

The background is a dark blue gradient. On the left, there are two overlapping geometric shapes: a blue parallelogram and a light green parallelogram. Below these, there is a circular inset showing a detailed view of a circuit board. In the top right corner, there is a faint, stylized pattern of interconnected lines resembling a circuit or a neural network.

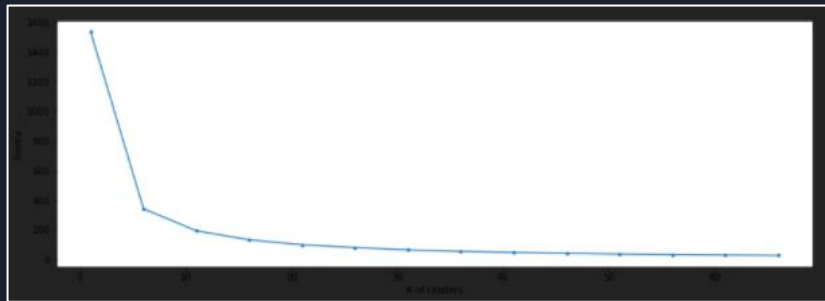
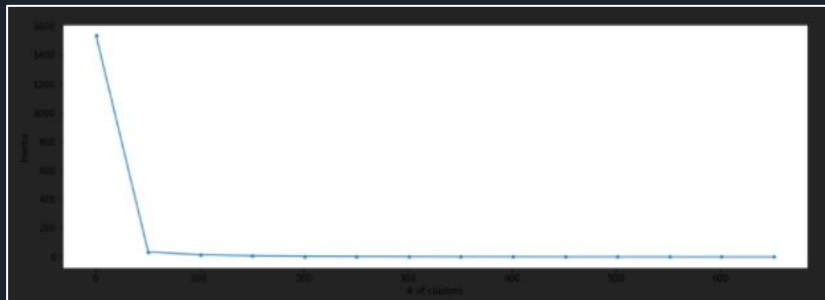
Artificial Intelligence

Laboratory 6

Isabella Quicaño
Raúl López

6.1 K-means algorithm : BMI and Glucose features

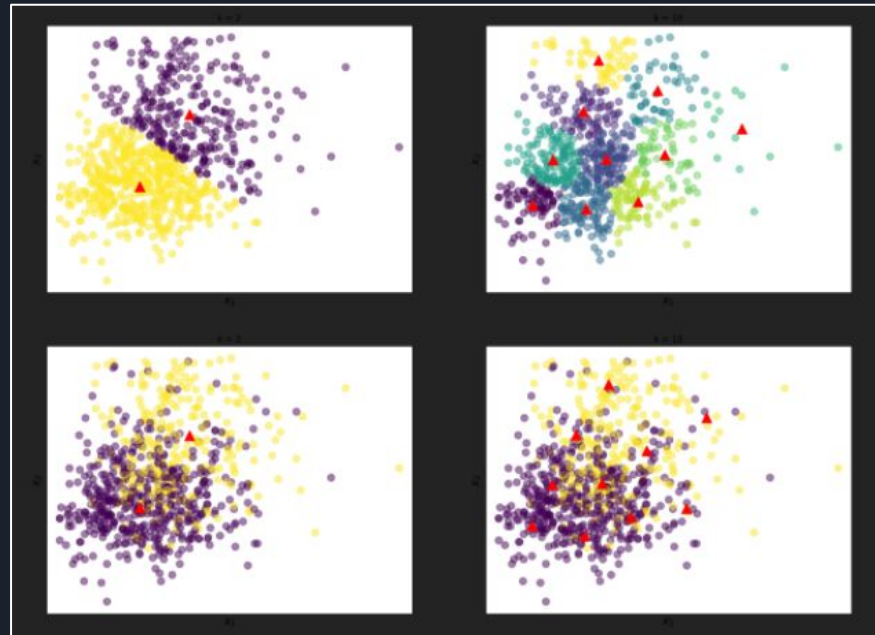
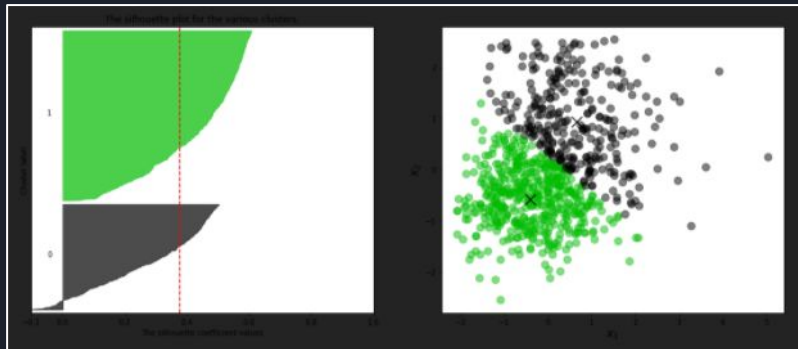
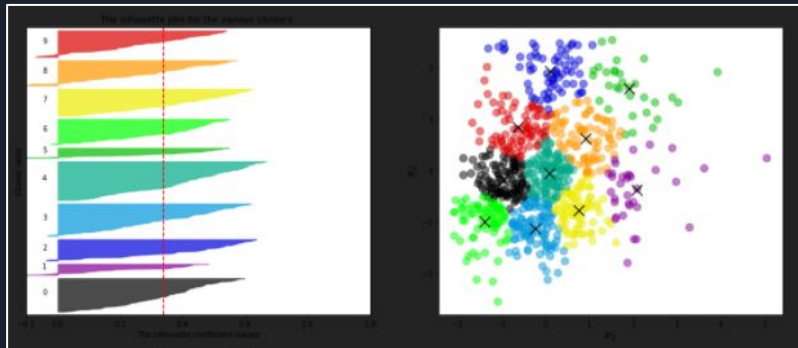
Inertia



