



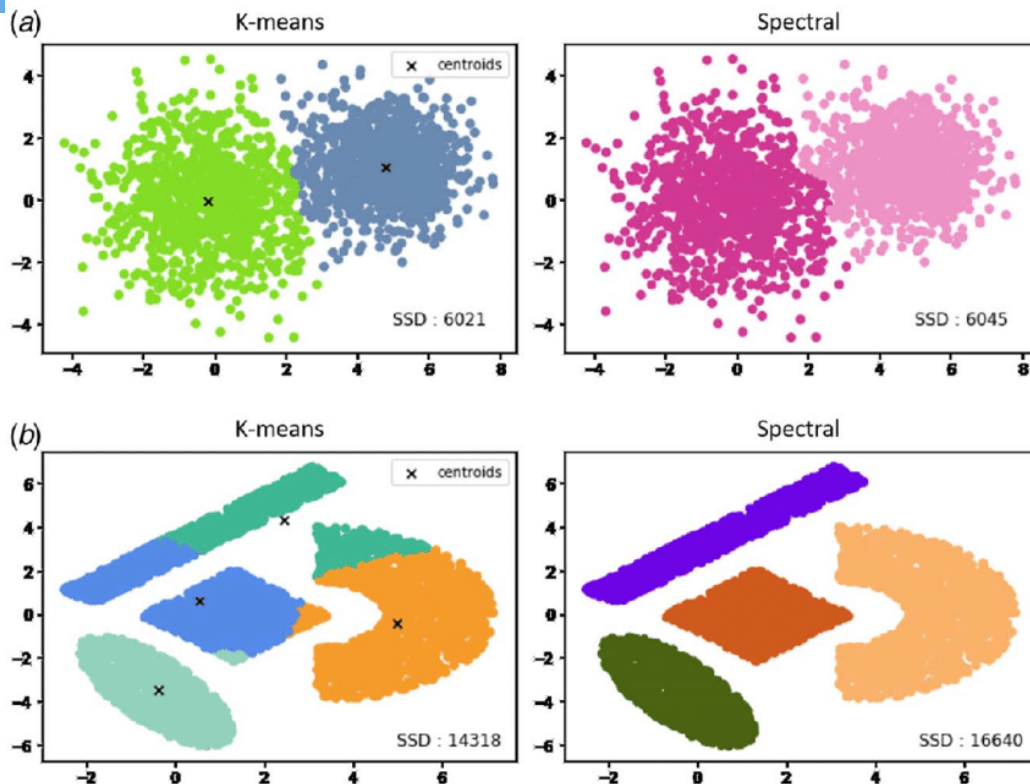
JOHNS HOPKINS

WHITING SCHOOL  
of ENGINEERING

# Algorithms for Data Science

Clustering and Dimensionality Reduction (Spectral Clustering)

# Spectral Clustering: Graph-Based Dimensionality Reduction & Clustering



Park, Seyoung & Kim, Harrison. (2022). Phrase Embedding and Clustering for Sub-Feature Extraction From Online Data. Journal of Mechanical Design. 144. 10.1115/1.4052904.

# Mathematical Foundation of Spectral Clustering

**Weighted  
Adjacency  
Matrix**

$A$

**Degree  
Matrix**

$$D_{ii} = \sum_j A_{ij}$$

**Laplacian Matrix**

$$L = D - A$$

- **Compute** the Eigenvectors of the Laplacian Matrix
- The **smallest nonzero eigenvalues** correspond to the most informative low-dimensional embedding
- **Apply K-Means clustering** to the Eigenvectors. (The spectral embedding finds a smooth manifold structure that captures natural clusters.)



# JOHNS HOPKINS

WHITING SCHOOL  
*of* ENGINEERING