

# Actividad Módulo 18: Visualización de datos Para Analytics

Utilizando el csv fifa\_eda.csv

1. Generación de un wireframe sobre un reporte con 5 elementos:
  - A. Matriz de correlación como heatmap
  - B. Gráfico de correlación que responda cuál es la relación entre la edad y el overall, explicar
  - C. Gráfico de barras por club que indique el número de jugadores
  - D. Gráfico multipanel que indique la relación altura vs skill moves siendo la variable de panel si es zurdo o es diestro
  - E. Gráfico a elección para mostrar a los potenciales cracks

```
In [ ]: import pandas as pd
import plotly.express as px
```

```
In [ ]: df = pd.read_csv('d:/Documentos/DataAnalysis/EBAC/Python/Modulo14/fifa_eda.csv')
df.sample(5)
```

```
Out[ ]:
```

	ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Preferred
<b>1531</b>	229092	A. Umar	23	Nigeria	76	81	Çaykur Rizespor	10500.0	13.0	Ri
<b>4512</b>	166838	M. Pugh	31	England	71	71	Bournemouth	2100.0	35.0	Ri
<b>17625</b>	227975	H. Fosssdal	19	Norway	53	67	FK Haugesund	100.0	1.0	Ri
<b>4431</b>	176790	S. Aigner	30	Germany	71	71	KFC Uerdingen 05	2200.0	6.0	Ri
<b>8610</b>	233436	A. Mbekile	31	South Africa	67	67	Orlando Pirates	500.0	1.0	Ri

```
In [ ]: # 1 Matriz de correlación como heatmap
fig = px.imshow(df.corr(numeric_only=True))
fig.show()
```

Encontramos unas correlaciones interesantes:

- Los jugadores con mejor Overall, Potential y Value tienden a tener mejor paga
- Los jugadores con mejor paga tienen un Release Clause mas alto, por ende tambien tiene correlación con Overall, Potential y Value
- A mayor altura, mayor peso
- A mayor altura, menor skill moves

```
In [ ]: # 2 Gráfico de correlación y explicacione entre la relación edad y overall
fig = px.scatter(df, x='Age', y='Overall')
fig.show()
```

En el scatterplot se puede ver como los jugadores más jóvenes tienden a tener menos calificación de overall que los jugadores mayores a 25, esto se puede deber a que los jugadores van adquiriendo experiencia conforme va progresando su carrera deportiva

```
In [ ]: # Gráfico de barras en donde muestre el número de jugadores por club
df_players_by_club = df.groupby('Club').size().rename('players').reset_index()
fig = px.bar(df_players_by_club, x='Club', y='players')
fig.show()
```

```
In [ ]: # Gráfico multipanel que indique la relación altura vs skill moves siendo la variab
fig = px.scatter(df, x='Height', y='Skill Moves', facet_col='Preferred Foot')
fig.show()
```

Los jugadores más altos suelen tener menor nivel de Skill moves independientemente de su pie preferido

```
In [ ]: # Gráfico para mostrar a los potenciales cracks
fig = px.box(df.sort_values('Position'), x='Position', y='Potential', title='Gráfico
fig.show()
```

Los posibles cracks serían los outliers para cada posición de jugadores dentro de la base de datos