Average silhouette score

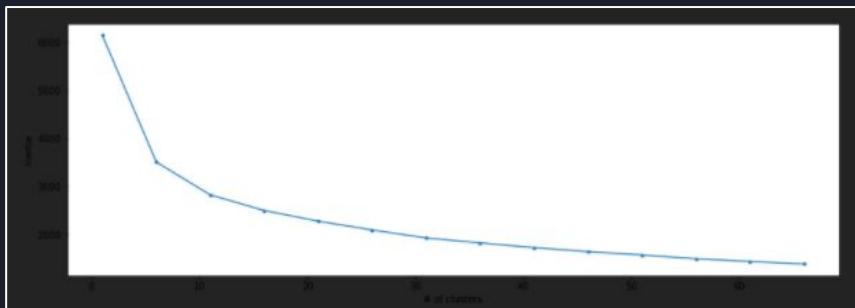
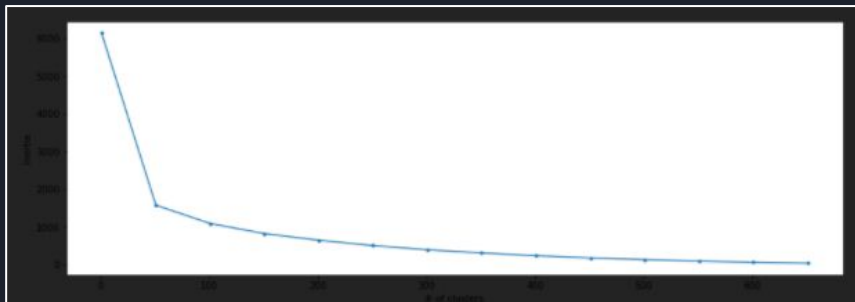
```
For n_clusters = 5 The average silhouette_score is : 0.3422693141177146  
For n_clusters = 10 The average silhouette_score is : 0.3464115882719585  
For n_clusters = 15 The average silhouette_score is : 0.34339536159762307  
For n_clusters = 100 The average silhouette_score is : 0.34838770977850525
```

6.1 K-means algorithm : BMI and Glucose features



6.2 K-means algorithm : All features

Inertia



Average silhouette score

```
For n_clusters = 5 The average silhouette_score is : 0.17929258229150447  
For n_clusters = 10 The average silhouette_score is : 0.14313193548130443  
For n_clusters = 15 The average silhouette_score is : 0.12229619280489552
```

