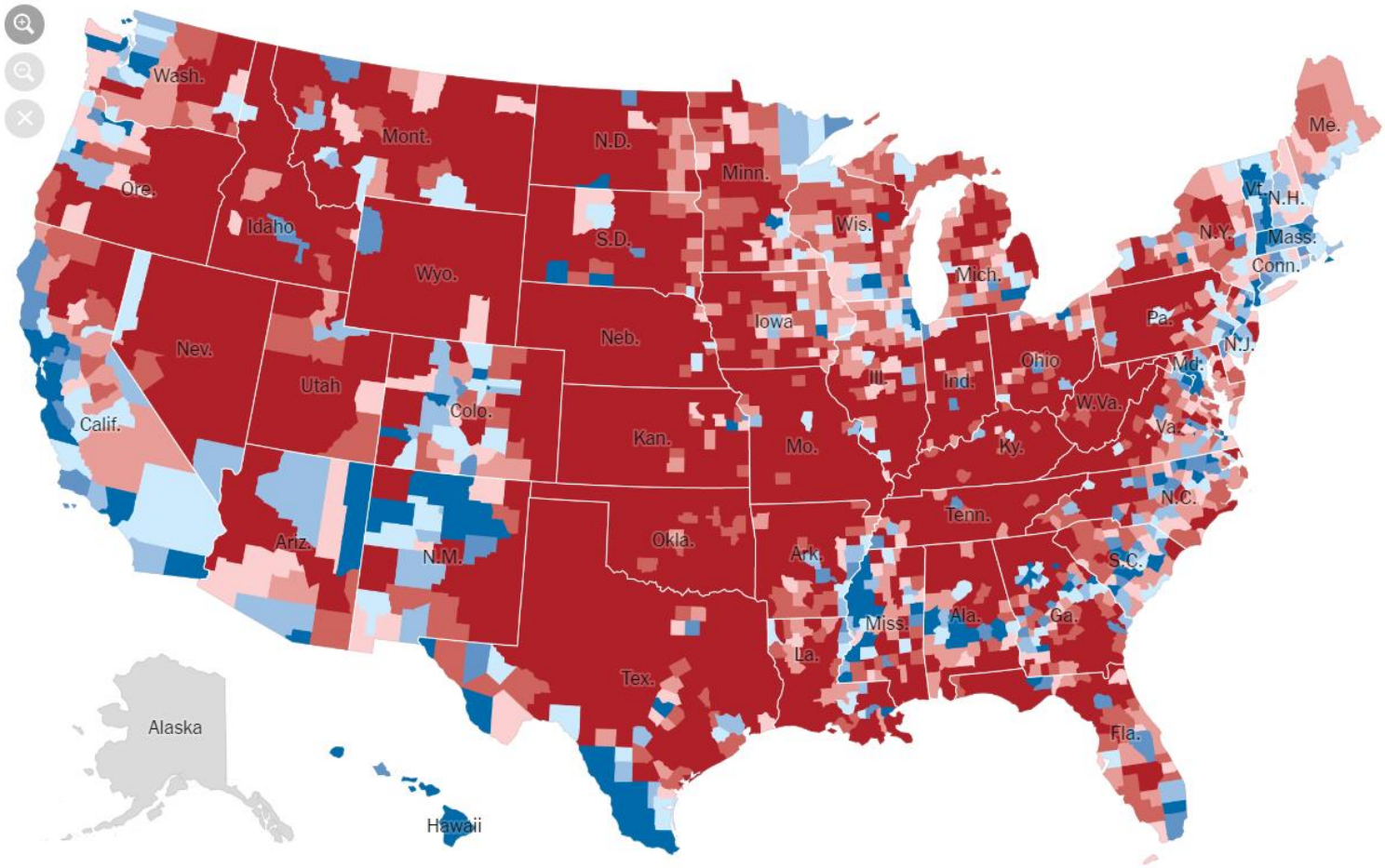
Machine Learning applied to Planetary Sciences

