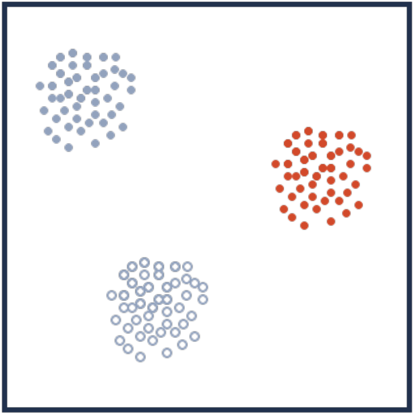
# Unsupervised Learning

Creating our own truth



Clustering

# Hierarchical Clustering

- Agglomerative clustering
- Divisive clustering

# Agglomerative Clustering Algorithm

**Algorithm 1:** Agglomerative Clustering Solution Algorithm

**Input:** Features from data

**Output:** $k$ cluster labeling for each point

Initialize $n$ clusters where every point is a cluster

**while** *cluster count reached greater than k* **do**

  Determine the two clusters closest to each other based on distance metric and linkage criteria

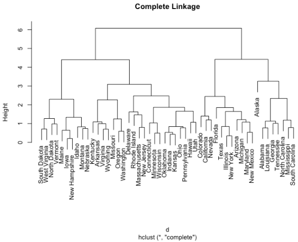
  Merge the two clusters into one

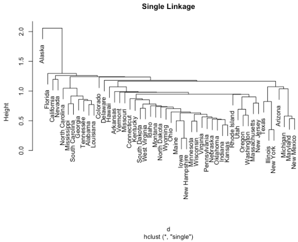
**end**

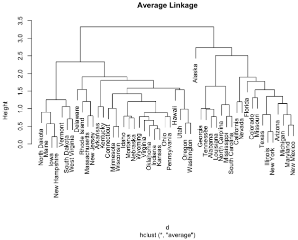
# Linkage Criteria

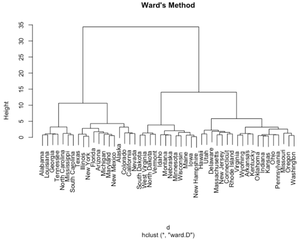
- Complete linkage - $\max(\text{dist}(c_i, c_j))$
- Single linkage - $\min(\text{dist}(c_i, c_j))$
- Average linkage - $\text{mean}(\text{dist}(c_i, c_j))$
- Ward - $\text{variance}(\text{merge}(c_i, c_j))$

