# What is Unsupervised Learning?

**Definition**:  
Unsupervised learning is a type of machine learning where the model is not given labeled data.  
That means there are no target values or outcomes. The goal is to find hidden patterns, structures, or relationships in the data.

**Examples of Unsupervised Learning**

1. **Customer Segmentation** – Grouping customers based on their buying behavior.
2. **Market Basket Analysis** – Finding items frequently bought together.
3. **Anomaly Detection** – Detecting frauds or unusual activities.
4. **Document Clustering** – Grouping similar news articles or research papers.
5. **Image Segmentation** – Separating objects within an image.

**Types of Unsupervised Learning**

1. **Clustering** – Grouping similar data points together.  
   Example: K-Means, Hierarchical Clustering
2. **Association** – Finding rules that describe data relationships.  
   Example: Apriori Algorithm (used in market basket analysis)
3. **Dimensionality Reduction** – Reducing data features while keeping important information.  
   Example: PCA (Principal Component Analysis)

## **What is K-Means Clustering?**

**Definition:**  
K-Means is one of the most popular **clustering algorithms**.  
It divides data into **K distinct, non-overlapping groups (clusters)** based on **similarity**.

**When to Use K-Means**

You can use K-Means when:

* You want to group similar data points.
* Data is **numerical and continuous**.
* You know (or can estimate) the number of clusters (K).
* Clusters are roughly **spherical and evenly sized**.

Avoid when:

* Data has different scales or shapes (not spherical).
* Contains outliers or categorical features.

## Code

# Step 1: Import libraries

import numpy as np

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

from sklearn.datasets import make\_blobs

# Step 2: Create sample data

X, y = make\_blobs(n\_samples=300, centers=4, cluster\_std=0.6, random\_state=0)

# Step 3: Visualize data before clustering

plt.scatter(X[:, 0], X[:, 1])

plt.title("Unclustered Data")

plt.show()

# Step 4: Apply KMeans

kmeans = KMeans(n\_clusters=4)

kmeans.fit(X)

y\_kmeans = kmeans.predict(X)

# Step 5: Visualize clustered data

plt.scatter(X[:, 0], X[:, 1], c=y\_kmeans, cmap='rainbow')

plt.scatter(kmeans.cluster\_centers\_[:, 0], kmeans.cluster\_centers\_[:, 1],

            s=200, c='black', marker='X', label='Centroids')

plt.title("K-Means Clustering Result")

plt.legend()

plt.show()

A diagram of blue dots

AI-generated content may be incorrect.A diagram of different colored dots

AI-generated content may be incorrect.