

▼ Exercício avaliativo

```
1 import io
2 import sqlite3
3 import pandas as pd
```

Neste exercício, vamos trabalhar com dados de uma empresa fictícia de consultoria denominada IBD.

Considere o seguinte esquema relacional:

Cliente (<u>CodCli</u> , NomeCli, UFcli) Projeto (<u>Sigla</u> , NomePro, DataInicio, CodCli) Consultor (<u>CPF</u> , NomeCon, DataNascCon, UFNascCon) Equipe (<u>Sigla</u> , CPF) Viagem (<u>CPF</u> , <u>Sigla</u> , <u>DataViagem</u> , Origem, Destino) Local (<u>NomeLoc</u> , UFLoc)	Projeto.CodCli → Cliente.CodCli Equipe.Sigla → Projeto.Sigla Equipe.CPF → Consultor.CPF Viagem.CPF → Consultor.CPF Viagem.Sigla → Projeto.Sigla Viagem.Origem → Local.NomeLoc Viagem.Destino → Local.NomeLoc
---	--

assumindo as seguintes premissas:

- Nenhum atributo admite valor nulo.
- Apenas as seguintes chaves primárias são numéricas: Cliente.CodCli, Consultor.CPF

▼ Download dos dados da empresa

```
1 !wget --quiet -P /tmp/ https://raw.githubusercontent.com/filipefbn/jupyter\_intro\_class\_
```

▼ Carregando o banco

```
1 conn = sqlite3.connect('/tmp/consult.db')
2 cursor = conn.cursor()
3
4 f = io.open('/tmp/exercicio.sql', 'r', encoding='utf-8')
5 sql = f.read()
6 cursor.executescript(sql)
```

```
-----
OperationalError                                Traceback (most recent call last)
<ipython-input-115-ad3cf5f366cd> in <module>()
      4 f = io.open('/tmp/exercicio.sql', 'r', encoding='utf-8')
      5 sql = f.read()
----> 6 cursor.executescript(sql)
```

▼ Consultas

Responda as perguntas abaixo através de consultas SQL no banco de dados carregado. Caso tenha alguma dúvida sobre `sqlite3` e `pandas`, revise as **seções 2 e 3** do [Tutorial: SQL, Google Colab e Jupyter Notebooks](#).

O formato da célula de resposta pode ser visto na resposta da consulta exemplo 0.

Você pode usar mais de uma célula para responder, desde que a última célula antes da próxima consulta contenha a resposta final no formato mostrado no exemplo 0.

▼ 0. [exemplo/não alterar] Qual o nome e o número de viagens dos 5 consultores que mais viajaram a trabalho pela empresa IBD?

```
1 query = """
2 SELECT
3     C.NomeCon as Nome, COUNT(C.CPF) as NViagens
4 FROM
5     Consultor as C
6 INNER JOIN Viagem as V
7     ON C.CPF = V.CPF
8 GROUP BY
9     C.CPF
10 ORDER BY
11     NViagens DESC
12 LIMIT 5
13 """
14
15 df = pd.read_sql_query(query, conn)
16 df
```

	Nome	NViagens
0	Melissa Carvalho Cunha	4
1	Larissa Fernandes Araujo	3
2	Rodrigo Araujo Cardoso	3
3	Diogo Cavalcanti Rocha	2
4	Julieta Maciel Sousa	1

1. Que consultores (CPF, nome e estado de nascimento) não nasceram em São Paulo (SP)?

```
1 query = ""
2 SELECT
3   C.CPF as CPF, C.NomeCon as Nome, C.UFNascCon as estado_de_nascimento
4 FROM
5   Consultor as C
6 WHERE
7   estado_de_nascimento != 'SP'
8 ""
9
10 df = pd.read_sql_query(query, conn)
11 df
```



