

Public Data Mechanical Properties Prediction Model

Mohamed Elsayed

February 1st, 2022

```
#####
# NOTE: Data in this report has been altered to protect information privacy
#####
```

I. Introduction

This is a project to build a machine learning algorithm using Linear Regression to predict mechanical properties based on 3 main elements (A, B & J). The goal is to be able to decrease B & J as much as possible without compromising product properties like One, Two & Three. Current production utilizes the following compositions for the target elements during production.

1. A : 0.22
2. B : 1.10
3. J : 0.04

Target Mechanical Properties:

1. One: MIN 405
2. Two: MAX 545
3. Three: MIN 12

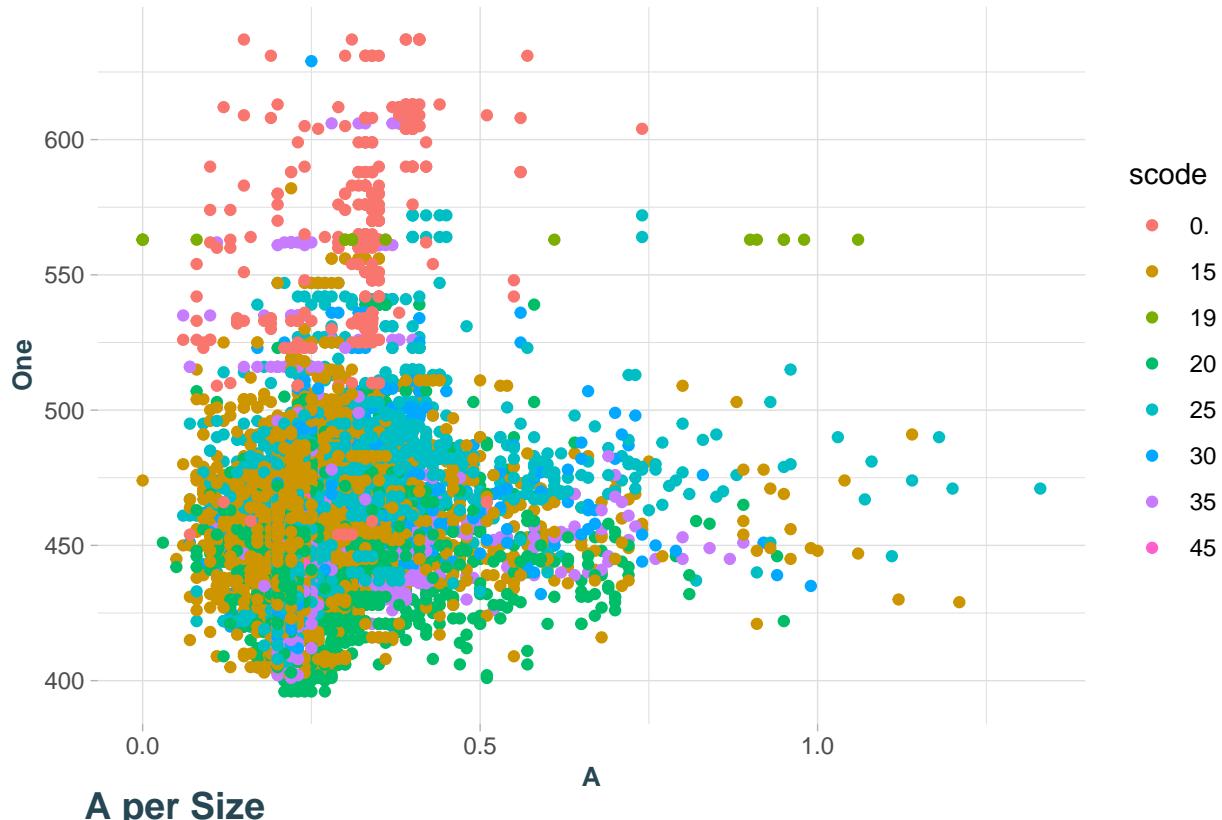
Raw data has been collected for product test results starting February 2003 until December 2021. Test data includes various grades for multiple product sizes. After data exploration, some of these grades will be filtered out to keep only the grade & sizes relevant to this project.

