### 对数据进行筛选

今天在做实验也发现很奇怪的现象，由于是做回归预测，所以数据的分布十分重要，我们此时将很奇怪的值去掉，然后看看实验结果。因为有时候有些奇异值真的很令人讨厌。

子文件夹metadata中的是老师发给我的未修改之前的数据，Model主文件夹中是修改后的数据.

# A\_train

## C

## MN

## SI

## P

## S

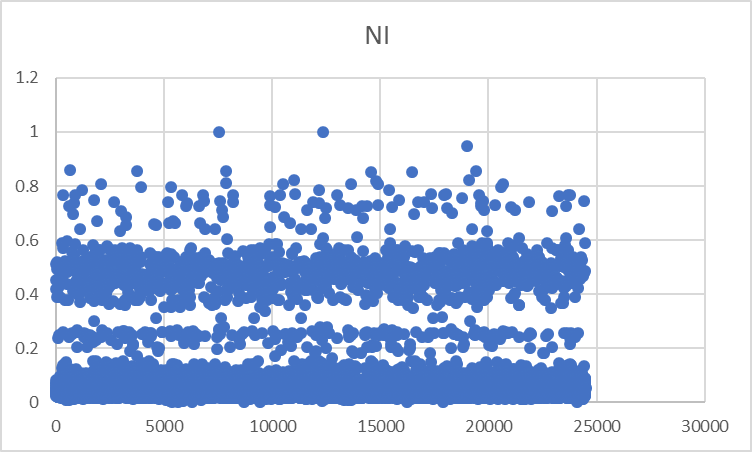
## CU

## AL

此时我们发现AL的估计有很大的误差，此时我们需要消除这些坏点。（其中AL>0.3的点都删除）

