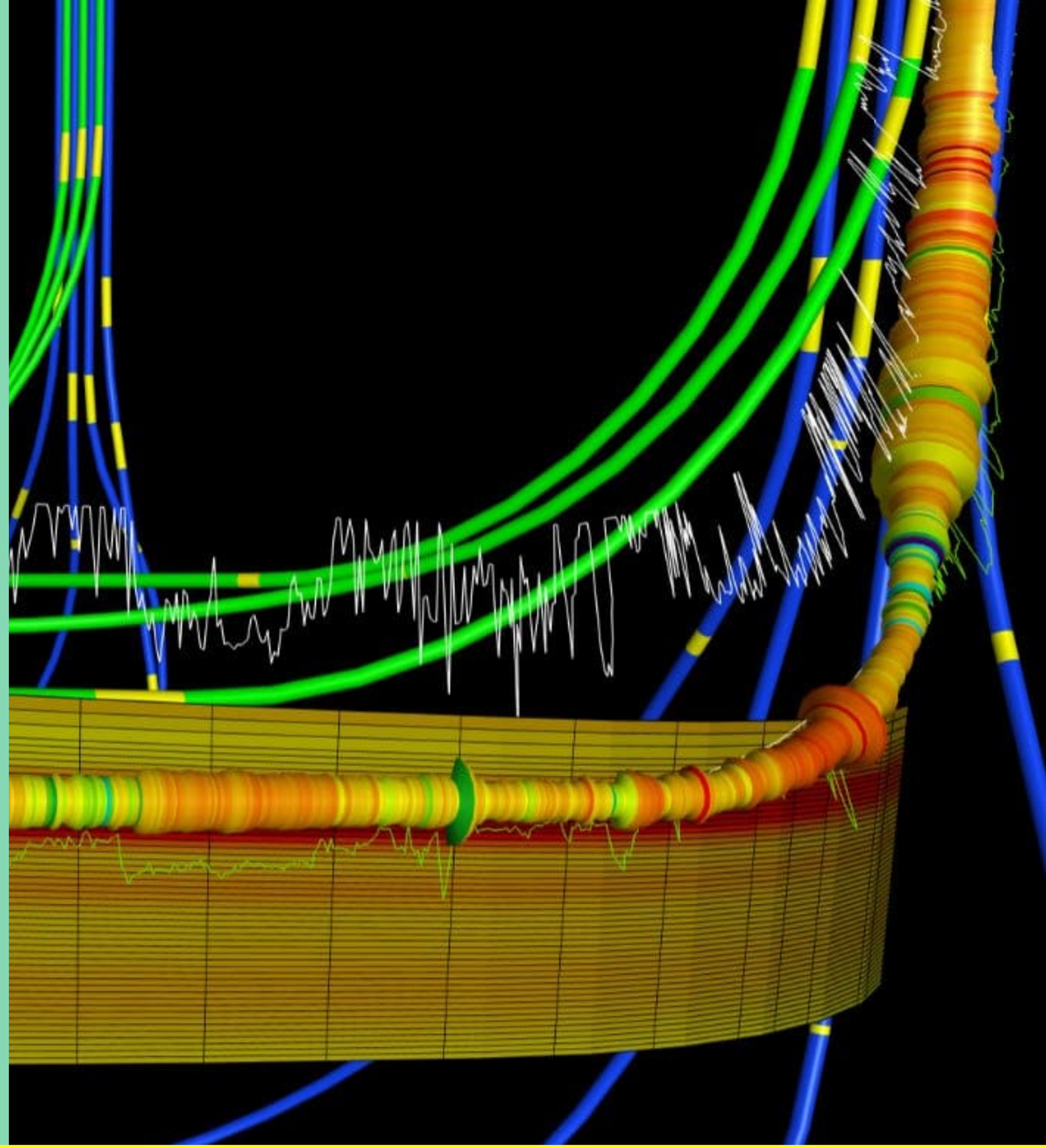
WLEFA

GROUP CASE STUDY

# Leveraging Machine Learning to Identify Lithofacies using Well Logs

Under the guidance of:-
**Prof. Partha Pratim Mandal**



# GROUP 1

## Group Members

Aditya Chowdhury

Shashwat Srivastava

Anmol Vishwakarma

Rajkumar Mondal

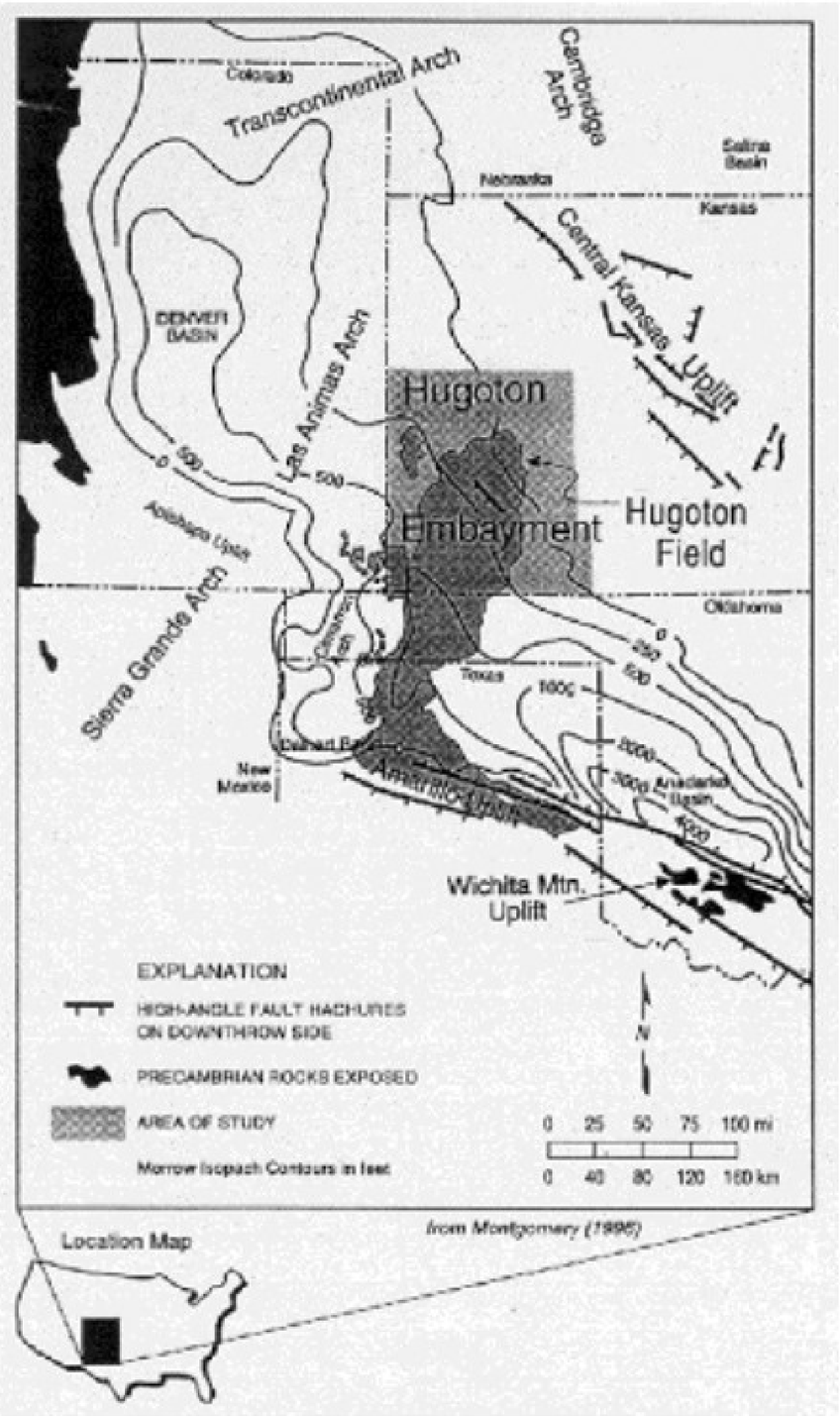
Anwesha Kodal

# Objectives:

## *Identifying the underlying geological problem*

| Facies | Description | Label | Adjacent facies |
|---|---|---|---|
| 1 | Nonmarine sandstone | SS | 2 |
| 2 | Nonmarine coarse siltstone | CSiS | 1,3 |
| 3 | Nonmarine fine siltstone | FSiS | 2 |
| 4 | Marine siltstone and shale | SiSh | 5 |
| 5 | Mudstone | MS | 4,6 |
| 6 | Wackestone | WS | 5,7,8 |
| 7 | Dolomite | D | 6,8 |
| 8 | Packstone-grainstone | PS | 6,7,9 |
| 9 | Phylloid-algal bafflestone | BS | 7,8 |

