

**EE16B, Spring 2018  
UC Berkeley EECS**

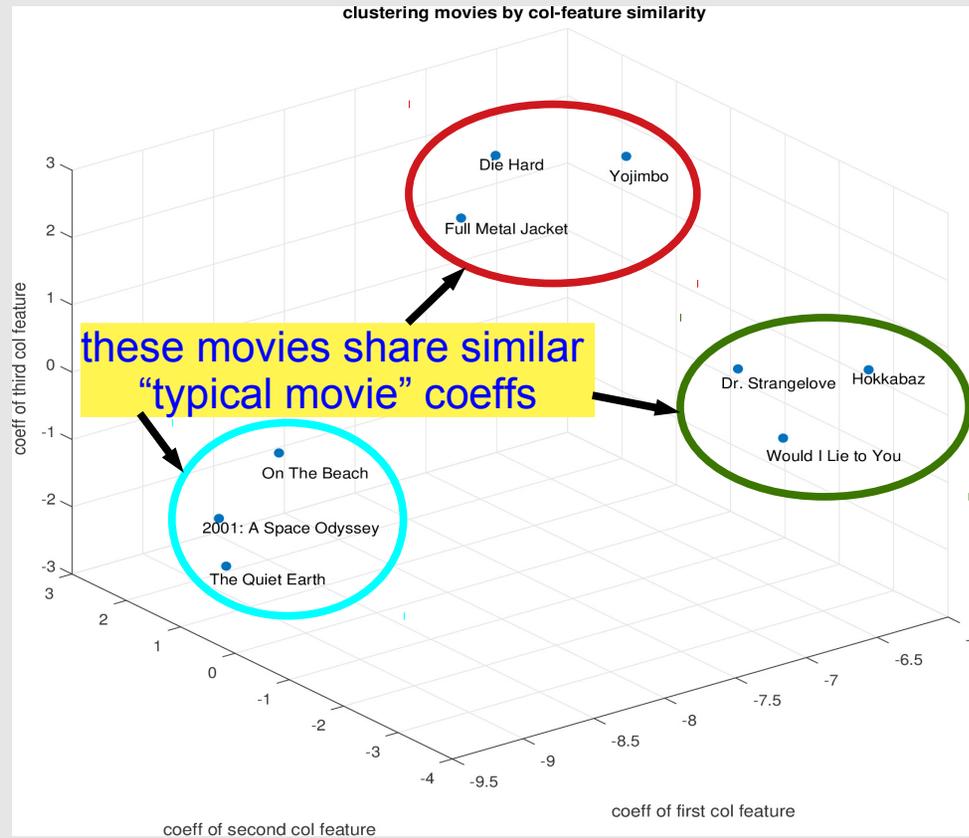
**Maharbiz and Roychowdhury**

**Lecture 9B**

**Data Analysis**

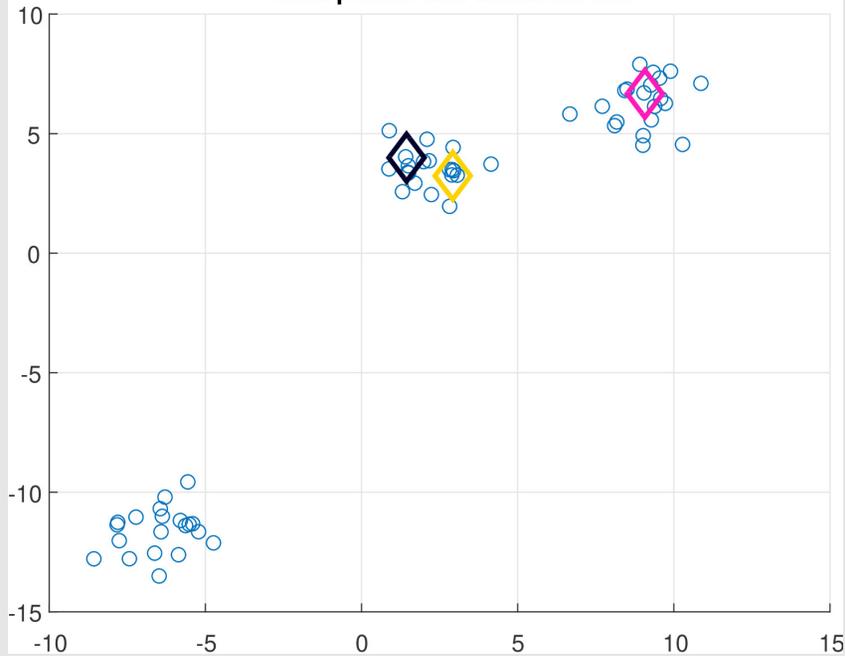