II. Exploring RAW data to identify outliers & irrelevant data

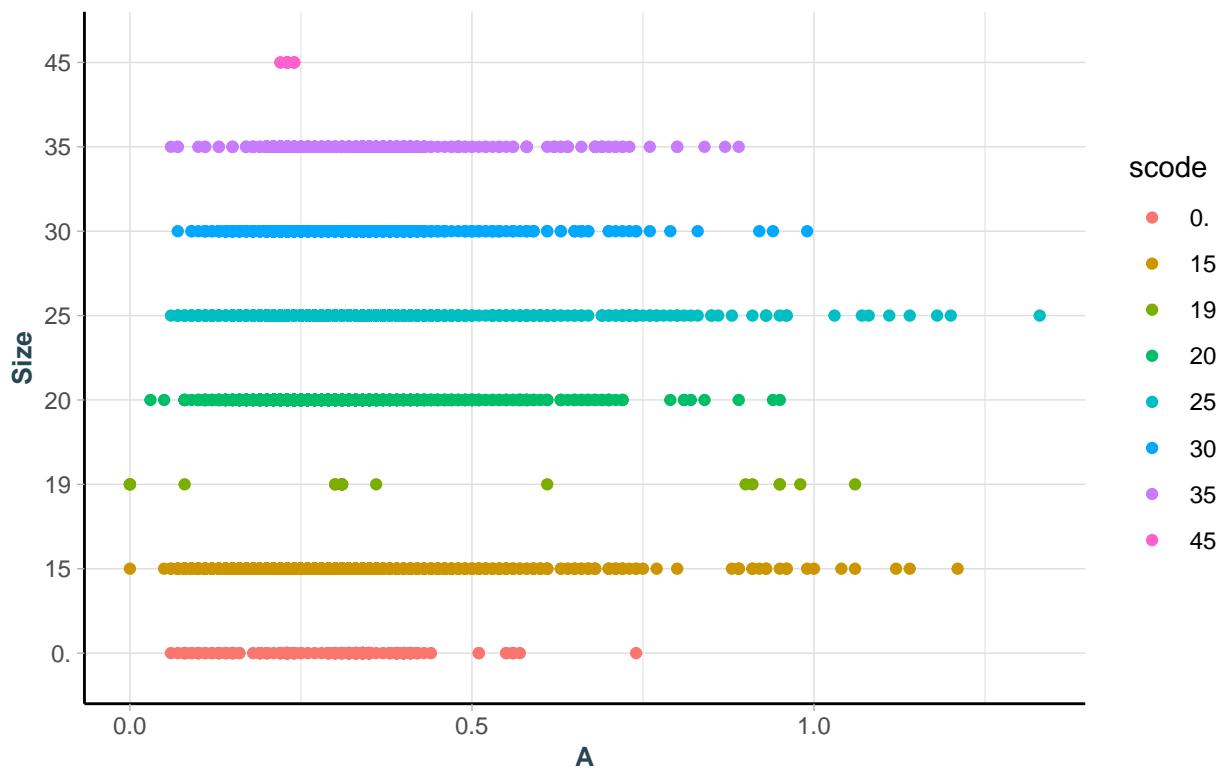
Below are plots showing that there is no clear distinction between product sizes when it comes to chemistry composition or mechanical properties. This shows that there is no need to build a separate prediction model for every product size. Also, based on the size distribution expressed in our data below there appears to be an advantage of using this method as there are some sizes with MUCH less observations to train the model than others.