6.2 K-means algorithm : All features

Column of interest : pregnancies

```
Cluster : 0 with mean = 2.7222222222222223
Cluster : 1 with mean = 1.6967213114754098
Cluster : 2 with mean = 3.65625
Cluster : 3 with mean = 7.7
Cluster : 4 with mean = 1.7555555555555555
Cluster : 5 with mean = 3.52
Cluster : 6 with mean = 2.0387096774193547
Cluster : 7 with mean = 9.120689655172415
Cluster : 8 with mean = 2.7450980392156863
Cluster : 9 with mean = 5.276923076923077
```

Column of interest : glucose

```
Cluster : 0 with mean = 109.01851851851852
Cluster : 1 with mean = 102.89907182605333
Cluster : 2 with mean = 152.90625
Cluster : 3 with mean = 104.51248659597283
Cluster : 4 with mean = 127.97777777777777
Cluster : 5 with mean = 163.6
Cluster : 6 with mean = 100.74434532617427
Cluster : 7 with mean = 155.41379310344828
Cluster : 8 with mean = 150.91176470588235
Cluster : 9 with mean = 134.6769230769231
```

Column of interest : blood_pressure

```
Cluster : 0 with mean = 66.3263099388611
Cluster : 1 with mean = 67.49533692662088
Cluster : 2 with mean = 79.28125
Cluster : 3 with mean = 75.10587870519659
Cluster : 4 with mean = 84.11111111111111
Cluster : 5 with mean = 69.52
Cluster : 6 with mean = 64.20400475289354
Cluster : 7 with mean = 84.42776497153879
Cluster : 8 with mean = 73.85928095658456
Cluster : 9 with mean = 78.267772064225
```

Column of interest : skin_thickness

```
Cluster : 0 with mean = 23.82470733210105
Cluster : 1 with mean = 32.404063513226845
Cluster : 2 with mean = 34.50231053604436
Cluster : 3 with mean = 31.16973617879345
Cluster : 4 with mean = 44.22904087081536
Cluster : 5 with mean = 32.64
Cluster : 6 with mean = 21.01481128137856
Cluster : 7 with mean = 32.788992287590034
Cluster : 8 with mean = 27.89801022072414
Cluster : 9 with mean = 27.789080051187273
```

Column of interest : insulin

```
Cluster : 0 with mean = 143.24008272231617
Cluster : 1 with mean = 127.51235749355072
Cluster : 2 with mean = 178.35072969543145
Cluster : 3 with mean = 146.37055837563435
Cluster : 4 with mean = 159.45730400451208
Cluster : 5 with mean = 496.48
Cluster : 6 with mean = 112.51606353364984
Cluster : 7 with mean = 173.98267110099763
Cluster : 8 with mean = 168.2892903354234
Cluster : 9 with mean = 155.0385786802829
```

Column of interest : bmi

```
Cluster : 0 with mean = 29.21217525319242
Cluster : 1 with mean = 36.21628949477012
Cluster : 2 with mean = 38.346875000000004
Cluster : 3 with mean = 32.4624762819743
Cluster : 4 with mean = 43.162222222222222
Cluster : 5 with mean = 34.836000000000006
Cluster : 6 with mean = 25.679354838709678
Cluster : 7 with mean = 38.015517241379314
Cluster : 8 with mean = 32.450980392156886
Cluster : 9 with mean = 28.078691189919706
```

