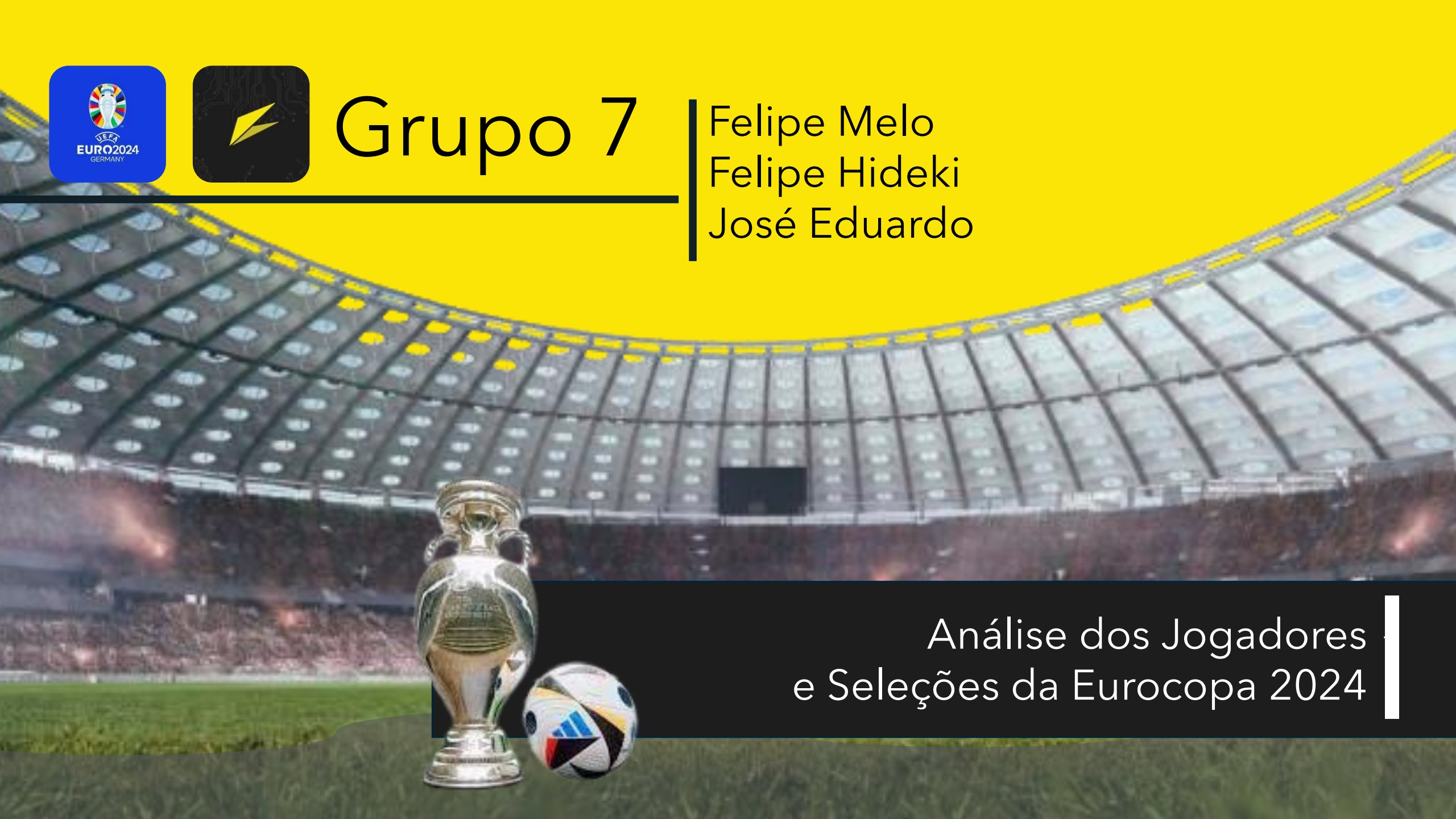
Grupo 7

Felipe Melo
Felipe Hideki
José Eduardo



Análise dos Jogadores e Seleções da Eurocopa 2024

FEA Dev – Análise dos Jogadores e Seleções da Eurocopa– Grupo 7

Introdução

Apresentação Grupo 7

Euro 2024 - Germany

A UEFA Euro 2024, também conhecida como Campeonato Europeu de Futebol, será realizada na Alemanha de 14 de junho a 14 de julho de 2024. Este torneio contará com 24 seleções nacionais competindo em 10 cidades-sede, desde Munique no sul até Hamburgo no norte.

O torneio começa com uma fase de grupos, onde 24 equipes são divididas em seis grupos de quatro. As duas melhores equipes de cada grupo, além das quatro melhores terceiras colocadas, avançam para as oitavas de final. A partir daí, o torneio segue um formato de eliminação direta até a final em Munique no dia 14 de julho

Euro 2024 - Germany

As seleções se classificam para a Eurocopa 2024 através de um processo de qualificação que envolve várias etapas:

Fase de Grupos das Eliminatórias:

Participação: Todas as seleções nacionais membros da UEFA participam das eliminatórias.
Formato: As seleções são divididas em grupos de cinco ou seis times.
Critério de Qualificação: As duas melhores equipes de cada grupo se classificam diretamente para o torneio.
Play-offs:

Participação: As vagas remanescentes são decididas através dos play-offs.
Critério de Participação: Equipes que não se classificaram diretamente através da fase de grupos, mas tiveram bom desempenho na Liga das Nações da UEFA, competem nos play-offs.
Formato: As equipes são divididas em caminhos (A, B, C, D) e competem em jogos eliminatórios.
Equipes Sede:

A Alemanha, como país anfitrião, está automaticamente classificada para a Eurocopa 2024

Apresentação Grupo 7

Euro 2024 - Germany (Cidades sedes)

Apresentação Grupo 7

Dados do Dataframe da Eurocopa, info() e extraindo os dados:

Código:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
import statsmodels.api as sm
import geopandas
nome='euro2024_players.csv'
df=pd.read_csv(nome)
df.info()
```

Output:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 623 entries, 0 to 622
Data columns (total 10 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   Name         623 non-null    object
 1   Position     623 non-null    object
 2   Age          623 non-null    int64
 3   Club         623 non-null    object
 4   Height       623 non-null    int64
 5   Foot         620 non-null    object
 6   Caps         623 non-null    int64
 7   Goals        623 non-null    int64
 8   MarketValue  623 non-null    int64
 9   Country      623 non-null    object
dtypes: int64(5), object(5)
memory usage: 48.8+ KB
```

