

### 3.Pandas2SQL\_futebol

March 4, 2025

```
[1]: # Importar bibliotecas
from pyspark.sql import SparkSession

# Criar sessão do Spark
spark = SparkSession.builder \
    .appName("Futebol_Analysis") \
    .getOrCreate()

# Criar dados dos jogadores
dados = [
    (1, 'Cristiano Ronaldo', 'Al Nassr', 'Avançado', 38, 20, 5),
    (2, 'Lionel Messi', 'Inter Miami', 'Avançado', 36, 18, 12),
    (3, 'Neymar', 'Al Hilal', 'Avançado', 31, 10, 8),
    (4, 'Kevin De Bruyne', 'Manchester City', 'Médio', 32, 5, 15),
    (5, 'Virgil van Dijk', 'Liverpool', 'Defesa', 33, 2, 1)
]

# Definir os nomes das colunas
colunas = ["id_jogador", "nome", "clube", "posicao", "idade", "golos", "assistencias"]

# Criar DataFrame
df = spark.createDataFrame(dados, colunas)

# Criar uma tabela temporária
df.createOrReplaceTempView("jogadores")

# Perguntas para os alunos:
# 1. Liste todos os jogadores que marcaram mais de 10 golos.
# 2. Ordene os jogadores por número de assistências em ordem decrescente.
# 3. Calcule a média de idade dos jogadores por posição.
# 4. Identifique qual clube tem o maior número total de golos.
# 5. Filtre os jogadores que tenham idade superior a 30 e que tenham feito pelo menos 5 assistências.

# a) Listar todos os jogadores que marcaram mais de 10 golos
spark.sql("""
```

```

SELECT * FROM jogadores
WHERE golos > 10
""").show()

# b) Ordenar os jogadores por número de assistências em ordem decrescente
spark.sql("""
SELECT * FROM jogadores
ORDER BY assistencias DESC
""").show()

# c) Calcular a média de idade dos jogadores por posição
spark.sql("""
SELECT posicao, AVG(idade) AS media_idade
FROM jogadores
GROUP BY posicao
""").show()

# d) Identificar qual clube tem o maior número total de golos
spark.sql("""
SELECT clube, SUM(golos) AS total_golos
FROM jogadores
GROUP BY clube
ORDER BY total_golos DESC
LIMIT 1
""").show()

# e) Filtrar os jogadores que tenham idade superior a 30 e que tenham feito
↳ pelo menos 5 assistências
spark.sql("""
SELECT * FROM jogadores
WHERE idade > 30 AND assistencias >= 5
""").show()

```

id_jogador	nome	clube	posicao	idade	golos	assistencias
1	Cristiano Ronaldo	Al Nassr	Avançado	38	20	5
2	Lionel Messi	Inter Miami	Avançado	36	18	12

id_jogador	nome	clube	posicao	idade	golos	assistencias
4	Kevin De Bruyne	Manchester City	Médio	32	5	15
2	Lionel Messi	Inter Miami	Avançado	36	18	12
3	Neymar	Al Hilal	Avançado	31	10	8
1	Cristiano Ronaldo	Al Nassr	Avançado	38	20	5

5	Virgil van Dijk	Liverpool	Defesa	33	2	1
---	-----------------	-----------	--------	----	---	---

posicao	media_idade
Avançado	35.0
Médio	32.0
Defesa	33.0

clube	total_golos
Al Nassr	20

id_jogador	nome	clube	posicao	idade	golos	assistencias
1	Cristiano Ronaldo	Al Nassr	Avançado	38	20	5
2	Lionel Messi	Inter Miami	Avançado	36	18	12
3	Neymar	Al Hilal	Avançado	31	10	8
4	Kevin De Bruyne	Manchester City	Médio	32	5	15

[ ]: