



ITESO

Universidad Jesuita
de Guadalajara

Introduction to the exploratory data analysis

Orange data mining

Luis Escobar

Dec., 2022

Outline

- Correlations
- Distributions
- Hierarchical clustering
- Principal Component Analysis (PCA)



Introduction



Data Mining Fruitful and Fun

Open source machine learning and data
visualization.

Build data analysis workflows visually, with
a large, diverse toolbox.

[Download Orange](#)



Figure 1: Orange Data Mining IDE vs Library

References

- M. Subhashree, Comparative Study of Different Orange Data Mining Tool-Based AI Techniques in Image Classification, May 2021.
- <https://docs.biolab.si/orange/2/widgets/rst/index.html>



Backup Slides

Hierarchical clustering

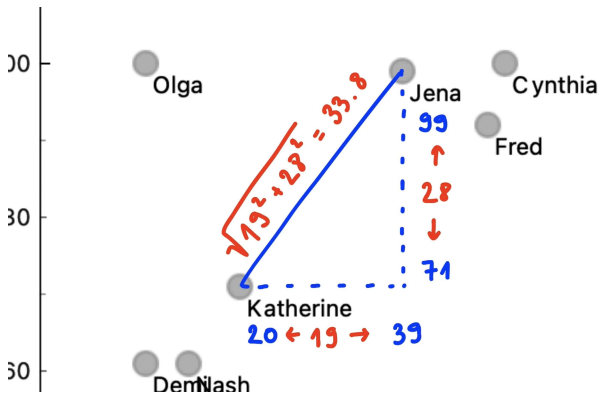


Figure 2: Euclidean distance

Principal component analysis PCA

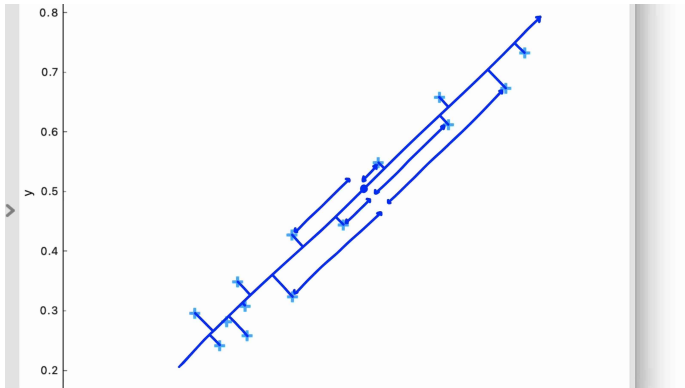


Figure 3: PC1 - transformation

Principal component analysis PCA

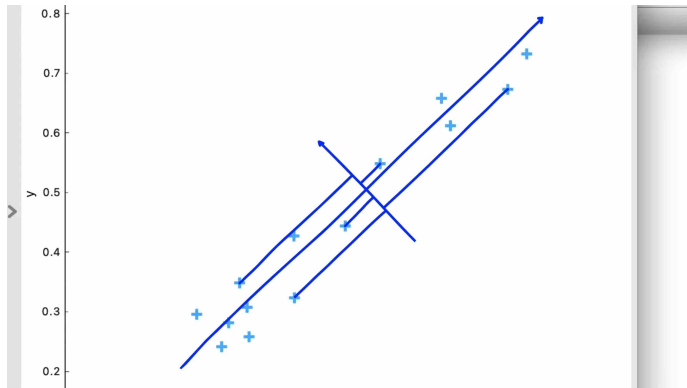


Figure 4: PC2 - transformation