FEA Dev – Análise dos Jogadores e Seleções da Eurocopa– Grupo 7

Contas Básicas

Valor de mercado das seleções

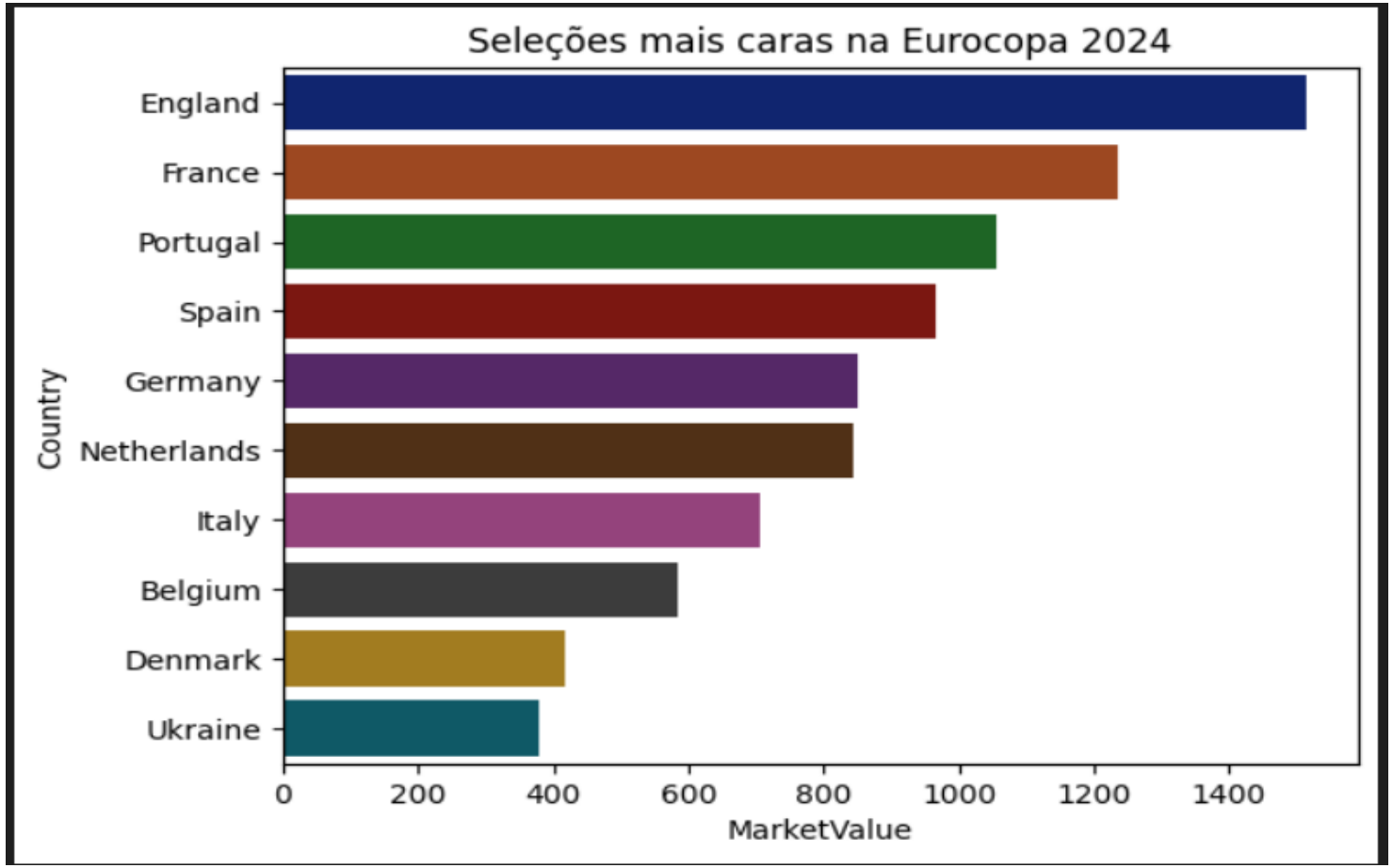
Código:

```
a = df.groupby('Country').agg({'MarketValue': sum})
a['MarketValue']=(a['MarketValue']/1000000).round(0)
a.sort_values(by='MarketValue',ascending=False).head(10).reset_index()

ax = sns.barplot(data=a.sort_values(by='MarketValue',ascending=False).head(10), x='MarketValue', y='Country', palette='dark')
plt.title('Seleções mais caras na Eurocopa 2024')
```

Apresentação Grupo 7

Valor de mercado das seleções

Output:



