### **1. Introduction**

#### 1.1 Dataset Overview

The dataset under analysis contains information about football players, including attributes such as nationality, overall rating, club affiliation, playing position, age, contract details, and various skill ratings. The dataset aims to provide insights into the distribution and characteristics of football players within the context of the FIFA video game.

### **2. Initial Exploration**

#### 2.1 Data Inspection

* Checked the structure of the dataset using functions like head(), info(), and describe() to understand the types of variables, missing values, and basic statistics.(BY Python)
* used power query editor by power Bi to explore the data

#### 2.2 Preliminary Observations

* Identified key variables such as 'Nationality,' 'Overall Rating,' 'Club,' 'Position,' 'Age,' 'Contract Details,' and 'Skill Ratings' as primary areas of interest for analysis.

### **3. Data Cleaning**

#### 3.1 Handling Missing Values

* Checked for and addressed missing values in critical columns, ensuring data integrity.
* full steps in powerBI:

1. PhotoURL' and 'PlayerURL' columns contain links that can be used to access the image and details of all the players represented in the data. These columns were removed from the data.
2. The 'Contract' column had inconsistent values and the wrong data type. The column has 3 categories of values in the format, '23rd July 2020 on loan’, 'Free' and '2018 ~ 2024’.To tackle this issue, I had the column split using the space delimiter to give two columns containing the Contract start date and the Contract end date.Process: I split the column using 'space' as the delimiter. Over 4 columns were created, got rid of all except the columns with Contract start and Contract end details. The columns were renamed, and the entries which were not in year format were replaced with 'Null’. This was done to allow the conversion of these two new columns to date types and also nullify records of players on 'Loan' and 'Free' from the Contracts column.
3. 'Positions' column was removed because there is another column bearing the 'Best\_Position' of each player. The 'Positions' column had the details in the 'Best\_Position' column in addition to the extra positions for players that had more than one position.
4. 'Weak\_foot’, 'Skill\_moves’, and 'International\_reputation' columns all contain ratings on a scale of 1-5. The values in these columns had the special character '☆' in front and the data type was text I replaced the special character with space, wrote the column names in full and then converted the column data types from text to whole number
5. 'Value’, 'Wage' and 'Release\_clause' columns had data type issues and values with suffixes in front. These columns had suffixes 'M' for millions, 'K' for thousands and '€' for euro (currency) The suffixes were replaced with space, data types were converted, and the custom column was created for the multiplication of values in these original columns and the values in the conditional columns. The new columns containing the normalized values were converted to whole number data types and renamed

#### 3.2 Data Type Conversion

* The data types were checkmated, and changes were made for some that had wrong data types and column names corrected.

3.3 unwanted columns were removed.

### **4. Exploratory Data Analysis (EDA)**

#### 4.1 Nationality Analysis

* Explored the distribution of players by nationality.
* Calculated the average overall rating by nationality.
* Identified the top nationalities represented in the dataset.

#### 4.2 Club Analysis

* Analyzed the distribution of players across different clubs.
* Calculated the average overall rating by club.
* Identified top clubs with the highest-rated players.

#### 4.3 Positional Analysis

* Explored average ratings for different playing positions.
* Investigated positional preferences in terms of left or right foot dominance.

#### 4.4 Age Analysis

* Examined the distribution of player ages.
* Calculated the average overall rating by age group.
* Identified players with the highest potential in different age groups.

#### 4.5 Contract Analysis

* Explored average wage by club.
* Analyzed contract length distribution.
* Identified players with the highest release clauses.

#### 4.6 Skill Analysis

* Explored the distribution of specific skill ratings.
* Identified players with exceptional skills in certain areas.

#### 4.7 Work Rate Analysis

* Investigated the distribution of attacking and defensive work rates.
* Explored the impact of work rates on overall ratings.

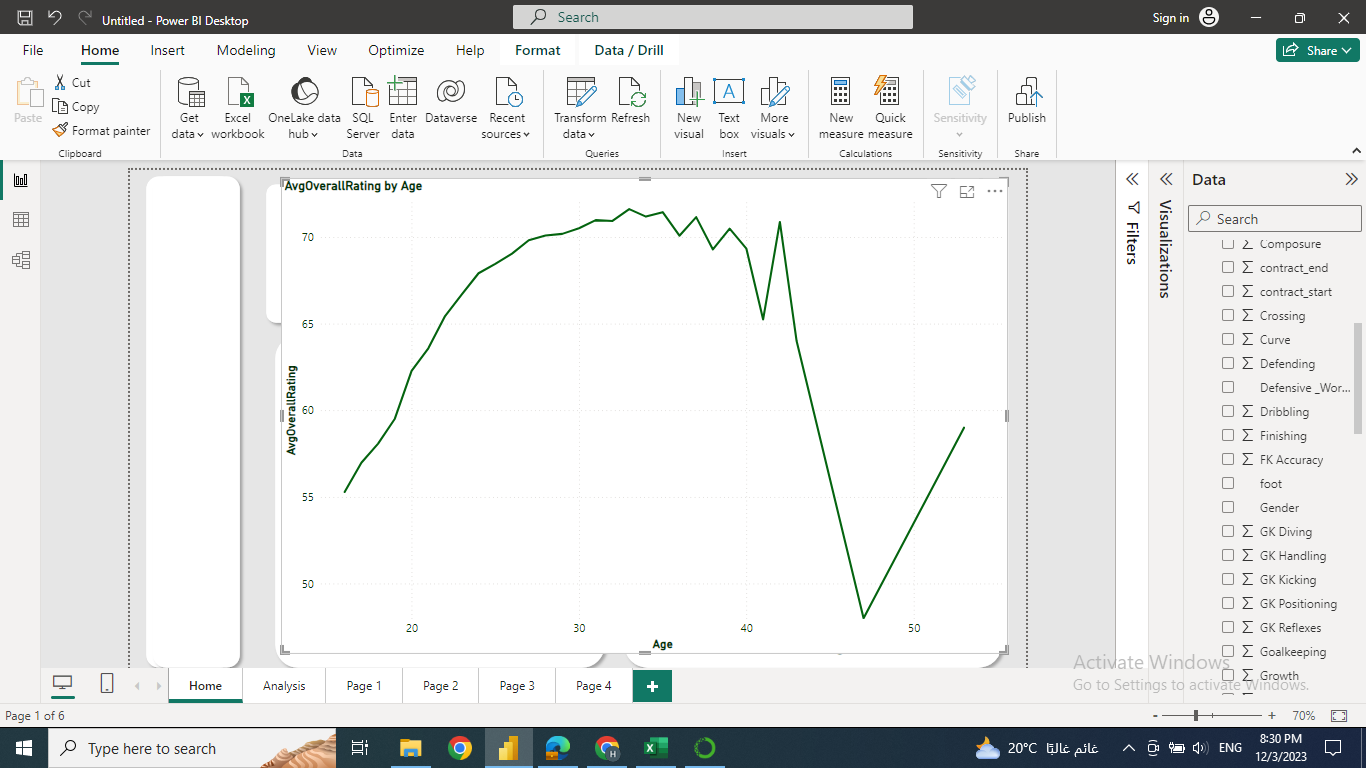
#### 4.8 Physical Attributes Analysis

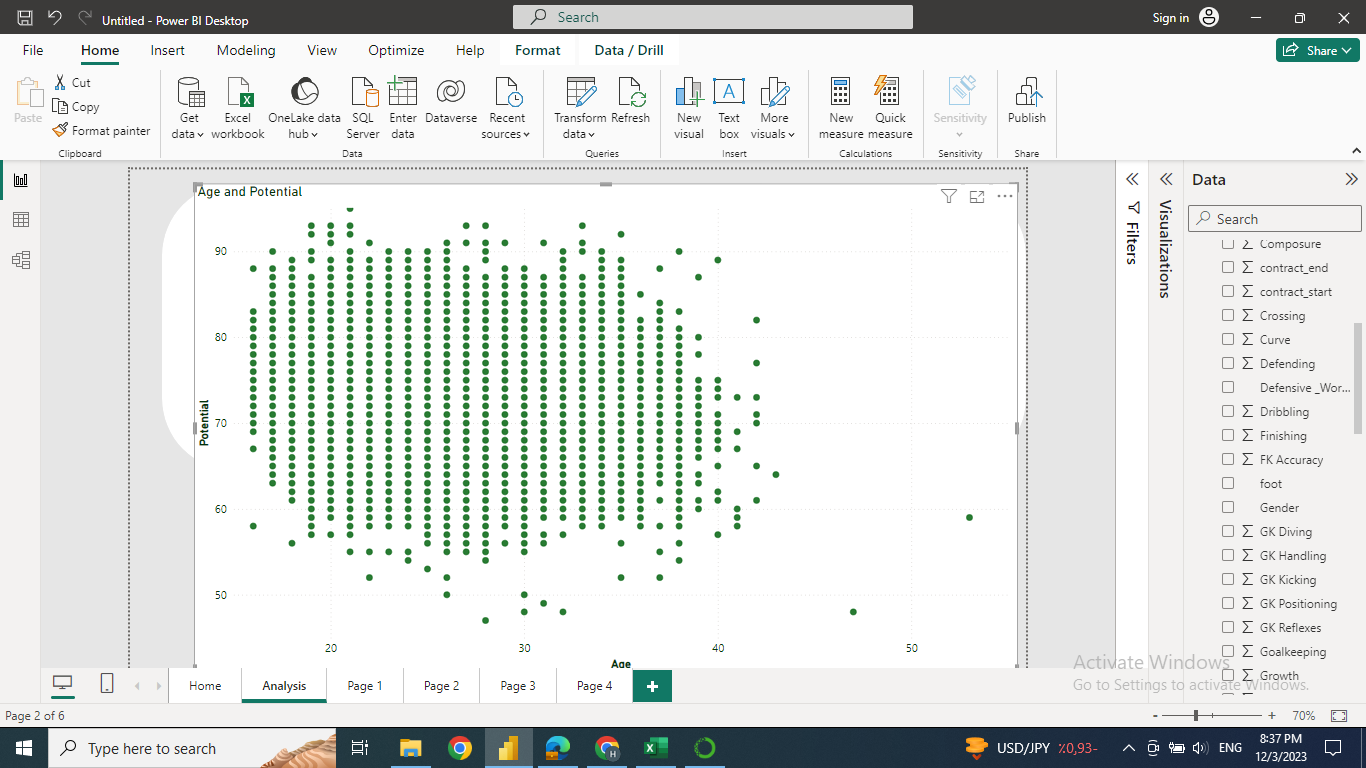
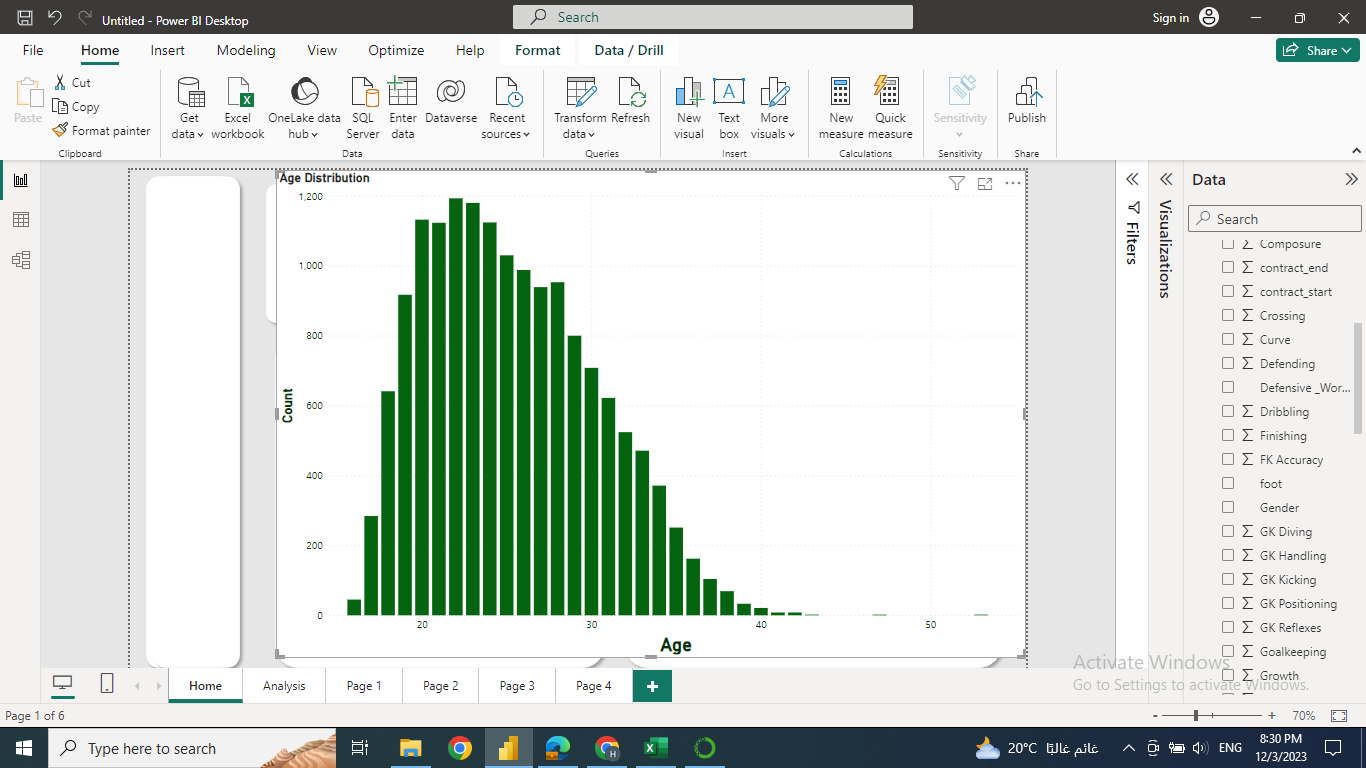
* Explored the distribution of physical attributes like height, weight, and strength.

