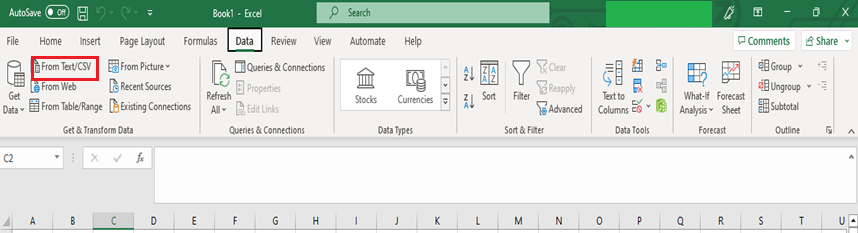
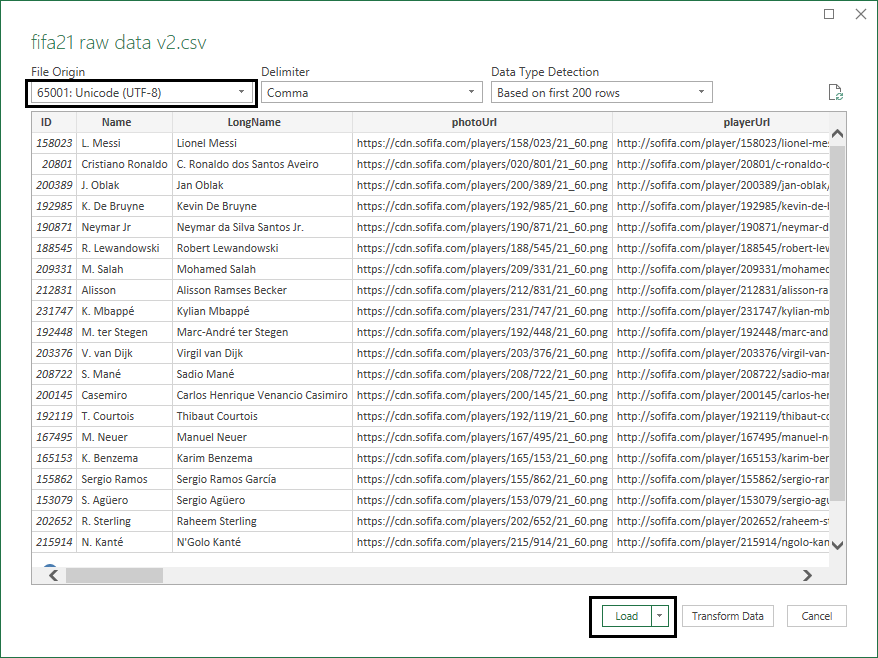
**FIFA 2021 DATA CLEANING WALKTHROUGH.**

**WRITER: MICHAEL OLAGOKE**

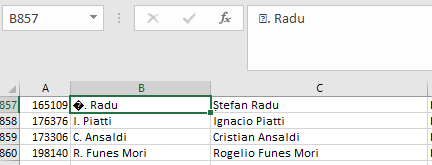
1. Load the fifa21 v2 csv dataset into excel through the csv connector in the data tab.



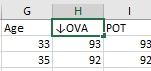
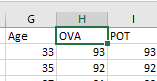
1. Select the file origin as “**Unicode(UTF-8)**” to avoid displaying none English alphabets as strange character. Click on load thereafter.



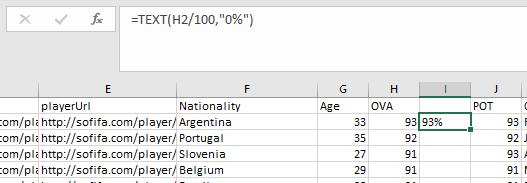
1. The data is loaded as a table into Fifa raw data v2 sheet, select all cell, copy and paste as values into cell A1 in sheet 1.
2. Scroll down to cell B857, delete the first symbol and replace as ‘**Ș’.**



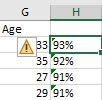
1. In cell H1, Remove the column heading from ‘**↓OVA**’ to ‘**OVA’.**

1. **OVA** column should be graded in percentage so insert a new column in column **I** o that column I becomes blank and POT is moved to column **J**. In cell I2, enter the formula ‘**=TEXT(H2/100,"0%")**’. This format the content of cell H2 into percentage. Then flash fill the cells below by double clicking on right bottom corner of cell **I2**.