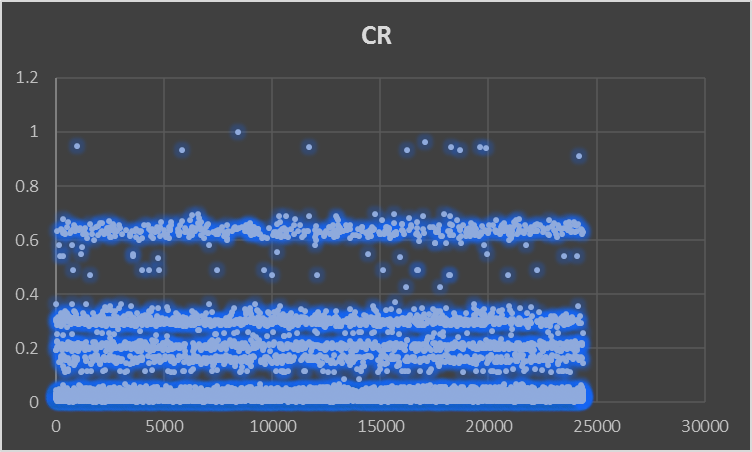
## ALS

## NI



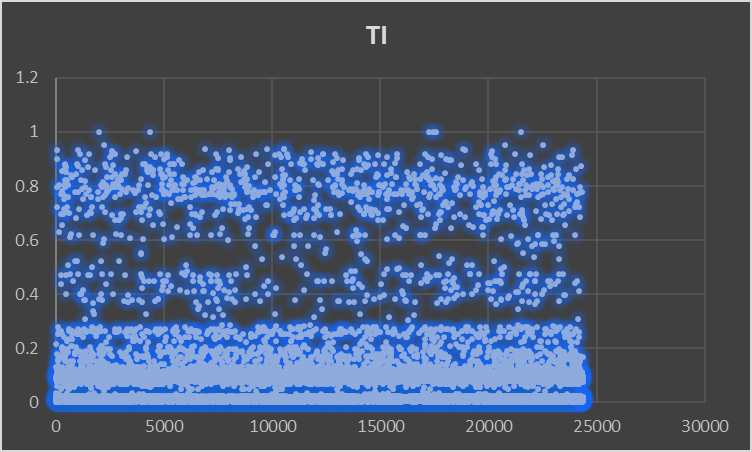
3W条数据中0.6以上分布的只有150条数据，我们将其删除

## CR

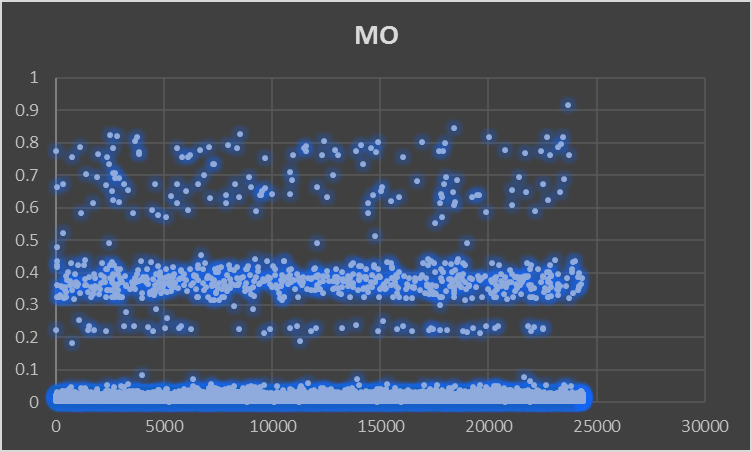


**我们删除其所有大于0.8的数据**

## TI

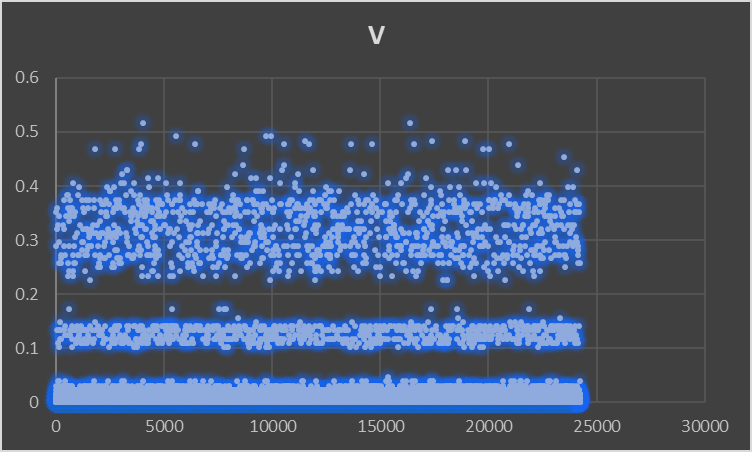