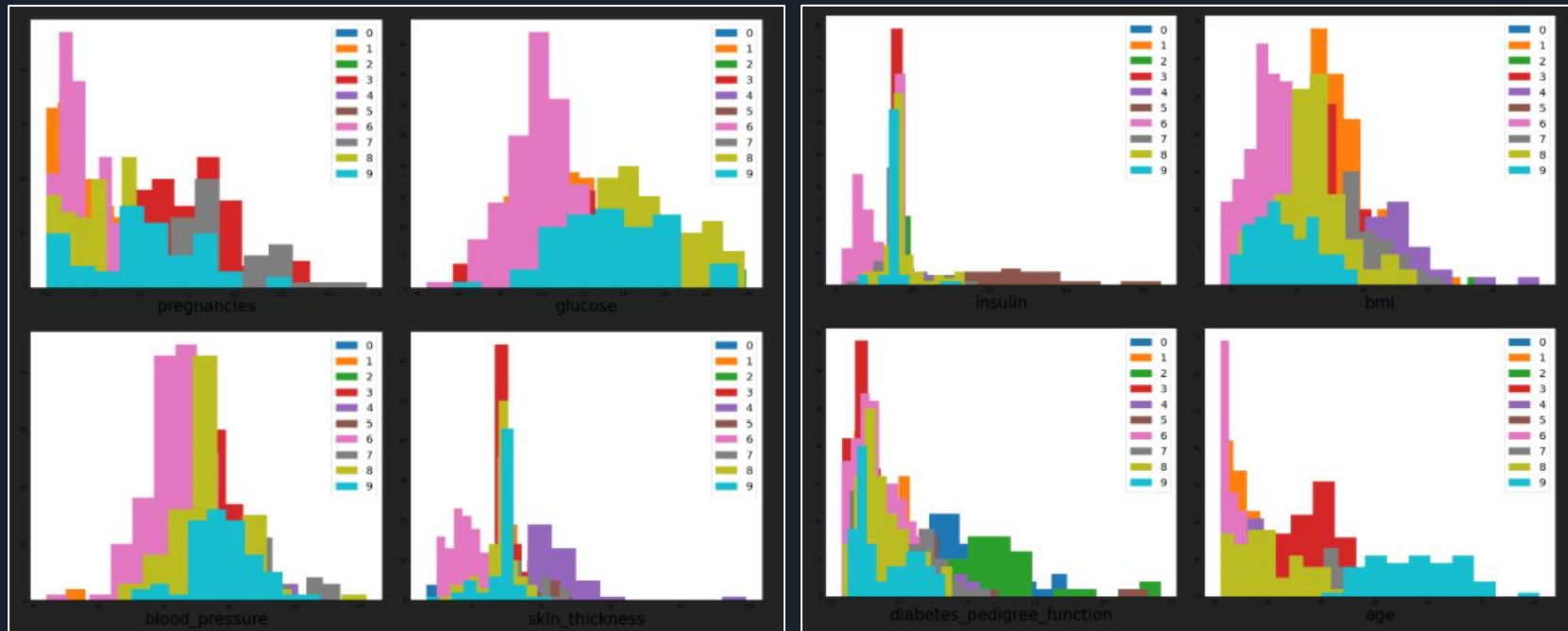
Column of interest : diabetes_pedigree_function

```
Cluster : 0 with mean = 0.9957037037037034
Cluster : 1 with mean = 0.3788934426229511
Cluster : 2 with mean = 1.3129687499999998
Cluster : 3 with mean = 0.338809090909091
Cluster : 4 with mean = 0.4975777777777777
Cluster : 5 with mean = 0.5535999999999999
Cluster : 6 with mean = 0.3426645161290321
Cluster : 7 with mean = 0.4862586206896552
Cluster : 8 with mean = 0.3923627450980392
Cluster : 9 with mean = 0.39316923076923066
```

