

## Algorithm: K-Means Clustering

### Step 1: Import required libraries

- Import `numpy`, `matplotlib.pyplot`, `pandas`, and `KMeans` from `sklearn.cluster`.
- Suppress warnings for cleaner output.

### Step 2: Load the dataset

- Read the file `Mall_Customers.csv` using `pandas.read_csv()`.
- Display the first few rows to verify the data.

### Step 3: Select features for clustering

- Extract columns **Annual Income (k\$)** and **Spending Score (1–100)** for analysis.
- Store them in variable `x`.

### Step 4: Determine the optimal number of clusters using the **Elbow Method**

- Initialize an empty list `wcss` (Within Cluster Sum of Squares).
- Loop `i` from 1 to 10:
  - Apply `KMeans` with `i` clusters.
  - Fit the model and append the inertia (WCSS value) to the list.
- Plot the curve of `wcss` vs Number of Clusters.
- The “elbow point” in the graph suggests the optimal number of clusters (commonly 5).

### Step 5: Apply K-Means clustering

- Create a `KMeans` model with `n_clusters = 5`, `init = 'k-means++'`, and `random_state = 0`.
- Fit and predict cluster labels for `x` using `fit_predict()`.

### Step 6: Visualize the clusters

- Plot data points for each cluster with different colors.
- Plot the **centroids** of the clusters in yellow.
- Label the axes and provide a legend and title.