scode	n
15	4563
20	2382
25	2209
35	1109
30	919
45	11

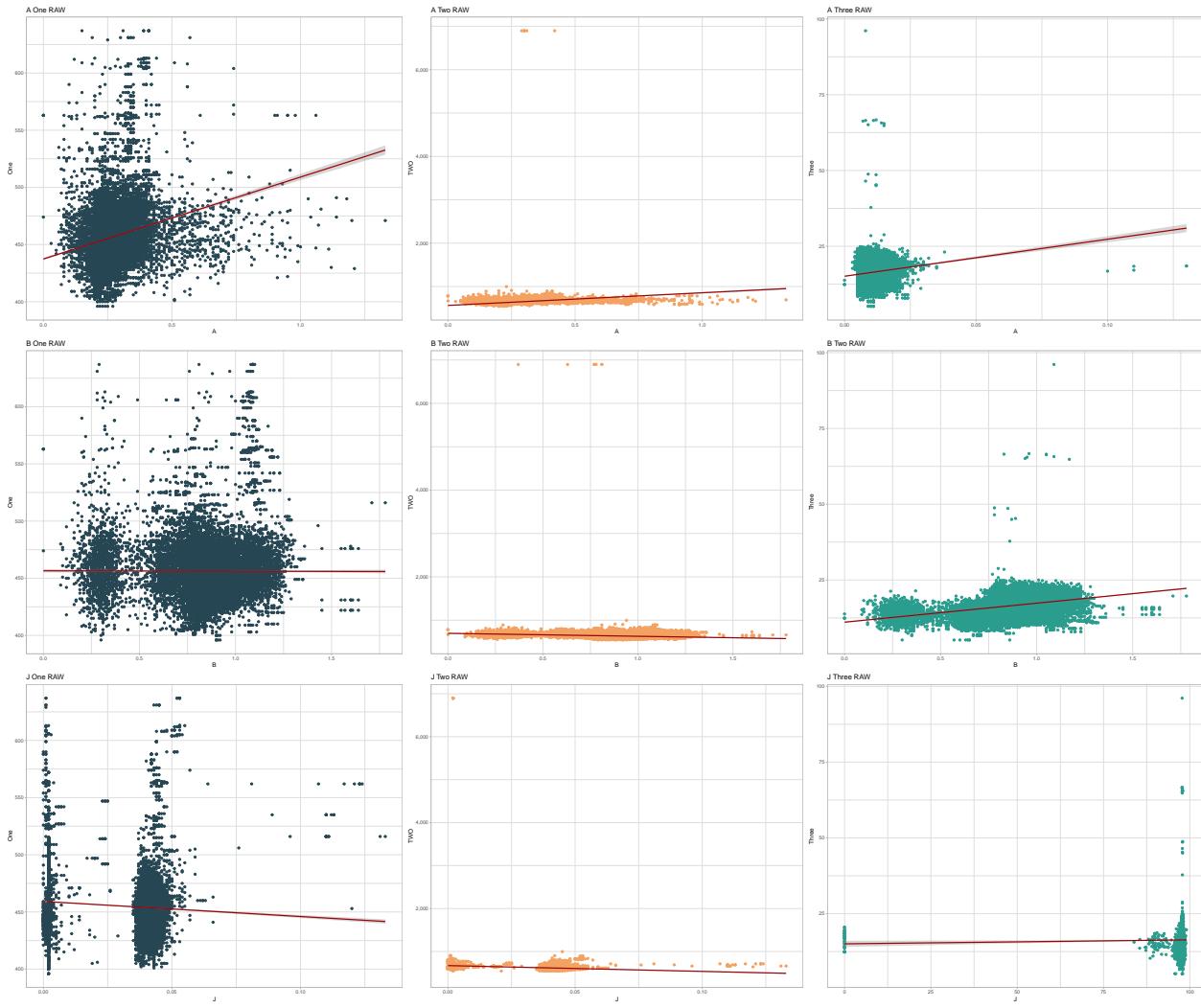
One Values per A % for each size



A per Size



- RAW Data plots below show several outliers & unrealistic data. **EX:** A above 0.5, Mechanical properties like TWO > 6,000 & Three > 50%



