:[161] In

```
import requests
import datetime
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from pandas import json_normalize
from flatten_json import flatten
pd.set_option('display.max_rows', 500)
pd.set_option('display.max_columns', 500)
pd.set_option('display.width', 1000)
```

:[19] In

```
# Find the relevant links in the Premier League website
player_data ="https://cdnapi.bamboo-video.com/api/football/stats?format=json&iid=5738
player_id ="https://cdnapi.bamboo-video.com/api/football/player?format=json&iid=57388
club_id =  "https://cdnapi.bamboo-video.com/api/football/team?format=json&iid=573881b

# A function that convert the json files that were scraped from the website into data
def js_to_df(URL):
    page = requests.get(URL)
    js =page.json()
    js = js['data']
    ls_keys = list(js.keys())
    df = pd.DataFrame(data=[js[ls_keys[0]].values()])
    for i in range(1,len(ls_keys)):
        df= df.append([js[ls_keys[i]].values()])

    df.columns = js[ls_keys[0]].keys()
    return(df)


```

:[24] In

```
df_player_data = js_to_df(player_data)
df_player_id = js_to_df(player_id)
df_club_id = js_to_df(club_id)
```

:[127] In

```
# select relevat data
df_player_filter = df_player_id.loc[:,['name','stringBirthDate','position','instatId'
df_club_filter=df_club_id[df_club_id['leagueId']==902].loc[:,['name','hebrewName','he
df_player_filter.teamInstatId = df_player_filter.teamInstatId.astype("int")
df_club_filter.teamInstatId = df_club_filter.teamInstatId.astype("int")
```

:[139] In

```
# merge the 3 data frames
df_player_team=df_player_filter.merge(df_club_filter, on='teamInstatId')
df_player_data_comp = df_player_data.merge(df_player_team,on='playerInstatId')
df_player_data_comp.stringBirthDate = pd.to_datetime(df_player_data_comp.stringBirthD
df_player_data_comp['player_age']=(pd.to_datetime("2023") - df_player_data_comp.strir
```

:[172] In

```
# Results
display(df_player_data_comp.iloc[:,5:].head())
```

	seasonName	round	gameInstatId	gameId	teamInstatId_x	teamId	playerInstatId	playerId
0	22/23	0	0	0	1-	1-	26011	23016
1	22/23	0	0	0	1-	1-	16125	41154
2	22/23	0	0	0	1-	1-	28294	45708
3	22/23	0	0	0	1-	1-	19516	45720
4	22/23	0	0	0	1-	1-	68020	46116

:[174] In

```
# Example, plot the Histogram of playes age in this season
plt.hist(df_player_data_comp['player_age'],edgecolor="black",bins=20)
plt.title("Histogram of age in season 22/23")
plt.ylabel("Count")
plt.xlabel("Player Age")
plt.show()
```