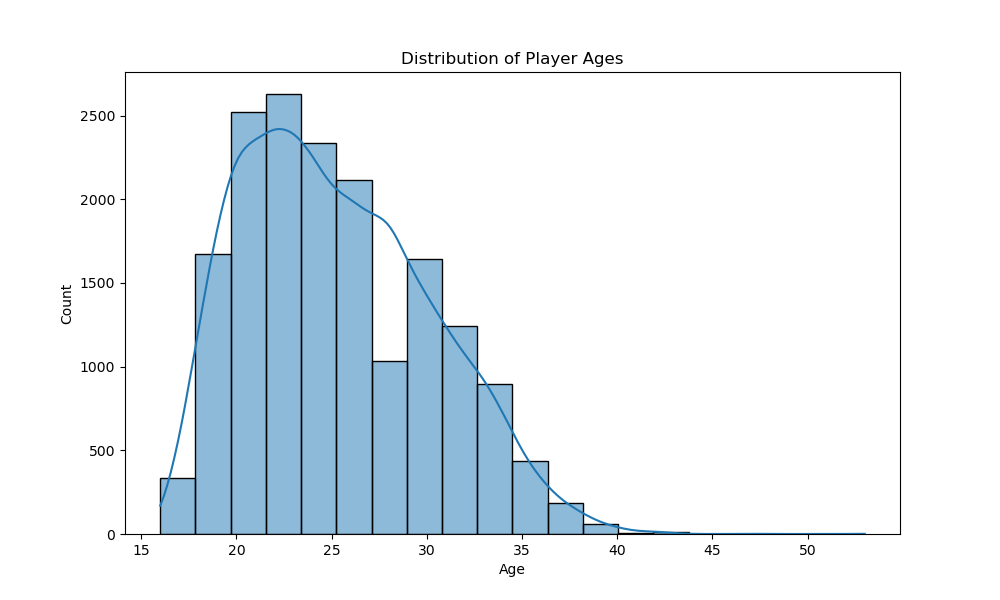
#### 4.9 Potential Growth Analysis

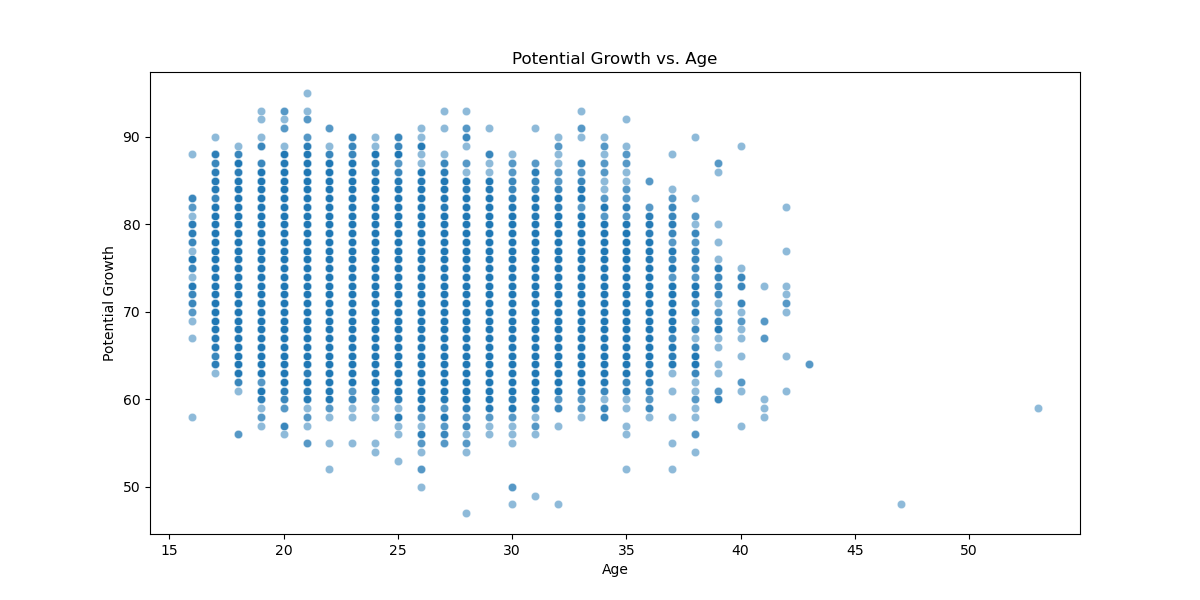
* Identified players with the highest potential for growth.
* Explored the relationship between age and potential growth.

