

## **Bio computing Assignment Documentation :**

<b>Names:</b>	<b>IDs:</b>
<i>Mohamed Mosaad Abdeltwab</i>	<i>20198119</i>
<i>Rewan Hisham Abdelaleem</i>	<i>201908039</i>
<i>Alyaa Khaled Ahmed</i>	<i>20198056</i>
<i>Amr Ahmed Atia</i>	<i>20198058</i>

**1) Selected Dataset:** [Mall Customer Segmentation Data | Kaggle](#) .

**'Mall\_Customers.csv' =**

**(CustomerID - Gender - Age - Annual Income (k\$) - Spending Score (1-100)) .**

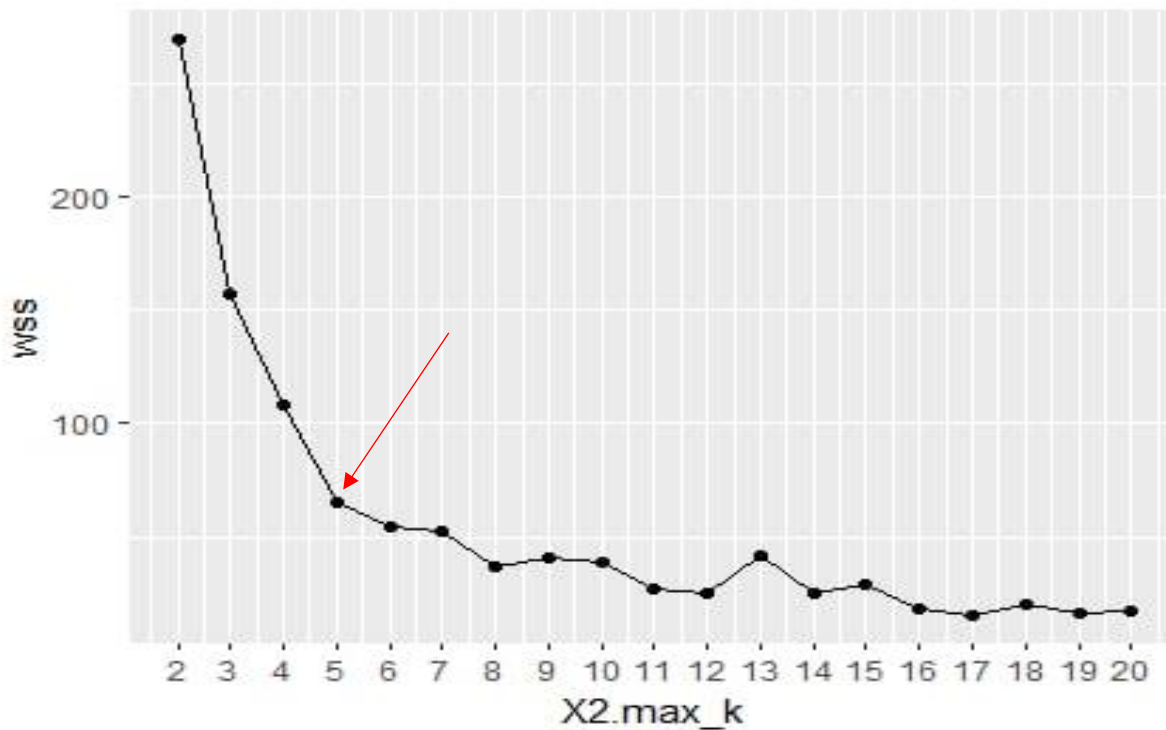
**2) Model:**

**In our model at first, we drop the age and gender `s columns Because they are not the subject of our study and will not help us in anything in representing the data at all, but they will provide a number of the clusters Because I want**

to study specific needs of my data, regardless of the age and gender of this person, so I dropped these two columns.

