

Clustering

How we deal with unknown data - A Lesson



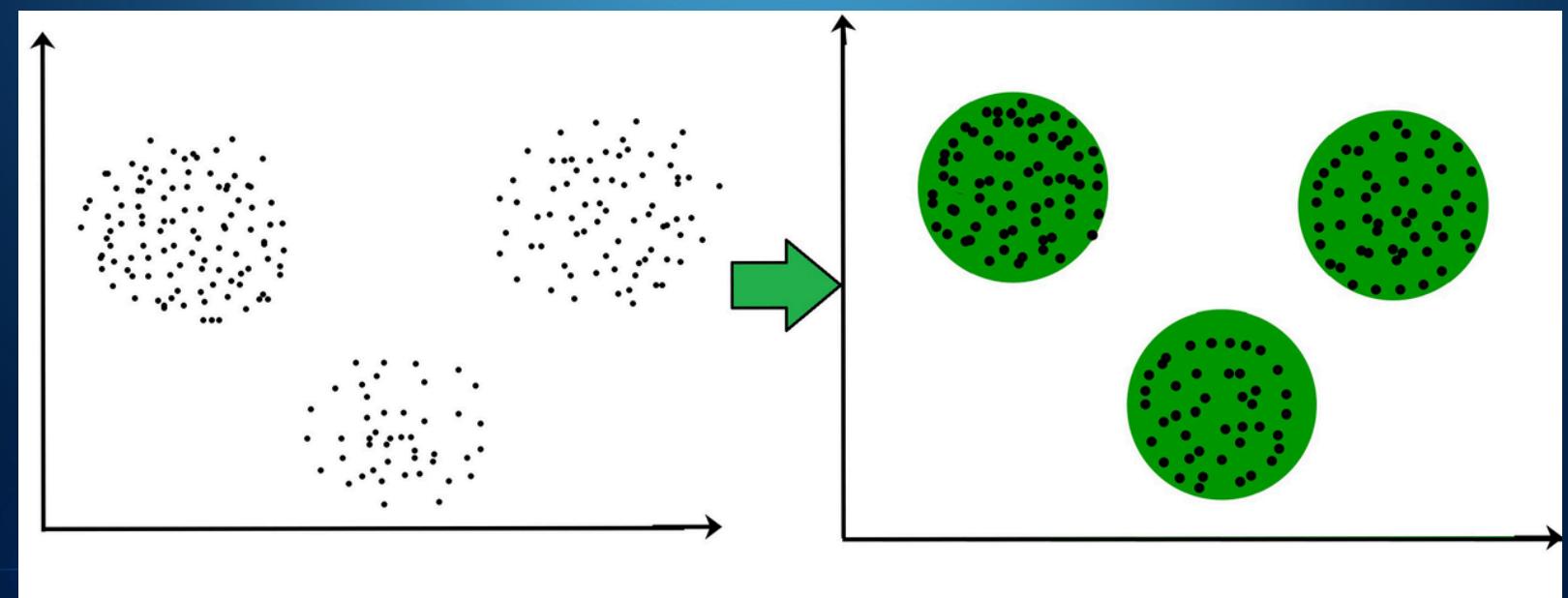
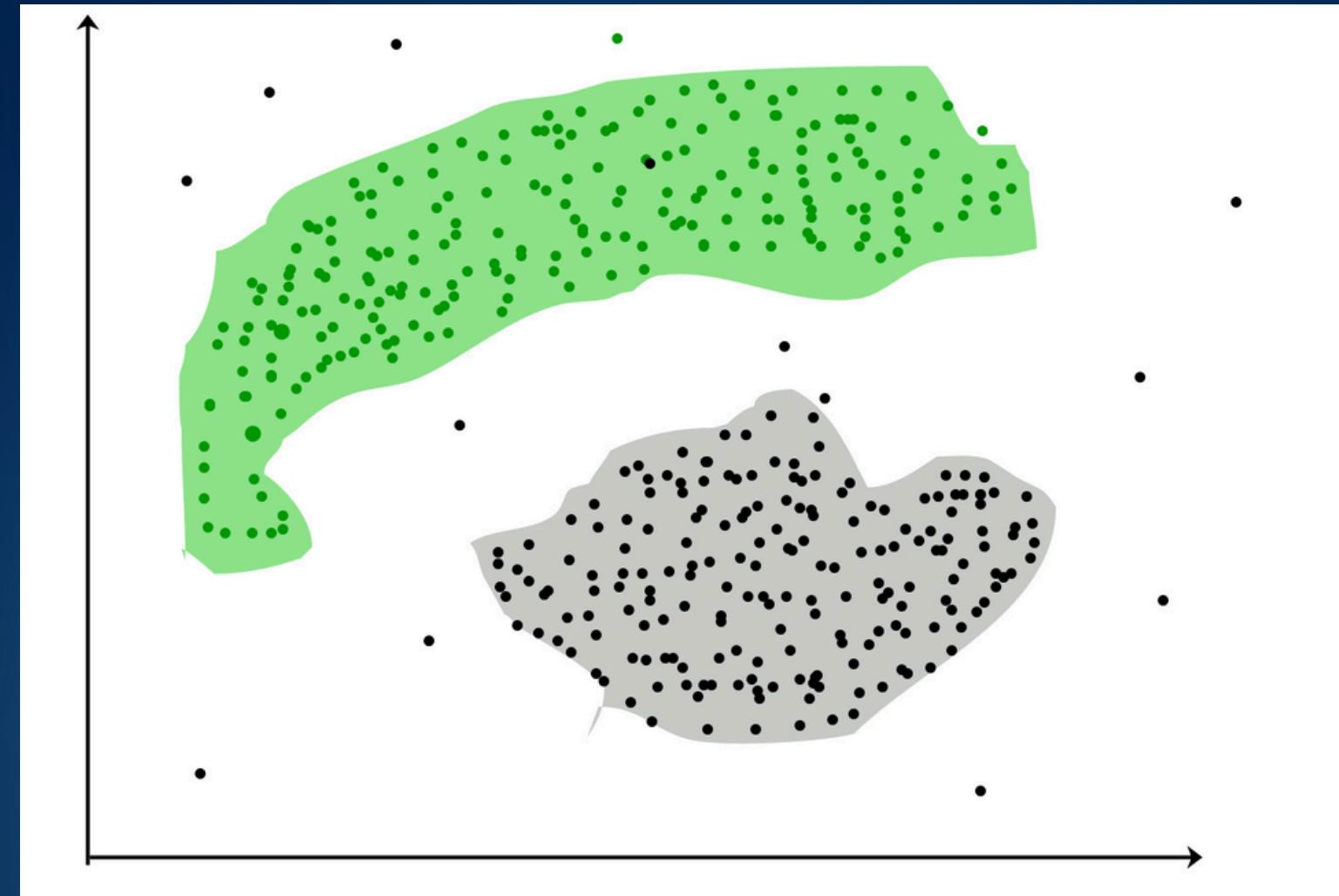


