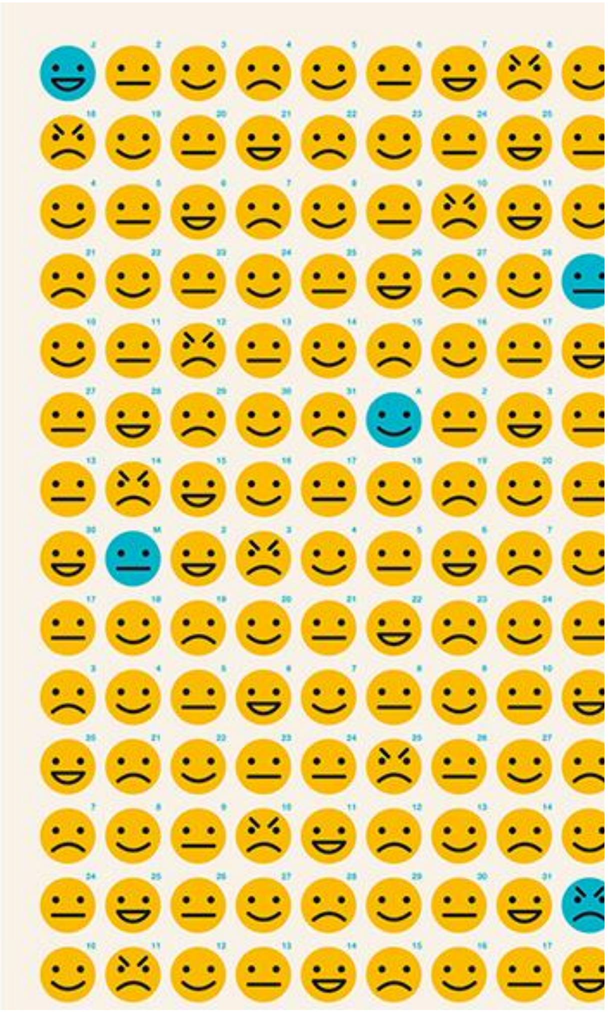
Business Data Analytics

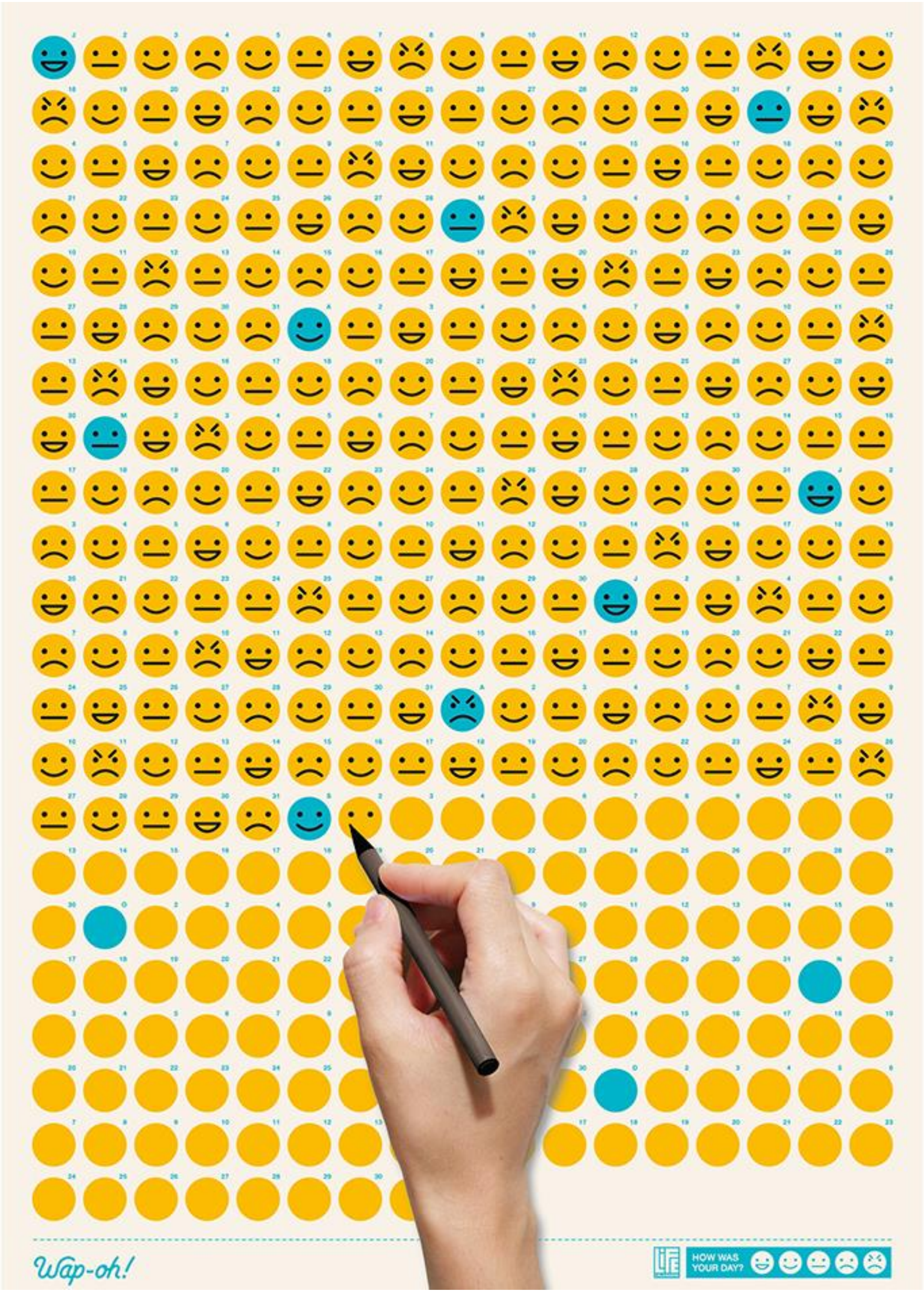
Lecture 3
Customer Segmentation

Rajesh Sharma
https://css.cs.ut.ee/

How'd you do it?

Imagine you have become a business owner of a big company
in the automobile sector and wish to redefine the focus of your production line.
You want to know what and how much to produce (SUV, 4x4, coupe, sedan, wagon, etc..)
You have a big client base.




Wap-oh!
HOW WAS YOUR DAY?

Marketing and Sales

Customer Segmentation

Lecture 1

Business Data Analytics

Repeatable, Decision, Mechanism (Approach), Objective, Segmentation, Classification and Prediction

Descriptive Analysis

Lecturer 2

Data Exploration

Descriptive (numbers) & Visualization

Lecture 3

Descriptive Analysis

Customer (Data) Segmentation

Outline

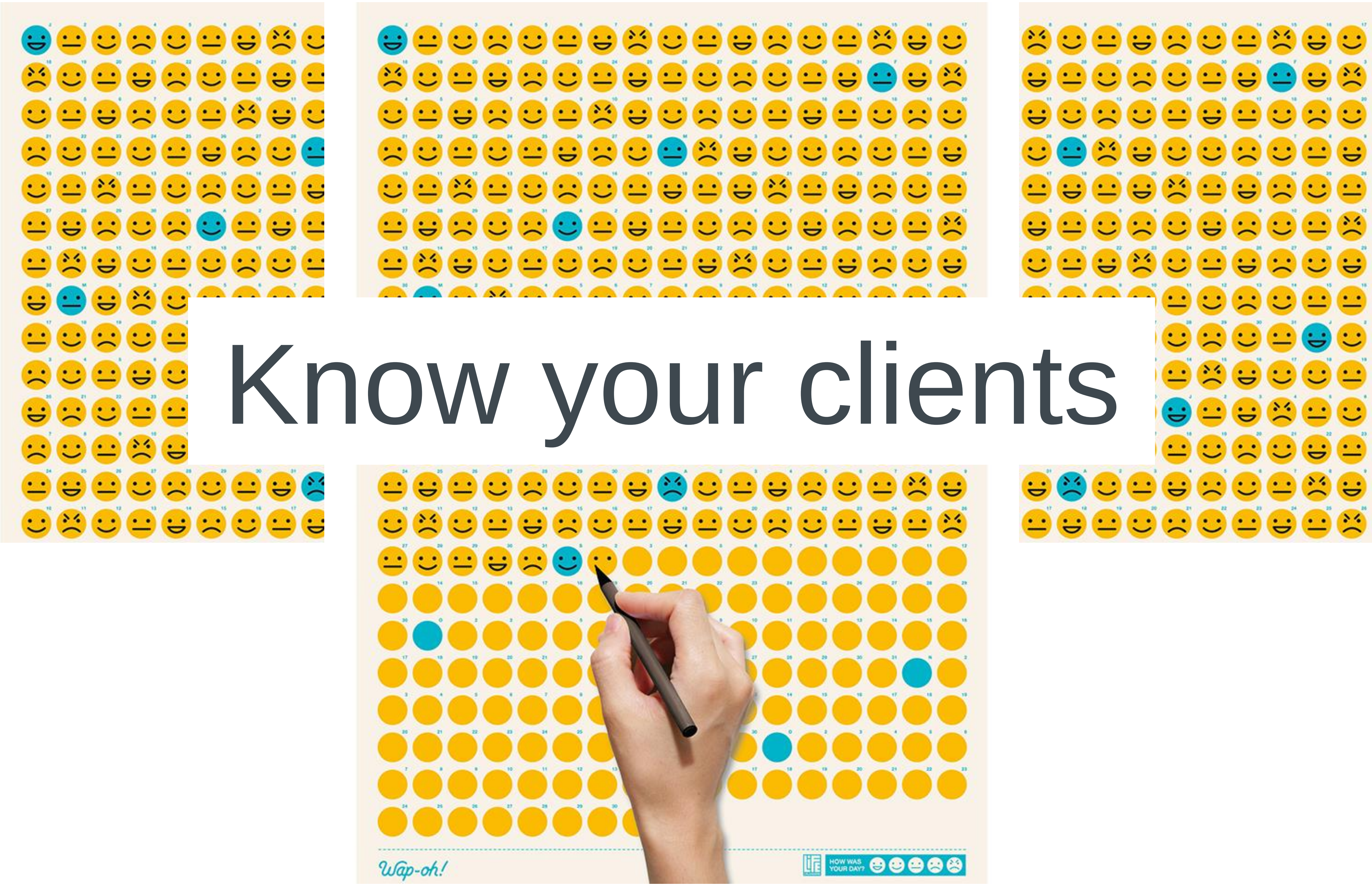
Techniques for Customer Segmentation

- Intuition Based

- Historical/Behavior based
  - RFM
  - Value Tier
  - Life Cyclestage

- Data Driven
  - K-Means
  - Hierarchical Clustering
  - DBSCAN

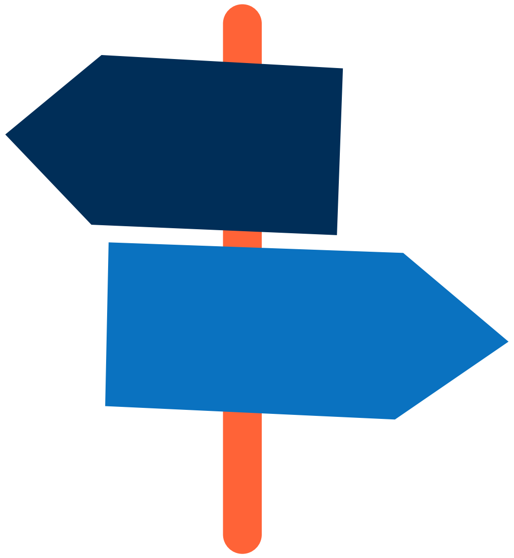
Know your clients



Customer segmentation

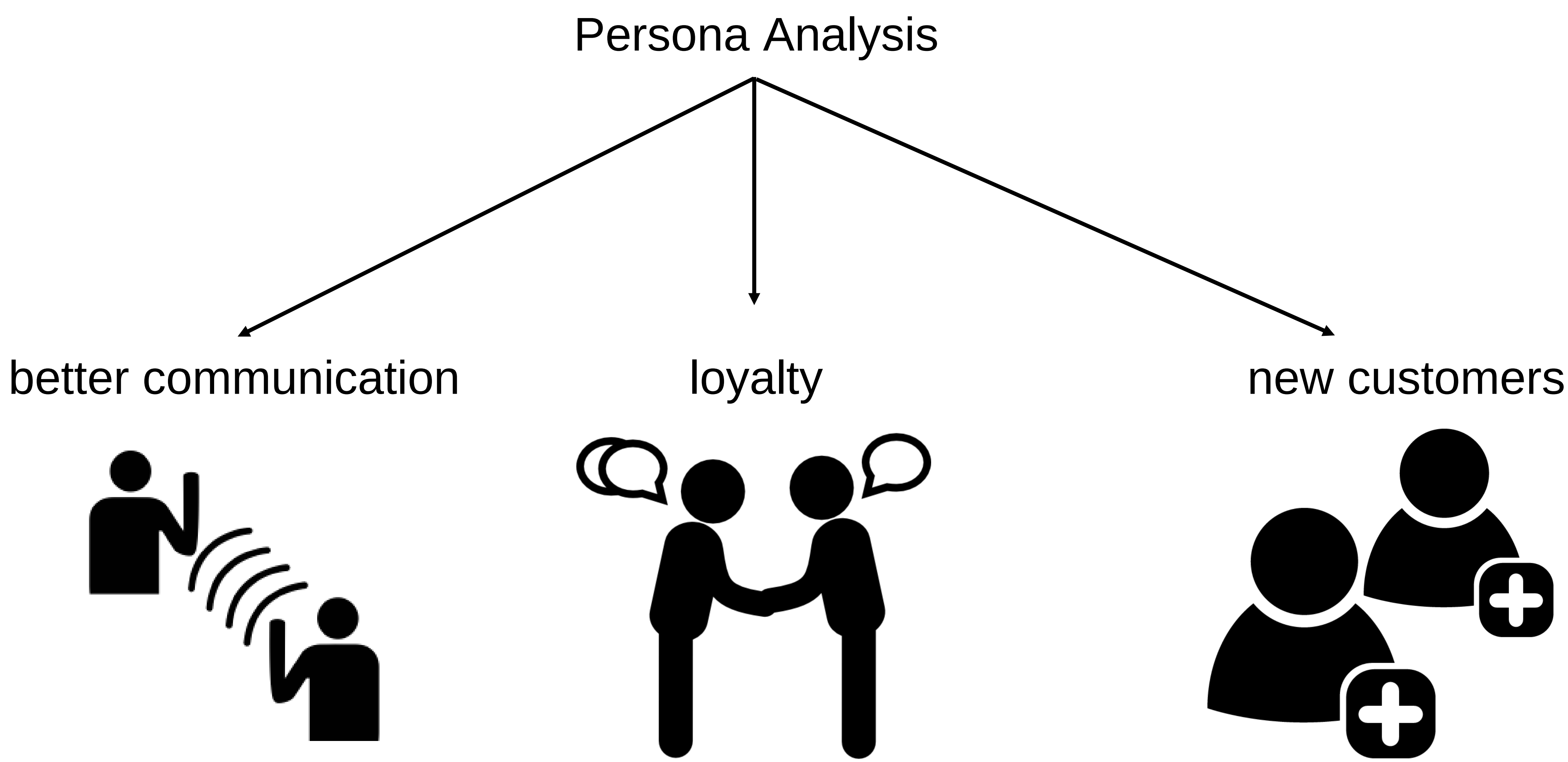
Persona analysis and
Intuition-based

Historical/behavioral-based

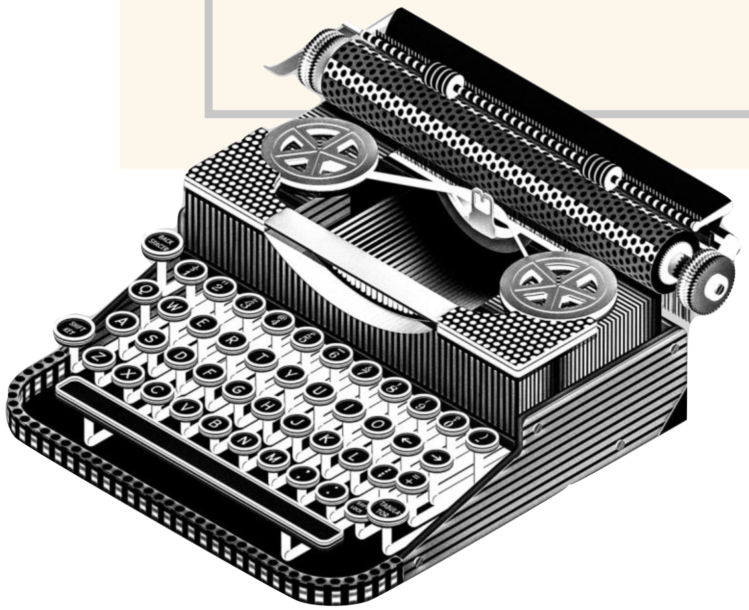


Data-driven

Persona analysis



The most attractive customers are usually those
for which there is a big gap between their needs
and the current satisfaction of these needs



Intuition-based

Demographic

grouping customers
based on their demographic
characteristics

Working men in 30s from
Tartu with children

Attitudinal

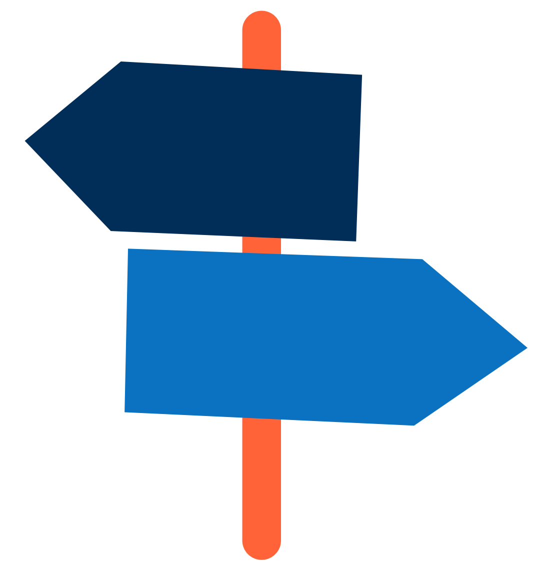
grouping customers
based on their needs

Women who wish to increase
sport activity
and need motivation

Customer segmentation

Intuition-based

Historical/behavioral-based



Data-driven

Behavioral-based

grouping customers based of what they done in the past: purchases,
website browsing, comments

Behavioral-based

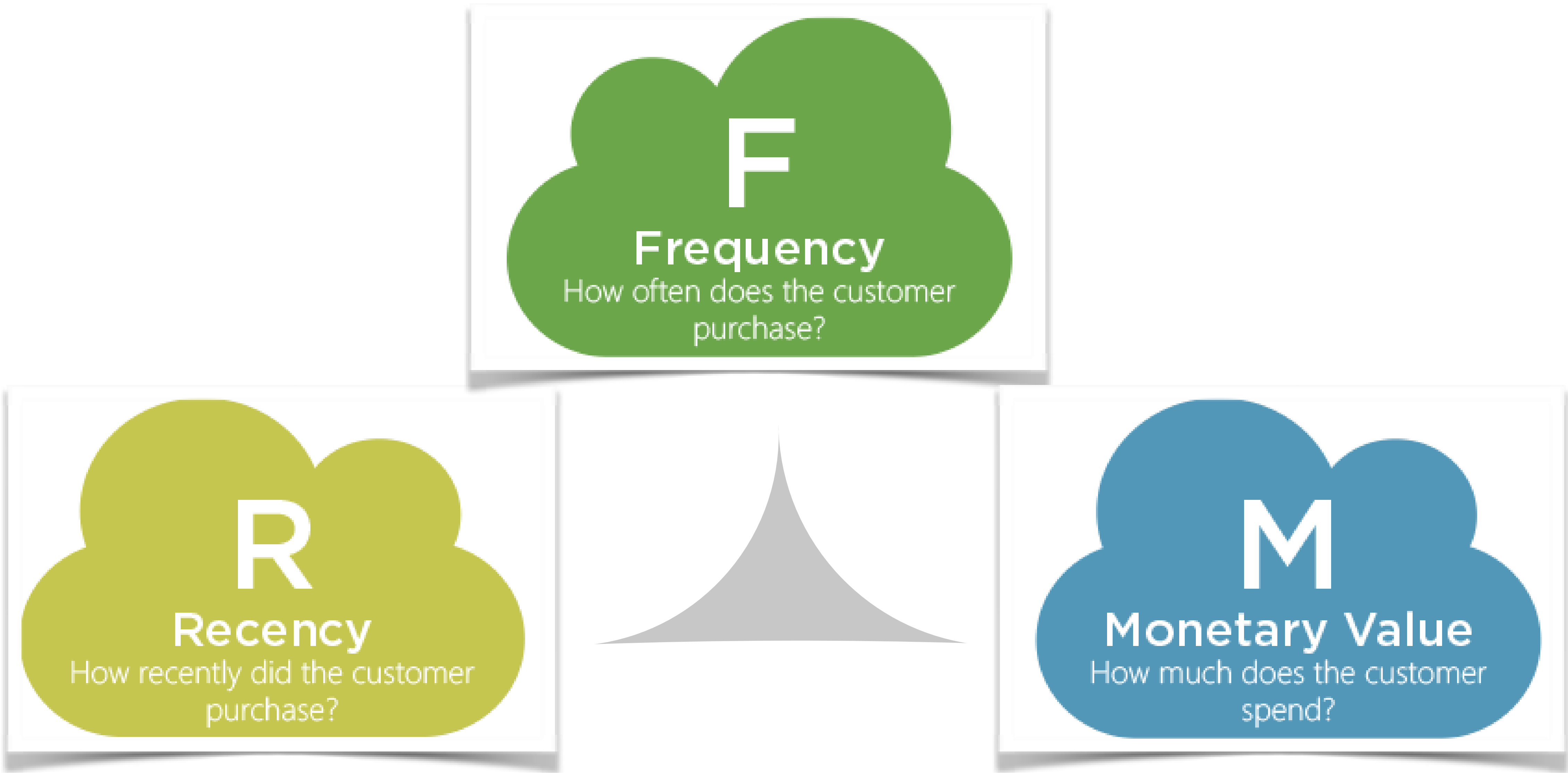
grouping customers based of what they done in the
past: purchases,
website browsing, comments

RFM

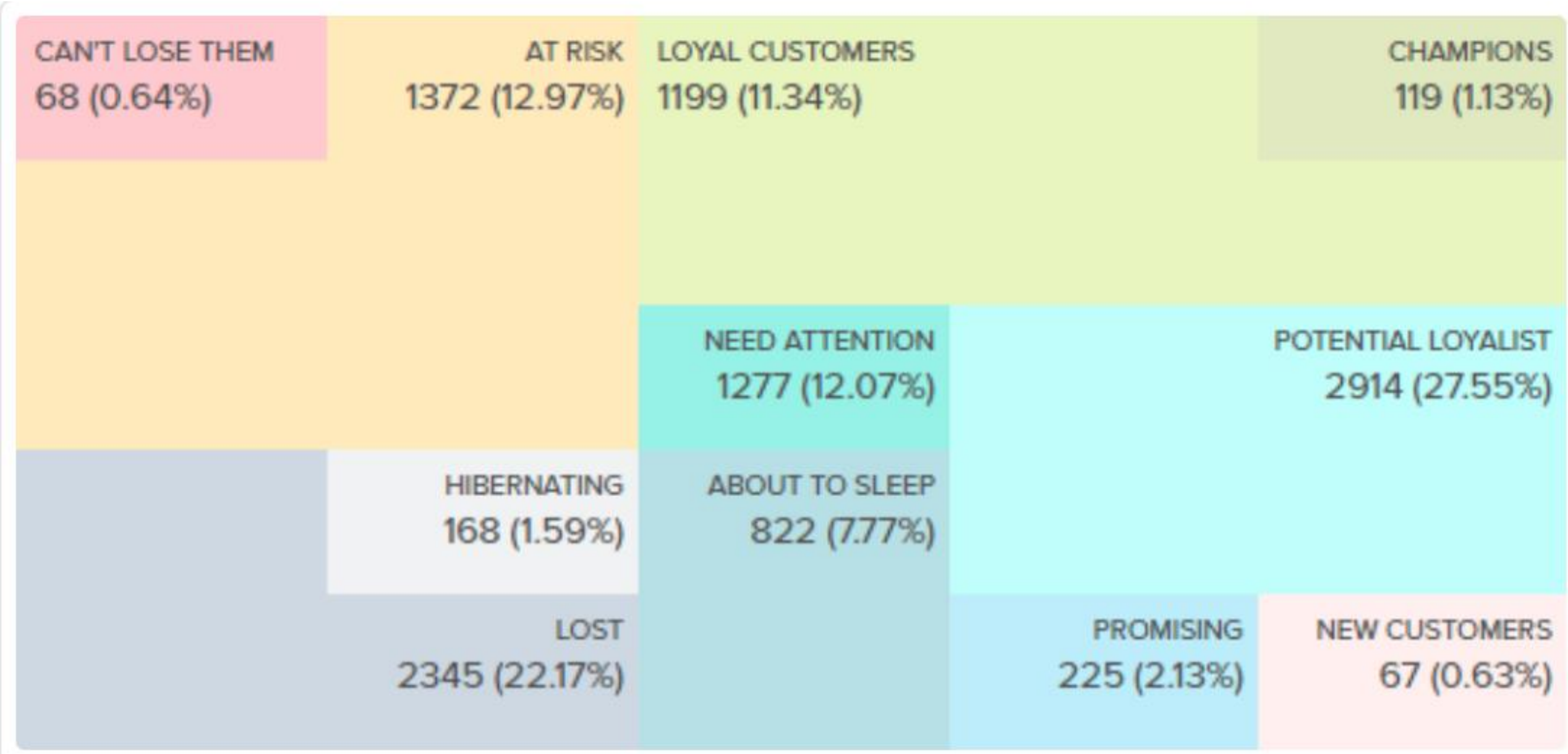
Value tier

Lifecycle stage

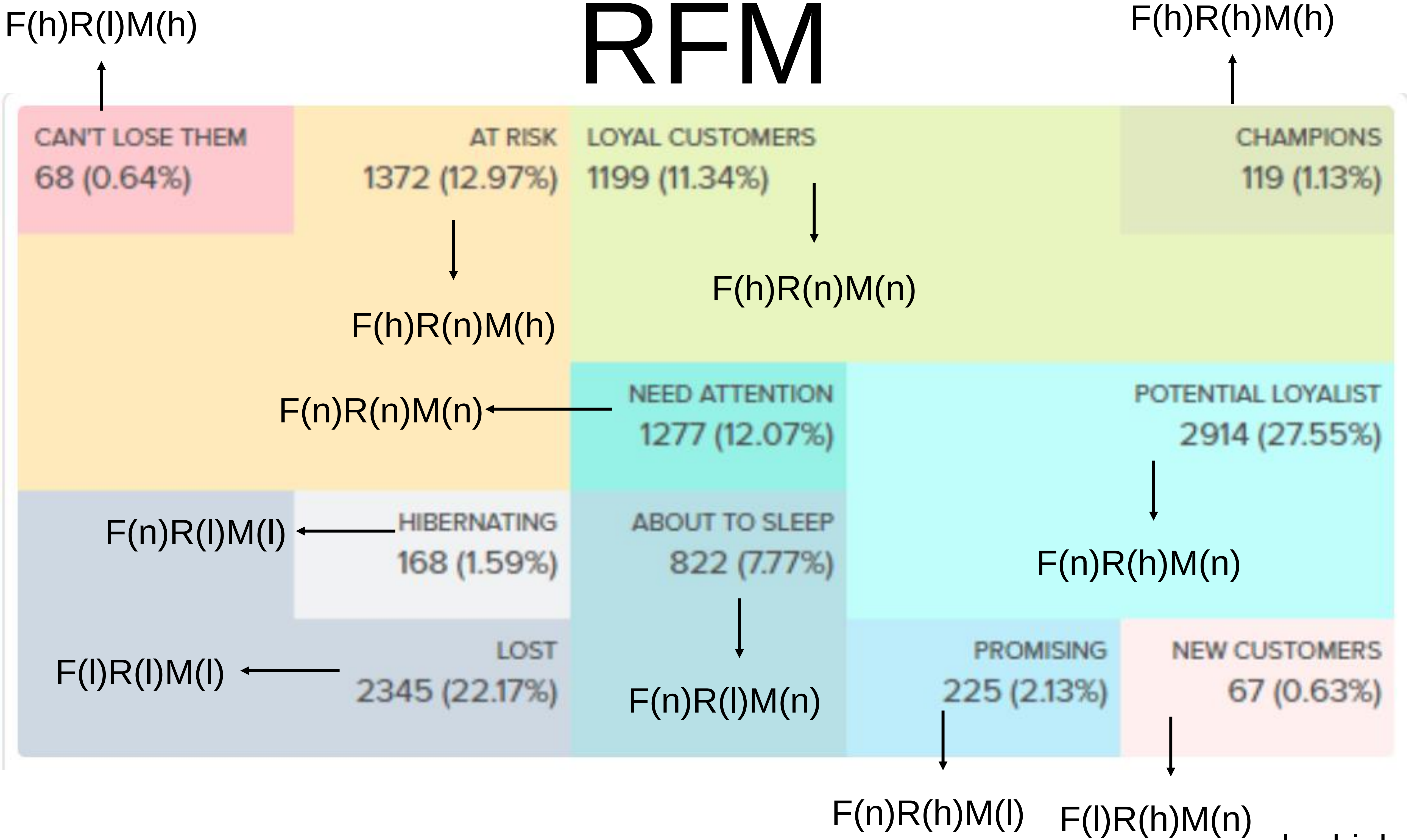
RFM


F
Frequency
How often does the customer purchase?
R
Recency
How recently did the customer purchase?
M
Monetary Value
How much does the customer spend?

RFM


CAN'T LOSE THEM
68 (0.64%)
AT RISK
1372 (12.97%)
LOYAL CUSTOMERS
1199 (11.34%)
CHAMPIONS
119 (1.13%)
NEED ATTENTION
1277 (12.07%)
POTENTIAL LOYALIST
2914 (27.55%)
HIBERNATING
168 (1.59%)
ABOUT TO SLEEP
822 (7.77%)
LOST
2345 (22.17%)
PROMISING
225 (2.13%)
NEW CUSTOMERS
67 (0.63%)

RFM

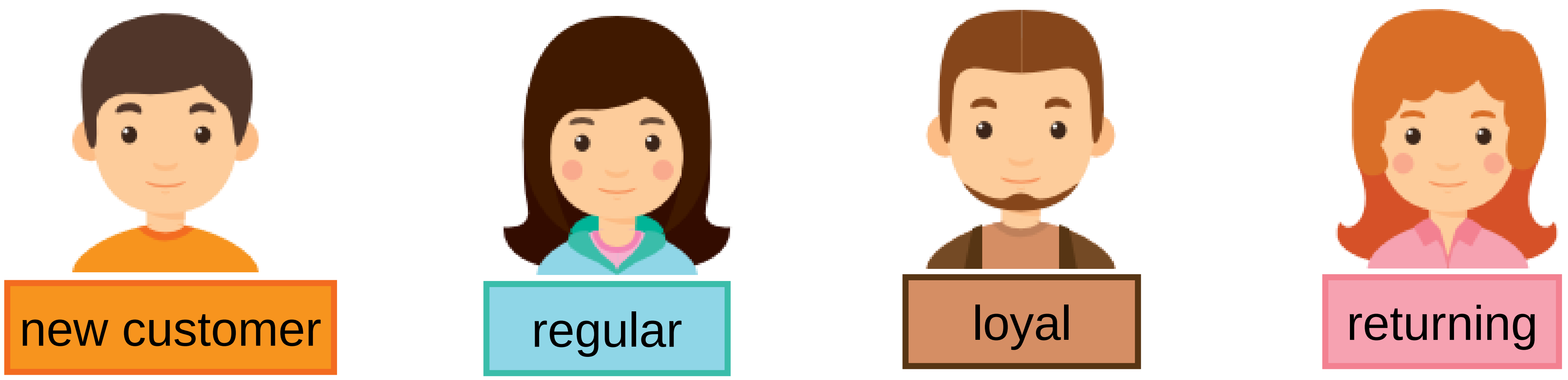

F(h)R(l)M(h)
F(h)R(h)M(h)
CAN'T LOSE THEM
68 (0.64%)
AT RISK
1372 (12.97%)
LOYAL CUSTOMERS
1199 (11.34%)
CHAMPIONS
119 (1.13%)
F(h)R(n)M(n)
F(h)R(n)M(h)
NEED ATTENTION
1277 (12.07%)
F(n)R(n)M(n)
POTENTIAL LOYALIST
2914 (27.55%)
HIBERNATING
168 (1.59%)
F(n)R(l)M(l)
ABOUT TO SLEEP
822 (7.77%)
F(n)R(h)M(n)
LOST
2345 (22.17%)
F(l)R(l)M(l)
F(n)R(l)M(n)
PROMISING
225 (2.13%)
NEW CUSTOMERS
67 (0.63%)
F(n)R(h)M(l)
F(l)R(h)M(n)
h - high
n - neutral
l - low

Value tier

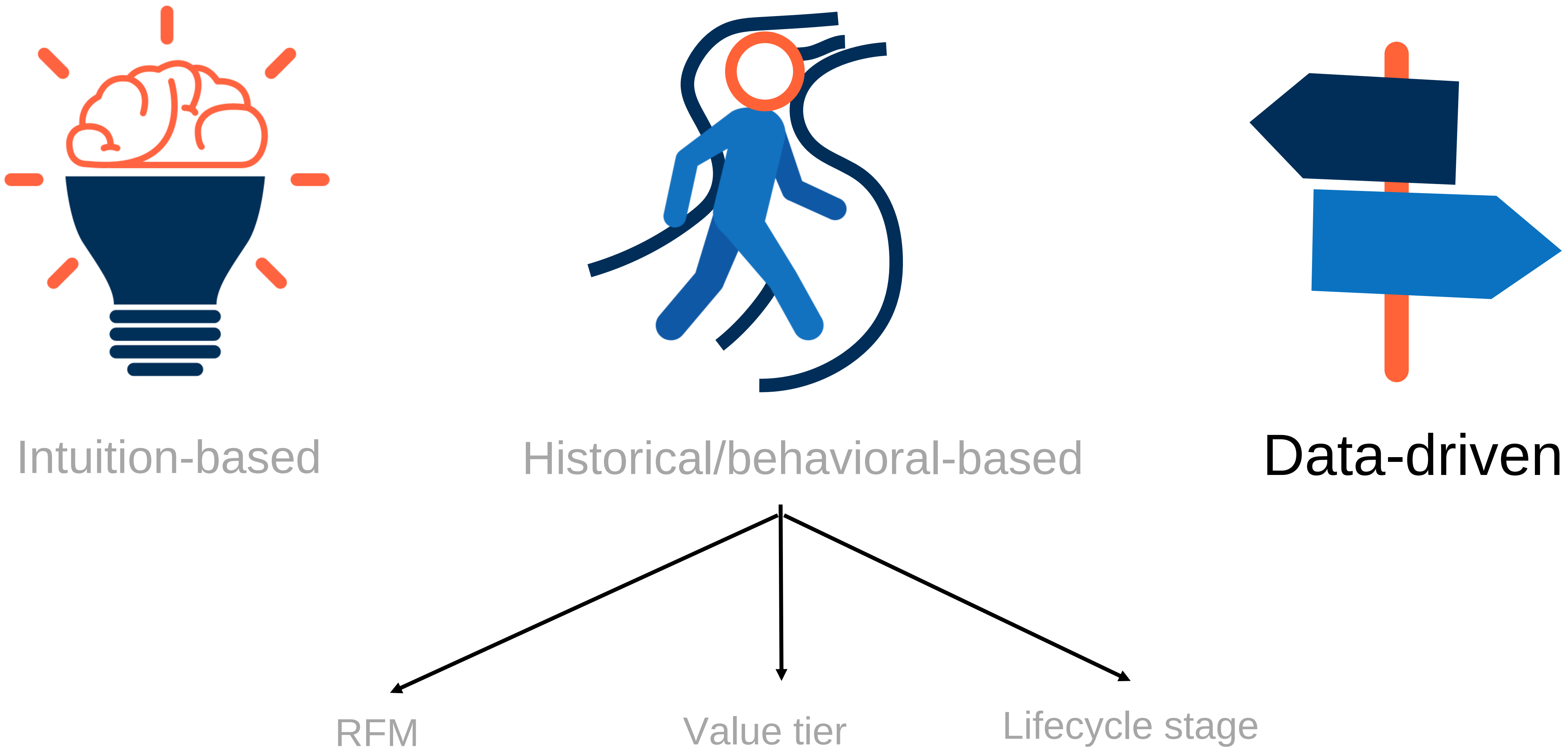
Grouping customers based on the value they deliver to your business. Top 1%, top 5 % etc of generated revenue.