**K-Means Clustering (Figures)**

# Movie Rating Clusters

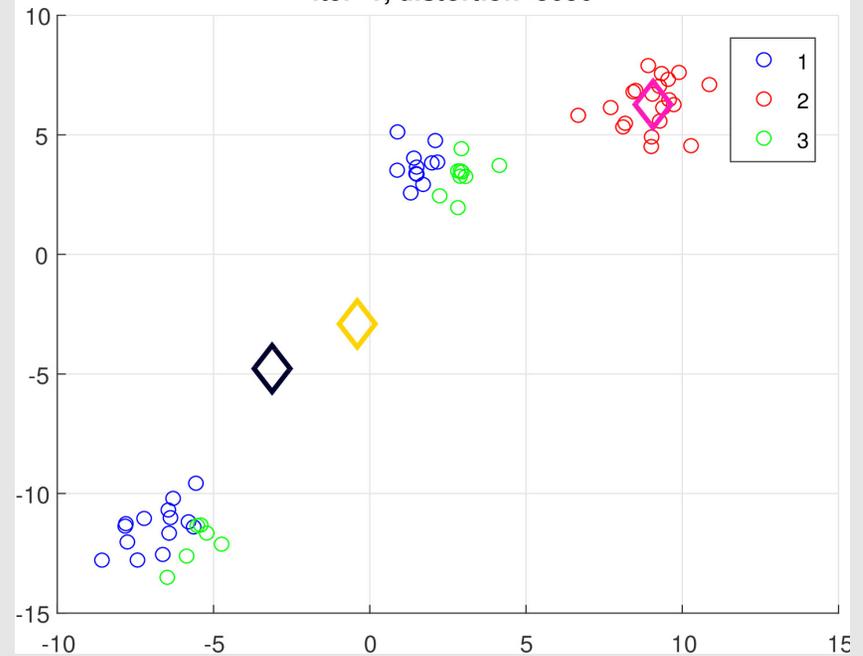


# 2D: 60 data points in 3 clusters

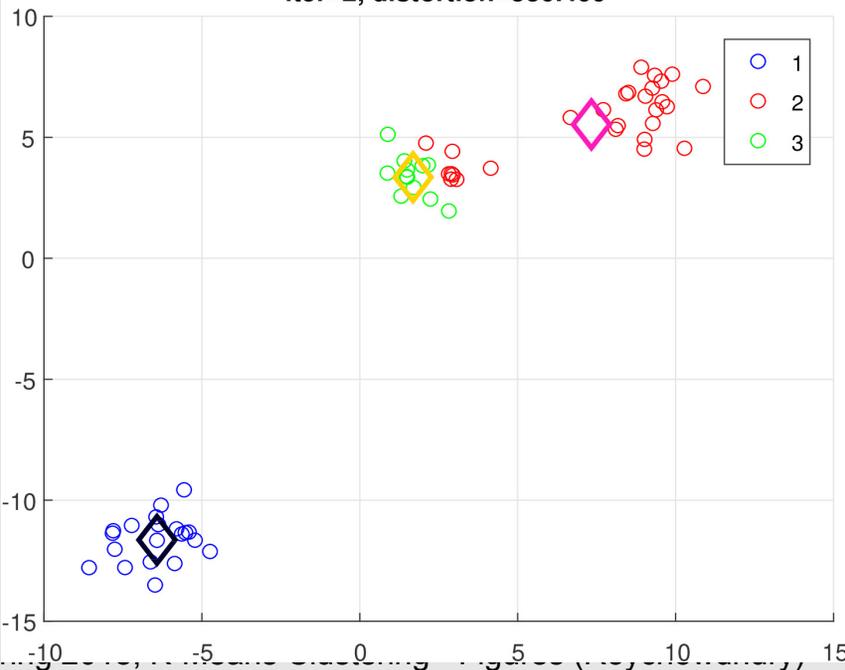
data points and initial means