PTYS 595B/495B

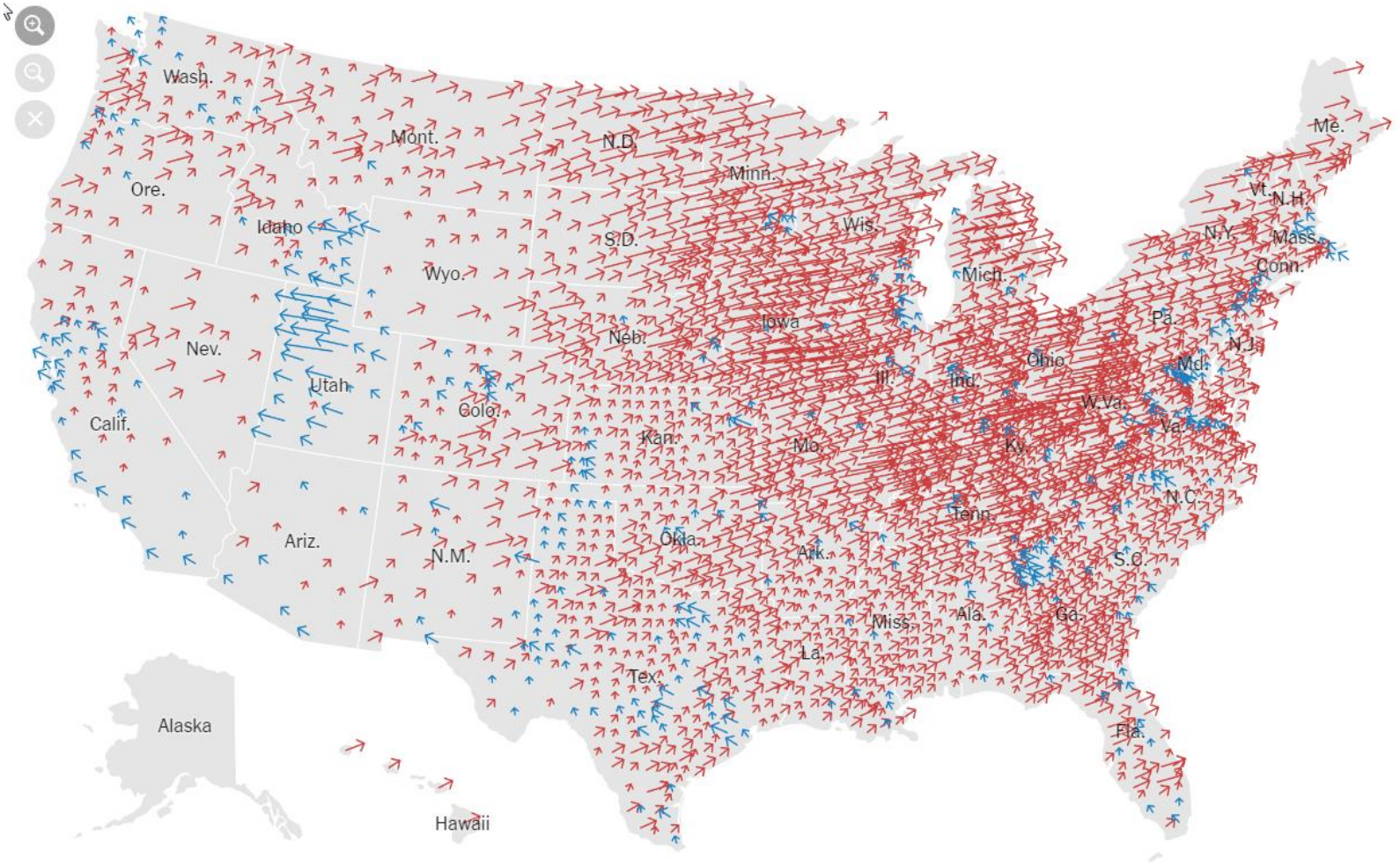
Leon Palafox

https://leonpalafox.github.io/MLClass/

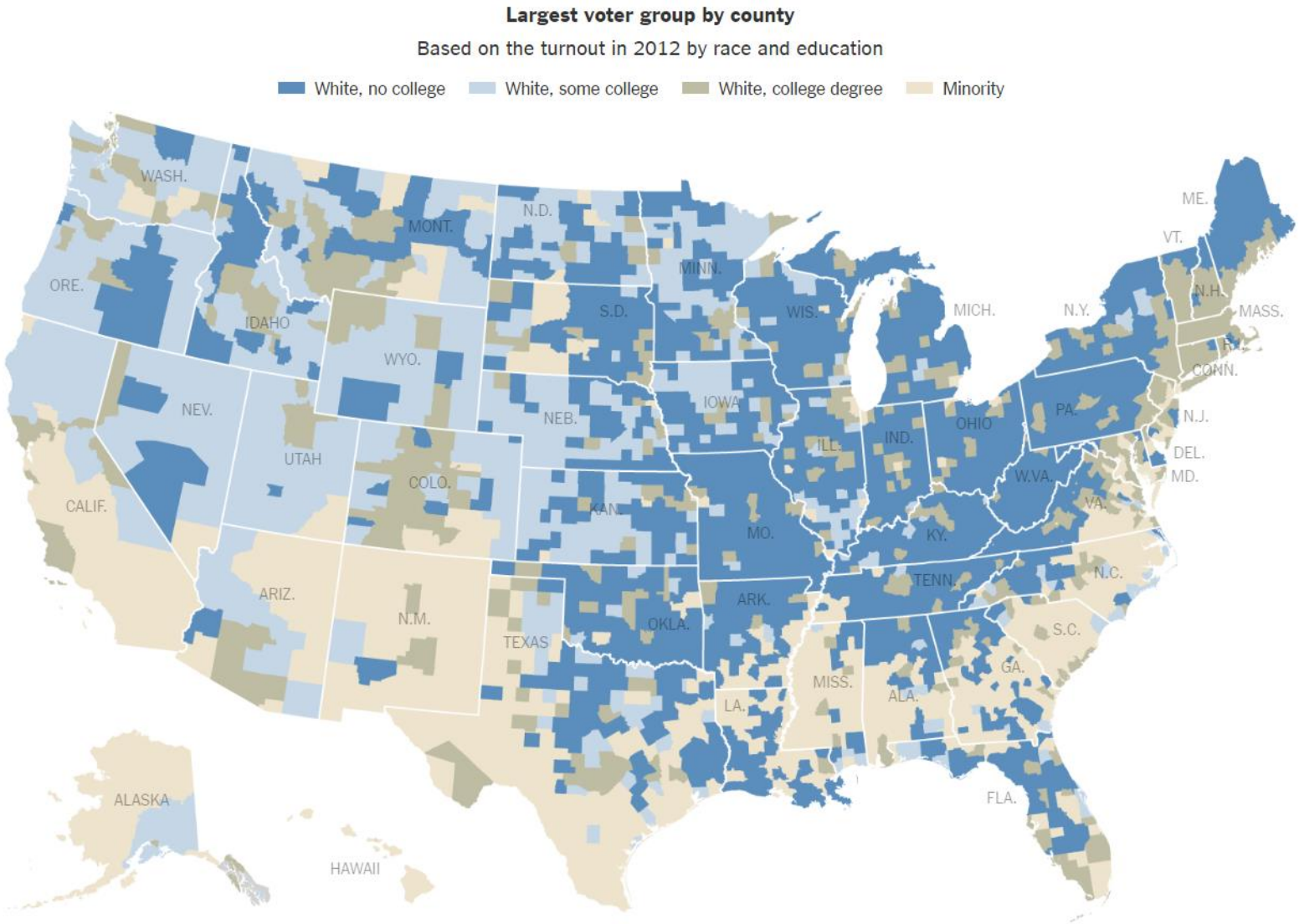
Group Activity



Group Activity



Group Activity



Knowledge discovery

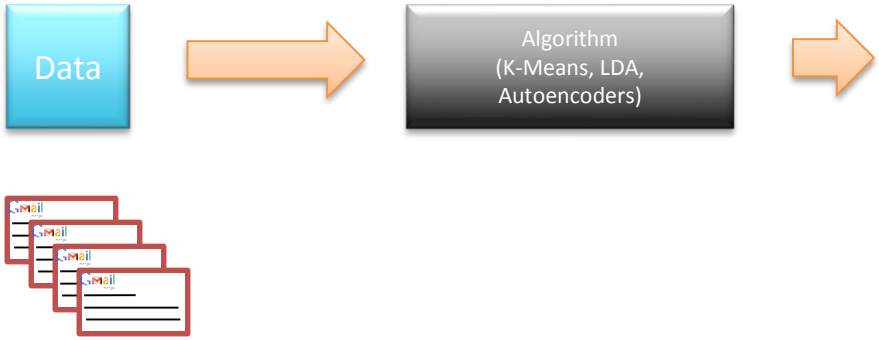
- We don’t need labels to discover patterns.
- Data organizes itself(closeness)
- Most algorithms just figure out that organization.

Manifolds and Data

- http://cs.stanford.edu/people/karpathy/tsnejs/

- In a manifold data points that are close share characteristics.
- Closeness is defined based on the features
  - Pixels
  - Words
  - Data

Unsupervised Learning



Mail Inbox

The set of elements that describe a single datum are called features, in this case, the features are the words in the e-mails.

Each topic (clusters) will have features that will characterize them.