Objectives for learning

1. What is Clustering?
2. How does it work?
3. Where can you apply it?

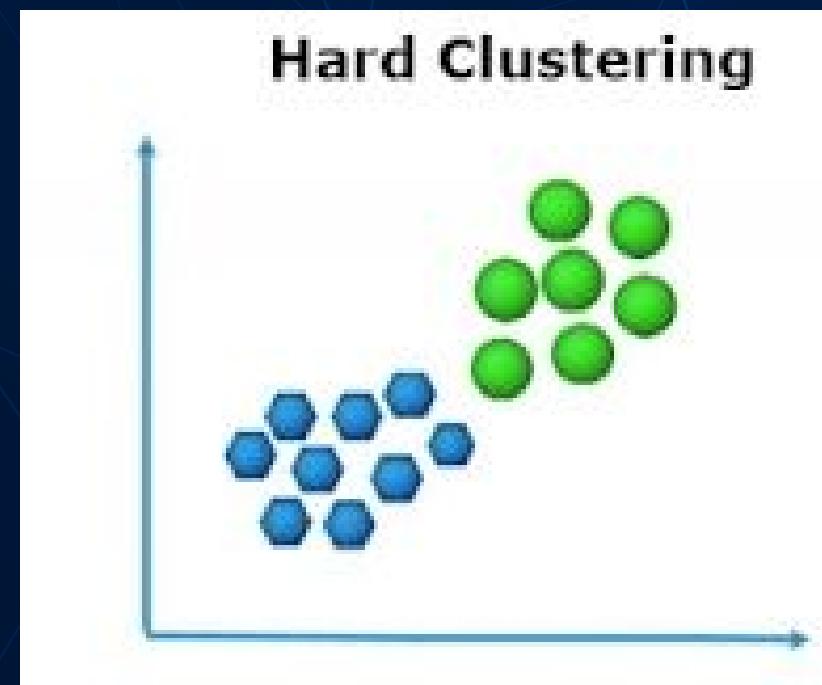
What is Clustering?

- Working with unlabelled data points
- Grouping homogenous data (data that has similarity) points from heterogenous data
- These similarities are found using metrics (e.g. Euclidean distance)
- Shape of clusters can be irregular