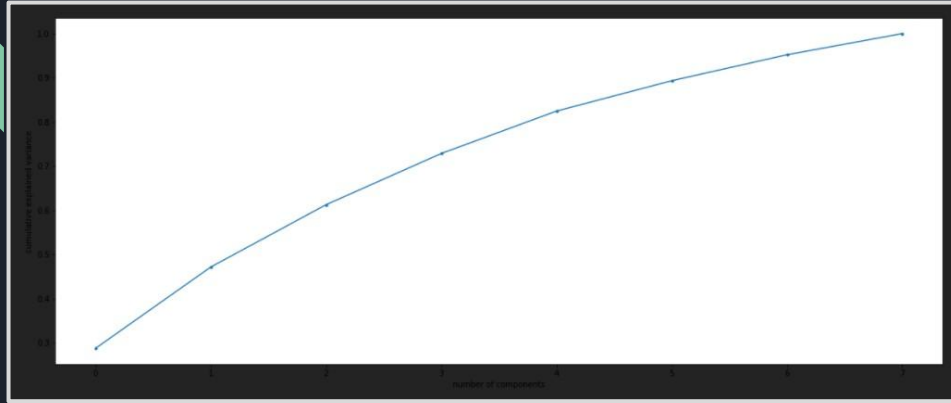
Column of interest : age

```
Cluster : 0 with mean = 29.12962962962963
Cluster : 1 with mean = 25.75409836065574
Cluster : 2 with mean = 39.15625
Cluster : 3 with mean = 38.763636363636365
Cluster : 4 with mean = 28.68888888888889
Cluster : 5 with mean = 33.36
Cluster : 6 with mean = 24.419354838709676
Cluster : 7 with mean = 44.327586206896555
Cluster : 8 with mean = 30.068627450980394
Cluster : 9 with mean = 57.676923076923075
```

6.2 K-means algorithm : All features

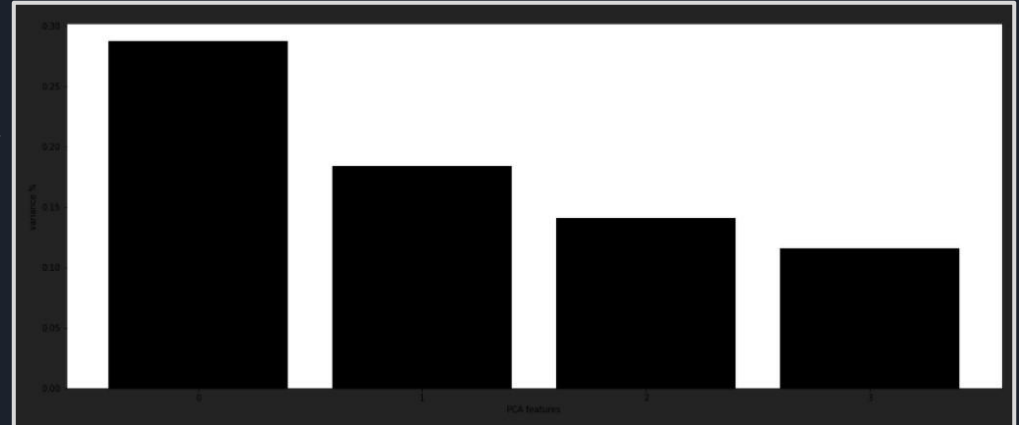
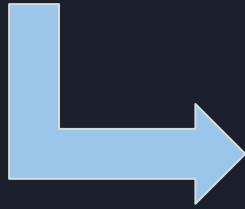