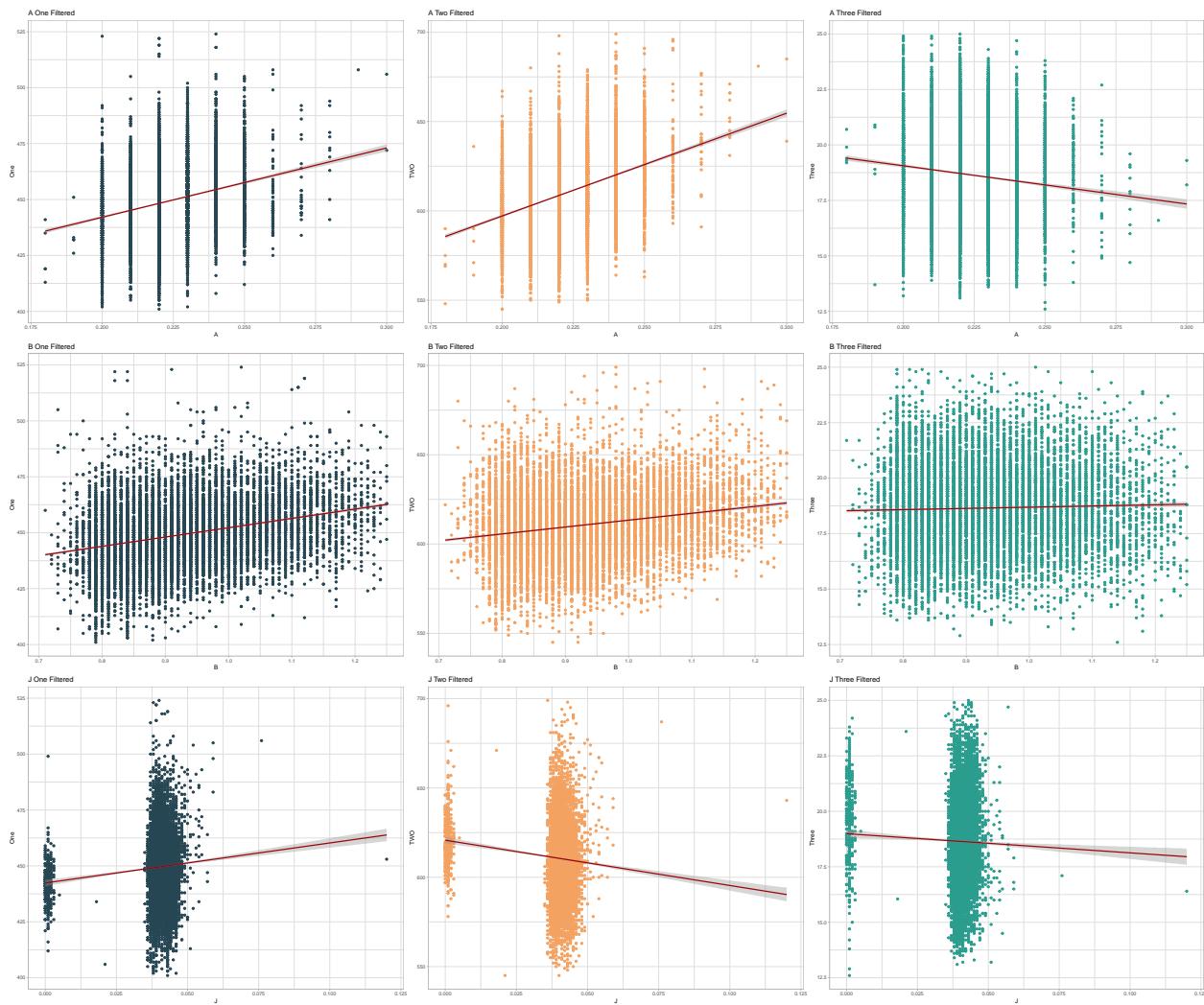
III. Data Preparation

Based on the data exploration done above, the following filters will be applied to the data to exclude irrelevant chemistry & mechanical properties outside of the desired range for this project.

