***#Load all necessary libraries that could be use along the journey of cleaning the data***

***I load my datasets and name it "fifa\_dirty" to check what columns needed to be cleaned up***

**The dataset, from brief information given from above shows we have 77 columns to work with. this is large and my notebook cannot give me a clearer picture of every column at a glance.**

**Because i want to get myself familiar with every column's data, i had divide my data along the columns to four datasets, then check for missing values and know those with sprecial characters that needs to be dropped**

##### after checking all my columns, the next is to extract the columns that needs cleaning into a single dataset and start working on them from column to column.

**let's import regular expression and write a function to drop special characters in Club's column**

**to clean the Contract column, we call our column as a string and split along ~. so let's do it**

##### to clean the Height and Weight, i should check for unique to see if it is only cm and kg respectively, are the only characters i want to remove or other characters are there.

**--- to clean this i can do two things: ----1. drop index from the back or -----2. add the characters to my regex and drop. but i will do 1**

#### **now, let's write a function to change feets and inches to centimetres**

#### **we will put the Weight column in the same process as that of Height in our cleaning**

##### The unique() above shows us we have pounds(lbs) to convert to kg. we will write function just as we did with Height column to do this

##### let's write a function to change our "Joined" to datetime

**Let's work on the value column**

**Let's write a python function to write the figures in full but before that,i will add the euro character to my regex function above, so that i can use it to drop it. we will put the sign back in the column's name**

##### We should be able to apply the full\_figure() function to our next two columns i.e the Wage and Release Clause. so let's do it....

##### The next is to drop the stars characters in "W/F", "SM", and "IR"

***# we will use the lambda function in cleaning all of W/F, SM, IR . so let's go***

##### our dataset is almost attaining the cleaned status . the only column left is the Hits column. so let's check it out and work on it...

##### from the hitmap above, Hits column has few unclean data. the null value is small. that shouldn't affect the outcome of our analysis if we want to go ahead with analysis. therefore, we will drop them..but before then let's check the amount of NaN, then we decide either to drop them or fill with something else...

##### it looks as if we have more null value in our hand than expected, let's see if we can fill with mean or better still, ZERO. but before that, let check the total amount of null value

##### instead of dropping it and missed out of the information for the other columns attached to the null value, the best thing to do is fill the missing value with zero

##### we need to convert the K to 1000 as display in the unique(). the best is to call our full\_figure() function

#### **we have, to the best of our knowledge, clean up our dirty columns. the next thing to do is to use our ID column as index and return all the data back to the main dataset. we have to drop the URLS column in the dataset because it has zero analysis value. so, let's go....**

#### **So we come to the end of our data cleaning processses. the next thing to do is to:**

* **merge our cleaned data to the dataset we scooped them from**
* **reset our index to ID column**
* **and rearrange our columns position if necessary**

#### **And we come to the end of our fifa\_2 data cleaning challenge. thank you for walking this journey with me. i will leave you with the heatmap showing if we still have any missing value in our overall data**

##### This is a clean and ready-to-work-with data. we can progress with asking some few questions with the data and doing analysis from there. Bye for now......AND SAVING MY DATASET BACK TO THE DIRECTORY