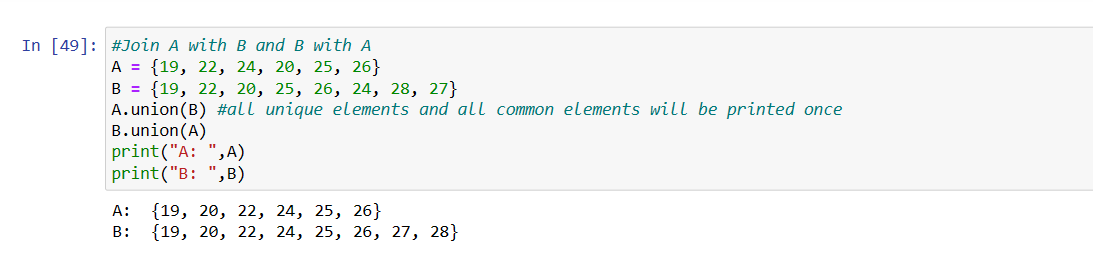
Output:



Description: If the two sets do not have any common elements in them,then the isdisjoint() function returns true else false so, the output will be “A and B are disjoint” else “No “

* **Join A with B and B with A:**

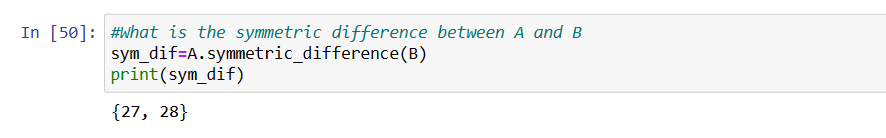
Output:



Description: Two sets can be joined using |= or union or update which returns the joined set of both AUB.

* **What is the symmetric difference between A and B :**

Output:

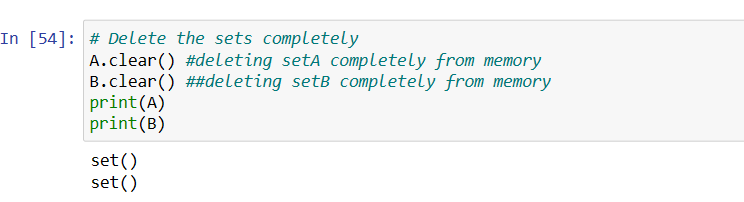


Description:

The set returned by symmetric\_difference(),contains all the elements of both sets except for those that are shared by both sets.

* **Delete the sets completely**

Output:

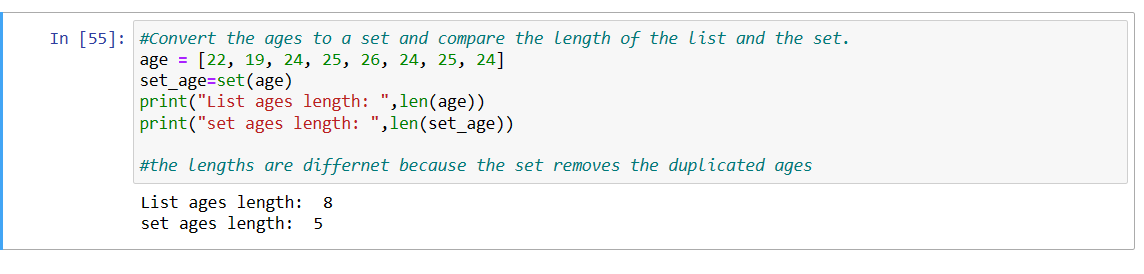


Description:

Clear command is used to delete/clear the declared set completely.

* **Convert the ages to a set and compare the length of the list and the set:**

Output:



Description:

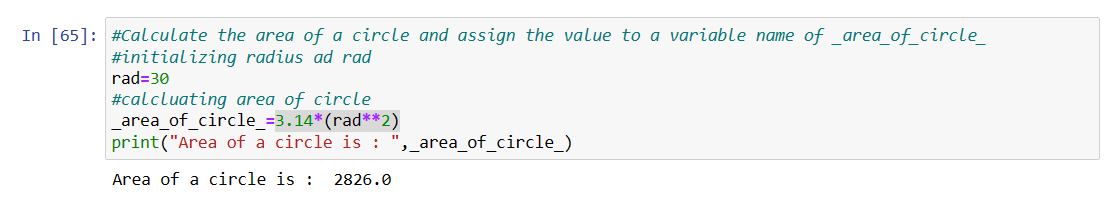
The length of list will be more because it has duplicate values whereas the set doesn’t allow the duplicate value,so the length will be less compared to the list.