Apresentação Grupo 7

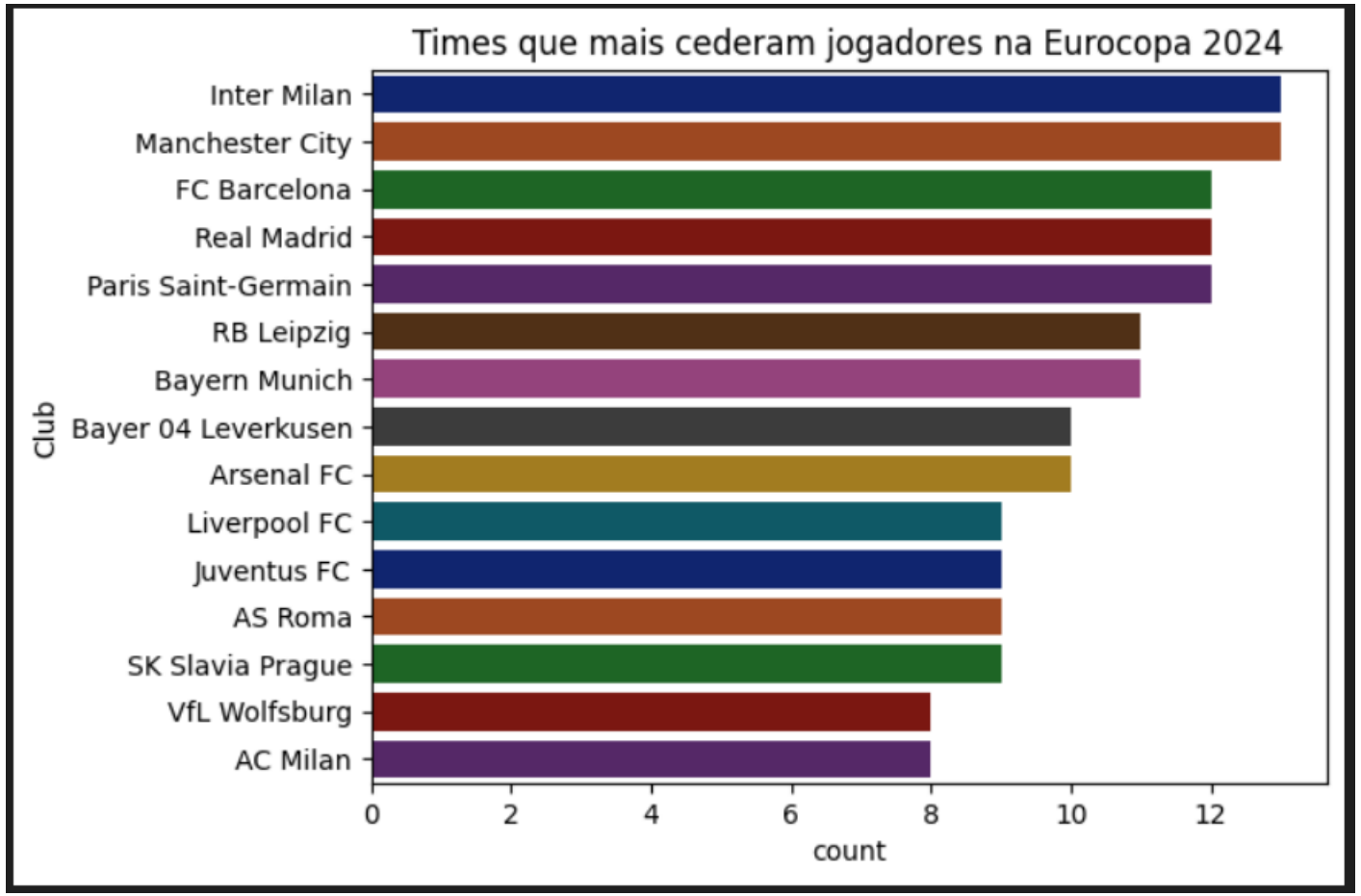
Quais times tiveram mais jogadores cedidos para as seleções

Código:

```
# Plota os 20 times com mais jogadores cedidos para a Eurocopa
contagem_clubes = df['Club'].value_counts()
contagem_clubes = contagem_clubes.to_frame().head(15).reset_index()

ax = sns.barplot(data=contagem_clubes, x='count', y='Club', palette='dark')
plt.title('Times que mais cederam jogadores na Eurocopa 2024')
```

Quais times tiveram mais jogadores cedidos para as seleções

Output:



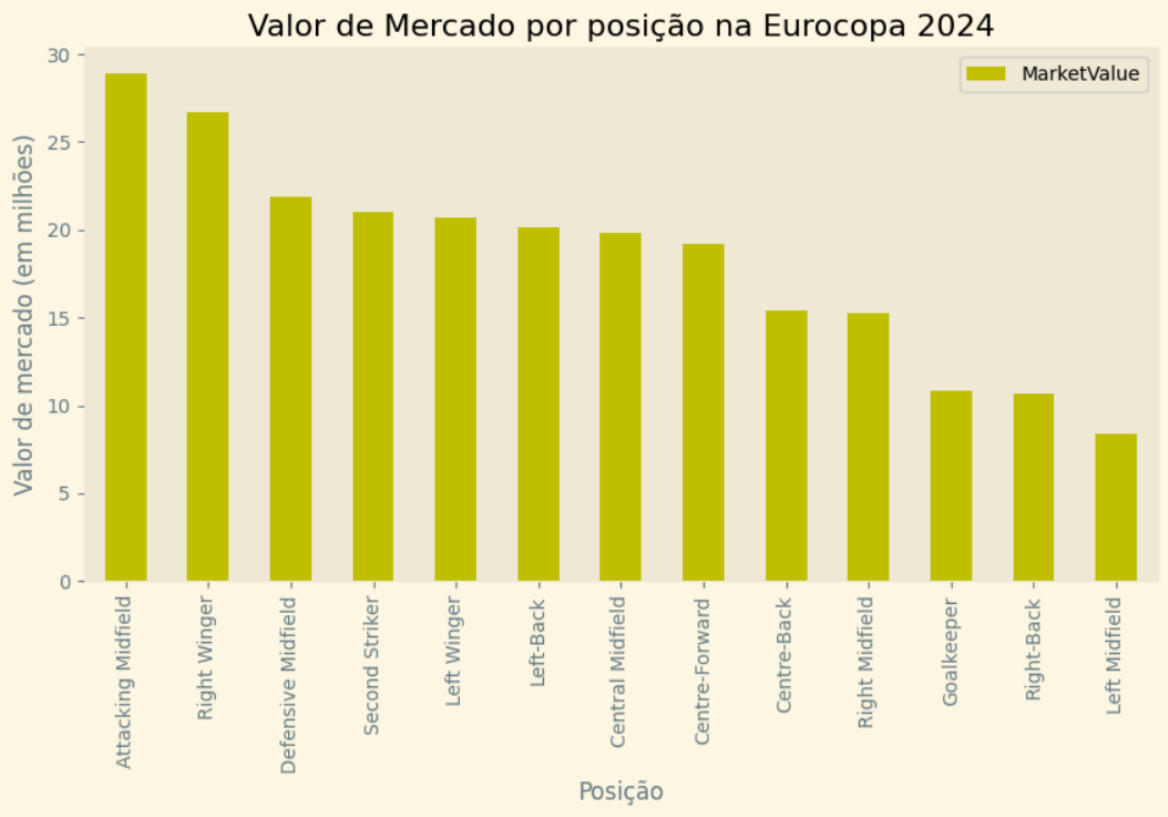
Valor de mercado por posições

Código:

```
plt.style.use('Solarize_Light2')
markValue_por_pos = df.groupby('Position').agg({'MarketValue': np.mean}).round(2).sort_values(by='MarketValue', ascending=False)
markValue_por_pos['MarketValue'] = markValue_por_pos['MarketValue']/1000000
markValue_por_pos.plot(kind='bar', subplots=True,ylabel= 'Valor de mercado (em milhões)', xlabel='Posição', color='y', figsize=(10,5))
plt.grid(False)
plt.title('Valor de Mercado por posição na Eurocopa 2024')
plt.show()
```

Valor de mercado por posições

Output:



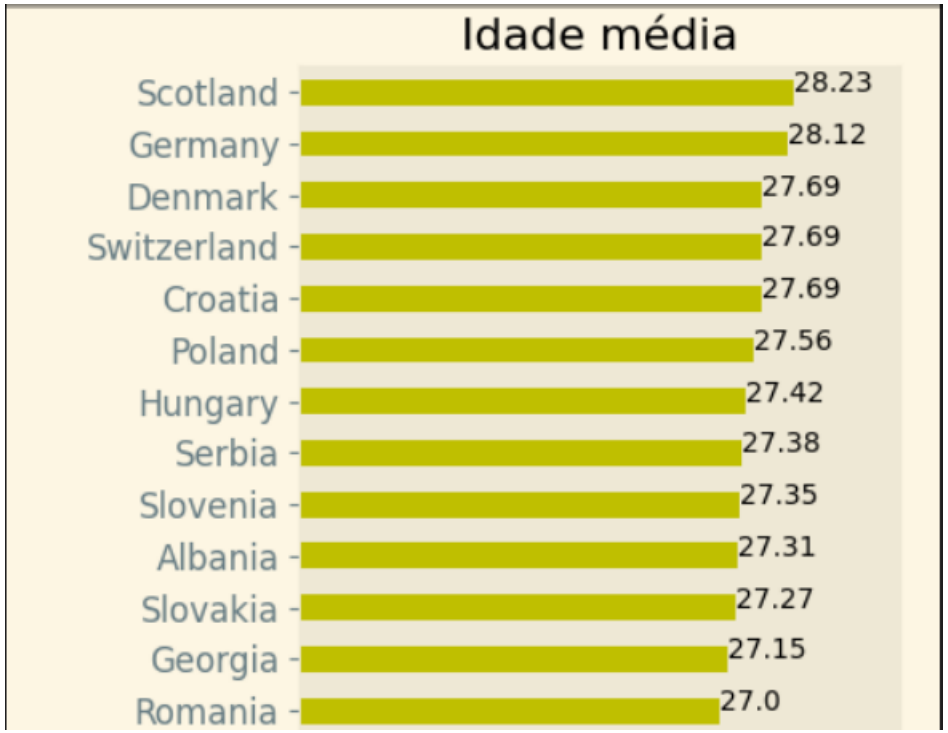
Idade média de cada seleção

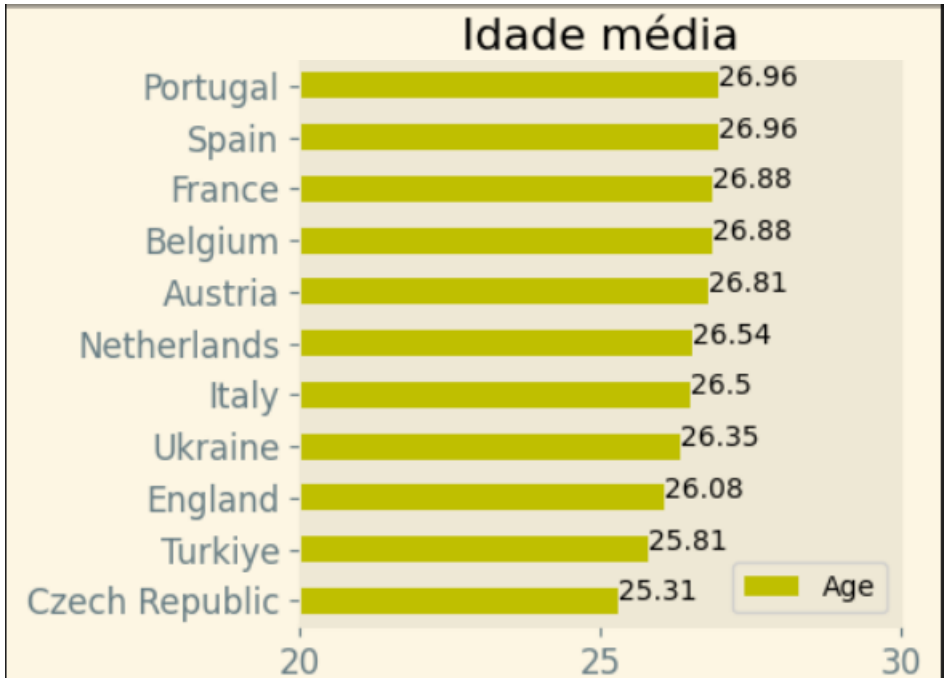
Código:

```
plt.style.use('Solarize_Light2')
markValue_por_pos = df.groupby('Position').agg({'MarketValue': np.mean}).round(2).sort_values(by='MarketValue', ascending=False)
markValue_por_pos['MarketValue'] = markValue_por_pos['MarketValue']/1000000
markValue_por_pos.plot(kind='bar', subplots=True,ylabel= 'Valor de mercado (em milhões)', xlabel='Posição', color='y', figsize=(10,5))
plt.grid(False)
plt.title('Valor de Mercado por posição na Eurocopa 2024')
plt.show()
```

Idade média de cada seleção

Output:





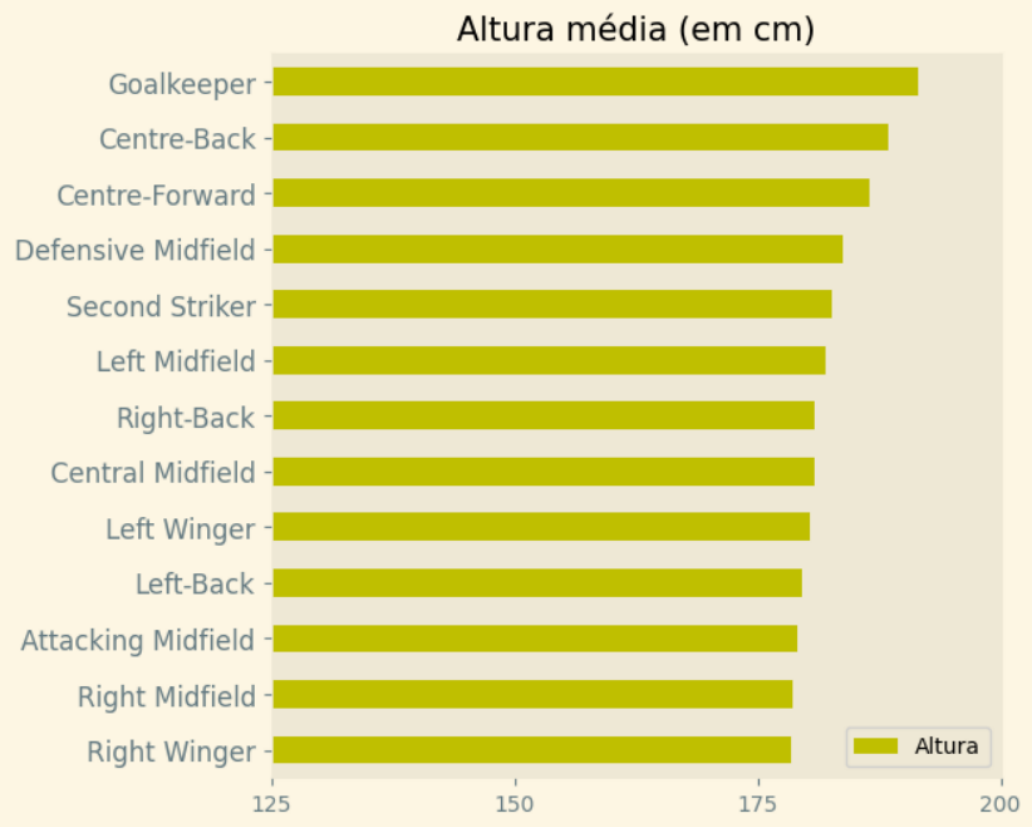
Altura média por posição

Código:

```
altura_por_pos = df.groupby('Position').agg({'Height': np.mean}).round(2).sort_values(by='Height', ascending=True)
altura_por_pos.rename(columns={'Height': 'Altura'},inplace=True)
altura_por_pos.plot(kind='barh', subplots=True, color='y',xticks=[125,150,175,200],xlim=(125,200), figsize=(6,6),legend= 'Altura')
plt.yticks(size = 12)
plt.grid(False)
plt.ylabel('')
plt.title('Altura média (em cm)',fontsize=15)
plt.show()
```

Altura média por posição

Output:



População e Ratio do Market Value

Código:

```
import requests
import pandas as pd

# Fazer uma solicitação à API REST Countries para obter informações sobre todos os países
response = requests.get("https://restcountries.com/v3.1/all")
data = response.json()

# Extrair o nome do país e a população
countries_population = []
for country in data:
    name = country.get('name', {}).get('common', None)
    population = country.get('population', None)
    if name and population is not None:
        countries_population.append({'Country': name, 'População': population})

# Criar um DataFrame do Pandas com os dados extraídos
df_population = pd.DataFrame(countries_population)

df_merged = pd.merge(a,df_population, on="Country", how="inner")
df2=df_merged.sort_values(by='População',ascending=False)
df2['Ratio']=((df2['MarketValue']/df2['População'])*1000000).round(0)
df2['População']=(df2['População']/1000000).round(0)
df2.sort_values(by='Ratio',ascending=False).head(10)
```

✓ 2.3s

Apresentação Grupo 7

População e Ratio do Market Value

Output:

	Country	MarketValue	População	Ratio
12	Portugal	1054.0	10.0	102.0
3	Croatia	328.0	4.0	81.0
4	Denmark	416.0	6.0	71.0
16	Slovenia	140.0	2.0	67.0
2	Belgium	584.0	12.0	51.0
10	Netherlands	845.0	17.0	51.0
14	Serbia	312.0	7.0	45.0
6	Georgia	160.0	4.0	43.0
0	Albania	112.0	3.0	39.0
18	Switzerland	282.0	9.0	33.0

FEA Dev – Análise dos Jogadores e Seleções da Eurocopa– Grupo 7

Regressão

Regressão Linear Múltipla:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + u_i$$

- Y_i é a nossa variável dependente
- β_1 é onde a reta de regressão linear corta o eixo y
- β_n são os coeficientes angulares parciais
- X_{ni} são as variáveis explicativas
- u_i é o erro, ou seja, o quanto nossa equação não conseguiu estimar corretamente

Regressão Linear Múltipla:

$$\text{Valor de Mercado} = \beta_0 + \beta_1(\text{Idade}^2) + \beta_2(\text{Jogos pela Seleção}) + \beta_3(\text{Gols pela Seleção}) + \beta_4(\text{Altura}) + \epsilon$$

Apresentação Grupo 7

Teste de Hipóteses:

Um teste de hipóteses é uma metodologia estatística usada para tomar decisões sobre uma população com base em uma amostra de dados. Envolve formular duas hipóteses:

- **Hipótese Nula (H_0):** Afirma que não há efeito ou diferença significativa.
- **Hipótese Alternativa (H_1):** Afirma que há um efeito ou diferença significativa.

P-Valor: O p-valor é a probabilidade de obter os resultados observados, ou mais extremos, assumindo que a hipótese nula é verdadeira. Ele ajuda a determinar a significância dos resultados:

- **P-valor ≤ α (nível de significância, geralmente 0,05):** Rejeitamos a hipótese nula, indicando que há evidências suficientes para apoiar a hipótese alternativa.
- **P-valor > α:** Não rejeitamos a hipótese nula, indicando que não há evidências suficientes para apoiar a hipótese alternativa.

Apresentação Grupo 7

Regressão Linear Multipla

Código:

```
# Set quais são os Y e Xs regressão
y = df["MarketValue"]

df = df.assign(Squared_Age = lambda x: x.Age ** 2) # Expressão Lambda para criar uma coluna com idade ao quadrado

x = df[['Squared_Age', 'Caps', 'Goals', 'Height']]

x = sm.add_constant(x)

resultados = sm.OLS(y, x).fit()

print(resultados.summary()) # Print os resultados alcançados
```

Apresentação Grupo 7

Resultados

Output:

OLS Regression Results

Dep. Variable:	MarketValue	R-squared:	0.192
Model:	OLS	Adj. R-squared:	0.187
Method:	Least Squares	F-statistic:	36.78
Date:	Fri, 05 Jul 2024	Prob (F-statistic):	1.34e-27
Time:	20:13:30	Log-Likelihood:	-11411.
No. Observations:	623	AIC:	2.283e+04
Df Residuals:	618	BIC:	2.285e+04
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	5.923e+07	2.47e+07	2.395	0.017	1.07e+07	1.08e+08
Squared_Age	-5.91e+04	5239.995	-11.279	0.000	-6.94e+04	-4.88e+04
Caps	2.389e+05	4.75e+04	5.033	0.000	1.46e+05	3.32e+05
Goals	2.258e+05	1.14e+05	1.973	0.049	1023.146	4.51e+05
Height	-2.6e+04	1.36e+05	-0.192	0.848	-2.92e+05	2.4e+05