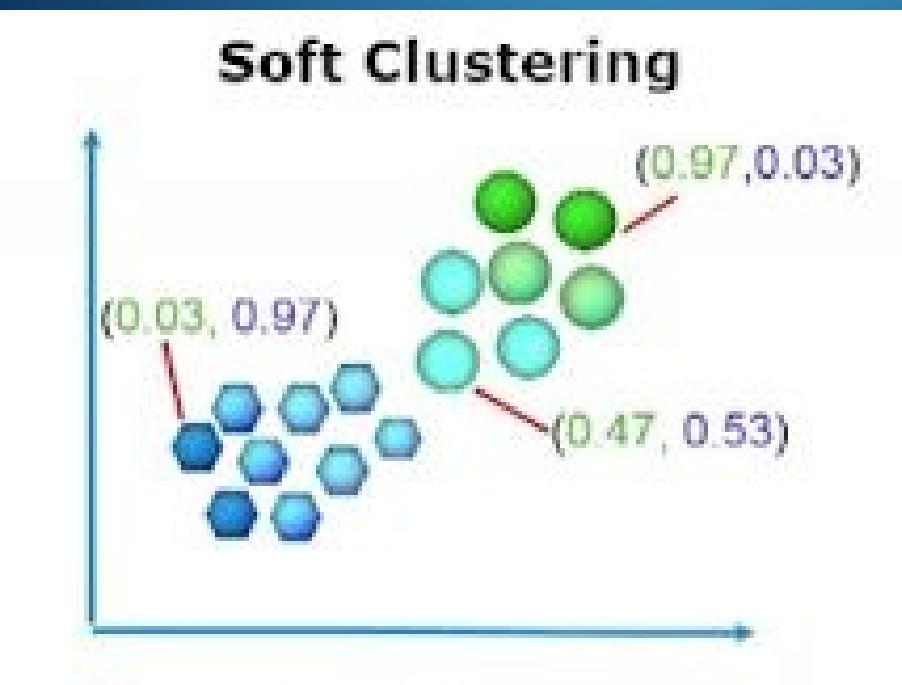
Hard Clustering

- Deciding whether a data point is in a cluster or not
- If there are 8 data points & 2 clusters
- Data points are either in Cluster 1 or 2



Soft Clustering

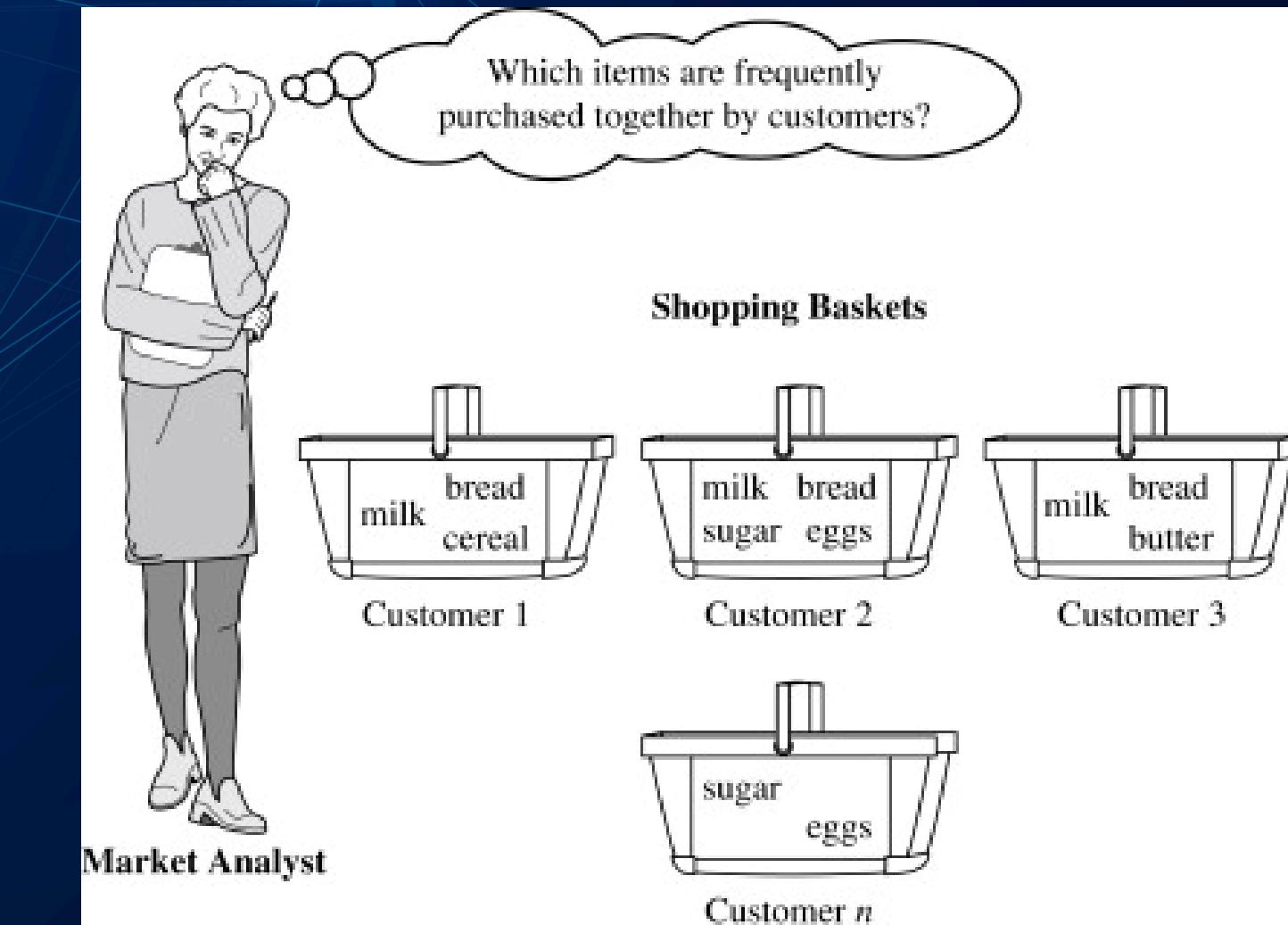
- Assigning a probability value to each data point
- The probability states the likelihood that a data point is in Cluster 1 or 2



Where is it used?



Market
Segmentation



Market Basket
Analysis

Where is it used?