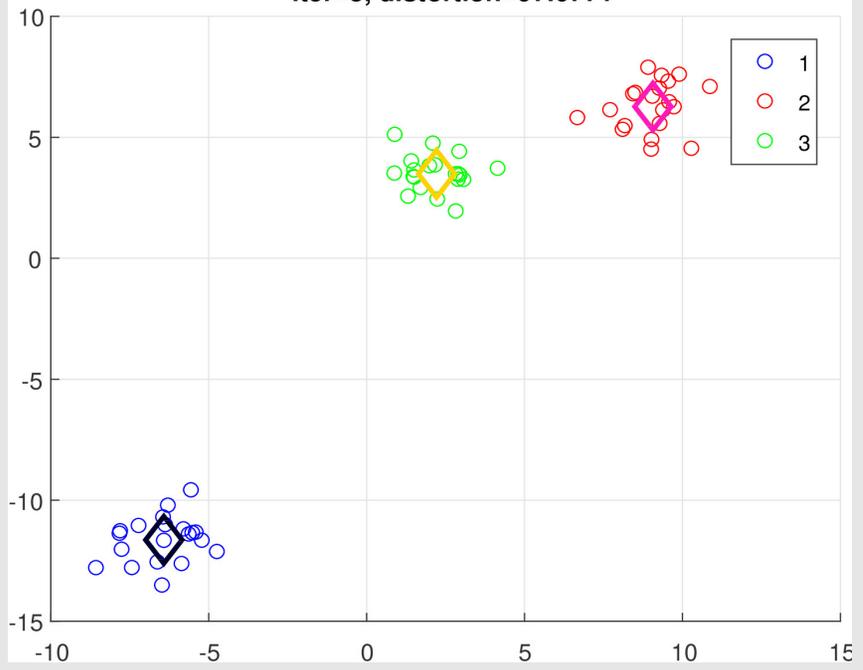
iter=1, distortion=3030



iter=2, distortion=336.499

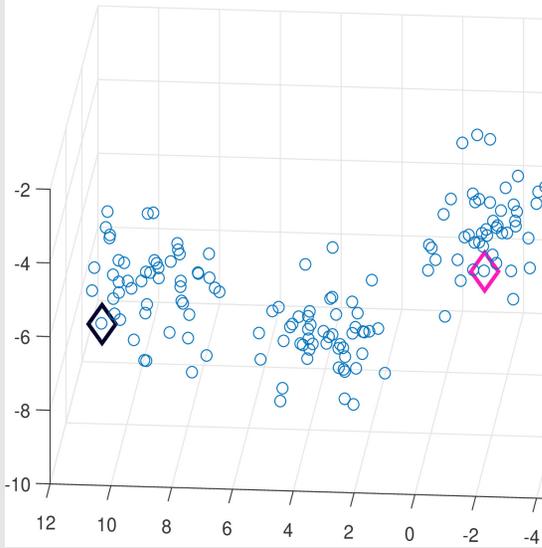


iter=3, distortion=97.6771



# 3D: 60 data points in 3 clusters

data points and initial means



