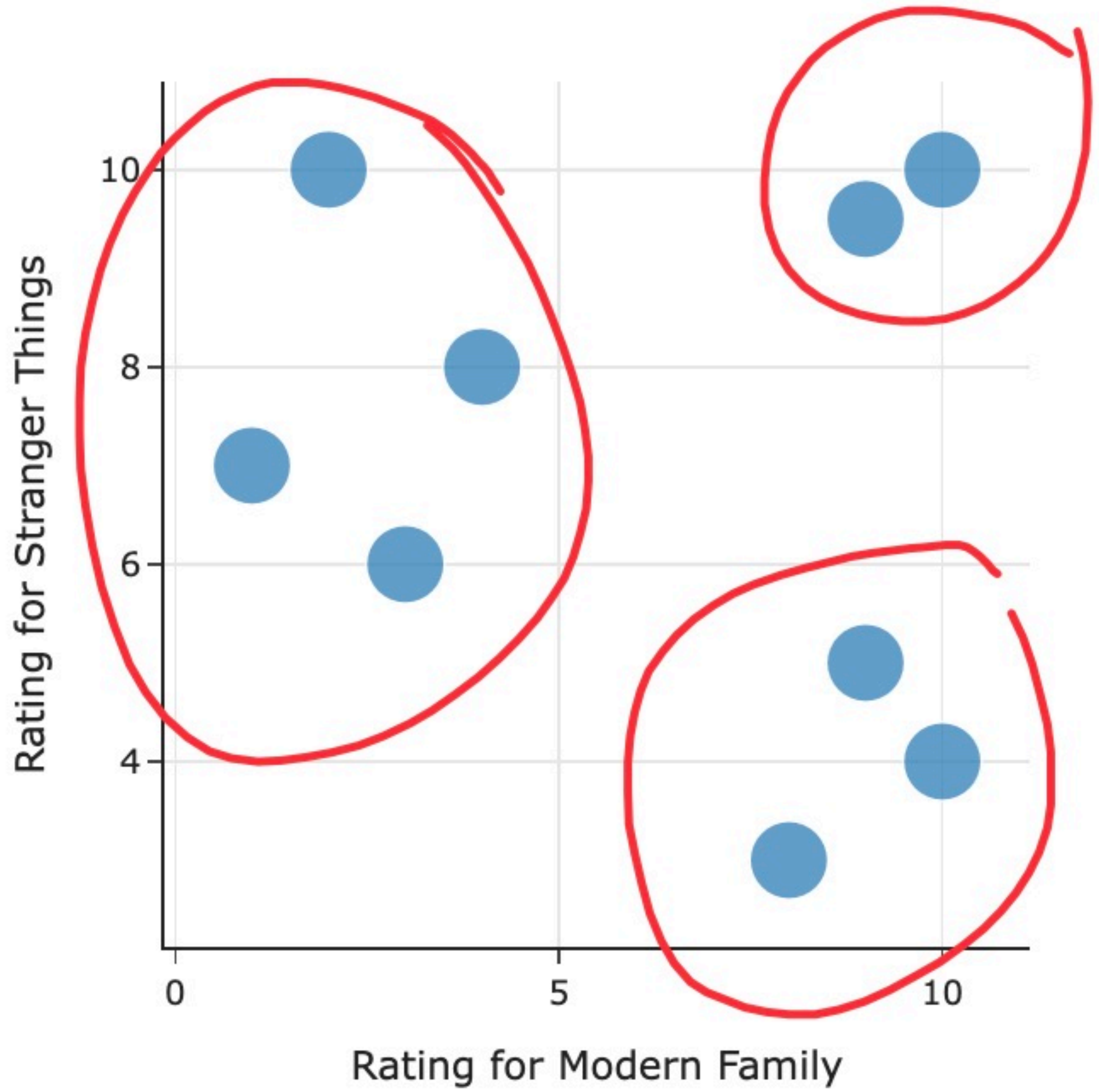
```
In [3]:  1 util.show_ratings()
```



3 natural groups

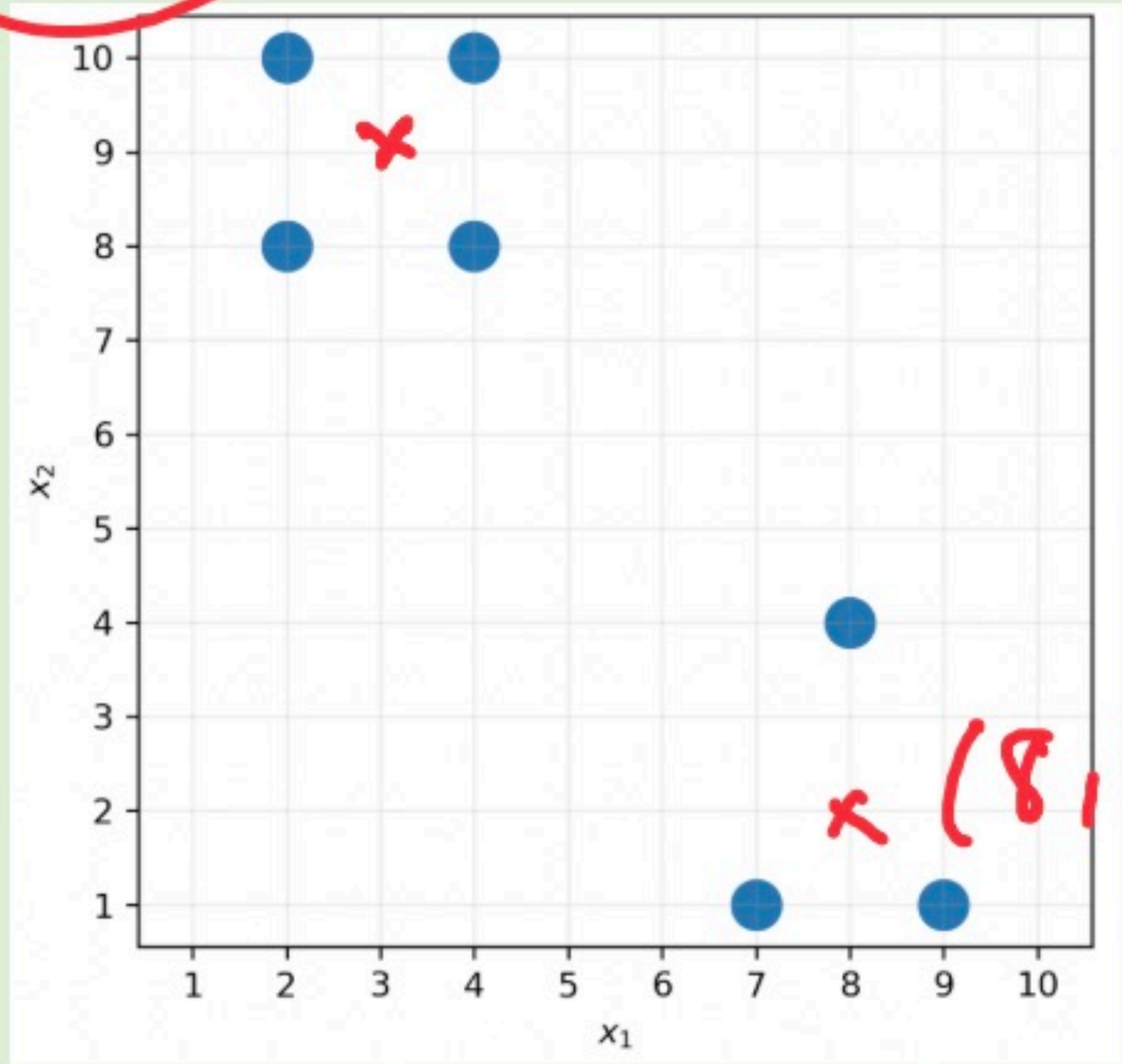
Activity

Recall, inertia is defined as follows:

$$I(\vec{\mu}_1, \vec{\mu}_2, \ldots, \vec{\mu}_k) = \text{total squared distance of each point } \vec{x}_i \text{ to its closest centroid } \vec{\mu}_j$$

why?
mean minimizes MSE.

Suppose we arrange the dataset below into $k = 2$ clusters. What is the **minimum possible inertia**?



