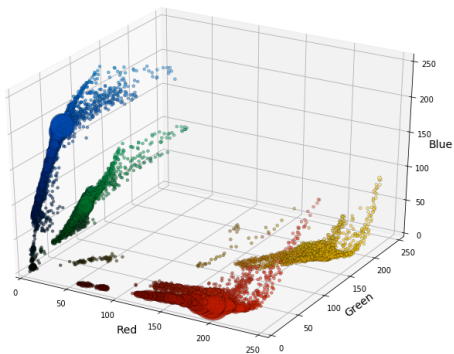


# Clustering - Intro

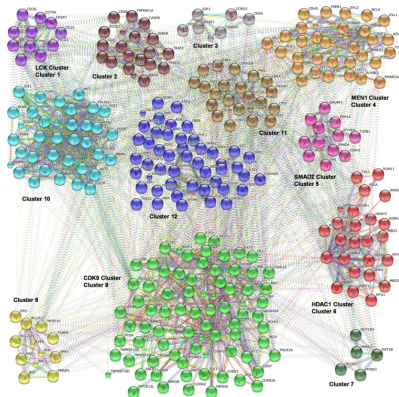


Frederik Mallmann-Trenn  
6CCS3AIN



- Goal: Points in the same cluster are ‘close’ and points of different clusters are ‘far’

# Application: Biological Networks



- Protein-Protein Interaction Networks (PPINs). See e.g., Rual et al Nature '05 (+2500 citations)

# Application: Recommendation Systems



© Rina Piccolo.

- Similar users will buy similar items
- High dimensional data, how to find similar users? Clustering!

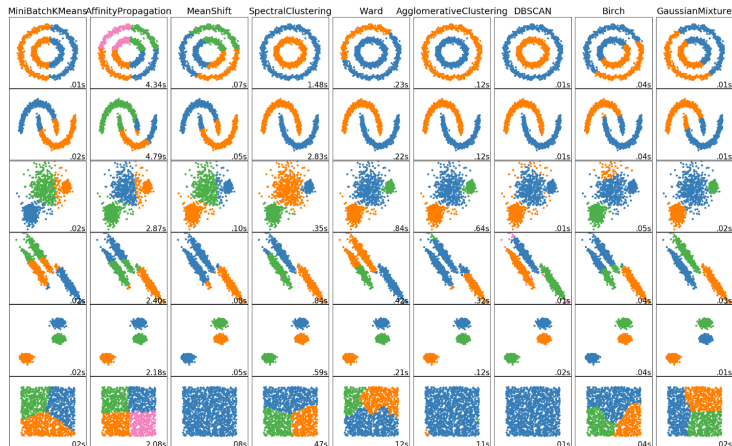
# Other Applications

- Unsupervised learning (e.g., PCA)
- encoder-decoder neural network
- outlier detection applications such as detection of credit card fraud
- ....

# Clustering Algorithm (incomplete)

- k-Means
- k-Median
- Agglomerative Clustering
  - Single-Linkage
  - Average-Linkage
  - Complete-Linkage
- Divisive Clustering
  - Sparsest-Cut

# Clustering Algorithm Comparison



Source: <https://towardsdatascience.com/>