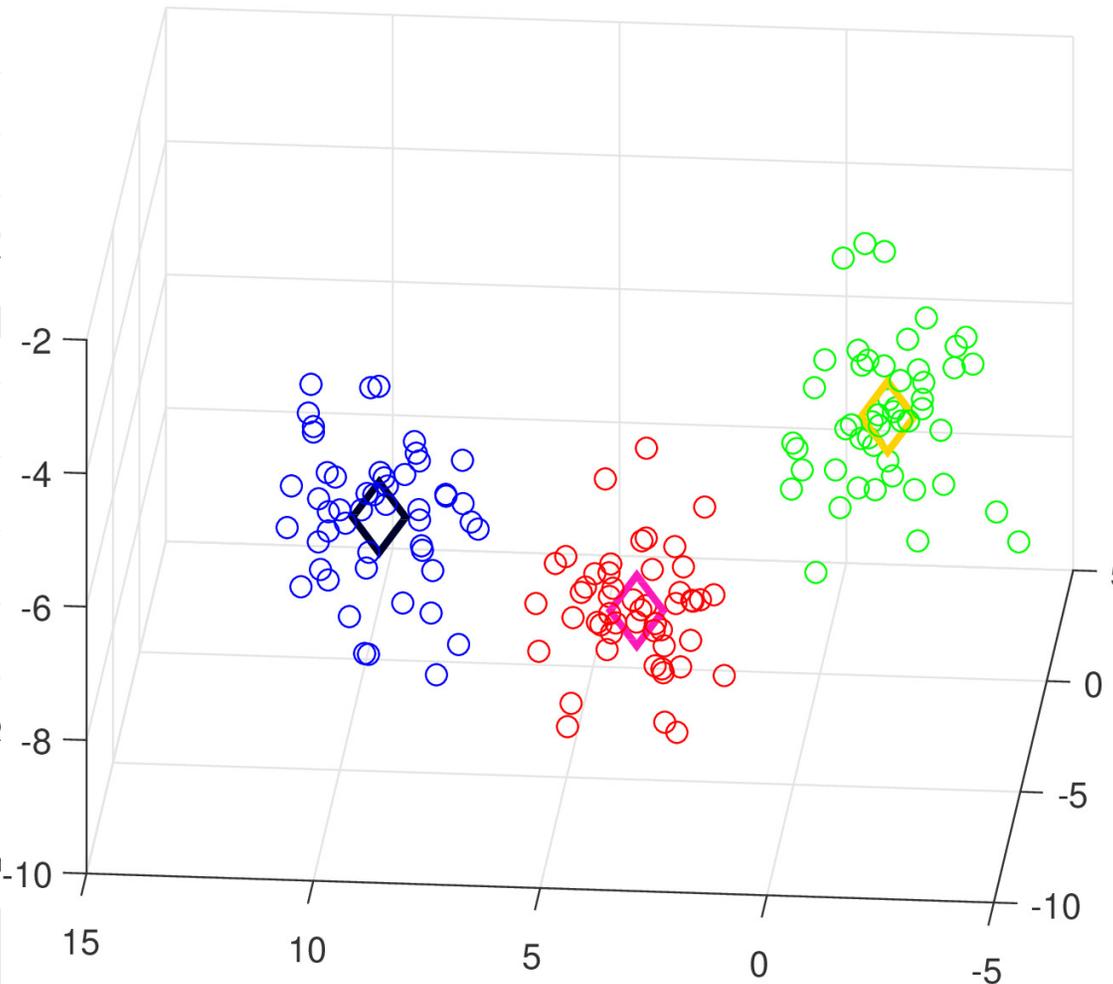
iter=1, distortion=1598.98



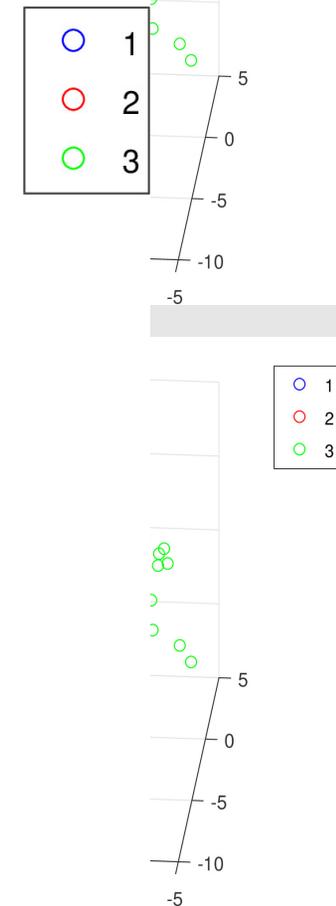
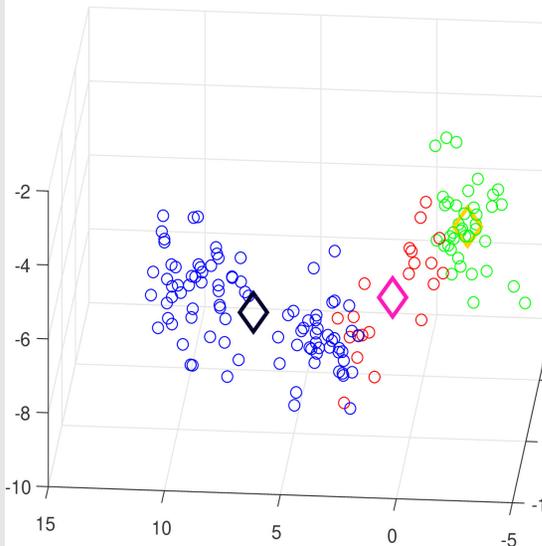
iter=2, distortion=1370.48