Lifecycle stage

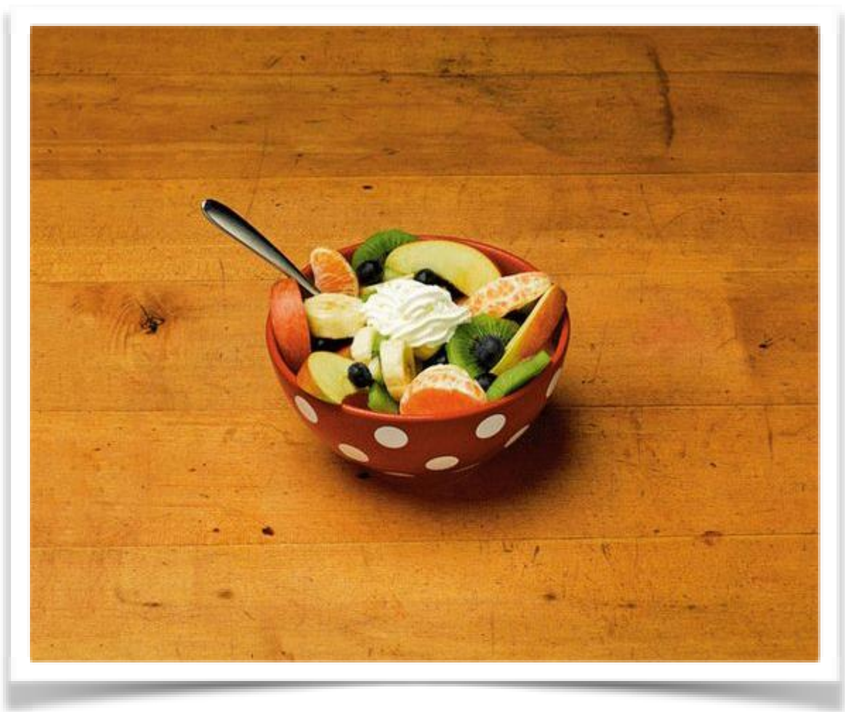
Grouping customers based on the type of relationships with the company/brand

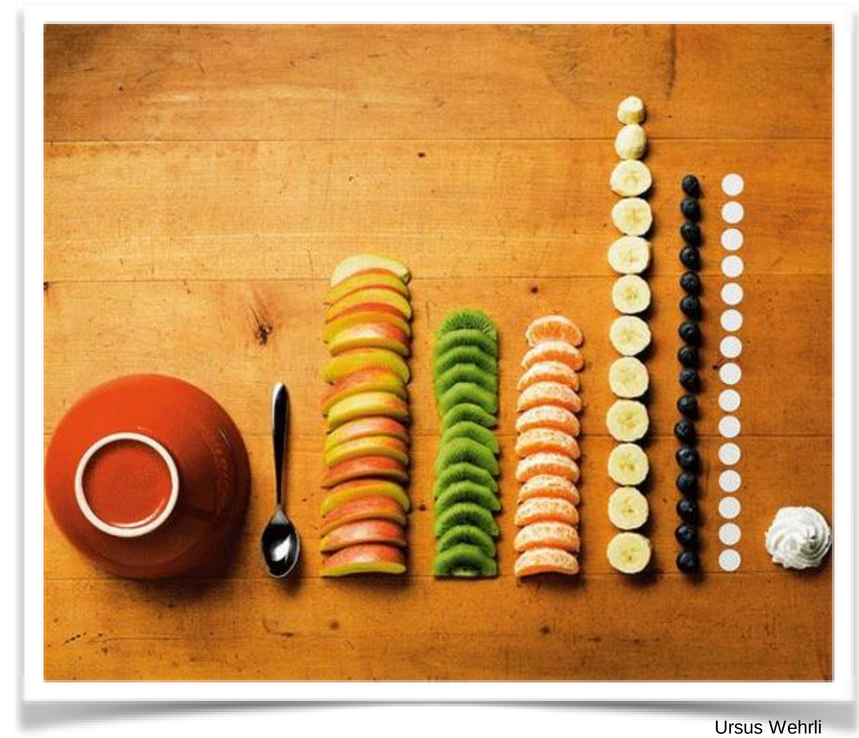


Customer segmentation



Data-driven segmentation



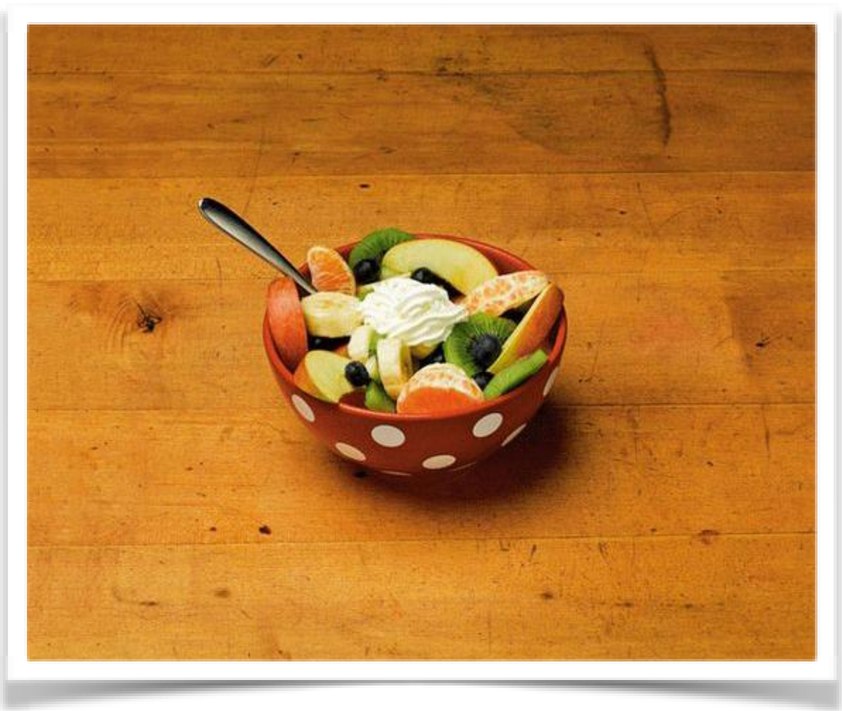


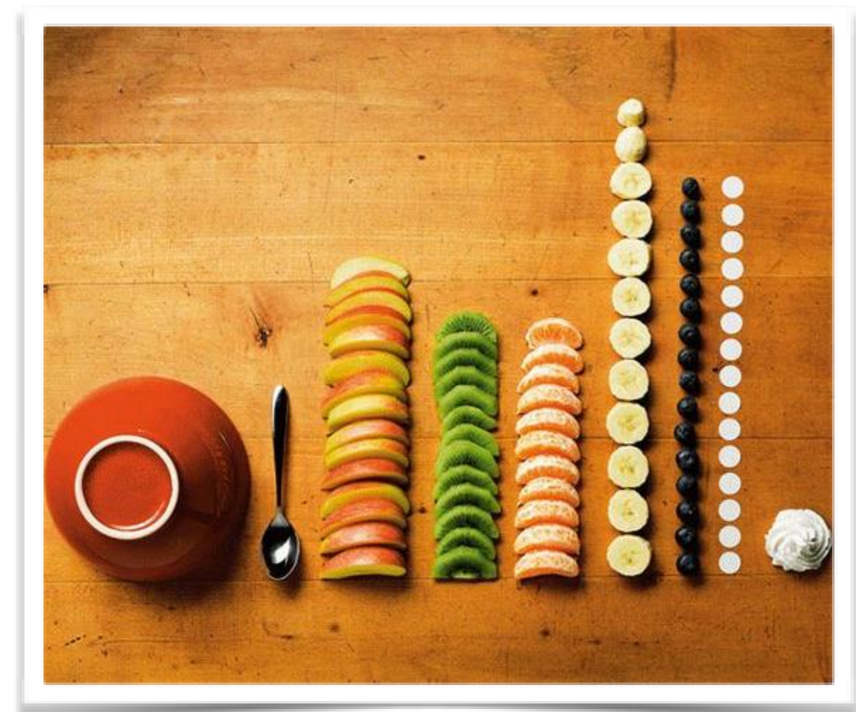
Ursus Wehrli

Automated discovery of the new segments

Depends on your data

Data-driven segmentation





Ursus Wehrli

Automated discovery of the new segments

email responses

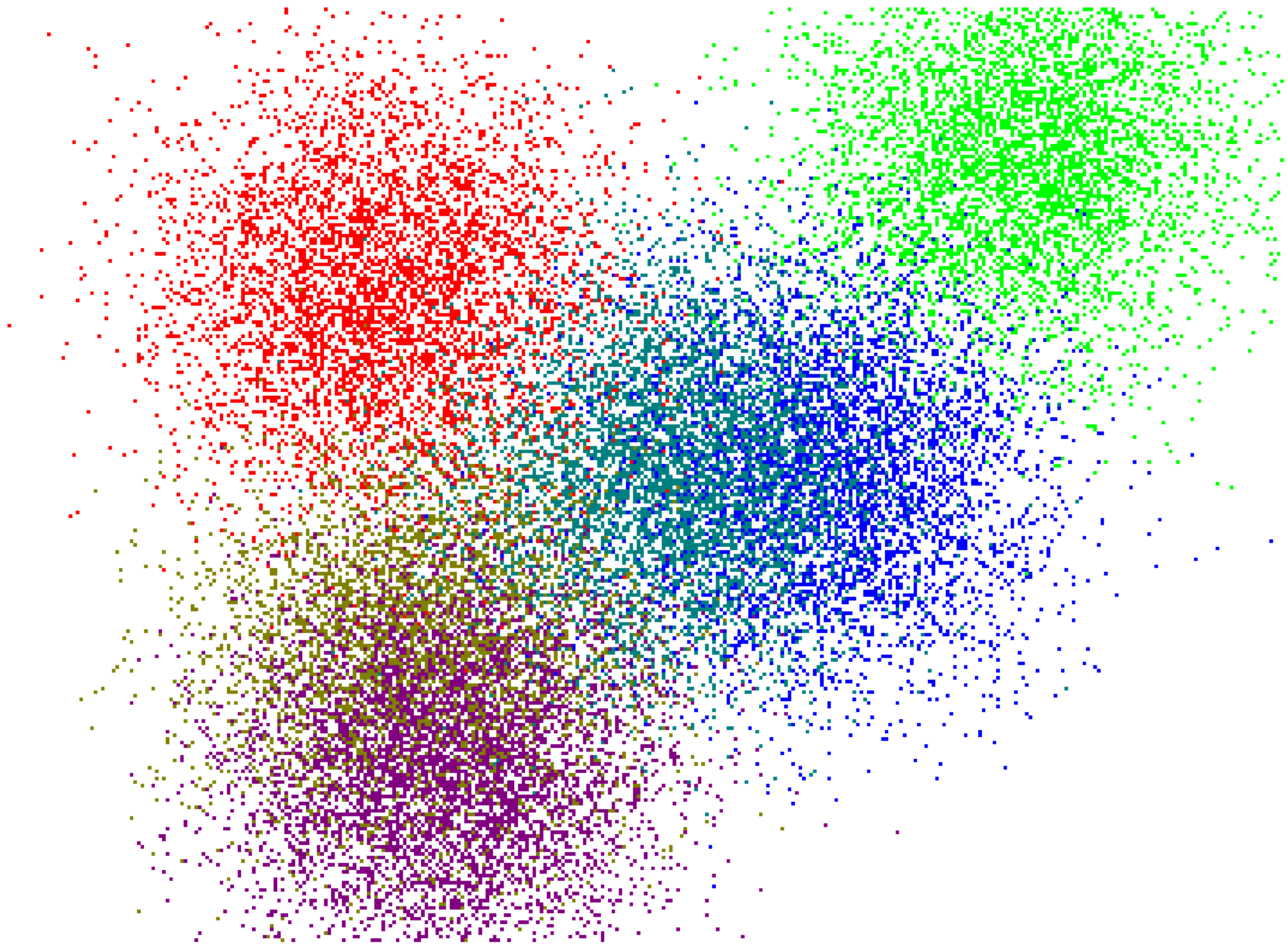
Depends on your data

Product affinity

Price sensitivity

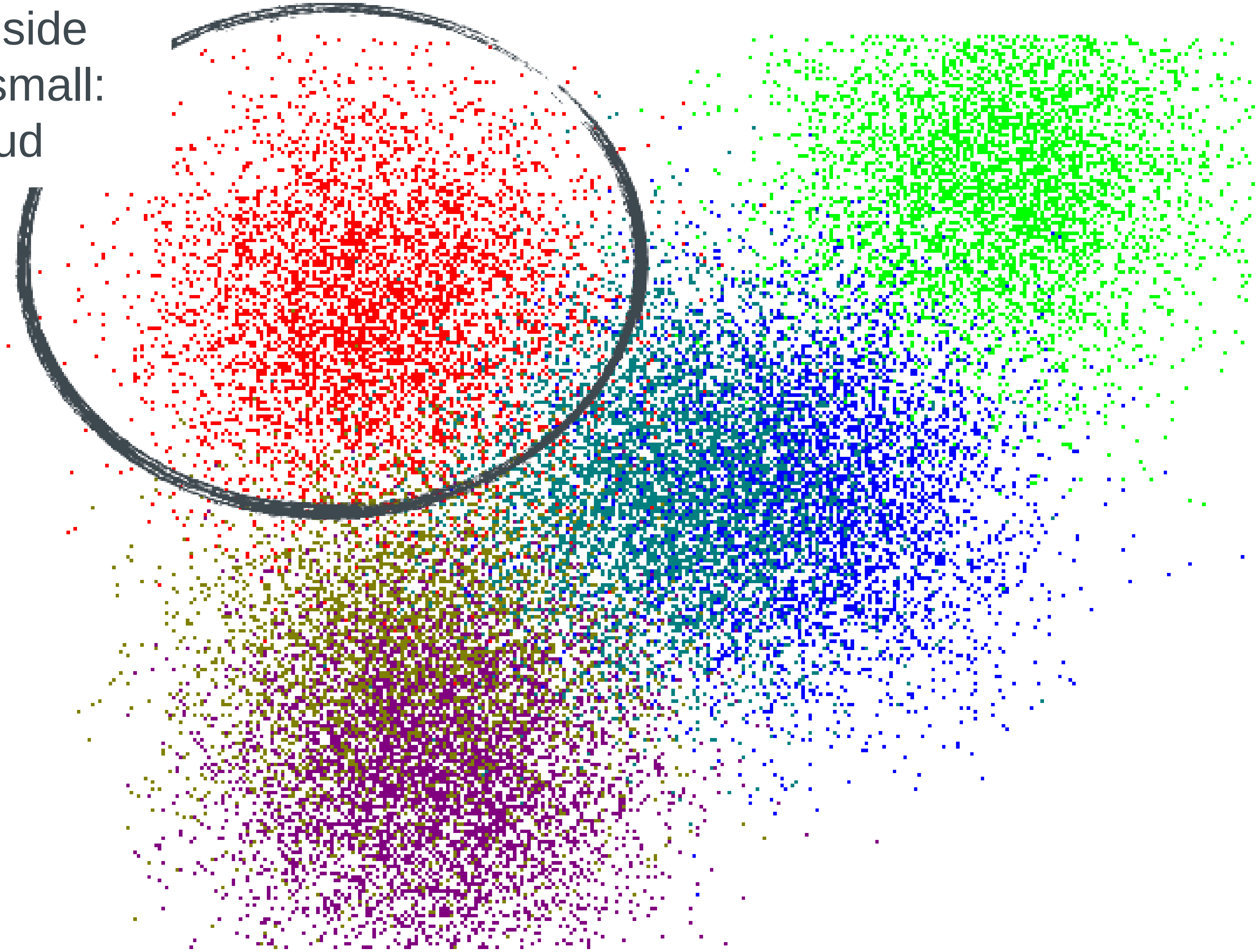
Promotion sensitivity

Automated segmentation

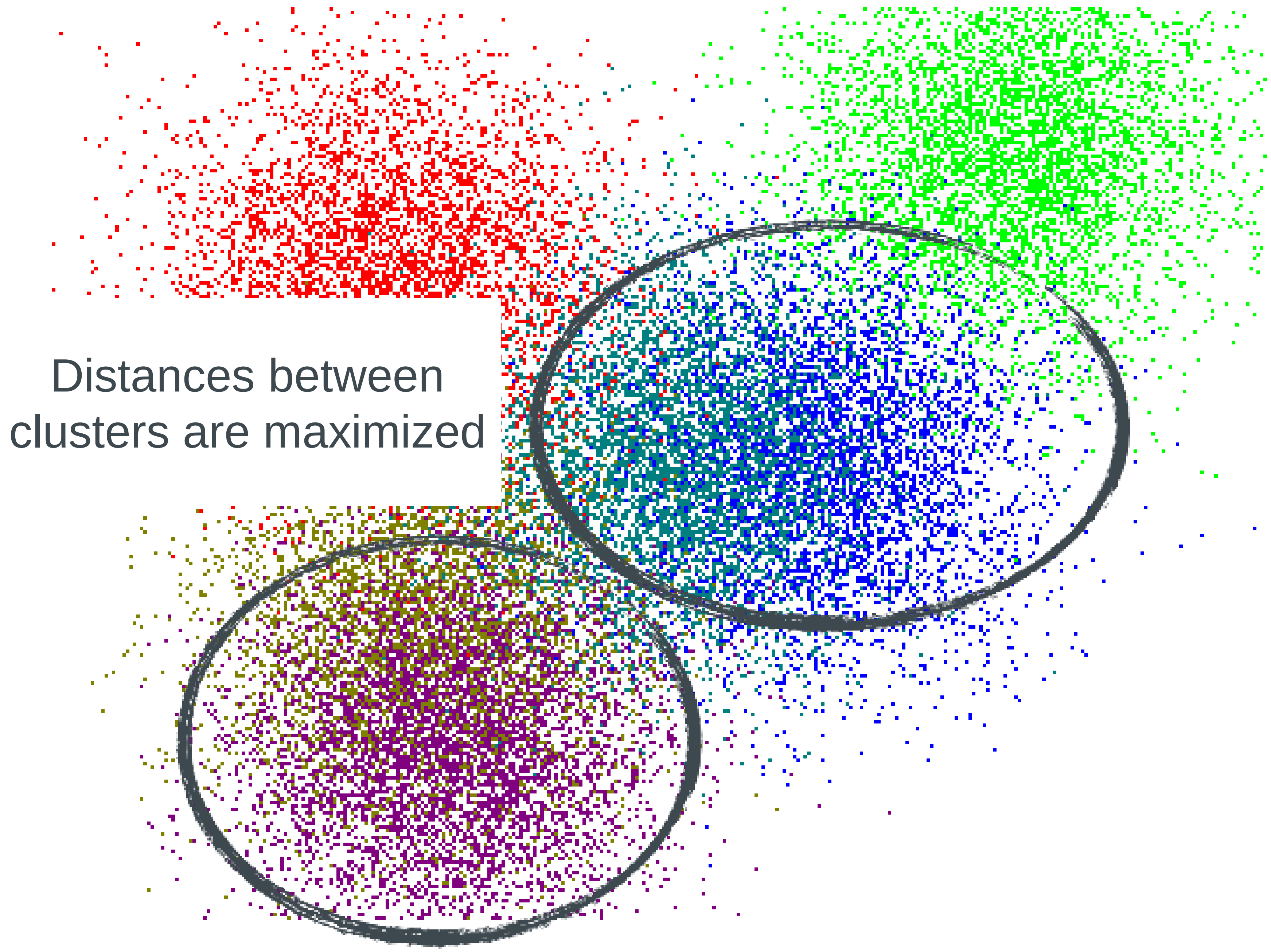


Automated segmentation

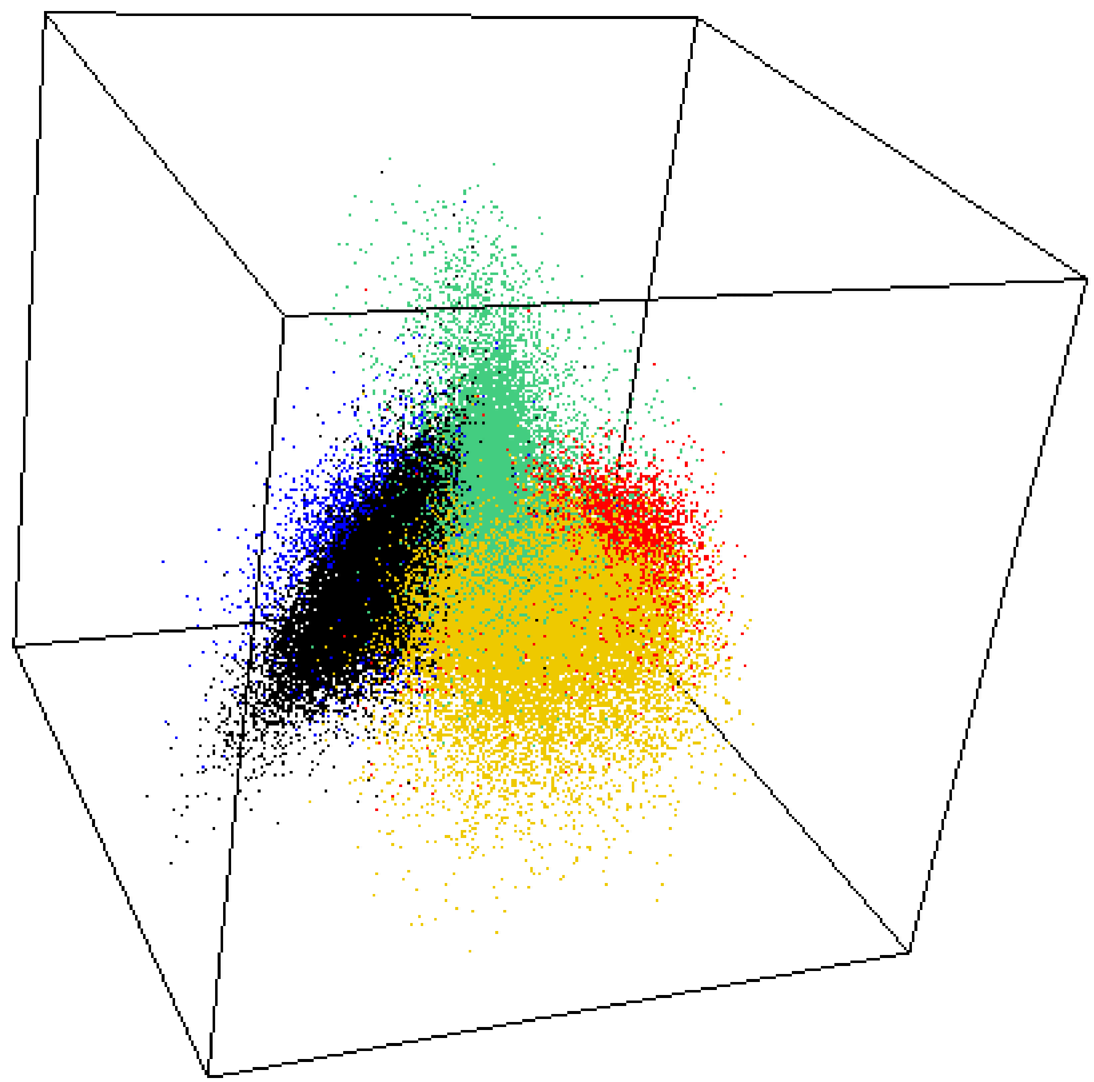
Distances inside clusters are small: dense cloud



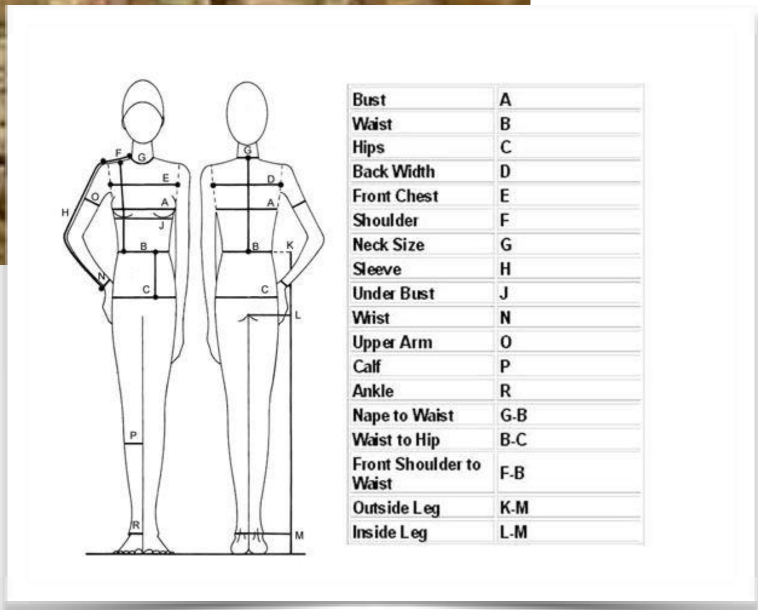
Automated segmentation

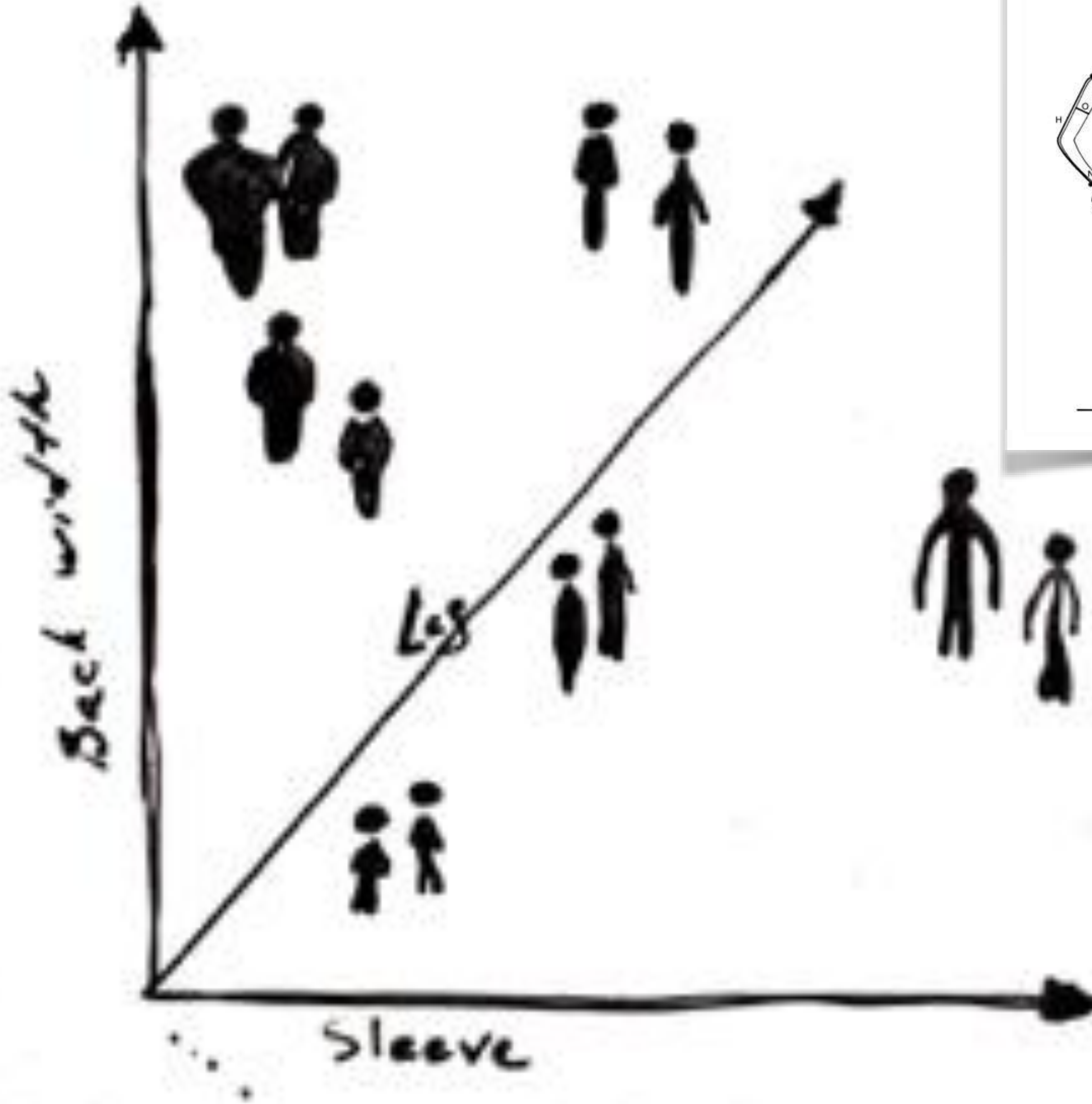


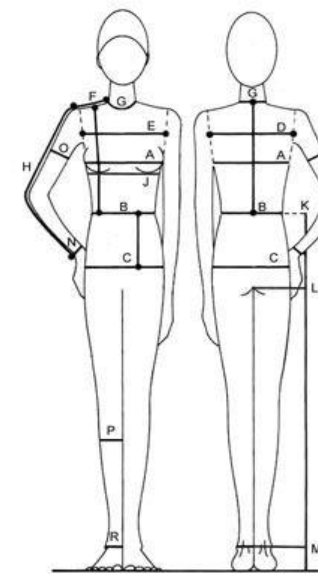
…in multidimensional space



Clustering in U.S Army







Bust	A
Waist	B
Hips	C
Back Width	D
Front Chest	E
Shoulder	F
Neck Size	G
Sleeve	H
Under Bust	J
Wrist	N
Upper Arm	O
Calf	P
Ankle	R
Nape to Waist	G-B
Waist to Hip	B-C
Front Shoulder to Waist	F-B
Outside Leg	K-M
Inside Leg	L-M

