## گزارش تمرین دوم داده کاوی

## :4-1

```
In [3]: iris_SP = iris[['sepal.length','sepal.width','petal.length','petal.width']]
            iris_SP.head()
Out[3]:
                 sepal.length sepal.width petal.length petal.width
                          4.9
                          47
                                       3.2
                                                     1.3
             2
                                                                     0.2
             3
                          4.6
                                        3.1
                                                       1.5
                                                                     0.2
                          5.0
                                        3.6
                                                      1.4
                                                                    0.2
In [8]: from scipy.cluster.hierarchy import dendrogram, linkage
            # generate the linkage matrix
Z = linkage(iris_SP, 'complete')
             [[1.01000000e+02 1.42000000e+02 0.00000000e+00 2.00000000e+00]
[7.00000000e+00 3.90000000e+01 1.00000000e-01 2.00000000e+00]
               [0.00000000e+00 1.70000000e+01 1.00000000e-01 2.00000000e+00]
              [9.00000000e+00 3.4000000e+01 1.0000000e-01 2.0000000e+00]
[1.28000000e+02 1.3200000e+02 1.0000000e-01 2.0000000e+00]
               [1.00000000e+01 4.80000000e+01 1.00000000e-01 2.00000000e+00]
              [4.00000000e+00 3.7000000e+01 1.41421356e-01 2.00000000e+00]
[1.90000000e+01 2.10000000e+01 1.41421356e-01 2.00000000e+00]
               [2.90000000e+01 3.00000000e+01 1.41421356e-01 2.00000000e+00]
              [5.70000000e+01 9.3000000e+01 1.41421356e-01 2.00000000e+00]
[8.00000000e+01 8.10000000e+01 1.41421356e-01 2.00000000e+00]
               [1.16000000e+02 1.37000000e+02 1.41421356e-01 2.00000000e+00]
              [8.00000000e+00 3.8000000e+01 1.41421356e-01 2.0000000e+00]
[3.0000000e+00 4.7000000e+01 1.41421356e-01 2.0000000e+00]
               [2.70000000e+01 2.80000000e+01 1.41421356e-01 2.00000000e+00]
              [8.20000000e+01 9.20000000e+01 1.41421356e-01 2.00000000e+00]
[9.50000000e+01 9.60000000e+01 1.41421356e-01 2.00000000e+00]
              [1.27000000e+02 1.38000000e+02 1.41421356e-01 2.00000000e+00]

[1.00000000e+00 4.50000000e+01 1.41421356e-01 2.00000000e+00]

[6.30000000e+01 9.1000000e+01 1.41421356e-01 2.00000000e+00]
```

## :4-7

```
In [5]: max_d = 7.08  # max_d as in max_distance

plt.figure(figsize=(25, 10))
 plt.title('Tris Hierarchical Clustering Dendrogram')
 plt.xlabel('Giseines')
 plt.ylabel('Giseines')
 dendrogram(

2,
 truncate_mode='lastp', # show only the last p merged clusters
 p=150, # Try changing values of p
 leaf_routation=90., # Try changing values of p
 leaf_font_size=8., # font size for the x axis labels
 )
 plt.axhine(y-max_d, c-'k')
 plt.show()

Inattenthold Clustering Dendrogram

Inattenthold Cluster
```

## :4-4:

```
In [6]: plt.figure(figsize=(10, 8))
plt.scatter(iris.data[:,0], iris.data[:,1], c=clusters, cmap='prism')
plt.show()
```