Omnibus:	322.396	Durbin-Watson:	0.984
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2296.356
Skew:	2.218	Prob(JB):	0.00
Kurtosis:	11.294	Cond. No.	2.27e+04

Apresentação Grupo 7

Gráficos de Regressão

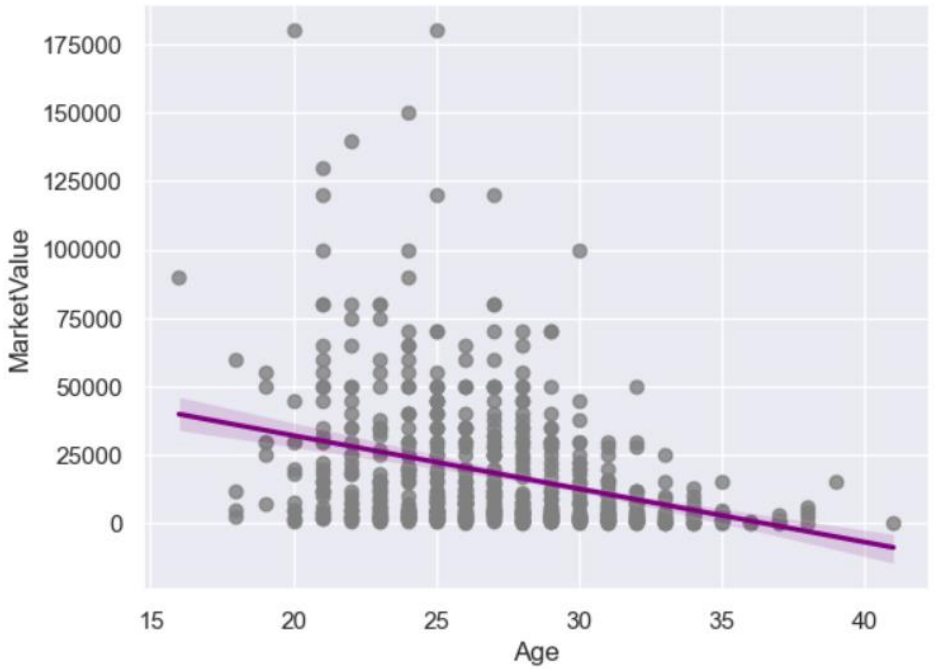
Código:

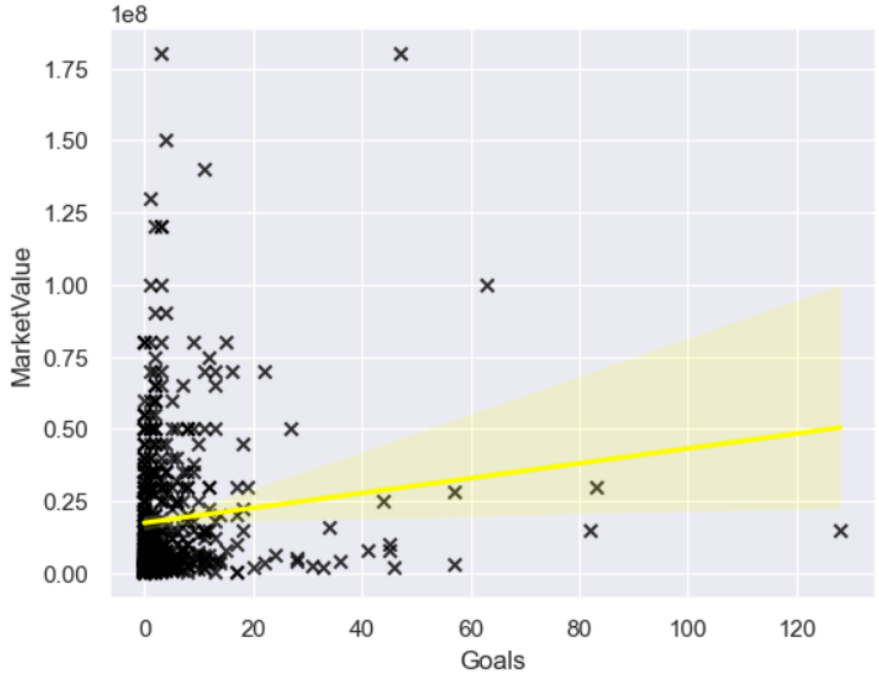
```
sns.set_theme()
sns.regplot(data=df, x='Age', y=df['MarketValue'].div(1000), fit_reg=True, marker='o', scatter_kws=dict(color='gray'), line_kws=dict(color="purple"))
```

```
sns.regplot(data=df, x='Goals', y='MarketValue', fit_reg=True, marker='x', scatter_kws=dict(color='Black'), line_kws=dict(color="Yellow"))
```

Idade média de cada seleção

Output:





FEA Dev – Análise dos Jogadores e Seleções da Eurocopa– Grupo 7

Convocação e Escalação

Convocação de cada país

Código:

```python
def retorna_convocação(pais):
    """
    Retorna a lista de convocação do país selecionado

    :param pais: o país que usuário deseja ver a convocação
    :return: Uma lista de tupla com as informações nome do jogador, posição e time que ele joga
    """
    if pais in df['Country'].values:
        dados = df.query(f"Country in '{pais}'") # Filtra os dados apenas para o pais selecionado

        lista_escalacao = [(jogador, posicao, time) for jogador,posicao,time in zip(dados['Name'], dados['Position'], dados['Club'])]

        return lista_escalacao
    else:
        return print(f'O país {pais} não participa da Eurocopa 2024')

retorna_convocação('France')
```