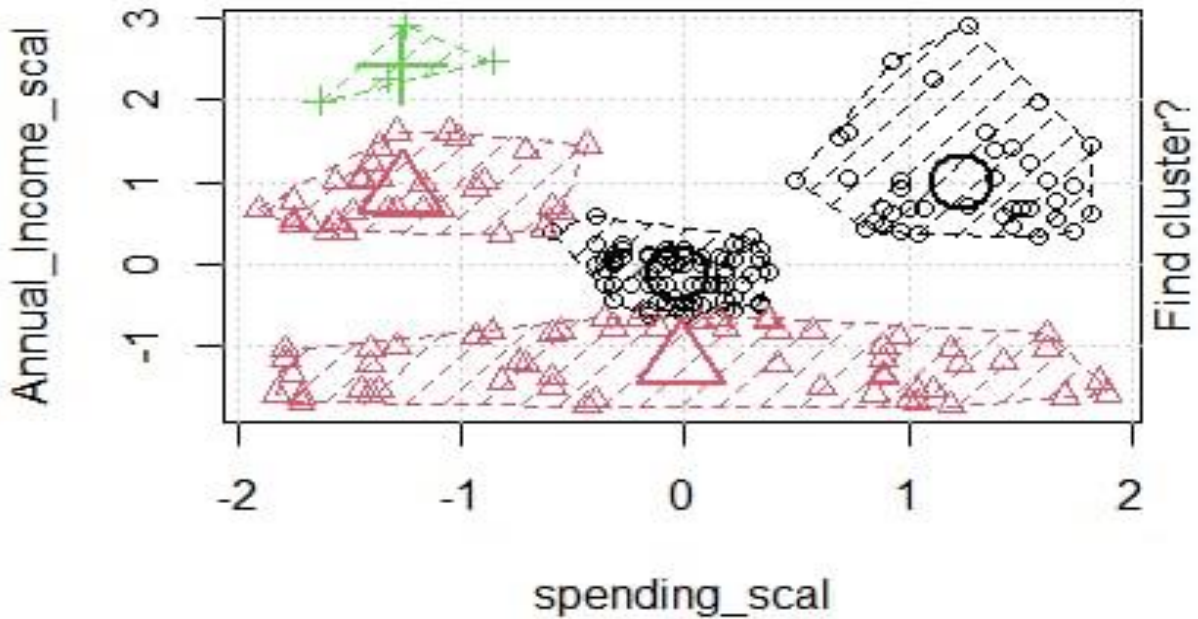
Then in our **k-means model we are using our k=5**, and the Euclidean Distance as an **error function** and it's = [Euclidean (rescale\_df [,1], rescale\_df [,2]))

Because we are doing a rescaling to the data and dropped the columns that I don't use it in my model (age – gender).



### 3) Plotting Graph of clustered data:

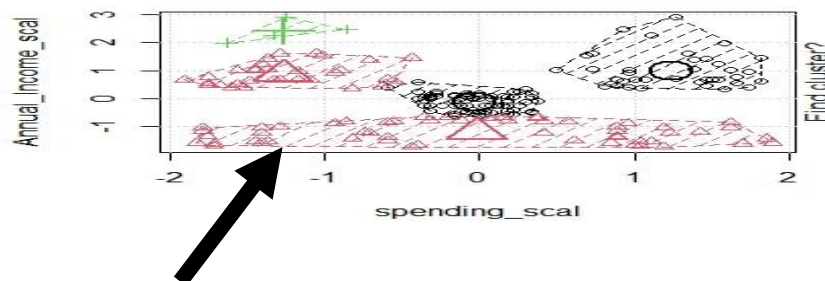
We are plotting the graph by using the animation library.



### 4) Justification of the clustered result:

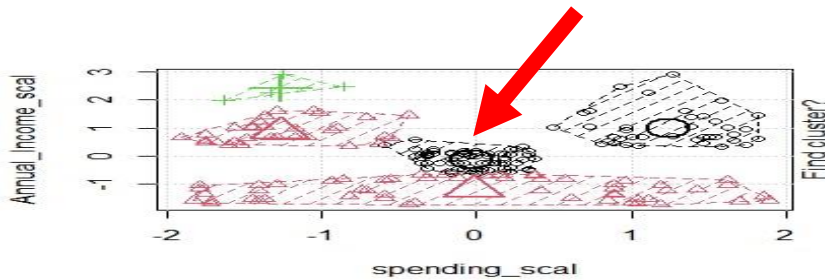
In our model here the Y-axis is representing the annual Incomes (The average income of humans) and the X-axis is representing the Spending Score (What do they spend).