3 Important Questions

1. How do you represent a cluster of more than one point ?

2. How do you determine the “nearness” of clusters?

3. When to stop combining clusters

Source: Stanford University – Hierarchical clustering

3 Important Questions

1. How do you represent a cluster of more than one point?

   centroid or clusteroid represents a set of points

2. How do you determine the “nearness” of clusters?

   Some Distance metric

3. When to stop combining clusters

   convergence or some threshold or cohesion

K-means clustering

K-means

fixed number of clusters -
you need to choose it
yourself

K-means

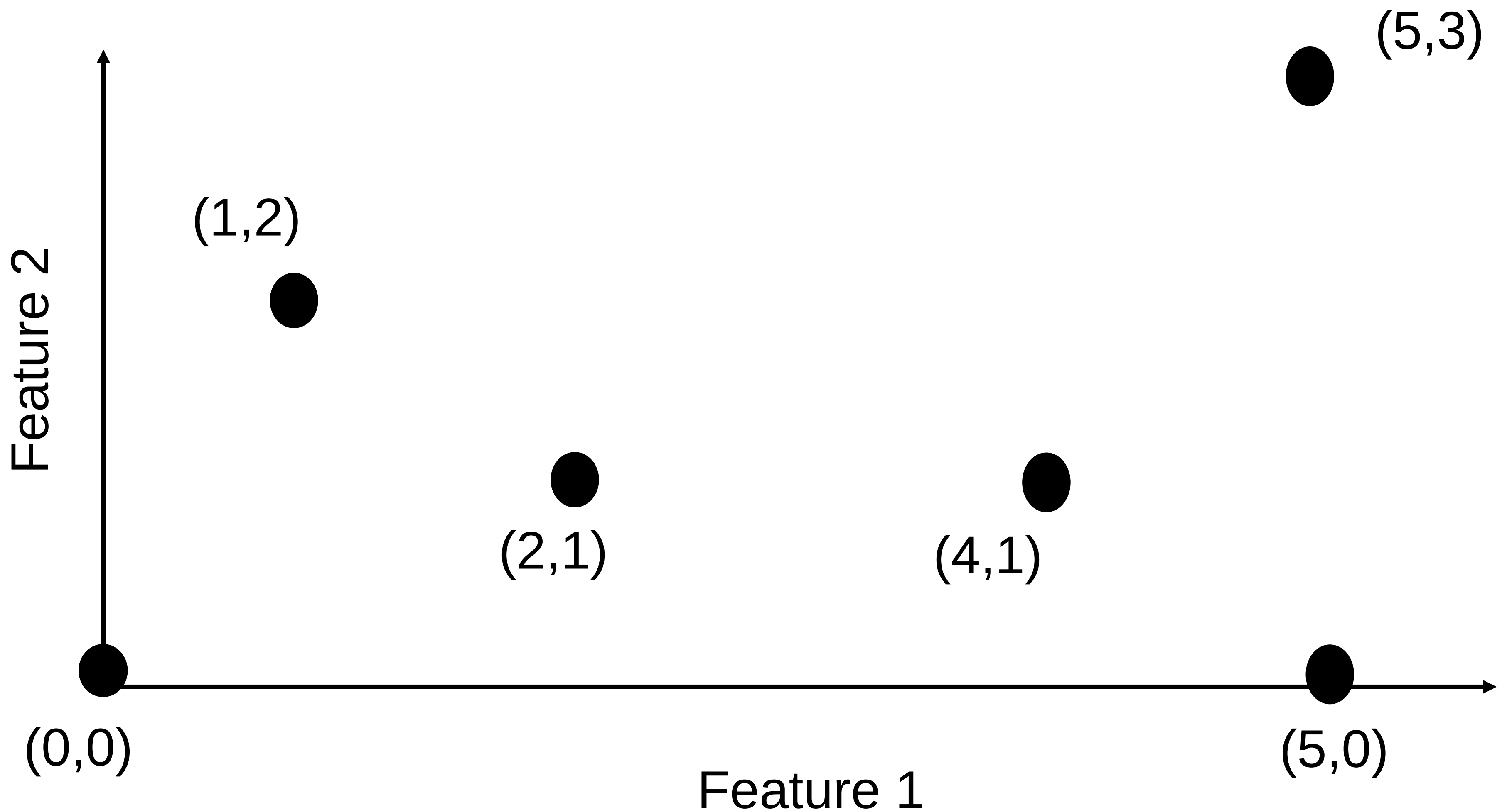
fixed number of clusters - you need to choose it yourself

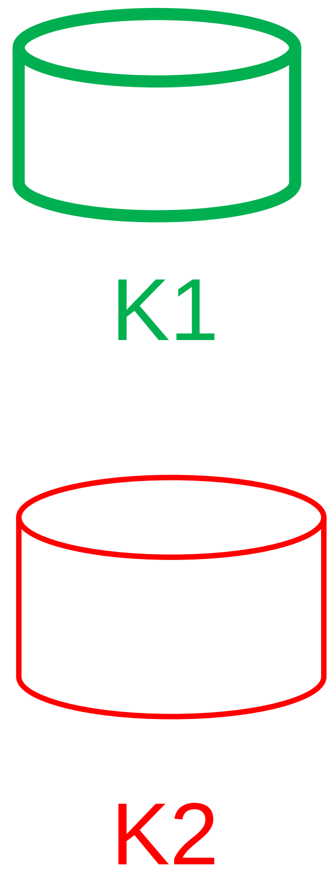
based on the calculation of averages

K-means Clustering

● Datapoint
X Centroid

K = 2 (we would like to create 2 clusters)



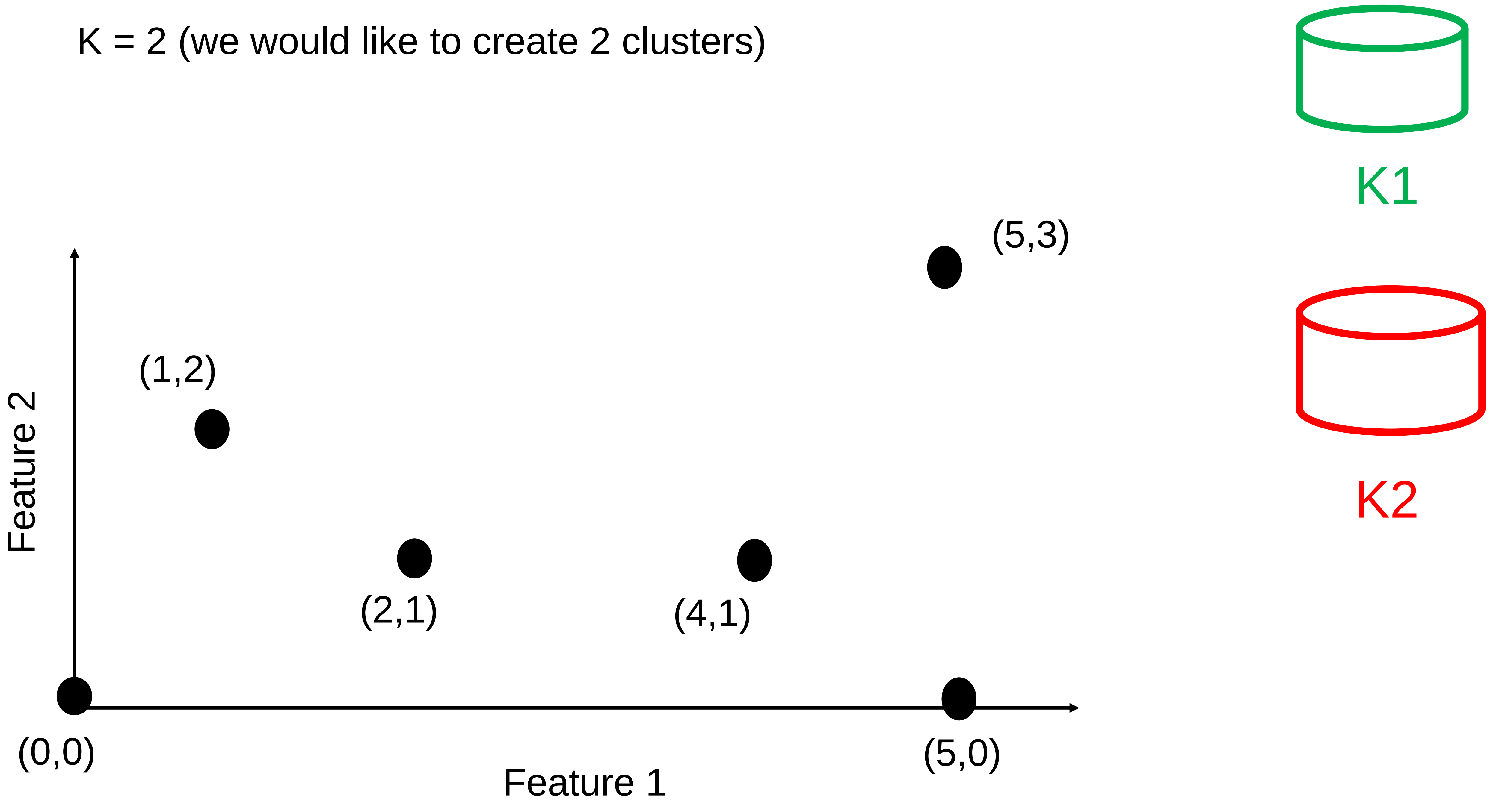


K-means Clustering

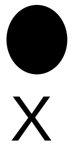
Each data point is a cluster. So, we have 6 clusters but we would like to create only 2 clusters (as K=2).

● Datapoint
X Centroid

K = 2 (we would like to create 2 clusters)

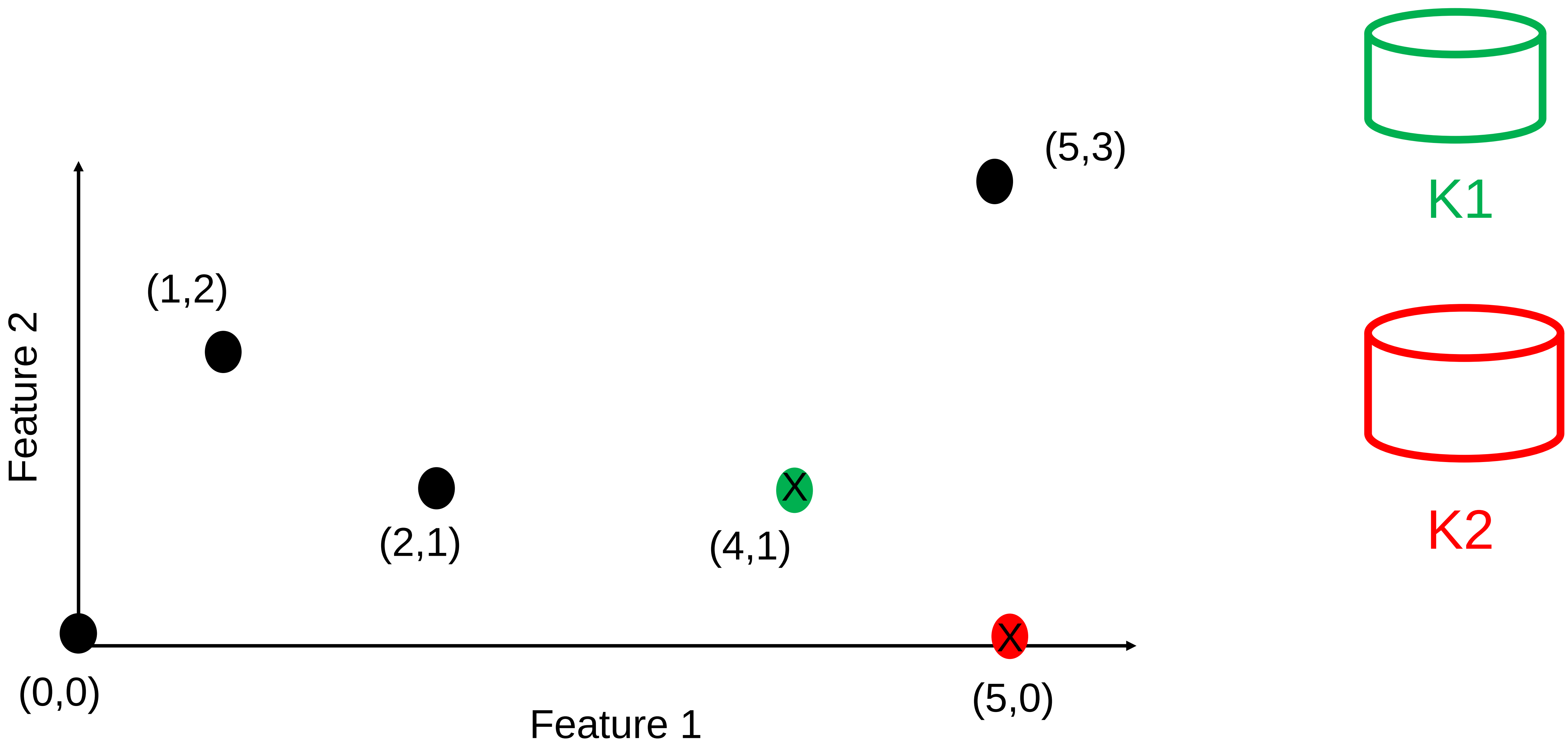


K-means Clustering



Step 1: Select the initiators

We have selected 2 initiators **randomly** (red and green points)



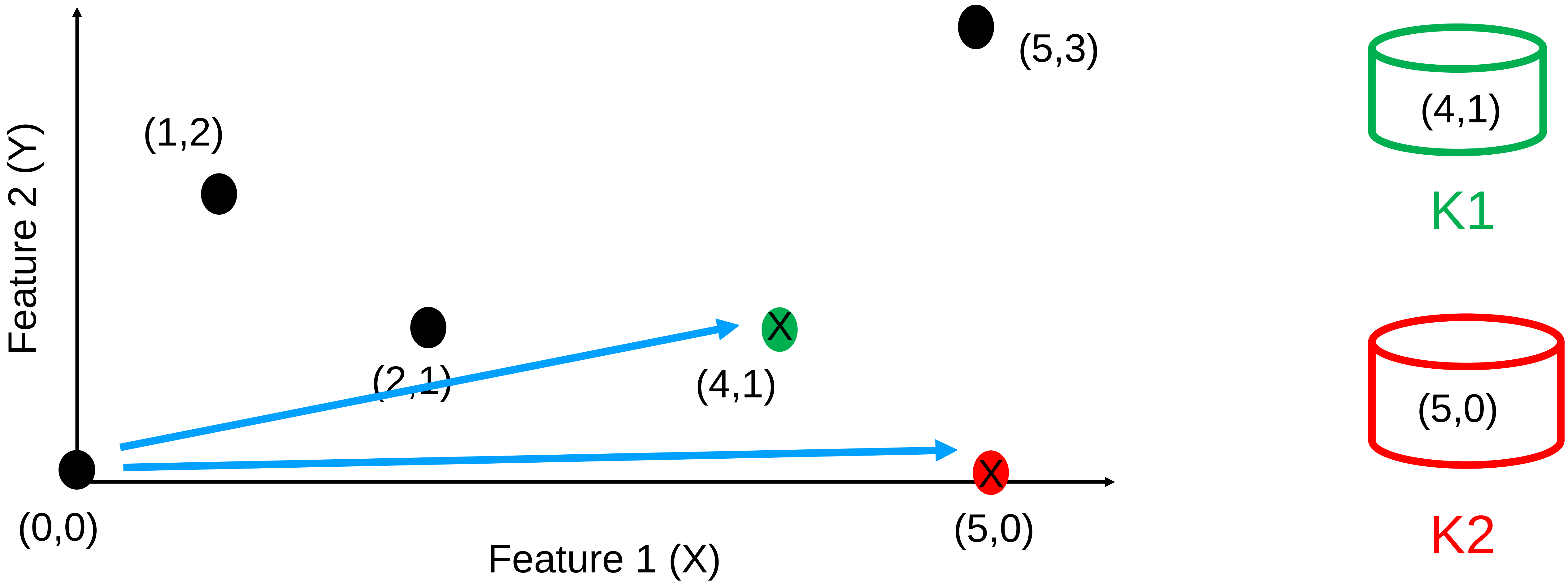
K-means Clustering

● Datapoint
X Centroid

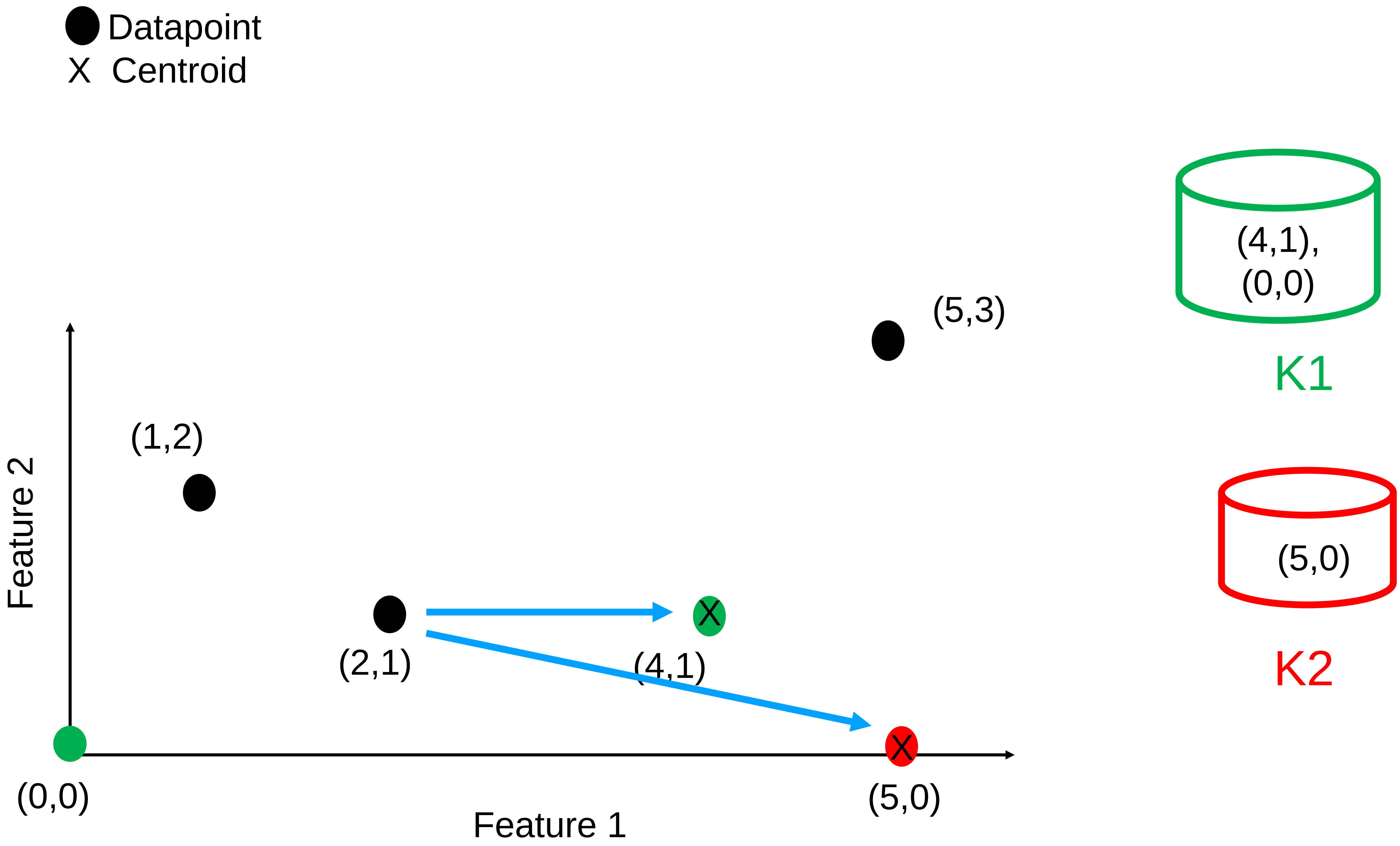
Step 2: Calculate the distance (Euclidean) of every point to all the initiators

$$d(a, b) = \sqrt{\sum_{i=1}^{n} (i_a - i_b)^2}$$

$$d((0,0), (4,1)) = \boxed{\sqrt{(x_4 - x_1)^2 + (y_4 - y_1)^2}}$$

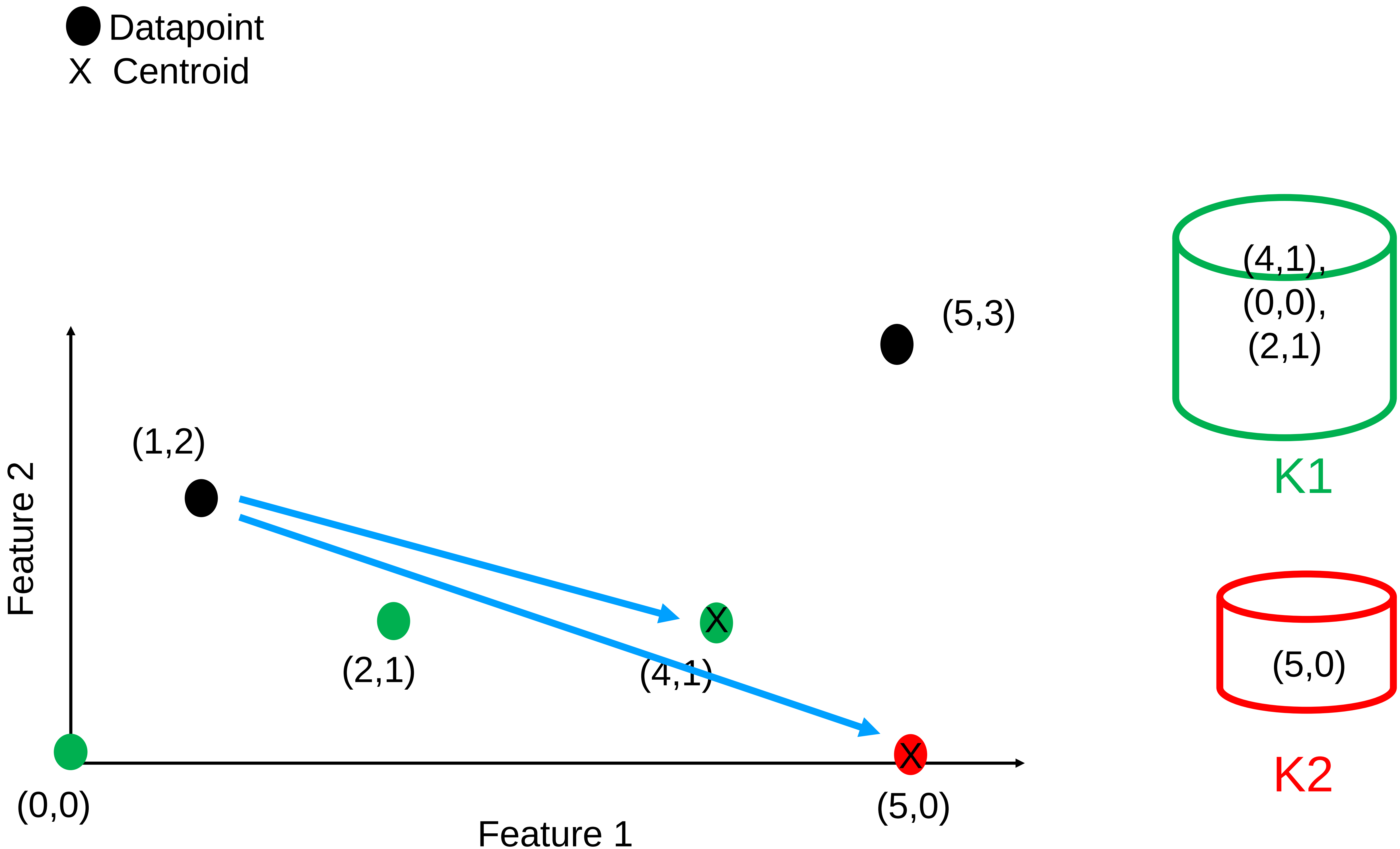


K-means Clustering



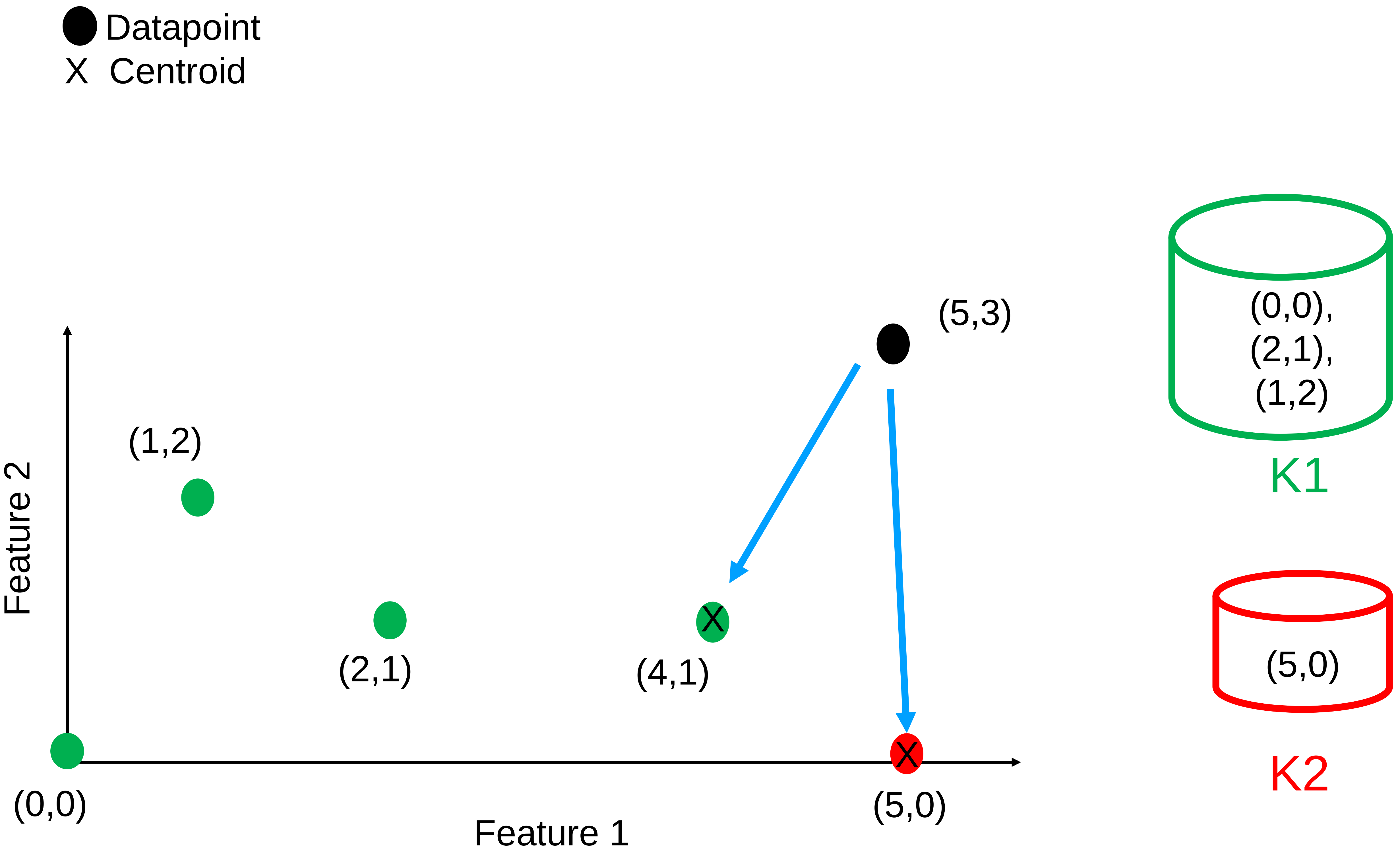
Source: Stanford – Hierarchical clustering

