K-Means Clustering: Math Behind It

K-Means is a widely used clustering algorithm that partitions a dataset into *K* clusters. It is an unsupervised machine learning algorithm that attempts to minimize the distance between data points and the centroid of their assigned cluster. Below is an explanation of the math behind K-Means:

1. **Initialization**:

   - Choose the number of clusters, *K*.
   - Randomly initialize *K* centroids (either randomly select *K* data points as centroids or randomly assign values in the feature space).

////// step 2and 3 in pic

4. **Repeat**:

```
-    Steps 2 and 3 are repeated until convergence. Convergence is achieved w
    -    The assignments of data points to clusters no longer change.
    -    The centroids no longer move significantly (or move at all).
```