**Lab1: Exploratory data analysis**

**Due** Sep 18 by 11:59pm

Darian Yang

We provided you with an obfuscated scientific dataset. Each row is an observation. The first column, titled “experimental\_proprty” is your target property of interest. All other columns encode features. There are no observation IDs.  
  
**You will have to perform EDA analysis:**

* Load data, prepare for analysis, process if necessary
* Analyze types of data
* Find and process missing and erroneous features
* Find outliers (if any)
* Find highly correlated variables (if any).
* Find if the target variable is correlated with any features.
* Use PCA to plot data in 2D and color code by the target property. Do you see any patterns?
* Prepare a short write-up describing your processing technics and choices above.

Bonus Qs:

* Use any non-linear dimensionality reduction method. Plot data in 2D and color code by the target property. Compare observed picture with PCA.
* Surprise me! Uncover hidden patterns and find non-trivial relationships in the data