Convocação de cada país

Output:

```
[('Mike Maignan', 'Goalkeeper', 'AC Milan'),
 ('Brice Samba', 'Goalkeeper', 'RC Lens'),
 ('Alphonse Areola', 'Goalkeeper', 'West Ham United'),
 ('William Saliba', 'Centre-Back', 'Arsenal FC'),
 ('Benjamin Pavard', 'Centre-Back', 'Inter Milan'),
 ('Dayot Upamecano', 'Centre-Back', 'Bayern Munich'),
 ('Jules Koundé', 'Centre-Back', 'FC Barcelona'),
 ('Ibrahima Konaté', 'Centre-Back', 'Liverpool FC'),
 ('Theo Hernández', 'Left-Back', 'AC Milan'),
 ('Ferland Mendy', 'Left-Back', 'Real Madrid'),
 ('Jonathan Clauss', 'Right-Back', 'Olympique Marseille'),
 ('Aurélien Tchouaméni', 'Defensive Midfield', 'Real Madrid'),
 ("N'Golo Kanté", 'Defensive Midfield', 'Al-Ittihad Club'),
 ('Eduardo Camavinga', 'Central Midfield', 'Real Madrid'),
 ('Warren Zaïre-Emery', 'Central Midfield', 'Paris Saint-Germain'),
 ('Adrien Rabiot', 'Central Midfield', 'Juventus FC'),
 ('Youssouf Fofana', 'Central Midfield', 'AS Monaco'),
 ('Kingsley Coman', 'Left Winger', 'Bayern Munich'),
 ('Bradley Barcola', 'Left Winger', 'Paris Saint-Germain'),
 ('Ousmane Dembélé', 'Right Winger', 'Paris Saint-Germain'),
 ('Kylian Mbappé', 'Centre-Forward', 'Paris Saint-Germain'),
 ('Marcus Thuram', 'Centre-Forward', 'Inter Milan'),
 ('Randal Kolo Muani', 'Centre-Forward', 'Paris Saint-Germain'),
 ('Antoine Griezmann', 'Centre-Forward', 'Atlético de Madrid'),
 ('Olivier Giroud', 'Centre-Forward', 'AC Milan')]
```

Escalação de cada país

Código:

```python
import pandas as pd
import matplotlib.pyplot as plt

# Supondo que o seu DataFrame tenha as colunas 'Nome', 'País', 'Posição', 'Valor de Mercado'
# Certifique-se de que df está carregado corretamente antes de usar
# df = pd.read_csv('seu_arquivo.csv')  # Exemplo de como carregar um DataFrame

# Lista de posições necessárias para a formação 4-3-3
positions_4_3_3 = ['Goalkeeper']
position_2cb= ['Centre-Back','Central Midfield']
position_else=['Right-Back', 'Left-Back', 'Defensive Midfield', 'Left Winger', 'Right Winger', 'Centre-Forward']

# Função para obter os jogadores para a formação 4-3-3
def get_best_players(df, country):
    selected_players = []
    for position in positions_4_3_3:
        players_in_position = df[(df['Country'] == country) & (df['Position'] == position)].nlargest(1, 'MarketValue')
        selected_players.append(players_in_position)
    for position in position_2cb:
        players_in_position = df[(df['Country'] == country) & (df['Position'] == position)].nlargest(2, 'MarketValue')
        selected_players.append(players_in_position)
    for position in position_else:
        players_in_position = df[(df['Country'] == country) & (df['Position'] == position)].nlargest(1, 'MarketValue')
        selected_players.append(players_in_position)
    return pd.concat(selected_players)
```

Escalação de cada país

Código:

```
def escalação(country):
    best_players = get_best_players(df, country)

    if best_players.empty:
        print(f"Não há jogadores disponíveis para o país selecionado: {country}")
    else:

        # Visualização da formação 4-3-3
        fig, ax = plt.subplots()
        ax.set_xlim(0, 10)
        ax.set_ylim(0, 10)
        ax.axis('off')

        positions_coordinates = {
            'Goalkeeper': (5, 1),
            'Centre-Back1': (3, 3),
            'Centre-Back2':(7, 3),
            'Left-Back': (1, 4),
            'Right-Back': (9, 4),
            'Defensive Midfield': (5, 5),
            'Central Midfield1': (3, 6),
            'Central Midfield2': (7, 6),
            'Left Winger': (2, 9),
            'Centre-Forward': (5, 8),
            'Right Winger': (8, 9)
```

Escalação de cada país

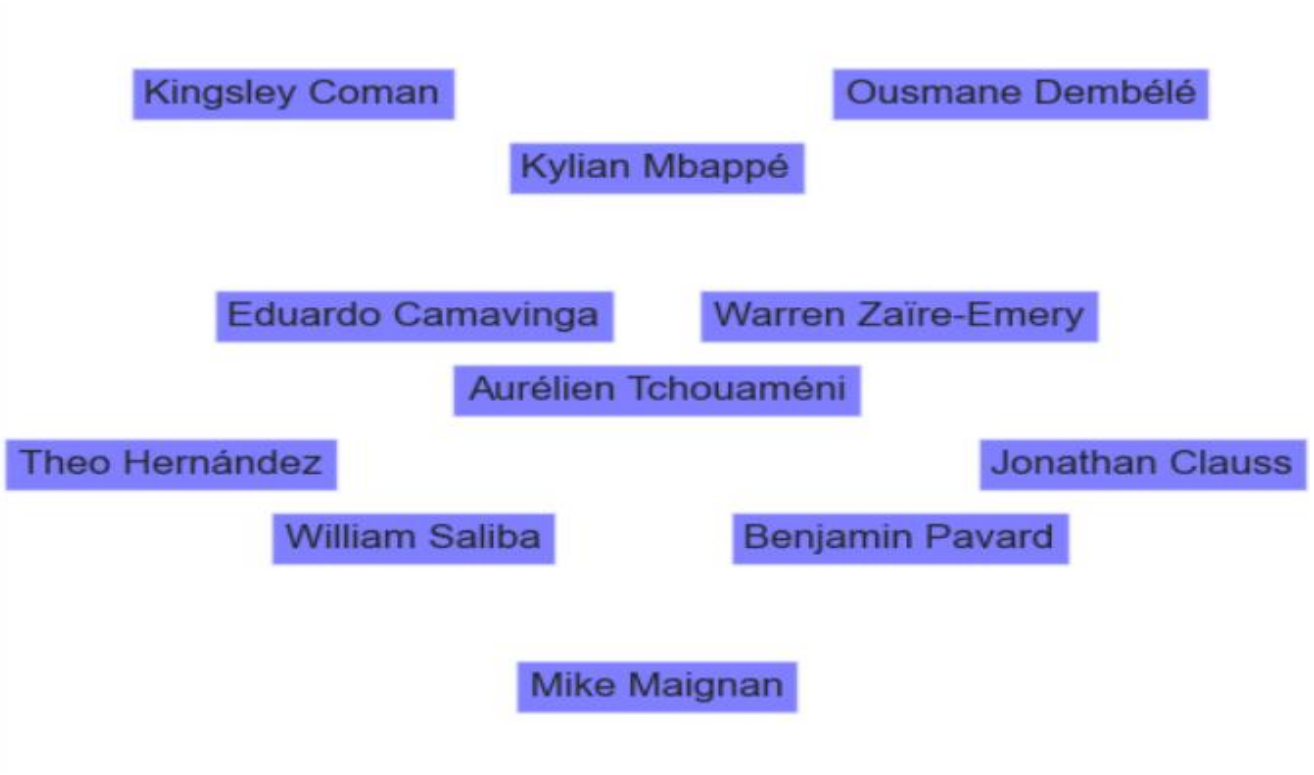
Código:

```
# Ajustar as posições para corresponder ao número de jogadores
best_players['PositionAdjusted'] = best_players['Position']
if 'Centre-Back' in best_players['Position'].values:
    indices = best_players[best_players['Position'] == 'Centre-Back'].index
    if len(indices) > 1:
        best_players.at[indices[0], 'PositionAdjusted'] = 'Centre-Back1'
        best_players.at[indices[1], 'PositionAdjusted'] = 'Centre-Back2'
if 'Central Midfield' in best_players['Position'].values:
    indices = best_players[best_players['Position'] == 'Central Midfield'].index
    if len(indices) > 1:
        best_players.at[indices[0], 'PositionAdjusted'] = 'Central Midfield1'
        best_players.at[indices[1], 'PositionAdjusted'] = 'Central Midfield2'

for _, player in best_players.iterrows():
    posicao = player['PositionAdjusted']
    nome = player['Name']
    if posicao in positions_coordinates:
        coord = positions_coordinates[posicao]
        ax.text(coord[0], coord[1], nome, ha='center', va='center', bbox=dict(facecolor='blue', alpha=0.5))

plt.show()
```