(avg x, avg y)

k-means clustering (i.e. Lloyd's algorithm)

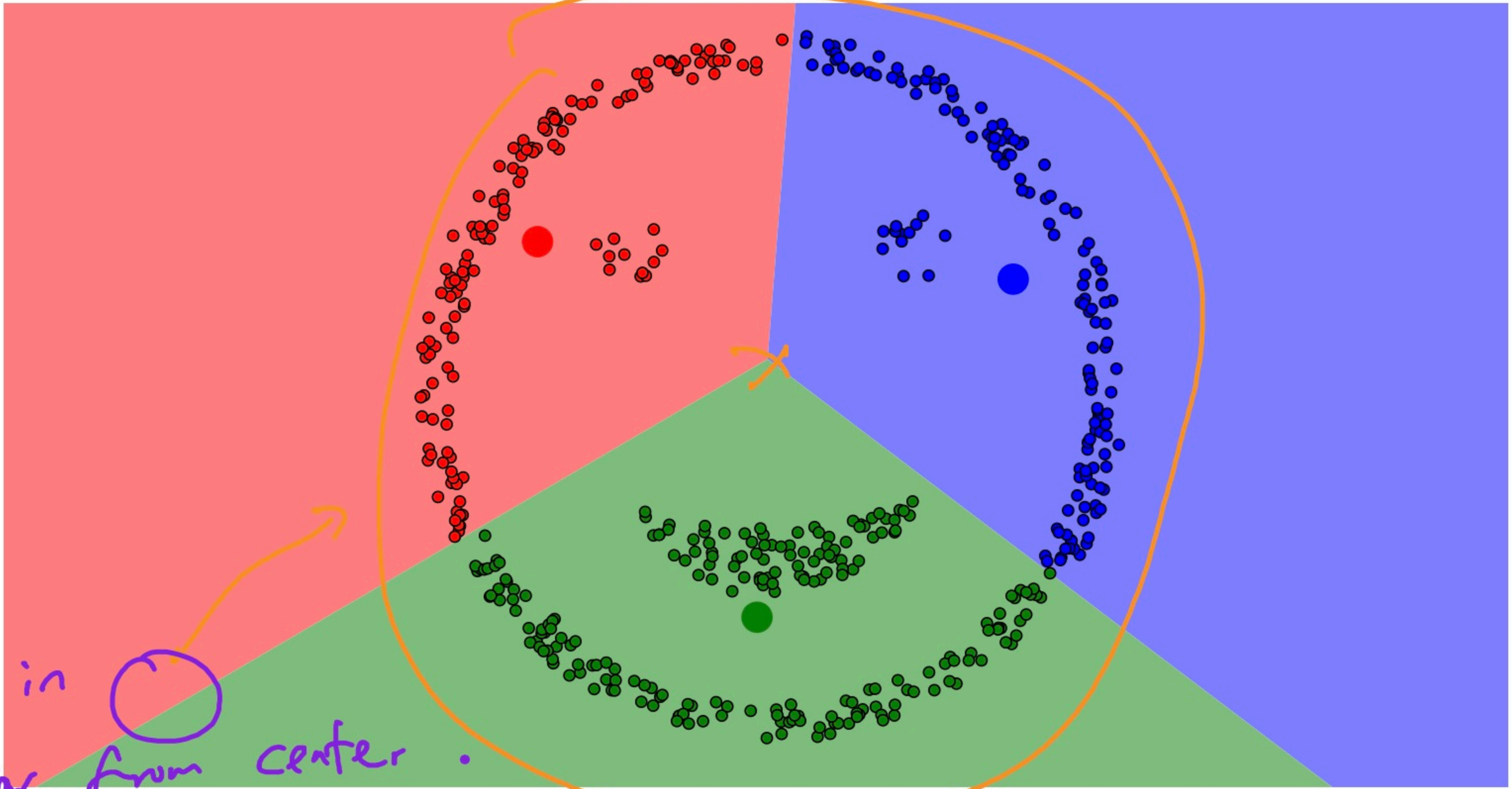
- Fortunately, there's an efficient algorithm that (tries to) find the centroid locations that minimize inertia. The resulting clustering technique is called **k-means clustering**.
  Note that this has no relation to k-nearest neighbors, which we used for both regression and classification. Remember that clustering is an unsupervised technique!

0. **Randomly** initialize k centroids.
   There are other ways of initializing the centroids as well.

1. Assign each point to the nearest centroid. "color"

2. Move each centroid to the **center** of its group. "move"
   We compute the center of a group by taking the mean of the group's coordinates.

