

K-means

- Centroid-based: describe each cluster by its mean
- Goal: assign data to K
- Algorithm objective: minimize the within-cluster variances of all clusters.

Assignment Project Exam Help

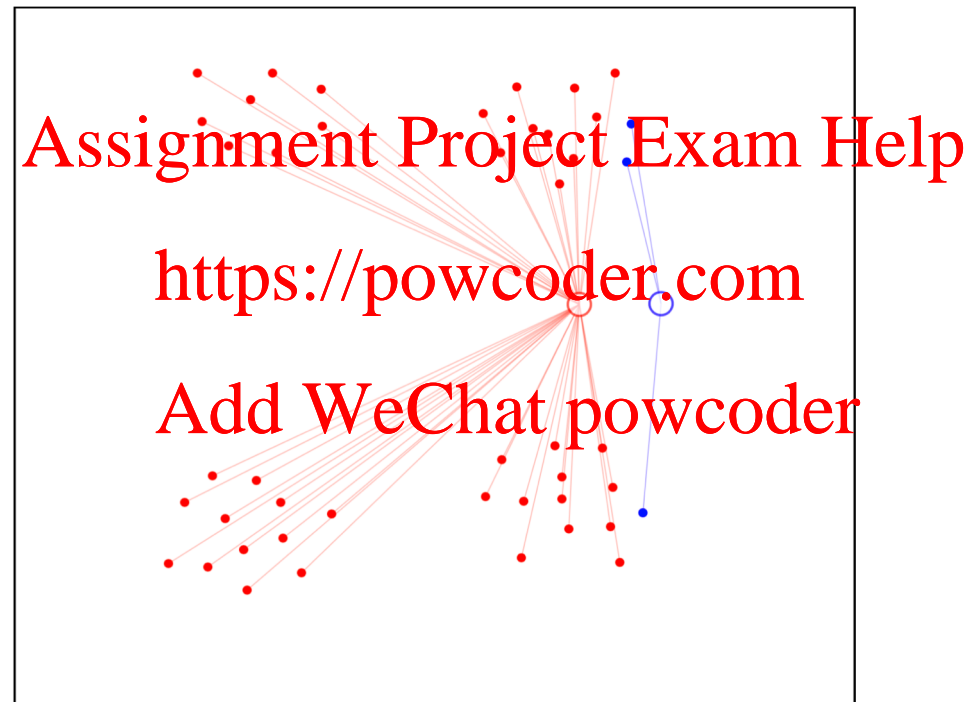
<https://powcoder.com>

Add WeChat powcoder



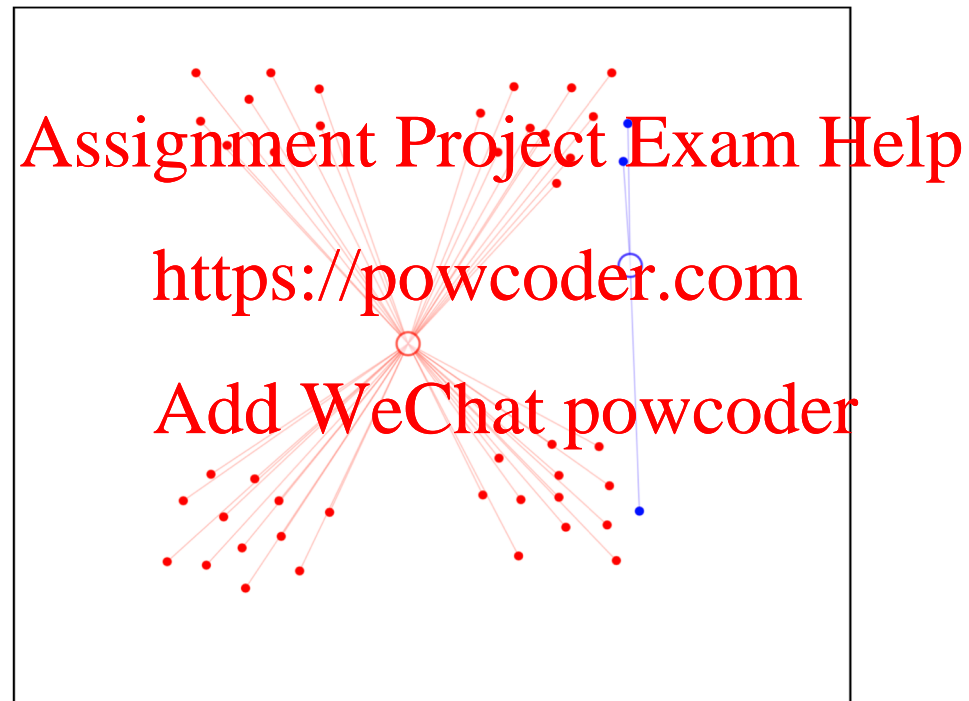
UNIVERSITY OF
BIRMINGHAM

