Dimensionality Reduction

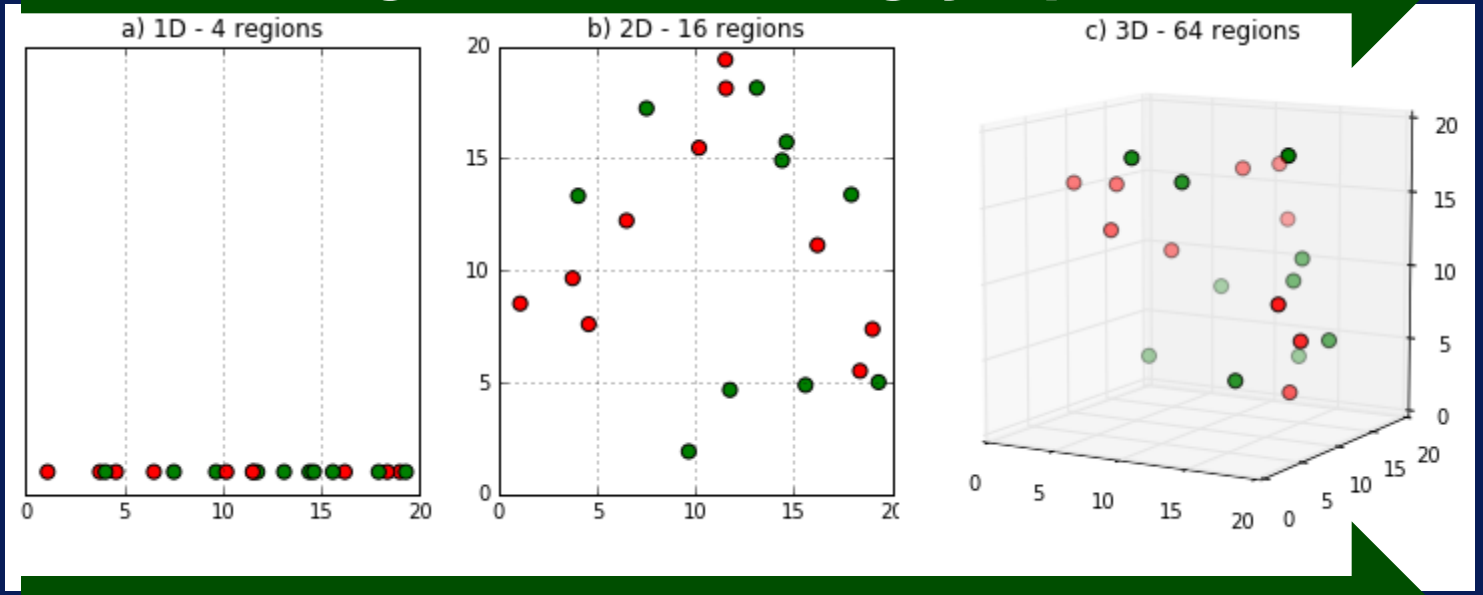
After this video you will be able to..

- Explain what dimensionality reduction is
- Discuss the benefits of dimensionality reduction
- Describe how PCA transforms your data