3. **Repeat** steps 1 and 2 until the centroids stop changing!
   This is an iterative algorithm!



points in all far from center.

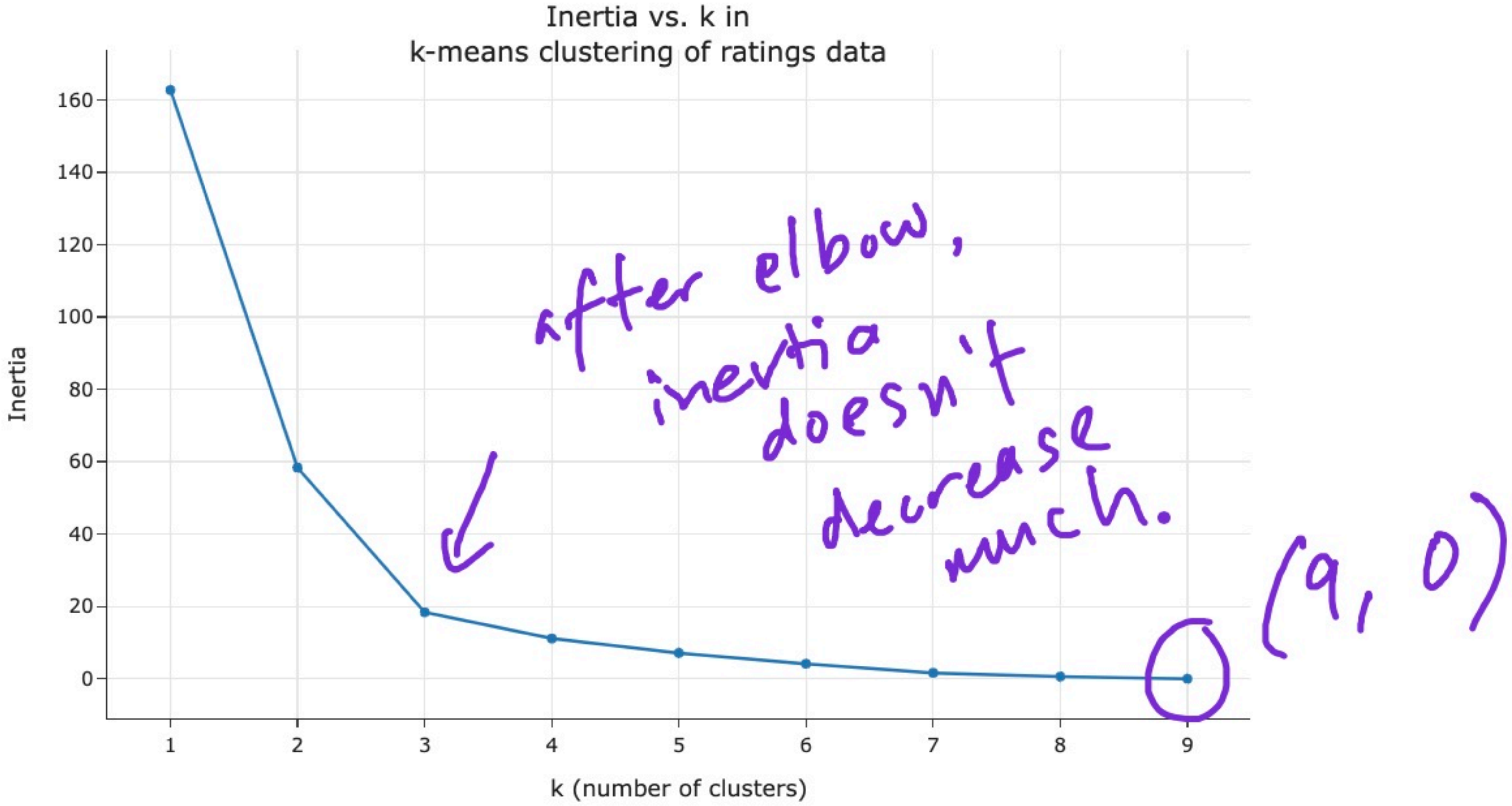
Restart Reassign Points

K-Means Algorithm

The k-means algorithm captures the insight that each point in a cluster should be near to the center of that cluster. It works like this: first we choose k, the number of clusters we want to find in the data. Then, the centers of those k clusters, called *centroids*, are