6.3 K-means algorithm: PCA extracted features

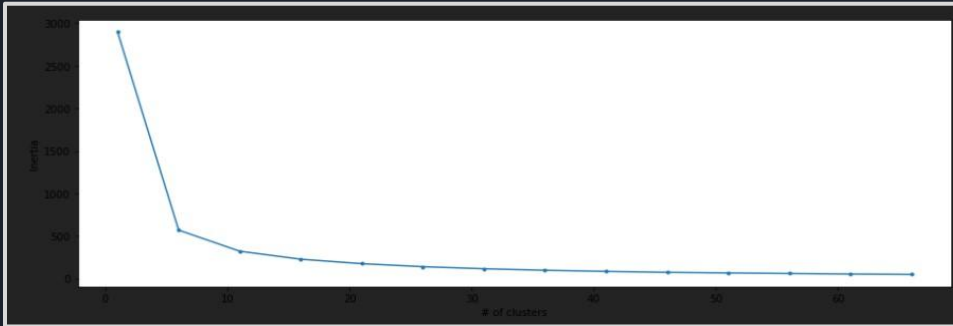
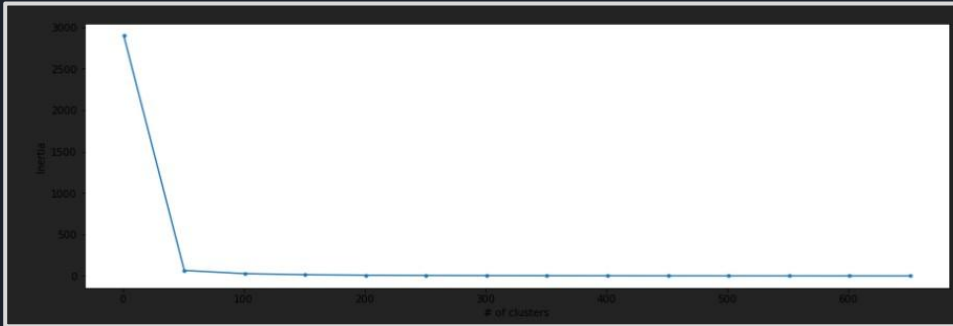
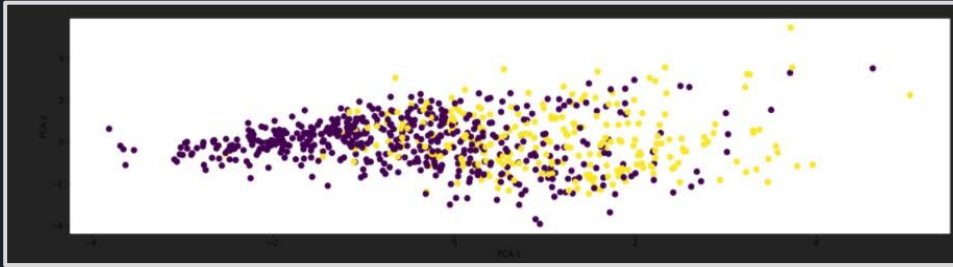


Cumulative explained
variance / n° of components

PCA features /
variance ratio



6.3 K-means algorithm: PCA extracted features



Cluster of interest : 0

Mean in first component = -0.44075916776284346
Mean in second component = -0.8365140772905645

Cluster of interest : 1

Mean in first component = -0.3491494556267769
Mean in second component = 1.3000685203261269

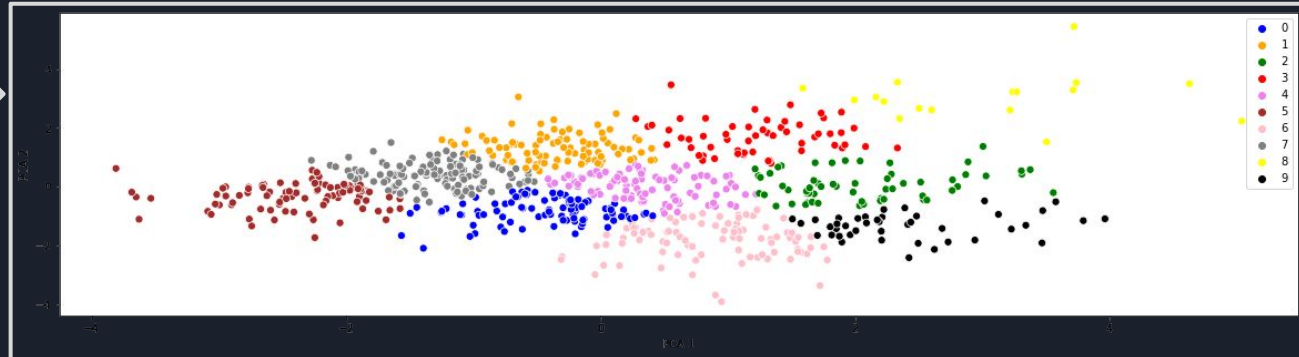
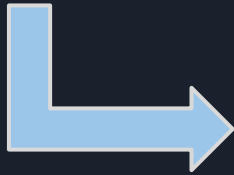
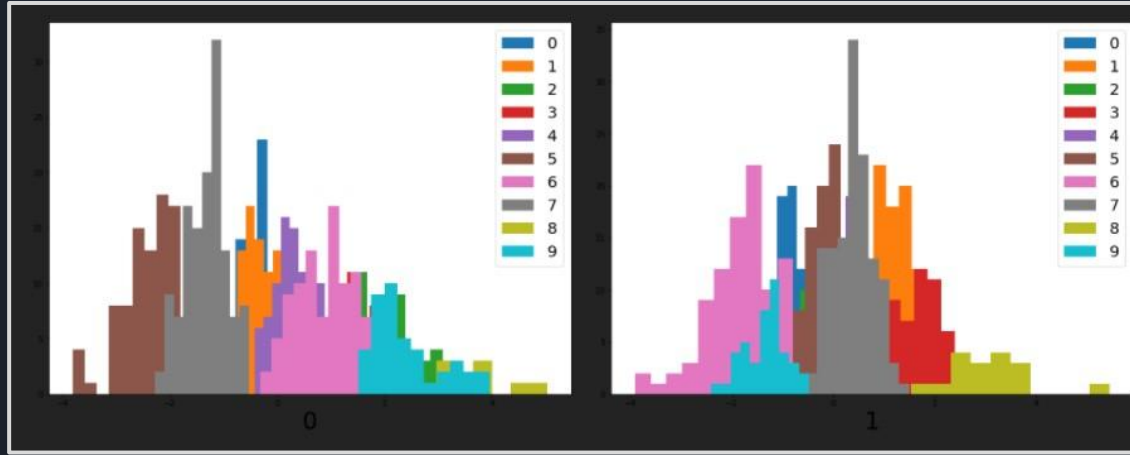
Cluster of interest : 2