### **5. Visualizations**

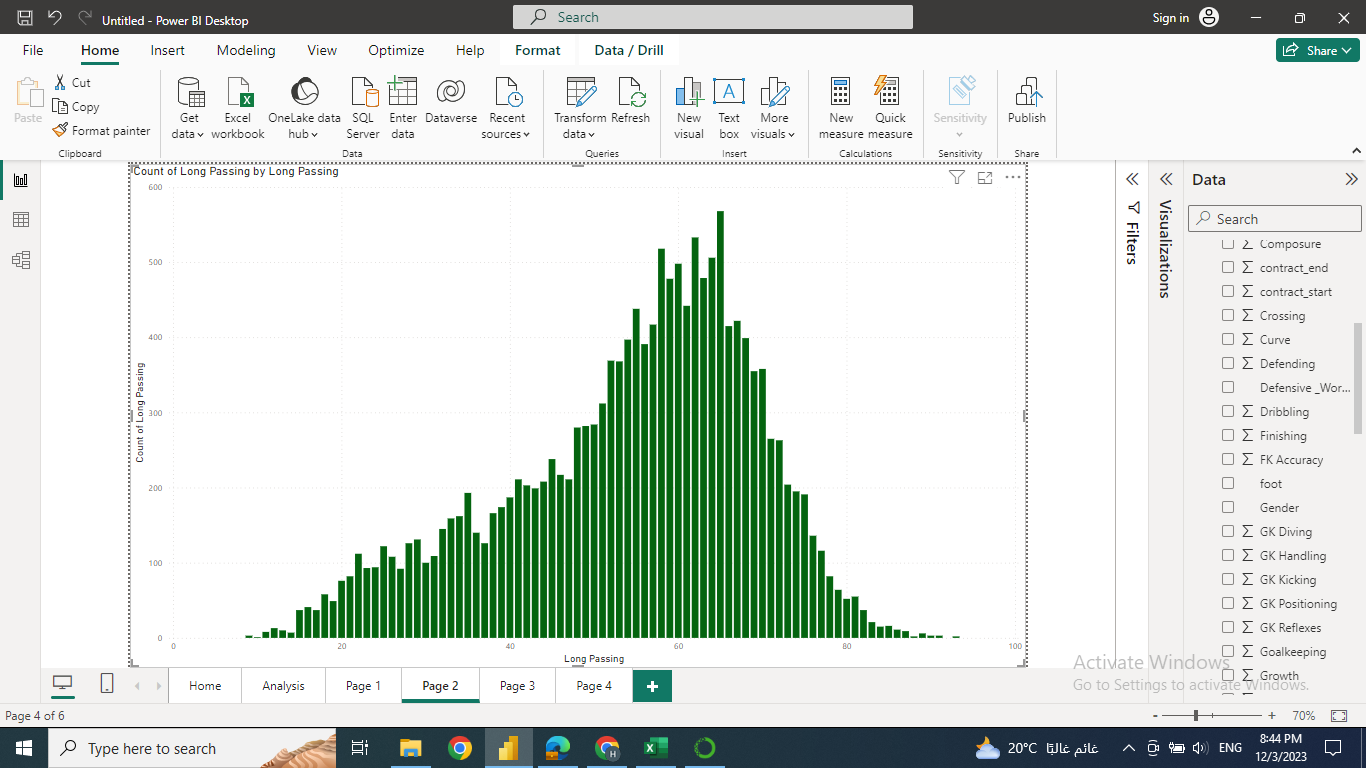
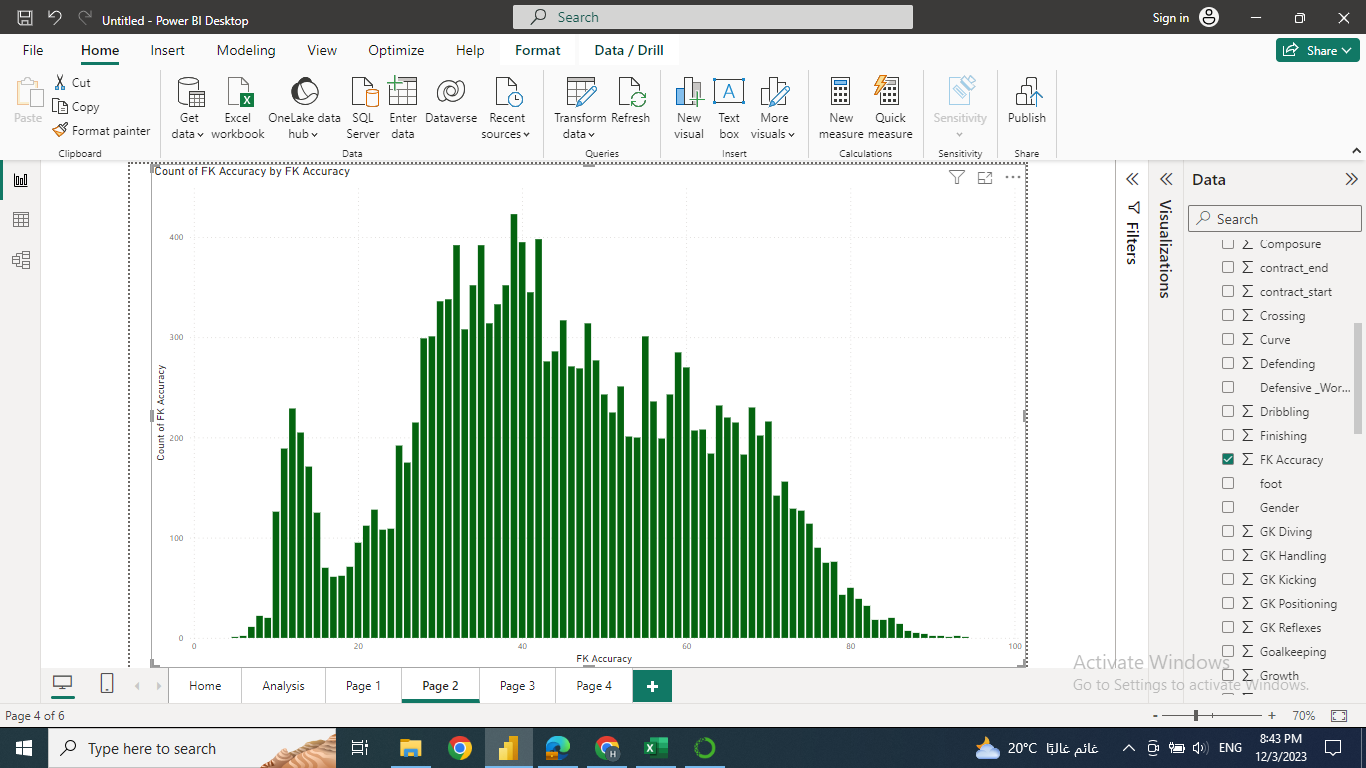
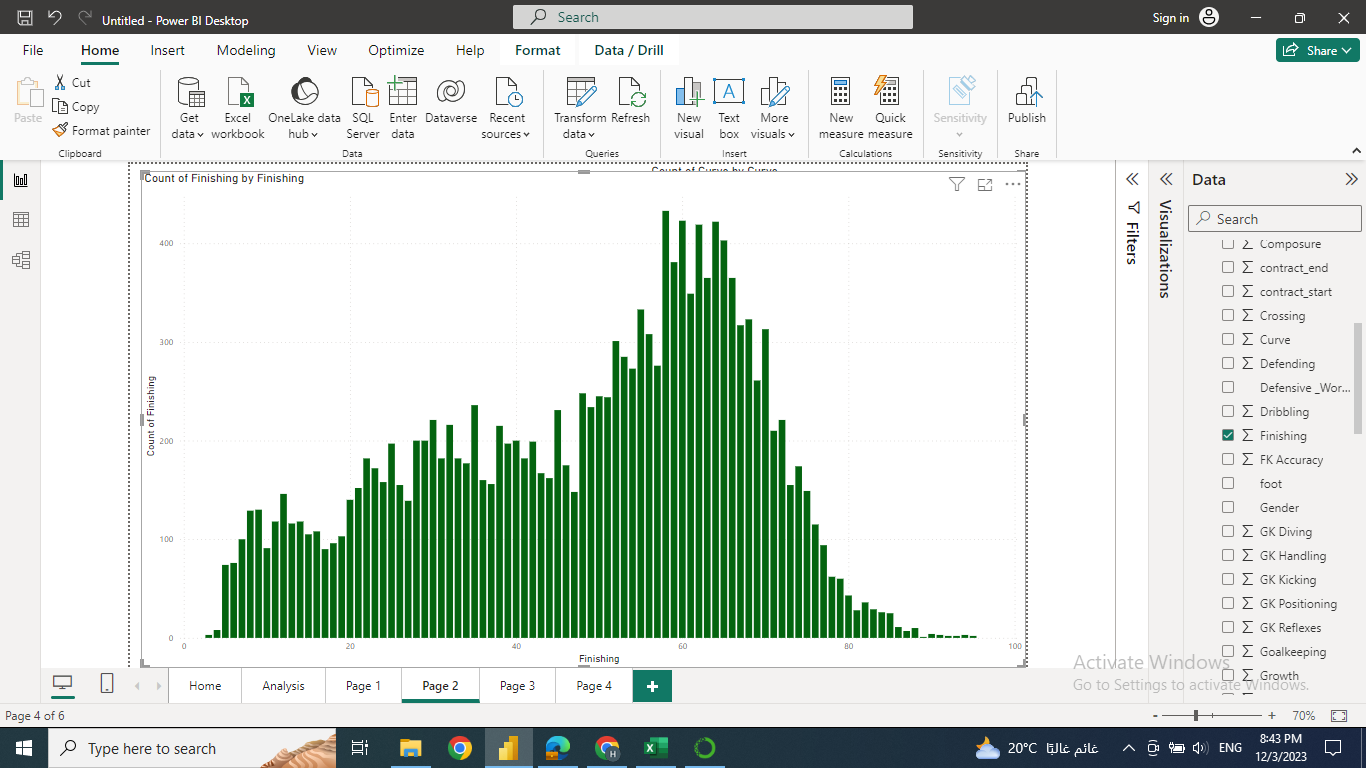
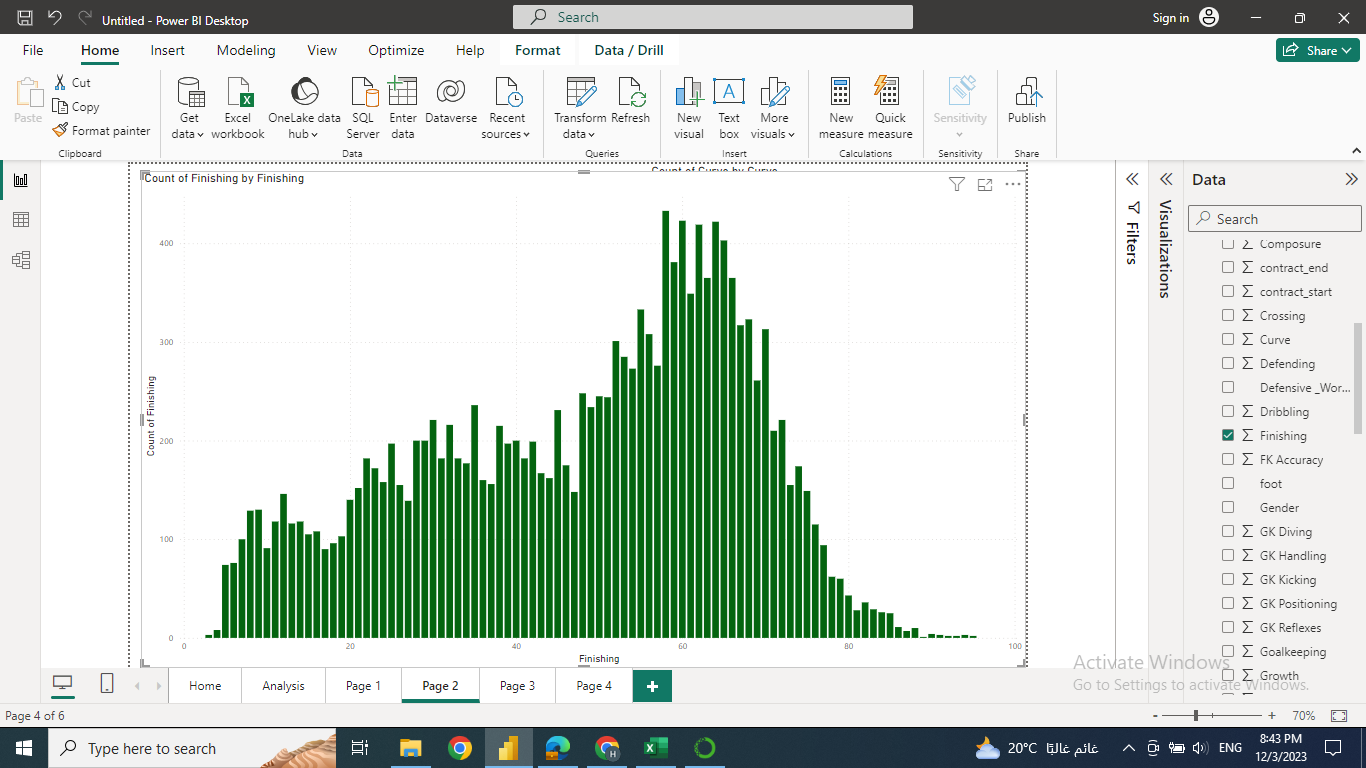
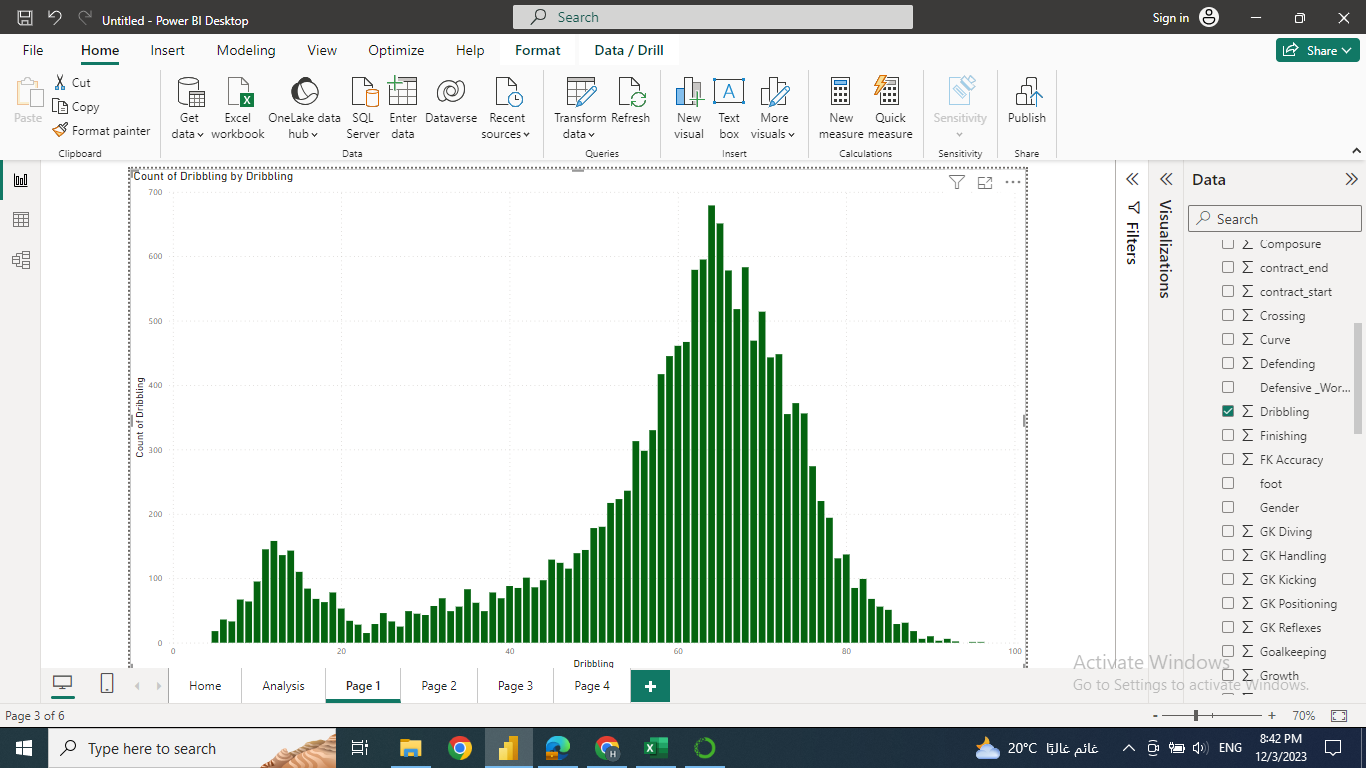
Age Analysis In Power Bi:



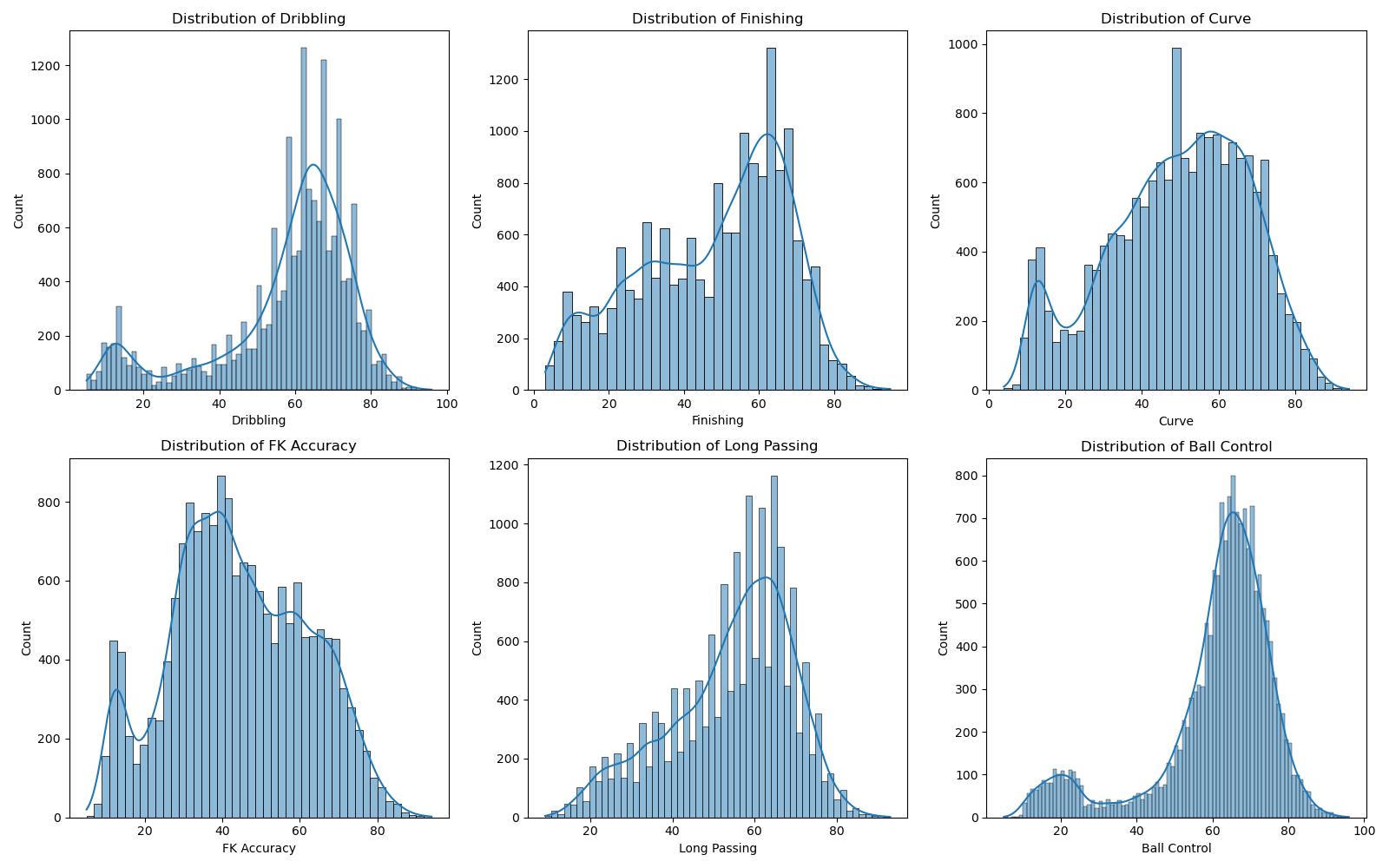
Age analysis using python



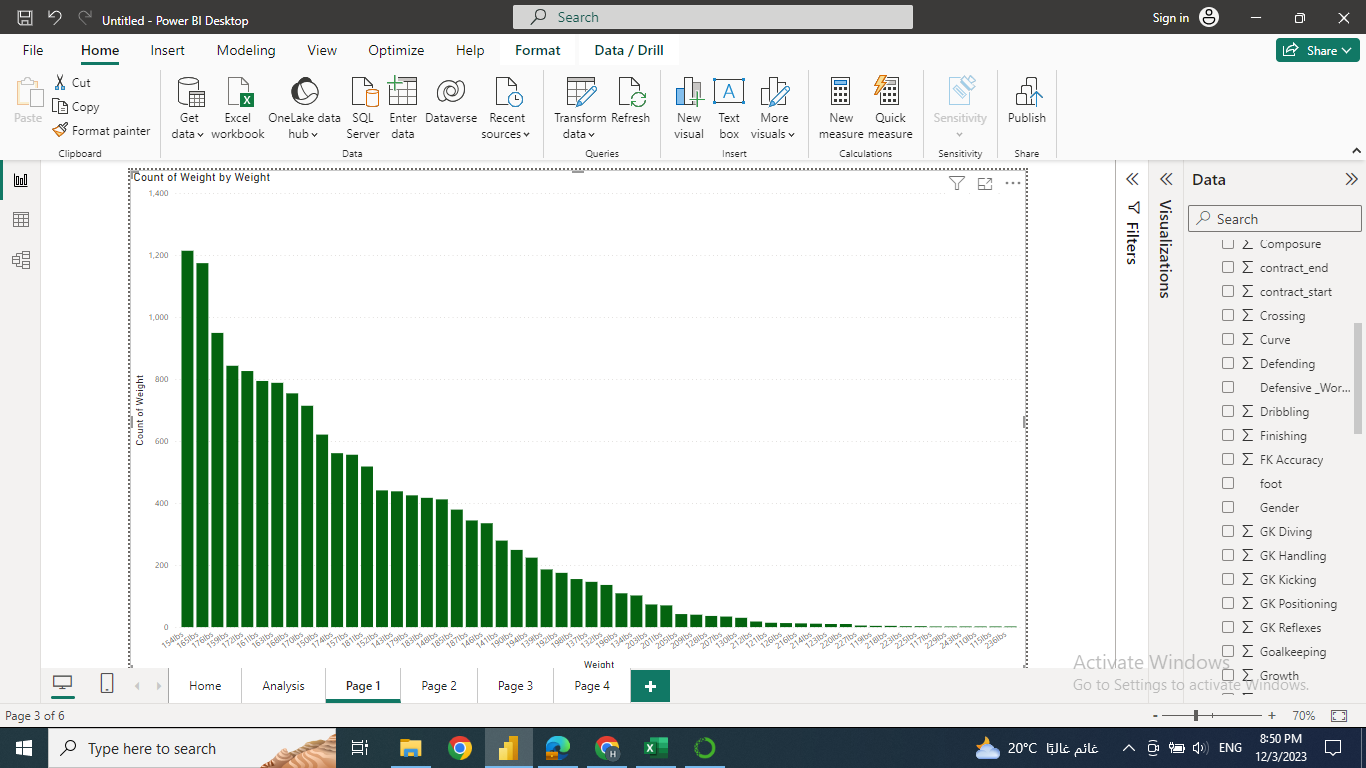
#### Skill Analysis in power Bi

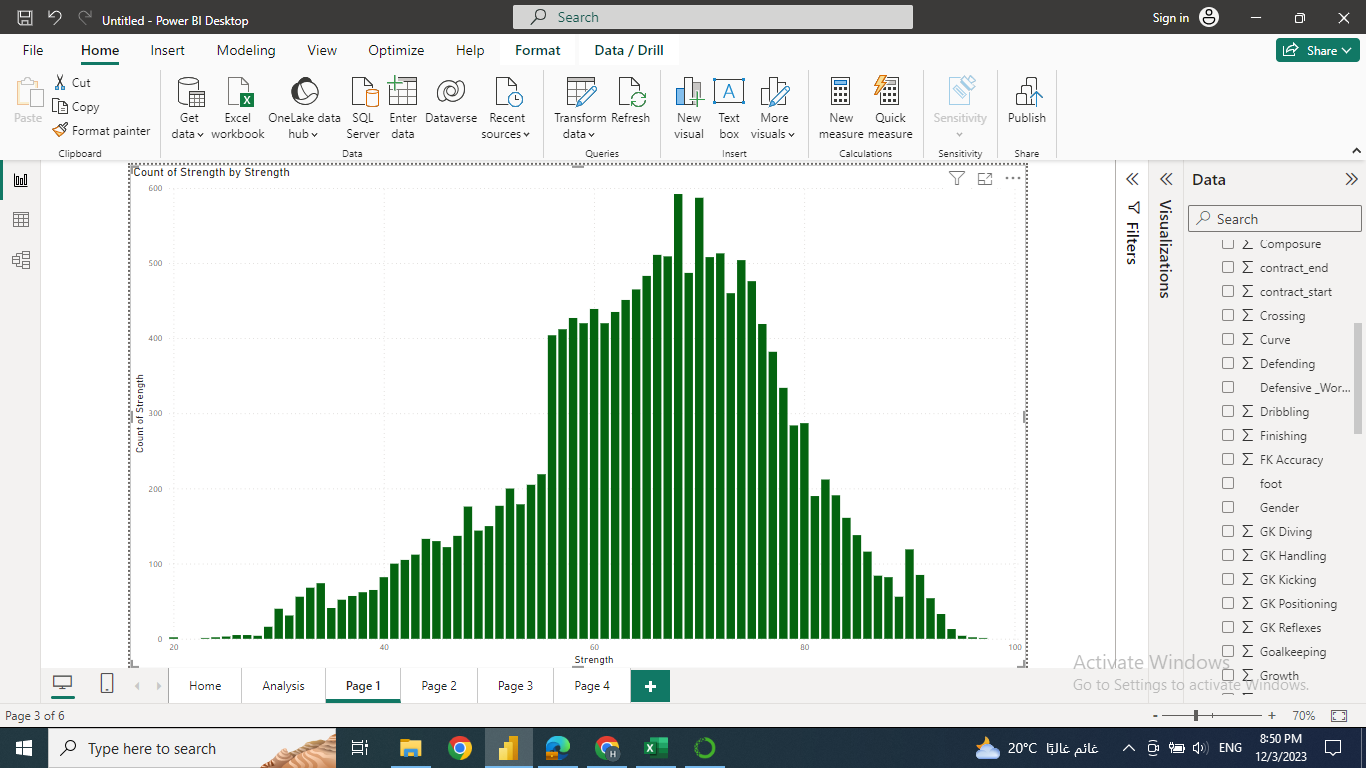


Skill analysis in python

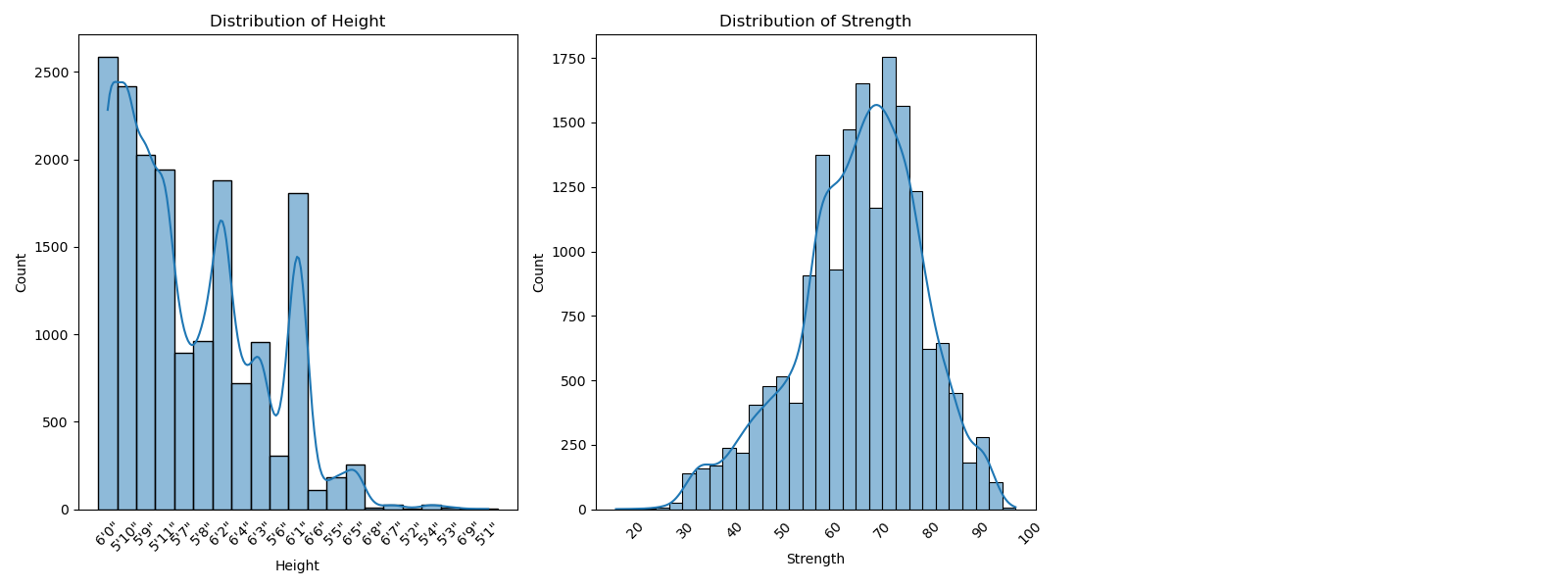


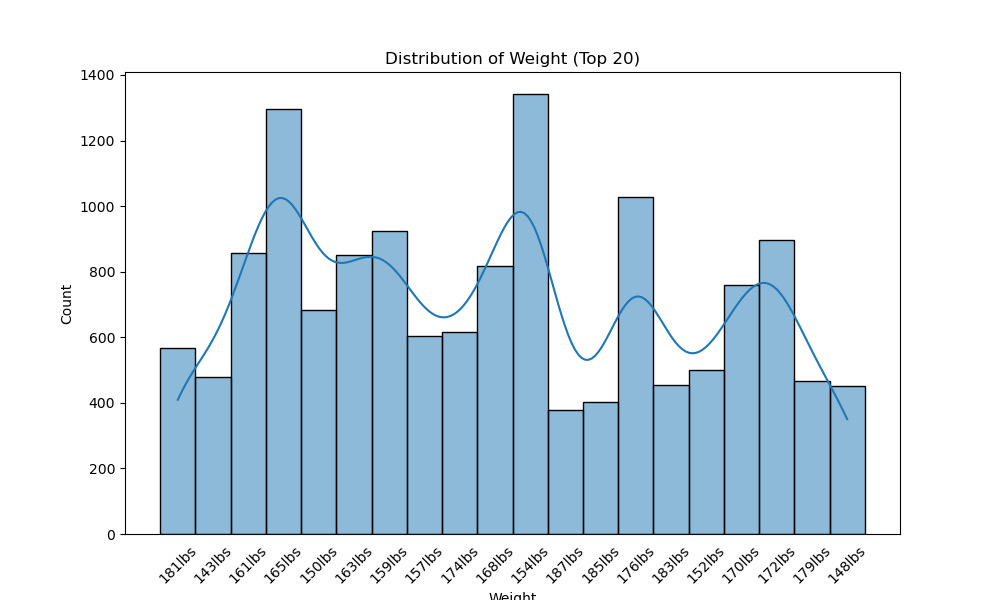
#### Physical Attributes Analysis in powerbi