K-means Clustering

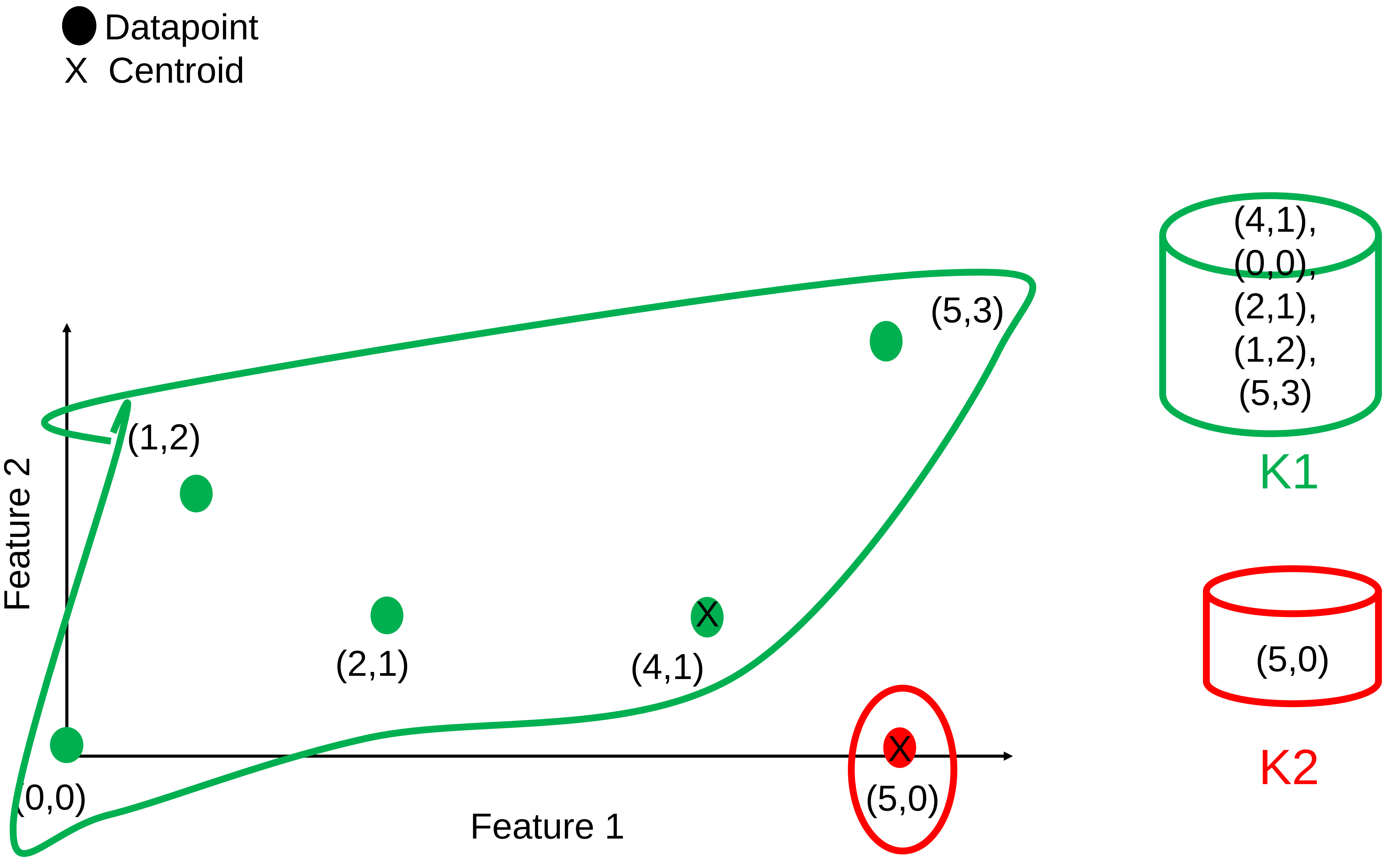


Source: Stanford – Hierarchical clustering

K-means Clustering



K-means Clustering

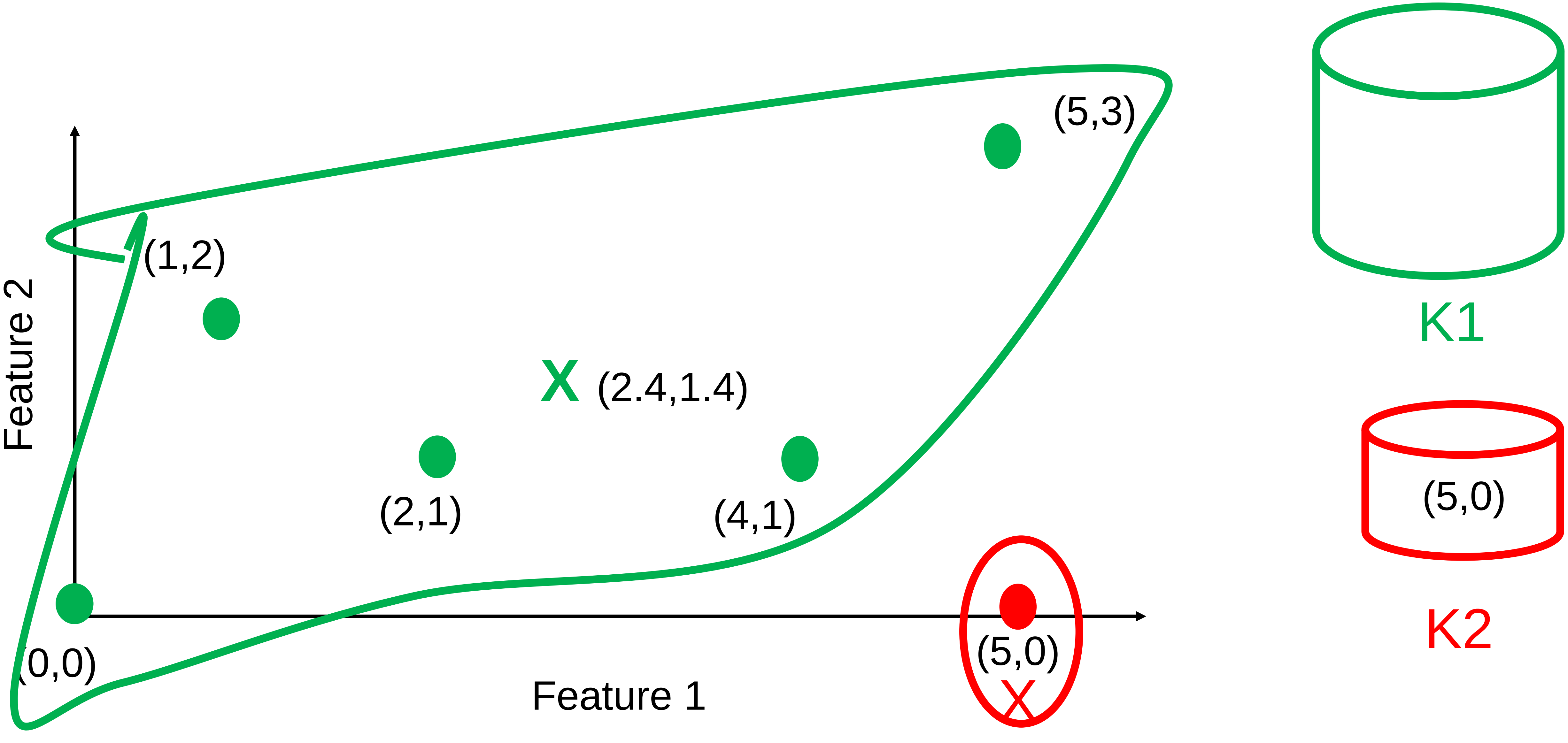


K-means Clustering

● Datapoint
X Centroid

End of Iteration 1: Calculate new centroids

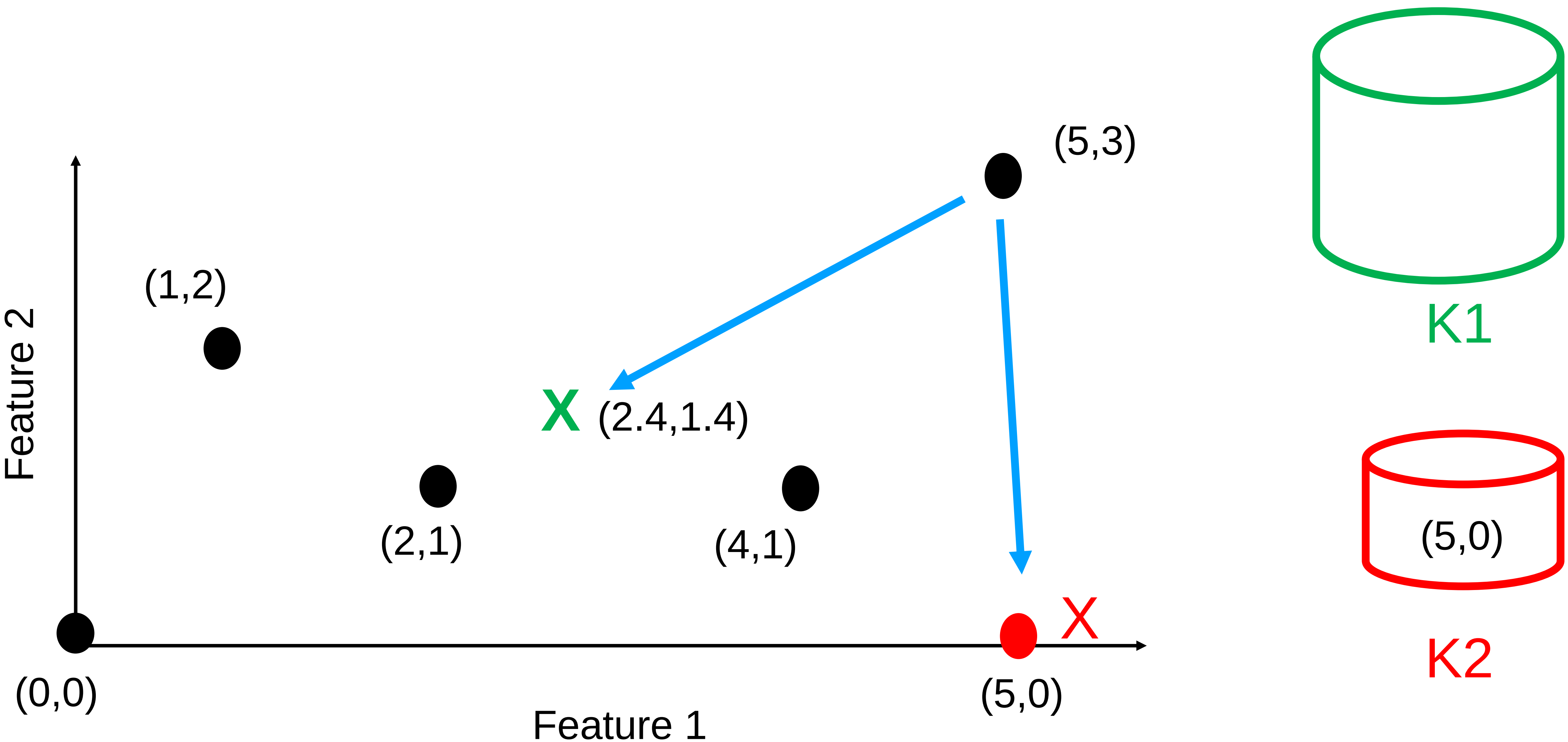
$$C = \left(\frac{1}{n}\sum_{i=1}^{i=n} x_i,\ \frac{1}{n}\sum_{i=1}^{i=n} y_i \right)$$



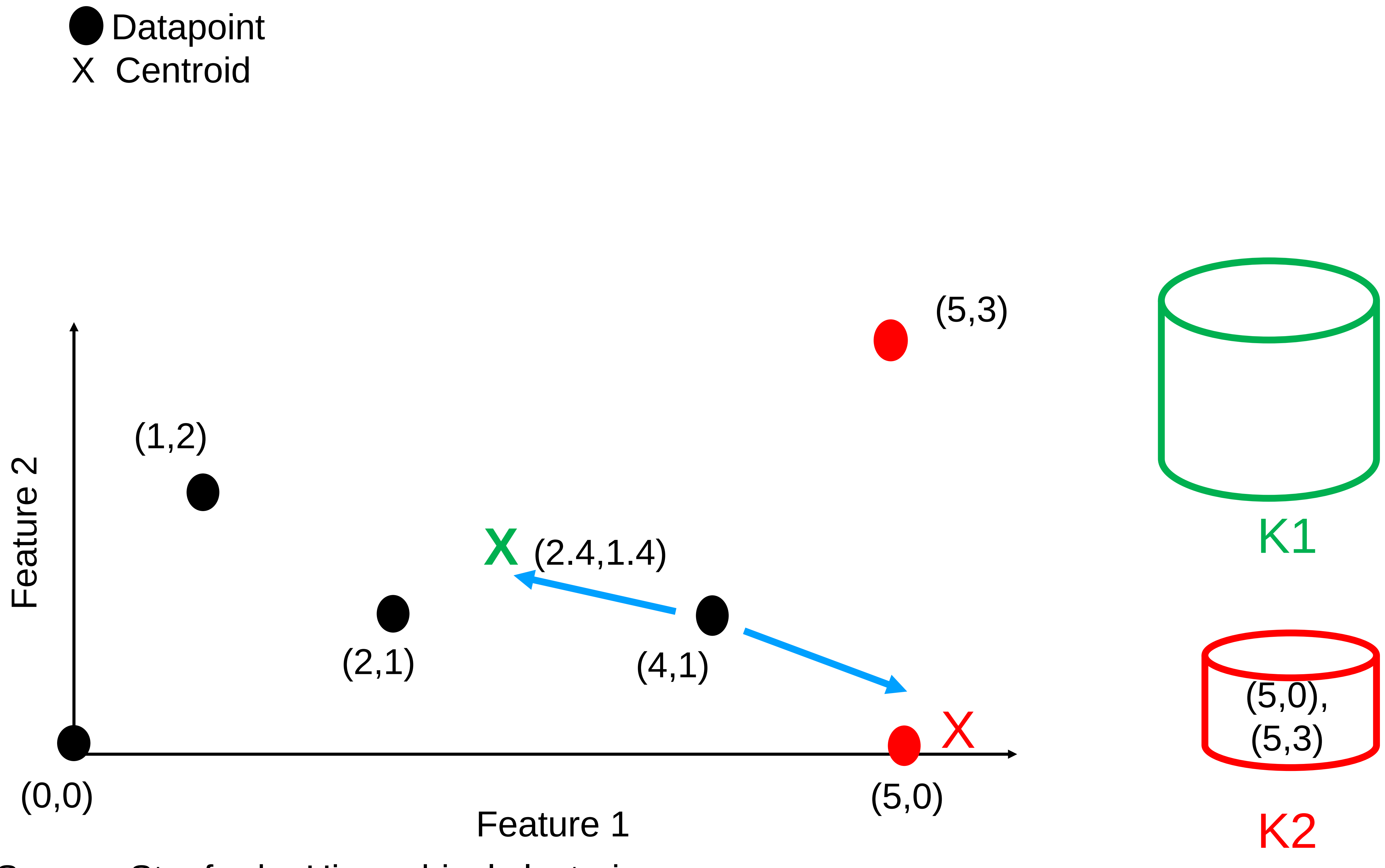
K-means Clustering

● Datapoint
X Centroid

Iteration 2

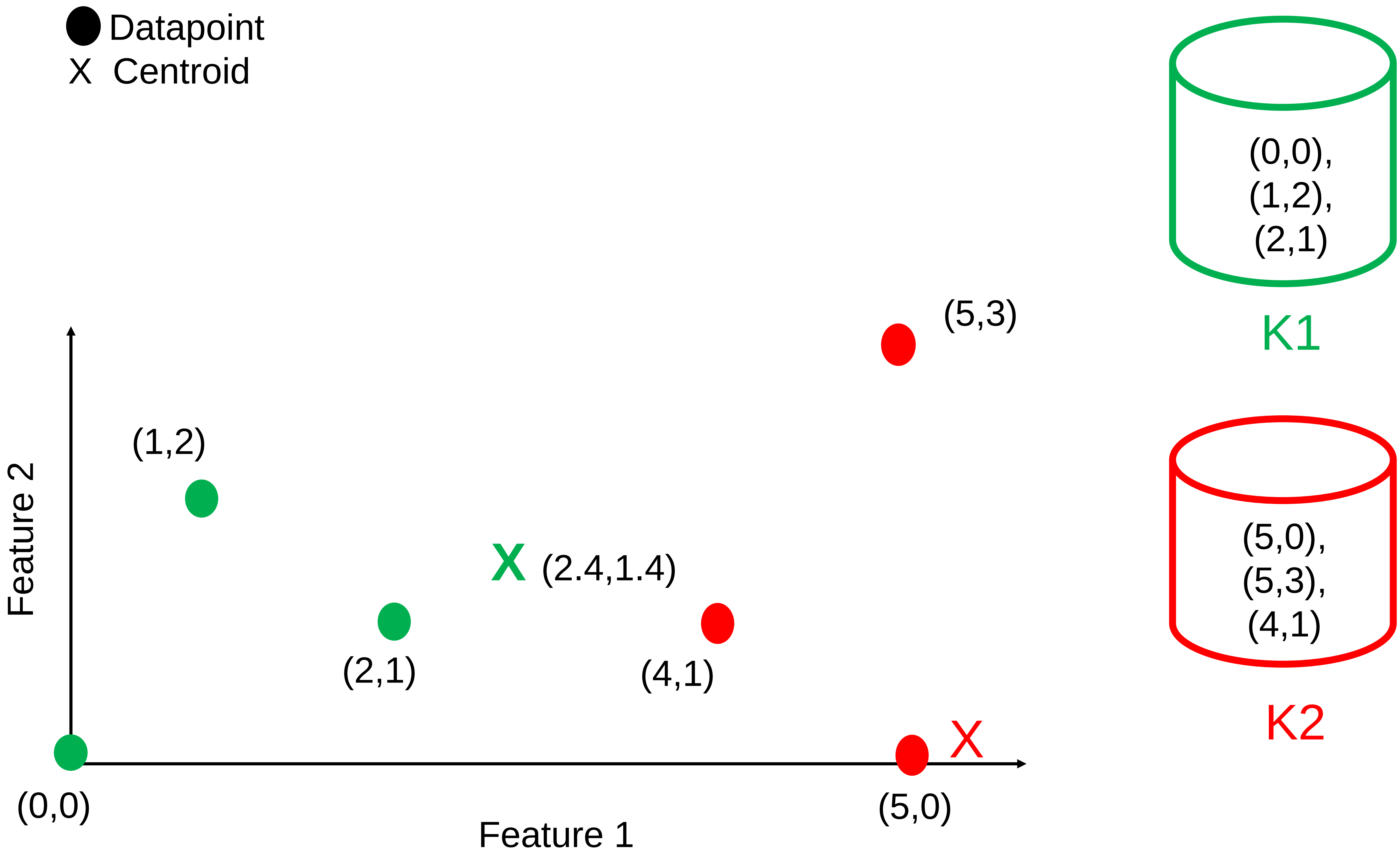


K-means Clustering



Source: Stanford – Hierarchical clustering

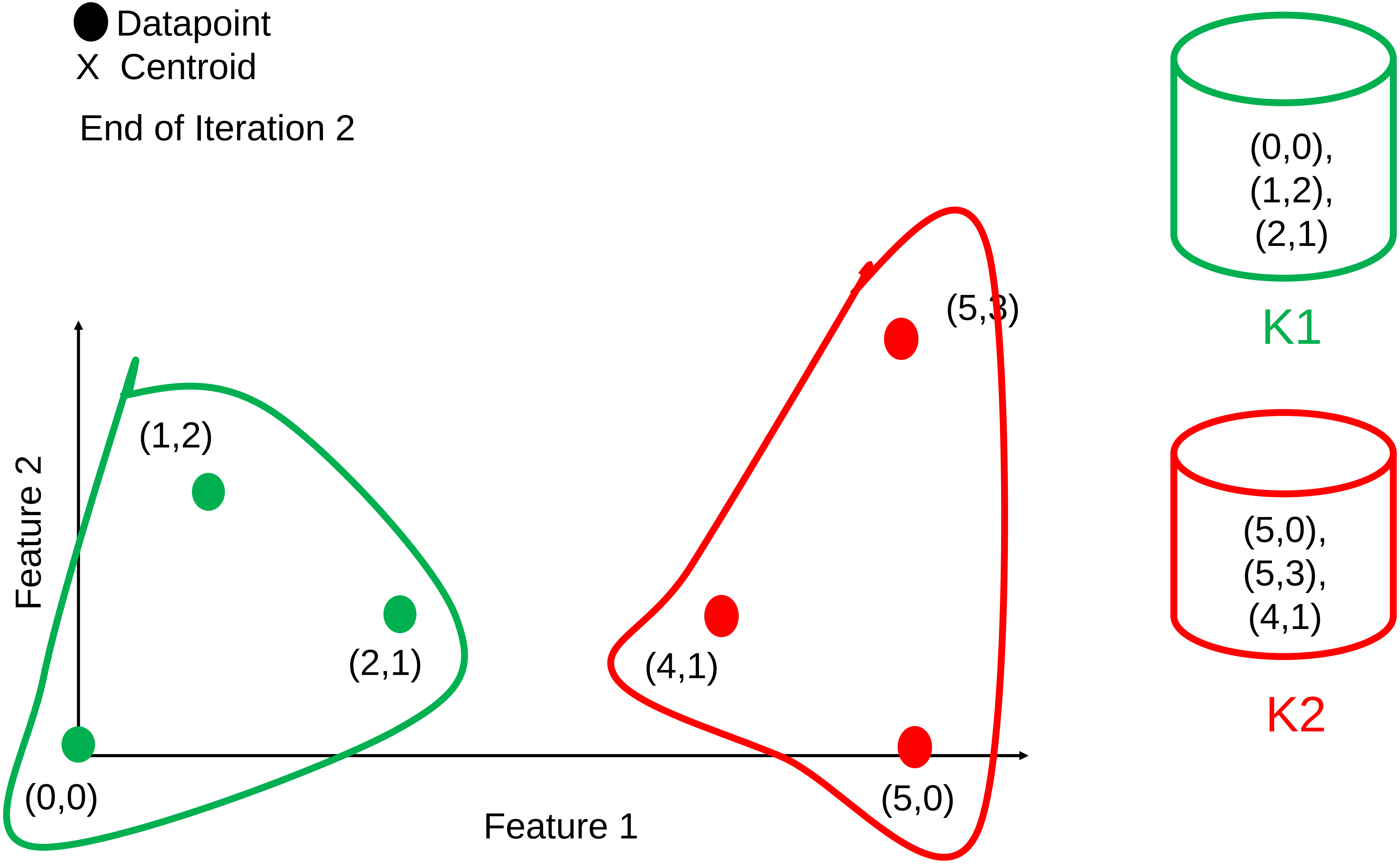
K-means Clustering



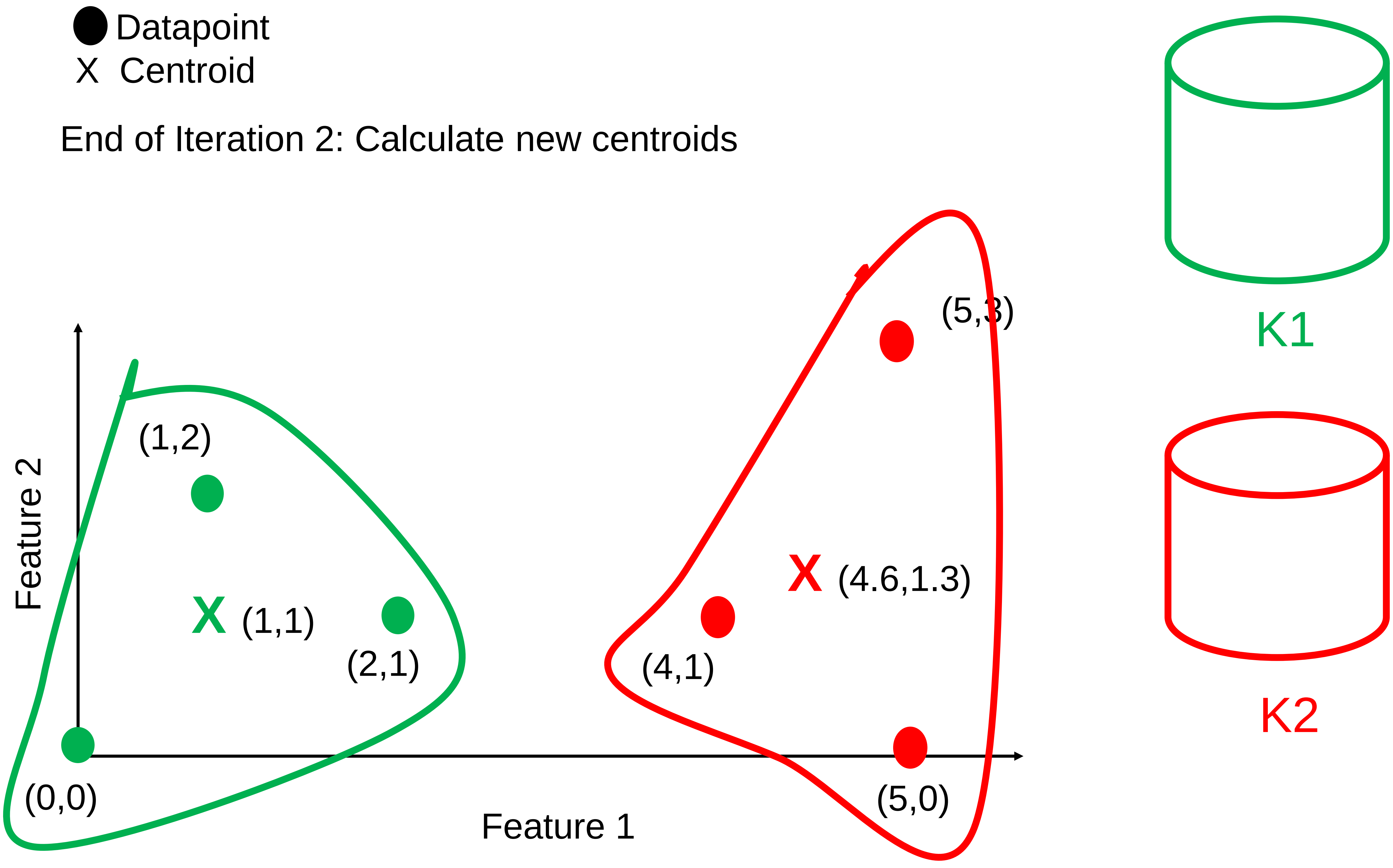
K-means Clustering

Datapoint
X Centroid

End of Iteration 2



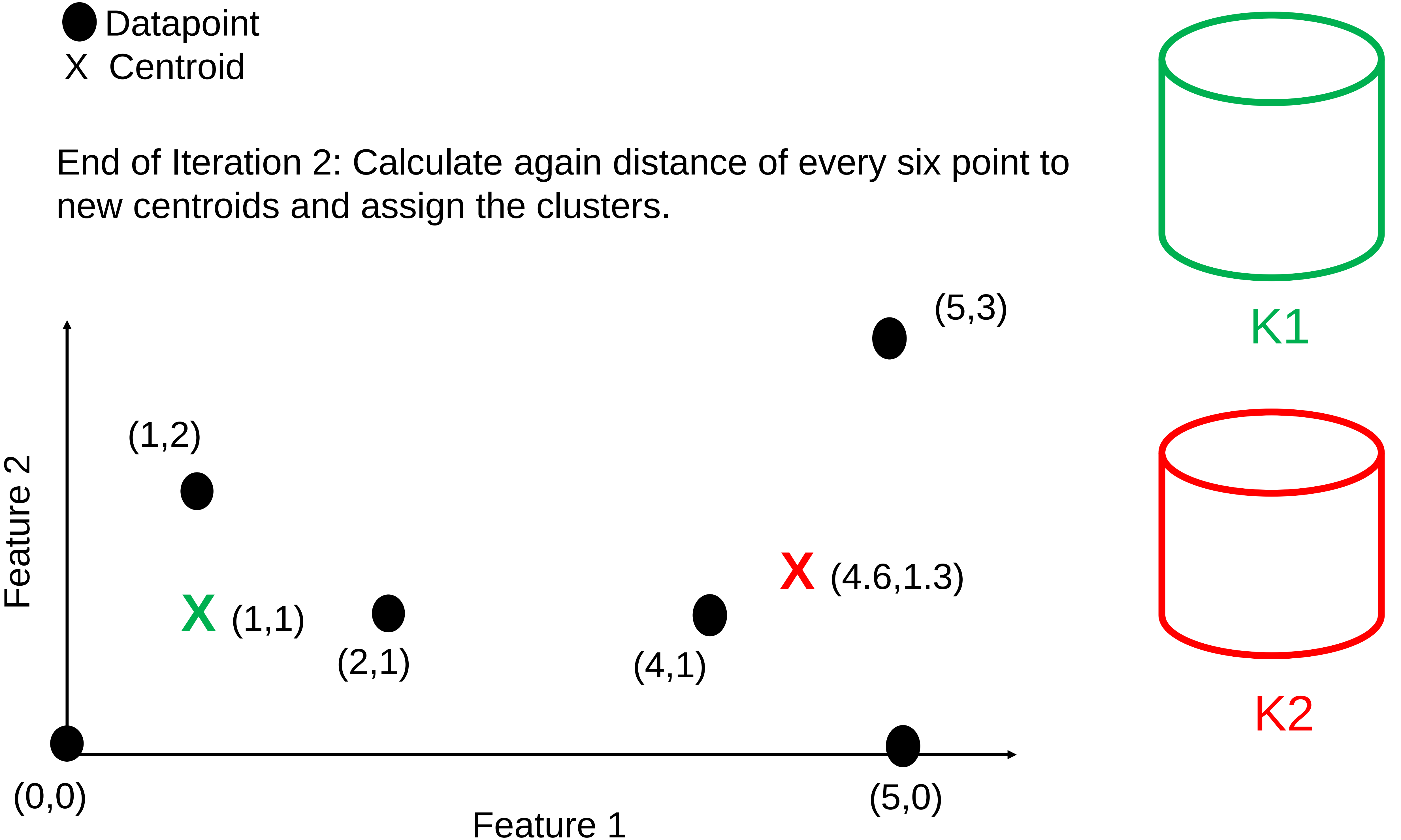
K-means Clustering



K-means Clustering

● Datapoint
X Centroid

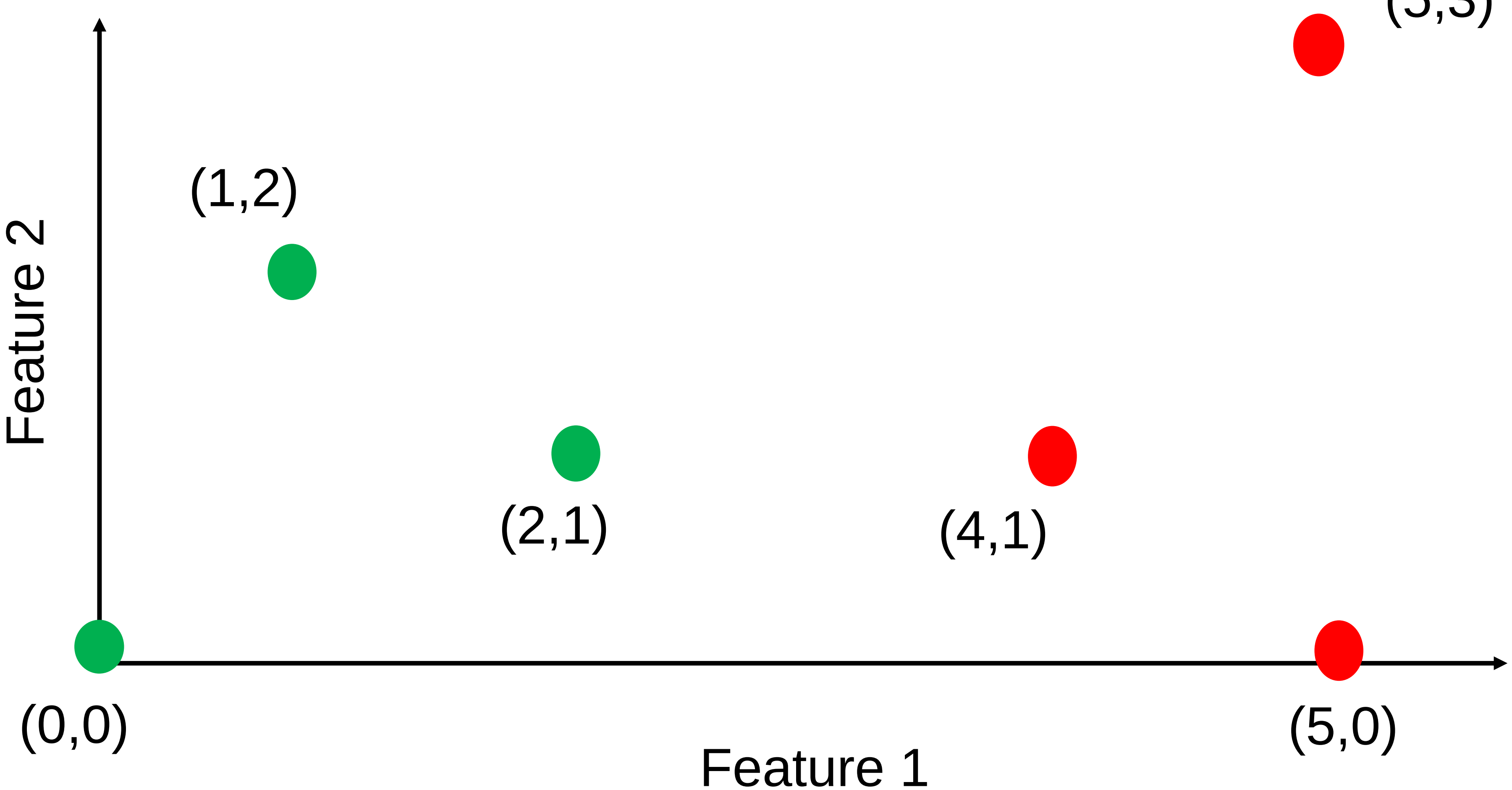
End of Iteration 2: Calculate again distance of every six point to new centroids and assign the clusters.

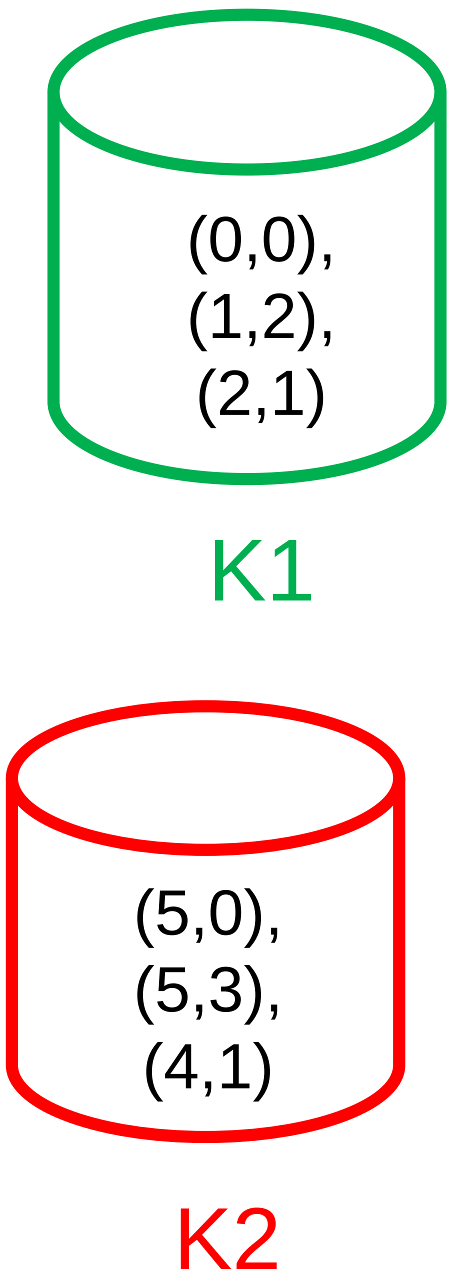


K-means Clustering

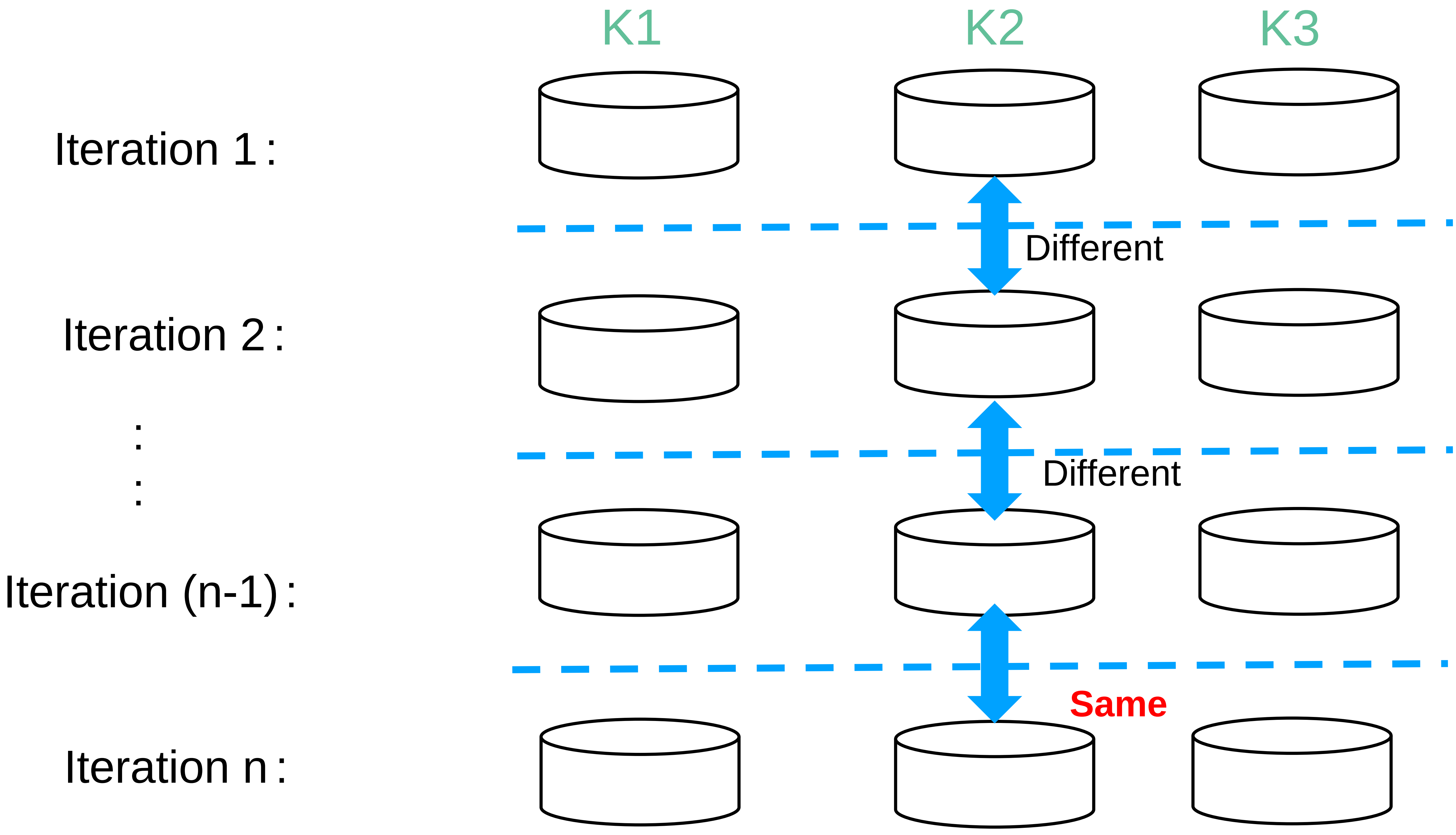
● Datapoint
X Centroid

Iteration 3: **Clusters remain the same (convergence has happened).**





Convergence: When to stop ?