Mean in first component = 2.060225204567502
Mean in second component = 0.06248092539692801

Cluster of interest : 3

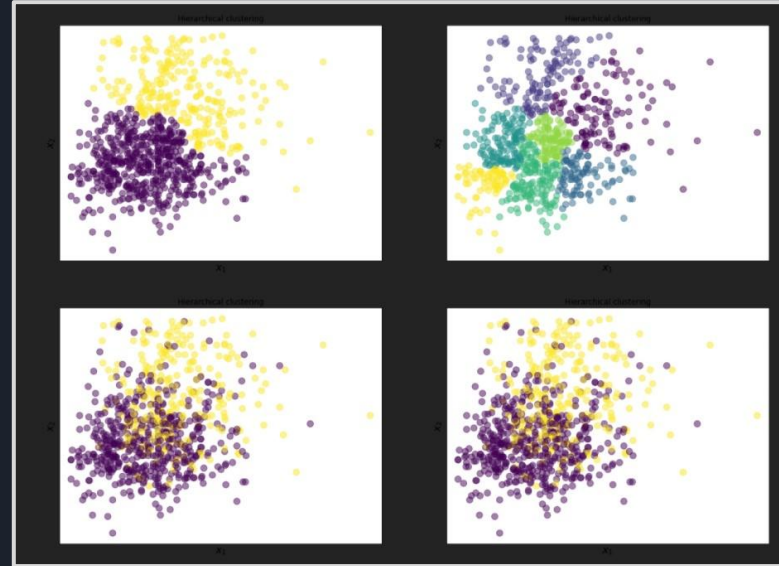
Mean in first component = 1.2314574884718448
Mean in second component = 1.7221308537638556