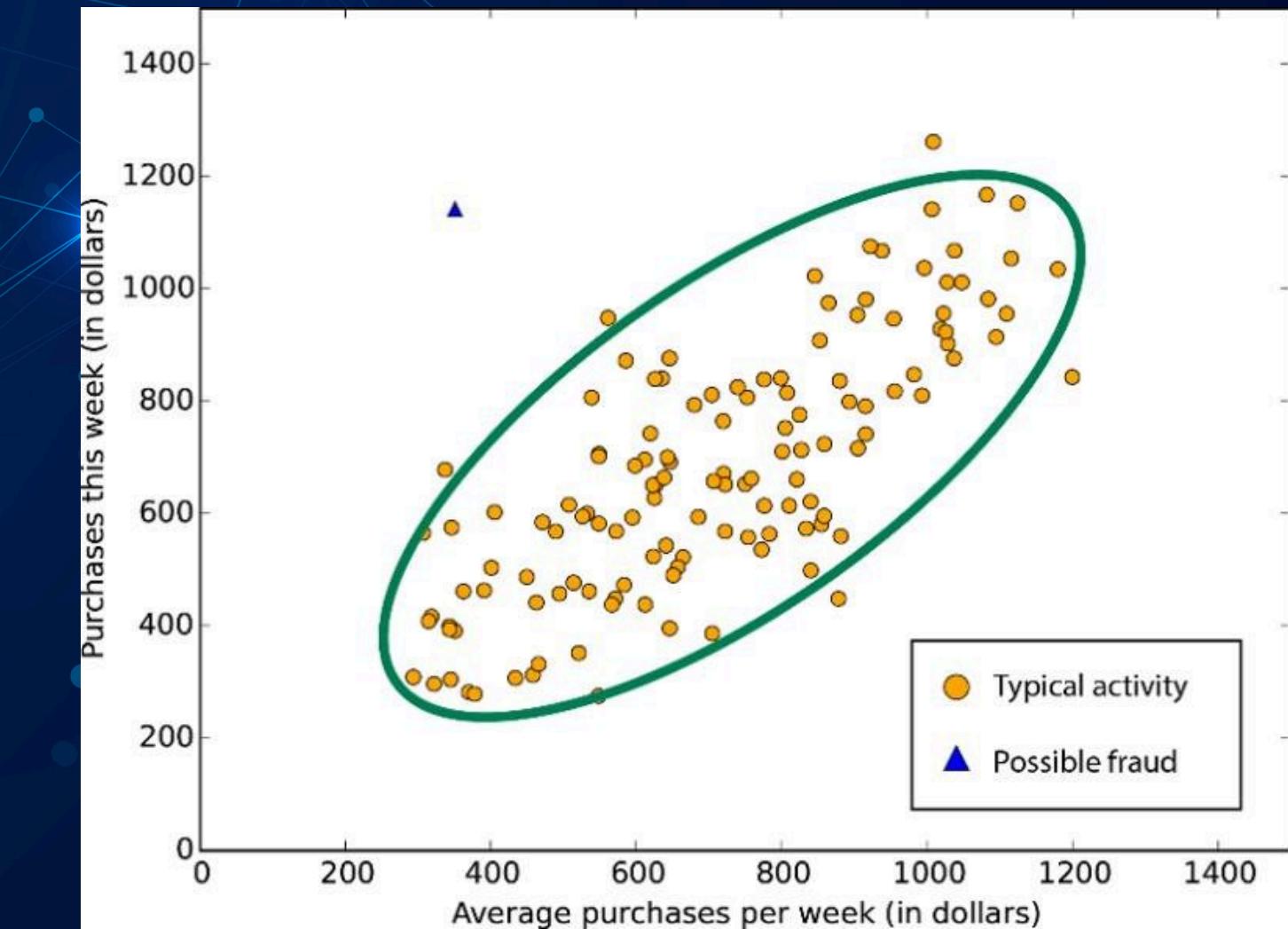


Browsing History



Content Recs

**Social Network
Analysis**



**Anomaly
Detection**

How does it work?