**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\_**

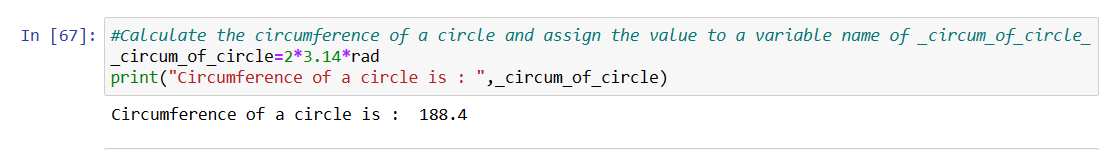
Output:



Description:

Area of the circle is calculated based on the given radius and assigned to \_area\_of\_circle\_

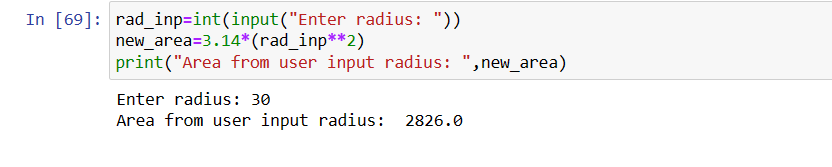
* **Calculate the circumference of a circle and assign the value to a variable name of \_circum\_of\_circle\_**

Output: 

Description:

Circumference of the circle is calculated based on the declared radius and assigned to \_circum\_of\_circle\_

* **Take radius as user input and calculate the area:**

Output: 

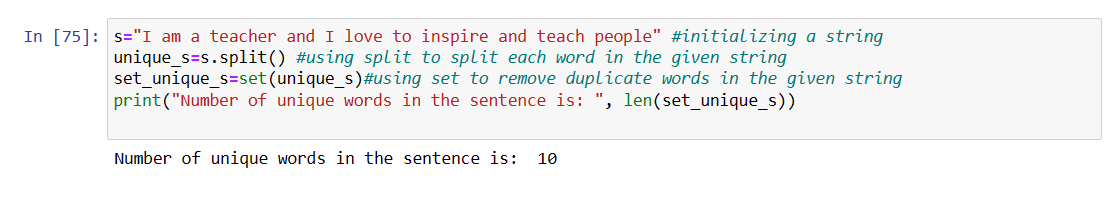
Description:User input can be read using input() function and int() function is used to convert the user input to interger data type. Based on the input , I have calculated the area and displayed the value.

**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:**

Output:



Description:

The split() function is used to split the text into a list. To get the unique words I have converted the list to set (as set does not allow duplicate elements) and calculated the length using len() function.

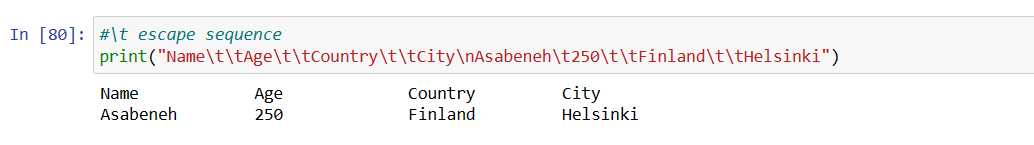
**Question 7:**

**Use a tab escape sequence to get the following lines**.

**Name Age Country City**

**Asabeneh 250 Finland Helsinki**

Output:



Description:\t and \n are used for tab space and next line in the output window.”/” is a special character and known as escape character.

**Question 8:**

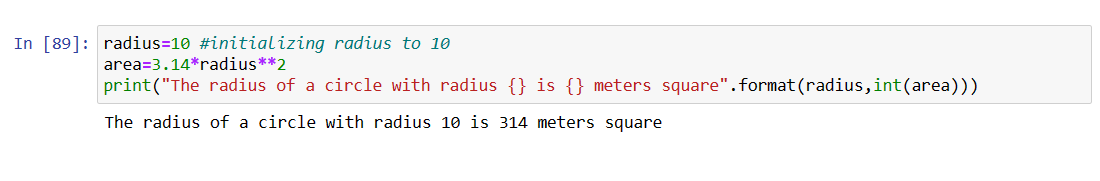
**Use the string formatting method to display the following:**

**radius = 10**

**area = 3.14 \* radius \*\* 2**

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

Output:



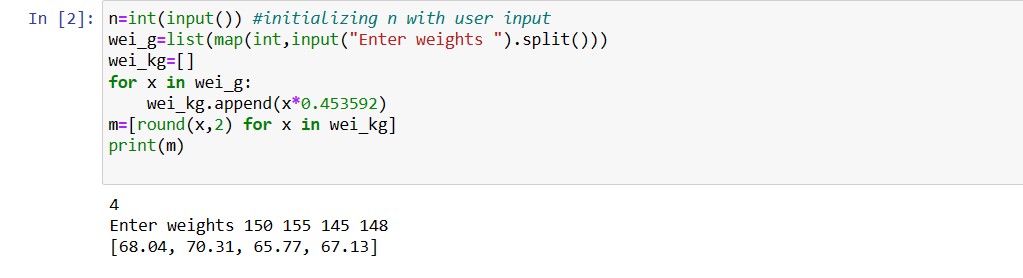
Description:

The process of dynamically incorporating elements into a string and presenting it is known as string formatting.

**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]**

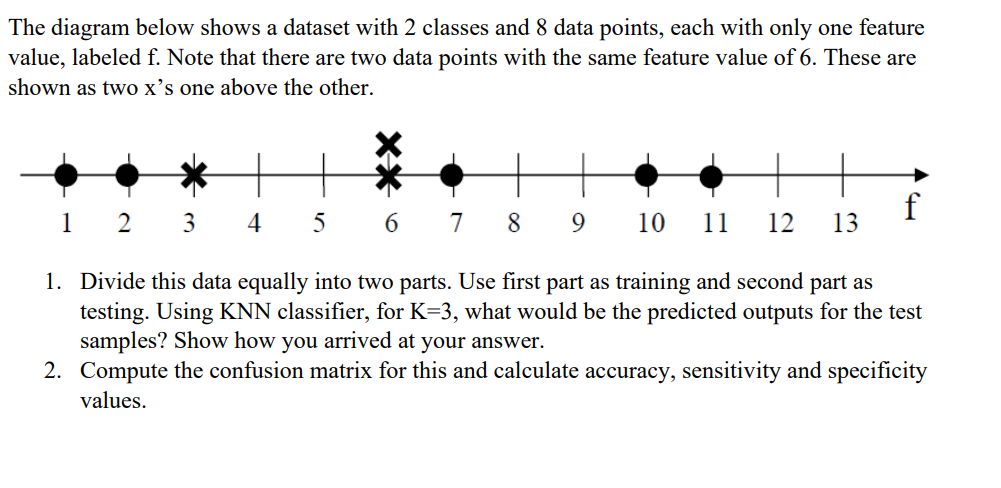
Output:



Description:

Here we converted a list of values in lbs to kilograms by taking the list elements from user input and we used round() function to get each element in the list rounded upto two decimal points.

**Question 10:**

****

Output:

Given data elements are taken in the tabular form as below,

|  |  |
| --- | --- |
| **Feature** | **Label** |
| 1 | O |
| 2 | O |
| 3 | X |
| 6 | X |
| 6 | X |
| 7 | O |
| 10 | O |
| 11 | O |

Here, the first four rows of data are considered to be the Training dataset and the next four rows are selected as the Testing dataset.

Now, according to the KNN Classifier we shall now consider K=3 and then the distance between the testing and training data is demonstrated below.

In the below table the columns are the training dataset and rows are the testing dataset.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 1(O) | 2(O) | 3(X) | 6(X) |
| 6 | 5 | **4** | **3** | **0** |
| 7 | 6 | **5** | **4** | **1** |
| 10 | 9 | **8** | **7** | **4** |
| 11 | 10 | **9** | **8** | **5** |

The highlighted rows are the distance values.

Let us now assume ‘O’ as negative and ‘X’ as positive values. Now the prediction on testing data is as below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | True label | Predicted label | O/P |
| 6 | X | X | Tp |
| 7 | O | X | Fp |
| 10 | O | X | Fp |
| 11 | O | X | Fp |

Confusion matrix for the above prediction is:

|  |  |
| --- | --- |
| TN | FP |
| FN | TP |

The final confusion matrix is :

**0 3**

**0 1**

Accuracy of the classifier = (TP + TN) / (P + N) = 1 / 4 = 0.25 Sensitivity of the classifier = TP / P = 1/1 = 1

Specificity of the classifier = TN / N = 0/3 = 0

Description:

The k-nearest neighbours algorithm, sometimes referred to as KNN or k-NN, is a supervised learning classifier that employs proximity to produce classifications or predictions about the grouping of a single data point.

A confusion matrix, sometimes referred to as an error matrix, is a particular table arrangement that enables visualization of an algorithm's performance.

By using KNN Classifier value as 3 predicted the values and based on them calculated the confusion matrix, based on the confusion matrix calculated the accuracy, sensitivity and specificity.