Research: Mars, Proposal, DTM, HiRISE, Machine Learning, Deep Nets, Bayesian

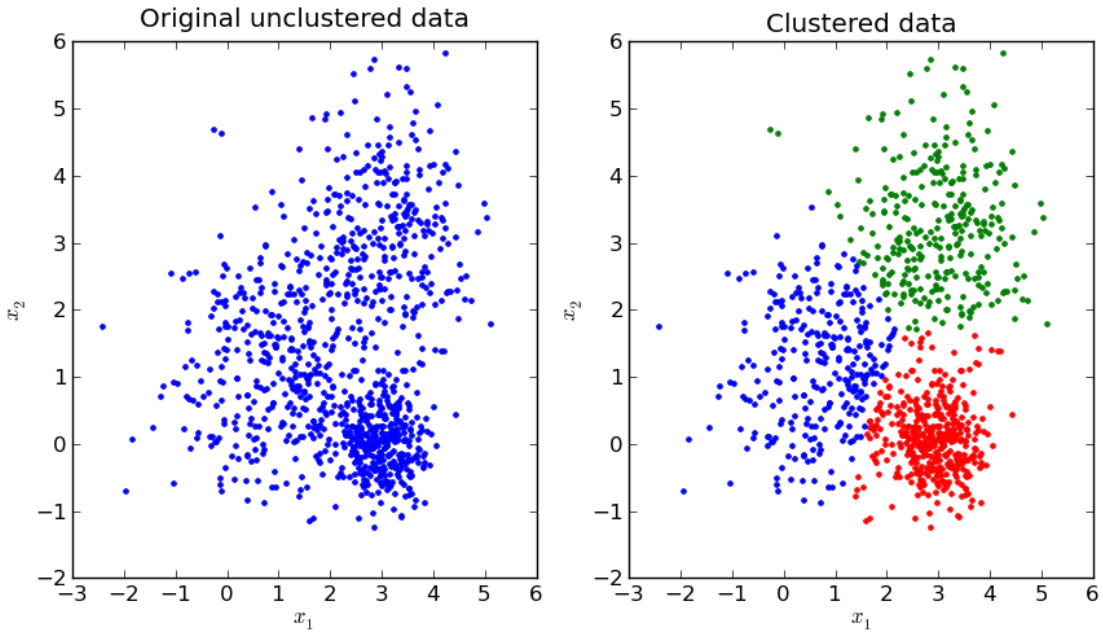
Family: Mom, House, Mexico

Promotions: Computer, PS4, Cheap, Amazon, Deal

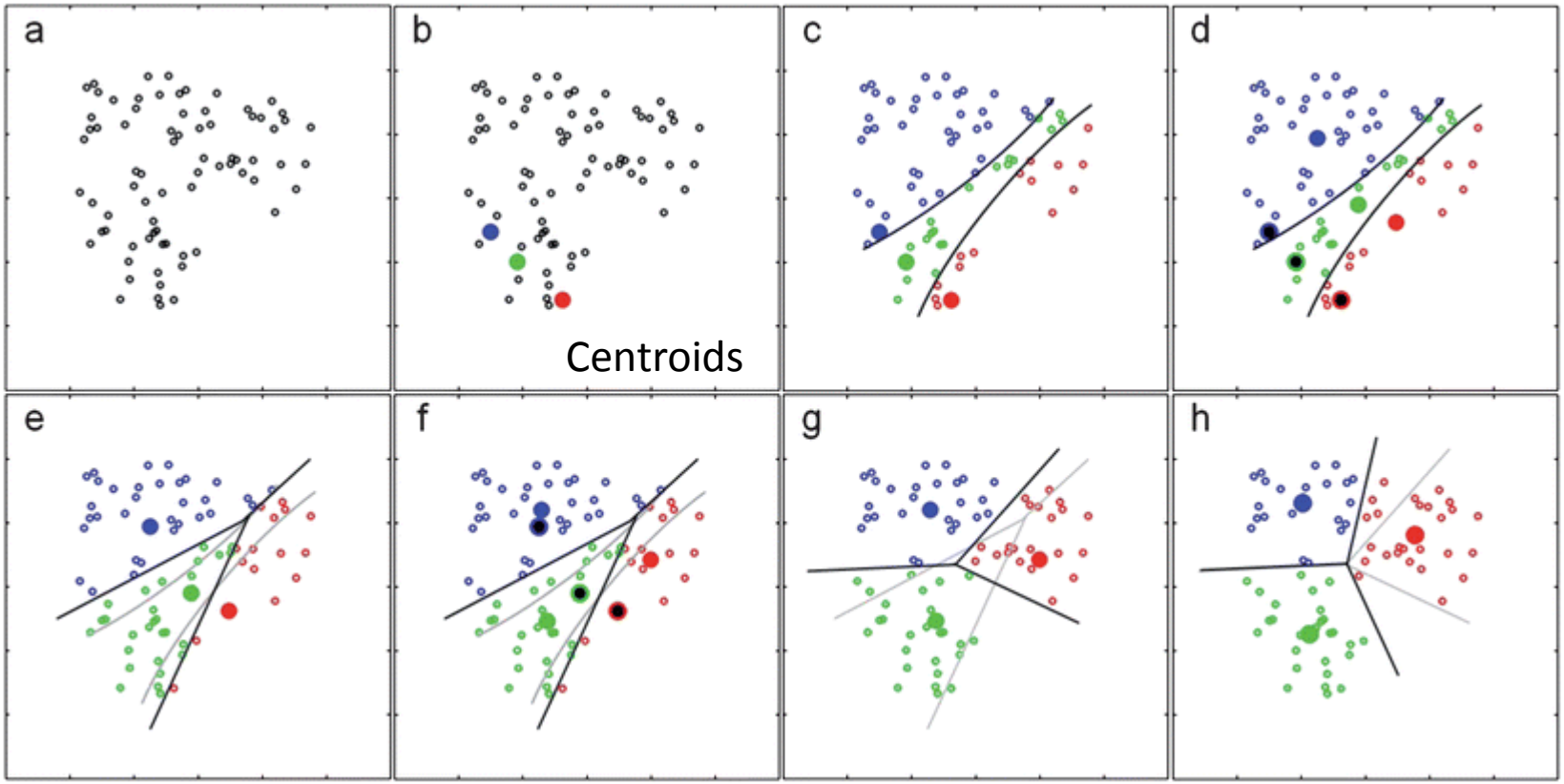
Classes: Grades, Homework, Questions, Office Time