## MO



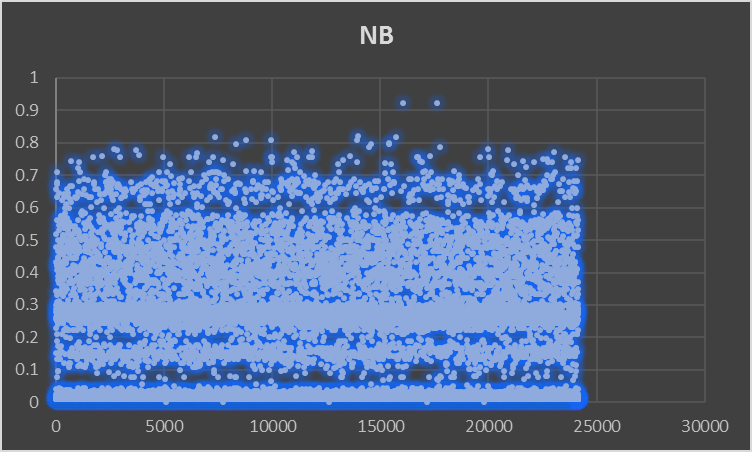
大于0.5的值有大约150个，所以我们将其删除

## V



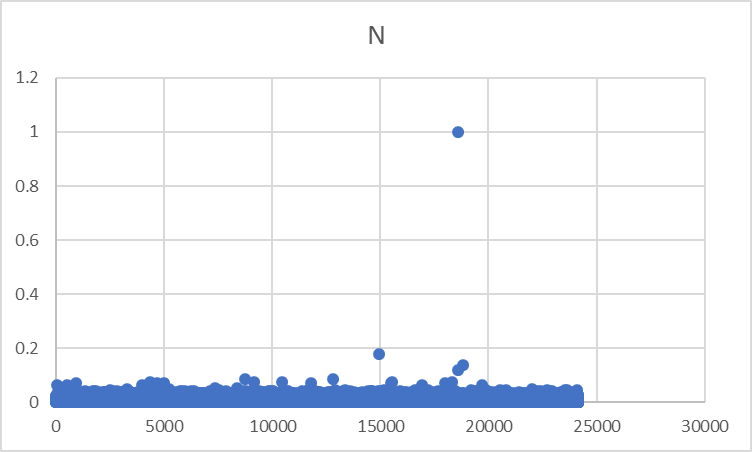
将大于0.4的值删除68个

## NB



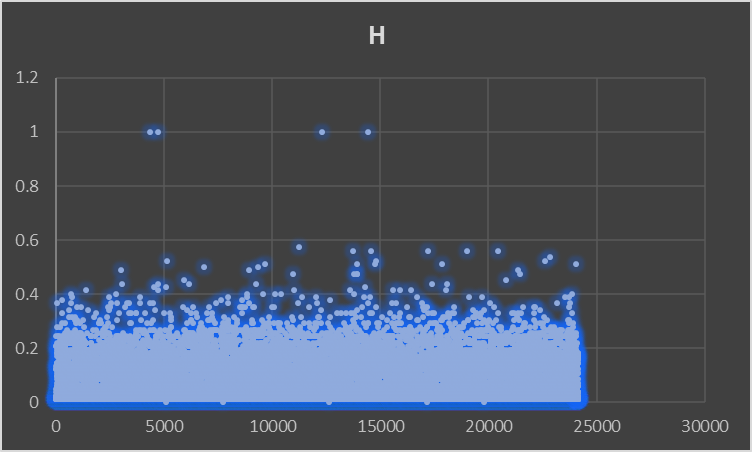
**删除大于0.9的值**

## N



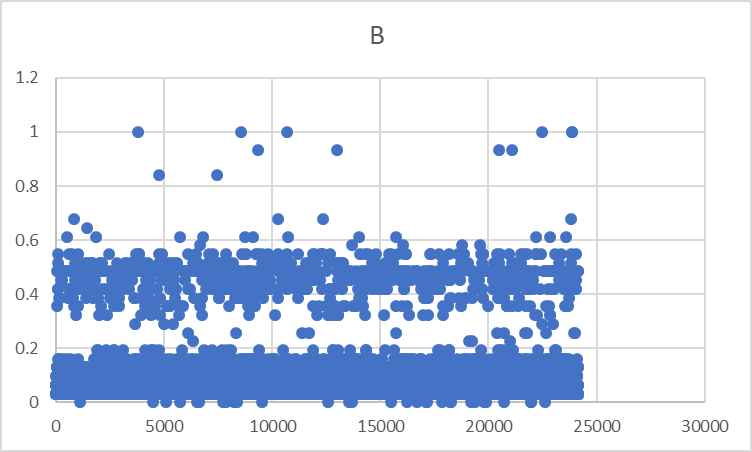