	CPF	Nome	estado_de_nascimento
0	13414263351	Samuel Araujo Azevedo	PR
1	71960563530	Diogo Cavalcanti Rocha	PB
2	3074973994	Julieta Maciel Sousa	RJ
3	29762035631	Caio Dias Ribeiro	MG
4	17787063122	Larissa Fernandes Araujo	MG
5	31266012516	Leonor Goncalves Rocha	MG
6	85726311353	Brenda Oliveira Goncalves	TO

2. Que projetos (sigla e nome) não são do cliente "Facebook"?

```
1 query = ""
2 SELECT
3   P.Sigla as Sigla, P.NomePro as Nome
4 FROM
5   Projeto as P
6 INNER JOIN Cliente as C
7   ON P.CodCli = C.CodCli
8 WHERE
9   C.NomeCli != 'Facebook'
10 ""
11
12 df = pd.read_sql_query(query, conn)
13 df
```

	Sigla	Nome
0	NMT	Neural Machine Translation (Chinês para Portug...
1	Search	Otimização Sistema de Busca
2	Spam	Detecção de Spam (Gmail)
3	Fake	Detecção de Fake News
4	SDC	Self Driving Car
-	-	-

3. Para que UFs viajaram os consultores que participam do projeto de sigla "RecSys"?

```

1 query = ""
2 SELECT DISTINCT
3   V.Destino as Destino, L.UFLoc as uf
4 FROM
5   Consultor as C, Viagem as V, Local as L
6 WHERE
7   C.CPF = V.CPF AND V.Destino = L.NomeLoc AND V.Sigla = 'RecSys'
8 ""
9
10 df = pd.read_sql_query(query, conn)
11 df

```

	Destino	uf
0	IBD-MG	MG
1	iFood	SP

4. Que projetos (sigla e nome) não têm equipe?

```

1 query = ""
2 SELECT DISTINCT
3   P.Sigla as Sigla, P.NomePro as Nome
4 FROM
5   Projeto as P
6 LEFT JOIN Equipe as E
7   ON P.Sigla = E.Sigla
8 WHERE
9   E.CPF IS NULL
10 ""
11
12 df = pd.read_sql_query(query, conn)
13 df

```

	Sigla	Nome
0	Fake	Detecção de Fake News
1	SDC	Self Driving Car

5. Que projetos (sigla e nome) não têm em sua equipe consultores nascidos antes de 1987?

```

1 query = """
2 SELECT DISTINCT
3   P.Sigla, P.NomePro
4 FROM
5   Projeto as P NATURAL JOIN Equipe as E NATURAL JOIN Consultor as C
6 WHERE
7   P.Sigla NOT IN (SELECT P.Sigla
8                   FROM Projeto as P NATURAL JOIN Equipe as E NATURAL JOIN Consultor as
9                   WHERE C.DaTaNascCon < 1987)
10
11
12 """
13
14
15 df = pd.read_sql_query(query, conn)
16 df

```

	SIGLA	NOMEPRO
0	NMT	Neural Machine Translation (Chinês para Portug...
1	Fraud	Detecção de Fraude
2	Spam	Detecção de Spam (Gmail)
3	VR	Hardware para Oculus VR

6. Que consultores (CPF e nome) viajaram por seu estado natal (origem ou destino)?

```

1 query = """
2 SELECT DISTINCT *
3 FROM Consultor as C INNER JOIN Viagem as V INNER JOIN Local as L
4   ON C.CPF = V.CPF AND ((V.Destino = L.NomeLoc AND C.UFNascCon = L.UFLoc) OR (V.Origem
5 ORDER BY
6   C.NomeCon
7 """
8
9
10 df = pd.read_sql_query(query, conn)
11 df

```

	CPF	NOMECON	DATANASCCON	UFNASCCON	SIGLA	CPF	DATAVIAGEM
0	29762035631	Caio Dias Ribeiro	1999-07-19 19:27:25	MG	RecSys	29762035631	2021-06-11 19:27:25
1	17787063122	Larissa Fernandes Araujo	1979-07-19 19:27:25	MG	Pay	17787063122	2021-06-29 19:27:25
2	62465487841	Lucas Fernandes Costa	2000-07-19 19:27:25	SP	RecSys	62465487841	2021-07-05 19:27:25
3	44434353608	Melissa Carvalho	1996-07-19 19:27:25	SP	NMT	44434353608	2021-06-04 19:27:25