*Hugoton & Panoma Gas Field*



# *Data Visualisation*

## AIM:

*To understand the distribution of log data*

01

**HISTOGRAM PLOTS**

02

**PAIR PLOTS**

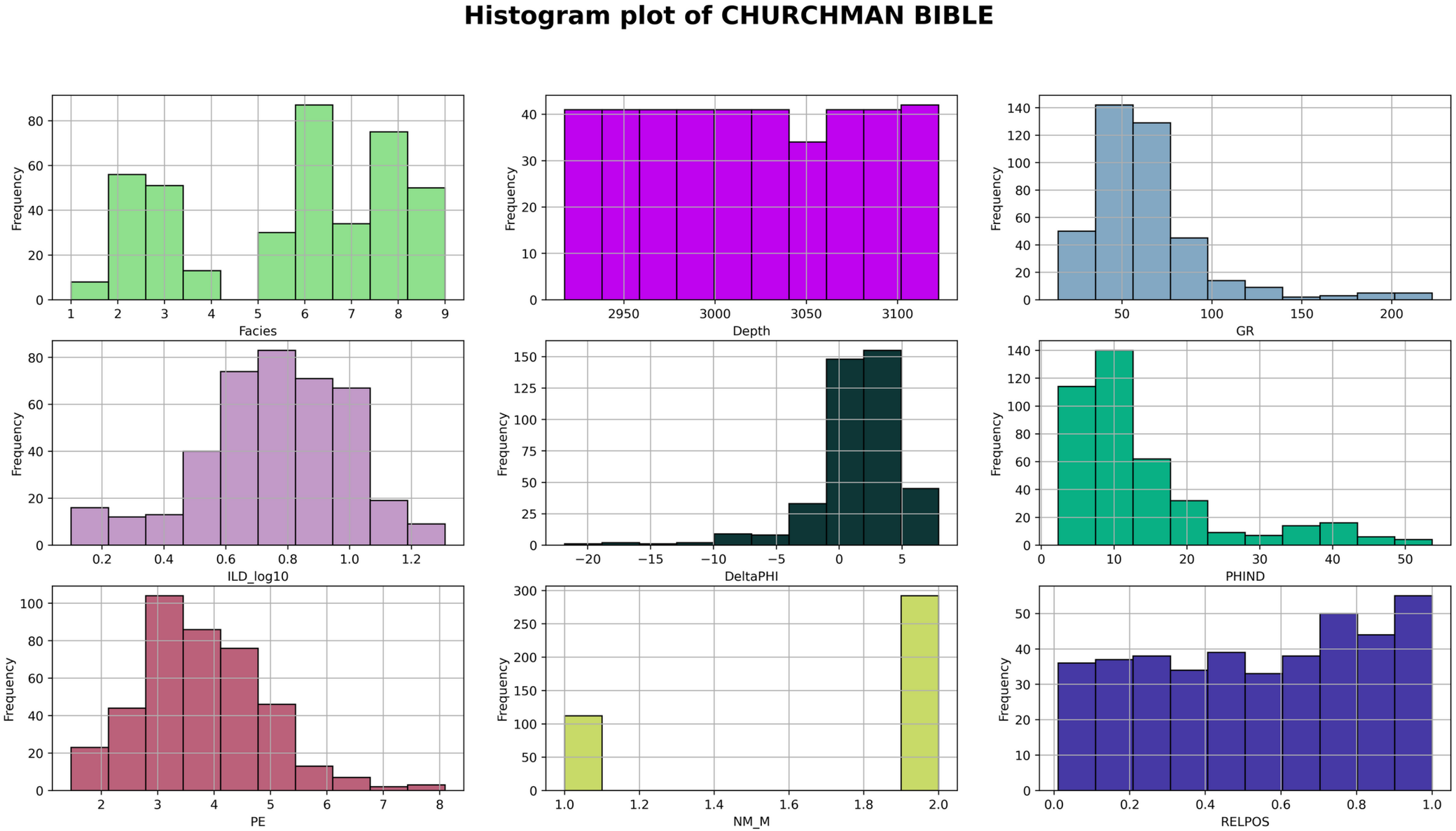
03

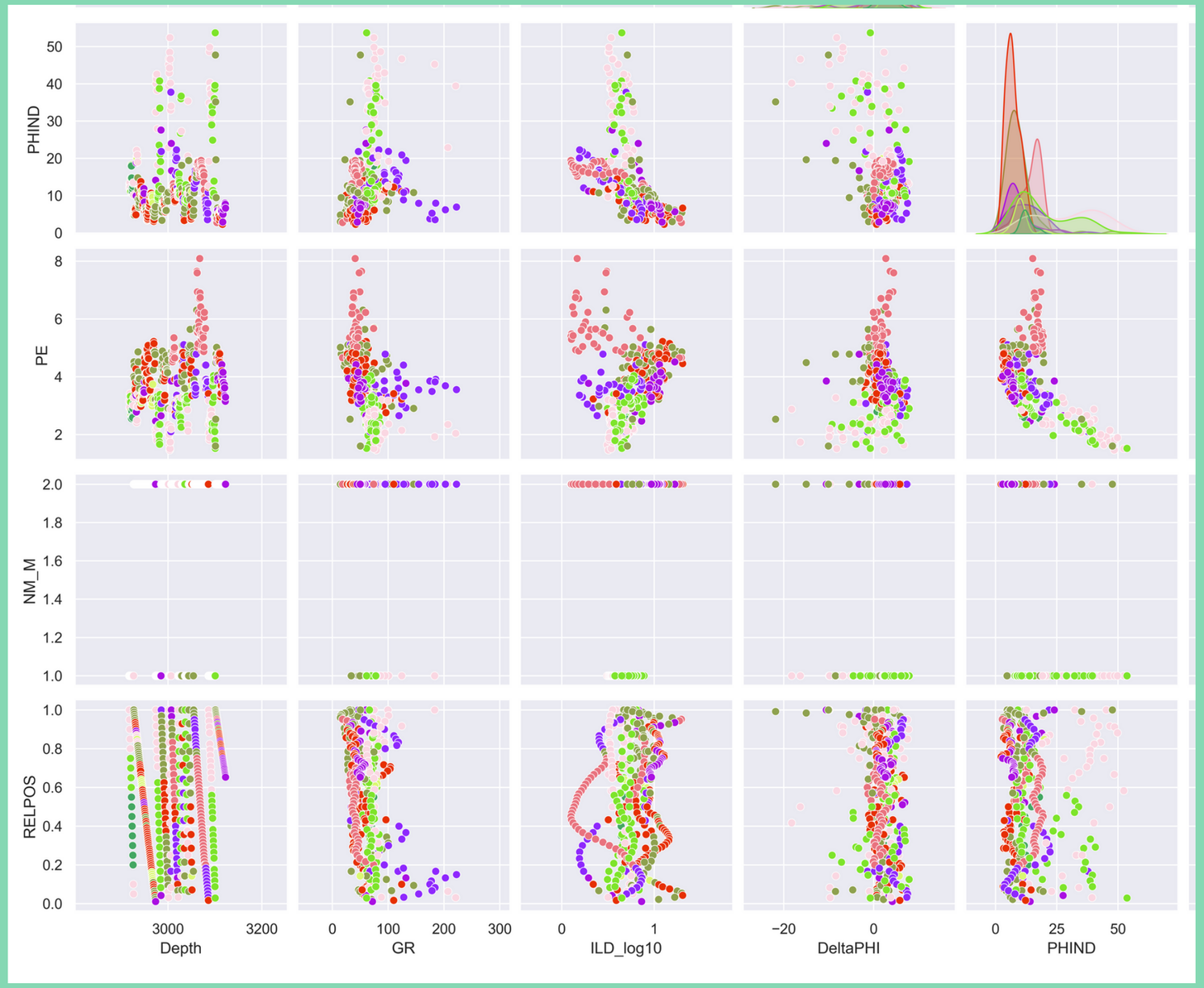
**VIOLIN PLOTS**

04

**STATISTICAL CORRELATIONS**

# **Histogram Plot:** *An example-*



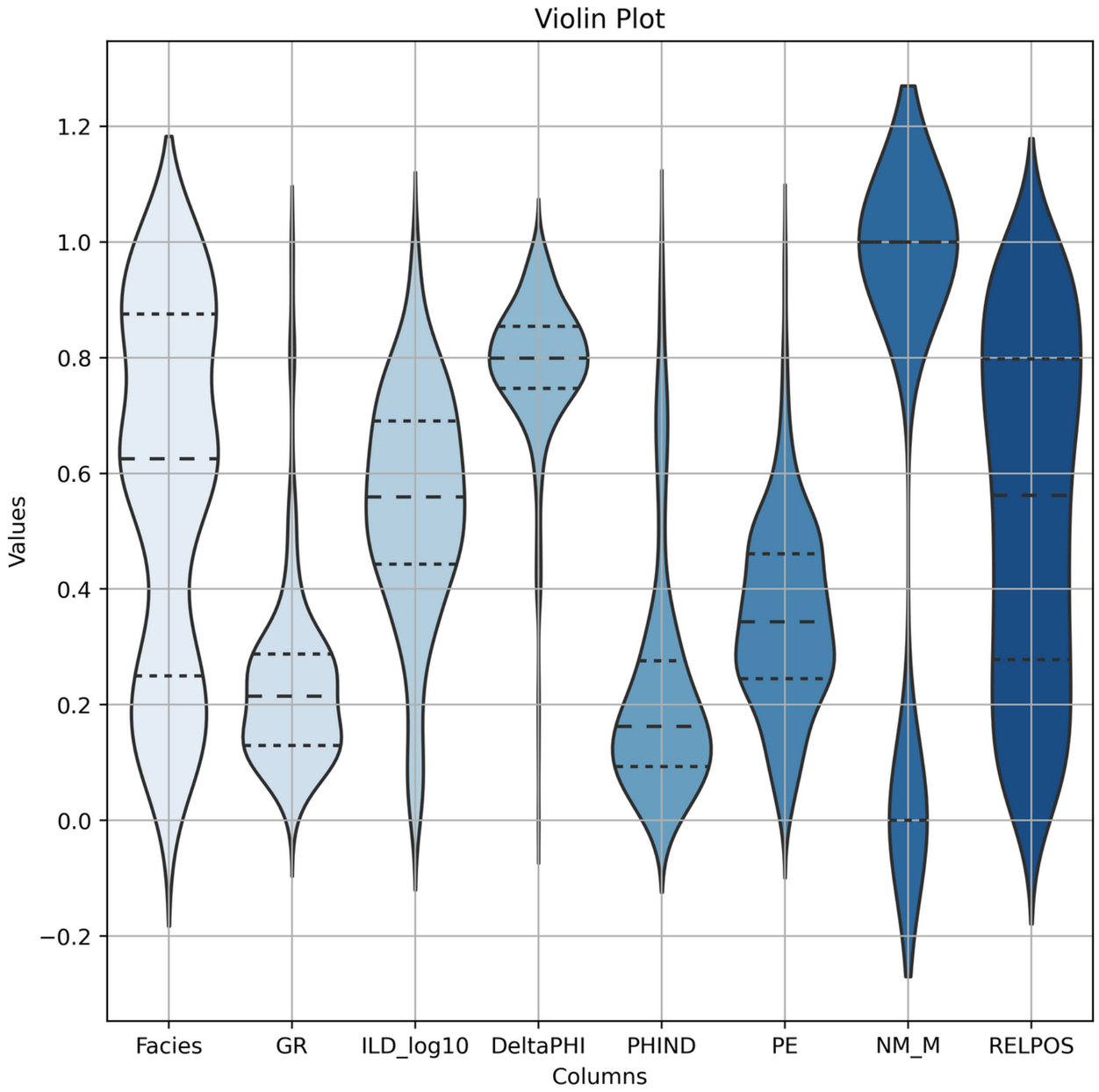