K-Means

- http://kluster.j38.net/



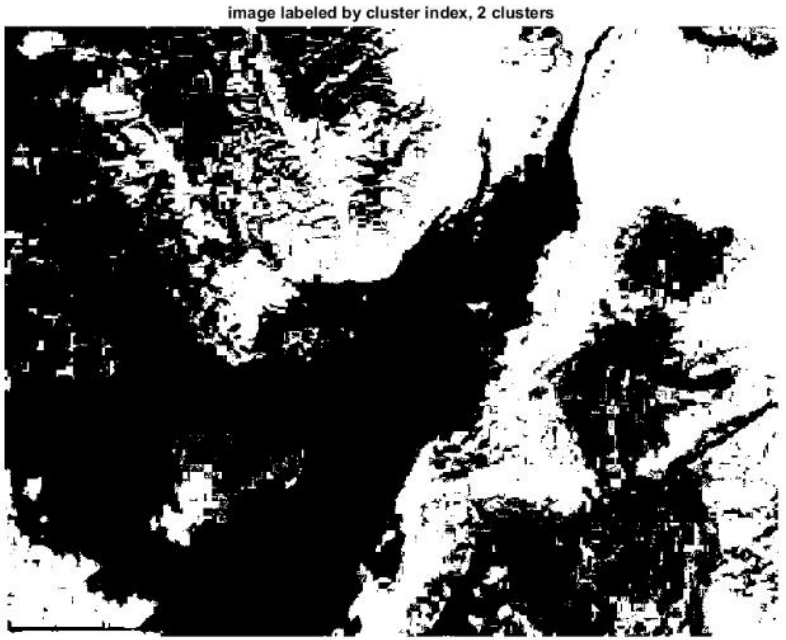
K-Means



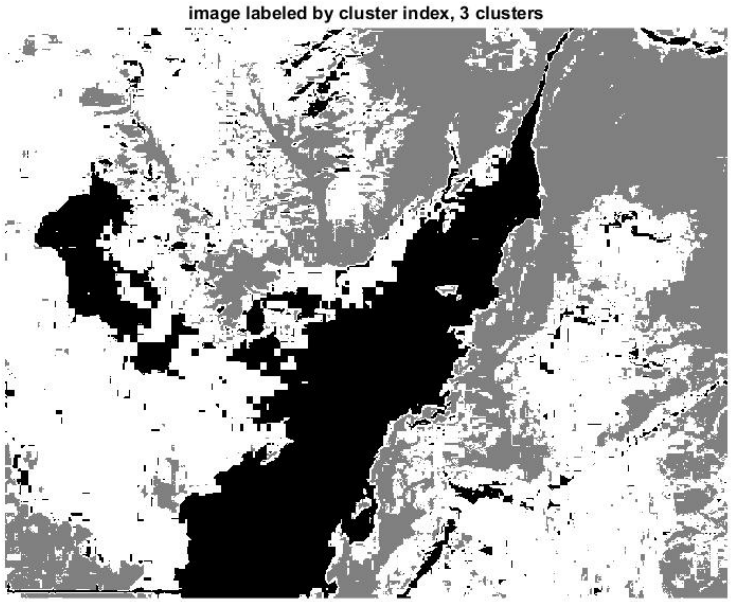
Applications to Geology

Scale Bar: 6.22 mi

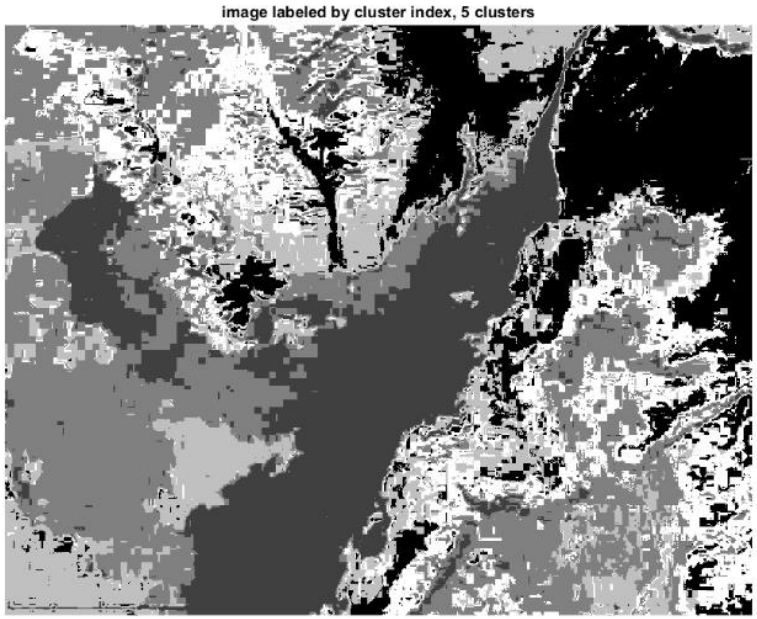
Applications to Geology