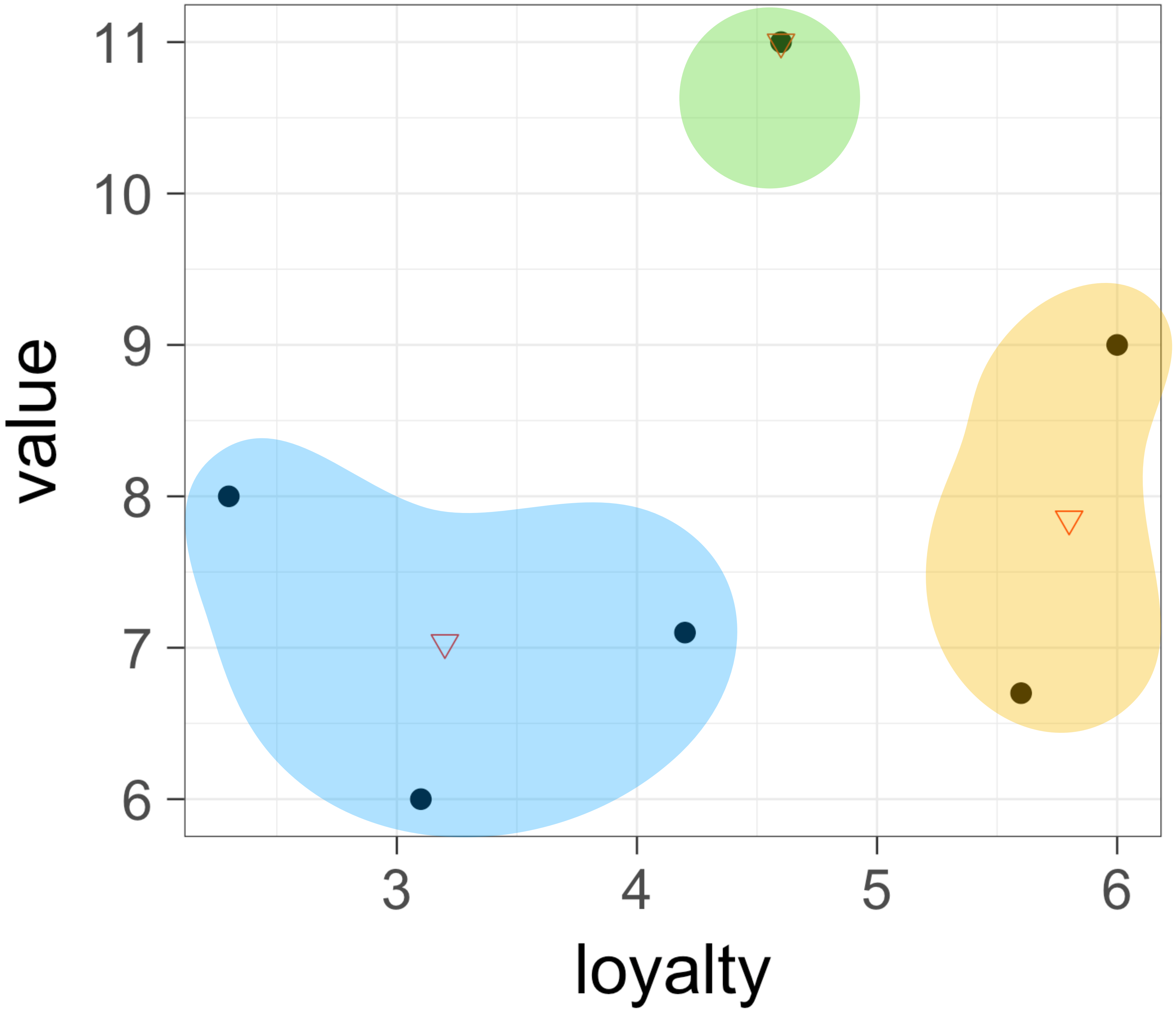


Exercise

You have a dataset with two dimensions, customer loyalty score and customer value score, and the following dataset for 6 customers:

	loyalty	value
1	2.3	8.0
2	5.6	6.7
3	4.2	7.1
4	3.1	6.0
5	4.6	11.0
6	6.0	9.0

Find 3 clusters using k-means



Python code:

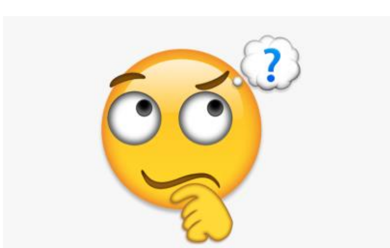
```
> model = KMeans(n_clusters=3)
> model.fit(df)
> clusters = model.predict(df)
```

Your result depends on K

Drawbacks

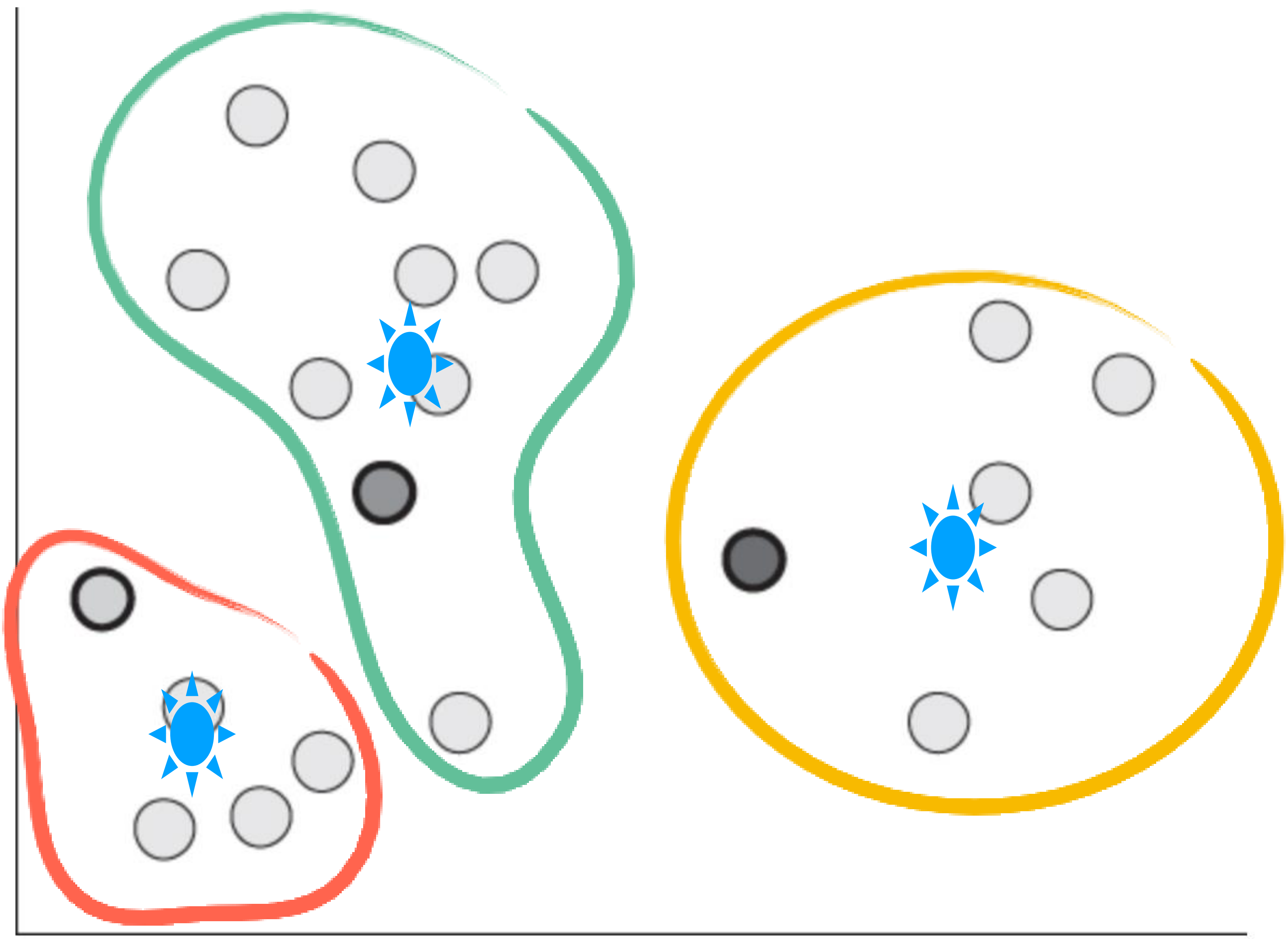
- How to select initial data points ?
  Pick the first seed randomly
  Pick the second seed as far as possible from the first seed.
  Pick the third seed as far as possible from the first two
  :
  :

- The number of clusters k is an input parameter:
  an inappropriate choice of k may yield poor results.



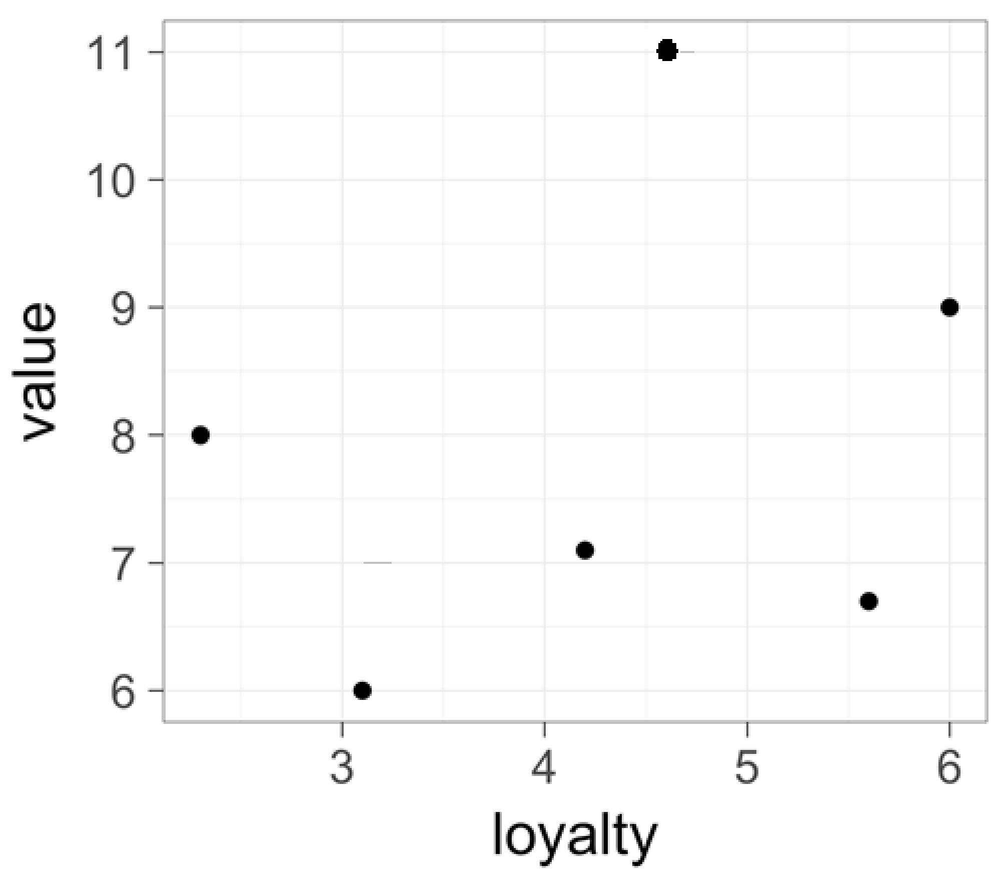
What is the best value of K ?

Y: Sum of Average of the distance from the centroid for each data point within that cluster.

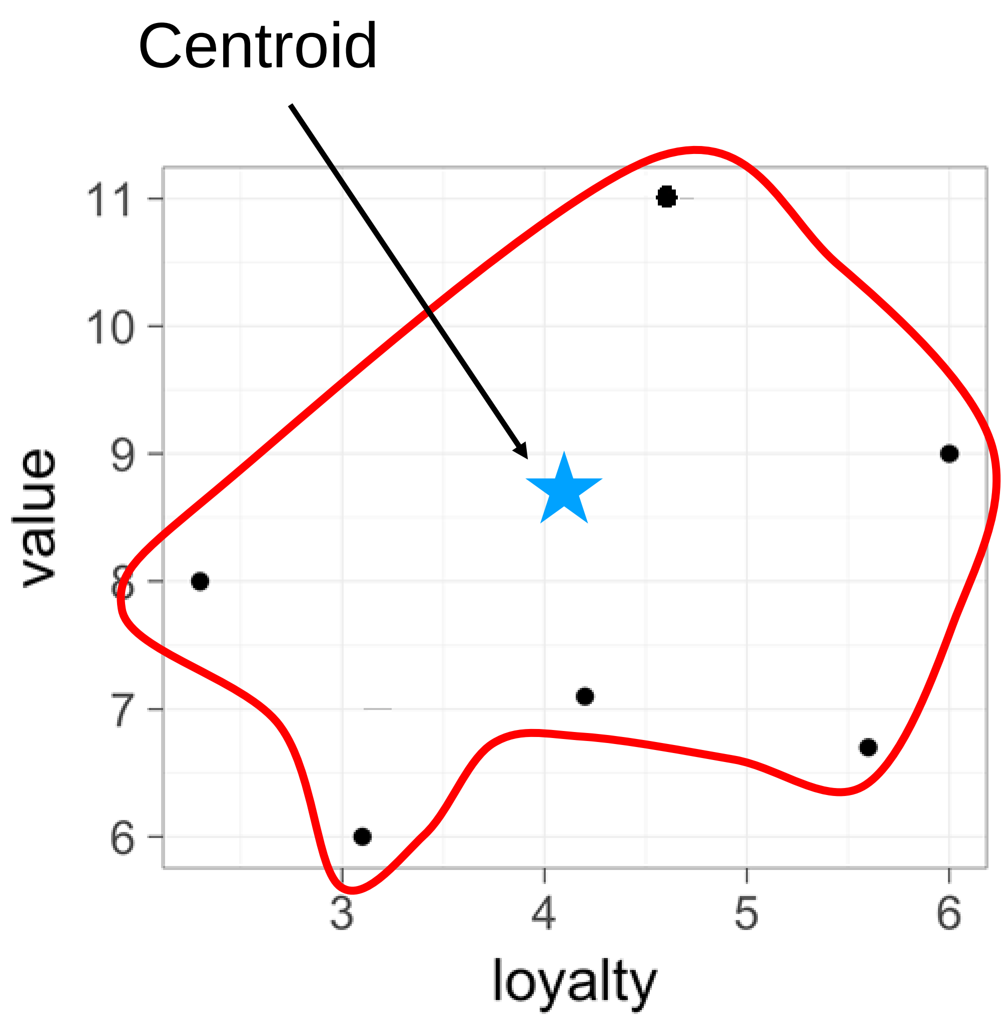


Which one has more Y?

Y: Sum of Average of the distance from the centroid for each data point within that cluster



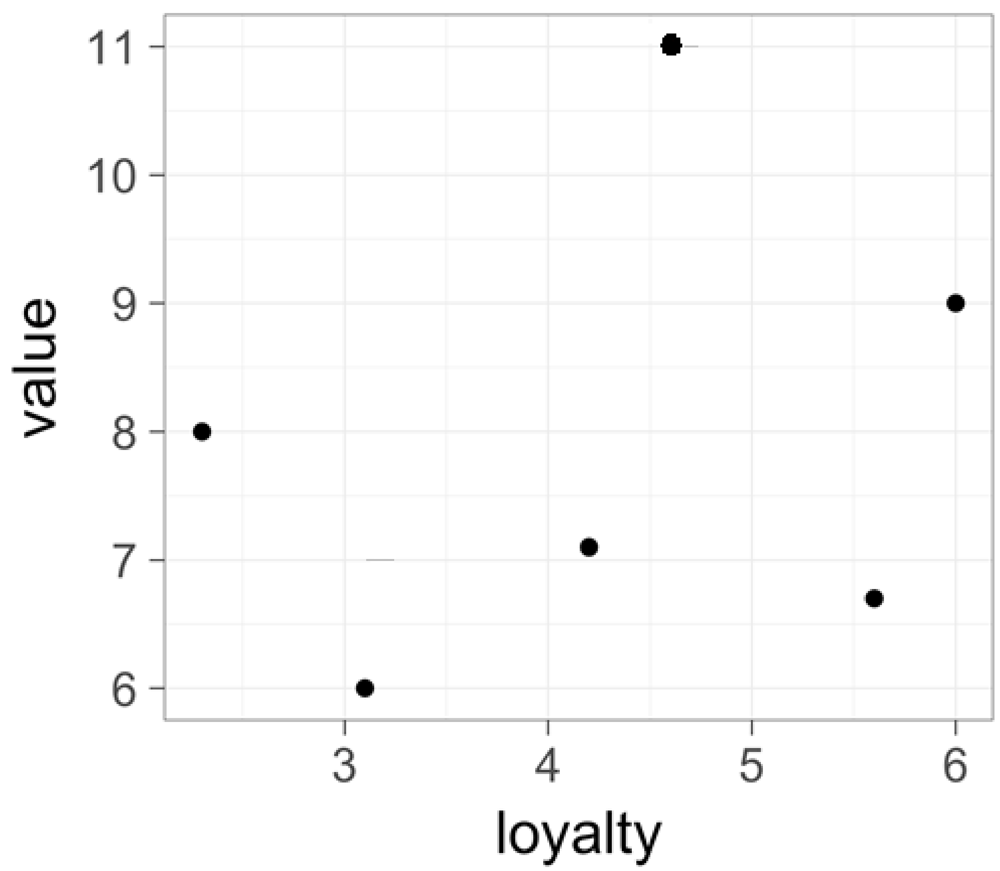
Individual Clusters



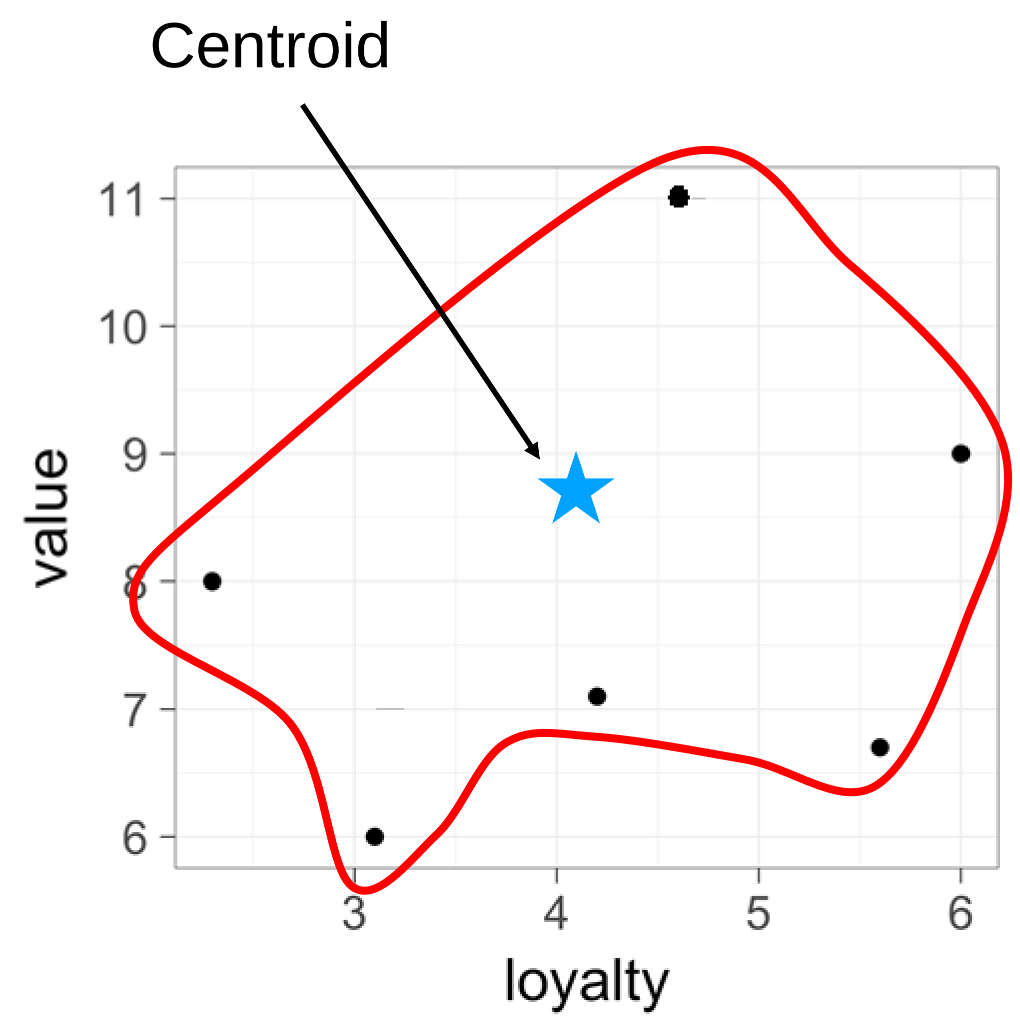
Only One Cluster

Which one has more Y?

Y: Sum of Average of the distance from the centroid for each data point within that cluster



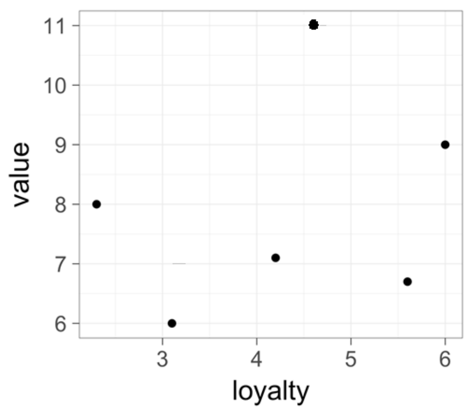
Individual Clusters = 0



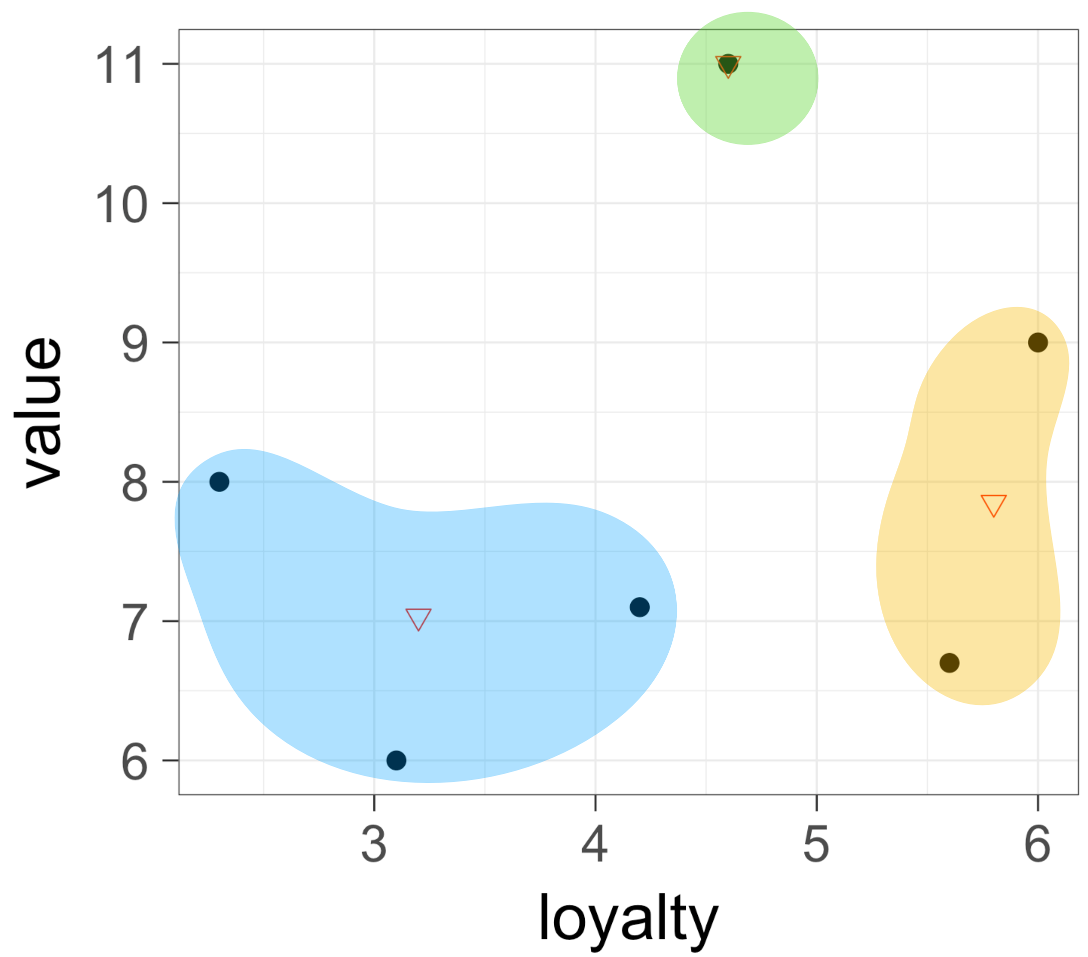
Only One Cluster = Very high

Which one has more Y?

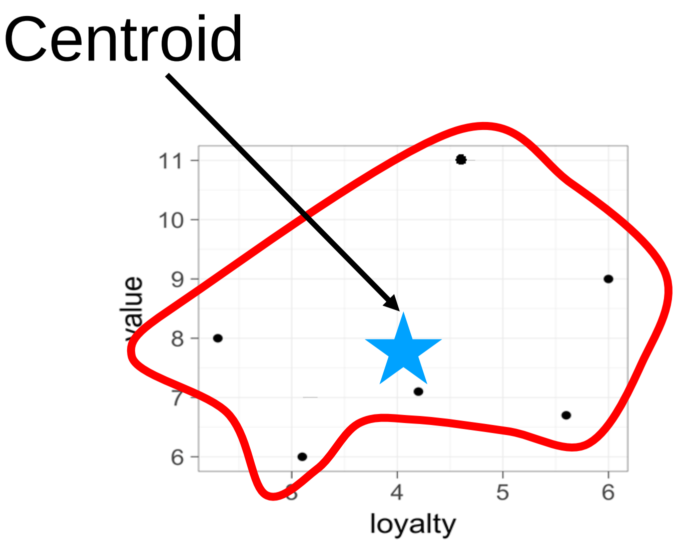
Y: Sum of Average of the distance from the centroid for each data point within that cluster



Individual Clusters = 0

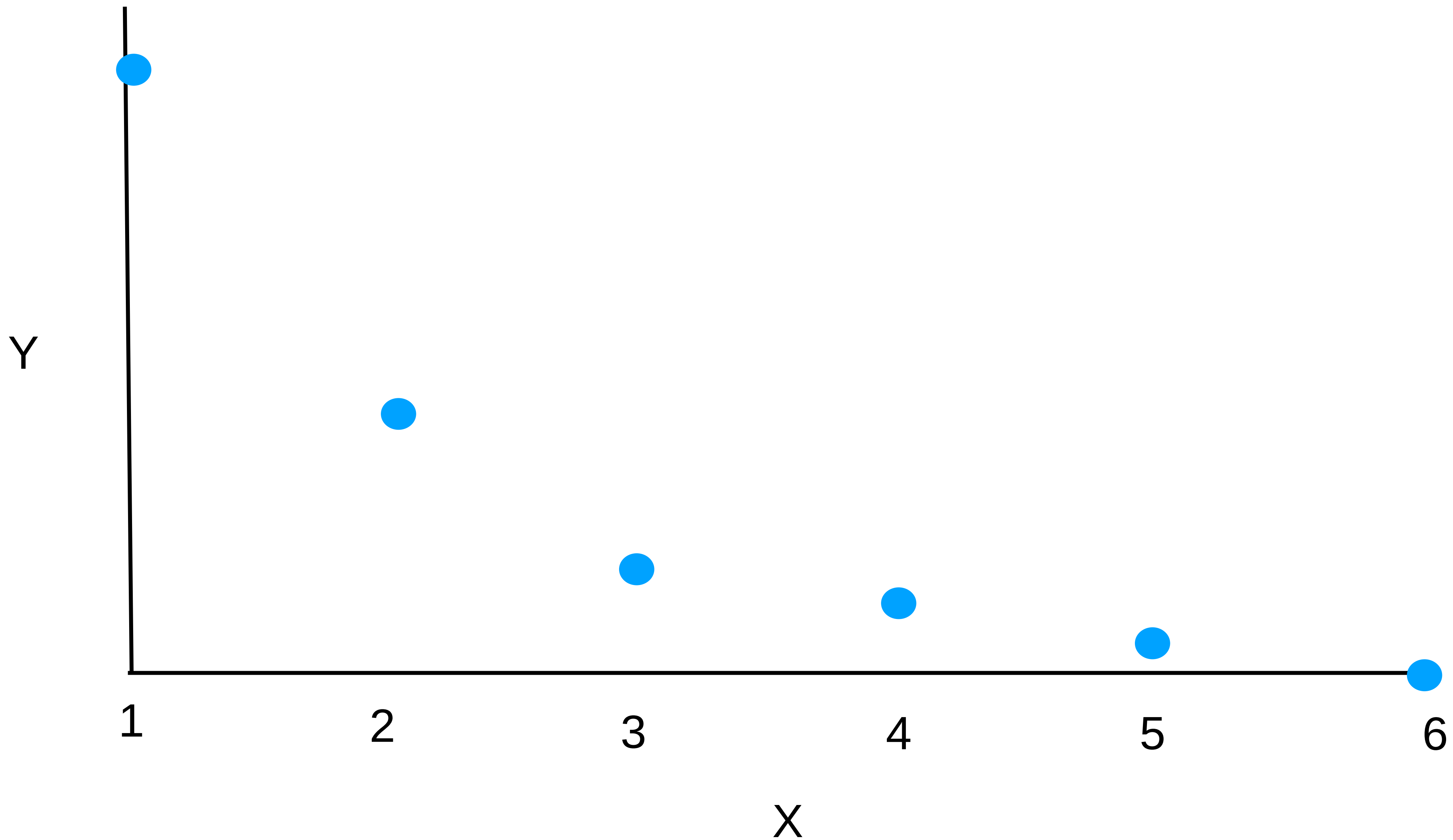


3 Clusters = ?



Only One Cluster = very high

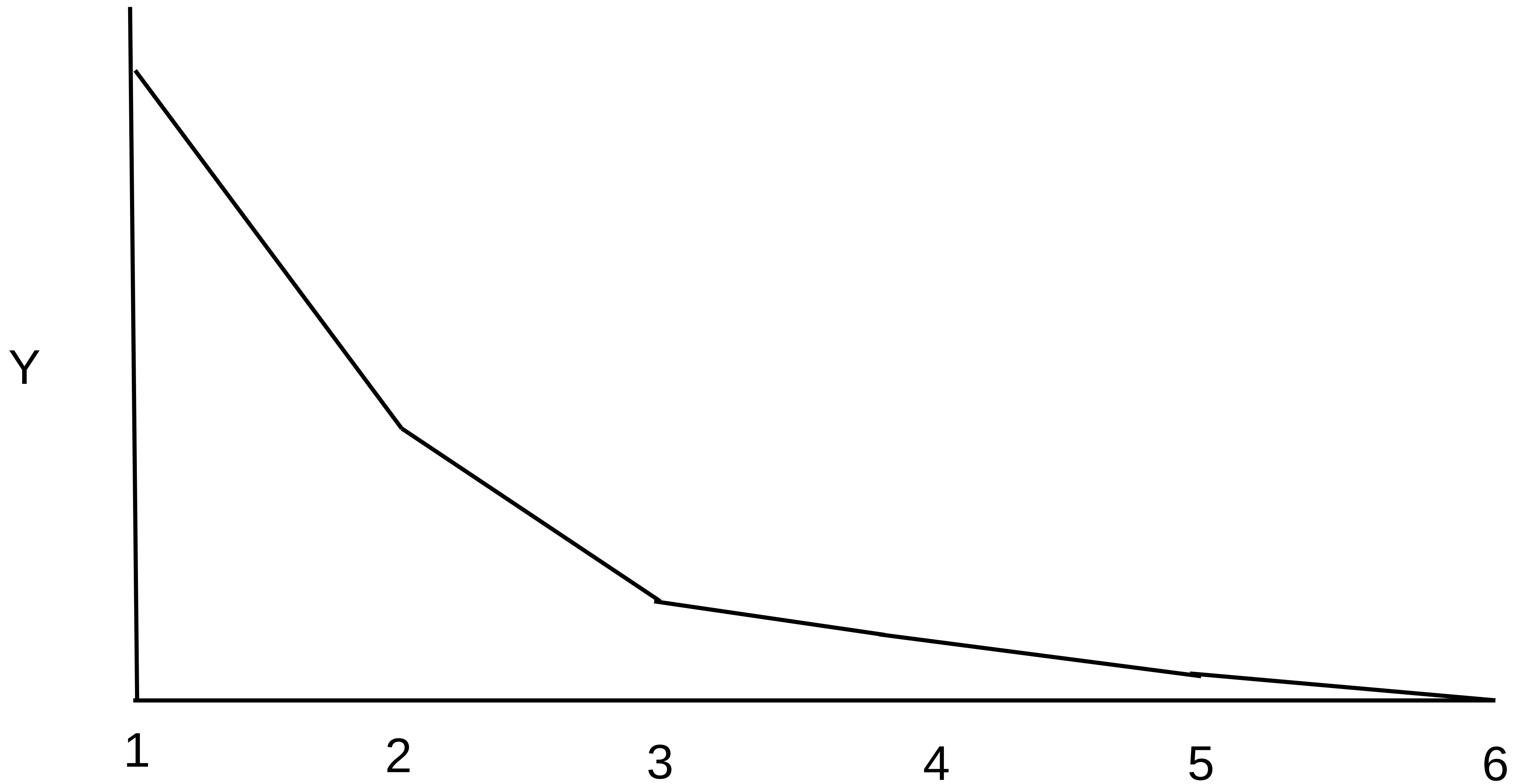
Let us plot Y (Y-axis) Vs. K (X-axis)



