

## ▼ Exercício avaliativo

```
import io
import sqlite3
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, UF Cli) 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

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

## ▼ Carregando o banco

```
conn = sqlite3.connect('/tmp/consult.db')
cursor = conn.cursor()

f = io.open('/tmp/exercicio.sql', 'r', encoding='utf-8')
sql = f.read()
cursor.executescript(sql)
```

```
↳ -----
OperationalError                                Traceback (most recent call last)
<ipython-input-13-ad3cf5f366cd> in <module>()
      4 f = io.open('/tmp/exercicio.sql', 'r', encoding='utf-8')
      5 sql = f.read()
----> 6 cursor.executescript(sql)

OperationalError: table [CLIENTE] already exists
```

SEARCH STACK OVERFLOW

## ▼ 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?

```

query = """
SELECT
    C.NomeCon as Nome, COUNT(C.CPF) as NViagens
FROM
    Consultor as C
INNER JOIN Viagem as V
    ON C.CPF = V.CPF
GROUP BY
    C.CPF
ORDER BY
    NViagens DESC
LIMIT 5
"""

```

```

df = pd.read_sql_query(query, conn)
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)?

```

query = """
SELECT CPF,
        NomeCon,
        UFNascCon
FROM Consultor
WHERE UFNascCon <> 'SP'
"""

```

```
df = pd.read_sql_query(query,conn)
df
```

	CPF	NOMECON	UFNASCCON
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"?

```
query = """
SELECT P.Sigla,
       P.NomePro
FROM Projeto as P, Cliente as C
WHERE P.CodCli = C.CodCli AND C.NomeCli <> 'Facebook'
"""
```

```
df = pd.read_sql_query(query,conn)
df
```

	<b>SIGLA</b>	<b>NOMEPRO</b>
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
5	Pay	Sistema de Pagamento
6	RecSys	Sistema de Recomendação
7	Fraud	Detecção de Fraude

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

```
query = """
SELECT L.UFLoc
FROM Local AS L
WHERE L.NomeLoc IN (
    SELECT V.Destino
    FROM Viagem AS V JOIN Consultor AS C
    ON V.CPF = C.CPF
    WHERE V.Sigla = 'RecSys'
)
"""
```

```
df = pd.read_sql_query(query,conn)
df
```

```

↳
   UFLOC
0      MG
1      SP

```

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

```

query = """
SELECT P.Sigla,
       P.NomePro
FROM Projeto as P LEFT JOIN Equipe as E
ON P.Sigla = E.Sigla
WHERE E.Sigla IS NULL
"""

```

```
df = pd.read_sql_query(query, conn)
df
```

```

↳
   SIGLA      NOMEPRO
0  Search  Otimização Sistema de Busca
1   Fake      Detecção de Fake News
2   SDC      Self Driving Car

```

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

```

query = """
SELECT P.Sigla, P.NomePro
FROM Projeto AS P
WHERE P.Sigla NOT IN(
    SELECT E.Sigla
    FROM Equipe AS E JOIN Consultor AS C
    ON E.CPF = C.CPF
    WHERE C.DataNascCon < 1987
)
"""

```

```

df = pd.read_sql_query(query,conn)
df

```

	<b>SIGLA</b>	<b>NOMEPRO</b>
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
5	VR	Hardware para Oculus VR
6	Fraud	Detecção de Fraude

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

```

query = """
SELECT C.CPF, C.NomeCon
FROM Consultor AS C
WHERE C.UFNascCon IN (
    SELECT L.UFLoc
    FROM Local AS L, Viagem AS V
    WHERE L.NomeLoc = V.Origem AND V.CPF = C.CPF
)
OR
C.UFNascCon IN (
    SELECT L.UFLoc
    FROM Local AS L, Viagem AS V
    WHERE L.NomeLoc = V.Destino AND V.CPF = C.CPF
)
"""

```

```

df = pd.read_sql_query(query,conn)
df

```

	CPF	NOMECON
0	44434353608	Melissa Carvalho Cunha
1	62465487841	Lucas Fernandes Costa
2	29762035631	Caio Dias Ribeiro
3	34646193892	Rodrigo Araujo Cardoso
4	17787063122	Larissa Fernandes Araujo

## ▼ 7. Para que locais não houve nenhuma viagem?

```

query = """
SELECT L.NomeLoc
FROM Local as L LEFT JOIN Viagem as V
ON L.NomeLoc = V.Destino
WHERE V.Destino IS NULL
"""

```



```
df = pd.read_sql_query(query,conn)
df
```

	<b>NOMELOC</b>
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?

```
query = """
SELECT P.Sigla, P.NomePro
FROM Projeto as P
WHERE P.Sigla IN (
    SELECT V.Sigla
    FROM Viagem AS V
    WHERE strftime('%Y', V.DataViagem) >= '2019' AND V.Destino IN(
        SELECT L.NomeLoc
        FROM Local AS L
        WHERE L.UFLoc = 'CA'
    )
)
"""
```

```
df = pd.read_sql_query(query,conn)
df
```

	<b>SIGLA</b>	<b>NOMEPRO</b>
0	VR	Hardware para Oculus VR
1	WhatsApp	Cliente desktop para o WhatsApp

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

```
query = """
SELECT DISTINCT C.CPF, C.NomeCon
FROM Viagem AS V JOIN Consultor AS C
ON v.CPF = C.CPF
WHERE (V.CPF,V.Sigla) NOT IN(
    SELECT E.CPF, E.Sigla
    FROM Equipe AS E JOIN Consultor AS C
    ON E.CPF = C.CPF
)
"""
```

```
df = pd.read_sql_query(query,conn)
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?

```
query = """
SELECT C.CPF, C.NomeCon
FROM Consultor AS C
WHERE C.CPF NOT IN(
    SELECT E.CPF
    FROM Equipe As E JOIN Projeto AS P
    ON E.Sigla = P.Sigla
    WHERE P.DataInicio >= 2019
)
"""
```

```
df = pd.read_sql_query(query,conn)
df
```

	<b>CPF</b>	<b>NOMECON</b>
<b>0</b>	13414263351	Samuel Araujo Azevedo
<b>1</b>	71960563530	Diogo Cavalcanti Rocha
<b>2</b>	85726311353	Brenda Oliveira Goncalves