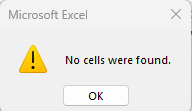
The [FIFA 21 dataset](https://www.kaggle.com/datasets/yagunnersya/fifa-21-messy-raw-dataset-for-cleaning-exploring/data), publicly available on Kaggle, consists of web-scraped raw messy data from sofifa.com. It includes detailed information on over 18,000 football players from around the world.

**Objective**The main objective for this project is to effectively clean and transform the dataset to enable proper data analysis using Excel.

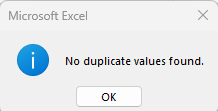
Cleaning Data and transforming it

1. Checking for Null or missing values and duplicates, The “Go To” feature is used to identify blank

cells. To access this feature, follow these steps: *Home > Find & Select > Go To Special* in the ribbon. Then, select the **Blanks** option from the **Go To Special** pop-up window and click **OK** to apply the selection.



Removing duplicates, To do this, navigate to *Data > Remove Duplicates* in the **Data Tools** group.



1. Contract, this column contains data in 3 categories one is “2004 ~ 2021”, other is “On Loan” and one more is “Free”. Using this information deriving new column “Loan Contract” using below formula:

= IFS(ISNUMBER(FIND(“On Loan”,O1)),”On Loan”,ISNUMBER(FIND(“~”,O1)),“Active”,ISNUMBER(FIND(“Free”,O1)), “Free”)

**Note: O1 is the relative cell addressing.**

Further separating contract data to Contract Start Year and Contract End Year using “Text to Column” feature of Excel.

1. Weight and Height data, converting weight into KG from lbm and height to inches using below formula

For Weight: =CONVERT(O1,"lbm","kg")

**Note: O1 is the relative cell addressing.**

For Height: =Convert(convert(left(O1,Find("'",O1)-1),"ft","in")+Substitute(mid(O1,Find("'",O1)+1,Len(O1)),"""",""),"in","cm")

**Note: O1 is the relative cell addressing.**

1. Value, Wage and Release Clause, these columns contain unnecessary characters such as ‘â’, ‘‚’ and ‘¬’. These are removed using Find and Replace feature in Excel.
2. W/F, SM and IR, these columns contain unnecessary characters such as “â˜….” These are removed using substitute function as mentioned below:

*=SUBSTITUTE(A1,“â˜…”,)*

1. Position this column is having comma separated values, which are separated using “Text to Column” feature in excel.

Below can be found excel with cleaned data.