Initialize 2 clusters and assign points to clusters



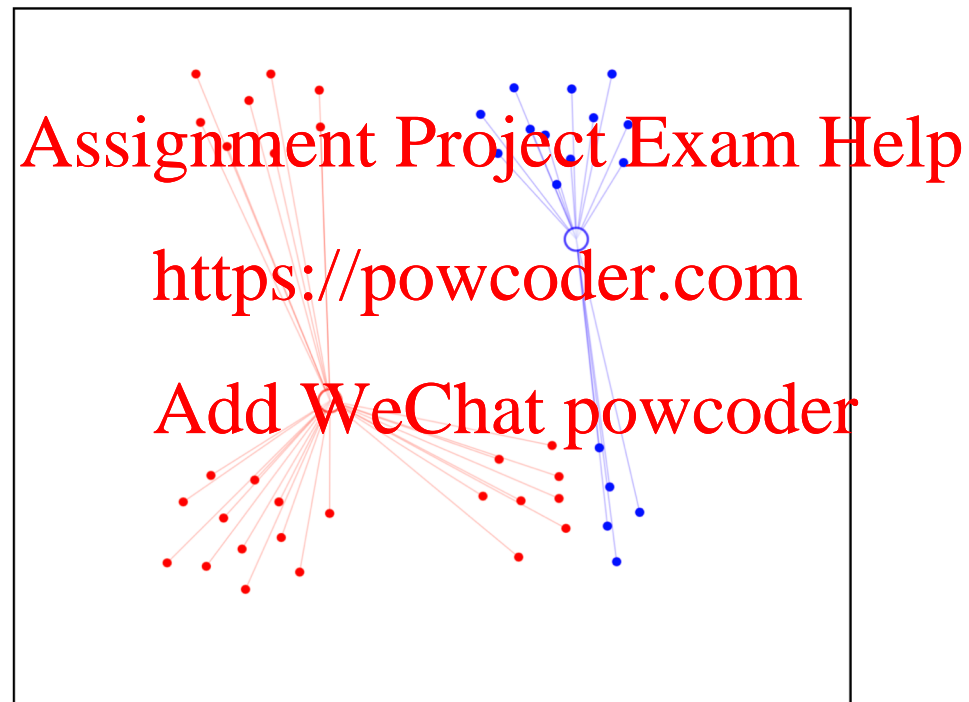
UNIVERSITY OF
BIRMINGHAM

Adjust mean



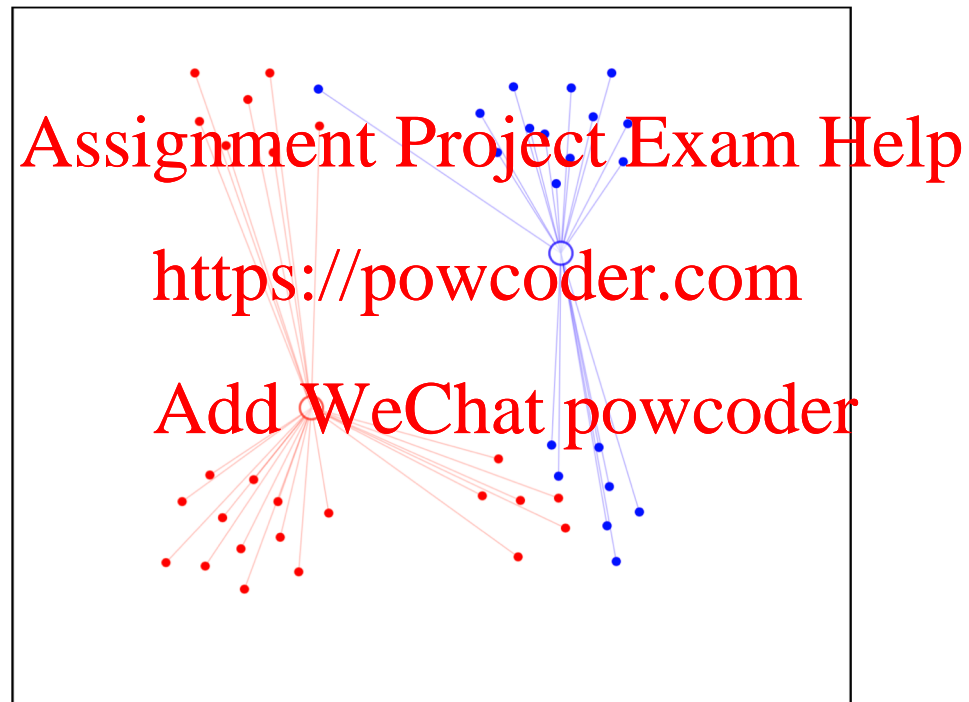
UNIVERSITY OF
BIRMINGHAM

Reassign points to clusters and adjust mean



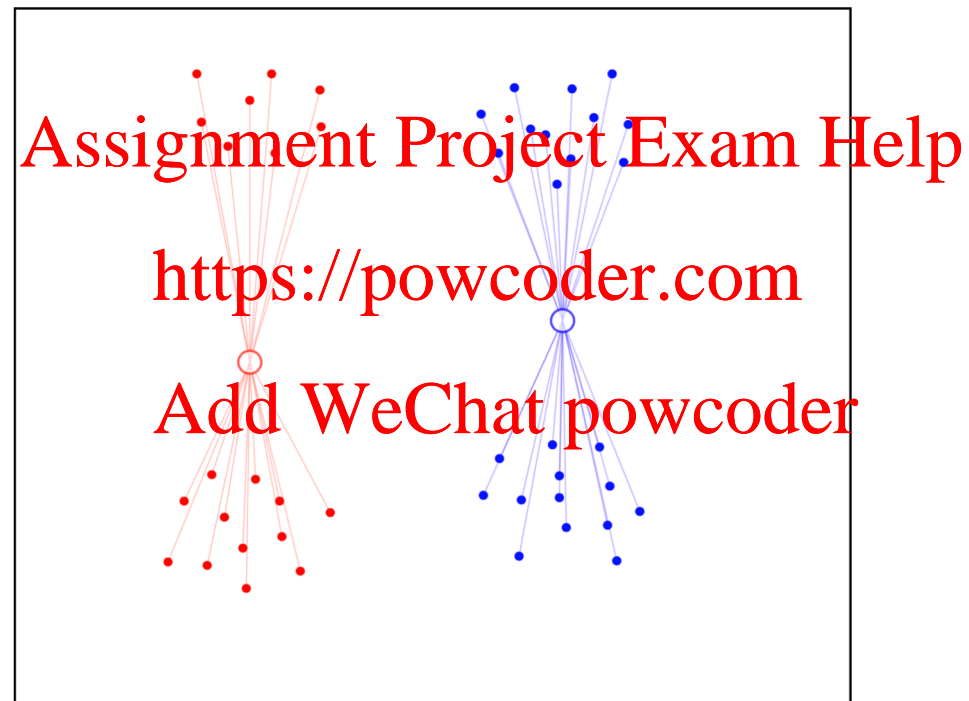
UNIVERSITY OF
BIRMINGHAM

Reassign points to clusters and adjust mean