iter=6, distortion=461.95

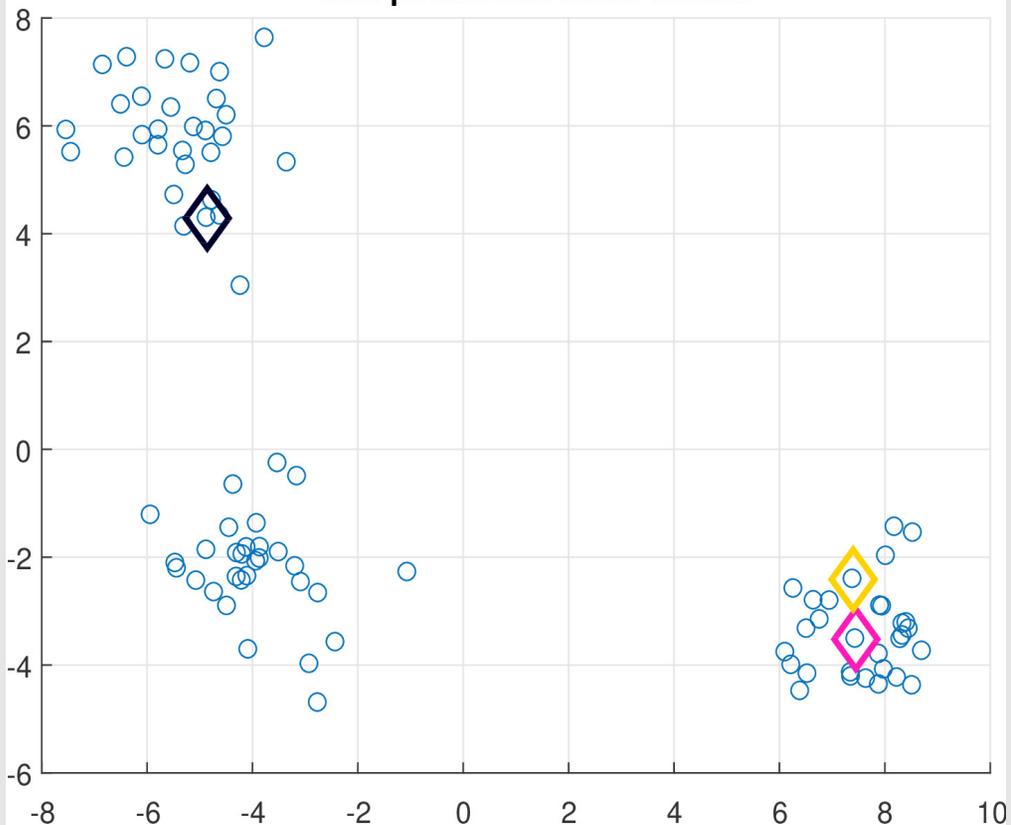


iter=3, distortion=1266.32

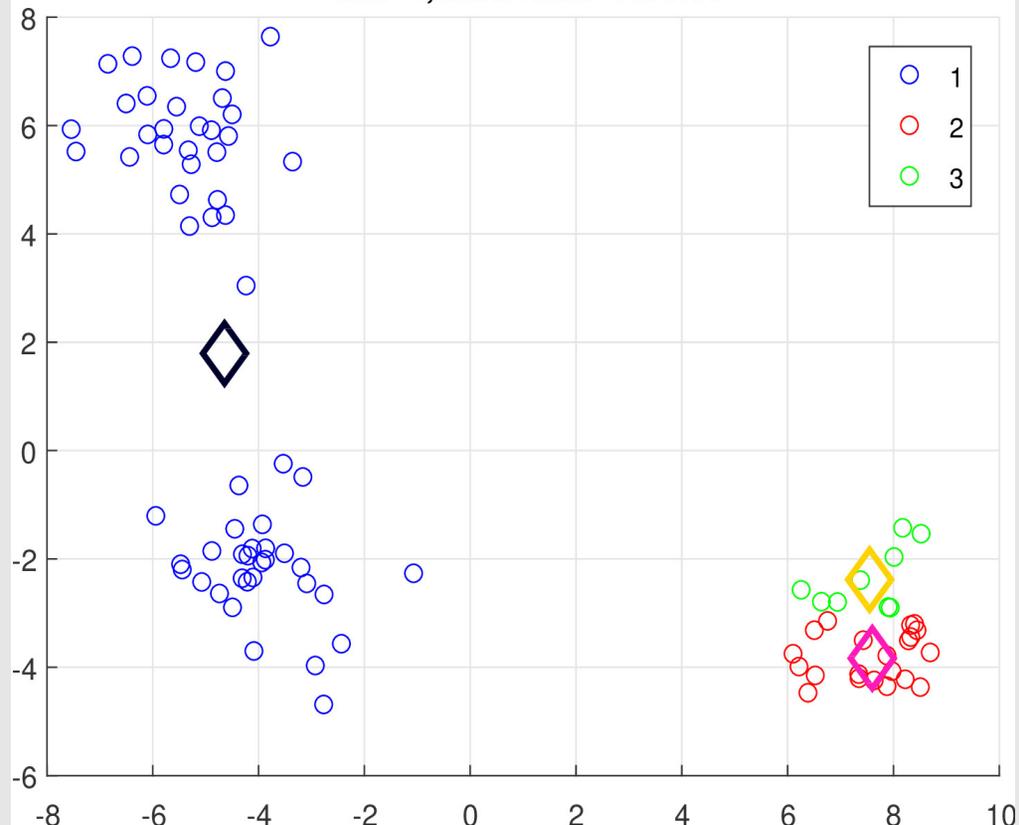


# 2D: another 60 data points in 3 clusters

data points and initial means