# Pair Plot:

*An example-*

# CHURCHMAN BIBLE

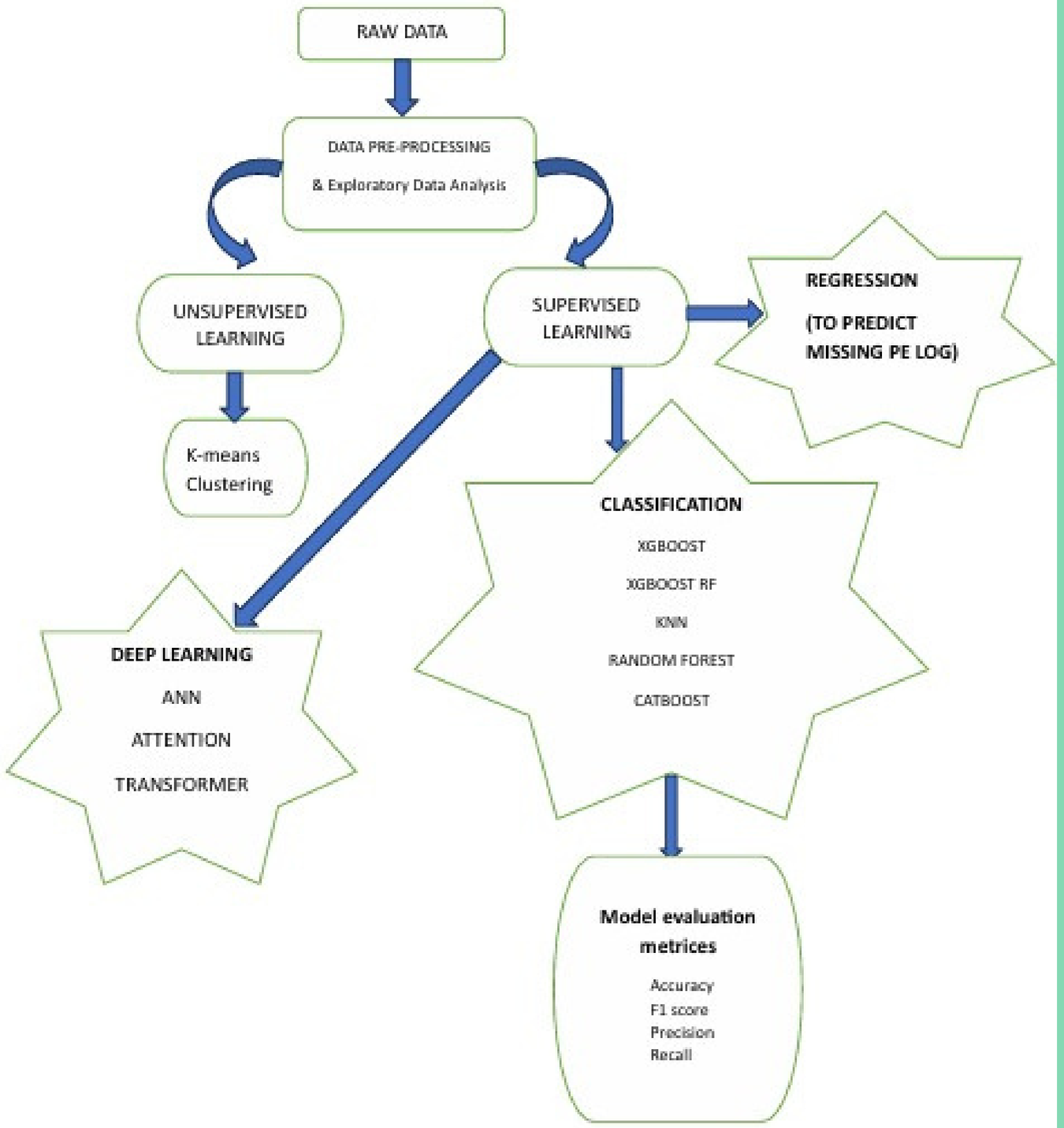
**We will use this <u>well</u> for our model validation due to presence of almost all lithofacies.**




# Violin Plot:

*An example-*

# CHURCHMAN BIBLE




# WORKFLOW USED

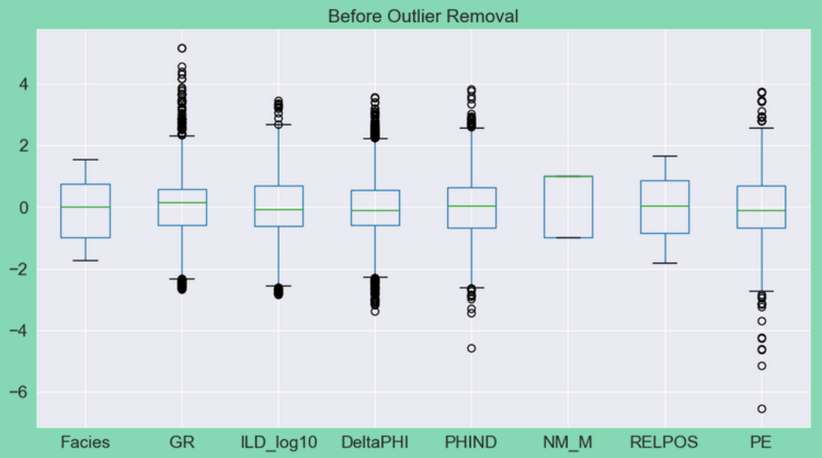
# Data Pre-processing:
*A step-by-step approach*