Filters Applied are as follows:

1. Grade: 400B
2. A =< 0.3
3. B =< 1.25
4. J: None Needed
5. One: < 525
6. Two: =< 700
7. Three: =<25

- FILTERED Data Plots Below are the plots representing the data after being filtered. This is the data that will be used to feed the machine learning algorithm for training.



Correlation Plots

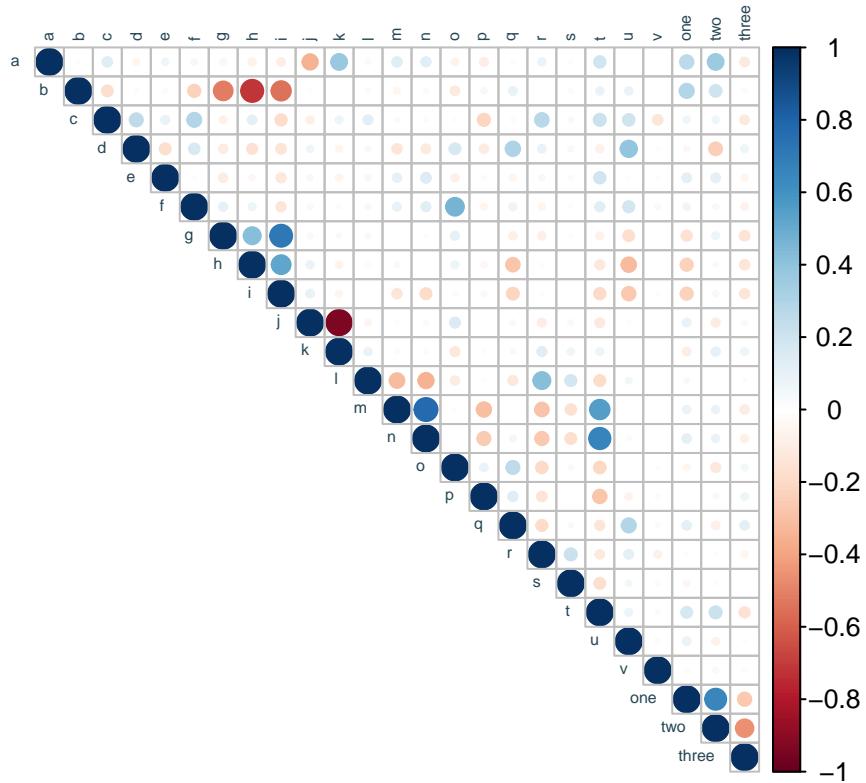
Below are two correlation plots showing how each of the elements of interest correlate to each of the mechanical properties. The darker the dot, the higher the correlation between both variables. **Blue** dot means +Ve correlation, **Orange** dot means -Ve correlation.