# Dendrogram Examples



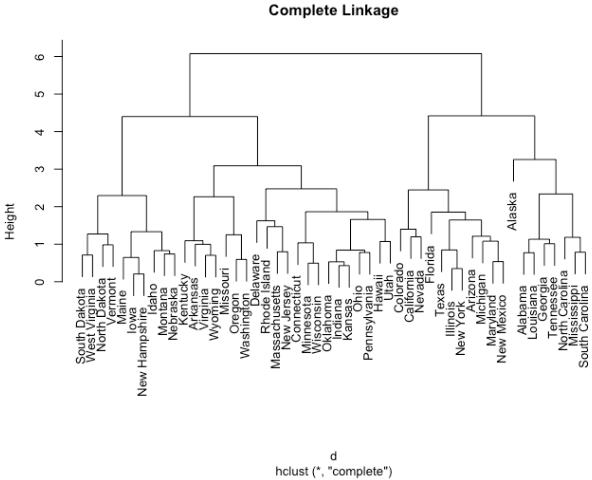



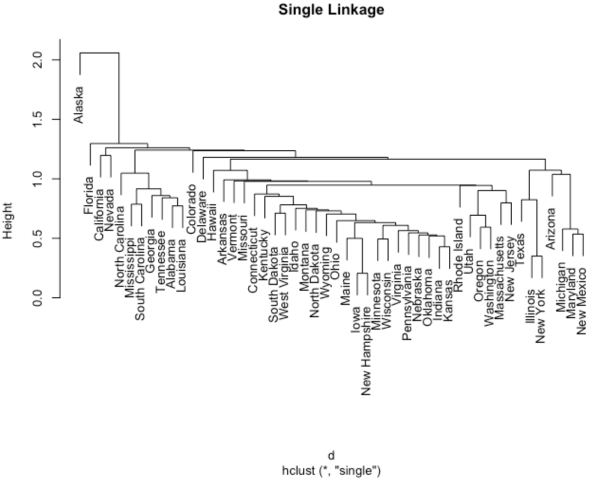
# Complete Linkage Dendrogram

$$\max(\text{dist}(c_i, c_j))$$



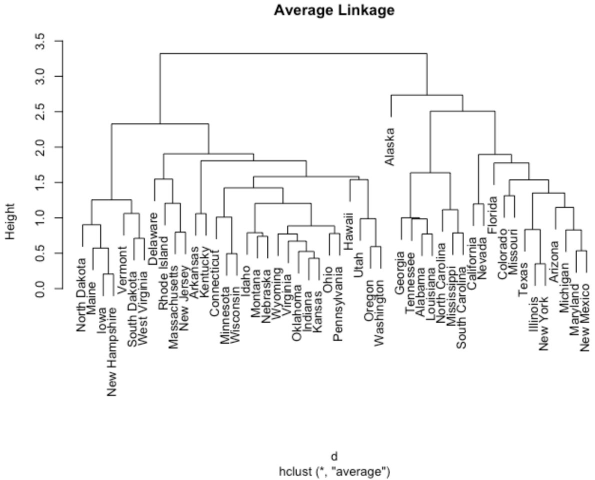
# Single Linkage Dendrogram

$$\min(\text{dist}(c_i, c_j))$$



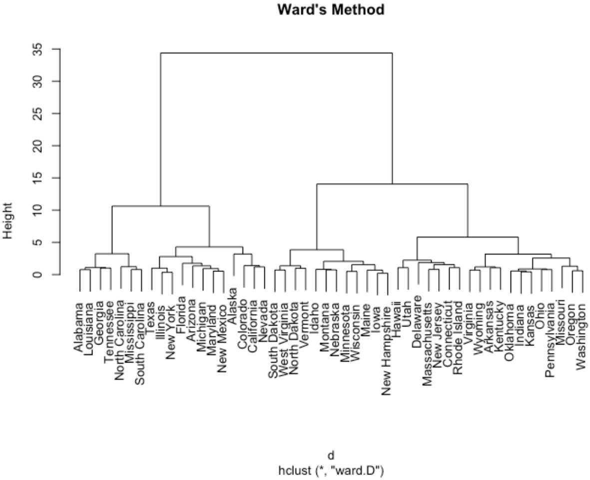
# Average Linkage Dendrogram

$$\text{mean}(\text{dist}(c_i, c_j))$$



# Ward Dendrogram

$$\text{variance}(\text{merge}(c_i, c_j))$$



# Agglomerative Clustering Algorithm

**Algorithm 1:** Agglomerative Clustering Solution Algorithm

**Input:** Features from data
**Output:** $k$ cluster labeling for each point
Initialize $n$ clusters where every point is a cluster
**while** *cluster count reached greater than* $k$ **do**
    Determine the two clusters closest to each other based on distance metric and linkage criteria
    Merge the two clusters into one
**end**

# Non-negative Matrix Factorization

$$\mathbf{X} = \mathbf{F}\mathbf{G}.$$

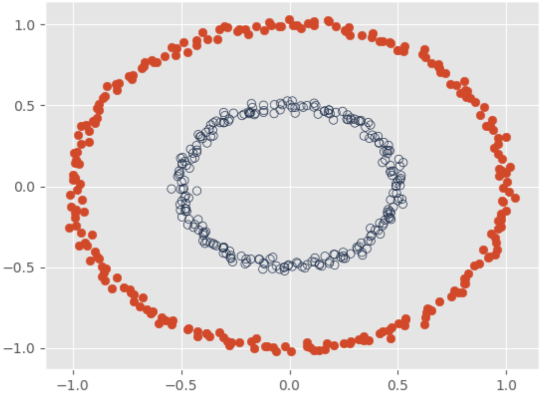
If $\mathbf{X}$ is a $d \times n$ matrix, we can produce
a $d \times k$, $\mathbf{F}$ matrix,
and $k \times n$, $\mathbf{G}$ matrix.

# NMF

The objective for NMF is,

$$\min_{\mathbf{F},\mathbf{G}} ||\mathbf{X} - \mathbf{F}\mathbf{G}||_F^2$$
$$s.t.\ \mathbf{F} \geq 0, \mathbf{G} \geq 0, \mathbf{G}\mathbf{G}^T = \mathbf{I}.$$

# Euclidean vs Manifold



# Manifold

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

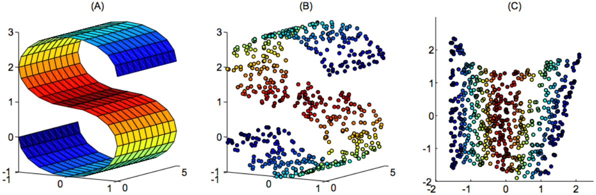
# Manifold

$$\mathbf{D} = \begin{bmatrix} \sum_{i=1}^{n} a_{1i} & 0 & 0 & \dots & 0 \\ 0 & \sum_{i=1}^{n} a_{2i} & 0 & \dots & o \\ 0 & 0 & \sum_{i=1}^{n} a_{3i} & \dots & 0 \\ \hdotsfor{5} \\ 0 & 0 & 0 & \dots & \sum_{i=1}^{n} a_{ni} \end{bmatrix}$$

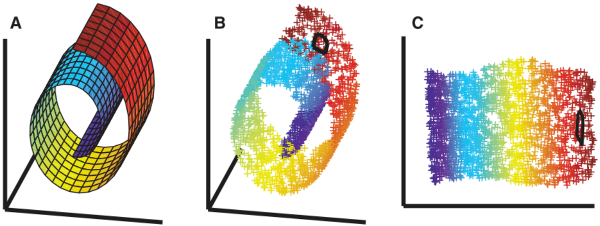
$$
\mathbf{L} = \mathbf{D} - \mathbf{A}
$$

$$
= \begin{bmatrix}
(\sum_{i=2}^{n} a_{1i}) & -a_{12} & -a_{13} & \dots & -a_{1n} \\
-a_{21} & (\sum_{i=1, i\neq 2}^{n} a_{2i}) & -a_{23} & \dots & -a_{2n} \\
-a_{31} & -a_{32} & (\sum_{i=1, i\neq 3}^{n} a_{3i}) & \dots & -a_{3n} \\
\hdotsfor{5} \\
-a_{n1} & -a_{n2} & -a_{n3} & \dots & (\sum_{i=1, i\neq n}^{n} a_{ni})
\end{bmatrix}
$$

# Spectral Clustering



# Spectral Clustering



# Questions

These slides are designed for educational purposes, specifically the CSCI-470 Introduction to Machine Learning course at the Colorado School of Mines as part of the Department of Computer Science.

Some content in these slides are obtained from external sources and may be copyright sensitive. Copyright and all rights therein are retained by the respective authors or by other copyright holders. Distributing or reposting the whole or part of these slides not for academic use is HIGHLY prohibited, unless explicit permission from all copyright holders is granted.