Spyder (Python 3.8)

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\h\Desktop\Statistical Programming 51100\Week2

C:\Users\h\Desktop\Statistical Programming 51100\Week2\kMeans.py

kMeans.py

```
print('DATA-51100, Fall 2022')
print('Name: Betul Mescioglu, Marianne Guieb, Satwik Reddy Yanala')
print('Programming Assignment #2\n')
#Get path of working directory
path = %pwd

#Path of the input file
input_file_path = path + '\\' + 'prog2-input-data.txt'

#Read input file's content and save numbers to file_content as floats
file_content = [float(i.rstrip()) for i in open(input_file_path)]
# Find each number's closest centroid:
def find_distance(centroids, numbers):
    '''
    DESCRIPTION:
        For a given set of numbers and centroids, this function
        assigns each number to the closest centroid.
    ----------
    Parameters
    ----------
    centroids : list
        List of centroids.
    numbers : list
        List of numbers.

    Returns
    -------
    centroids_numbers : dictionary
        A dictionary of clusters and numbers closest to those clusters.

    '''
    #Create dictionary that will hold cluster numbers and numbers
    centroids_numbers = {}
    #Initialize the dictionary with cluster numbers (k) as keys
    for centroid in range(len(centroids)):
        centroids_numbers[centroid] = []
    #Go through numbers list, for each number, calculate it's distance to
    #each centroid, find the centroid it is closest to, update centroids_numbers
    #dictionary with this centroid and number.
    for number in numbers:
        dist_to_centroid = []

        for centroid in centroids:
            dist_to_centroid.append(abs(number-centroid))
```

Source Console Object

Help Variable explorer Plots Files

Console 1/A

```
Desktop/Statistical Programming 51100/Week2')
DATA-51100, Fall 2022
Name: Betul Mescioglu, Marianne Guieb, Satwik Reddy Yanala
Programming Assignment #2


Enter the number of clusters: 5

Iteration 1
0 [1.8]
1 [4.5, 6.5]
2 [1.1, 0.5]
3 [2.1, 3.2]
4 [9.8, 7.6, 11.32]

Iteration 2
0 [1.8, 2.1]
1 [4.5, 6.5]
2 [1.1, 0.5]
3 [3.2]
4 [9.8, 7.6, 11.32]

Iteration 3
0 [1.8, 2.1]
1 [4.5, 6.5]
2 [1.1, 0.5]
3 [3.2]
4 [9.8, 7.6, 11.32]

Point 1.8 in cluster 0
Point 4.5 in cluster 1
Point 1.1 in cluster 2
Point 2.1 in cluster 0
Point 9.8 in cluster 4
Point 7.6 in cluster 4
Point 11.32 in cluster 4
Point 3.2 in cluster 3
Point 0.5 in cluster 2
Point 6.5 in cluster 1
```

IPython console History

LSP Python: ready | conda: base (Python 3.8.3) | new-branch [114] | Line 144, Col 22 | ASCII | CRLF | RW | Mem 75%

10:16 PM 9/9/2022