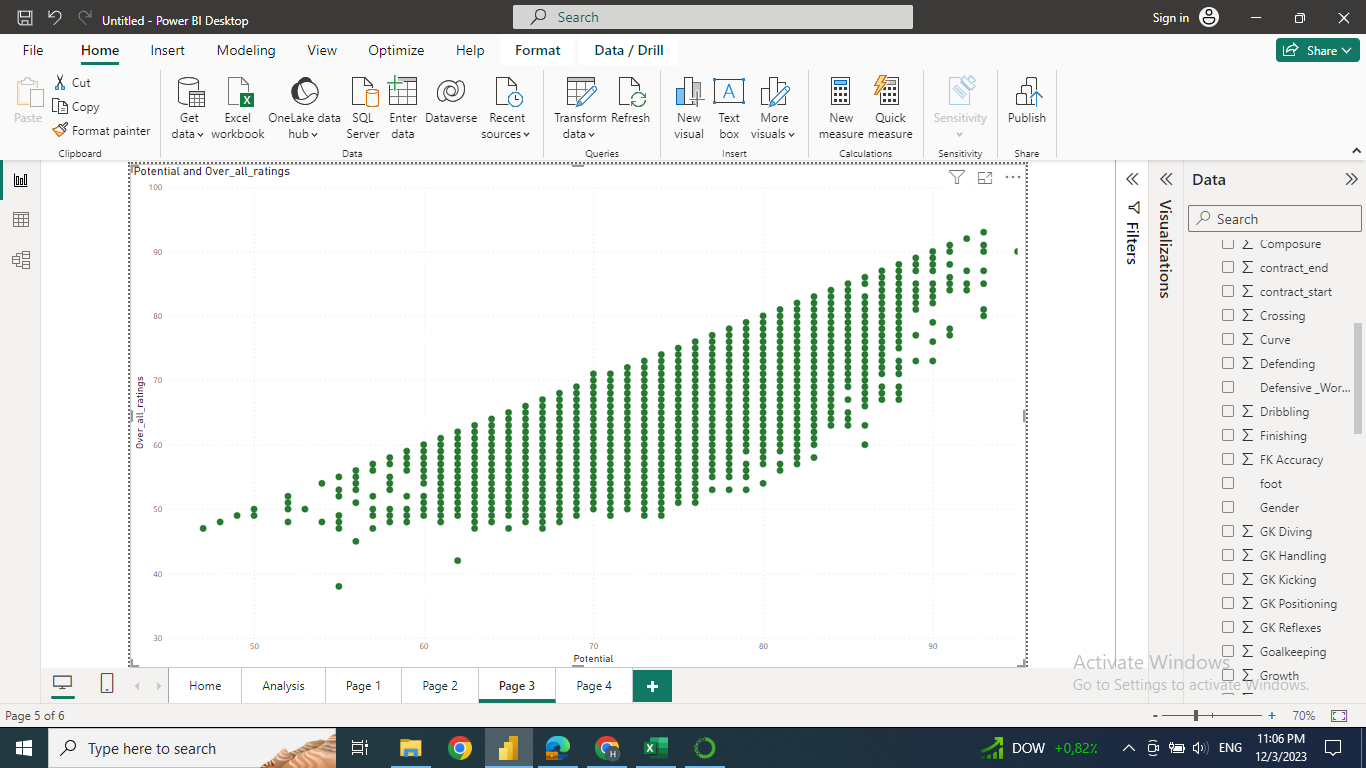


physical analysis in python:

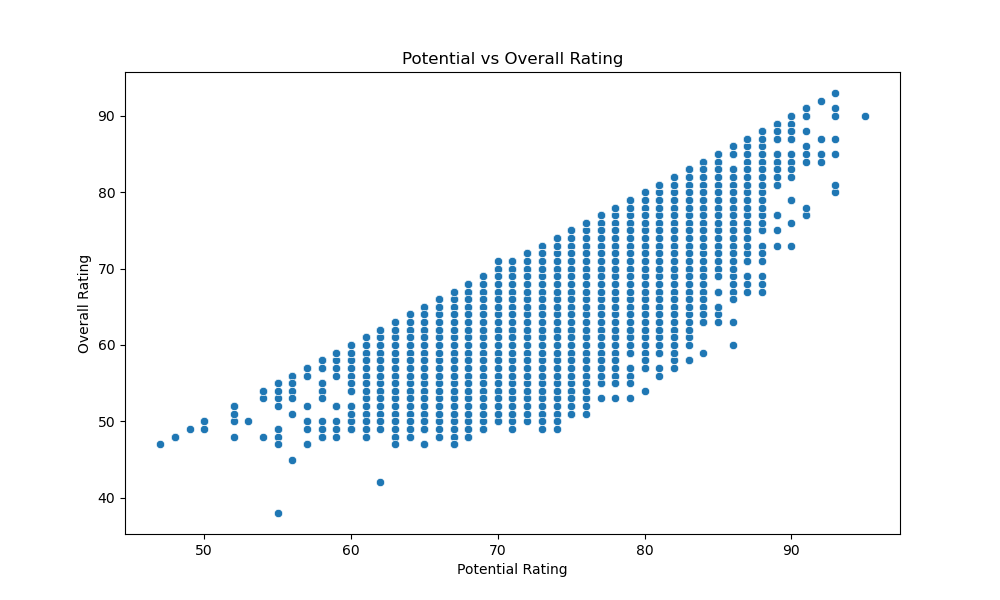




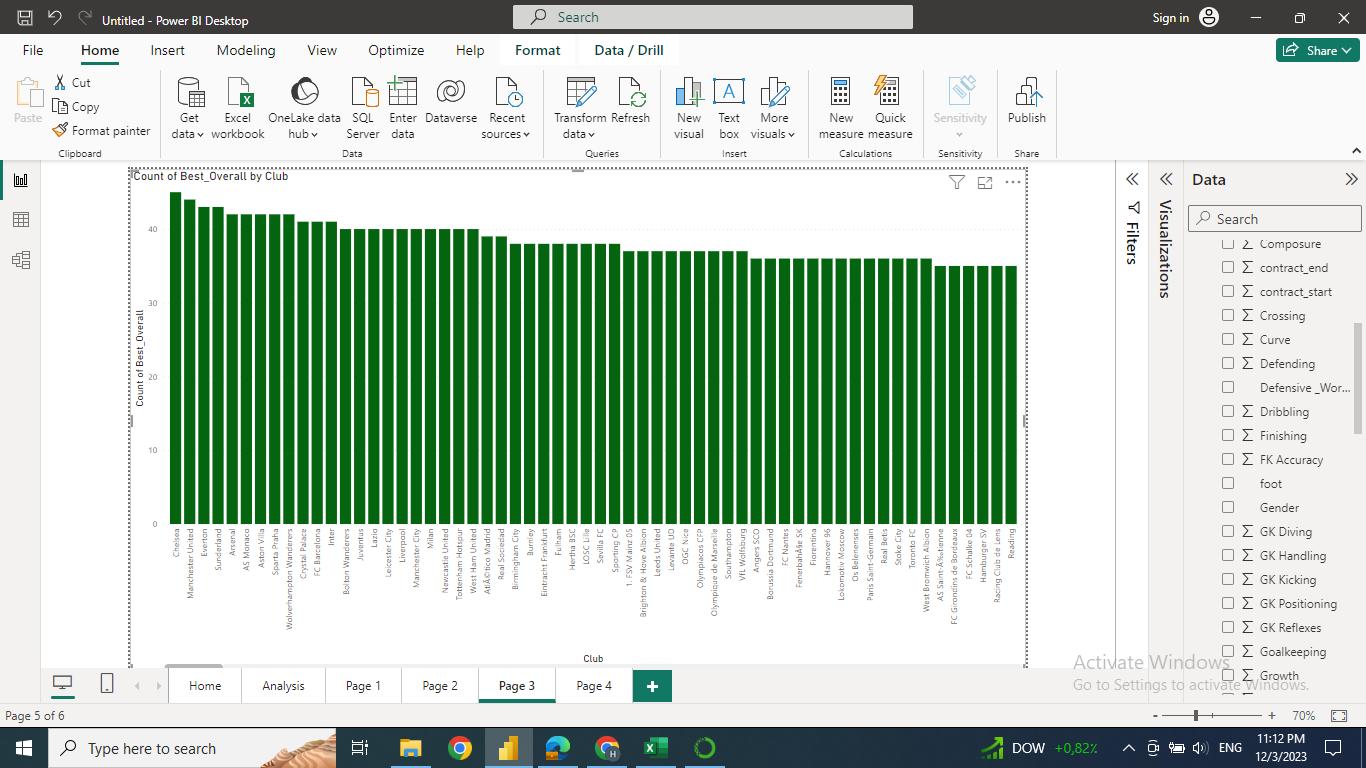
potential analysis in power bi

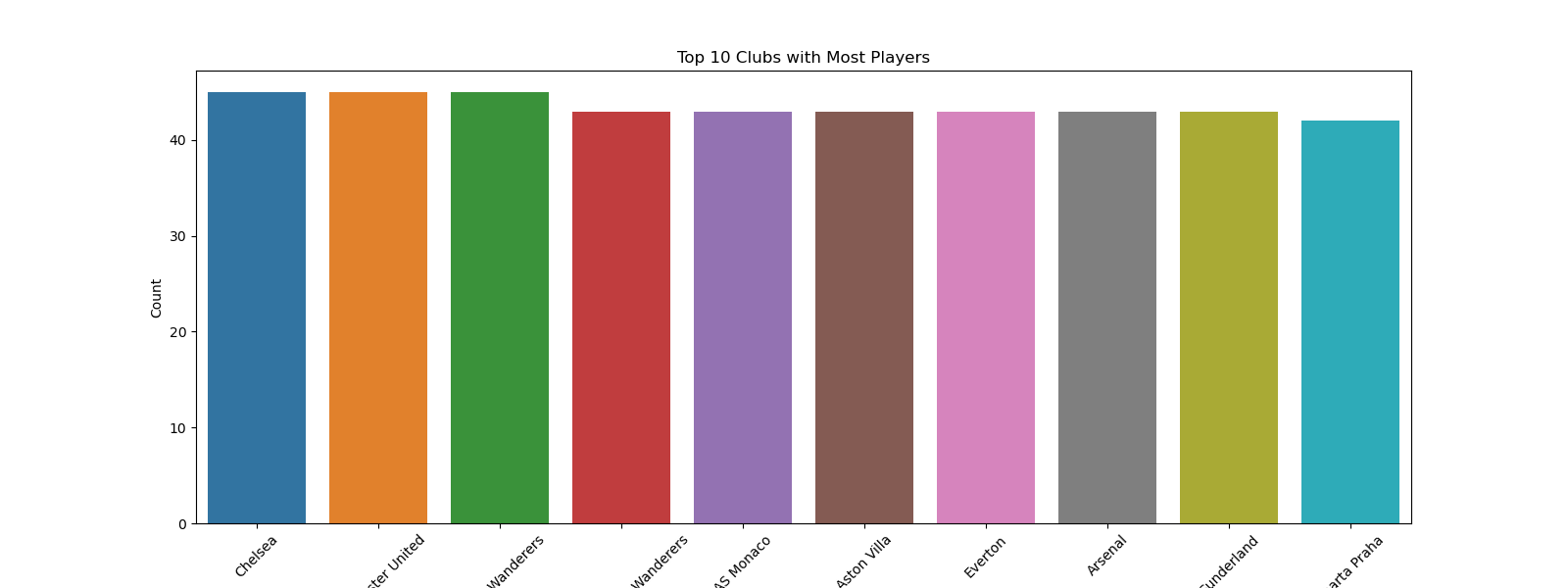


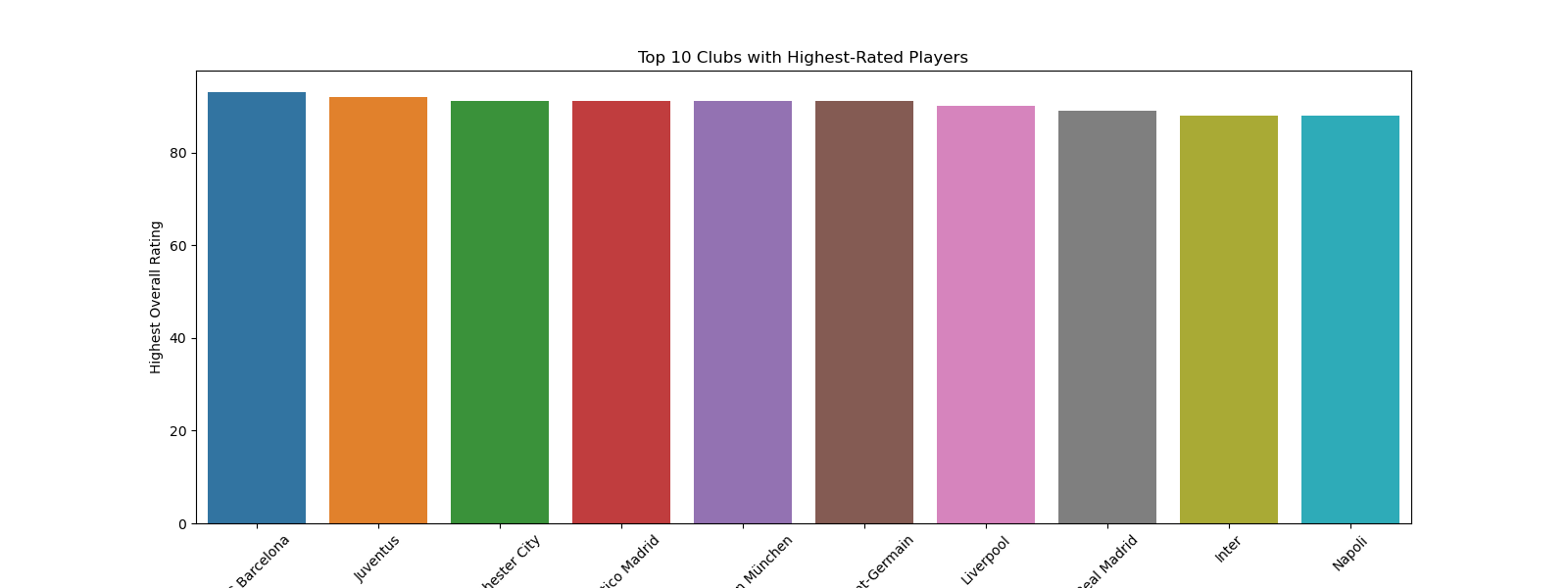
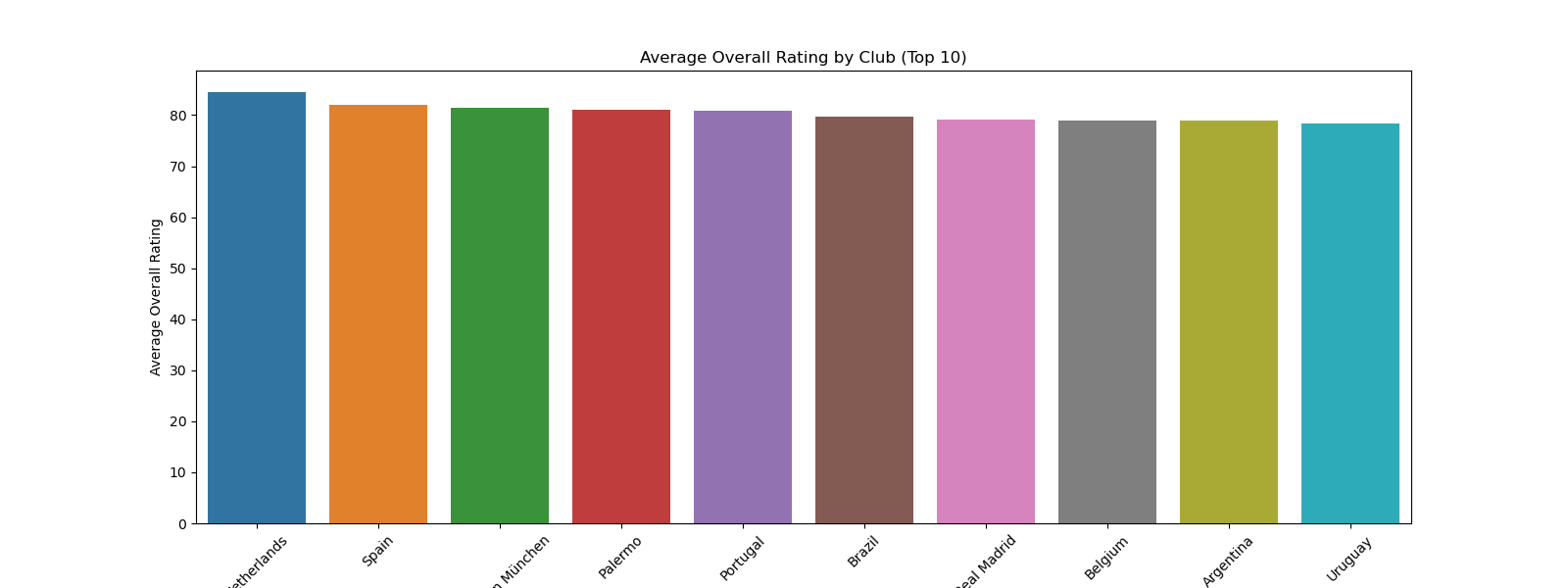
in python:



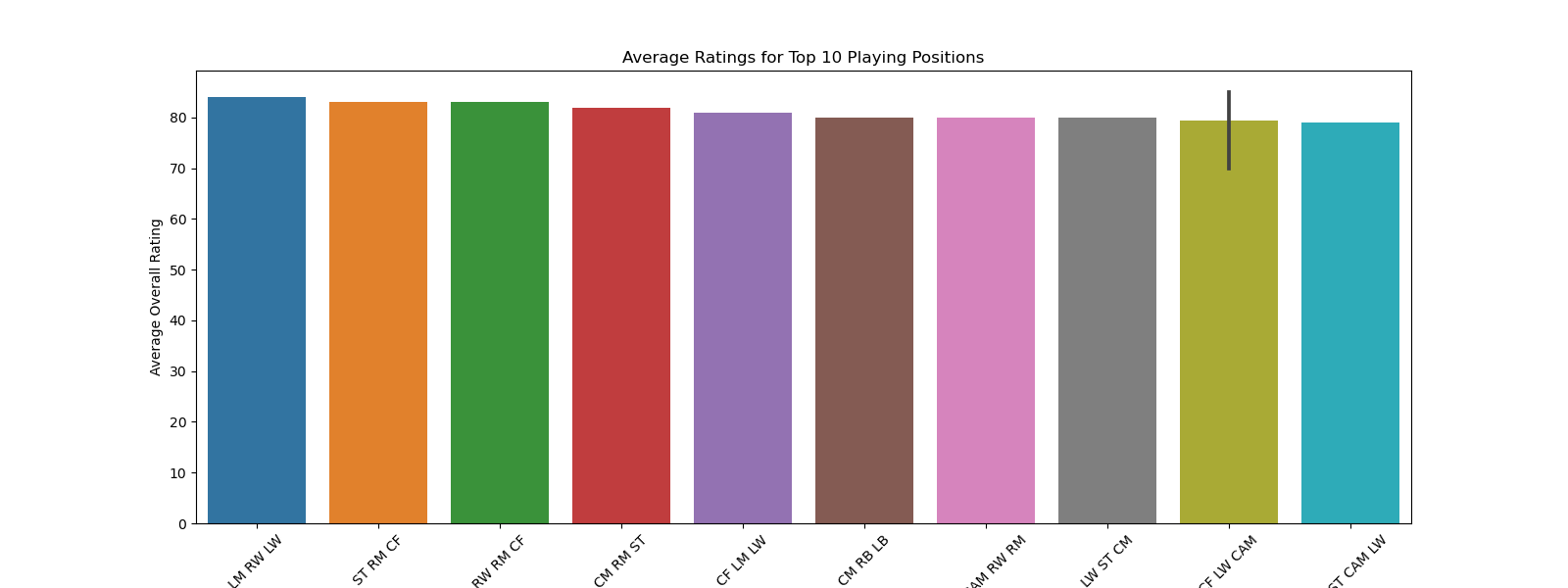
Club analysis powerbi and python:

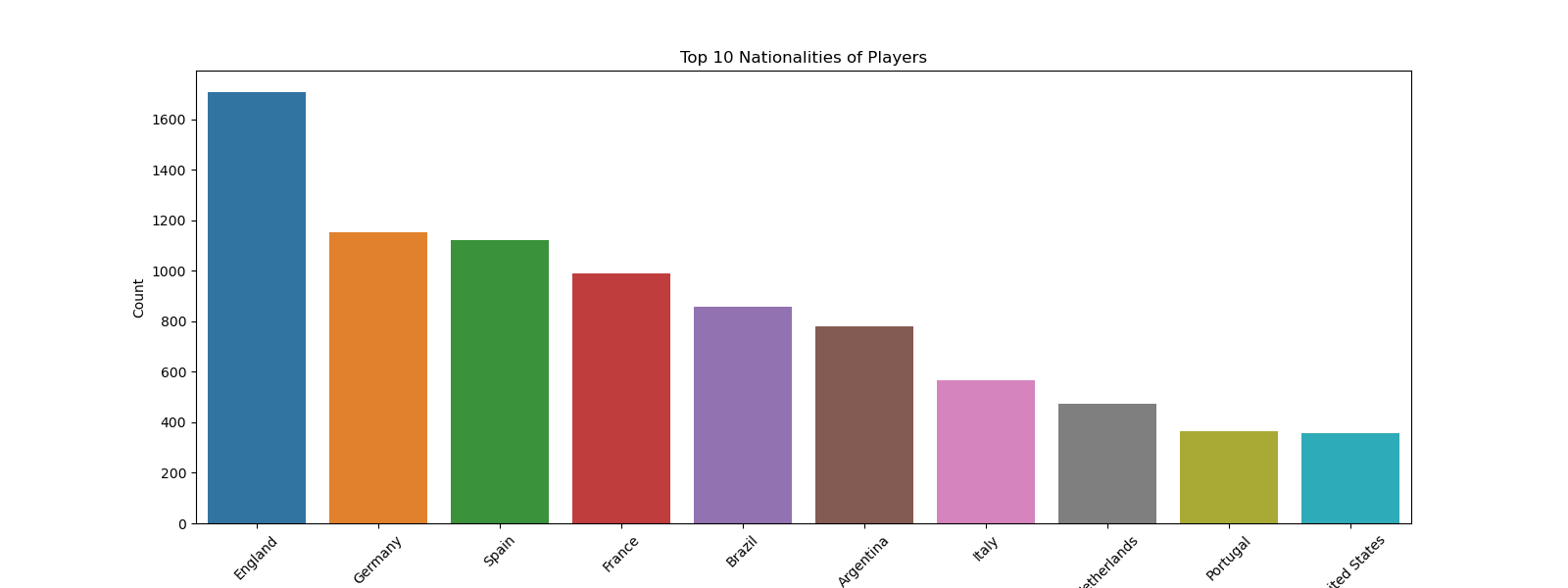
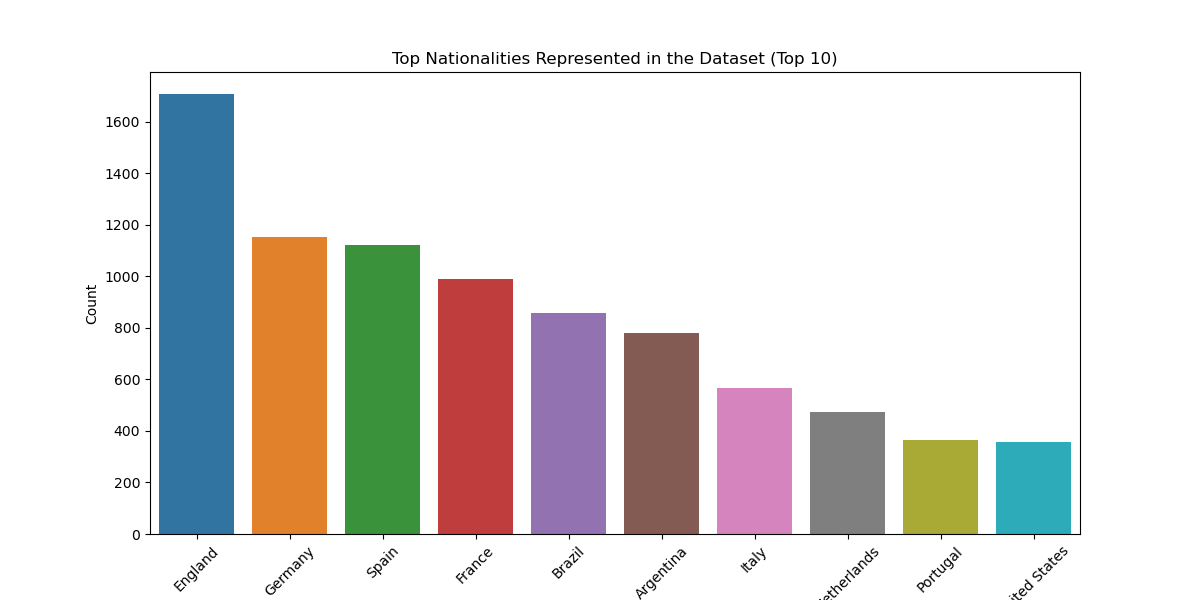


