

K Means and GMM Procedure Example

May 8, 2020

0.0.1 K Means and GMM Procedure Example

This notebook provides an example of using procedure.py for k Means and GMM clustering. 4 files, clustering with the above 2 methods with and without PCA are exported as a csv file.

```
In [2]: import pandas as pd
        from procedure import *
```

```
norm_roi_df = pd.read_csv("data\\BM_AM_180808_190401_norm.csv")
norm_roi_df.head(5)
```

```
Out[2]:
```

	Unnamed: 0	image	total cell index	ind cell index	Area	Circ.	AR	\
0	0	21_s1m1	0	0	0.049	0.857	1.450	
1	1	21_s1m1	1	1	0.058	0.878	1.135	
2	2	21_s1m1	2	2	0.068	0.708	1.663	
3	3	21_s1m1	3	3	0.067	0.763	1.336	
4	4	21_s1m1	4	4	0.099	0.799	1.435	

	Round	Solidity	% depth	...	z norm SCGN_median	z norm Kv2.2_median	\
0	0.690	0.897	5.259663	...	-0.020591	0.302211	
1	0.881	0.892	22.574779	...	2.700736	1.379877	
2	0.601	0.864	70.322646	...	-0.369889	0.786296	
3	0.749	0.863	54.642624	...	-0.155288	0.734430	
4	0.697	0.887	17.882495	...	-0.280852	0.550017	

	z norm DAPI3_median	z norm CAVIII_median	z norm Rec_median	\
0	-0.746300	-0.565612	1.181044	
1	0.631802	-0.585292	0.171825	
2	2.512683	-0.948704	-1.030963	
3	1.223703	-0.391122	1.201069	
4	0.442898	-0.593164	0.370732	

	z norm DAPI4_median	z norm Calb_median	z norm PKC_median	\
0	-1.485921	-1.086136	-0.721506	
1	-1.464316	-1.086136	-0.996451	
2	-1.198727	-0.961293	-0.966496	
3	-1.373327	-0.968633	-0.909795	
4	-1.464316	-1.053046	-0.811371	

	z norm Parv_median	z norm DAPI5_median
0	-0.334144	-1.151339
1	-0.508083	-0.387308
2	-0.540182	1.400713
3	-0.540182	0.586593
4	-0.431165	-0.597813

[5 rows x 197 columns]

```
In [3]: procedure(norm_roi_df, 'drop', ['Islet1', 'CD15'], 7, "hi_")
```

K means PCA: [1 4 2 ... 1 2 0]

K means No PCA: [2 2 8 ... 7 8 7]

GMM: [0 3 2 ... 0 5 3]

GMM No PCA: [4 2 1 ... 5 6 2]

```
In [ ]:
```