Get dataset and
clean it

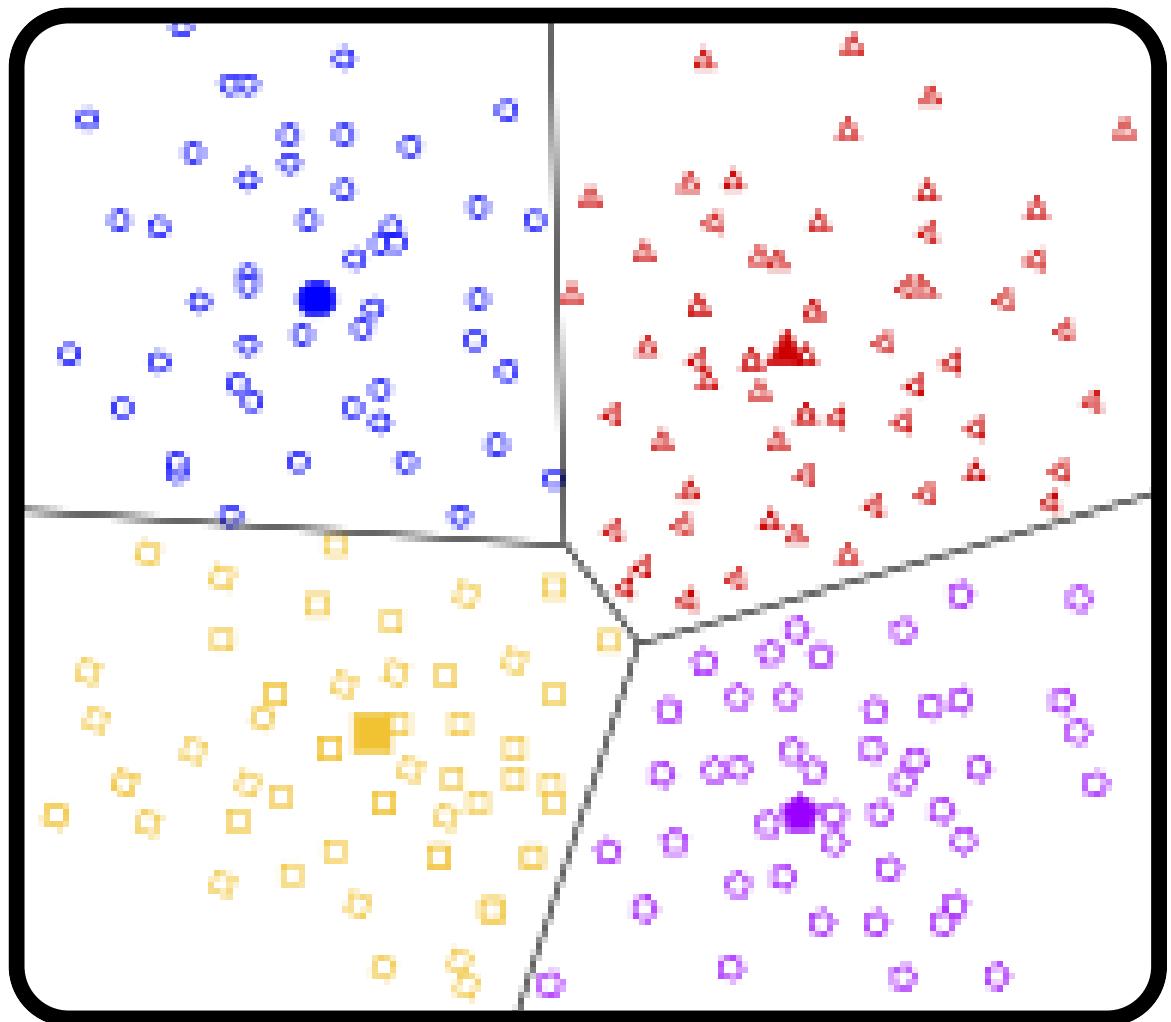


Choose a clustering
algorithm and apply
it.

