Kmeans Clustering

```
# (0) import Packages
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import Kmeans
from sklearn.metrics import silhouette_score

# (1) Read in your dataset

# (2) Pre-processing & cleaning

# (3) Standardize data
X_std = StandardScaler().fit_transform(data)

# (4a) Decide on K

# (4b) Run the Kmeans
kmeans = KMeans(n_clusters= K).fit(X_std)

# (5) Interprete results
#     If unsatisfactory, go back to (2) and iterate
```

Run your code on the following datasets:

(1) BirthDeathRates.

(2) BankLoan

Hand-in Workshop:

Perform Kmeans on the restaurant dataset.

You can use the PCs or start afresh

Do the cluster validation (ref: slide57-58)

Deliverables (19July, 2359hrs):

Python Code

Clusters profiles

Interpretation

Note: **hand in ONE .py file only**