Applications to Geology



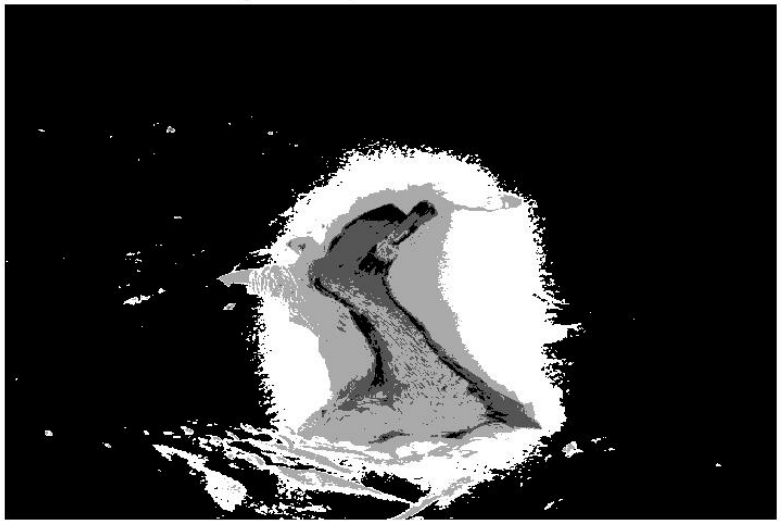
Applications to Geology



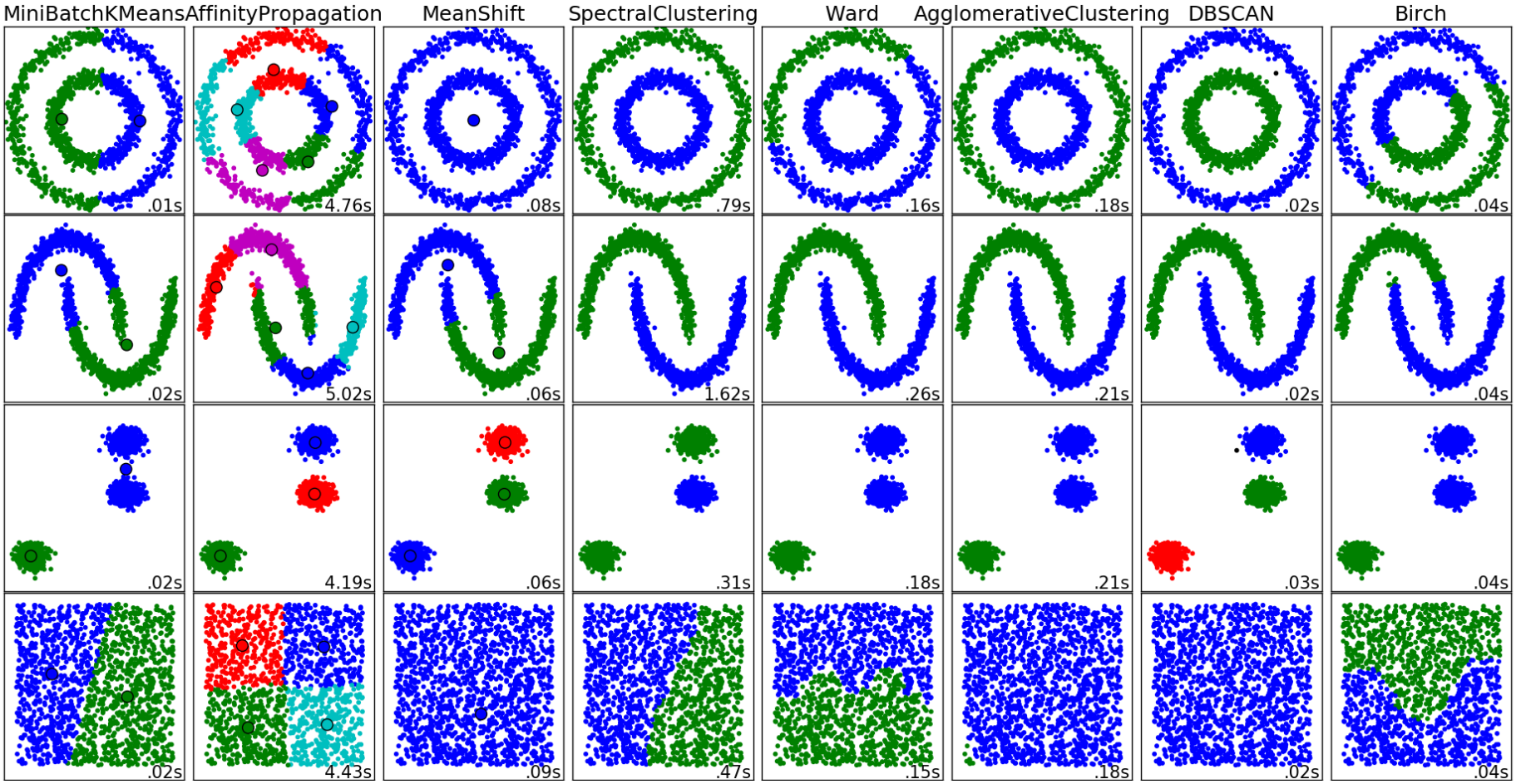
Applications to Geology

Lava Flow in Hawaii

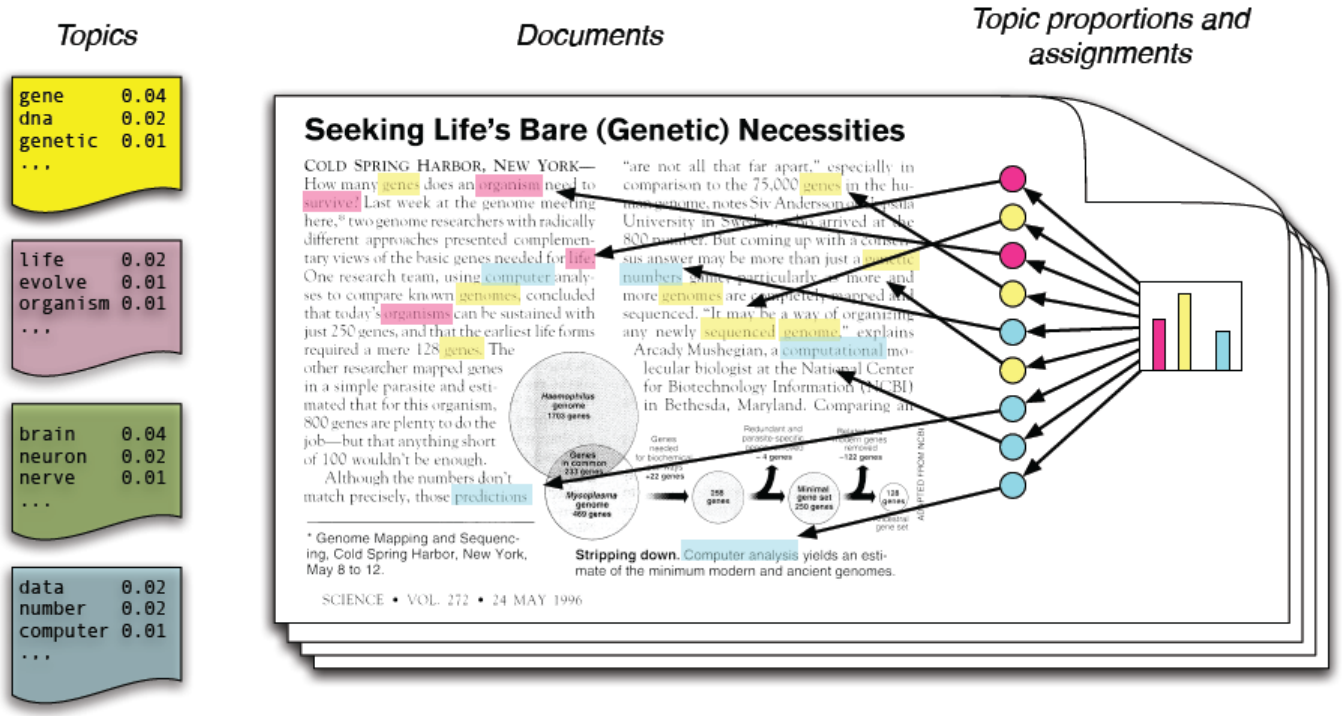
image labeled by cluster index, 4 clusters