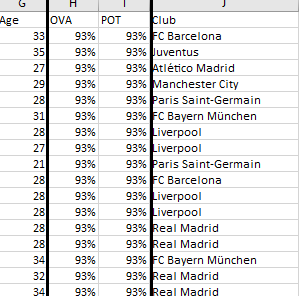


1. Copy column I and paste to column H as values. Column H2 shows an error hint icon as shown below.

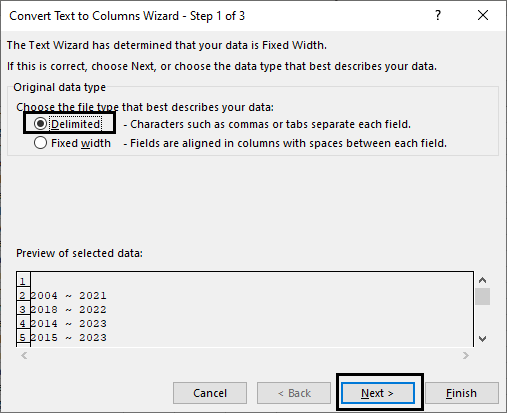
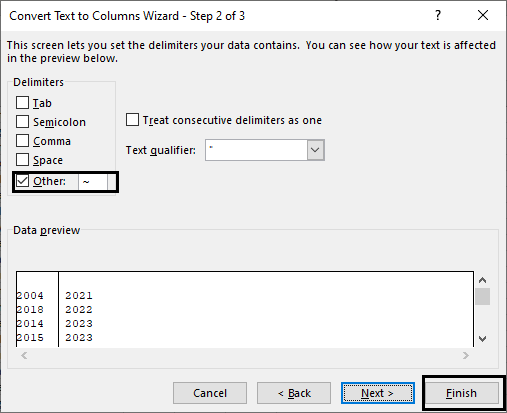


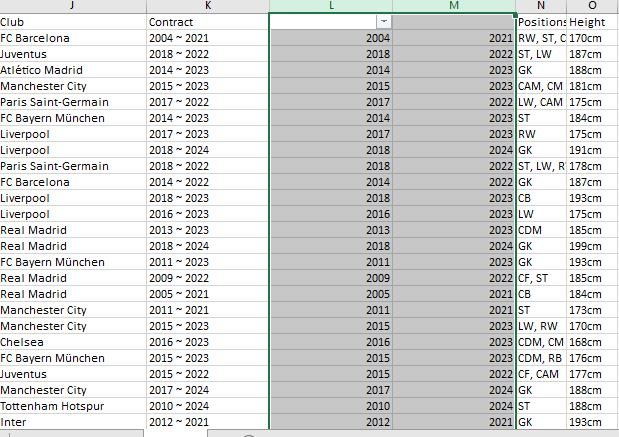
This is because Excel detects that a number value has been pasted as text, to correct this, click on the icon and select ‘**Convert to Number**’ and flash fill the cells below. Delete column I and rename H1 as ‘**OVA’.**

1. Repeat procedure in **6** and **7** to column H to clean up **POT**



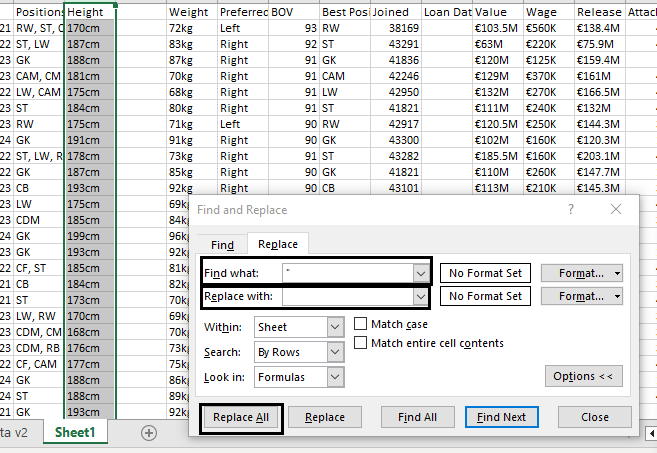
1. Insert a new column in column L, in cell L2, input the formula ‘**=IF(RIGHT(K2,4)="Loan", LEFT(RIGHT(K2,12),4) &" ~ " & RIGHT(K2,4),IF(RIGHT(K2,4) ="free", "Free ~ ",K2))**’ flash fill the cells below. Copy column L and paste as values into column L.
2. Select column L, click the Data Tab and select ‘**Text to columns**’, select ‘delimited’ in the wizard window and click ‘**Next**’. In the next window, tick only ‘other’ and input delimiter as ‘**~**’.