Apresentação Grupo 7

Escalação de cada país

Output:



FEA Dev – Análise dos Jogadores e Seleções da Eurocopa– Grupo 7

Mapa Interativo

Mapa europeu por valor de mercado

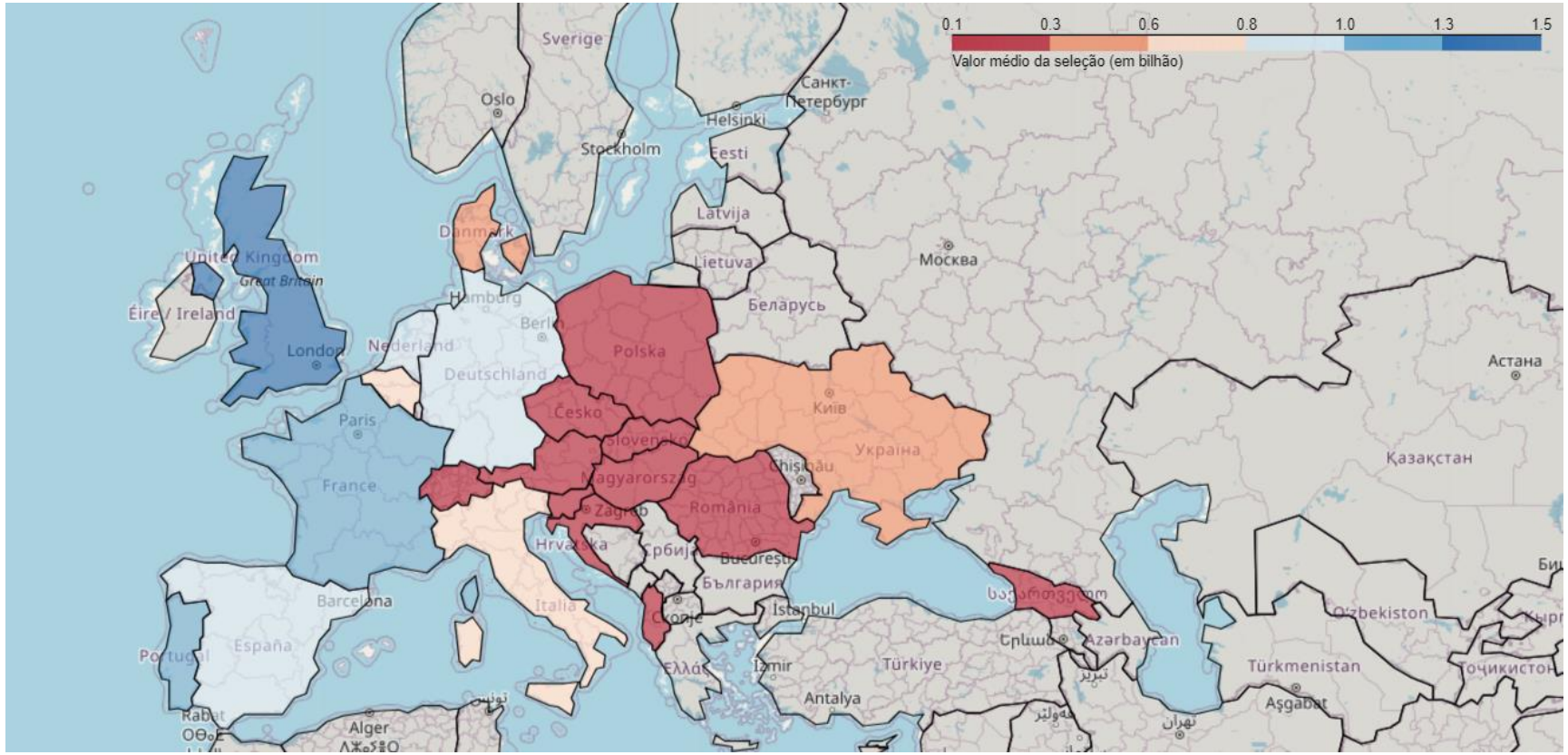
Código:

```
import folium
import requests
world = requests.get(
    "https://raw.githubusercontent.com/python-visualization/folium/main/examples/data/world-countries.json"
).json()
m = folium.Map([53, 9], zoom_start=4)

folium.Choropleth(
    geo_data=world,
    data=valor_selecao,
    columns=["name","MarketValue"],
    nan_fill_color="black",
    nan_fill_opacity=0.1,
    key_on="feature.properties.name",
    fill_color="RdBu",
    legend_name="Valor médio da seleção (em bilhão)"
).add_to(m)
m
```

Apresentação Grupo 7

Mapa europeu por valor de mercado

Output:



FEA Dev – Análise da Eurocopa, Seleções e Jogadores

Grupo 7

Q&A