X-axis (represented by K) = number of clusters

Y-axis: Sum of Average of the distance from the centroid for each data point within that cluster

Let us plot Y Vs. K (Elbow Plot)

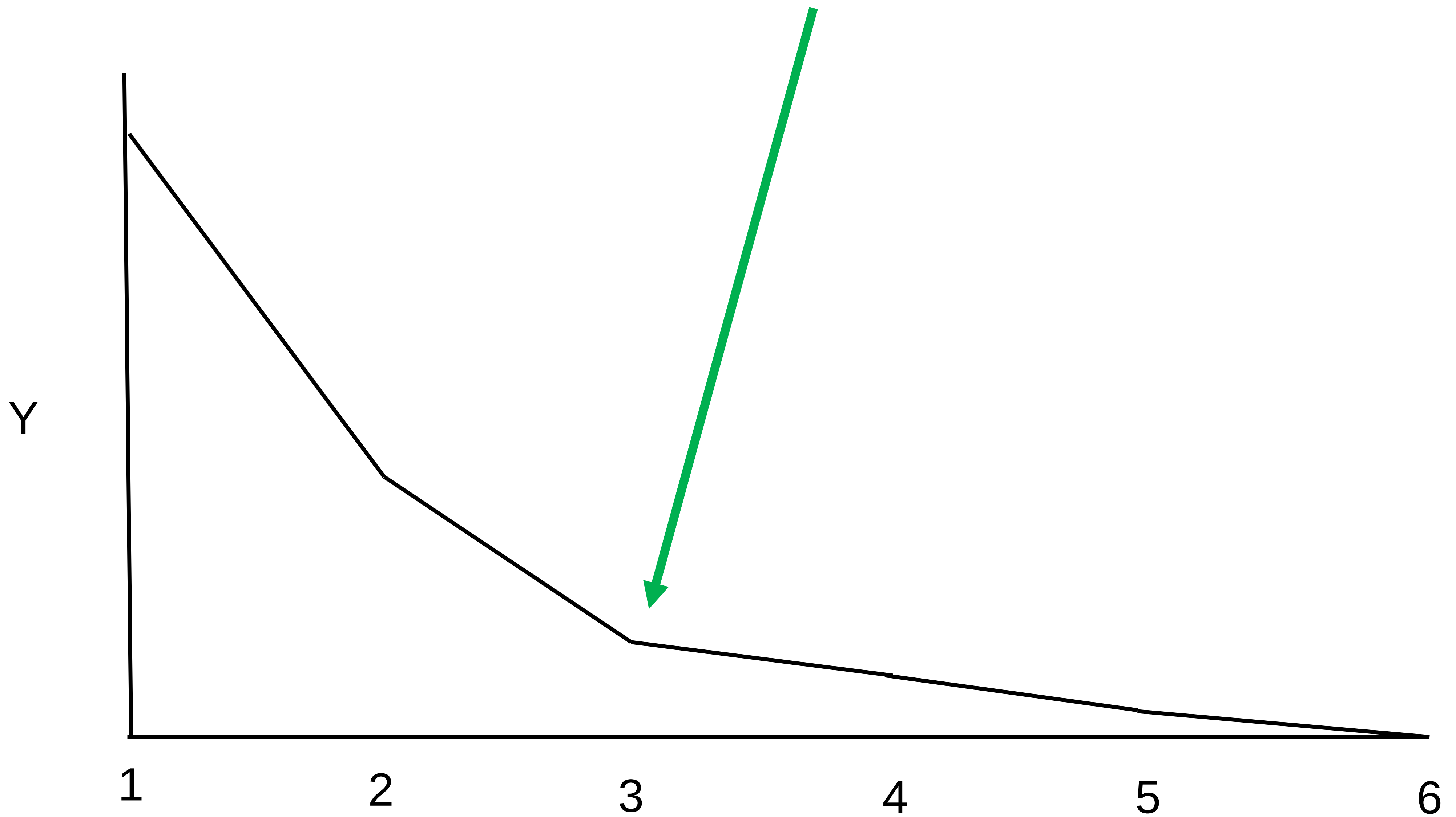


X-axis (represented by K) = number of clusters
Y-axis: Sum of Average of the distance from the centroid for each data point within that cluster

Let us plot Y Vs. K (Elbow Plot)

- **What is the best value of K ?**

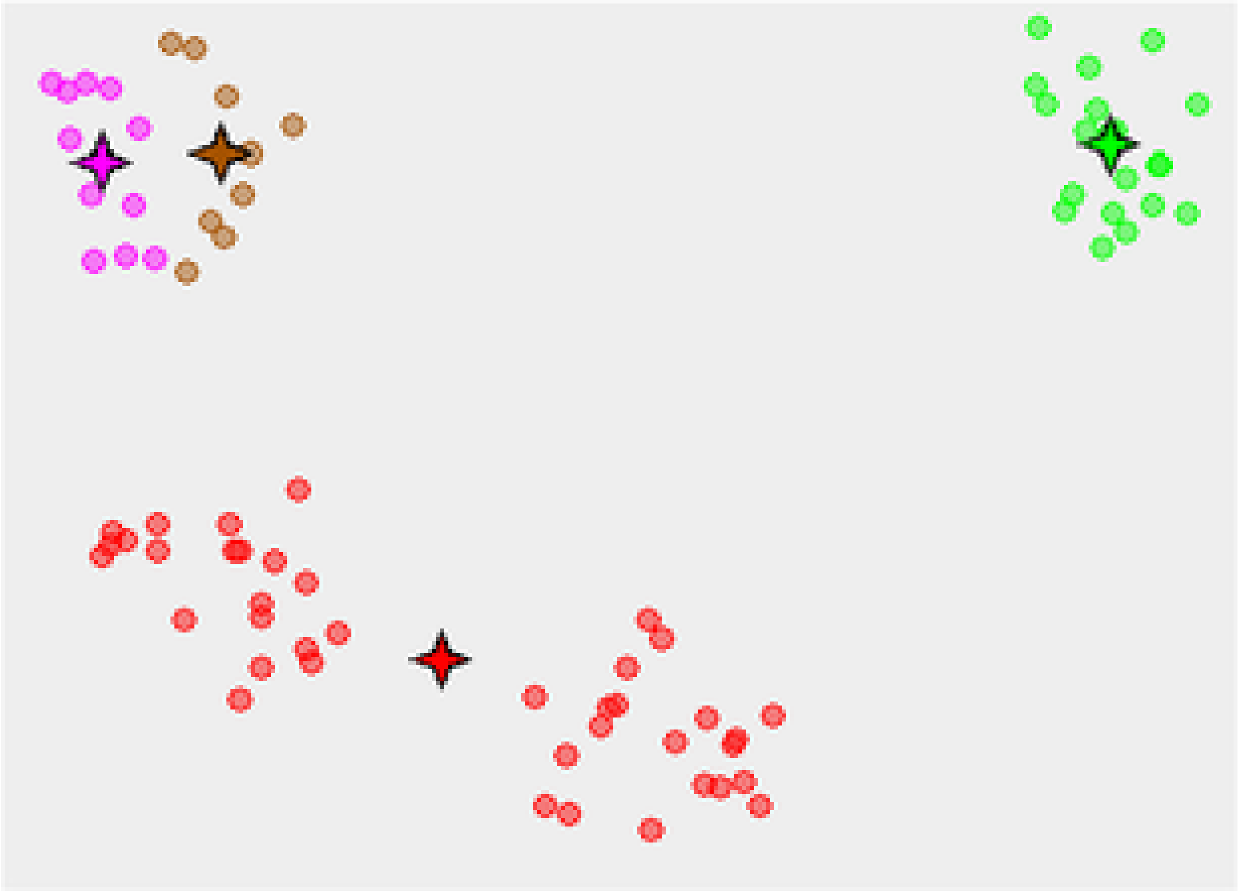
  Elbow point is the best value of K



X-axis (represented by K) = number of clusters
Y-axis: Sum of Average of the distance from the centroid for each data point within that cluster

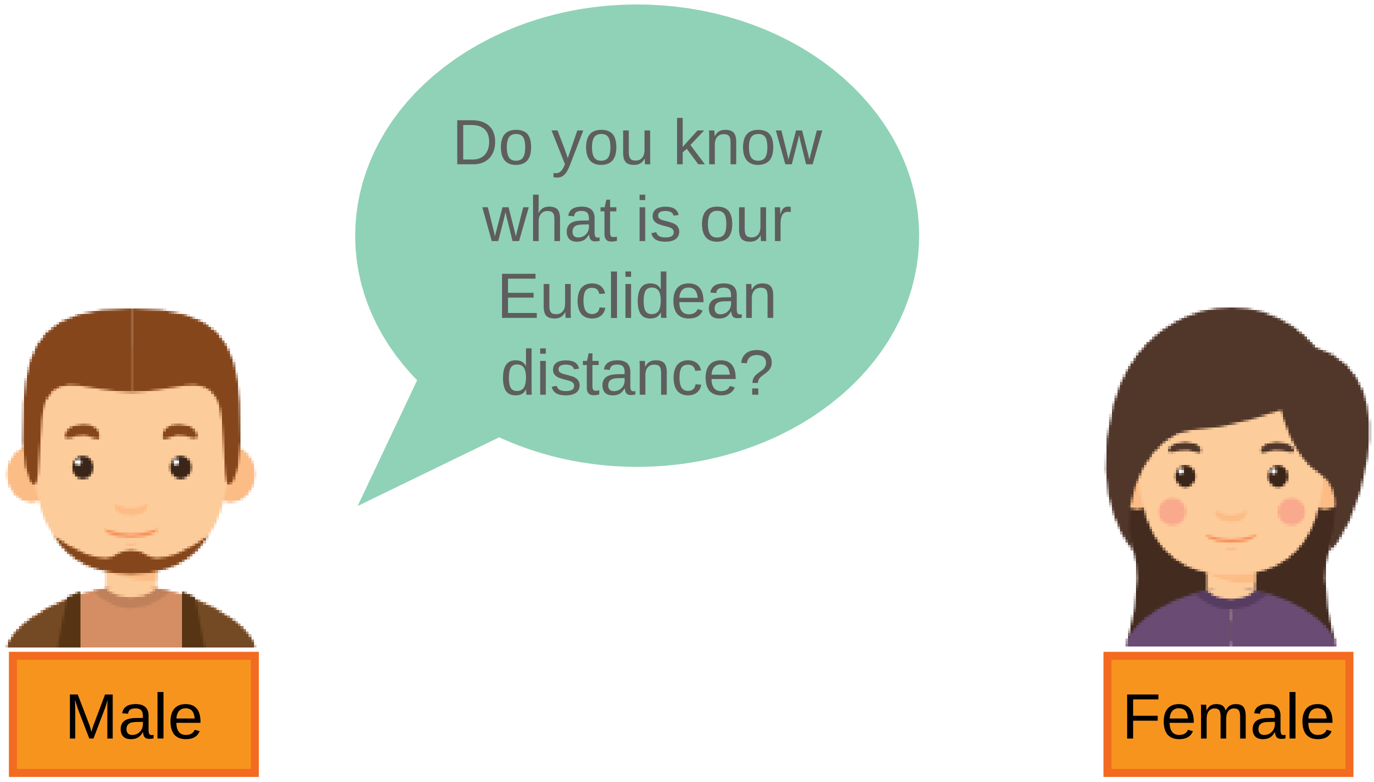
Drawbacks

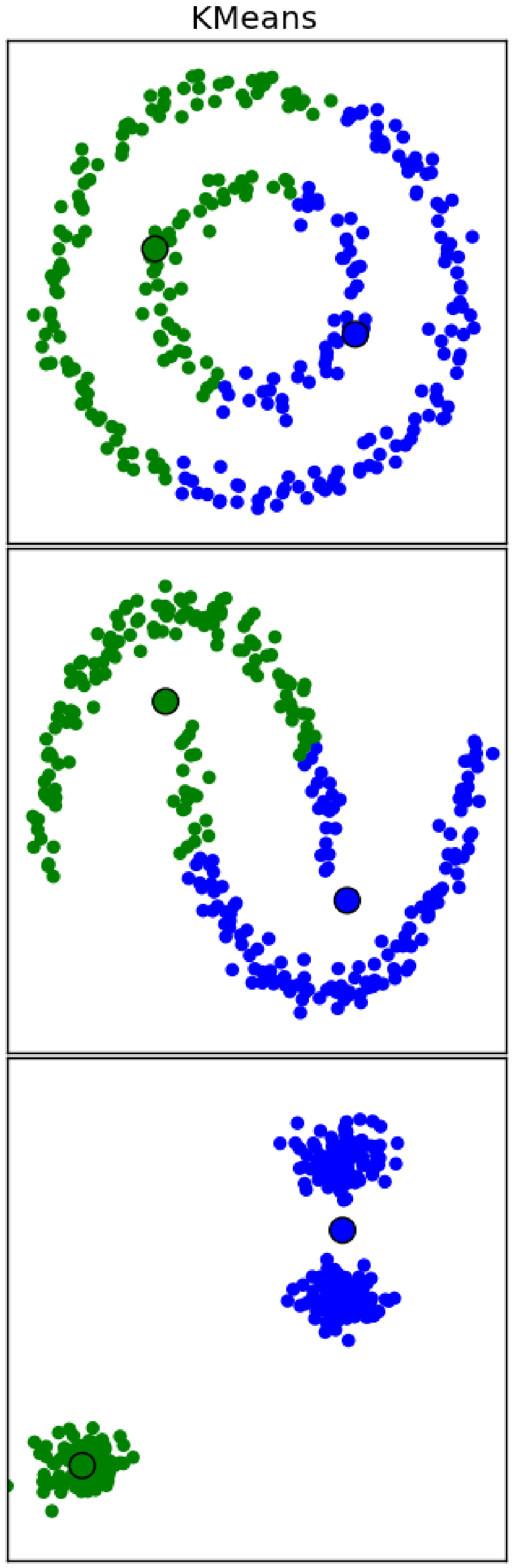
- Convergence to a local minimum may produce counterintuitive ("wrong") results
- Sensitive to noise and outliers



Drawbacks

- Can handle only numerical features



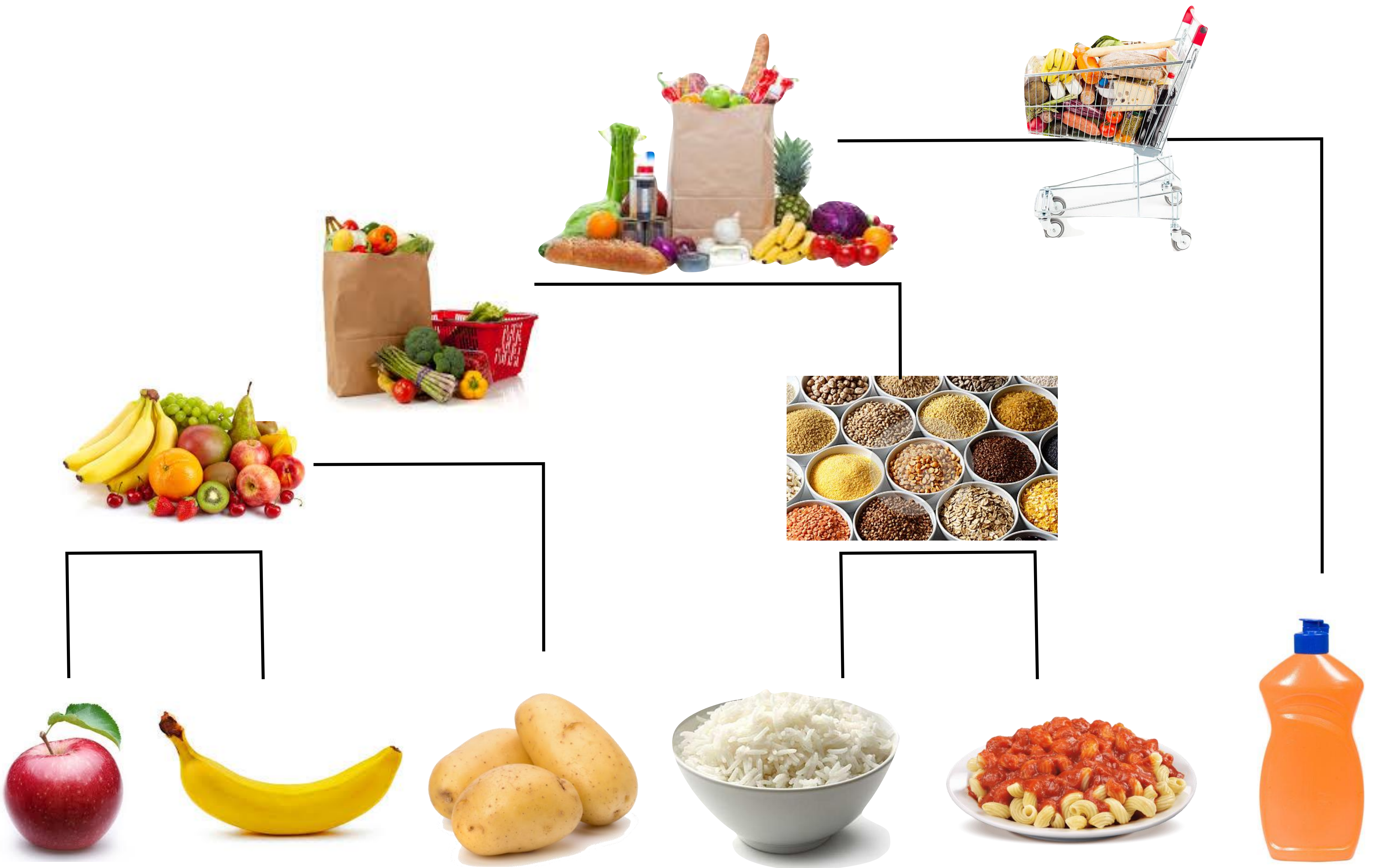


Drawbacks

- Cannot detect clusters of various shapes

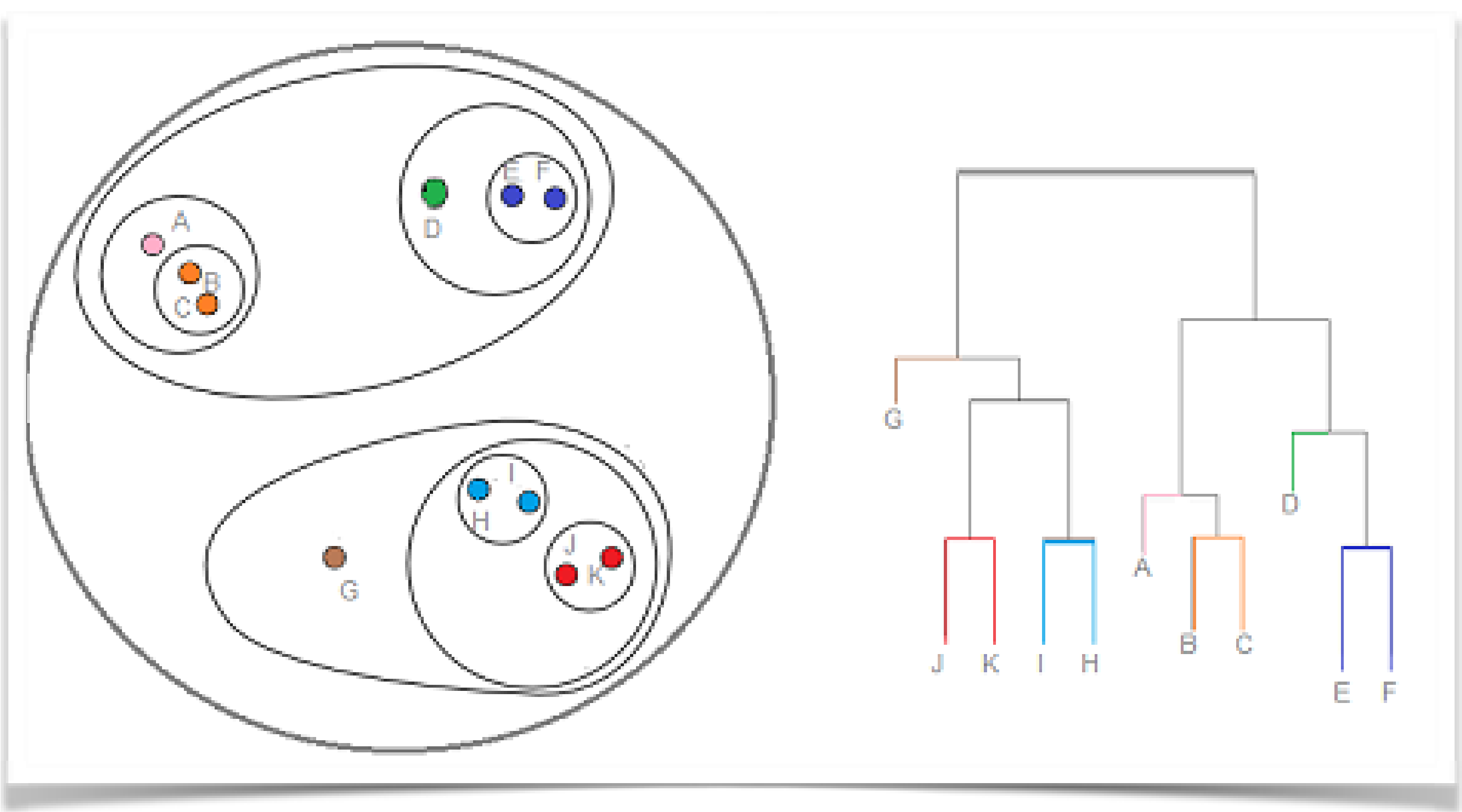
Hierarchical clustering

Adv: Clustering products by customer preferences



Why?

No need to choose number of clusters and worry about the initialization. Creates a tree, where lower levels are subclusters of higher levels

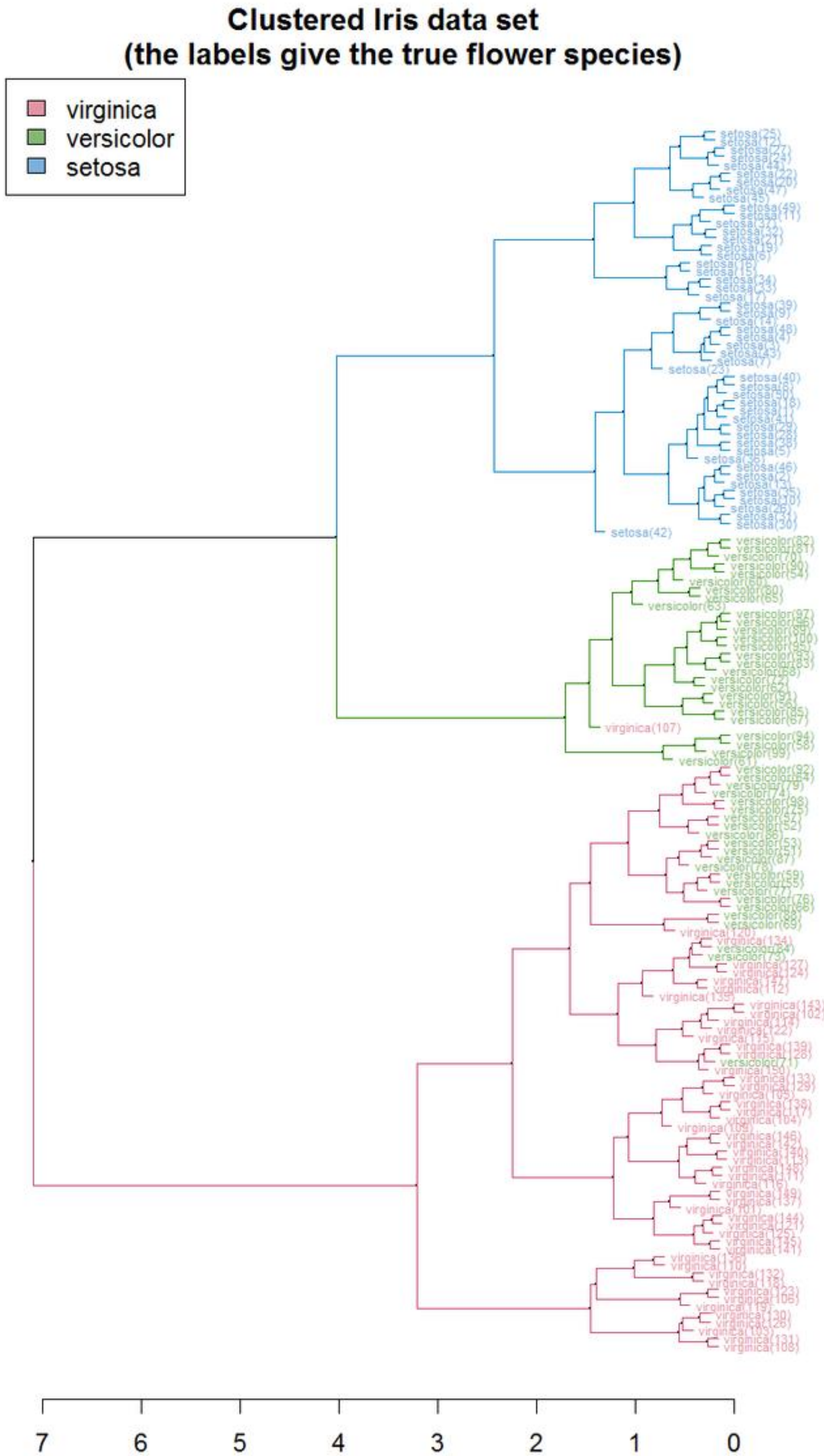


Why?

Any distance metric can be used

Names	Formula
Euclidean distance	$\|a-b\|_2 = \sqrt{\sum_i (a_i - b_i)^2}$
Squared Euclidean distance	$\|a-b\|_2^2 = \sum_i (a_i - b_i)^2$
Manhattan distance	$\|a-b\|_1 = \sum_i \lvert a_i - b_i \rvert$
maximum distance	$\|a-b\|_\infty = \max_i \lvert a_i - b_i \rvert$
Mahalanobis distance	$\sqrt{(a-b)^\top S^{-1} (a-b)}$ where S is the Covariance matrix

Why?



Easy to visualize, provides a good summary of the data structure in terms of clusters

The plot of hierarchical clustering is called dendrogram

Hierarchical clustering

- Agglomerative (bottom up)
  - Initially each point is a cluster
  - Repeatedly combine the two “nearest” clusters into one
- Divisive (Top Down)
  - Start with one cluster and recursively split it

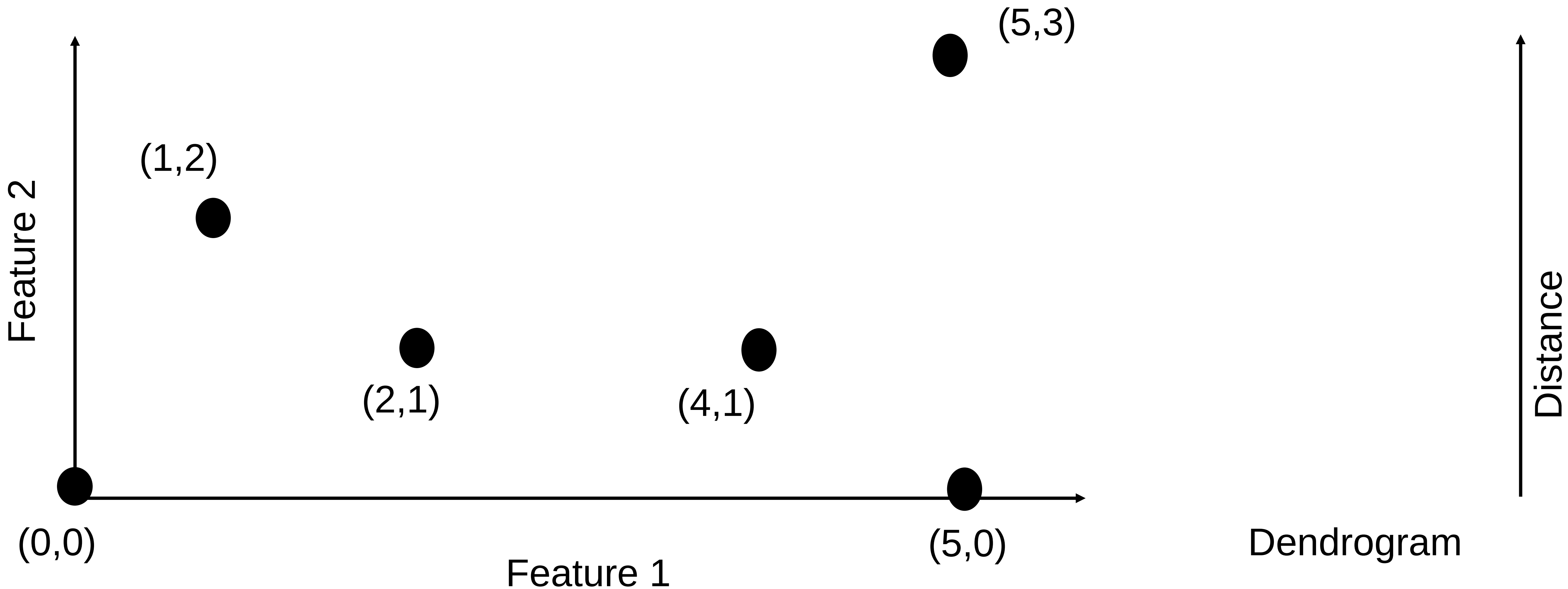
Hierarchical Clustering

Agglomerative (bottom up)

● Datapoint
X Centroid

We have 6 clusters. Each data point is a cluster.

Step 1: Calculate the distance (Euclidean) among very pair of points



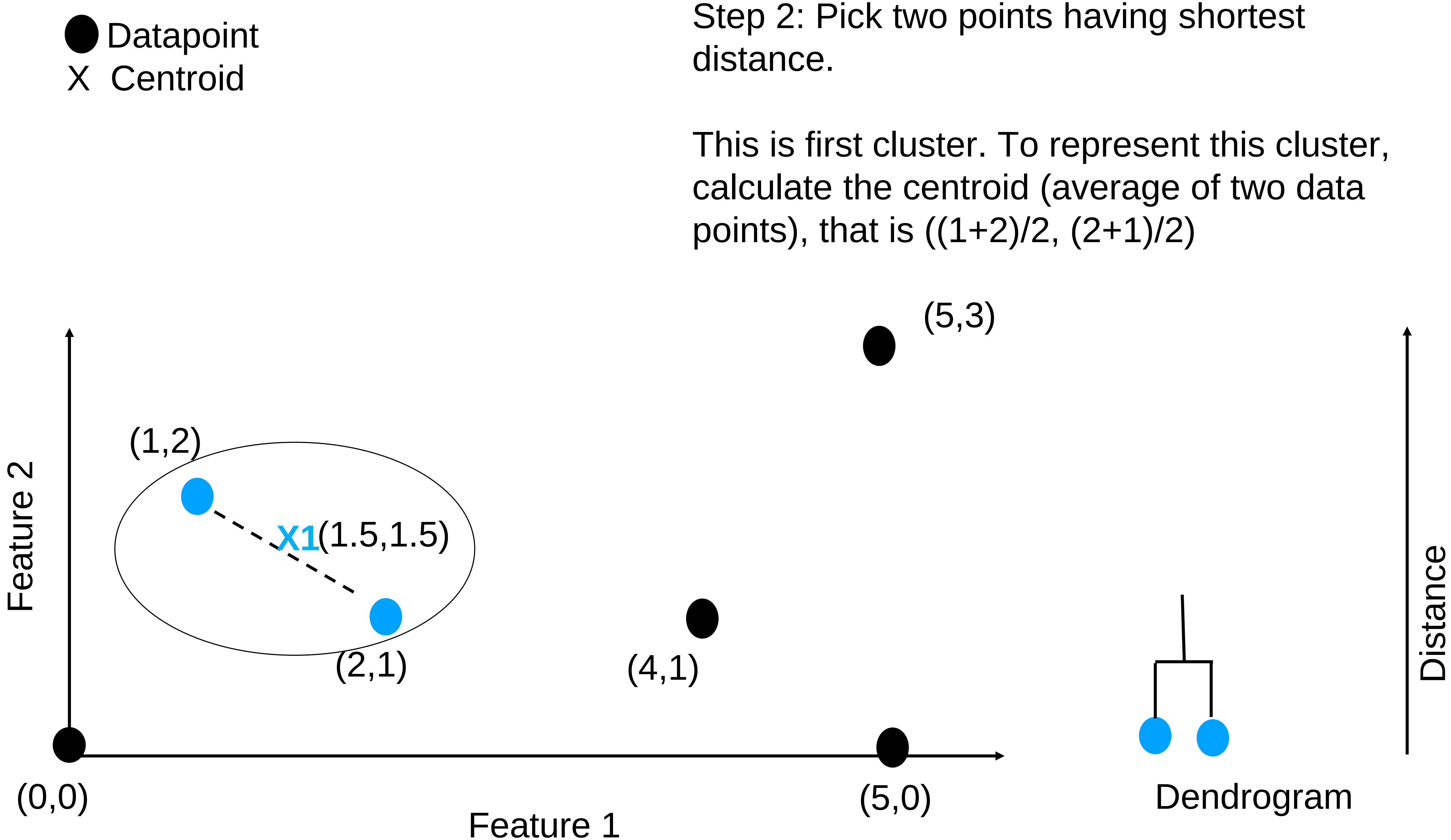
Source: Stanford – Hierarchical clustering

Hierarchical Clustering

Agglomerative (bottom up)

Step 2: Pick two points having shortest distance.