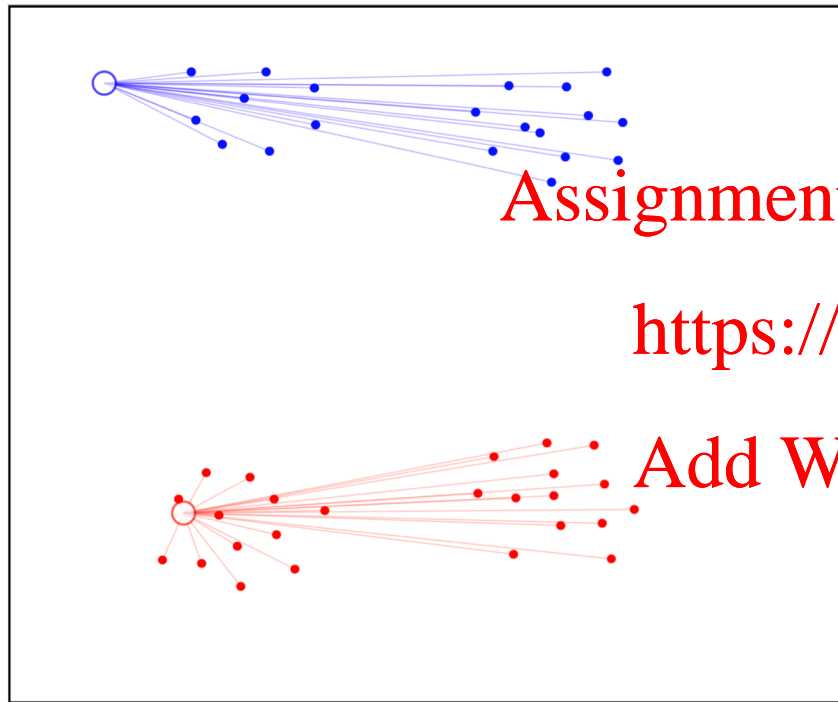
UNIVERSITY OF
BIRMINGHAM

Repeat this, until no cluster changes

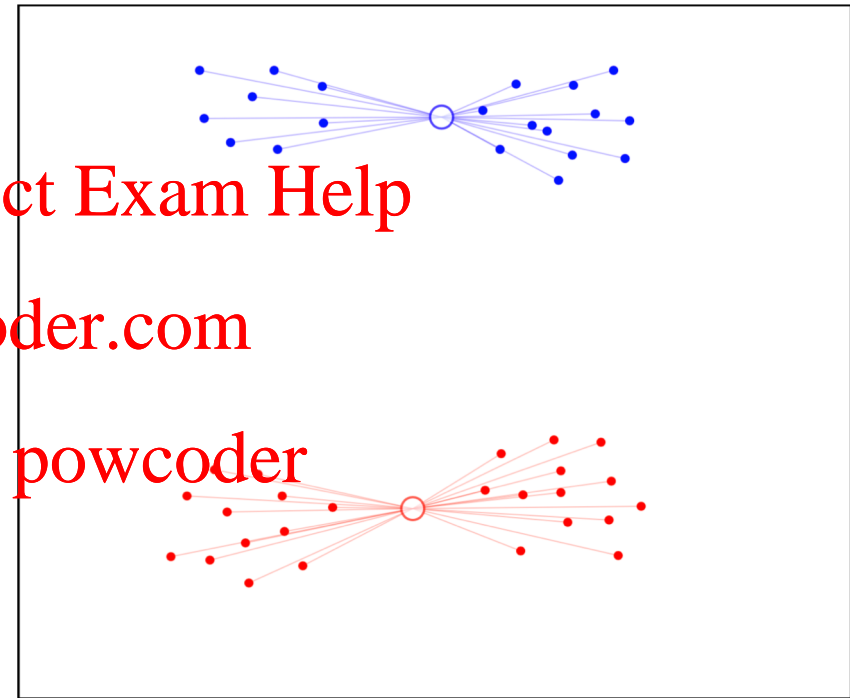


UNIVERSITY OF
BIRMINGHAM

If we have a different starting point



Initial clusters



Final clusters

Assignment Project Exam Help

<https://powcoder.com>

Add WeChat powcoder



UNIVERSITY OF
BIRMINGHAM

K-means

- A non-deterministic method
- Finds a local optimal result (multiple restarts are often necessary)

<https://powcoder.com>

Add WeChat powcoder



UNIVERSITY OF
BIRMINGHAM

Algorithm description

① Initialization

- Data are $\mathbf{x}_{1:N}$
- Choose initial cluster means $\mathbf{m}_{1:k}$ (same dimension as data).