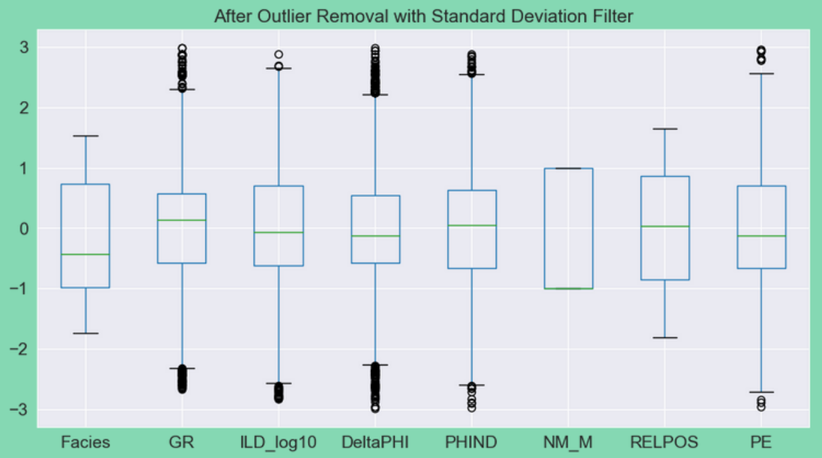
# Outlier Removal Methodologies

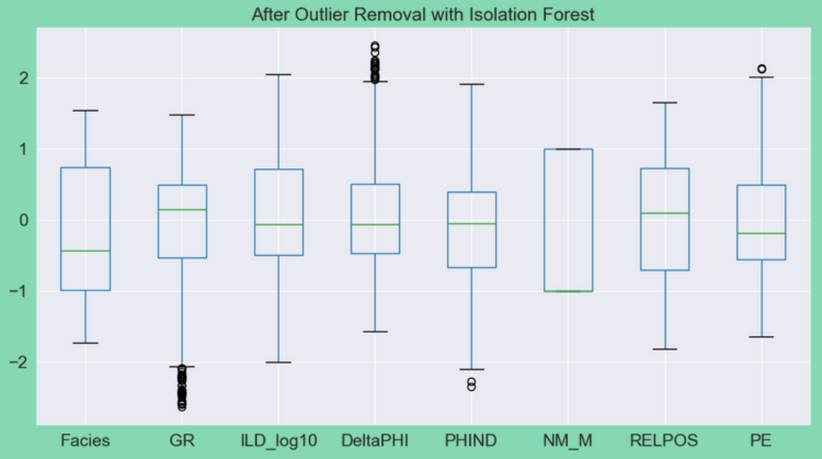
1. Standard Deviation Filter
2. Isolation Forest
3. Minimum Covariance
4. Local Outlier Factor
5. One-Class SVM

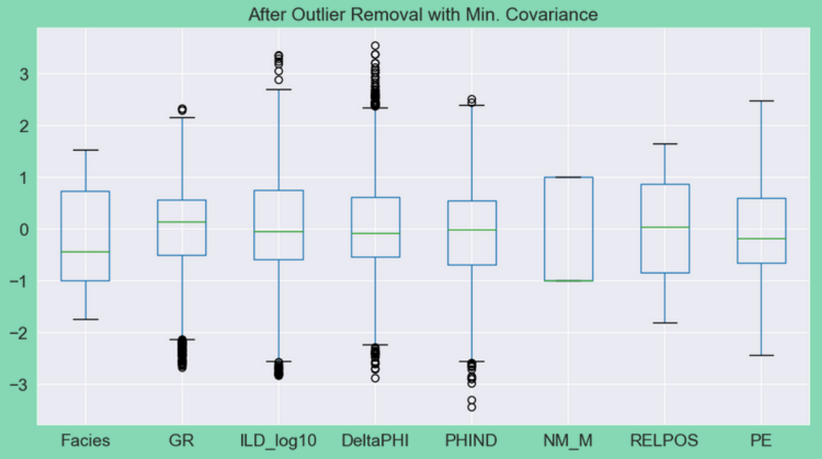
```
Number of points before outliers removed                    : 3244
Number of points after outliers removed with Standard Deviation: 3158
Number of points after outliers removed with Isolation Forest  : 1616
Number of points after outliers removed with Min. Covariance   : 2907
Number of points after outliers removed with One-class SVM     : 2910
```

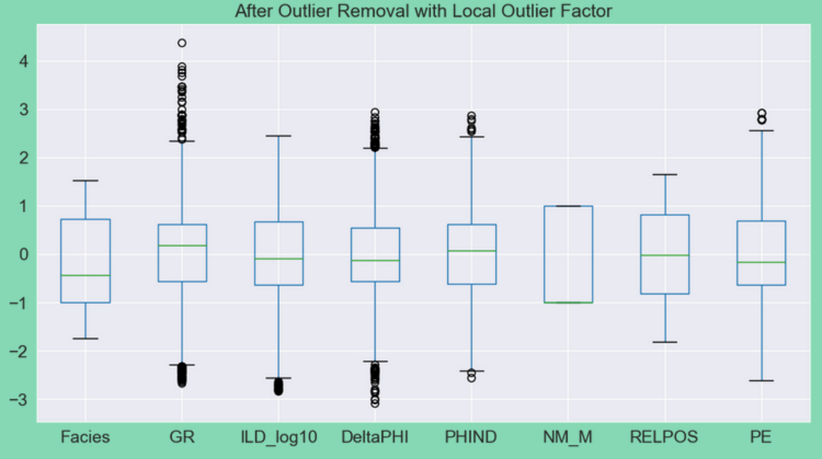
**SVM** used for outlier removal for all datapoints