Other Clustering Techniques



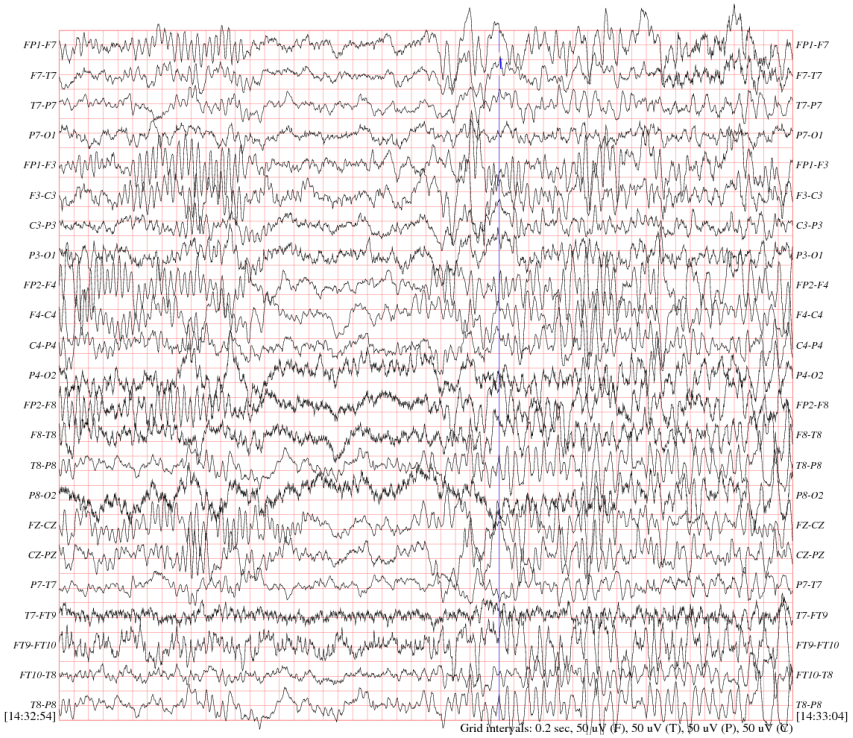
Latent Dirichlet Allocation (LDA)

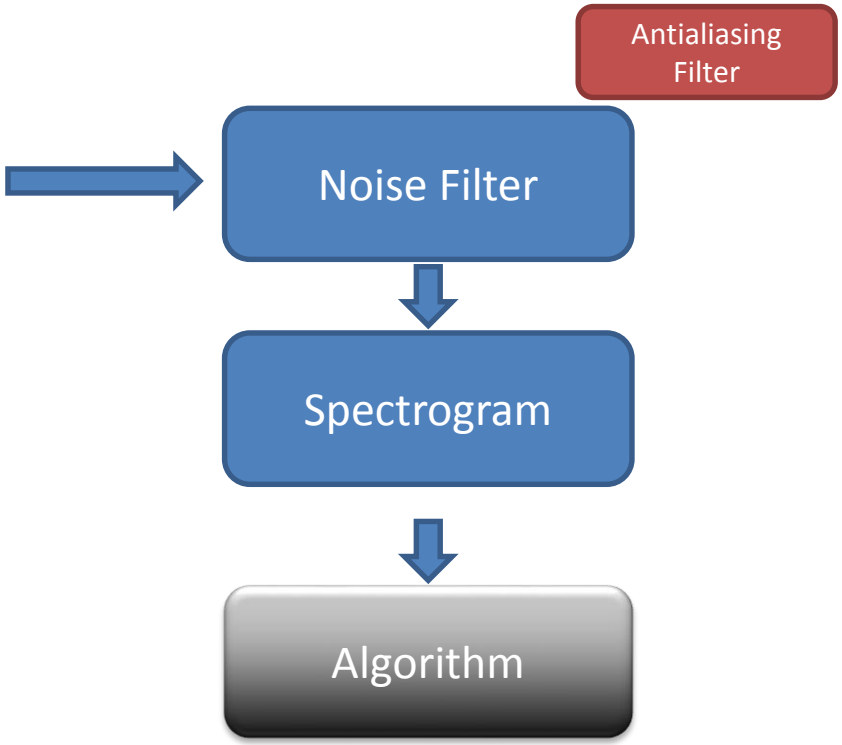


https://www.lpl.arizona.edu/~leonp/lplpapers2014/lpl2014lda5Topics.html
https://www.lpl.arizona.edu/~leonp/HiRISE/HiRISELDA5Topics.html