6.3 K-means algorithm: PCA extracted features



6.4 Hierarchical clustering algorithm

Section 1: BMI-Glucose

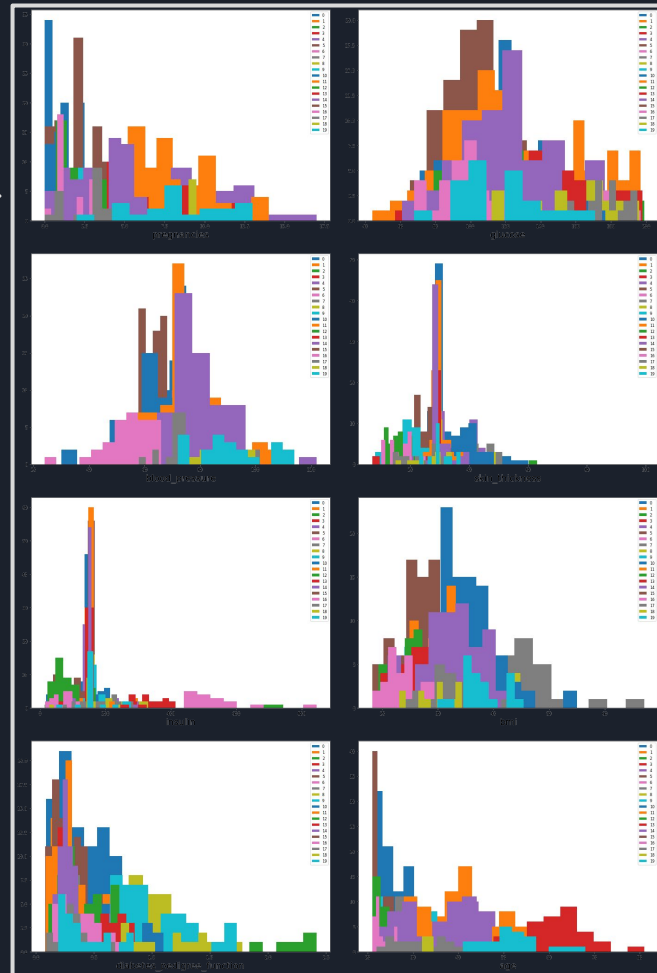
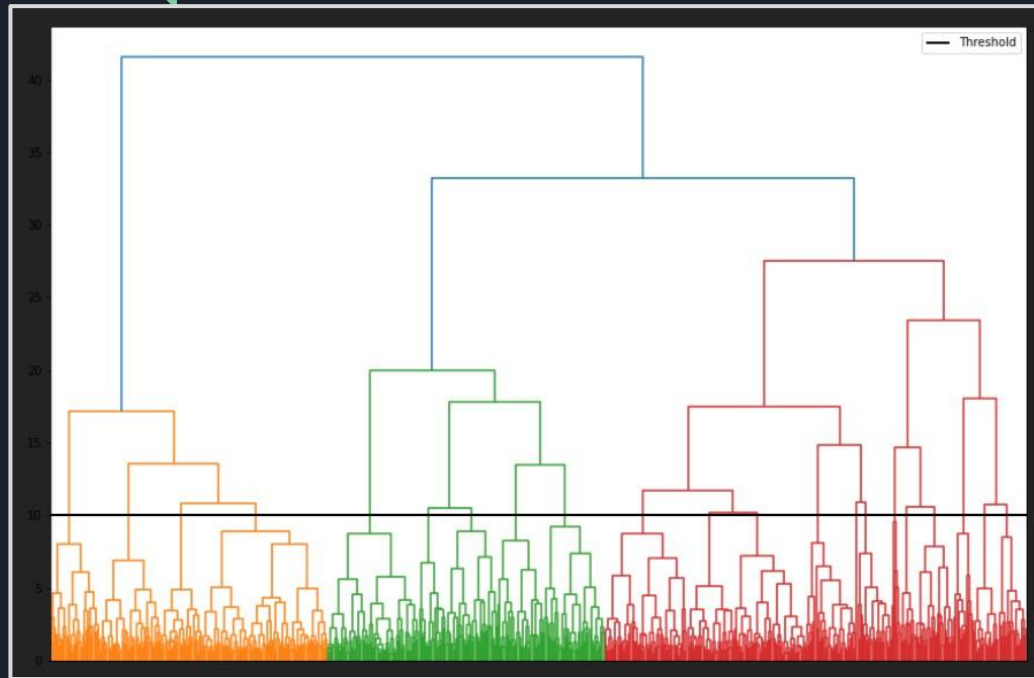


With $k=7$

6.4 Hierarchical clustering algorithm

Section 2: All features

With $k=20$



6.4 Hierarchical clustering algorithm

Section 3: PCA

With k=14

```
Cluster of interest : 0
Mean in first component = -0.19335678758141914
Mean in second component = 0.390114278879464

Cluster of interest : 1
Mean in first component = 1.675413451735883
Mean in second component = 1.9007086262455228

Cluster of interest : 2
Mean in first component = 0.4753634702296078
Mean in second component = 0.35425635573205716

Cluster of interest : 3
Mean in first component = -1.8487709040199816
Mean in second component = 0.053705855293669606
```