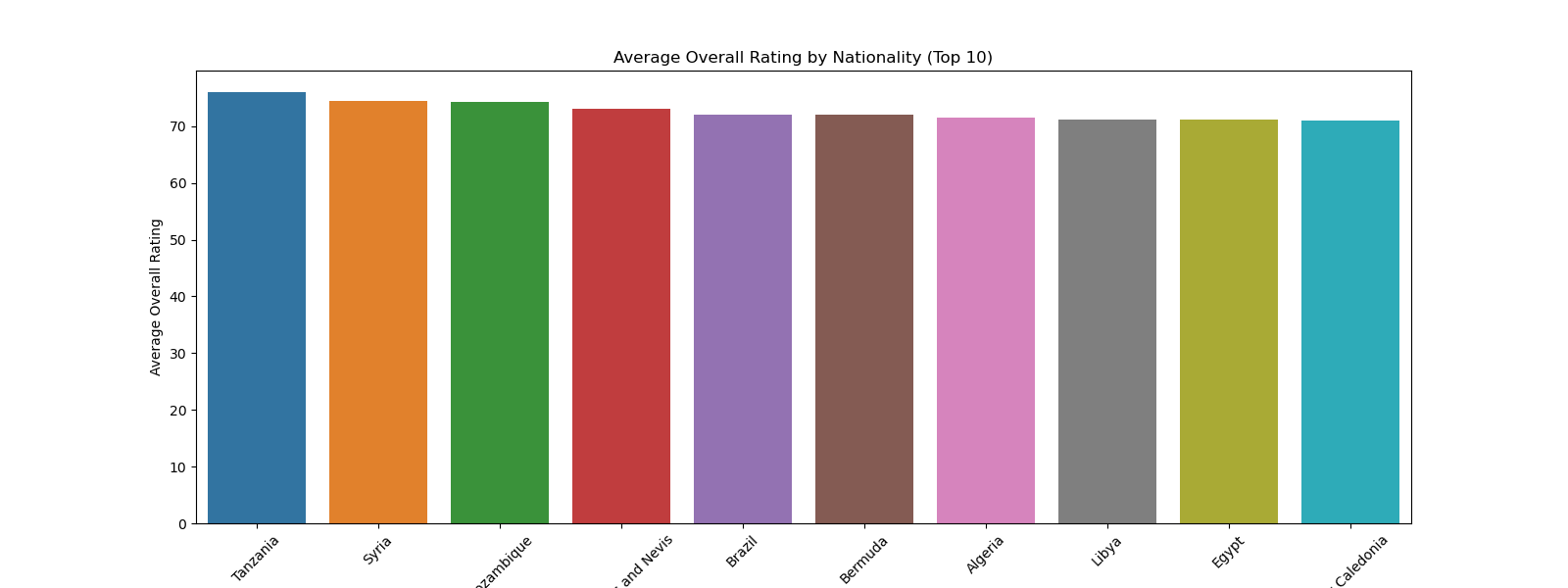




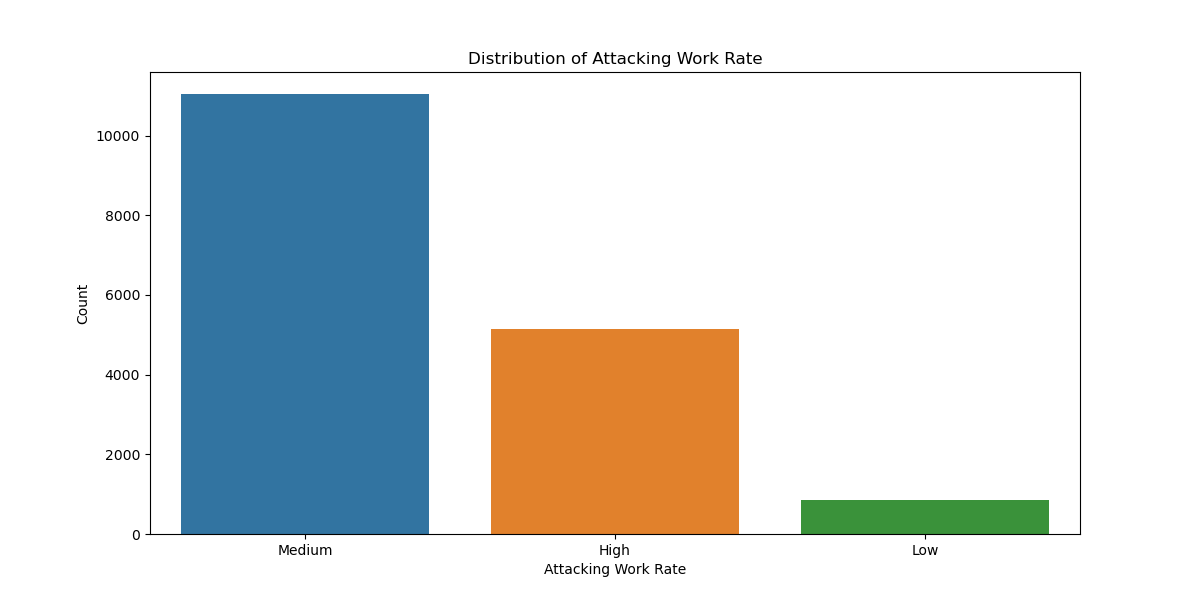
Nationality analysis

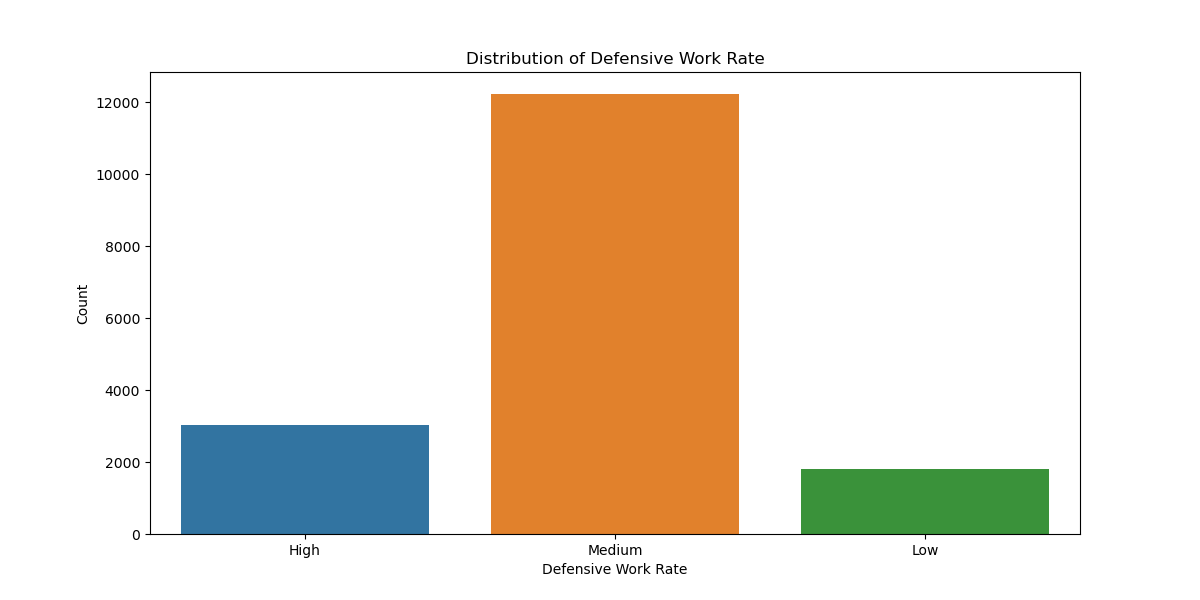


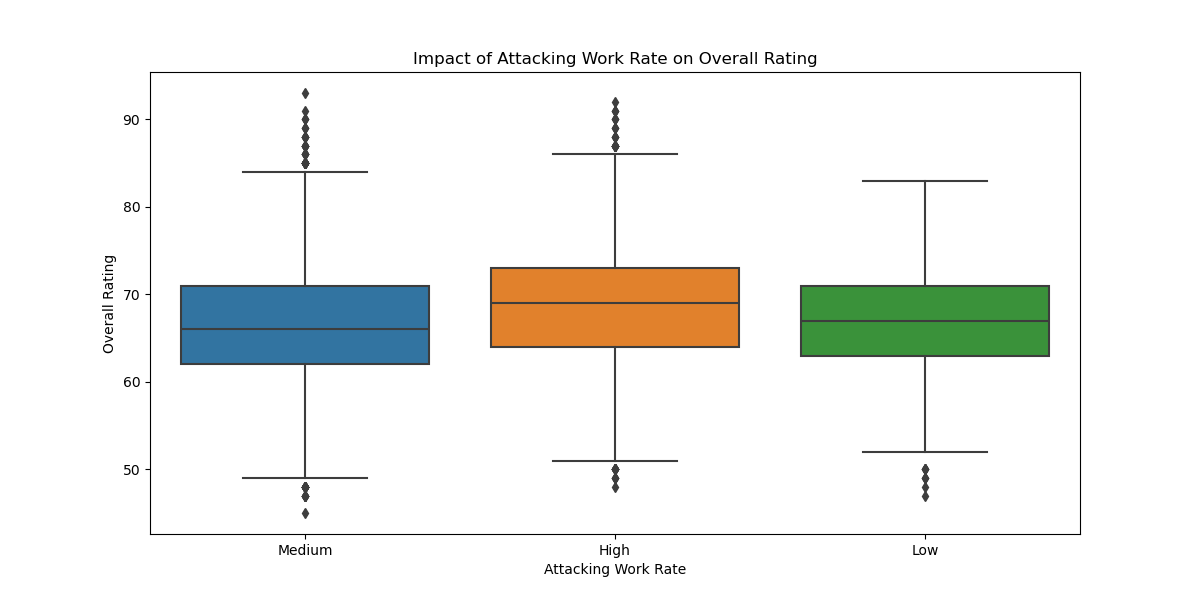


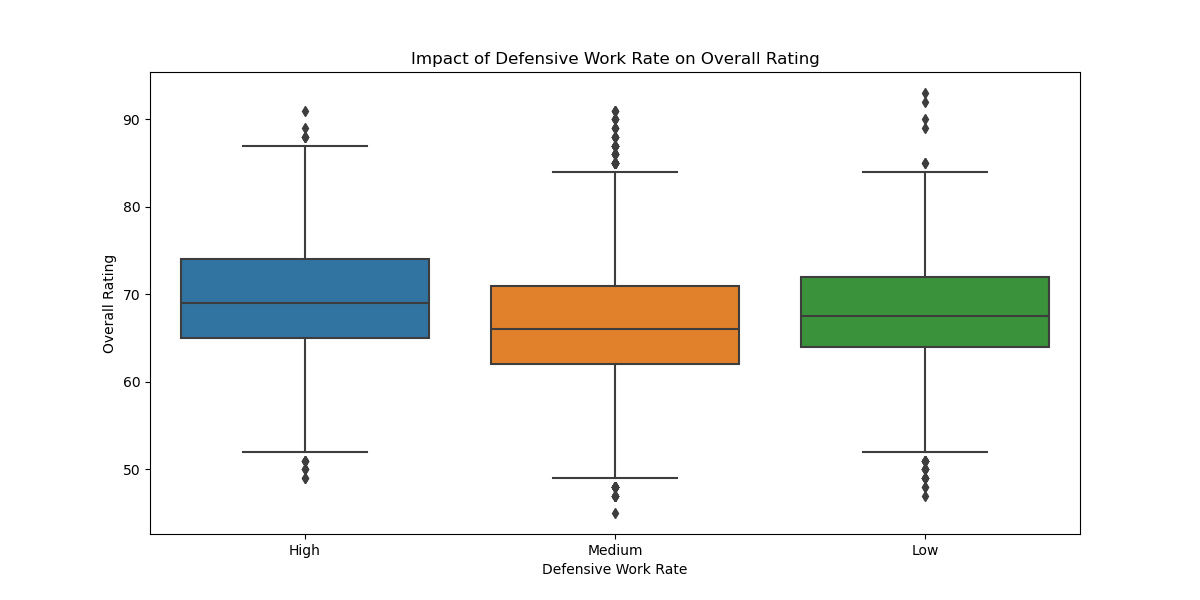


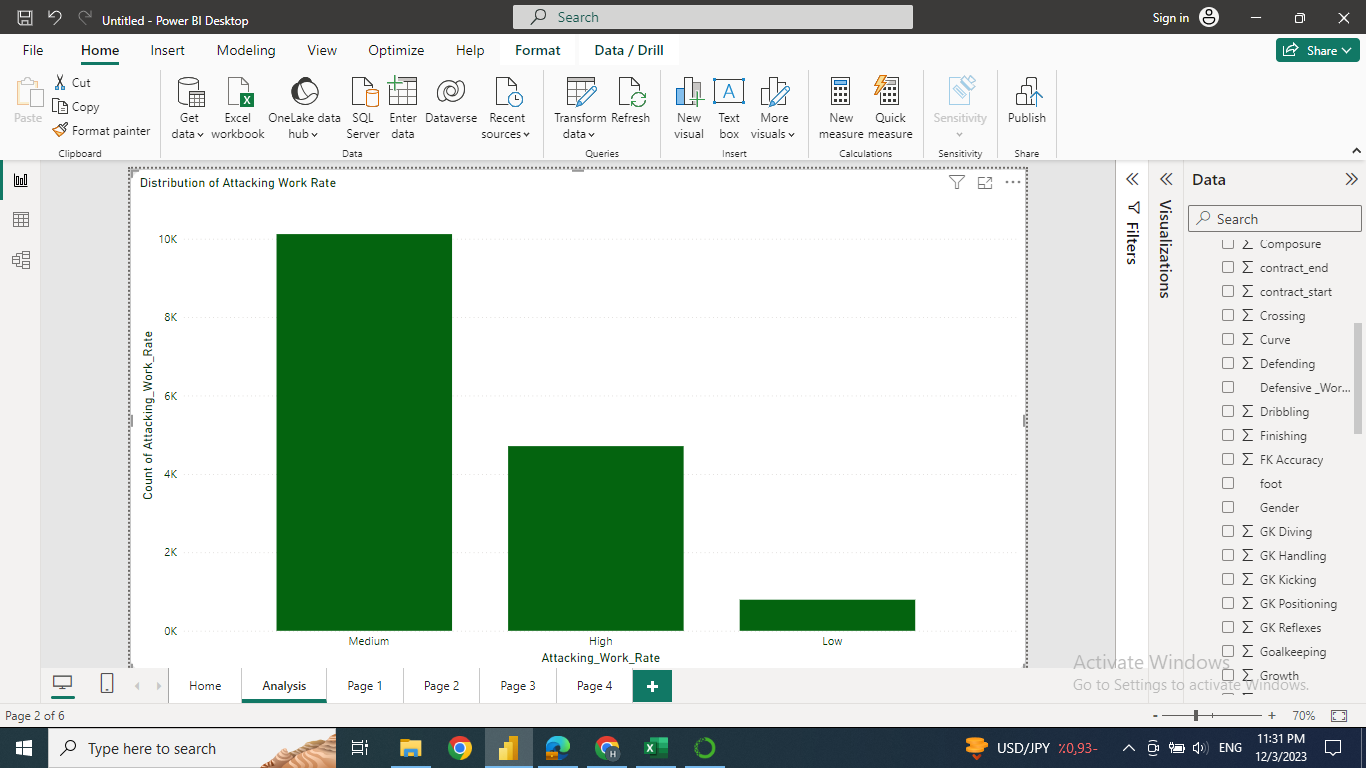
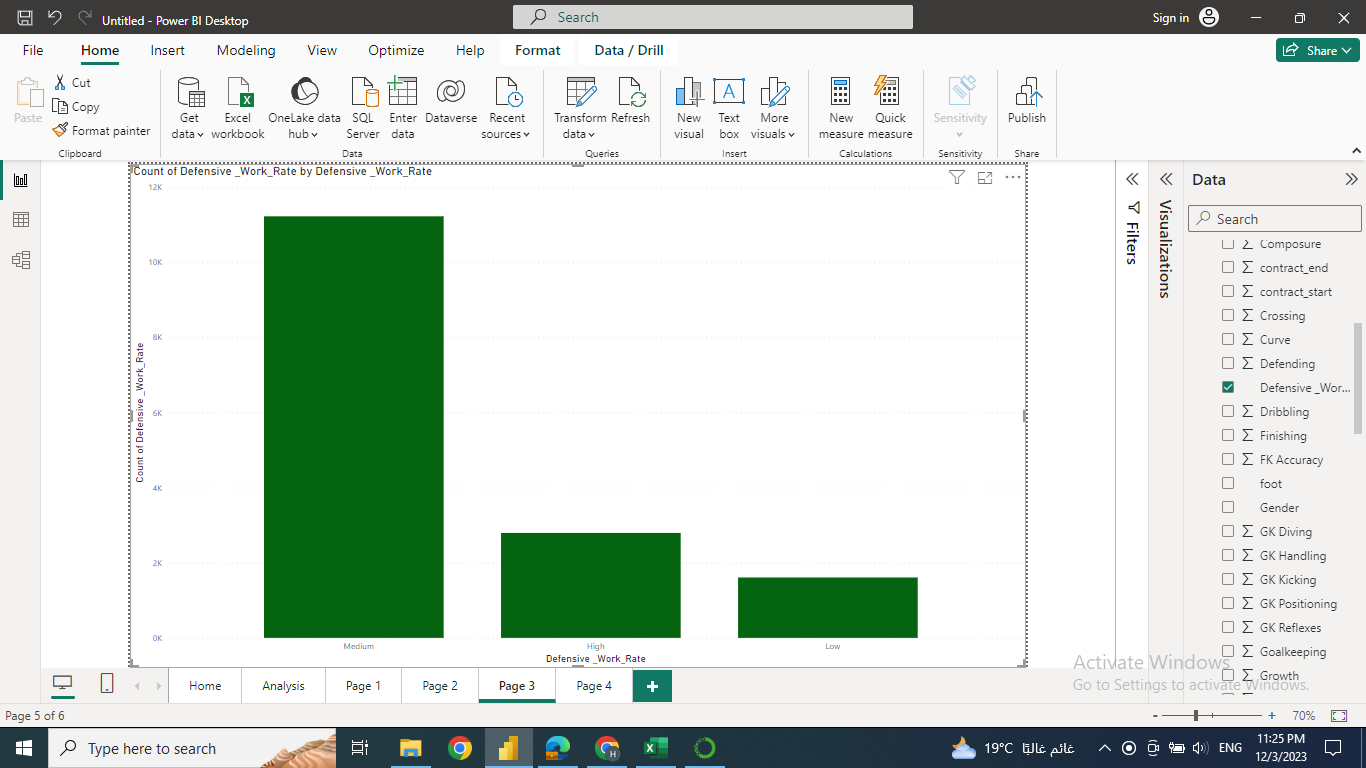
work rate analysis



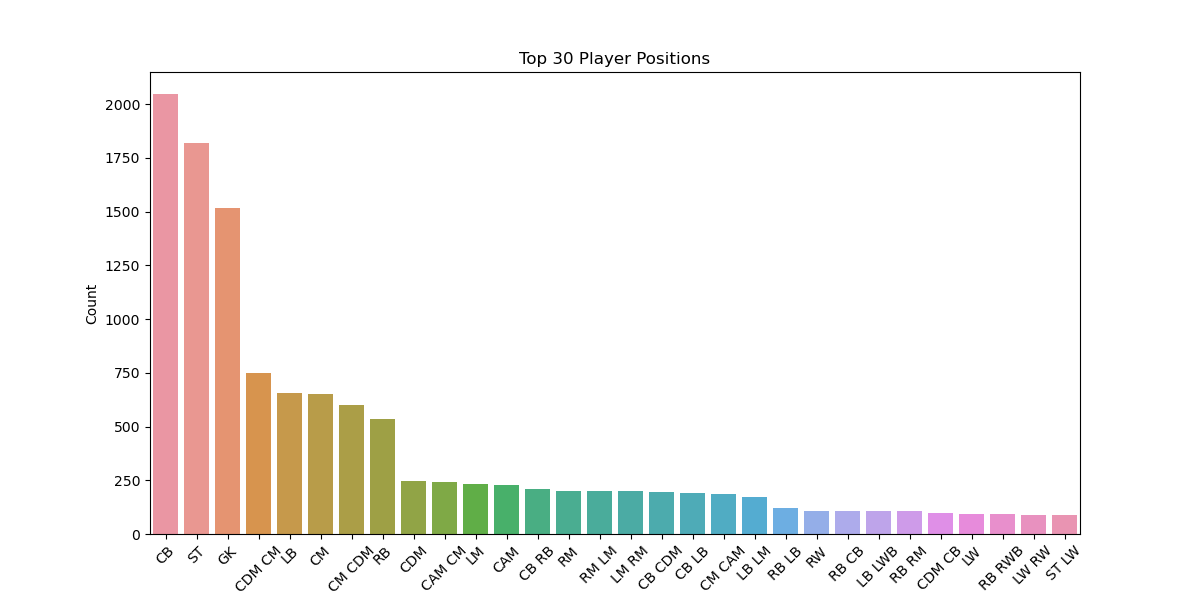