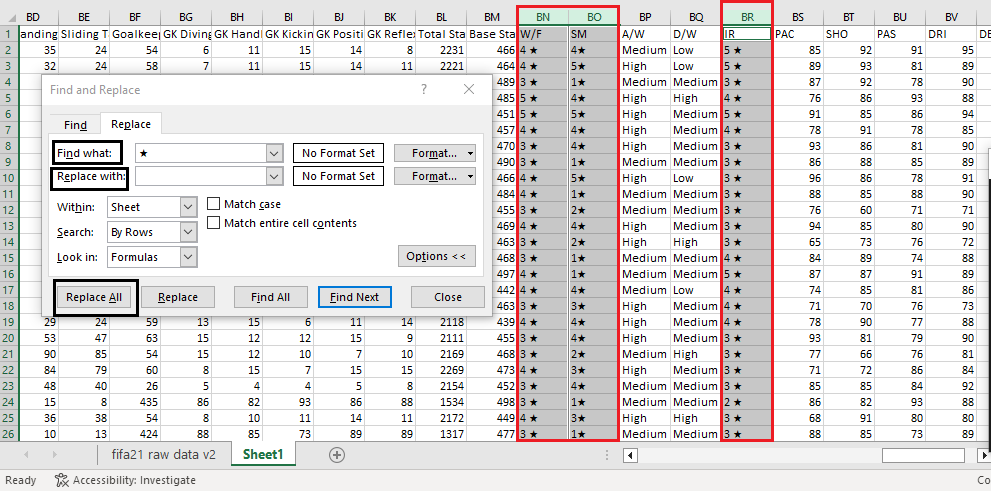


Rename cell L1 as ‘**Contract Start**’ and cell M1 as ‘**Contract End**’ then delete column K.

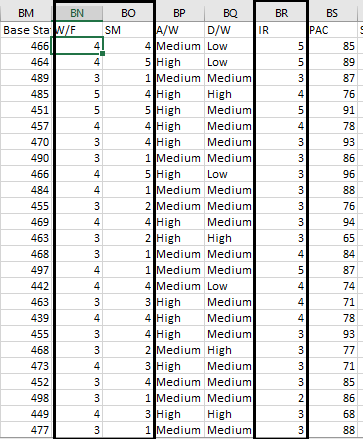
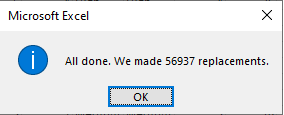
1. Select Column N and press **Ctrl** + **H**. Insert **”** into the ‘**Find what’** field, leave the ‘**Replace with**’ as blank. Click ‘**Replace all**’



1. Repeat procedure above and replace ‘ with ‘.’
2. Insert a new column in column O, in cell O1, input the formula ‘**=IF(RIGHT(N2,2)="cm",LEFT(N2,3)\*1, CONVERT(N2,"ft","cm")\*1)**’. Flash fill the formula to the cells below. Copy column O and paste into column N as values. Rename column N as ‘**Height in cm**’. Delete column O.
3. To clean the weight column, select column O and replace all ‘kg’ with blanks. Select column P and insert a new column. In cell P1, insert the formula ‘**==IF(RIGHT(O2,3)="lbs",LEFT(O2,3)\*1,ROUND((CONVERT(O2,"kg","lbm")\*1),0))**’ and flash fill to the cells below. Copy column P and paste as value to column O. Rename O1 as ‘**Weight in lbs**’. Delete column P
4. Select column T. replace all blanks with ‘**Null**’.
5. Select column V and insert a new column. Select cell V1 and insert the formula ‘**=IF(RIGHT(U2,1)="M",LEFT(U2,LEN(U2)-1)\*1000000,LEFT(U2,LEN(U2)-1)\*1000)**’ and flash fill the cells below. Copy column V and paste into column U as values. Delete column V and rename Column U header as **Value**.
6. Repeat the procedures above for **Wages** and **Release** column.
7. Select column **BN**, **BO** and **BR**. Press **Ctrl** + **H**, a dialog box appears. Input ‘★’ into the **Find what** field and leave the **Replace with** field as blank. Click on **Replace All**.

****

The window prompt below appears.



1. Select Cell CA2 and input the formula ‘**==IF(RIGHT(BZ2,1)="K",LEFT(BZ2,LEN(BZ2)-1) \* 1000,BZ2\*1)**’. Flash fill the cells below. Copy column CA and paste as value to column BZ. Delete column CA and rename cell BZ1 as ‘**Hit**’.