Dimensionality of Data

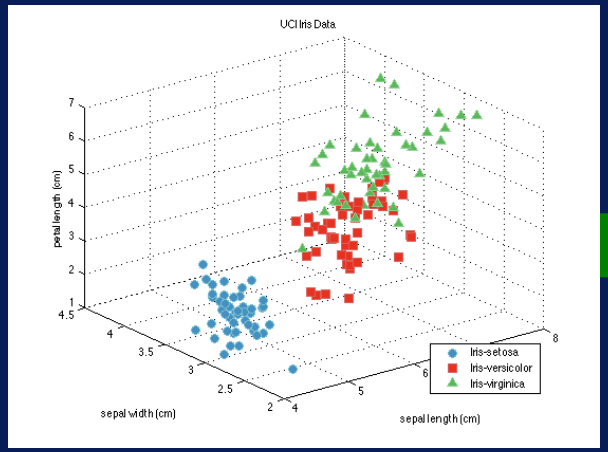
Curse of Dimensionality

Data gets increasingly sparse

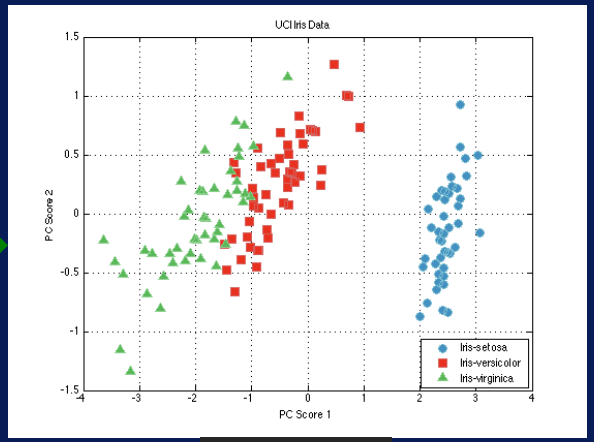


Analysis results degrade

Dimensionality Reduction

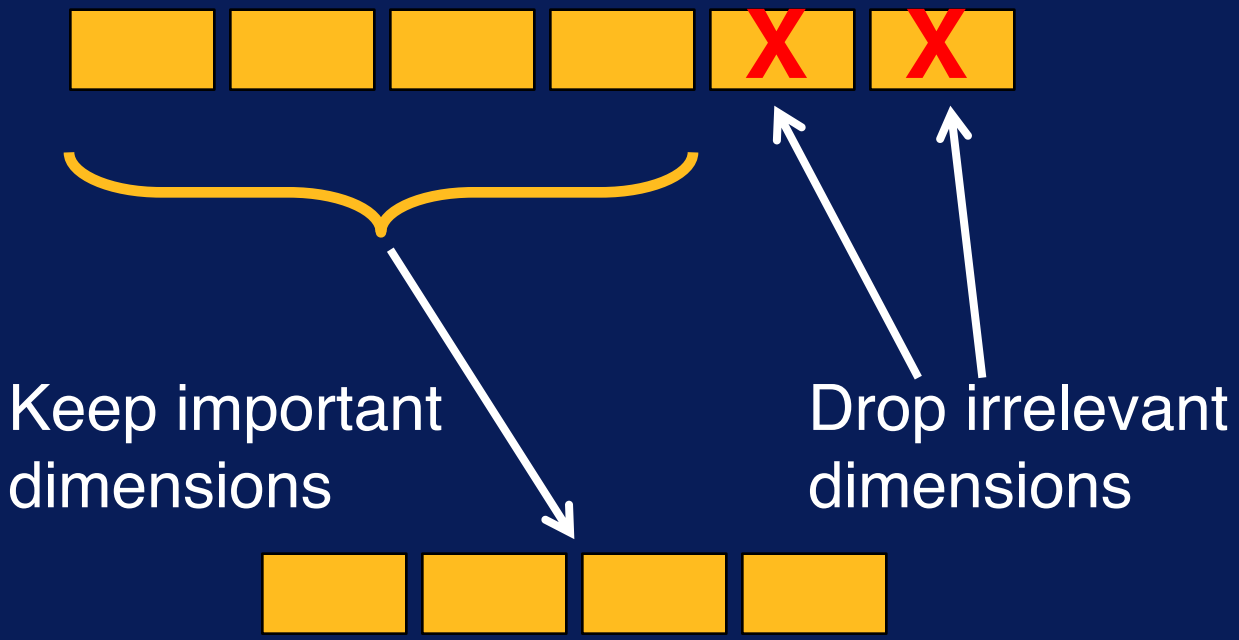