② Repeat

- ### ① Assign each data point to its closest mean

Euclidean distance

$$z_n = \arg \min_{i \in \{1, \dots, k\}} d(\mathbf{x}_n, \mathbf{m}_i) = \arg \min_{i \in \{1, \dots, k\}} \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_i - q_i)^2 + \dots + (p_n - q_n)^2}.$$

- ### ② Compute each cluster mean to be the coordinate-wise average over data points assigned to that cluster.

$$\mathbf{m}_k = \frac{1}{N_k} \sum_{\{n: z_n=k\}} \mathbf{x}_n$$



For each dimension
j of \mathbf{x}_i in cluster k:

$$(\sum_i x_{i,j}) / N_k$$

- ### ③ Until assignments $\mathbf{z}_{1:N}$ do not change



UNIVERSITY OF
BIRMINGHAM

Assignment Project Exam Help

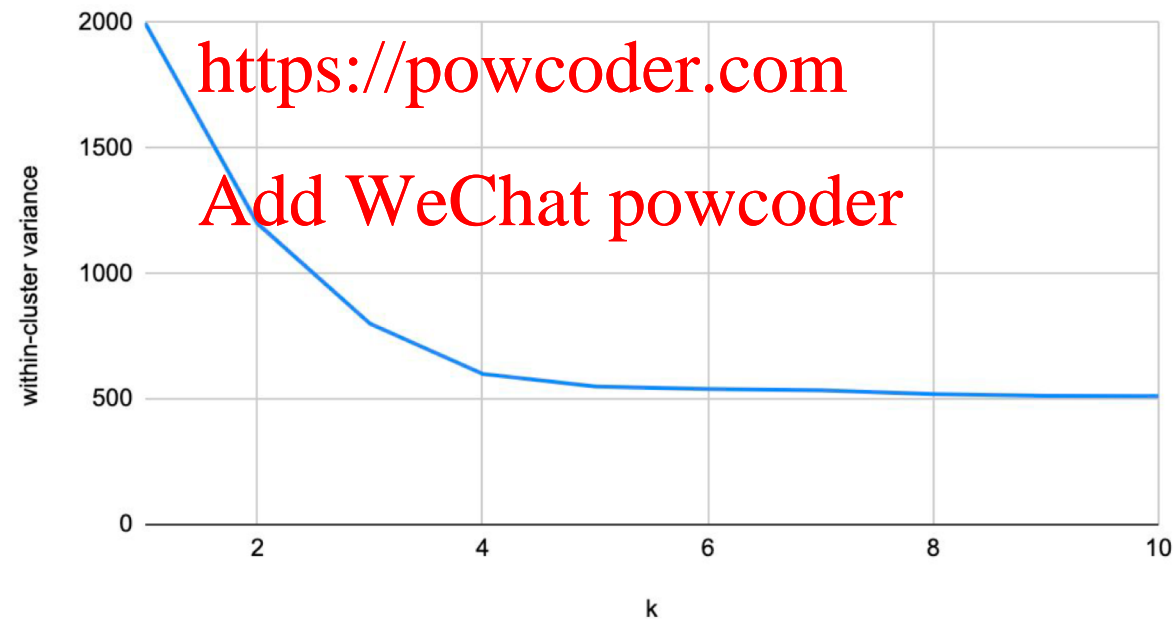
<https://powcoder.com>

Add WeChat powcoder

K-means: finding optimal k

- Plot the cost for each k and find the “Elbow”

Assignment Project Exam Help



UNIVERSITY OF
BIRMINGHAM