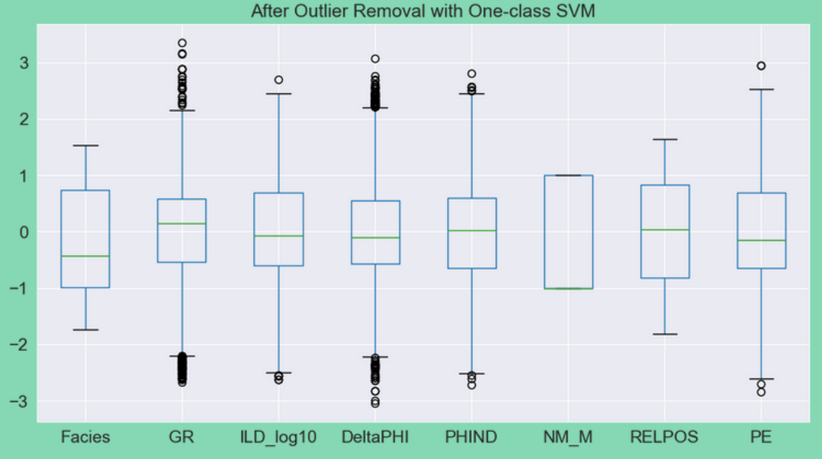


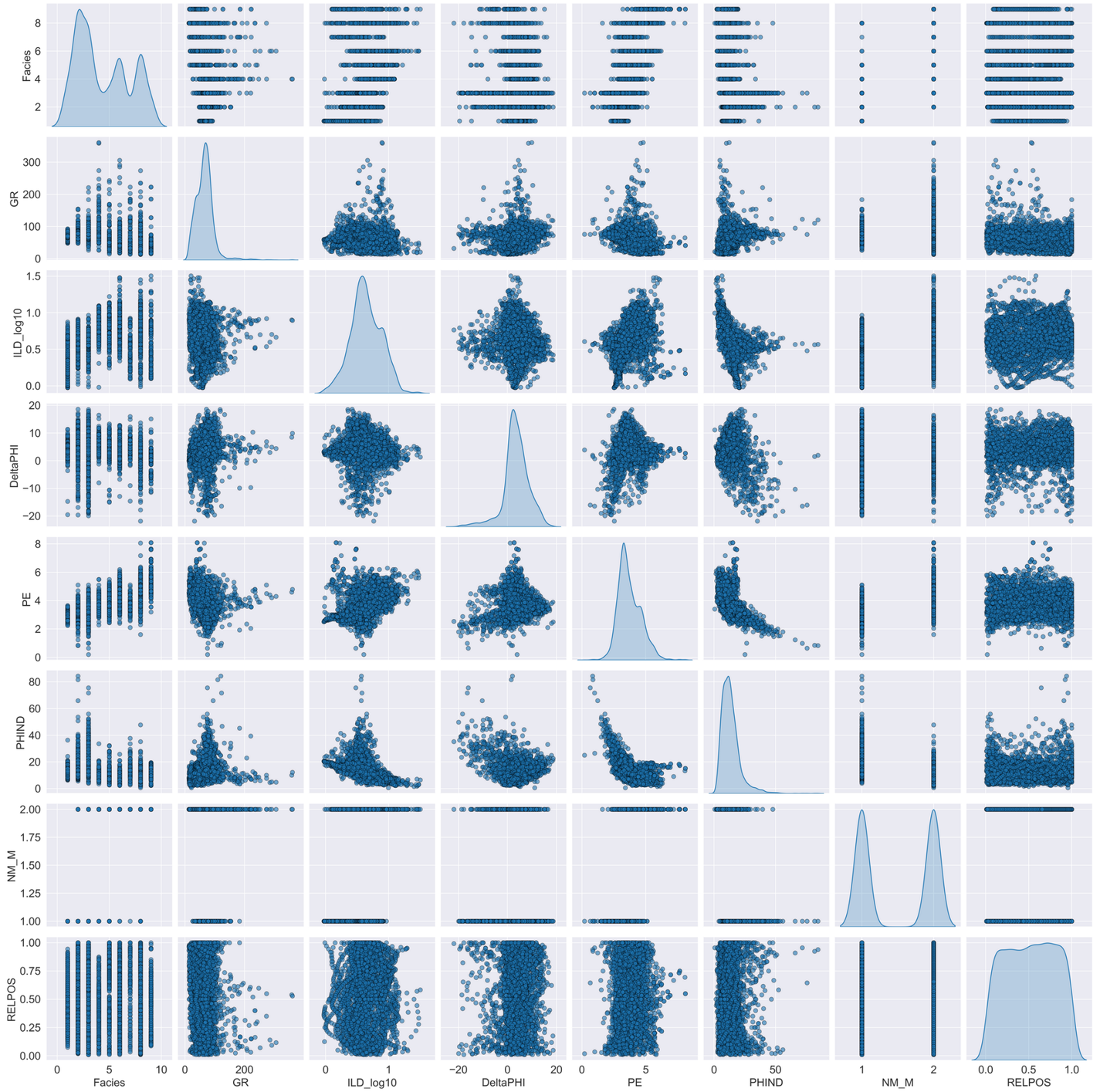




# **Box-Plot:** *Outlier visualisation*

# Training

# Normalisation

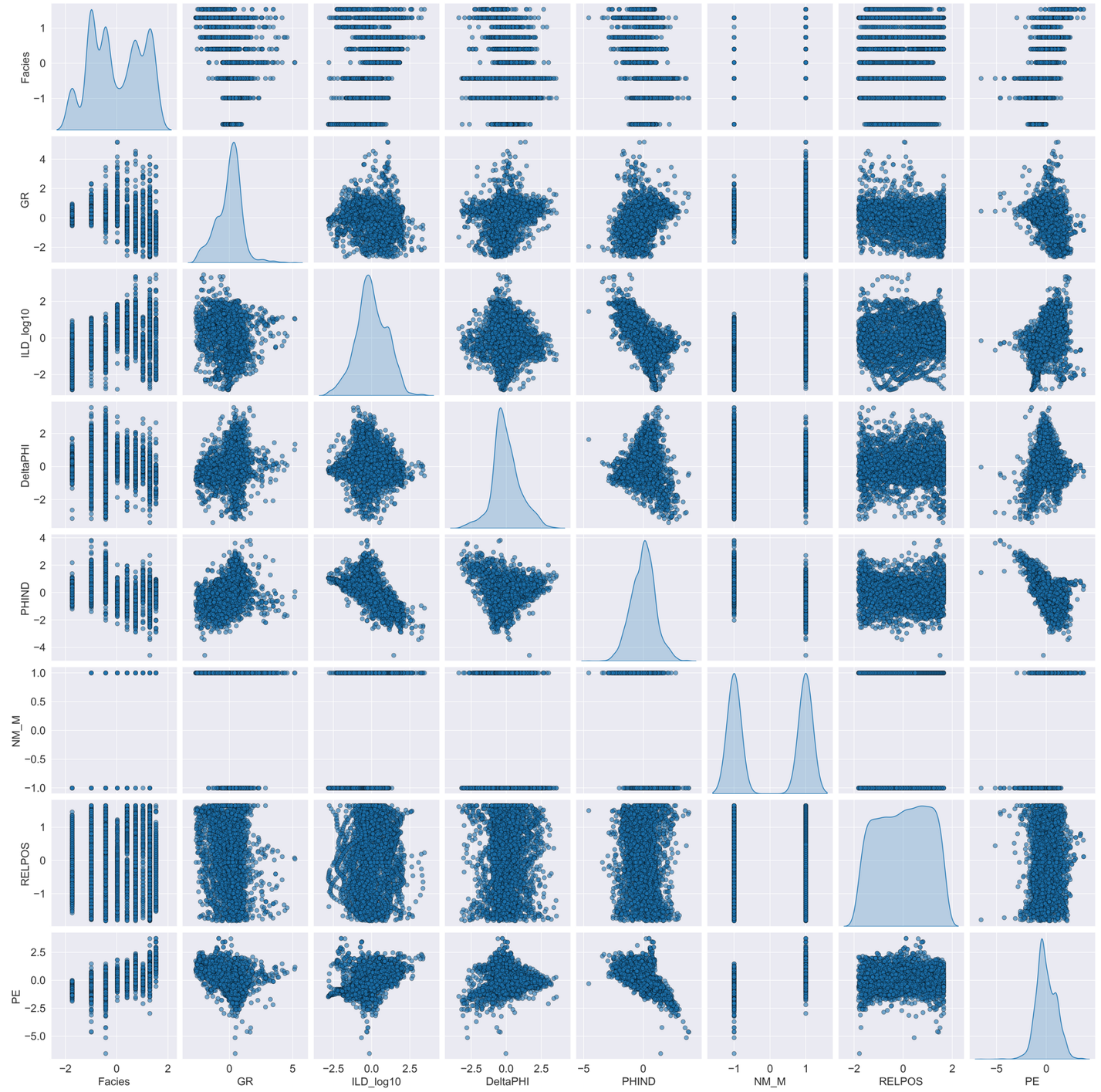
# Outlier-Removal



# Training

# Normalisation

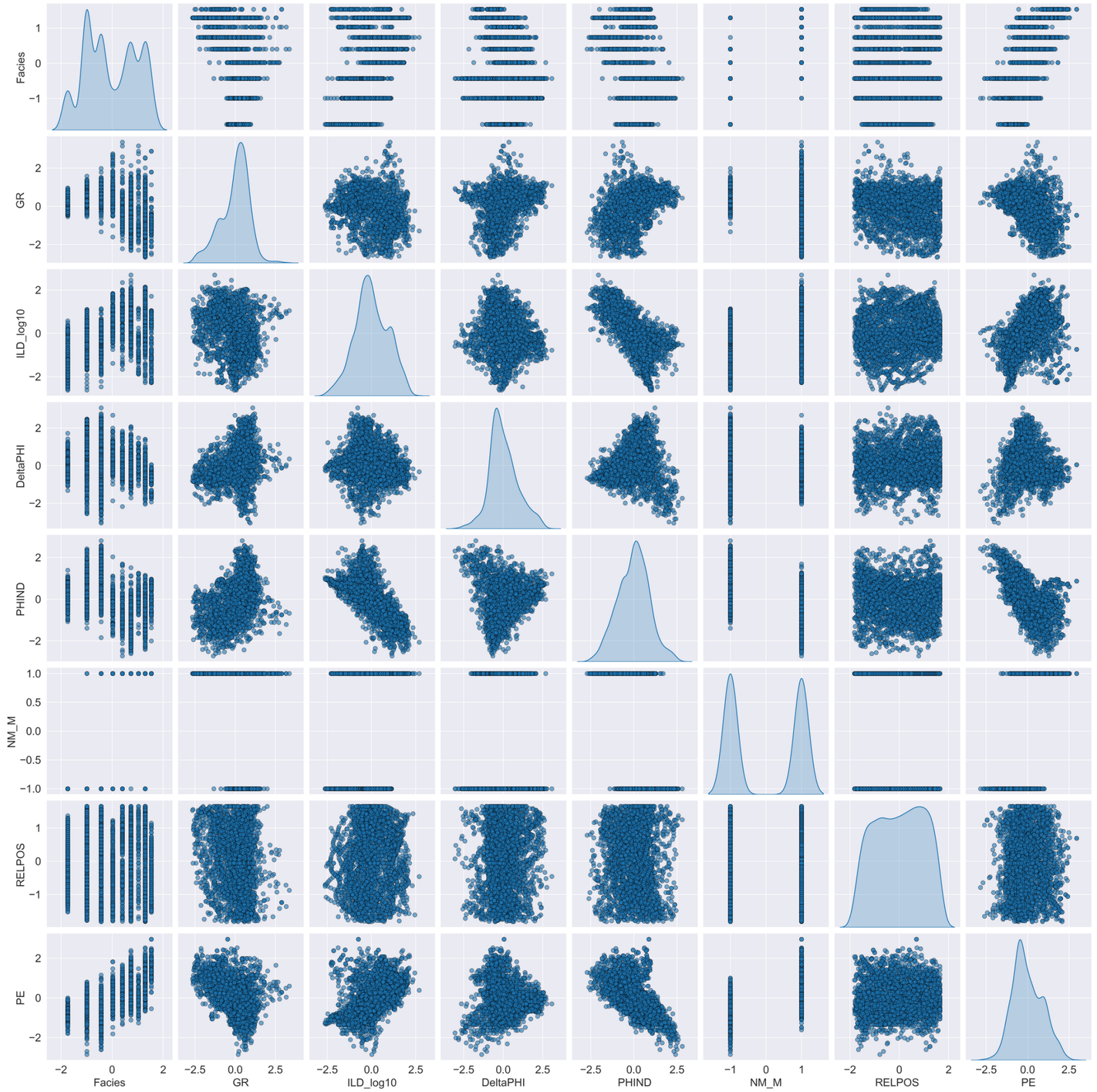
# Outlier-Removal



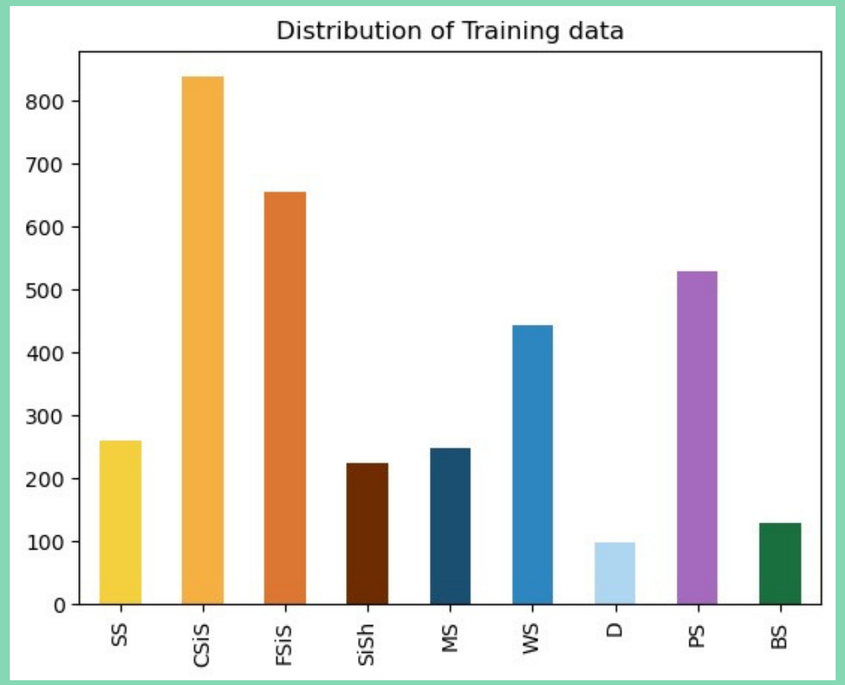
# Training

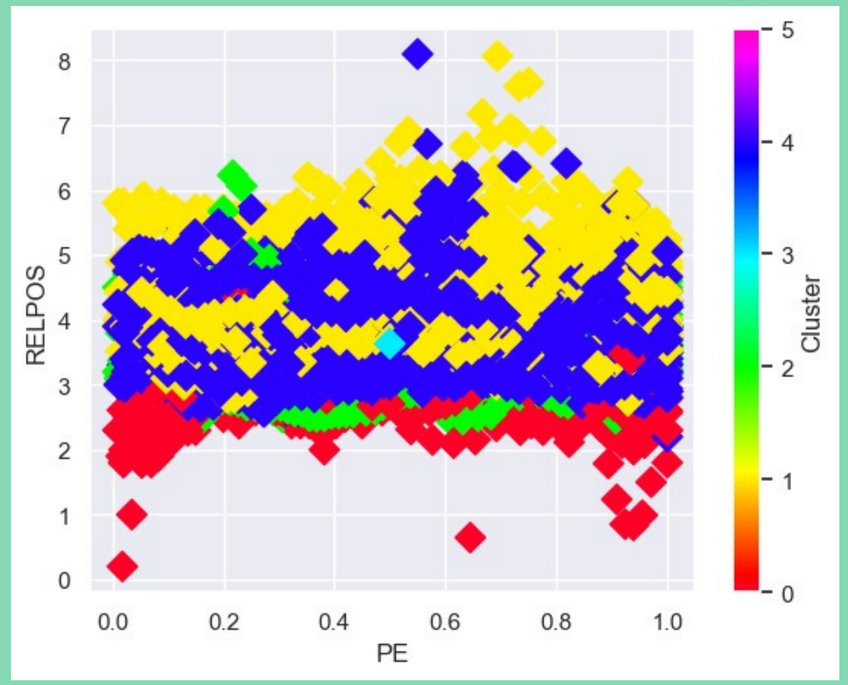
# Normalisation

# Outlier-Removal

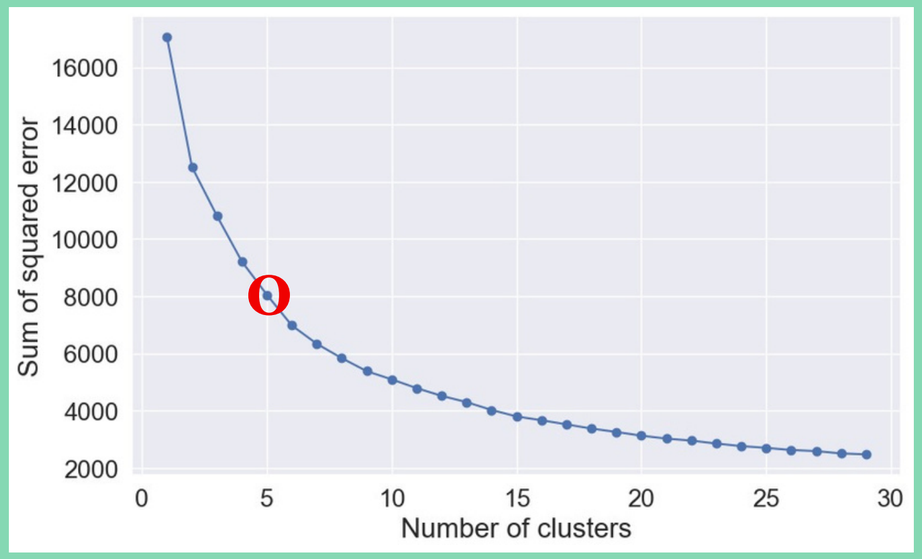