First Plot: Correlation of mechanical properties to ALL elements

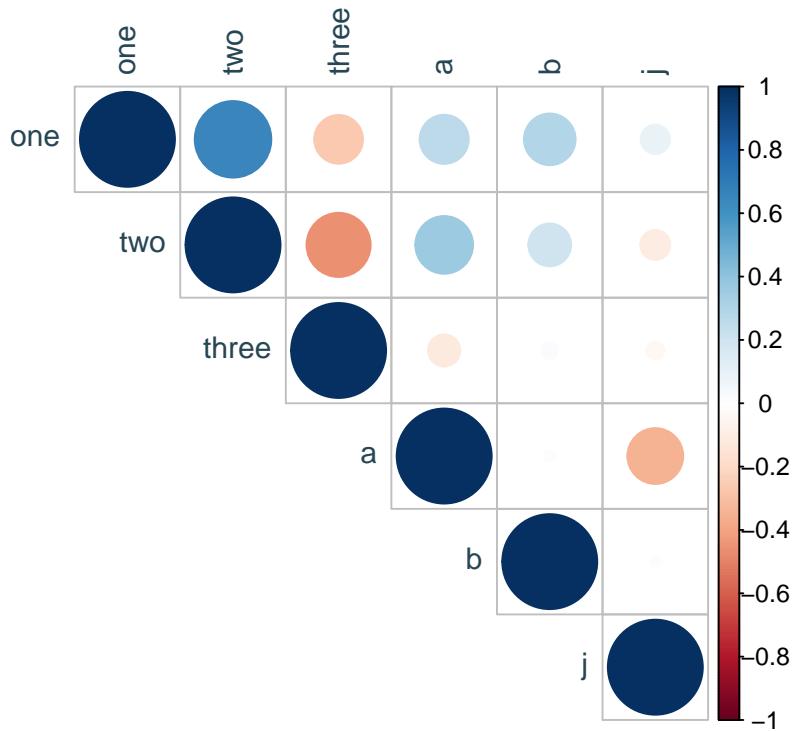
Second Plot: Correlation of mechanical properties to ONLY the 3 target elements (A, B, J)

It is clear from the plots below that no single element has a significant correlation value with any of the mechanical properties. The second plot shows a slight correlation between (A & Two), (B & One) & a very weak inverse correlation between(J & Two)

Correlation of mechanical properties to ALL minerals



Correlation of mechanical properties to ONLY the 3 target minerals (A, B, J)



IV. Building Algorithm

The dataset will be split in 2 portions. One portion will contain 80% of the data & this will be used for training the algorithm. The second portion will contain the remaining 20% of the data & this will be used for testing (verification) of the algorithm prediction results.

Linear Regression model data

- Linear Model predicting **ONE** based on chemistries. Model accuracy around 89% when limiting the difference between predicted & actual results to only 5%. *RMSE for ONE predicted by this model is 12.53.*
- Linear Model predicting **TWO** based on chemistries. Model accuracy around 95% when limiting the difference between predicted & actual results to only 5%. *RMSE for TWO predicted by this model is 15.22*
- Linear Model predicting **THREE** based on chemistries. Model accuracy around 67% when limiting the the difference between predicted & actual results to around 10%. *RMSE for THREE predicted by this model is 1.7*

V. Results

Below is a table containing multiple outputs of the prediction model.

The first three results are verification results for heats that were rolled through the BM after the data used for creating the prediction model was collected. This is a good test of the model's prediction capability. For ALL 3 heats verified, the mechanical properties predicted by the model based on C, Mn & V values were **within 95% Accuracy of the Actual results** achieved from testing.

Fourth & Fifth results are predictions based on an average of our current recipes & the recipe of a competitor

The Last result is the prediction based on the Proposed Chemistry for production

A	B	J	One	Two	Three	Actual_One	Actual_Two	Actual_Three	Remarks
0.23	1.04	0.040	449	591	17.22	454	607	17.05	Size 15 Results Verification
0.21	0.90	0.043	438	571	17.76	442	591	19.35	Size 35 Results Verification
0.22	0.74	0.051	435	559	18.35	417	556	18.05	Size 55 Results Verification
0.22	1.10	0.040	458	642	16.73	454	613	18.47	Current Chemistry
0.27	1.24	0.018	473	655	12.44	434	671	16.05	Competitor Chemistry
0.20	0.75	0.015	427	587	17.19	NA	NA	NA	Proposed Chemistry

VI. Recommendation

Based on the results above, Reducing A to 0.20, B to 0.75 & J to 0.015 would still result in a product that meets the minimum mechanical requirements for tensile testing.

This is a significant decrease in Two by almost 32% & Three by almost 64%.

NOTE: Safety margins, mineral cut-off values & other mechanical properties need to be metallurgically verified before production.