**过分了啊，删除所有大于0.1的值**

## H



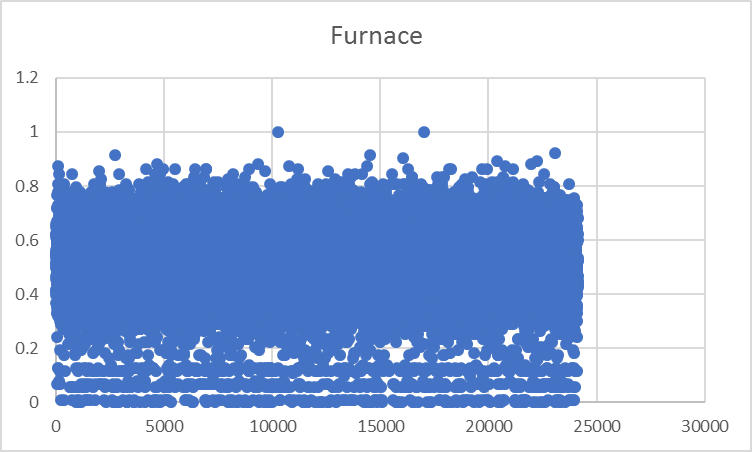
**删除所有大于0.5的值**

## B

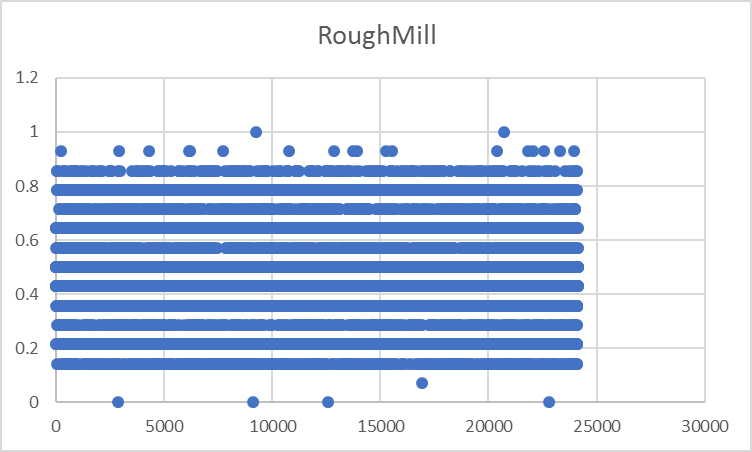


**删除所有>0.8的值**

## Furance

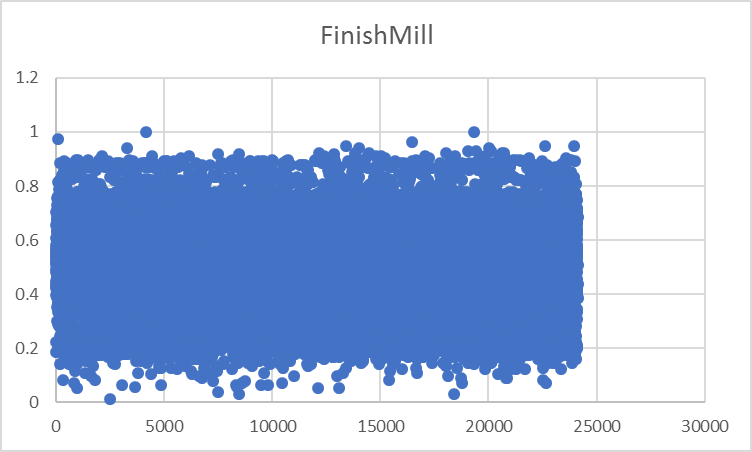


## RoughMILL

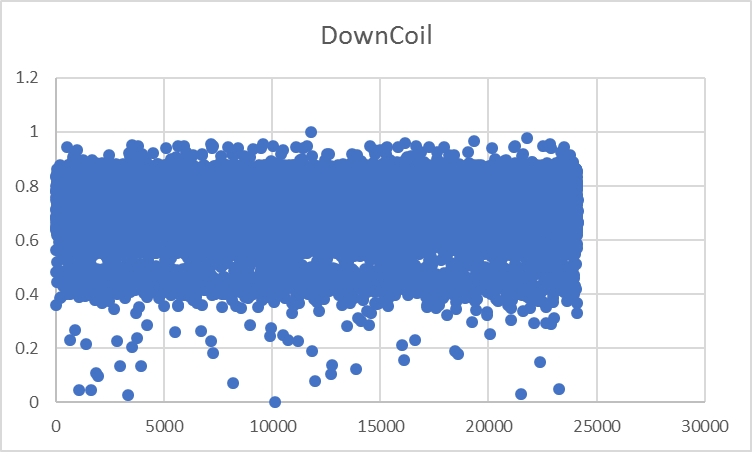


