**Preprocessing** :

* We discarded “id” , “name” , “full\_name” columns because they don’t have any effect on value and “birth\_date” column because we already have the age column which calculates the birth of date .
* Delete columns which contains more than 7000 missing value in it’s row .
* Delete rows which contain any null value .
* ‘position’ column contains 4 distinct values so we split them with ‘ , ’ and make a new column for each of them then replaced none values to zeros then replace any nominal values to ones and finally sum them and get them back into ‘position’ column and drop the new columns .
* ‘work\_rate’ column contains 2 distinct values so we split them with ‘ / ‘ and make a new column for each of them ‘Attacking Work Rate’ and ‘Defensive Work Rate’ then delete the main column and replace ‘Low 🡪 1’ , ‘Medium 🡪 2’ , ‘High 🡪 3’ .
* ‘body\_type’ column replace nominal values into numeric values .
* We used LabelEncoder on 4 nominal columns alphabetically to make it numeric with values .
* For ‘+2’ handling in LS column 🡪 RB column we used str function to get the only first 2 digits of each value .
* In ‘club\_join\_date’ and ‘contract\_end\_year’ to get the year we take the last 2 digits of each value because they have different date formats .
* We used correlation to filter the data and get the top features that affect more than 50% on the value so it became 8 only from 91 column .
* We used MinMaxScaler to normalize data and make the top feature columns within range 0 🡪 1 to avoid varieties between values .
* Splitting data into 80 % training and 20 % for testing .

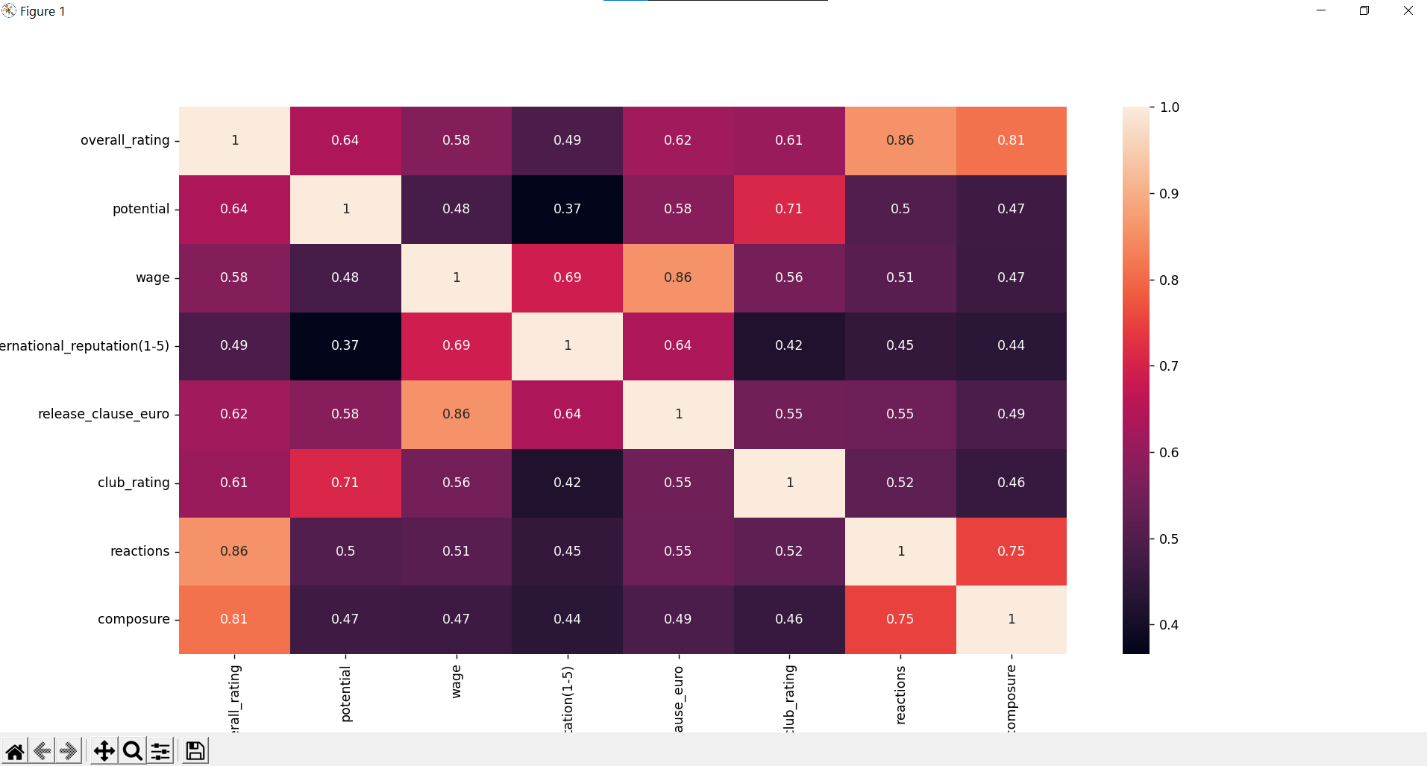
**Models** :