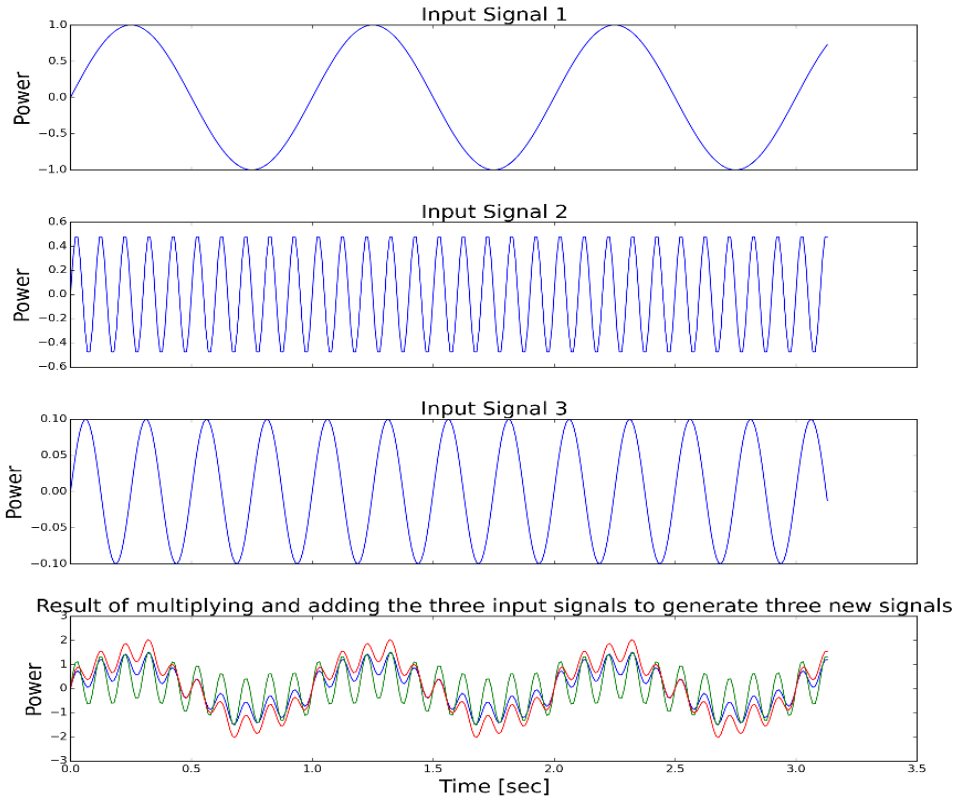
Preprocessing Data

- It’s a pain, but is needed

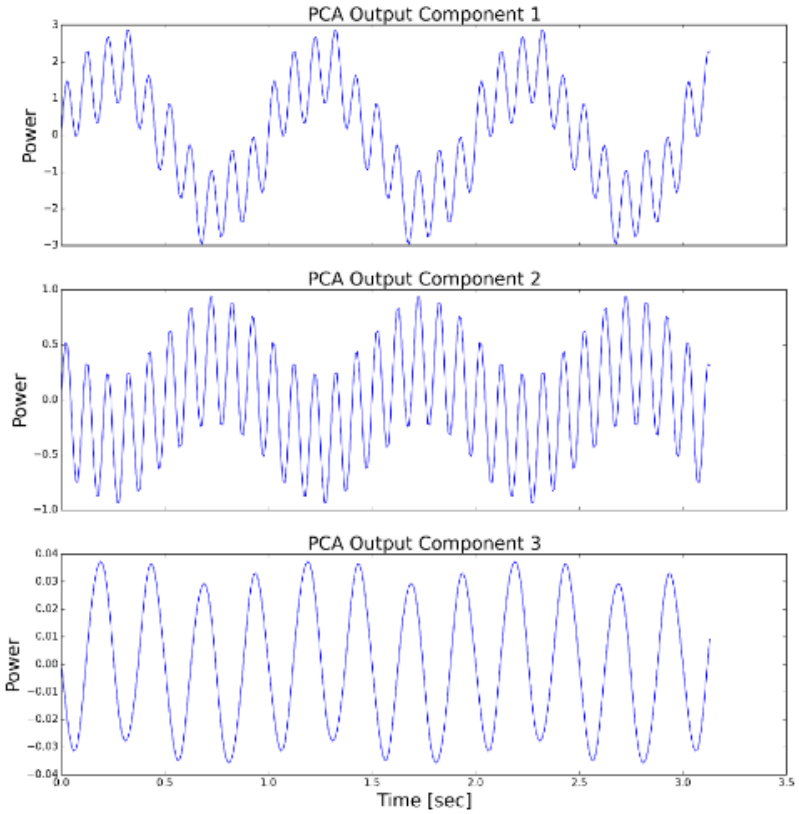


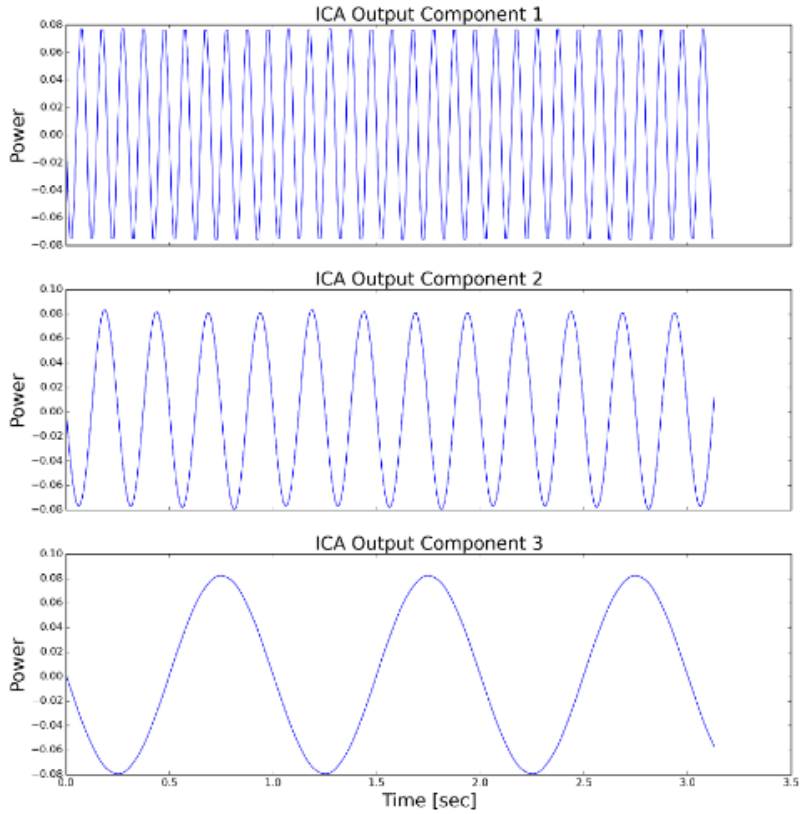


Data Decomposition



Data Decomposition





Independent Component Analysis