Clustering and Ordination

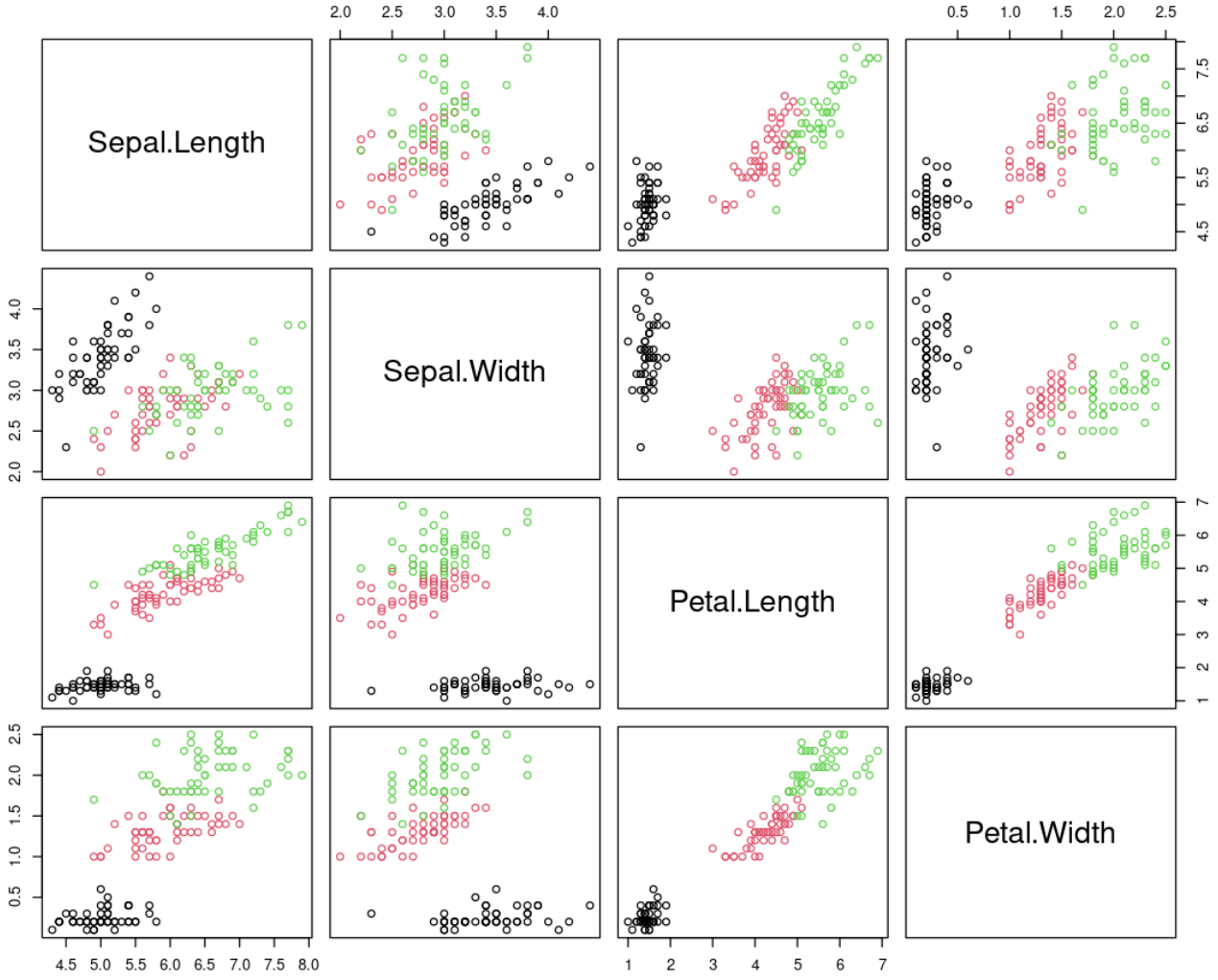
CSCI 297b, Spring 2023

May 10, 2023

Supervised vs. Unsupervised Algorithms

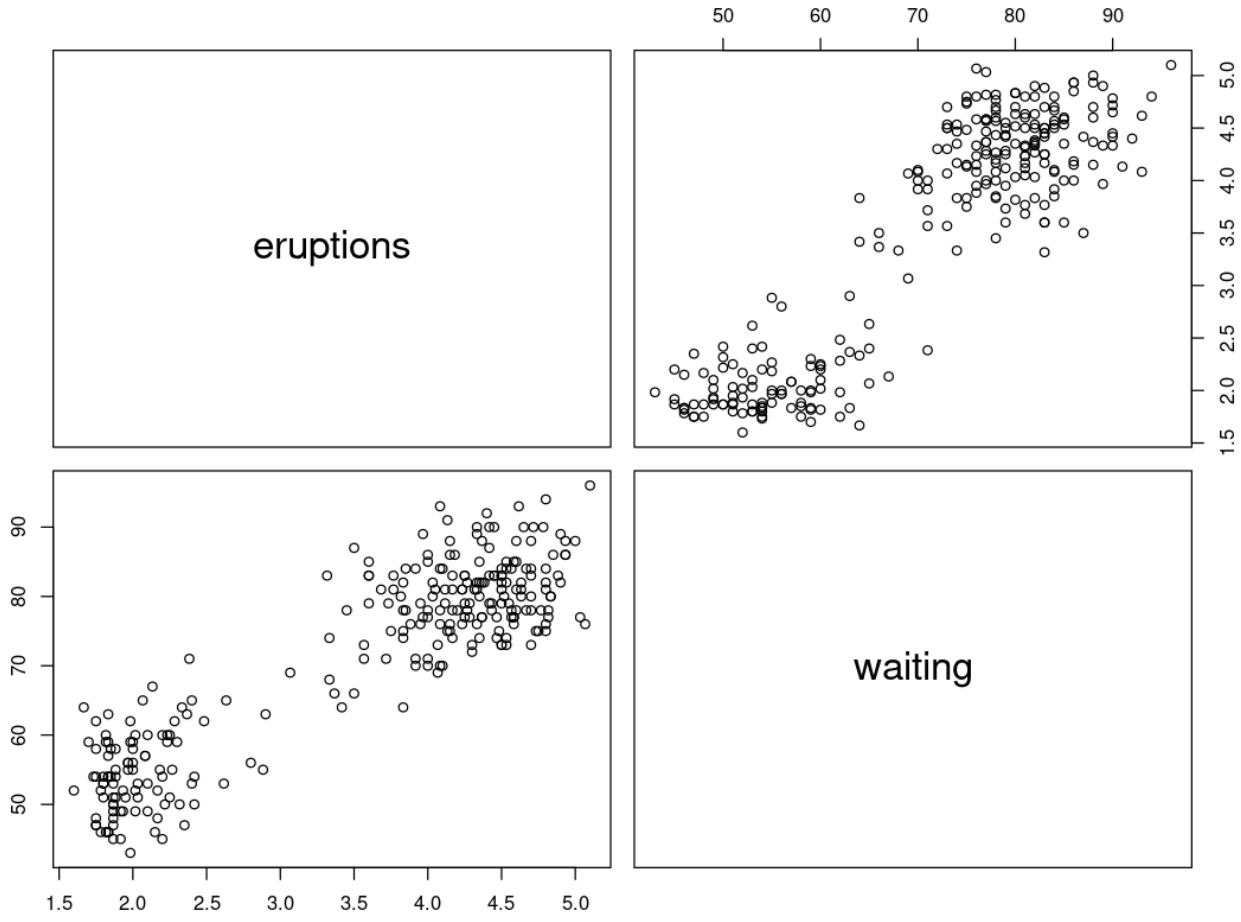
- Supervised algorithms answer questions based on known groups or trends:
  - Is treatment different from control?
  - Are males different from females?
- Unsupervised algorithms attempt to find unknown groups or trends:
  - What is the main trend in the data?
  - What are the two or three dominant groups in the data?

Supervised vs. Unsupervised Algorithms



data(iris)

Supervised vs. Unsupervised Algorithms



data(faithful)

Clustering and Ordination

- Ordination: put things in a line. Principal components:
  - Explanation: https://en.wikipedia.org/wiki/Principal_component_analysis
  - R Implementation: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/prcomp
- Clustering: put things in groups. EM Gaussian mixture:
  - Explanation: https://en.wikipedia.org/wiki/Expectation%E2%80%93maximization_algorithm
  - R implementation: https://cran.r-project.org/web/packages/mclust/vignettes/mclust.html