This is first cluster. To represent this cluster, calculate the centroid (average of two data points), that is ((1+2)/2, (2+1)/2)



Source: Stanford – Hierarchical clustering

Hierarchical Clustering

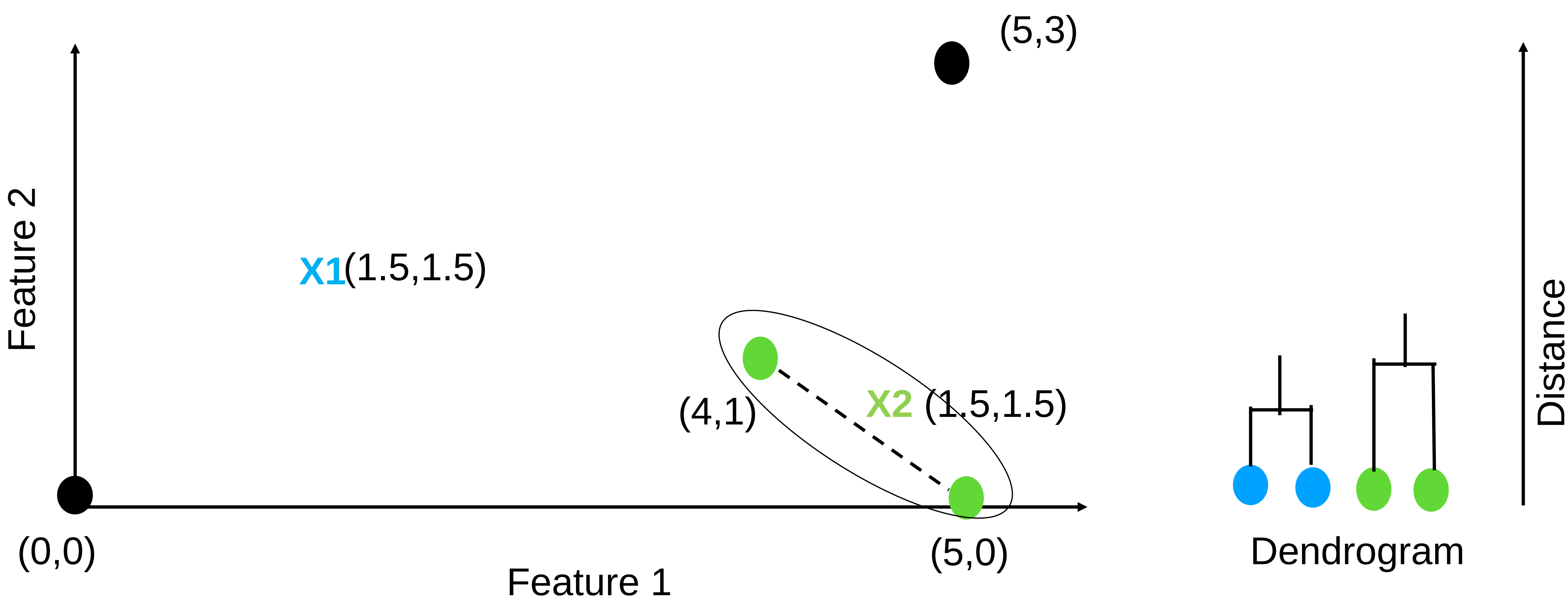
Agglomerative (bottom up)

● Datapoint
X Centroid

Now we have 5 clusters.

Calculate again the distance among all the data points, but now cluster of (1,2) and (2,1) is represented by X1(1.5, 1.5)

Pick with min distance.



Source: Stanford – Hierarchical clustering

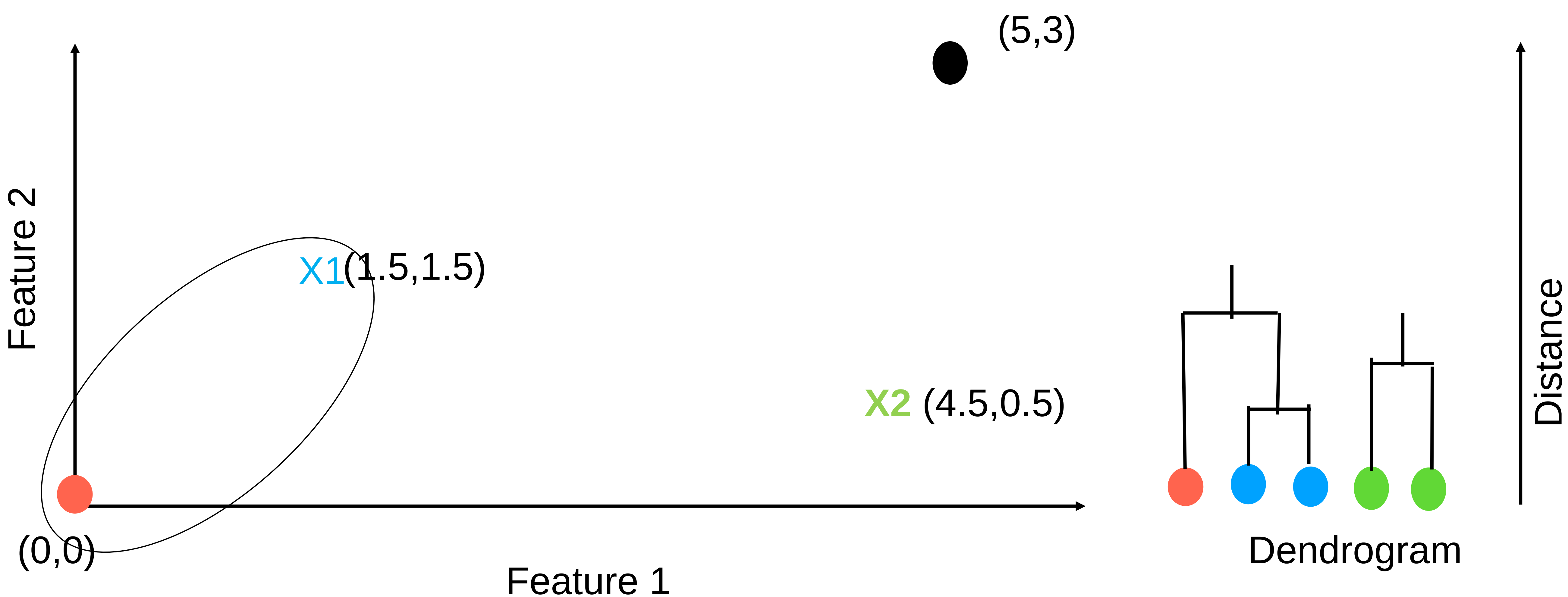
Hierarchical Clustering

Agglomerative (bottom up)

Now we have 4 clusters.

Distance between (0,0) and centroid X1 is the minimum

Pay Attention: How to calculate the new centroid ?

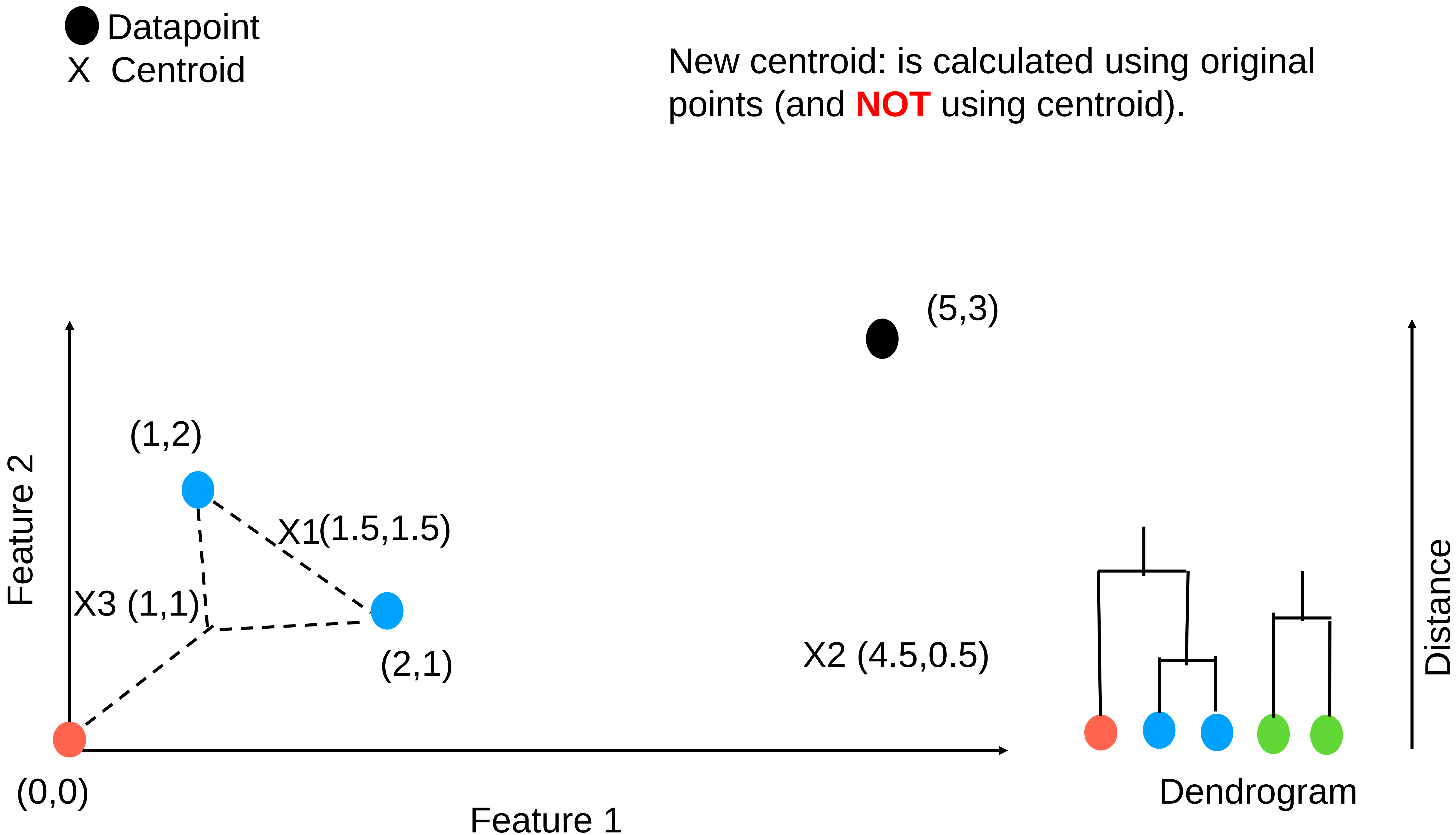


Source: Stanford – Hierarchical clustering

Hierarchical Clustering

Agglomerative (bottom up)

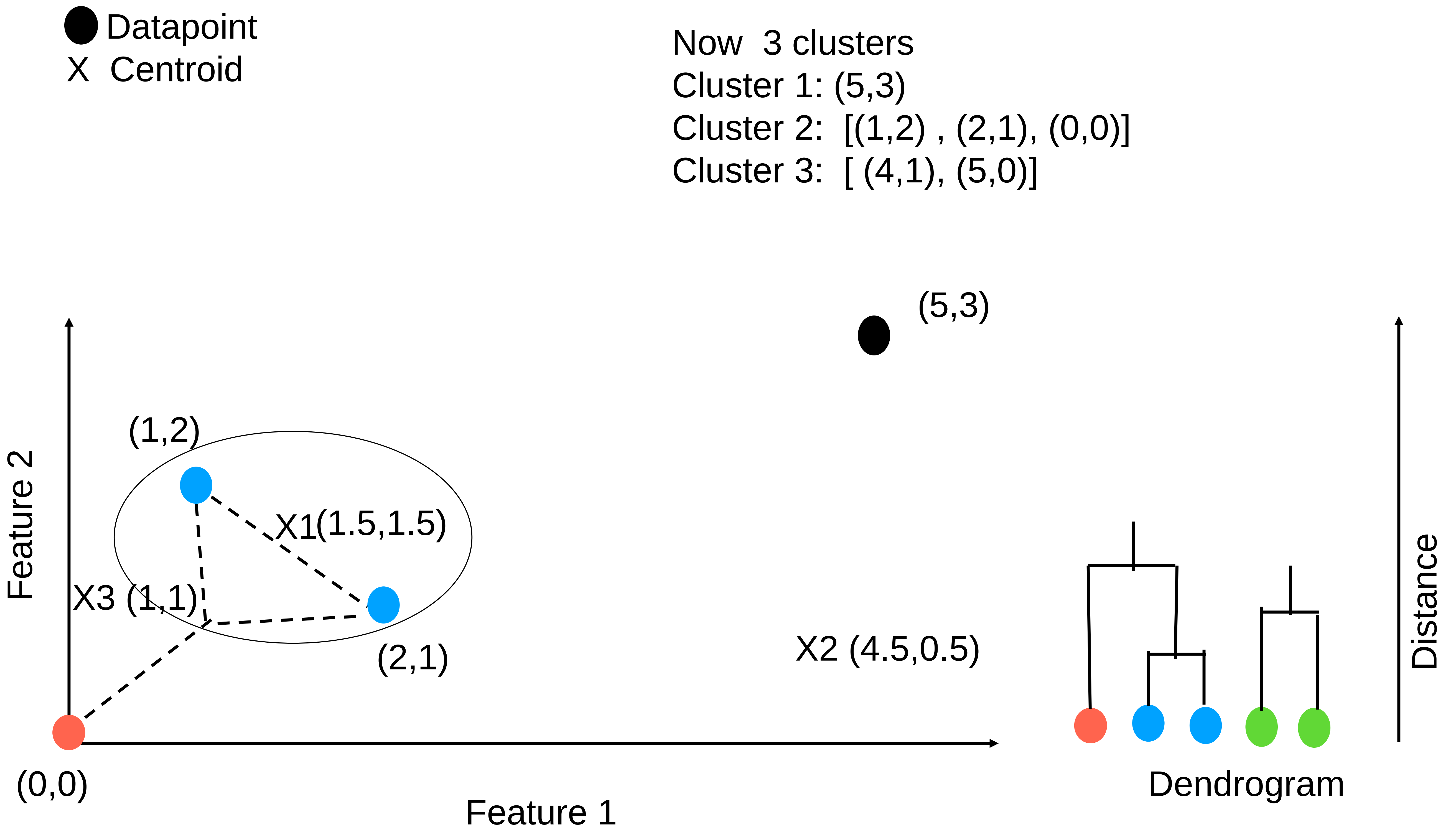
New centroid: is calculated using original points (and **NOT** using centroid).



Source: Stanford – Hierarchical clustering

Hierarchical Clustering

Agglomerative (bottom up)



Source: Stanford – Hierarchical clustering

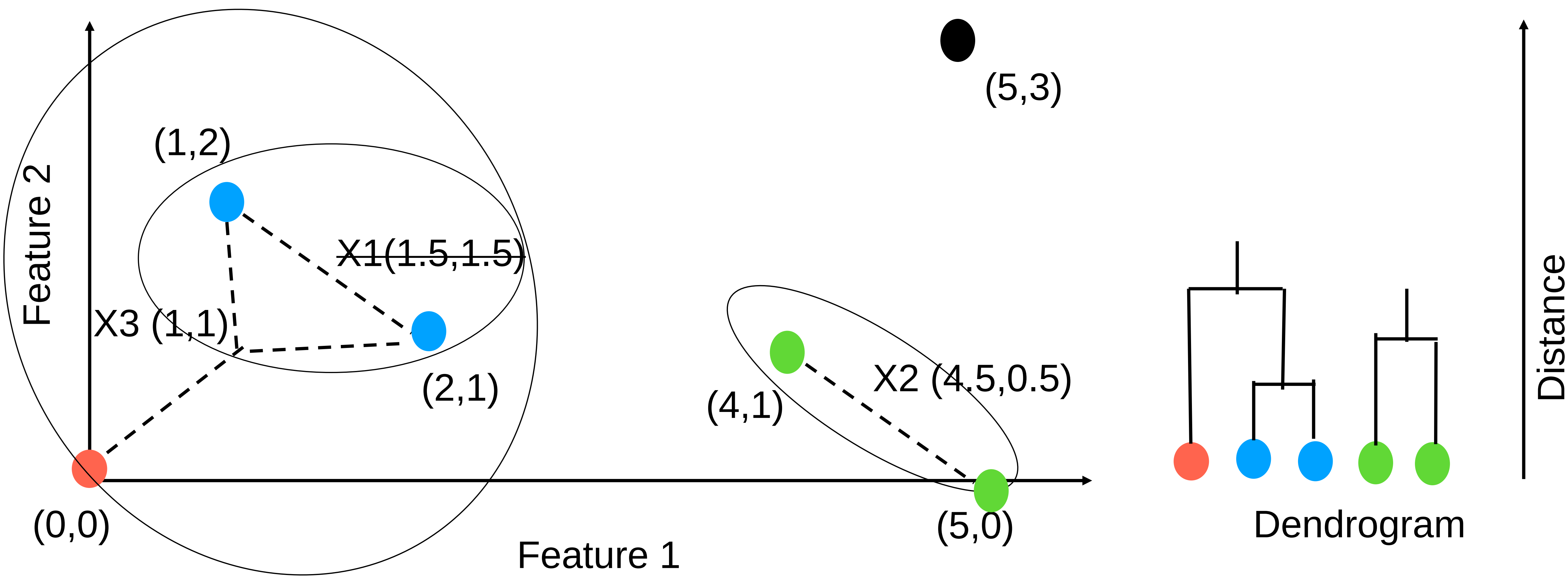
Now 3 clusters

Cluster 1: (5,3)
Cluster 2: [(1,2) , (2,1), (0,0)]
Cluster 3: [(4,1), (5,0)]

Datapoint
X Centroid

NOTE: To calculate distance: Always use recent centroids or data points, whichever applicable

Calculate Distance between
(5,3) and X2
X2 and X3
X3 and (5,3)



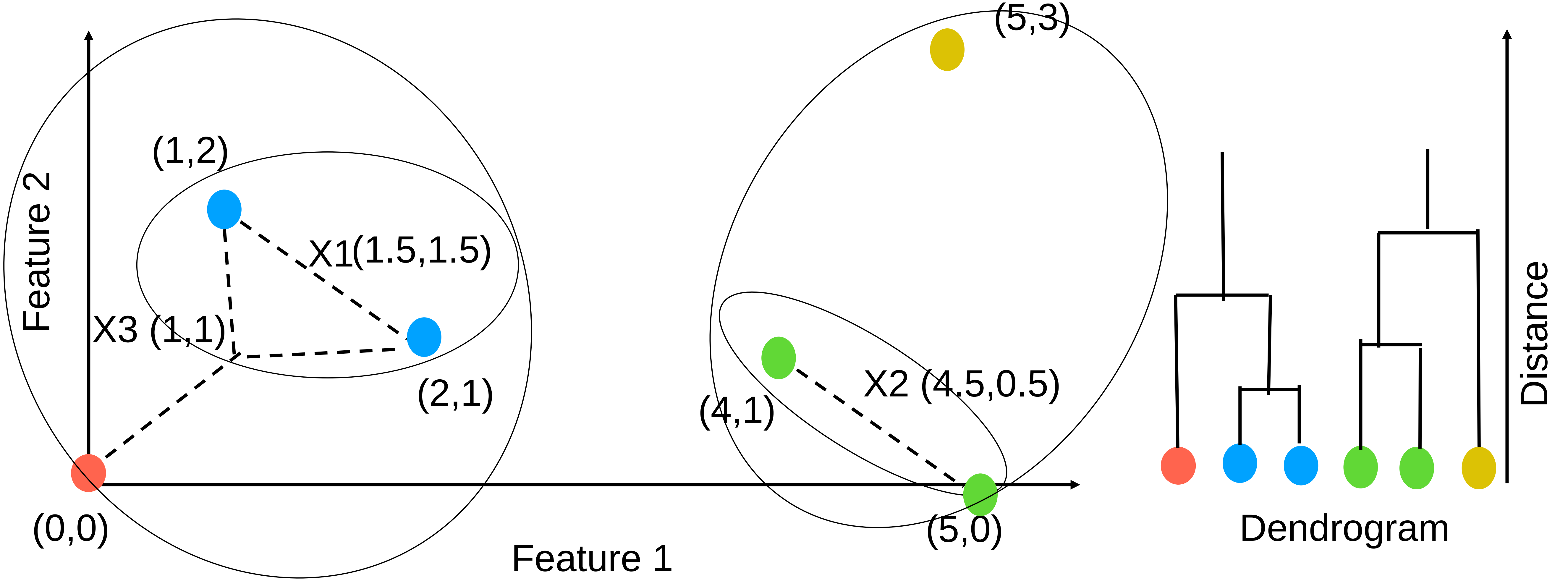
Source: Stanford – Hierarchical clustering

Hierarchical Clustering

Agglomerative (bottom up)

Datapoint
X Centroid

Now we have two clusters

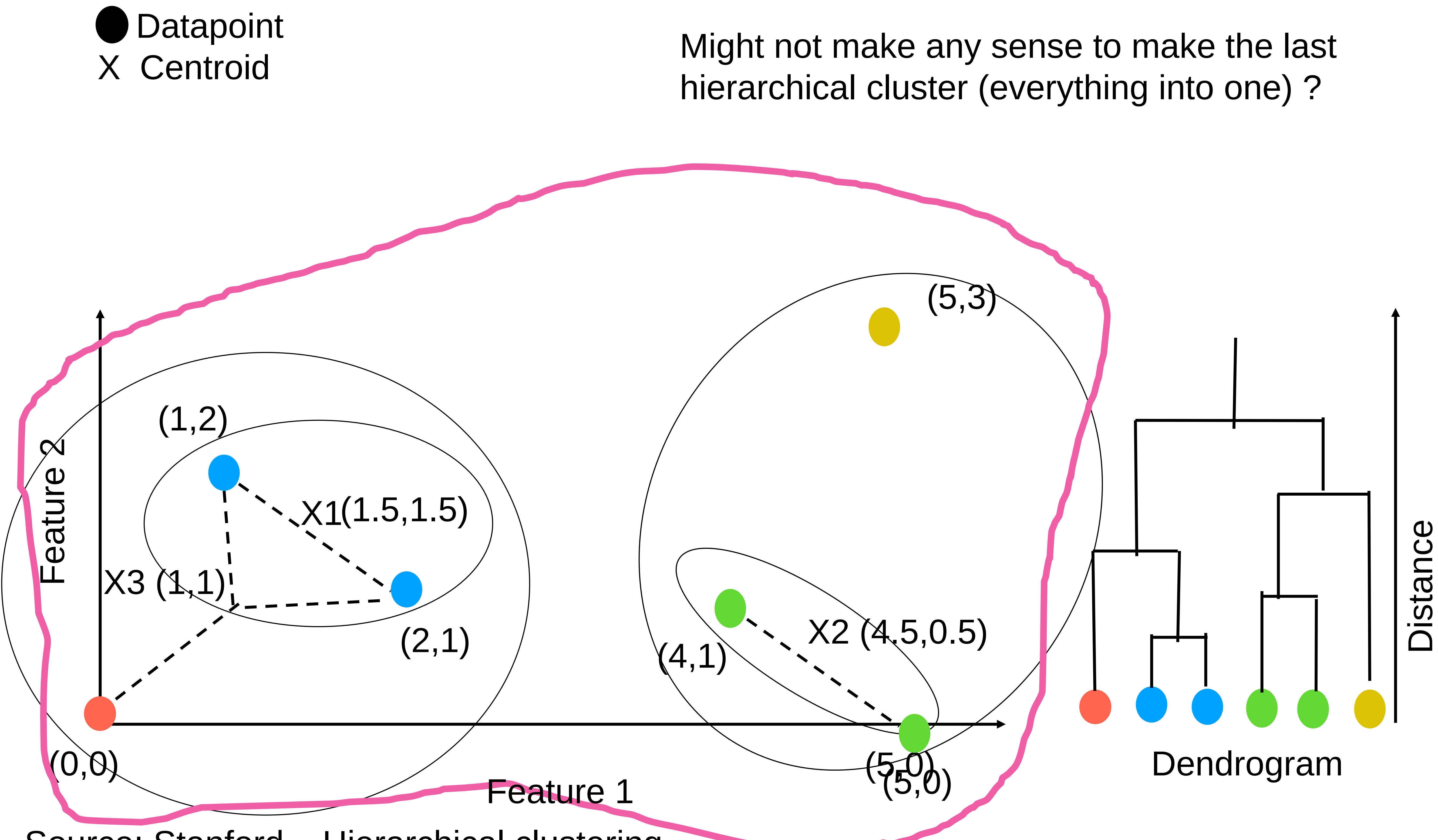


Source: Stanford – Hierarchical clustering

Hierarchical Clustering

Agglomerative (bottom up)

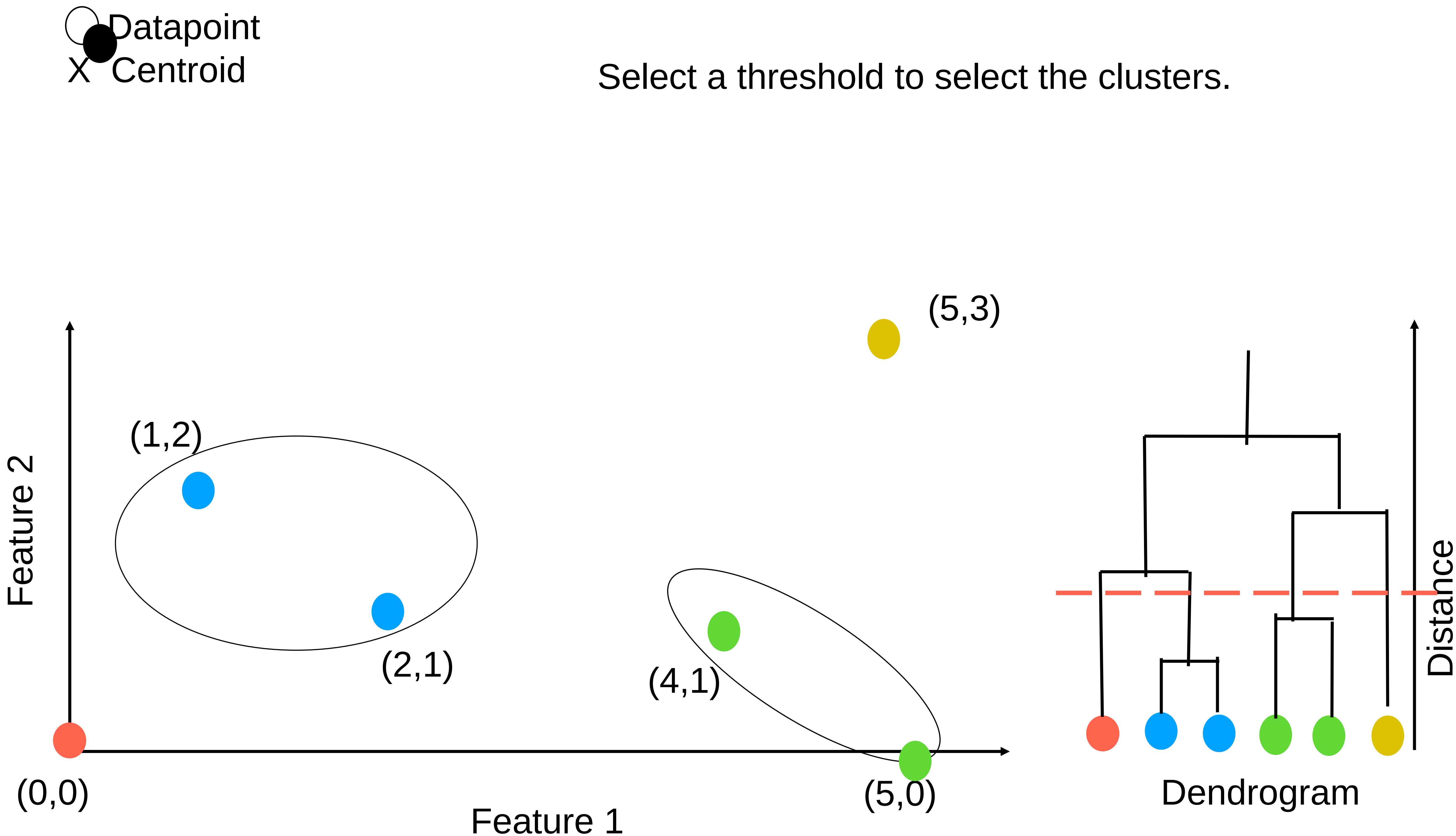
Might not make any sense to make the last hierarchical cluster (everything into one) ?



Source: Stanford – Hierarchical clustering

Convergence: When to stop ?

Select a threshold to select the clusters.

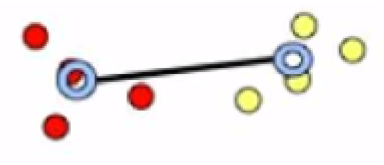


Source: Stanford – Hierarchical clustering

Did you notice what distance metric we used ?

Centroid Linkage

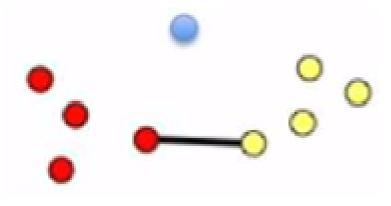
Distance between centroids (means) of two clusters



Pick clusters with minimum distance but how to calculate distance ?