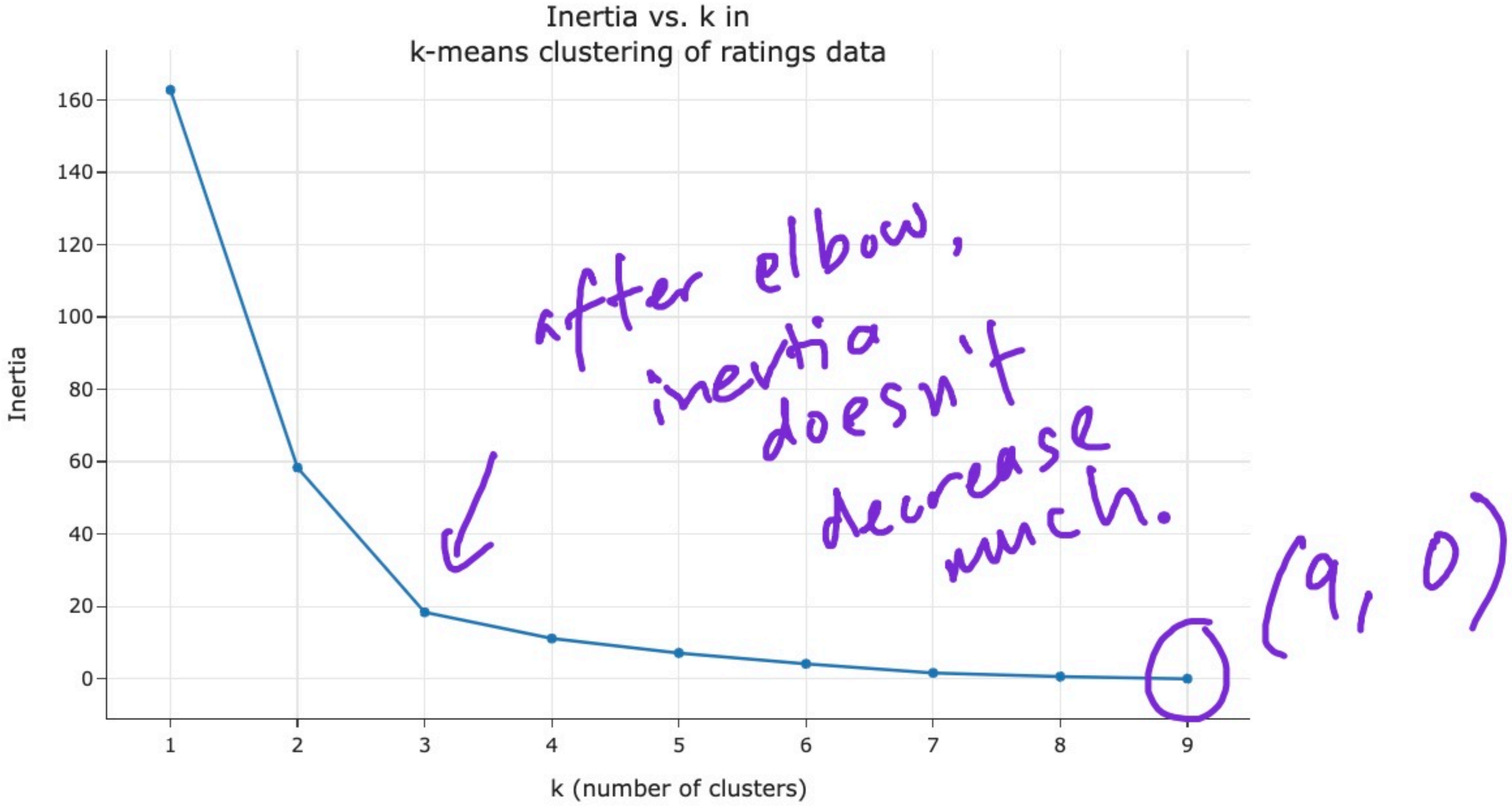
- For several different values of k, let's compute the inertia of the resulting clustering, using the scatter plot from the previous slide.

