**Lab Sheet 5**

**Check out:**

[**https://www.inferentialthinking.com/chapters/intro**](https://www.inferentialthinking.com/chapters/intro)

[**http://data8.org/**](http://data8.org/)

**2.**

Take a dataset of your choice.

Write your interpretation of the data like type of attribute, distribution, correrlation using the different plots like histogram, box-plots, scatter plots for each attribute

Perform the preprocessing steps –

1. Data cleaning by handling the missing values
2. Apply a normalization function to reduce your data
3. Find data correlation using covariance and heatmap thereby reducing the data to remove redundant rows and columns. You may apply sampling of your data for row reduction.
4. Apply PCA over the sampled data.