**删除所有0值**

## Finish Mill

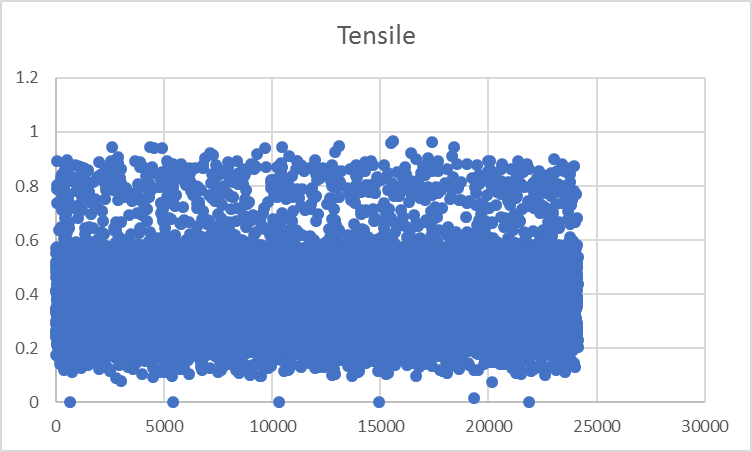


## DownCoil



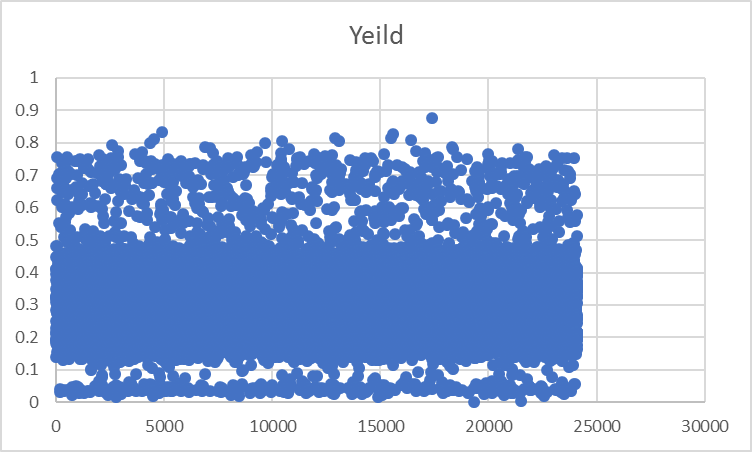
**去掉0.1以下所有值（9个点）**

## Tensile

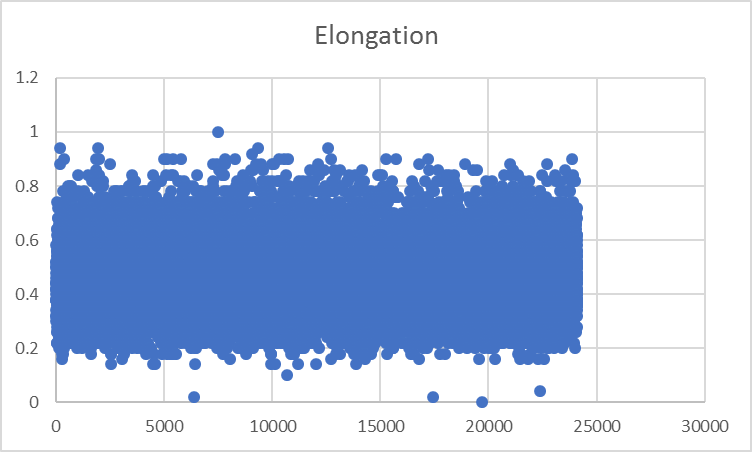


**去掉值为0所有点**

## Yeild



## Elongation



**去掉所有0值**