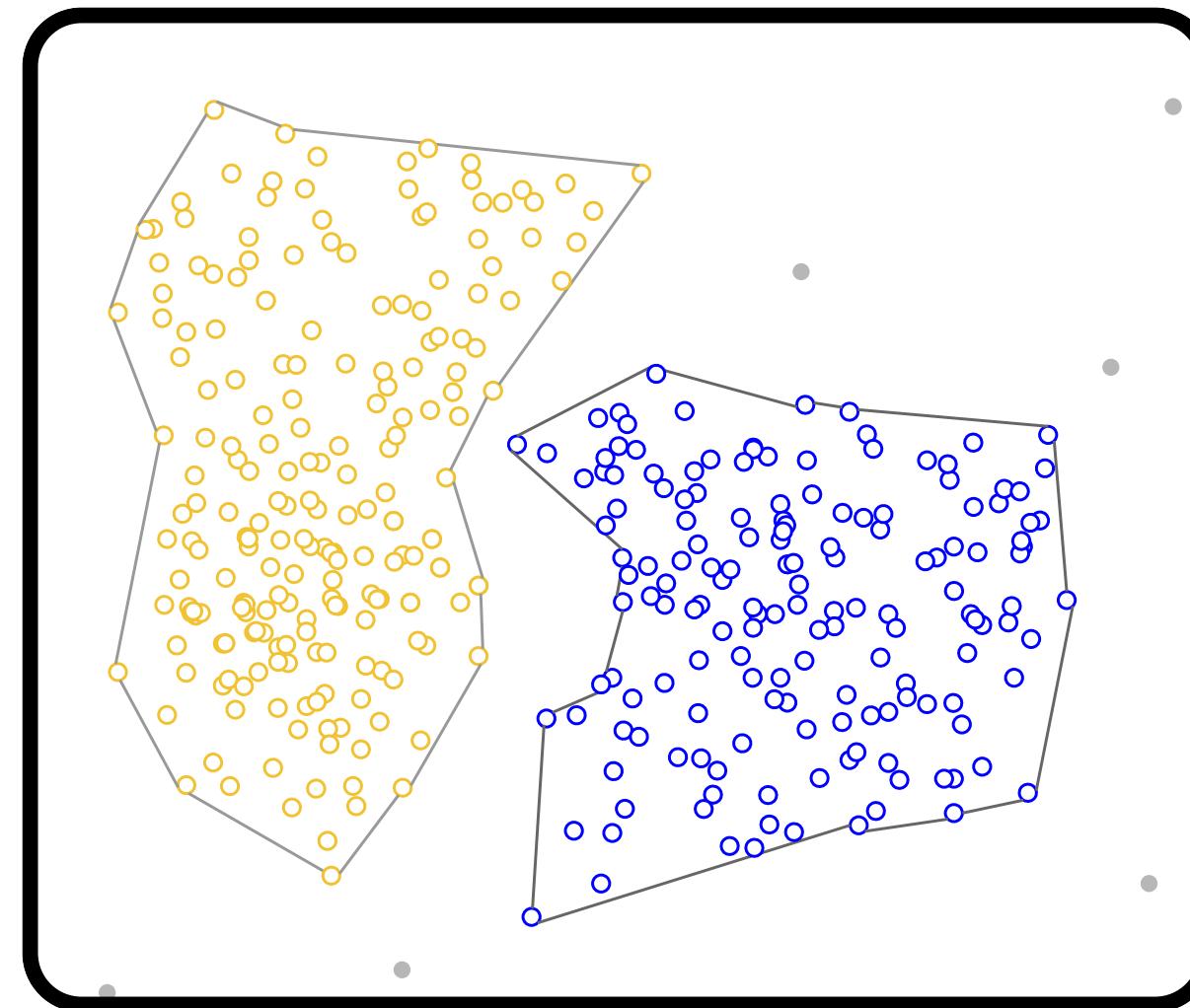


Look at the result
and create
recommends!

