**Excel – Soccer Transfers Analysis**

**Objective**

The main objective of this project is to showcase the use of Excel data cleaning tools like Remove duplicates, Find and Replace, and Text-to-columns; as well as functions such as VLOOKUP, SUMIFS, COUNTIFS, and RANK to analyse the data on our base data set and the newly created one.

This project was done by analysing a data set containing information related to two different country’s soccer season transfers, on our base dataset we find information like transfers incoming and outgoing by country and continent and the total club-to-club compensations of such transfers.

**Setup**

To be able to work with the data, the following data cleaning and processing tasks were applied:

* Transform a single column’s information into 2 different values/columns to be able to analyse information by country comfortably.
* Removing duplicates
* Finding and replacing errors for unnecessary spaces before the continent’s field and incorrect numbers regarding the season studied.
* Fill in the empty Continent column for both seasons.

A screenshot of a computer

Description automatically generated

Figure 1. Base dataset

Analysis & Data Exploration

With our dataset cleansed, the next step was to do some data exploration, the main aspect that we’re interested in, is knowing the net transfers within Europe for both seasons and which are the top 5 European countries regarding their spending in incoming transfers.

The steps needed to find the answer to that have been:

* Using COUNTIFS function to know what the net transfers are incoming and outcoming for Europe, this saves a lot of time in big datasets as one would need to count one by one the entries without it.
* Using SUMIFS to obtain the net incoming and outcoming transfers numbers and expenses, but this time doing it by country instead of continent like the previous step.

A screenshot of a computer

Description automatically generated

Figure 2. Newly created database

After working with the data, the following chart was obtained.

A screenshot of a graph

Description automatically generated

Figure 3. Final table and chart

**Key Findings**

1. By using the RANK function, we obtained the top 5 countries with the highest spendings regarding incoming transfers, we can see that the highest one, England, doubles the spendings of the fifth one, this means that English soccer incoming transfers are an important source of investing for the soccer industry since it is ready to spend important amounts on this field.
2. By looking at the Average transfer fee line, we can see that this value is not strongly correlated to the number of transfers, in other words, a lower number of transfers, does not necessarily mean a lower amount of incoming investment, quantity ≠ quality.
3. By understanding the two previous points, and doing further exploration on the data, we could tell important findings like the best continent or countries where to invest or if we had more data from previous seasons, we could do the same to know which are the countries that have been growing and would be a good investment opportunity.