3D

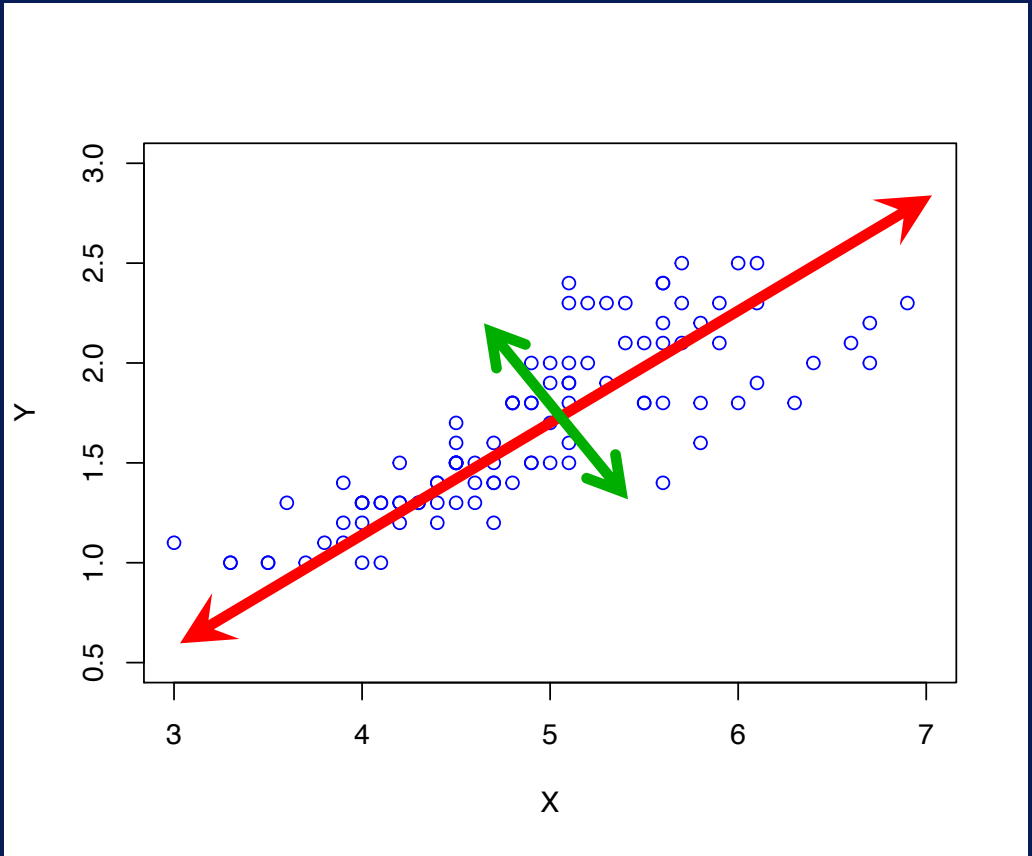


2D

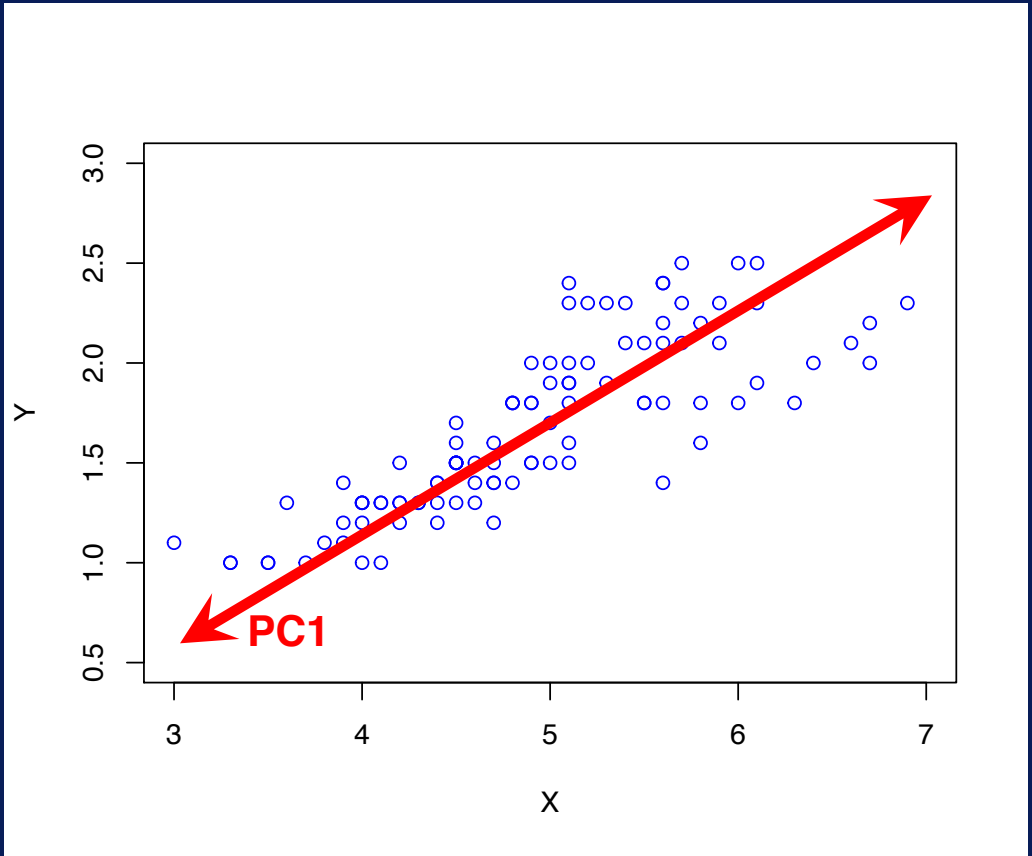
Dimensionality Reduction



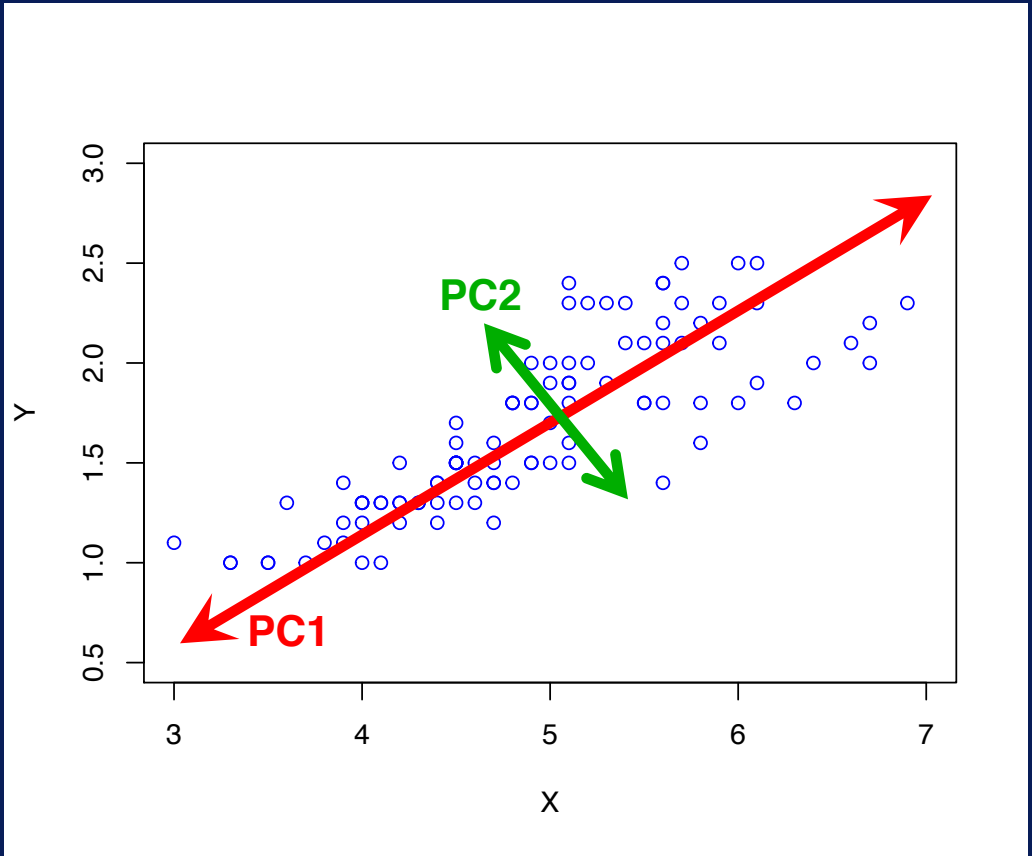
Principal Component Analysis



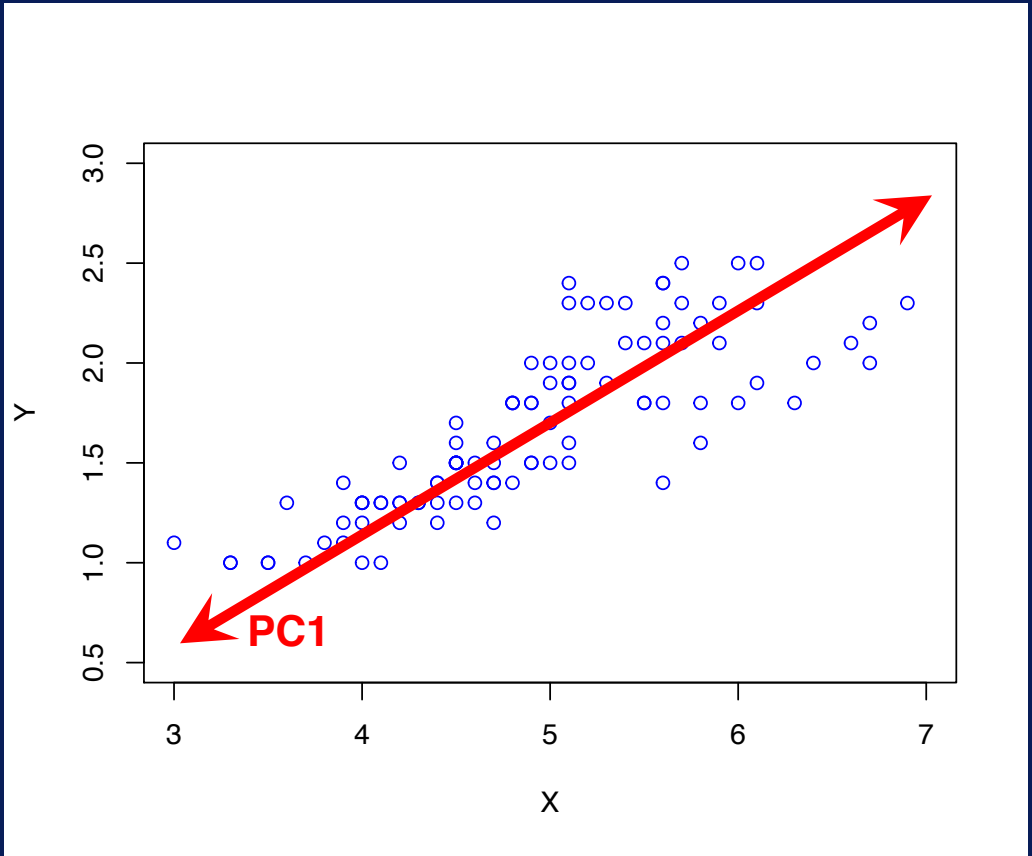
Principal Component Analysis



Principal Component Analysis



PCA for Dimensionality Reduction



PCA Main Points

- **Finds a new coordinate system such that**
  - PC1 captures greatest variance
  - PC2 captures second greatest variance, etc.
- **First few PCs capture most of variance**
  - Define lower-dimensional space for data.

PCA – To Note

- **Original dimensions**
  - Income, age, occupation, etc.
- **New dimensions**
  - PC1, PC2, PC3, etc.
- **More difficult to interpret!**