In [18]:
```
util.show_elbow()
```

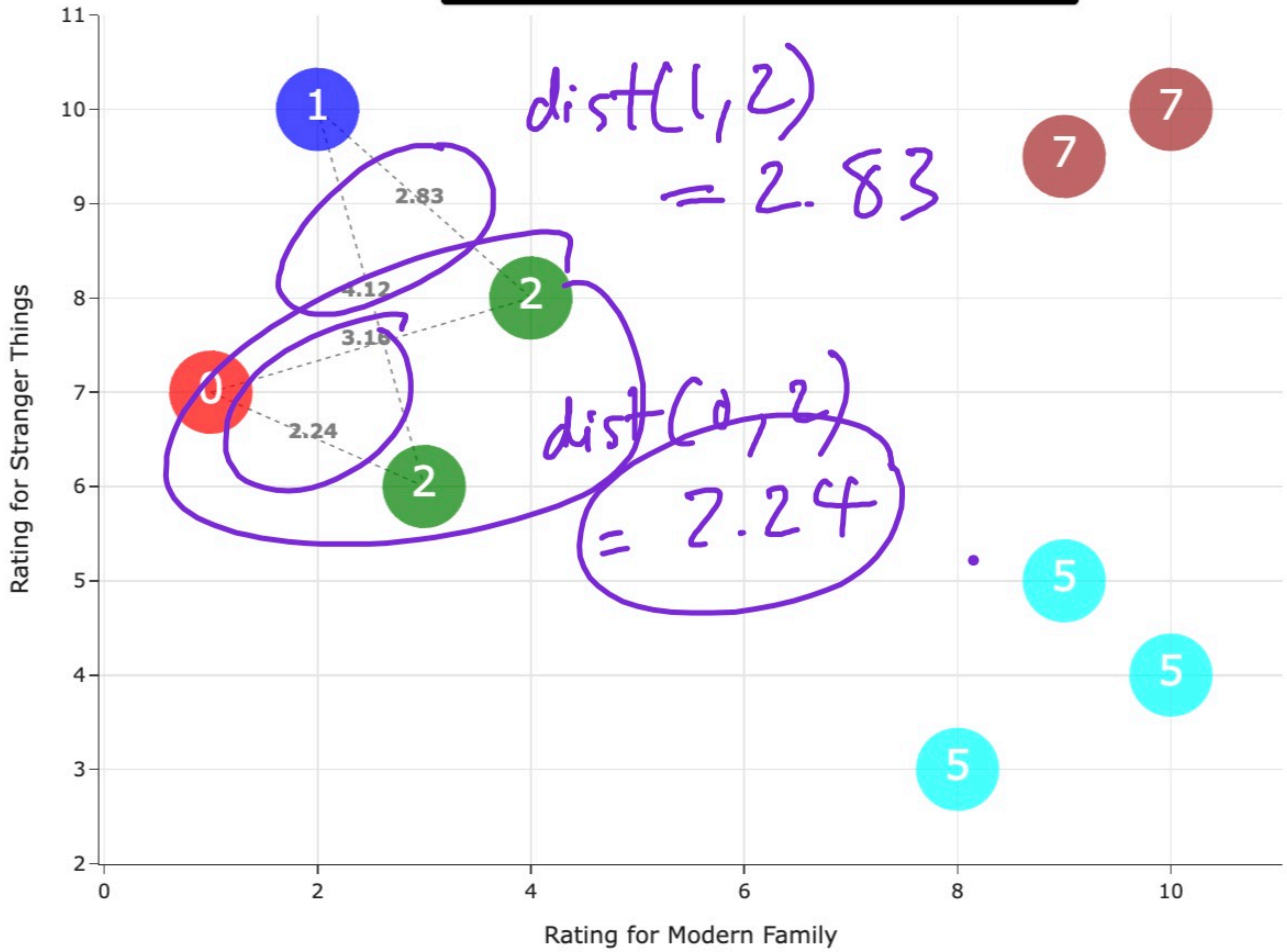


- For several different values of k, let's compute the inertia of the resulting clustering, using the scatter plot from the previous slide.

In [18]:
```
util.show_elbow()
```



32.1


localhost
dist(1,2)
= 2.83
dist(0,2)
= 2.24
2.83
4.12
3.16
2.24
Rating for Stranger Things
Rating for Modern Family
0
1
2
2
5
5
5
7
7