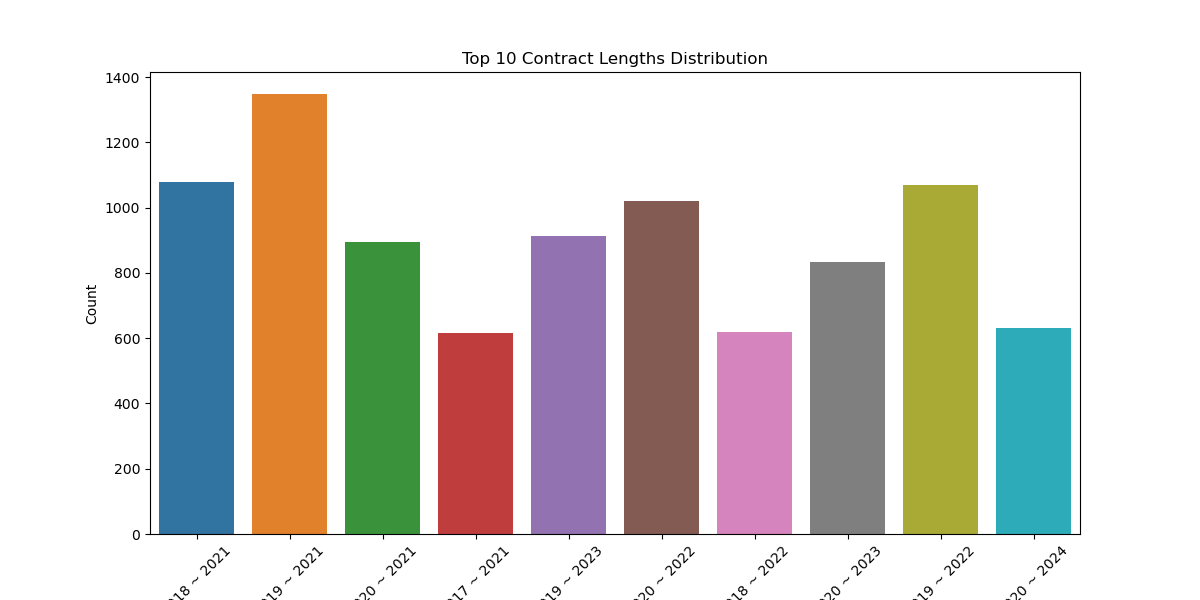




positional analysis



Contract length analysis



### **6. Calculated Measures**

* Implemented calculated measures to derive additional insights, such as average overall rating by nationality.
* Max release clause
* players by nationality

### **7. Tools**

Power bi

Python:<https://github.com/Hagar83/Python/tree/main/FIFA-21>