

# Break & Quiz

**Q 1.1:** We have two datasets: a social network dataset  $S_1$  which shows which individuals are friends with each other along with image dataset  $S_2$ .

What kind of clustering can we do? Assume we do not make additional data transformations.

- A. k-means on both  $S_1$  and  $S_2$
- B. graph-based on  $S_1$  and k-means on  $S_2$
- C. k-means on  $S_1$  and graph-based on  $S_2$
- D. hierarchical on  $S_1$  and graph-based on  $S_2$

# Break & Quiz

**Q 1.1:** We have two datasets: a social network dataset  $S_1$  which shows which individuals are friends with each other along with image dataset  $S_2$ .

What kind of clustering can we do? Assume we do not make additional data transformations.

- A. k-means on both  $S_1$  and  $S_2$
- **B. graph-based on  $S_1$  and k-means on  $S_2$**
- C. k-means on  $S_1$  and graph-based on  $S_2$
- D. hierarchical on  $S_1$  and graph-based on  $S_2$

# Break & Quiz

**Q 1.1:** We have two datasets: a social network dataset  $S_1$  which shows which individuals are friends with each other along with image dataset  $S_2$ .

What kind of clustering can we do? Assume we do not make additional data transformations.

- A. k-means on both  $S_1$  and  $S_2$  **(No: can't do k-means on graph)**
- **B. graph-based on  $S_1$  and k-means on  $S_2$**
- C. k-means on  $S_1$  and graph-based on  $S$  **(Same as A)**
- D. hierarchical on  $S_1$  and graph-based on  $S_2$  **(No:  $S_2$  is not a graph)**

# Break & Quiz

**Q 1.2:** The CIFAR-10 dataset contains 32x32 images labeled with one of 10 classes. What could we use it for?

(i) Supervised learning (ii) PCA (iii) k-means clustering

- A. Only (i)
- B. Only (ii) and (iii)
- C. Only (i) and (ii)
- D. All of them

# Break & Quiz

**Q 1.2:** The CIFAR-10 dataset contains 32x32 images labeled with one of 10 classes. What could we use it for?

(i) Supervised learning (ii) PCA (iii) k-means clustering

- A. Only (i)
- B. Only (ii) and (iii)
- C. Only (i) and (ii)
- **D. All of them**

# Break & Quiz

**Q 1.2:** The CIFAR-10 dataset contains 32x32 images labeled with one of 10 classes. What could we use it for?

(i) Supervised learning (ii) PCA (iii) k-means clustering

- (i) **Yes: train an image classifier; have labels)**
- (ii) **Yes: run PCA on image vectors to reduce dimensionality**
- (iii) **Yes: can cluster image vectors with k-means**
- **D. All of them**

# Break & Quiz

**Q 2.1:** Can we do t-SNE on NLP (words) or graph datasets?

- A. Never
- B. Yes, after running PCA on them
- C. Yes, after mapping them into  $\mathbb{R}^d$  (ie, embedding)
- D. Yes, after running hierarchical clustering on them

# Break & Quiz

**Q 2.1:** Can we do t-SNE on NLP (words) or graph datasets?

- A. Never
- B. Yes, after running PCA on them
- **C. Yes, after mapping them into  $\mathbb{R}^d$  (ie, embedding)**
- D. Yes, after running hierarchical clustering on them



# Break & Quiz

**Q 2.1:** Can we do t-SNE on NLP (words) or graph datasets?

- A. Never **(No: too strong)**
- B. Yes, after running PCA on them **(No: can't run PCA on words or graphs directly. Need vectors)**
- **C. Yes, after mapping them into  $R^d$  (ie, embedding)**
- D. Yes, after running hierarchical clustering on them **(No: hierarchical clustering gives us a graph)**