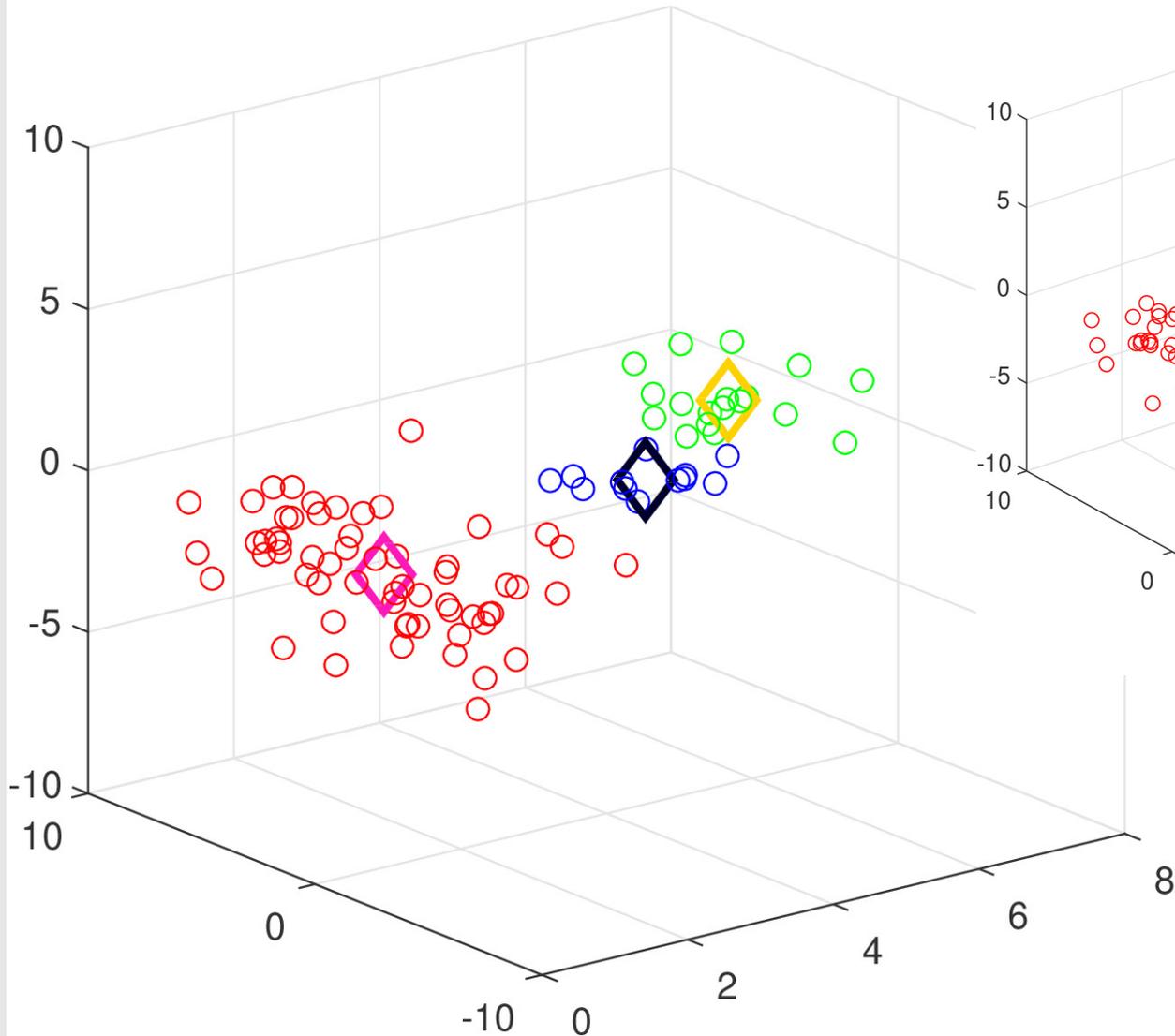


iter=1, distortion=1133.18

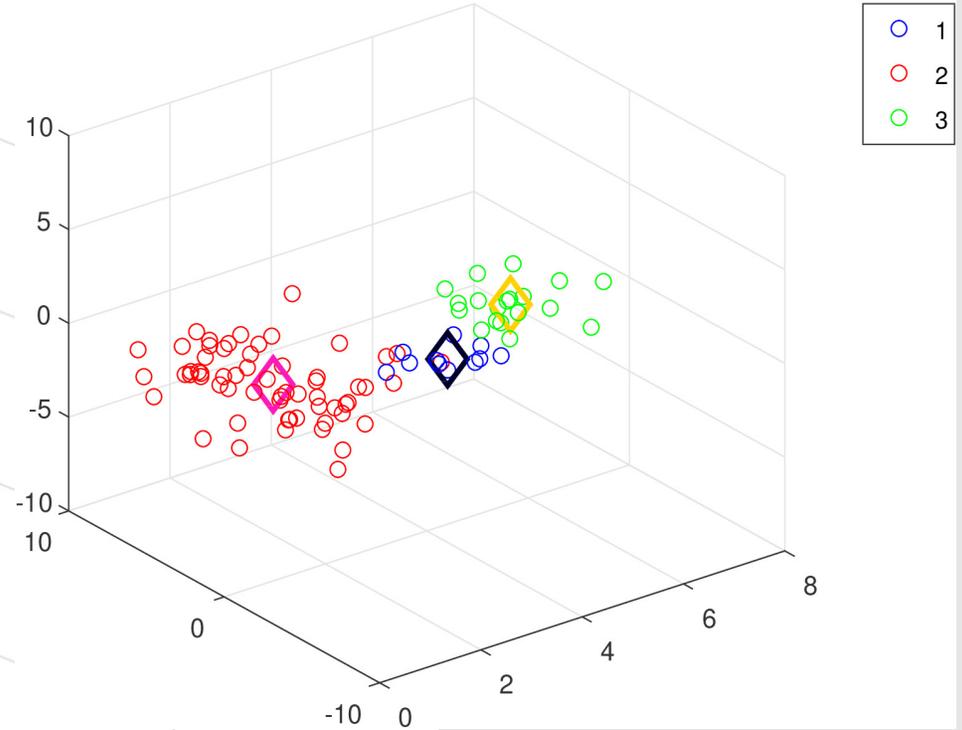


# 3D: another 60 data points in 3 clusters

iter=2, distortion=408.012



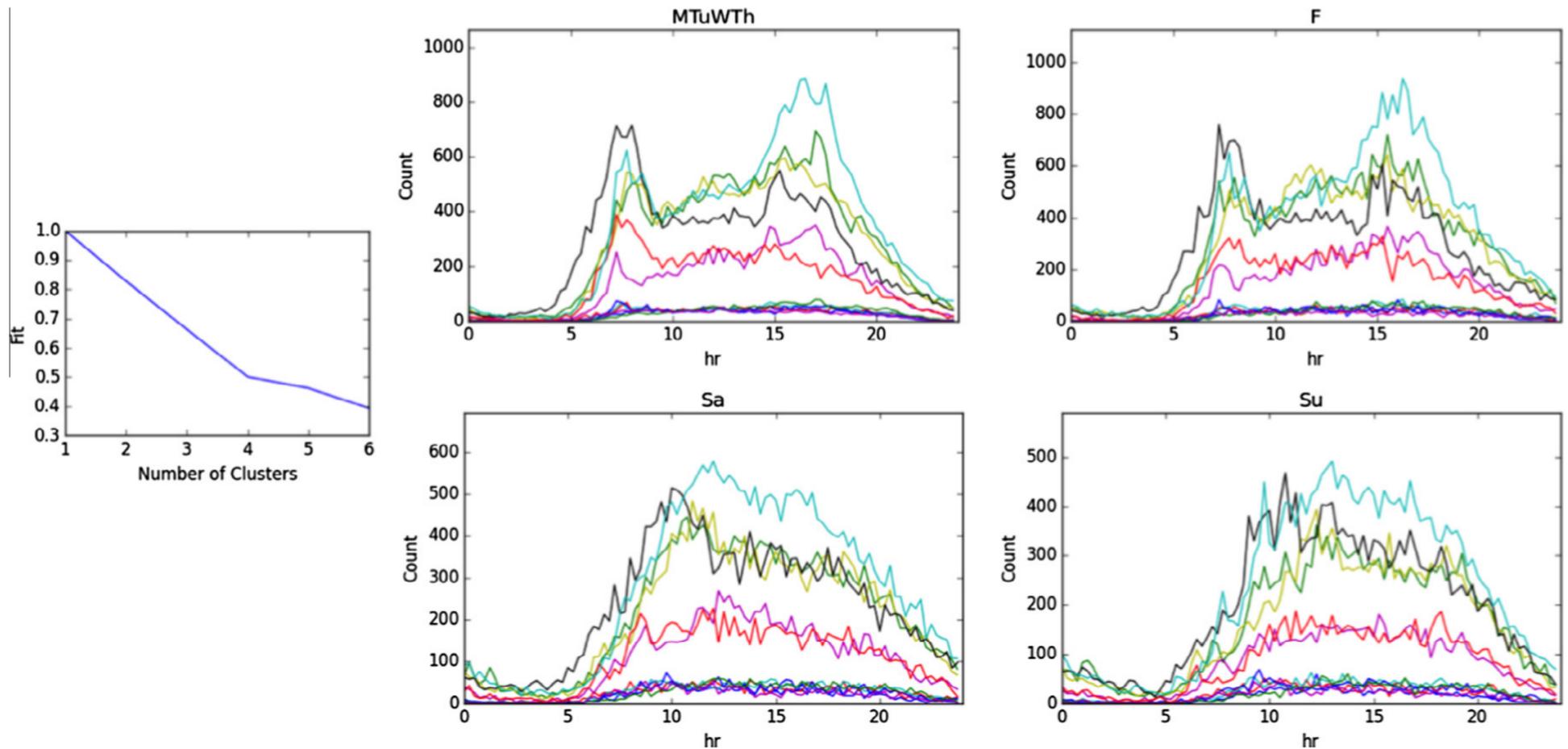
iter=1, distortion=409.134



- 1
- 2
- 3

# Clustering for Traffic Analysis

- Management of intersections with multi-modal high-resolution data
  - A. Muralidharan, S. Coogan, C. Flores, P. Varaiya
    - Transportation Research Part C: Emerging Technologies, July 2016



**Fig. 5.** Clustering of daily data for Dec 2014 to May 2015 in an intersection in Beaufort, SC.