# Unsupervised Learning | K-Means Clustering



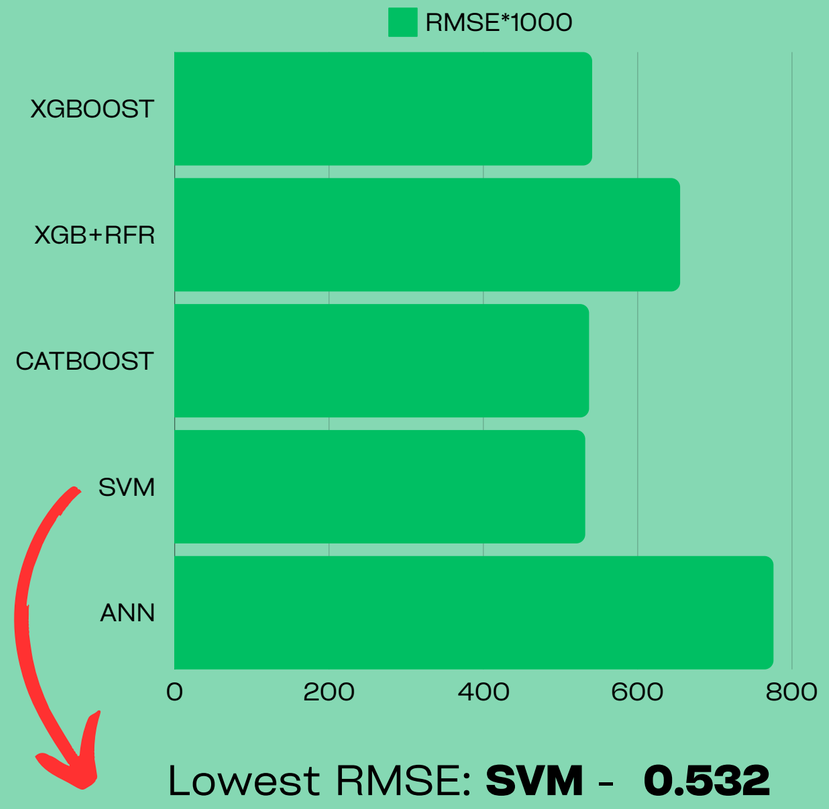




**Clustering done using 5 clusters**




**Optimal Number of Clusters:- 5**

# Predicting Missing PE Logs: *Regression*




Lowest RMSE: **SVM** - **0.532**

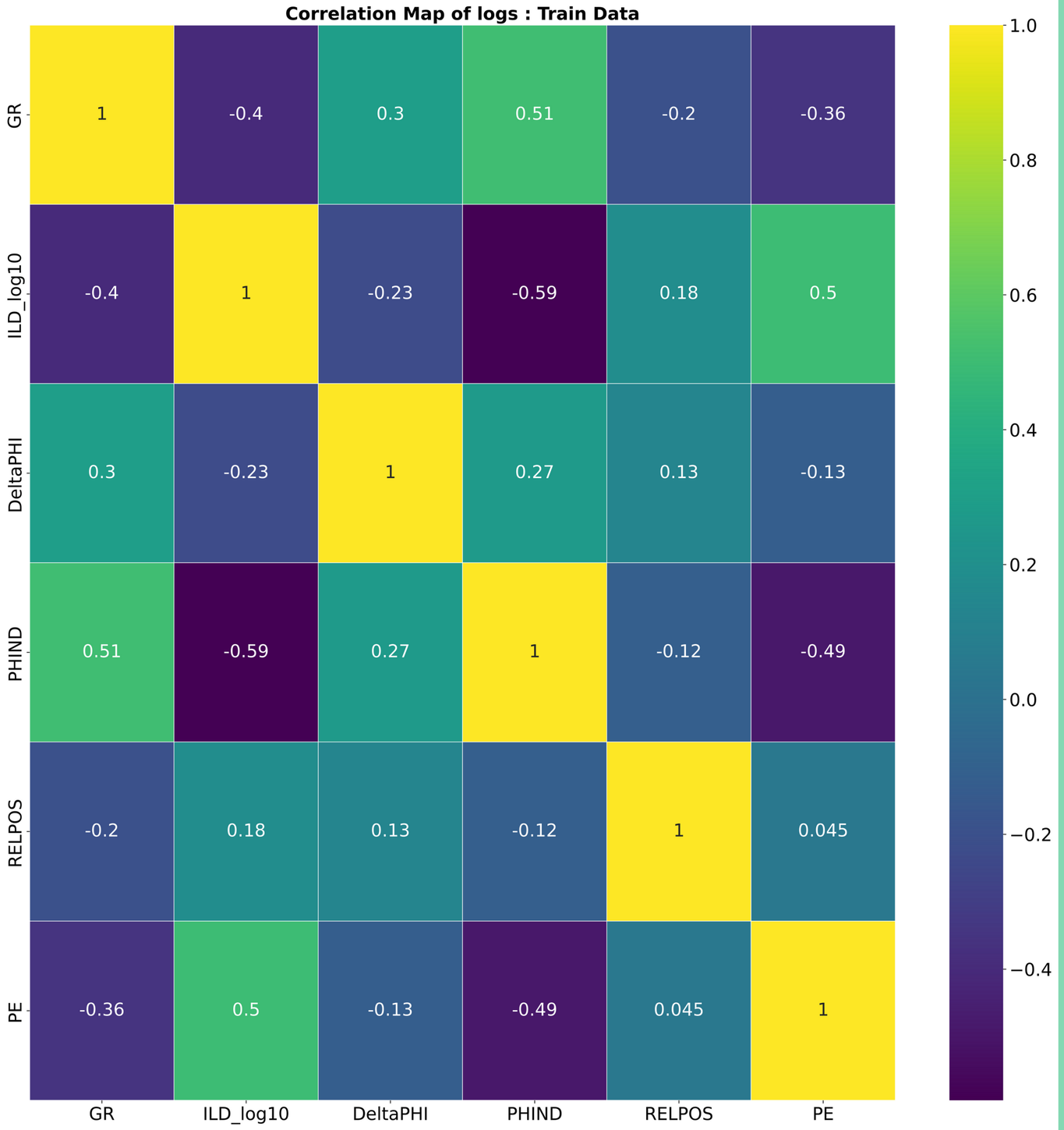
Wells not containing PE log: **ALEXANDER D, KIMZEY A**

**Scaler**: MinMax

MinModels validated from:

**CHURCHMAN BIBLE**

| **Regression Models** | XGBOOST | XGBOOSTRF | CATBOOST | SVM | ANN |
|---|---|---|---|---|---|