▼ 7. Para que locais não houve nenhuma viagem?

```

1 query = """
2 SELECT L.NomeLoc
3 FROM Local as L LEFT JOIN Viagem as V ON (L.NomeLoc = V.Destino)
4 WHERE V.Destino IS NULL
5 """
6
7
8 df = pd.read_sql_query(query, conn)
9 df

```

	NOMELOC
0	Amazon
1	IBD-PR
2	Netflix
3	OpenAI
4	Spotify

▼ 8. Em que projetos (sigla e nome) foram realizadas viagens para a California (CA) desde o início de 2019?

```

1 query = """
2 SELECT DISTINCT P.Sigla, P.NomePro
3 FROM Projeto as P INNER JOIN Viagem as V INNER JOIN Local as L
4   ON P.Sigla = V.Sigla AND V.Destino = L.NomeLoc
5 WHERE L.UFLoc = 'CA' AND V.DataViagem >= 2019
6 """
7
8
9 df = pd.read_sql_query(query, conn)

```

10 df

	SIGLA	NOMEPRO
0	WhatsApp	Cliente desktop para o WhatsApp
1	VR	Hardware para Oculus VR

9. Que consultores (CPF e nome) viajaram por um projeto de cuja equipe não fazem parte?

```
1 query = ""
2 SELECT DISTINCT C.NomeCon, E.Sigla
3 FROM Consultor as C INNER JOIN Equipe as E
4   ON C.CPF = E.CPF
5 ""
6
7
8 df = pd.read_sql_query(query, conn)
9 df
```

	NOMECON	SIGLA
0	Melissa Carvalho Cunha	NMT
1	Melissa Carvalho Cunha	Fraud
2	Diogo Cavalcanti Rocha	Fraud
3	Melissa Carvalho Cunha	Spam
4	Diogo Cavalcanti Rocha	Spam
5	Caio Dias Ribeiro	RecSys
6	Julieta Maciel Sousa	RecSys
7	Lucas Fernandes Costa	RecSys
8	Larissa Fernandes Araujo	WhatsApp
9	Leonor Goncalves Rocha	WhatsApp
10	Larissa Fernandes Araujo	Pay
11	Leonor Goncalves Rocha	Pay
12	Rodrigo Araujo Cardoso	VR

```
1 query = ""
2 SELECT C.CPF, C.NomeCon
3 FROM Consultor as C JOIN Viagem as V
4   ON C.CPF = V.CPF
5 WHERE (C.CPF, V.Sigla) NOT IN (SELECT
6                                   CPF, Sigla
7                                   FROM Equipe as E)
```

```

8 """
9
10 df = pd.read_sql_query(query, conn)
11 df

```

	CPF	NOMECON
0	34646193892	Rodrigo Araujo Cardoso
1	44434353608	Melissa Carvalho Cunha

10. Que consultores (CPF e nome) não estão em nenhuma equipe de projeto que tenha iniciado em 2019?

```

1 query = """
2 SELECT DISTINCT C.CPF, C.NomeCOn
3 FROM Consultor as C JOIN Equipe as E JOIN Projeto as P
4   ON C.CPF = E.CPF AND E.Sigla = P.Sigla
5 WHERE P.DataInicio < '2019' OR P.DataInicio > '2019'
6 """
7
8 df = pd.read_sql_query(query, conn)
9 df

```

	CPF	NOMECON
0	44434353608	Melissa Carvalho Cunha
1	71960563530	Diogo Cavalcanti Rocha
2	29762035631	Caio Dias Ribeiro
3	3074973994	Julieta Maciel Sousa
4	62465487841	Lucas Fernandes Costa
5	17787063122	Larissa Fernandes Araujo
6	31266012516	Leonor Goncalves Rocha
7	34646193892	Rodrigo Araujo Cardoso

✓ 0s conclusão: 23:14