Types of Clustering

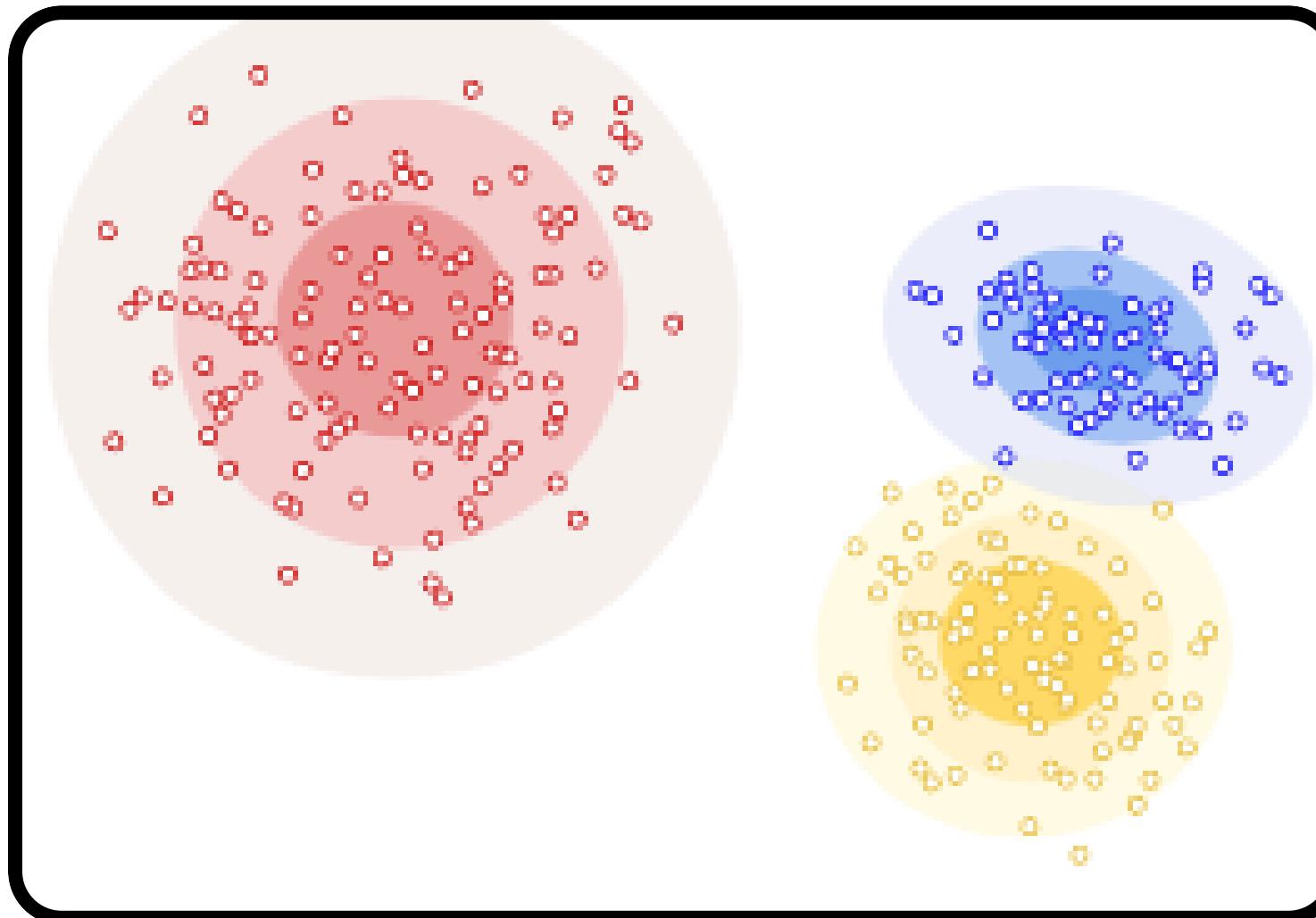


**Centroid-based
Clustering**

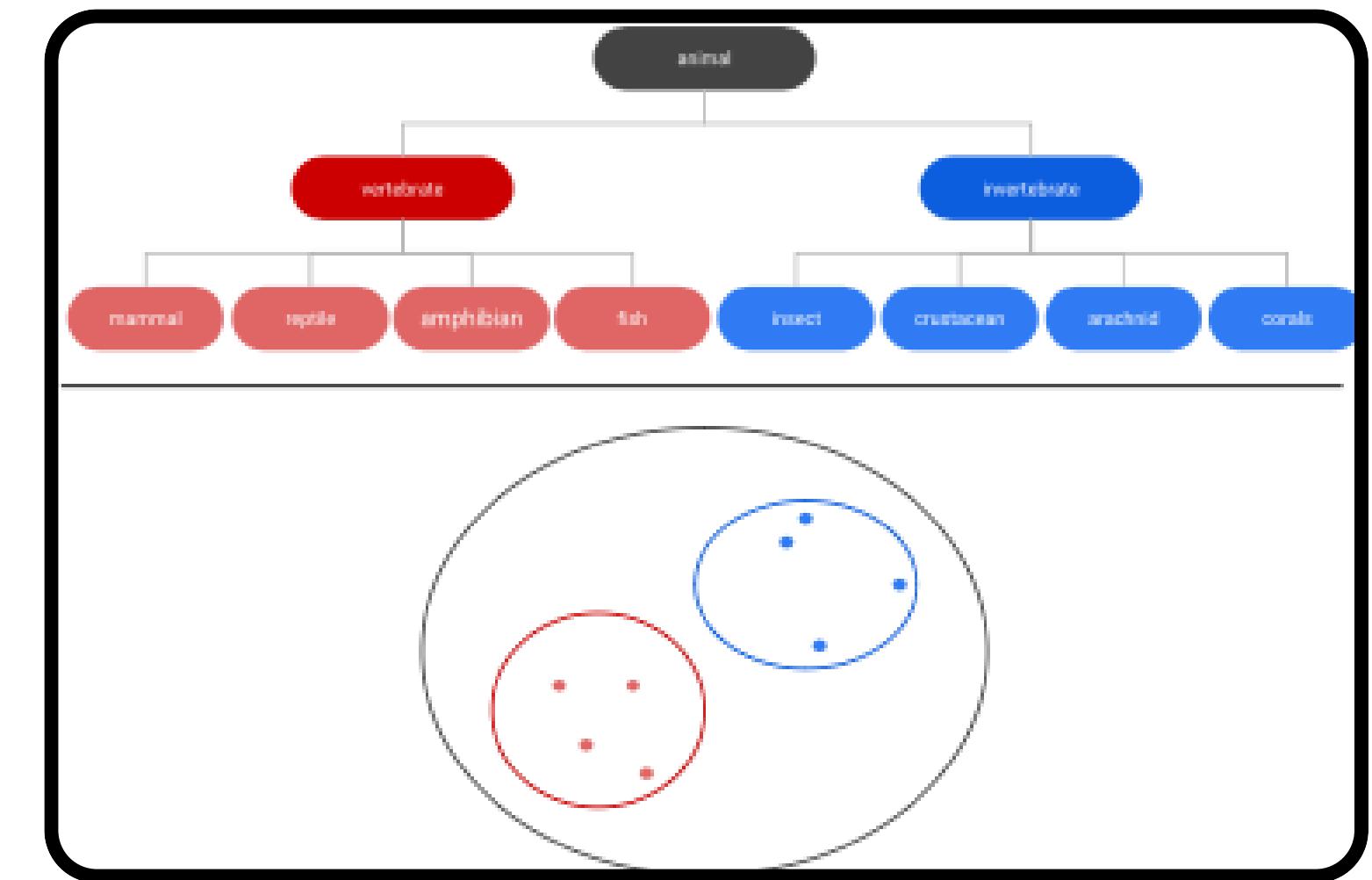


**Density-based
Clustering**

Types of Clustering



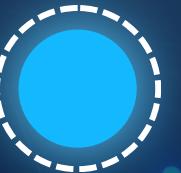
**Distribution-based
Clustering**



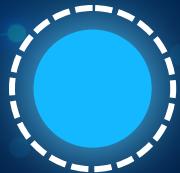
Hierarchical Clustering

**next: Continue
Working on
Project!**

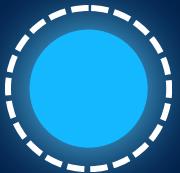
Invention of the Internet



In the 1960s, the United States government, businesses, and colleges collaborated to develop a system that would allow computers across the country to share information