| **Parameters** | Default | **objective-** squared error **lr-**0.1 **max_depth-**5 **estimators-**100 | **iterations-**100 **lr-**0.1 **depth-**6 **loss_func-**RMSE | **kernel-**RBF (Radial Basis Function) | **activation-**ReLU **hidden_layers-**2 100 neurons/layer **optimiser-**adam |

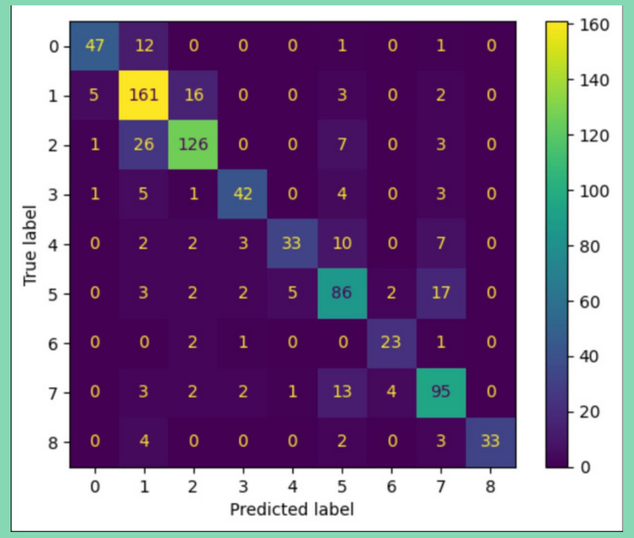



# Predicting Missing PE Logs: *Regression*

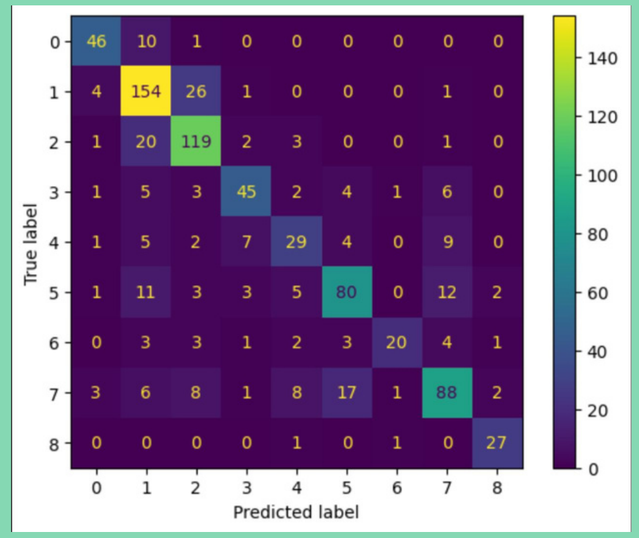
Correlation Heatmap of all the features used in the training dataset