Then we will explain the graph:



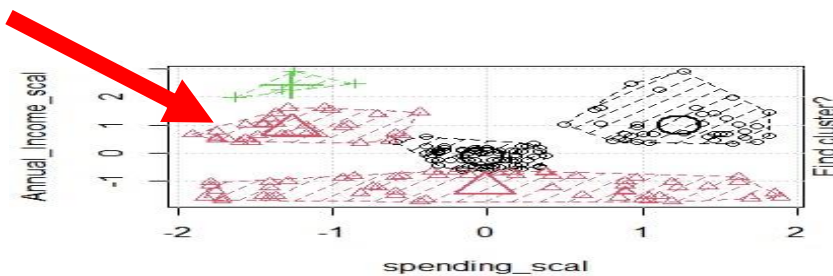
The following shape which the arrow refers to meaning that: Those are people who have little income and spend a lot, and this is a bad situation this is the first cluster.

---



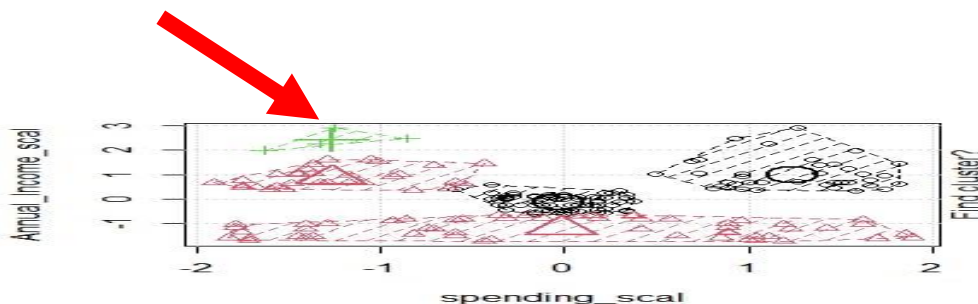
The following shape which the arrow refers to meaning that: Those are people who have average income and average cash, and this is a balanced situation this is the second cluster.

---



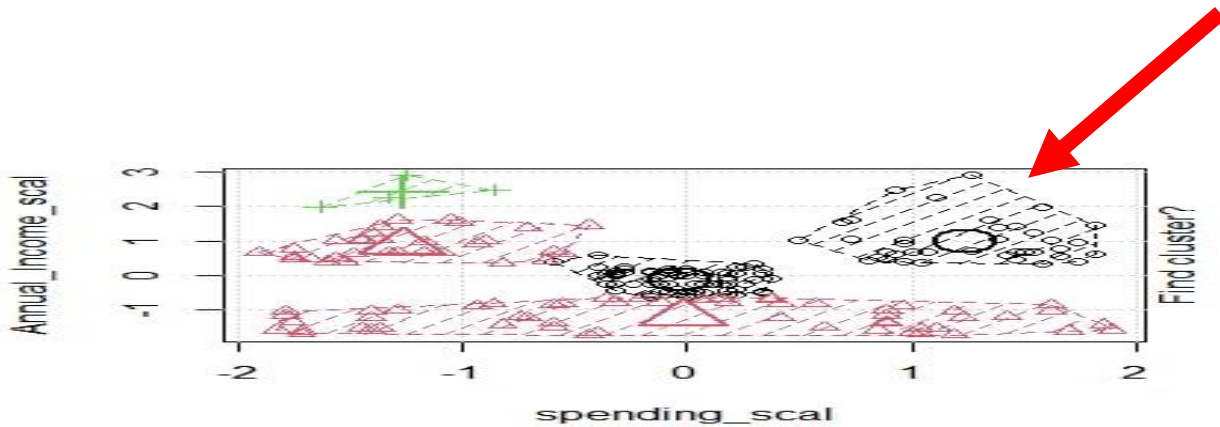
The following shape which the arrow refers to meaning that: Those are people who have very high income and spent little, and this is the third cluster.

---



The following shape which the arrow refers to meaning that: Those are people who have very high income and spent very little, and this is the fourth cluster.

---



The following shape which the arrow refers to meaning that: Those are people who have very high income and spent very high too, and this is the fifth cluster.

---

Thanks