**K-Means**

**Regression : Linear regression**

**Classification : Logistic regression**

**Decision tree**

**Naïve bayes**

**KNN**

**Classification metrics**

**Cluster algorithms: K-means**

**Hierarchical**

**K-means :**

**KNN k ============= ? hyper parameter**

**K-means : k is also a number , hyper parameter**

**Sourav went to B.tech first year group will start**

**4th ========= group**

**Sourav he is not happy with leader**

**Every group has one leader**

**Another group one leader**

1. **We divide the observations randomly into groups**
2. **Every group has mean value**
3. **Will calculate distance between each observation with mean of each group**
4. **Less distance that observation will move to that group**
5. **Again will calculate new means**
6. **Again calculate distance between observation with new means**
7. **Less distance that observation will move to that group**
8. **This process will continue till the means will not change**
9. **That means the group observations are tightly together**
10. **That groups are our final clusters**

| **Individual** | **Variable1** | **Variable2** |
| --- | --- | --- |
| **1** | **1** | **1** |
| **2** | **1.5** | **2** |
| **3** | **3** | **4** |
| **4** | **5** | **7** |
| **5** | **3.5** | **5** |
| **6** | **4.5** | **5** |
| **7** | **3.5** | **4.5** |
|  |  |  |

| **Observations** | **D with Mean-1 (1,1)** | **D with Mean-2 (5,7)** |  |
| --- | --- | --- | --- |
| **(1,1)** | **0** | **7.21** | **G-1** |
| **(1.5,2)** | **1.12** | **6.10** | **G-1** |
| **(3,4)** | **3.61** | **3.61** | **G-1** |
| **(5,7)** | **7.21** | **0** | **G-2** |
| **(3.5,5)** | **4.72** | **2.5** | **G-2** |
| **(4.5,5)** | **5.31** | **2.06** | **G-2** |
| **(3.5,4,5)** | **4.30** | **2.93** | **G-2** |

| **Observations** | **D with Mean-1 (1.83,2.33)** | **D with Mean-2 (4.12,5.38)** |  |
| --- | --- | --- | --- |
| **(1,1)** | **1.57** | **7.21** | **G-1** |
| **(1.5,2)** | **0.47** | **6.10** | **G-1** |
| **(3,4)** | **2.04** | **1.78** | **G-2** |
| **(5,7)** | **5.64** | **1.84** | **G-2** |
| **(3.5,5)** | **3.15** | **0.73** | **G-2** |
| **(4.5,5)** | **3.78** | **0.54** | **G-2** |
| **(3.5,4,5)** | **2.74** | **1.08** | **G-2** |

| **Observations** | **D with Mean-1 (1.25,1.5)** | **D with Mean-2 (3.9,5.1)** |  |
| --- | --- | --- | --- |
| **(1,1)** | **0.56** | **7.21** | **G-1** |
| **(1.5,2)** | **0.56** | **6.10** | **G-1** |
| **(3,4)** | **3.05** | **1.42** | **G-2** |
| **(5,7)** | **6.66** | **2.20** | **G-2** |
| **(3.5,5)** | **4.16** | **0.42** | **G-2** |
| **(4.5,5)** | **4.78** | **0.61** | **G-2** |
| **(3.5,4,5)** | **3.75** | **0.72** | **G-2** |