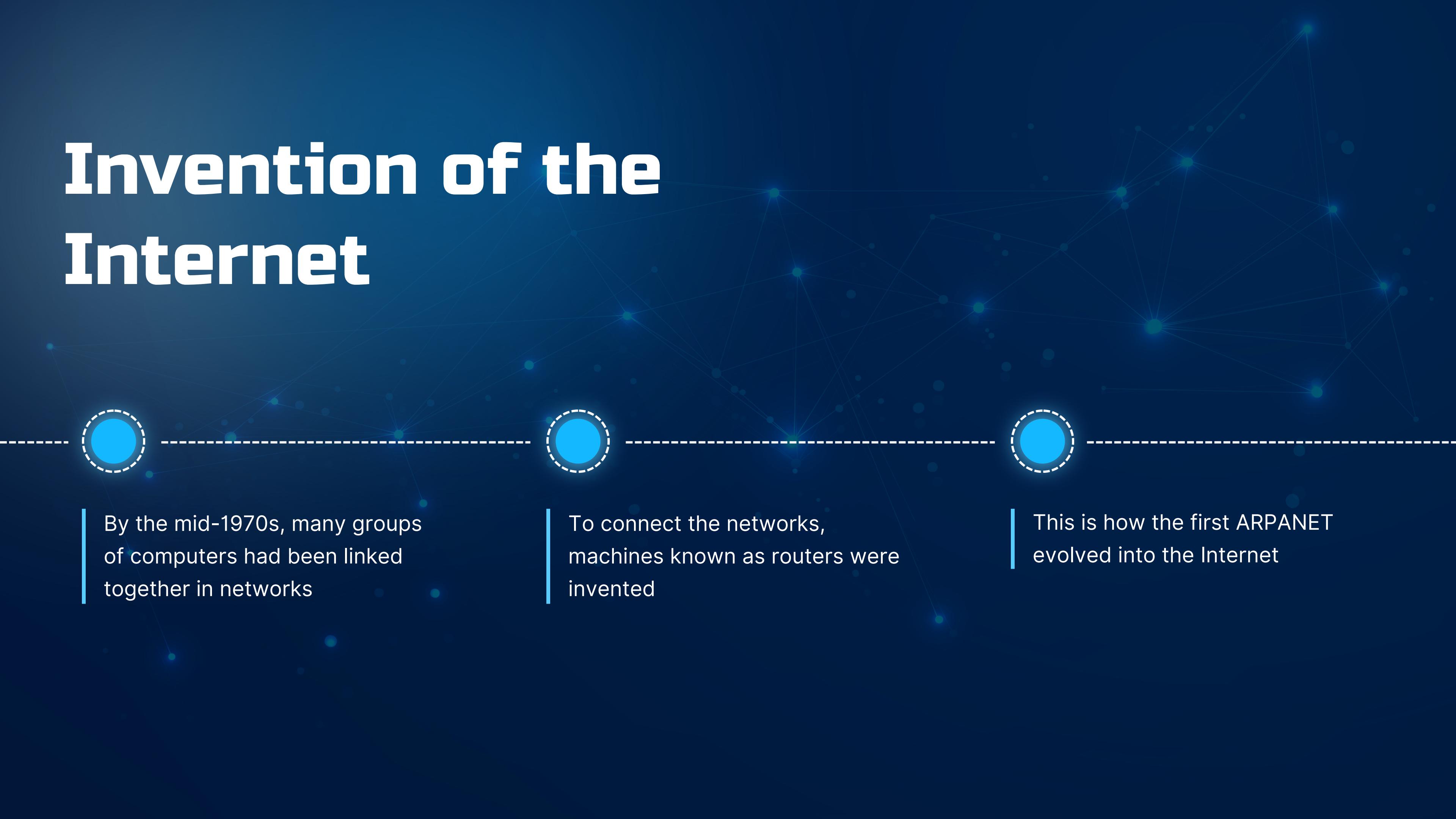


In 1969, they developed ARPANET, an early form of the Internet



Electronic mail, or e-mail, was invented in 1971 as a method of sending a message from one computer to another

Invention of the Internet



By the mid-1970s, many groups of computers had been linked together in networks

To connect the networks, machines known as routers were invented

This is how the first ARPANET evolved into the Internet

Electronic Mail (Email)



The sending and receiving of email is the second most common use of the internet

Email is a private communication medium that is sent from one user to another

Instant messaging (such as AIM or ICQ) is similar to email, but it allows two or more people to communicate much more quickly



Recommended Websites for Drawing

Art for Kids

Lesson plans for drawing, painting, printmaking, multicultural art, recycling projects, and coloring pages are included.

Paint By Me

Has over 400 drawing tutorials.
You can print them and learn how to draw

Fun Drawing

Learn interesting facts about works of art, travel back in time, and get ideas for your own creative projects.

Jiggy Jigsaw

Play puzzles in the largest collection of free challenging puzzle games online!

Twisty Blocks

Who doesn't enjoy a good game of blocks? For hours on end, solve the addictive shape-fitting puzzle.

Criss Cross Words

Boost your vocabulary while playing the daily crossword puzzle



Recommended Websites for Puzzle



Recommended Websites for Fun Games

Bike Rush

Discover a selection of our bike games, ranging from motorcycles and dirt bikes to mountain trials and stunt bikes

Is It There?

Look for hidden objects, solve puzzles, and collect items to advance.

Happy Mail

While attempting to make package deliveries, players must avoid objects on the road and other drivers.