## Unsupervised Learning Final Exam Review Questions

* What is principal component analysis (PCA)

Dimensionality reduction

* What is a loading vector points? Loading vector / First PCA points in the direction of maximum variance
* What is a PCA score

Linear combination of corresponding values of the feature is multiplied by loading vector

* What is a projection

Projecting the features in the direction of loading vector. It is equivalent to linear combination.

* What is a scree plot

Variance explained in each PCA.

* Name 3 use cases for PCA in data science

<https://www.linkedin.com/pulse/dos-donts-principal-component-analysis-syed-sadat-nazrul/>

* Assuming a training data matrix of size 100x3, what is the size of the resulting loading vector matrix assuming the maximum number of loading vectors are calculated.

We can calculate 3 vectors for 3 features. 3X3

* How would we calculate the score for the 5th row of data using the 1st principal component. What is the shape of the resulting score vector, matrix, or scaler.

M = min ((n-1), p) n: number of rows, p: number of columns

* How can we perform feature selection using PCA
* What data pre-processing is required to perform PCA analysis

Centering the data. In some cases, to also divide by std.

* How can we get the original data back from PCA scores