Lab 4 STAT 315-463: Multivariable Statistical Methods and Applications

Thursday 16 March 2023

The 'Data'

For this lab, the data is a black and white photo of Karl Perason, which is available from Learn. You will the code below helpful

```
library(jpeg)
# Importing Karl Pearson
pic <- readJPEG("Pearsonphoto.jpg")

#As the imported image appears as an array, the data we will use will be the first part of the array
picdata<-pic[,,1]
```

What to do.

Explore various methods of dimension reduction and clustering as looked at in the lectures in Weeks 3 and 4. These are:

- Principal Component Analysis (PCA).
  - When looking at PCA, you may find it interesting to see how well a low dimensional approximation of the data can replicate the photo. If you wish to re-export the image, use the function `writeJPEG(lddata, target="ldimage.jpeg", quality=1)`, where `lddata` is the matrix containing the low dimensional representation of the data, and `ldimage.jpeg` is the name of the exported image.
- hierarchical clustering
- K-means clustering
- Clustering based on multivariate normal mixtures.
  - For the three clustering techniques, it might be interesting to investigate if the clusters determined correspond to meaningful features in the photo.