# **Machine Learning:** *Comparative analyses of models*

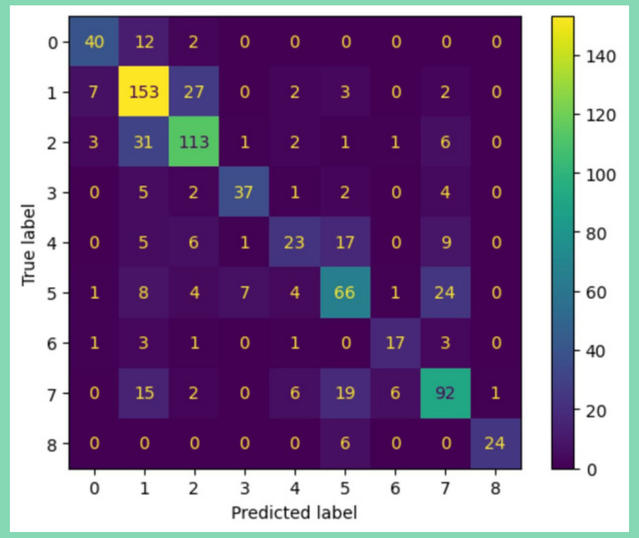
# Confusion Matrices



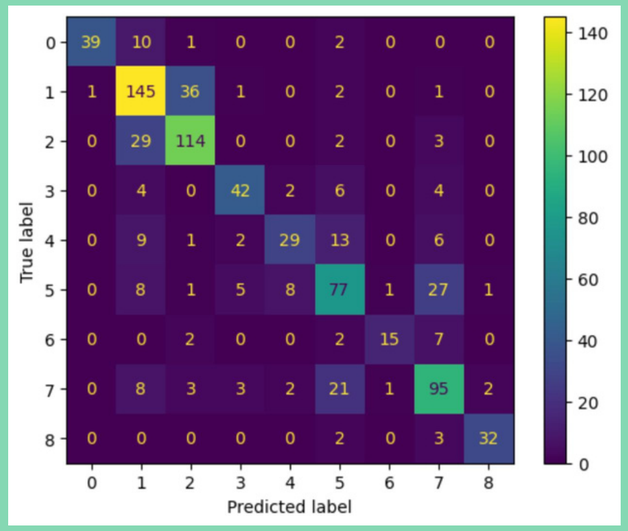

Random Forest



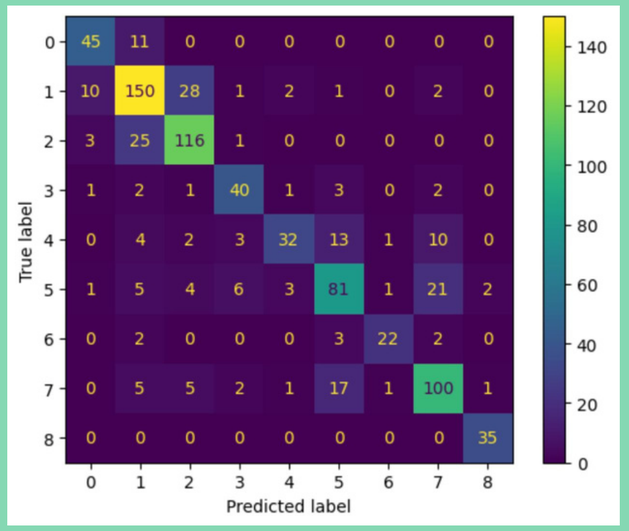

KNN




XGBoost




XGBoost-RF




CatBoost

# *Performance Metrics of all **Machine Learning** Models used:*

| | Models | Accuracy | F1 Score | Precision | Recall |
|---|---|---|---|---|---|
| 0 | XGBoost | 0.680723 | 0.687110 | 0.713182 | 0.669708 |
| 1 | XGBoost RF | 0.708434 | 0.724155 | 0.770074 | 0.694930 |
| 2 | CatBoost | 0.748193 | 0.764346 | 0.780170 | 0.759833 |
| 3 | KNN | 0.732530 | 0.727359 | 0.751021 | 0.715834 |
| 4 | Random Forest | 0.778313 | 0.784679 | 0.814630 | 0.766409 |

Metrics obtained from testing on **Churchman Bible** and **Kimzey A**

# **Deep Learning:** *Well is a sentence?*

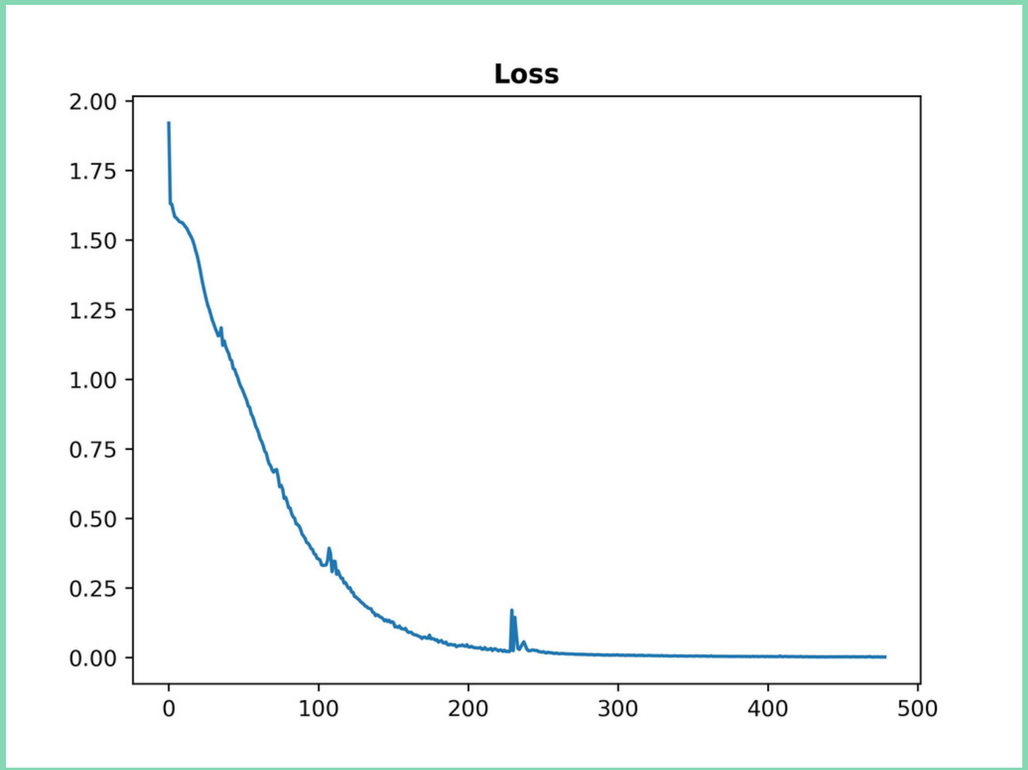
Canva

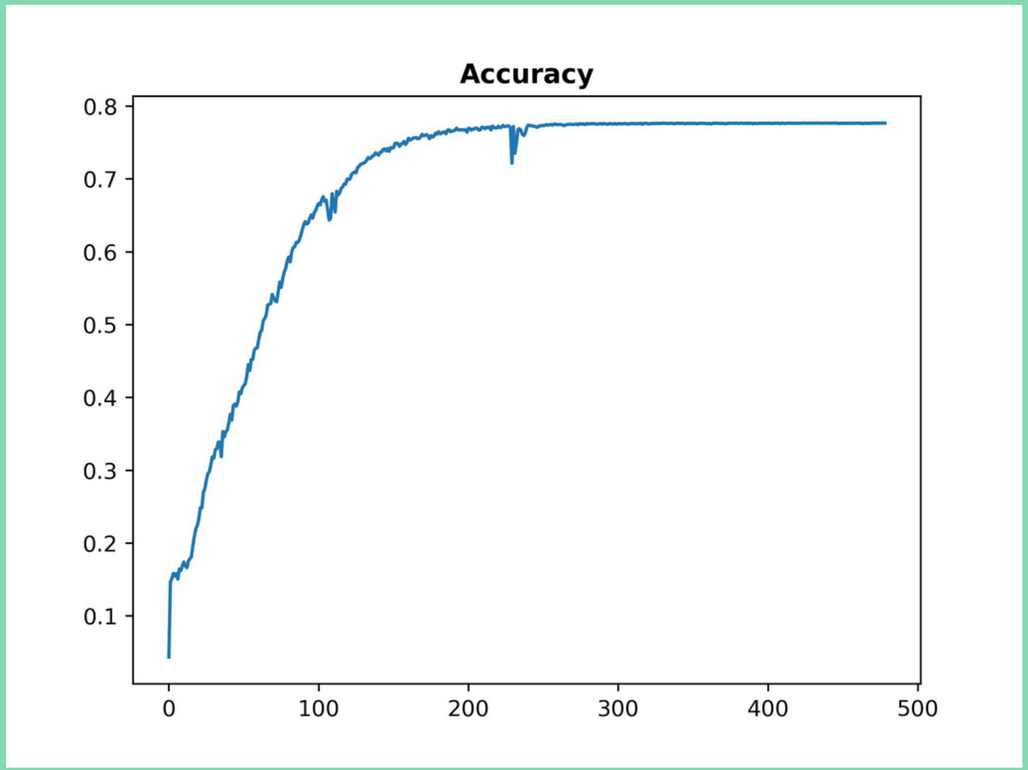
# Forbidden (403)

Sorry, you cannot access this page

**Please visit our Help Center** for more information. Error code: [**876ba3f198c739af**]

# *Training Curves*




## Evaluation metrics::

| **Headers** | **Start** | **End** |
|---|---|---|
| **Accuracy** | 4.36% | **92.8%** |
| **Loss** | 1.92 | **0.000102** |

# Thank you!