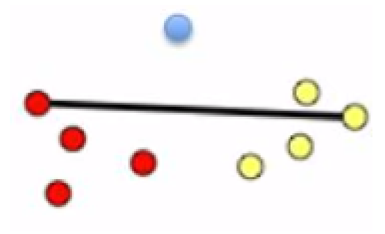
Single Linkage
Distance between closest elements in clusters

Produces chains



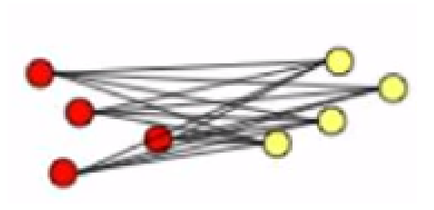
Complete Linkage
Distance between furthest elements in clusters

Forces “spherical” clusters with consistent diameter

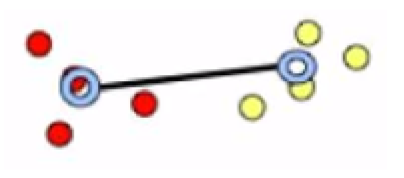


Average Linkage
Average of all pairwise distances

Less affected by outliers

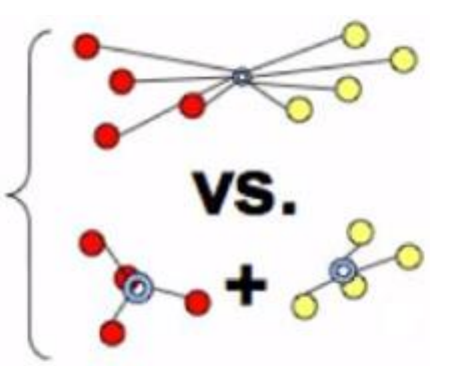


Centroid Linkage
Distance between centroids (means) of two clusters



Ward’s method
Consider joining two clusters, how does it changes the total distance (variance) from centroids

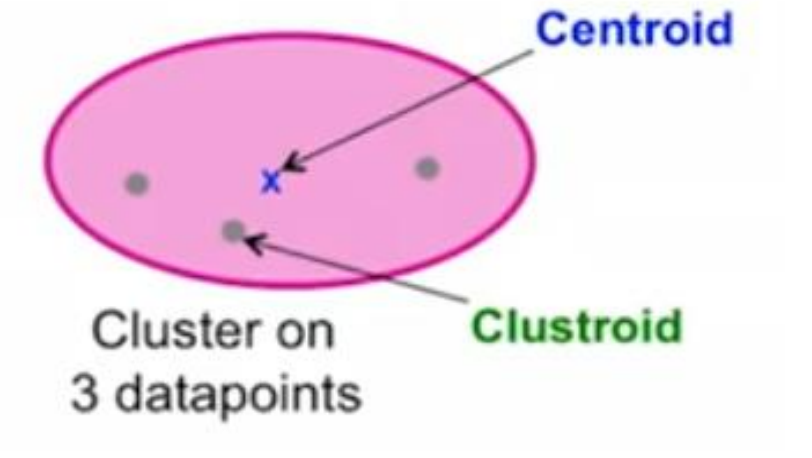
Variance is the key



Source: http://homepages.inf.ed.ac.uk/vlavrenk/

Non Euclidean space

How to represent a cluster of many points ?
Clustroid: data (point) closest to other points



How do you determine the “nearness” of clusters ?
Treat clustroid as if it were centroid when computing intercluster distances

Possible meaning of closest
Smallest maximum distance to other points
Smallest average distance to other points
Smallest sum of squares of distance to other points

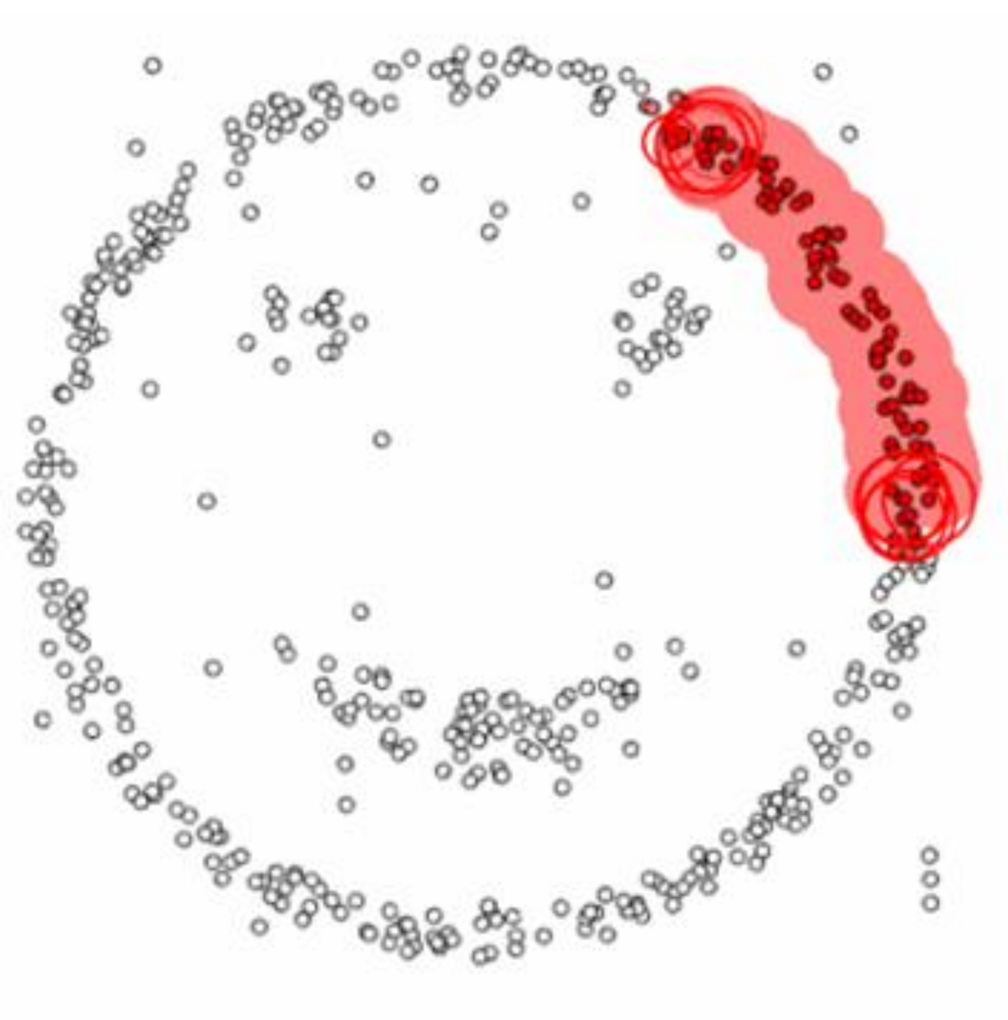
Drawbacks

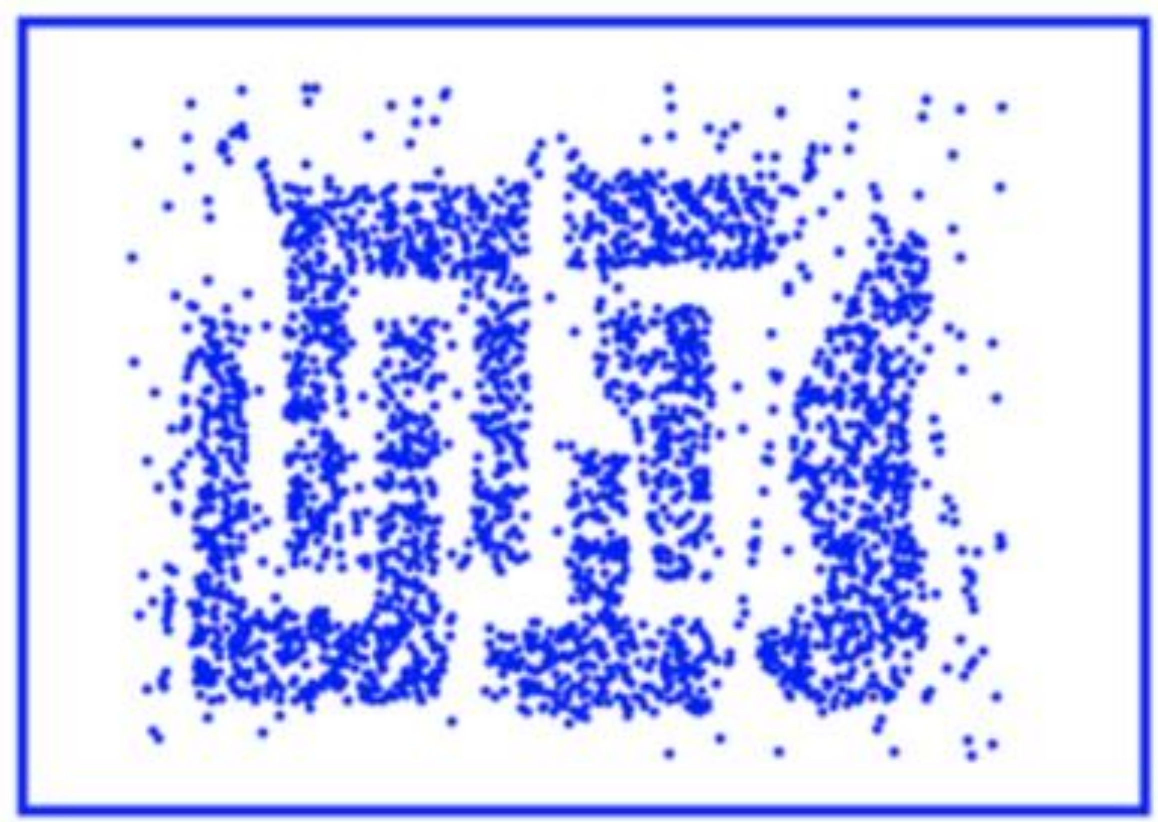
- Sensitivity to outliers
- A hierarchical structure is created even when such structure is not appropriate.
- Sometimes the dendrogram is too huge huge to infer anything meaningful

DBSCAN

Density-Based Spatial Clustering of Applications with Noise

How about following data points shapes





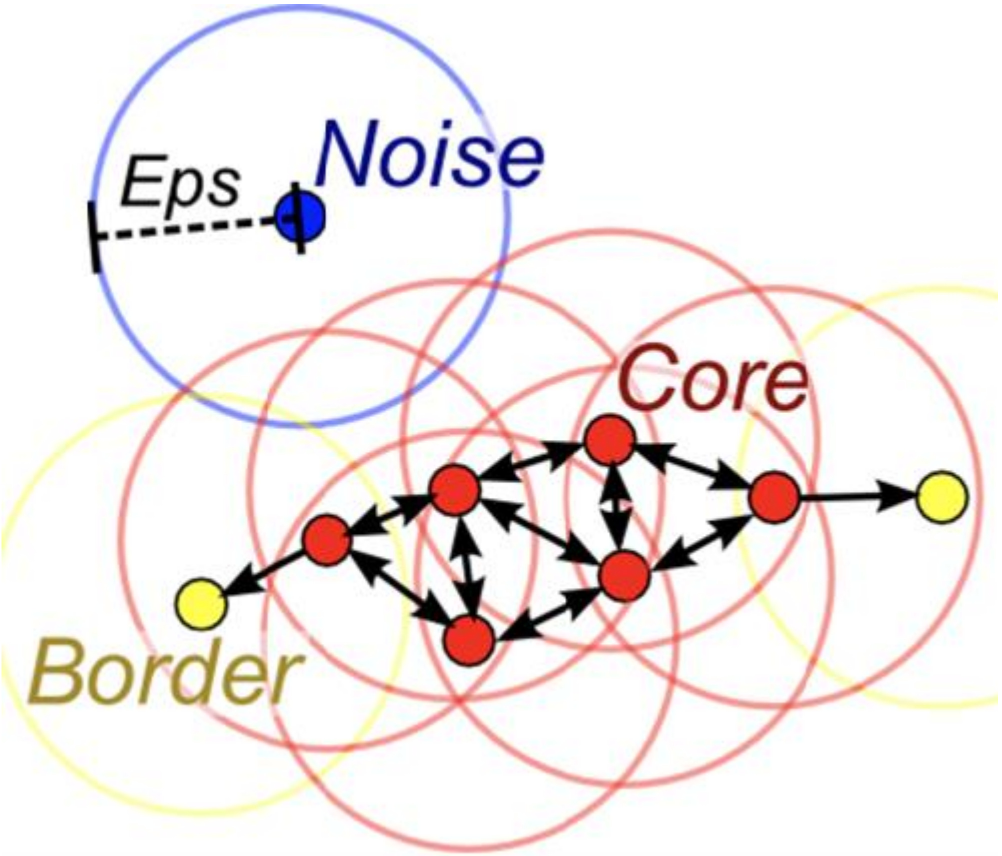
DBSCAN

Density-Based Spatial Clustering of Applications with Noise

Inputs:

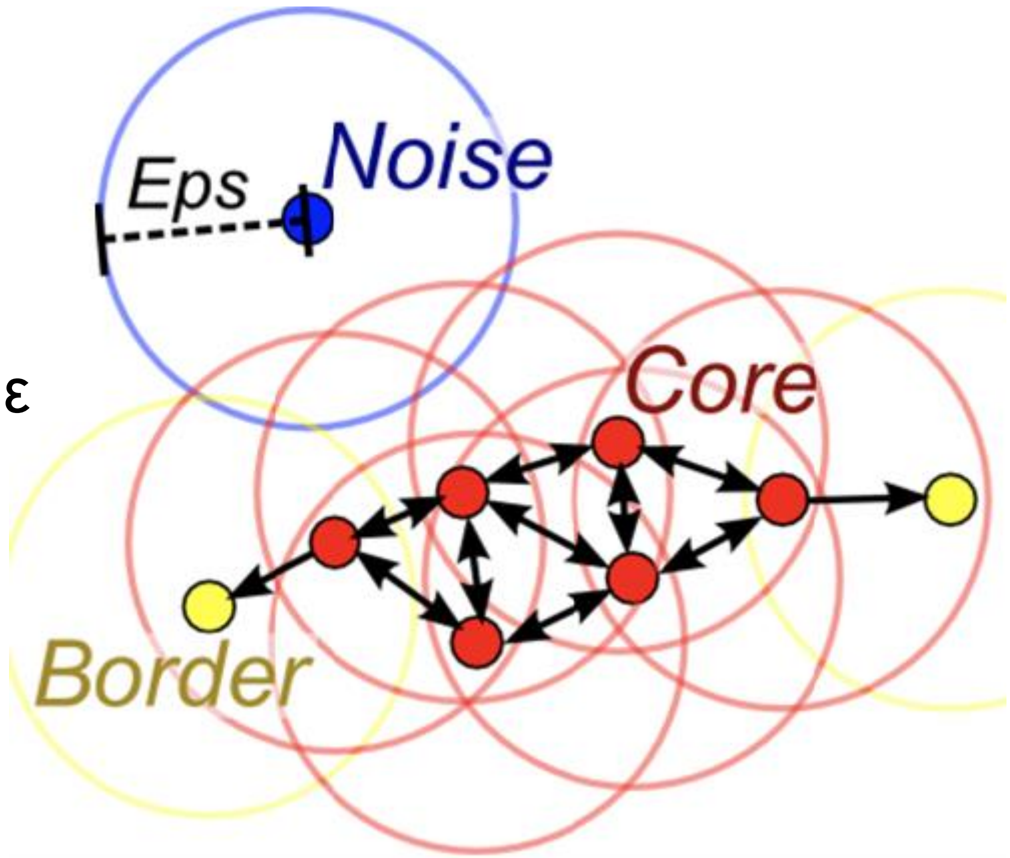
1) Radius distance/Eps (ε): Distance between two data points
   * *k-distance graph can help*

2) Minimum points: Minimum number of points which have to be within the distance so they can form a cluster.
   * *minPoints ≥ D + 1, where D is # of features*



Source:https://towardsdatascience.com/how-dbscan-works-and-why-should-i-use-it-443b4a191c80

DBSCAN



1) Select a data point that has not been visited. Neighborhood of this point is extracted using a distance ε

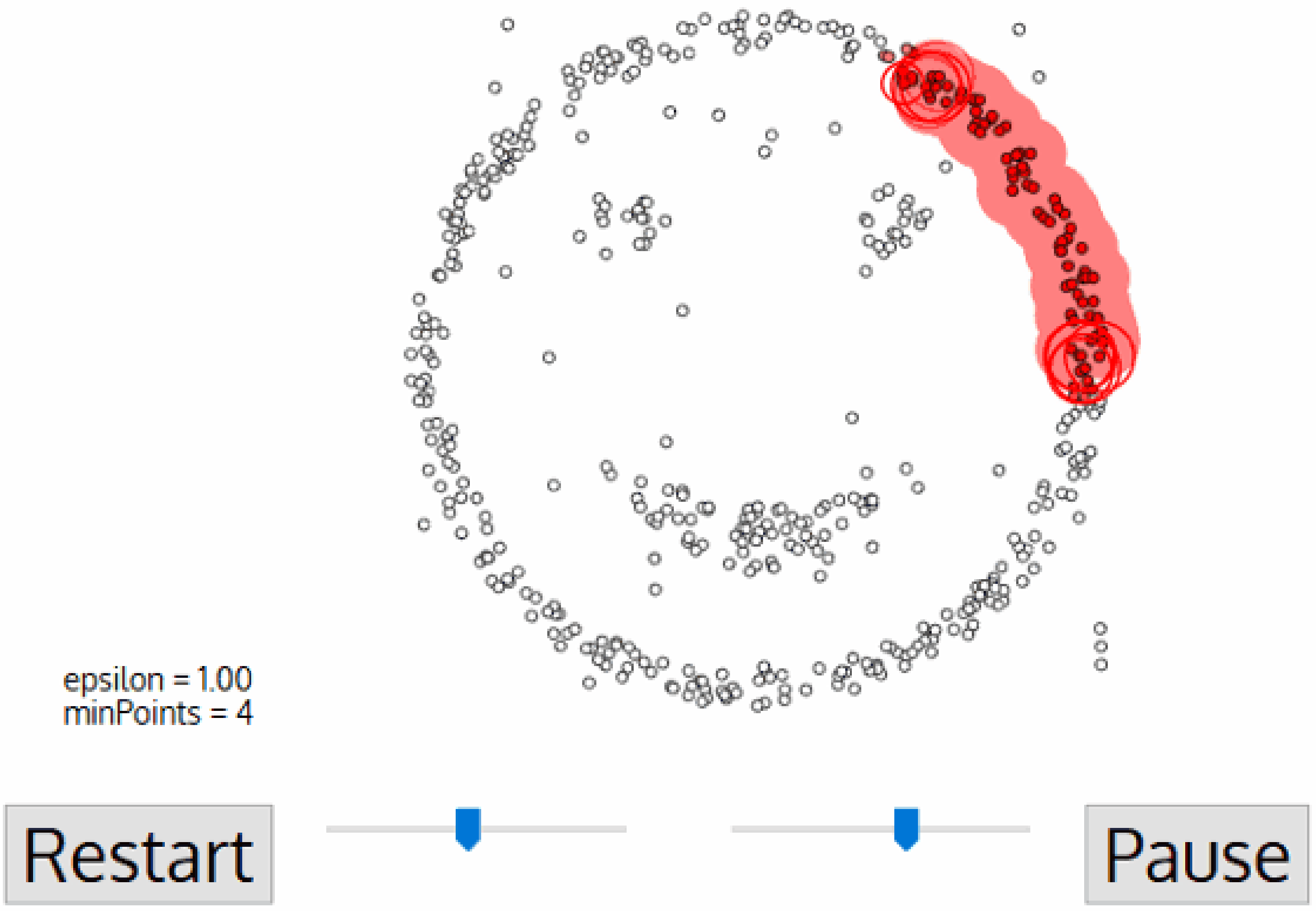
2) If there are a sufficient number of minPoints within this neighborhood, then the clustering process starts and the current data point becomes the first point in the new cluster. Otherwise, the point will be labeled as noise.

3) For this first point in the new cluster, the points within its ε distance neighborhood also become part of the same cluster. This procedure of making all points in the ε neighborhood belong to the same cluster is then repeated for all of the new points that have been just added to the cluster group.

4) steps 2 and 3 is repeated until all points in the cluster are determined

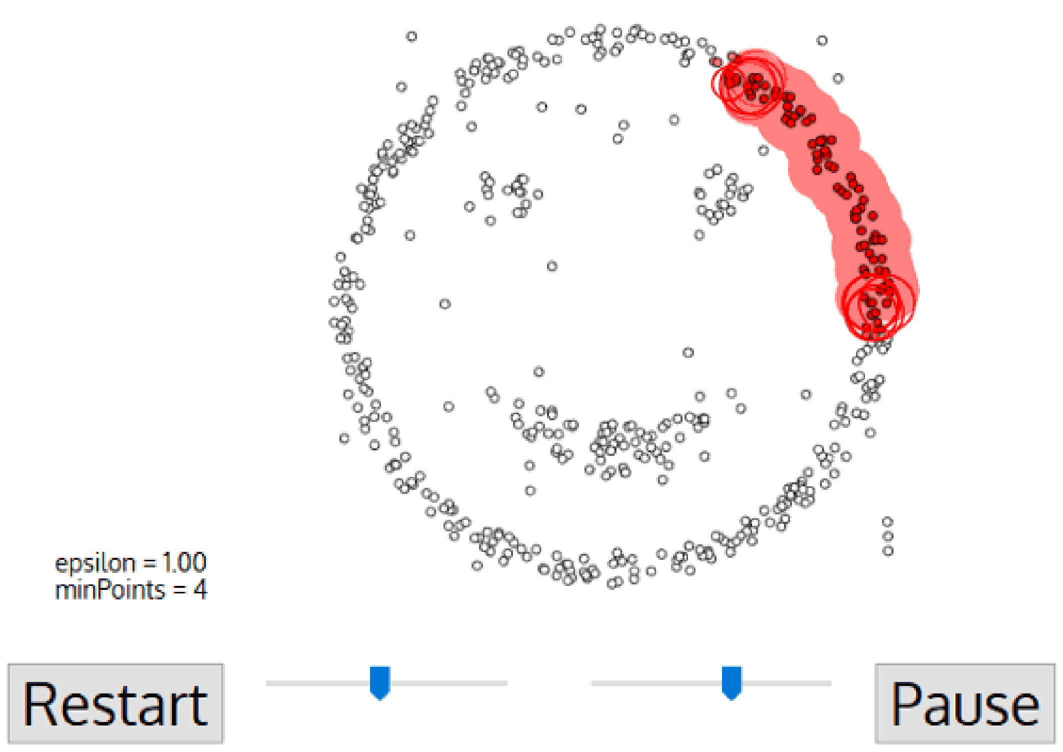
5) Once we are done with the current cluster, a new unvisited point is retrieved and processed, leading to the discovery of a further cluster or noise. This process repeats until all points are marked as visited.

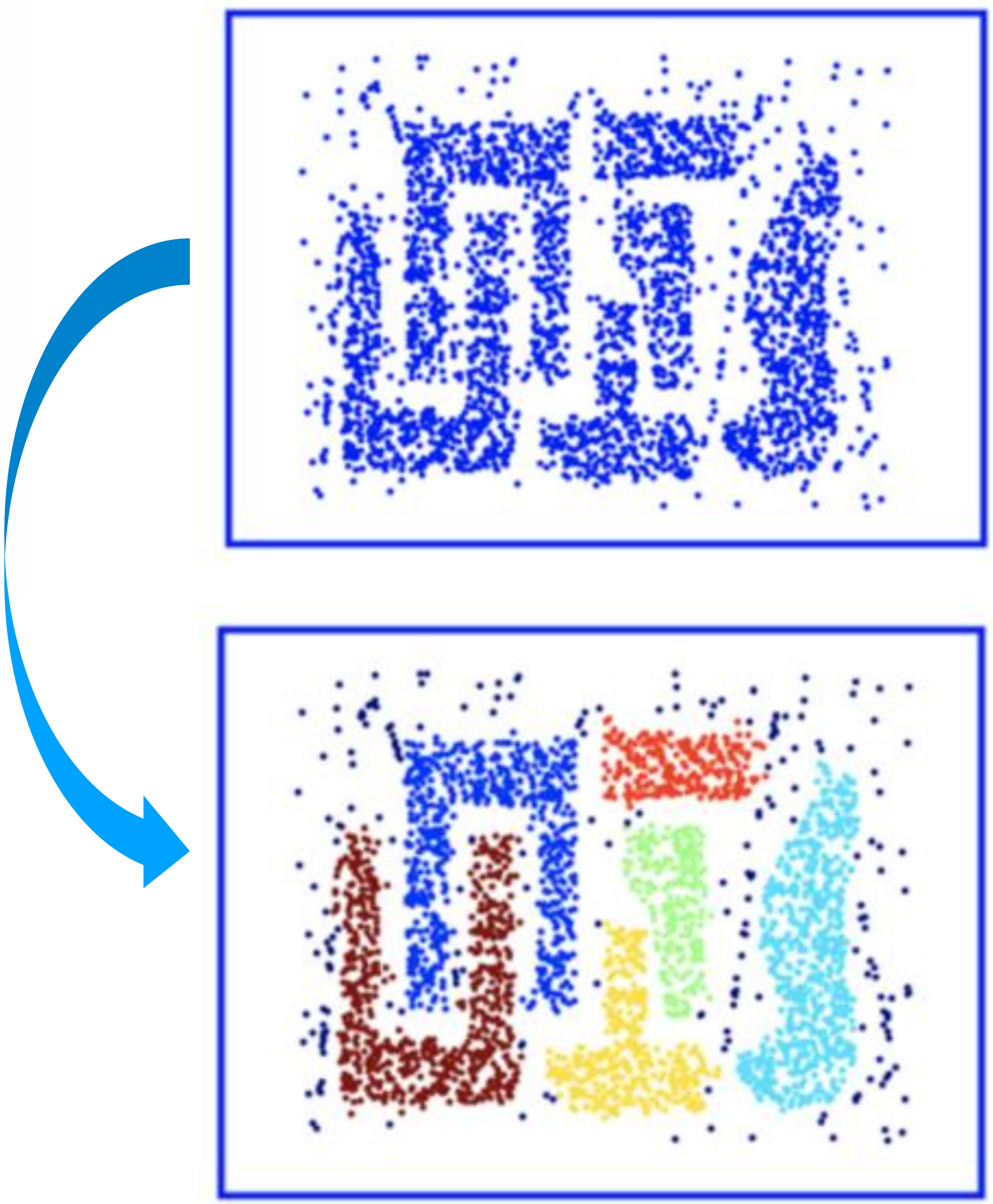
DBSCAN



Source: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

DBSCAN





DBSCAN: Pros & Cons

Advantages:

1) Great at separating clusters of high density versus clusters of low density within a given dataset.
2) Is great with handling outliers within the dataset.

Disadvantages:

1) Does not work well when dealing with clusters of varying densities.
2) Struggles with high dimensionality data*.

More on DBSCAN

https://towardsdatascience.com/dbscan-clustering-for-data-shapes-k-means-cant-handle-well-in-python-6be89af4e6ea

*** Isolation Forest and Robust Random Cut Forest are two algorithms which works good for high dimensionality data.**

3 Important Questions

1. How do you represent a cluster of more than one point?
   centroid or clusteroid represents a set of points

2. How do you determine the "nearness" of clusters?
   Some Distance metric

3. When to stop combining clusters
   convergence (k-means, GMM) or some threshold (Hierarchical) or cohesion (DBSCAN)

Clustering on large scale datasets

Mini Batch K-means:

The algorithm takes small randomly chosen batches of the dataset for each iteration

Each data in the batch is assigned to the clusters, depending on the previous locations of the cluster centroids.

It then updates the locations of cluster centroids based on the new points from the batch.

Please note:

As the number clusters and the number of data increases, the relative saving in computational time also increases.

The saving in computational time is more noticeable only when the number of clusters is very large.

Increasing the number of clusters, decreases the similarity of the mini batch K-means solution to the K-means

Source: https://www.geeksforgeeks.org/ml-mini-batch-k-means-clustering-algorithm/

More on this: https://upcommons.upc.edu/bitstream/handle/2117/23414/R13-8.pdf

(Internal) Validation
Silhouette coefficient/index/value

For data point $i \in C_i$ (data point i in the cluster C_i), let

$$a(i) = \frac{1}{|C_i| - 1} \sum_{j \in C_i, i \neq j} d(i, j)$$

mean dissimilarity of point i to some cluster C_k (where $C_k \neq C_i$).

$$b(i) = \min_{k \neq i} \frac{1}{|C_k|} \sum_{j \in C_k} d(i, j)$$

and

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \text{ if } |C_i| > 1$$

From the above definition it is clear that

$$-1 \leq s(i) \leq 1$$

https://en.wikipedia.org/wiki/Silhouette_(clustering)

External Validation will be discussed at the end of Lecture about (Classification) Evaluation metrics

Other clustering methods

- Clustering Algorithms
  - Average-linkage clustering: a simple agglomerative clustering algorithm
  - Canopy clustering algorithm: an unsupervised pre-clustering algorithm related to the K-means algorithm
  - Complete-linkage clustering: a simple agglomerative clustering algorithm
  - DBSCAN: a density based clustering algorithm
  - Expectation-maximization algorithm
  - Fuzzy clustering: a class of clustering algorithms where each point has a degree of belonging to clusters
    - Fuzzy c-means
    - FLAME clustering (Fuzzy clustering by Local Approximation of MEmberships): define clusters in the dense parts of a dataset and perform cluster assignment solely based on the neighborhood relationships among objects
  - KHOPCA clustering algorithm: a local clustering algorithm, which produces hierarchical multi-hop clusters in static and mobile environments.
  - k-means clustering: cluster objects based on attributes into partitions
  - k-means++: a variation of this, using modified random seeds
  - k-medoids: similar to k-means, but chooses datapoints or medoids as centers
  - Linde–Buzo–Gray algorithm: a vector quantization algorithm to derive a good codebook
  - Lloyd's algorithm (Voronoi iteration or relaxation): group data points into a given number of categories, a popular algorithm for k-means clustering
  - OPTICS: a density based clustering algorithm with a visual evaluation method
  - Single-linkage clustering: a simple agglomerative clustering algorithm
  - SUBCLU: a subspace clustering algorithm
  - Ward's method : an agglomerative clustering algorithm, extended to more general Lance–Williams algorithms
  - WACA clustering algorithm: a local clustering algorithm with potentially multi-hop structures; for dynamic networks

source:wikipedia

Additional examples:

http://inseaddataanalytics.github.io/INSEADAnalytics/BoatsSegmentationCaseSlides.pdf

1 Novices

% Total Population: 17%
% US Population: 18%
% Brazil Population: 6%
% Canada Population: 18%

Who they are
- Rely more on expert opinion than their own
- Boating helps them escape from everyday life and relax
- Boating gives me a feeling of adventure

Who they are not
- Not considered knowledgeable about boating
- Boating is not their true life passion
- Boating is not the #1 activity they do in their spare time

2 No Frills

% Total Population: 23%
% US Population: 29%
% Brazil Population: 5%
% Canada Population: 19%

Who they are
- Functionality is more important than style
- Perform repairs and maintenance on their boats
- Tend to prefer a boat with little to no frills

Who they are not
- Do not go for the latest and greatest boat
- Having a powerful boat is not as important to them
- Do not see the boat brand as saying a lot about who they are

3 Status Seekers

% Total Population: 20%
% US Population: 19%
% Brazil Population: 29%
% Canada Population: 18%

Who they are
- Willing to pay a premium for a brand with a reputation for high quality
- Buy the latest and greatest boats
- View their boat as a status symbol

Who they are not
- Do not choose functionality over style
- Do not prefer a basic boat with little to no thrills
- Do not perform repairs and maintenance on their boats

4 Active Family Boaters

% Total Population: 9%
% US Population: 10%
% Brazil Population: 7%
% Canada Population: 10%

Who they are
- Boating helps them stay active
- Boating allows them to excel in sports they're passionate about
- Boating gives me an outlet to socialize with family and/or friends

Who they are not
- The lowest price is not more important than boat brand
- Do not prefer a basic boat with little to no frills
- Do not rely on expert opinion other than their own

5 Price driven Lifestylers

% Total Population: 31%
% US Population: 25%
% Brazil Population: 53%
% Canada Population: 33%

Who they are
- Boating is their true passion in life
- Consider themselves more knowledgeable than their boating peers
- Boating is the #1 activity they do in their spare time

Who they are not
- Boating is not a means to escape from everyday life and relax
- Boating does not provide them with a sense of adventure
- They do not consider owning a boat as a way of rewarding themselves for hard work

SEGMENT 1:
BOATING DNA

Young, adventurous, active. This segment considers boating as part of their identity, and wants a boat they can customize — the better for entertaining and fishing on their own terms. They do extensive research on social media sites as well as with more traditional means like brochures and pro angler guides. While these boaters are price-conscious, they make up the most valuable segment of our population.

SEGMENT 2:
ACTIVE SOCIAL BOATERS

For Active Social Boaters, being on the water is not the time for relaxing; it's all about the activities. This segment uses their boat to engage in as many activities as possible with their spouse and kids. They look for durable boats that can handle frequent usage, and keep their precious cargo safe at the same time. Active Social Boaters prefer a dealer that will teach them how to operate and maintain their jet boat or cruiser. That way, this segment has the know-how to keep their boat performing at its best.

SEGMENT 3:
IMAGE CONSCIOUS BOATERS

Image Conscious Boaters love to cruise on their boat and entertain their significant others, family, and friends. A professional-grade, cutting edge, and prestigious boat shows the world that they have arrived. While they have owned boats in the past and consider themselves at an intermediate level, they prefer a trustworthy dealer who gives them the attention they deserve.

SEGMENT 4:
FUNCTION-FIRST BOATERS

To the Function-First Boaters segment, the most important element of a boat is that it works well. They steer clear of gadgets and accessories that get in the way of their enjoying the outdoors. They like to perform minor repairs and maintenance themselves on their boat, both because they enjoy it and because they want to get it back on the water. They tend toward sturdy boats that are durable and a good value.

SEGMENT 5:
CASUAL BOATERS

When it comes to boating, these consumers want their purchase to feel safe and easy to use. It's important that they buy a reputable brand from a knowledgeable salesperson. For Casual Boaters, boating is all about cruising with their families. They research boats online, and ultimately make the final purchase decision with their spouses.

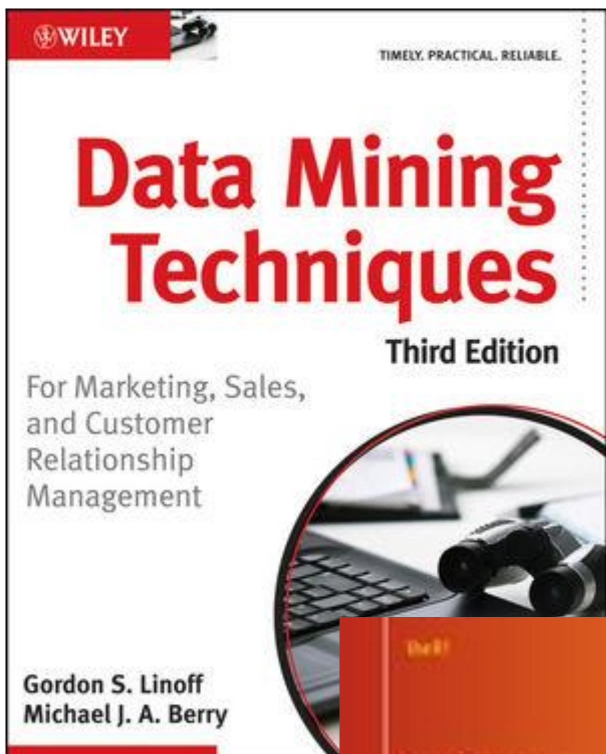
Summary

Techniques for Customer Segmentation

- Historical/Behavior based
  - RFM
- Data Driven
  - K-Means: How to Pick K ? Elbow Method
  - Hierarchical Clustering: Different Distance metrics
  - DBSCAN

Books and links

http://proquestcombo.safaribooksonline.com.ezproxy.utlib.ut.ee/book/databases/business-intelligence/9780470650936



About RFM:
https://www.putler.com/rfm-analysis/

About segmentation in B2B setting:

http://labs.openviewpartners.com/files/2012/10/Customer-Segmentation-eBook-FINAL.pdf