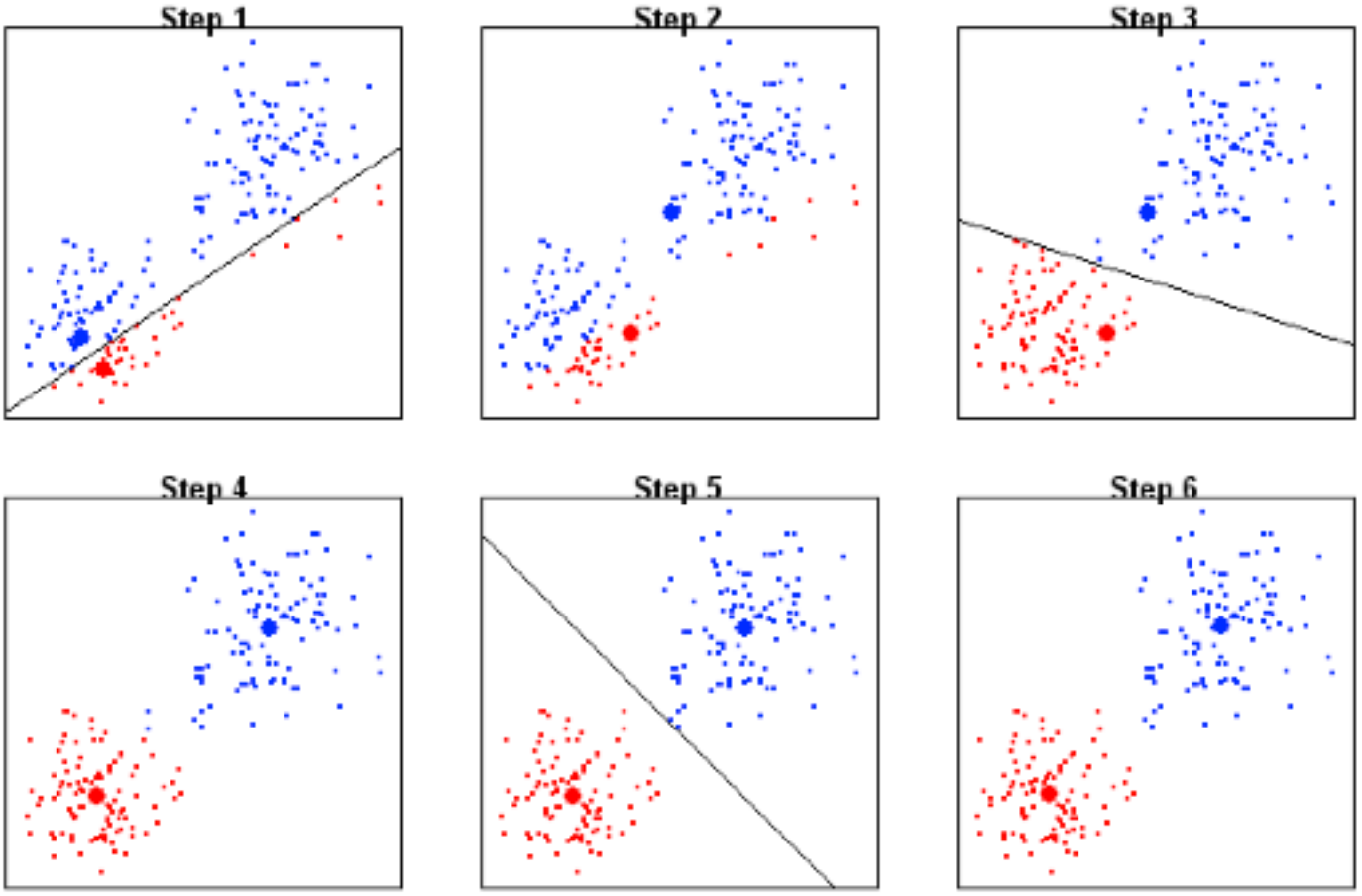
- Unsupervised learning

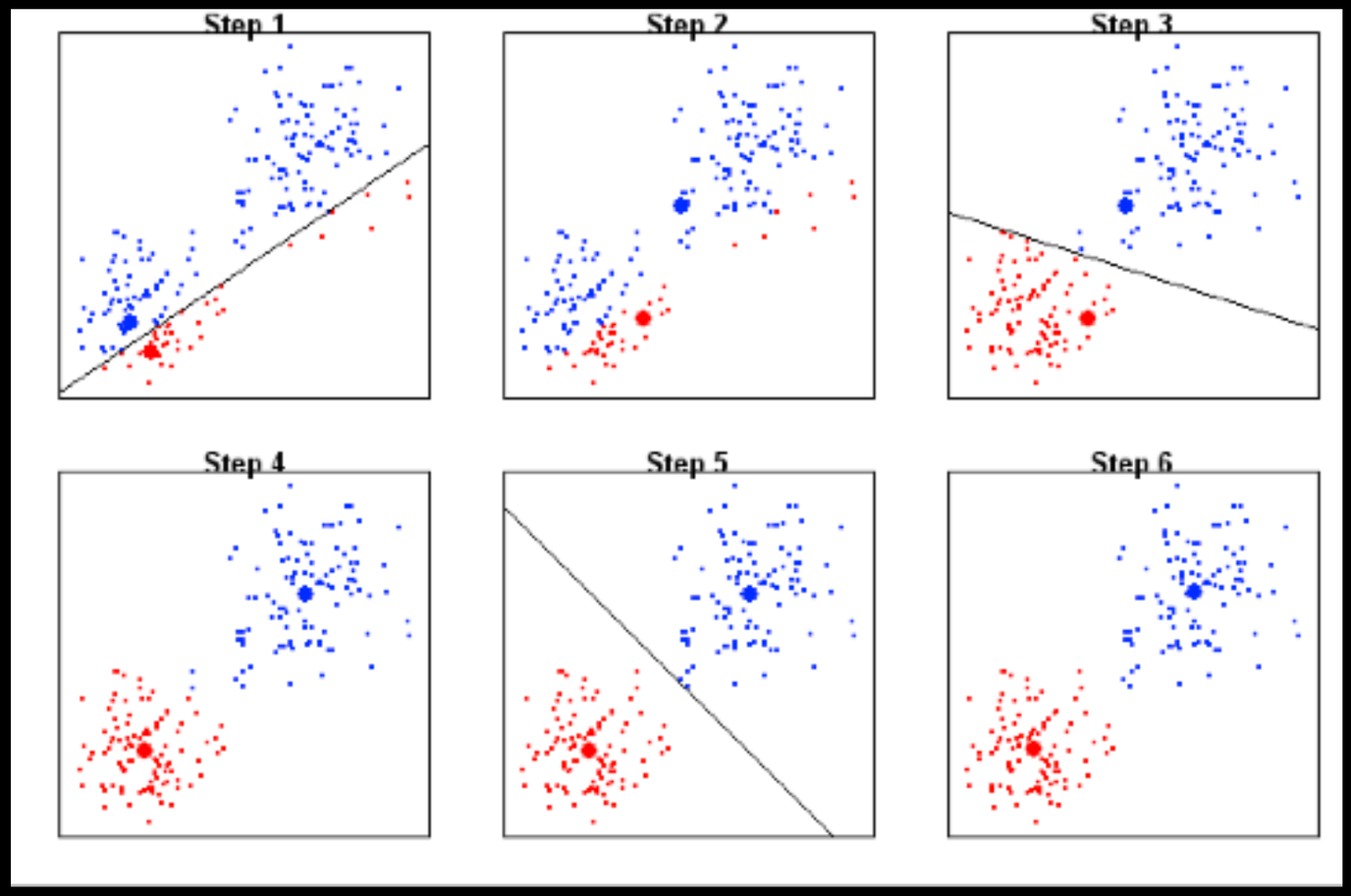
- Clustering

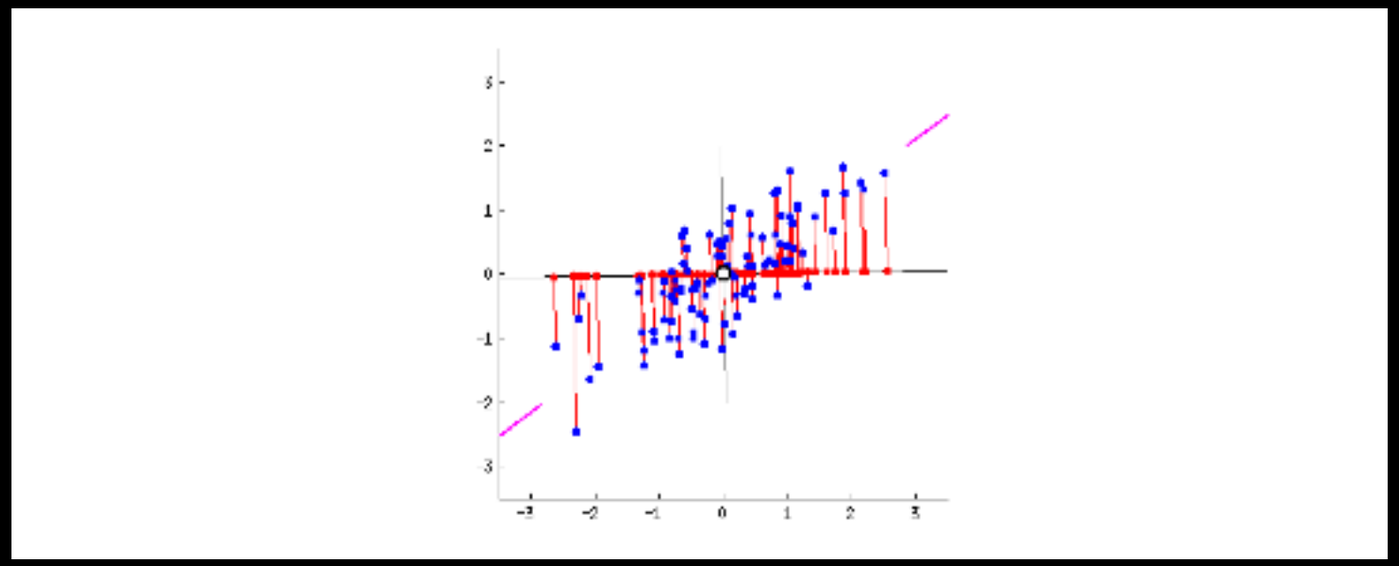
- Association

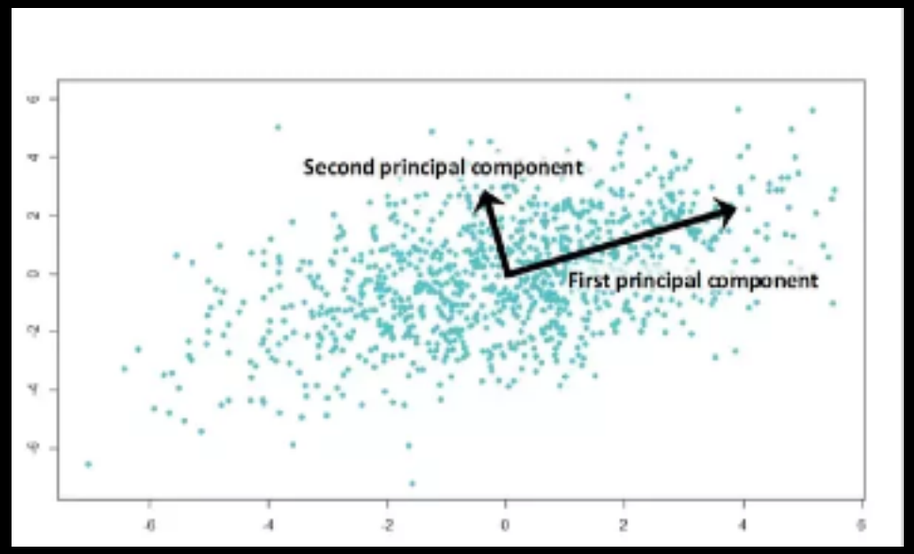

Step 1
Step 2
Step 3
Step 4
Step 5
Step 6

RECALL FROM LAST TIME…

- Unsupervised learning is where you only have input data (X) and no corresponding output variables (Y).
- You don't yet know what patterns you wish to find. There is no correct answers.
- Two types of unsupervised learning:
  - Clustering: Uncover the inherent groupings in the data, such as grouping customers by purchasing behavior.
  - Association: Uncover rules that describe large portions of the data, such as people that buy X also tend to buy Y.







Principal component analysis (**PCA**) is a classic example of unsupervised machine learning that **clusters** data together.

PCA reduces a large number of variables to a small set that still contains most of the information from the larger group of variables.

In other words, it's a way to compress a lot of data, but still keeps the essence of the original large data.

Watch the video on the next slide on **PCAs** and pay close attention to how they reduce the number of dimensions.

Watch the video on the next slide on **PCAs** and pay close attention to how they reduce the number of dimensions.

In other words, it's a way to compress a lot of data, but still keeps the essence of the original image data.

PCA reduces a large number of variables to a small set that still contains most of the information from the larger group of variables.