We used 2 models “Polynomial Regression” and “MultiVariable Regression” and the one which gives less mean square error was the “Polynomial Regression” with degree of 3 it takes “0.05787372589111328” second in training time with “197971539579.55865” in mean square error and the other model takes “0. 0019936561584472656” second in training time and “397445289609.2285” in mean square error .

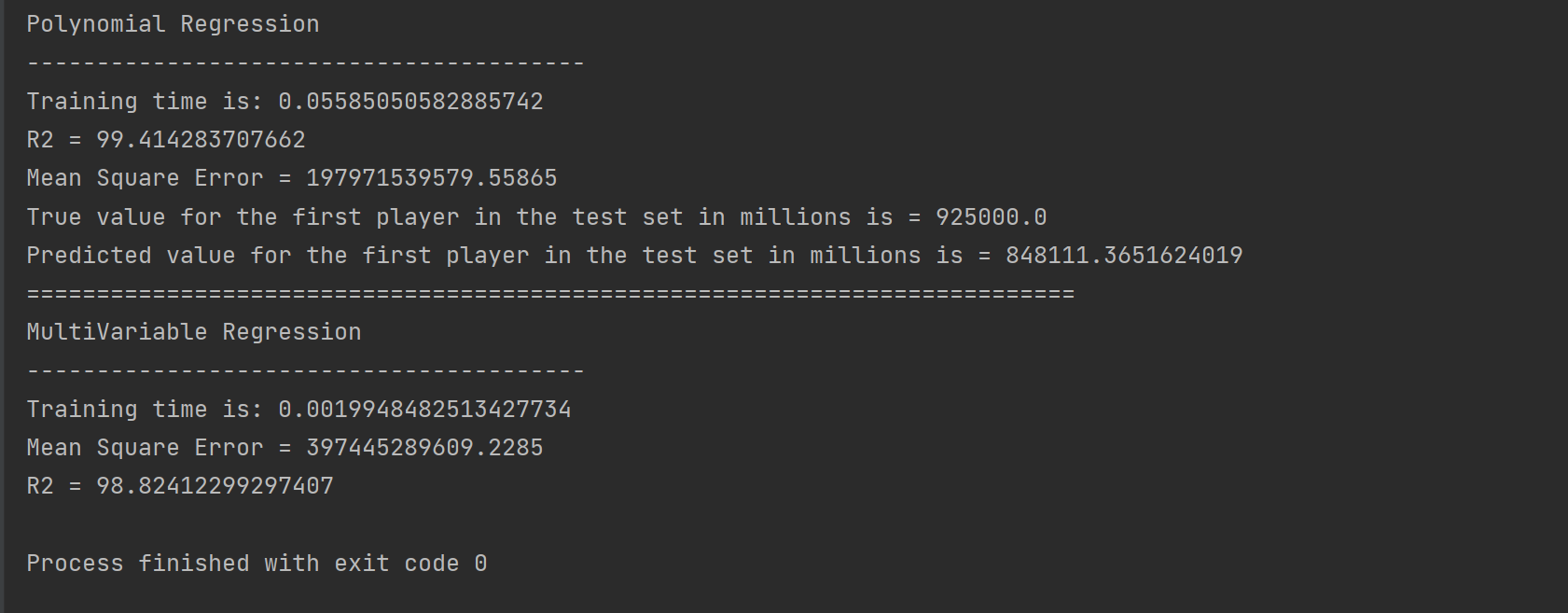
**Conclusion** :

Type your Bolbol here …

**ScreenShots** :

**Correlation Plotting**

**Results**

****