

▼ 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\_class\_ibd
```

▼ 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)
```

▼ 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
    C.CPF, C.NomeCon, C.DataNascCon
FROM
```

```

Consultor as C
WHERE
  C.UFNascCon != "SP"
"""

```

```

df = pd.read_sql_query(query, conn)
df

```

	CPF	NOMECON	DATANASCCON
0	13414263351	Samuel Araujo Azevedo	1992-07-19 23:41:25
1	71960563530	Diogo Cavalcanti Rocha	1993-07-19 23:41:25
2	3074973994	Julieta Maciel Sousa	1986-07-19 23:41:25
3	29762035631	Caio Dias Ribeiro	1999-07-19 23:41:25
4	17787063122	Larissa Fernandes Araujo	1979-07-19 23:41:25
5	31266012516	Leonor Goncalves Rocha	1981-07-19 23:41:25
6	85726311353	Brenda Oliveira Goncalves	1983-07-19 23:41:25

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

```

query = """
SELECT
  P.Sigla, P.NomePro
FROM
  Projeto as P
INNER JOIN Cliente as C
  ON P.CodCli = C.CodCli
WHERE
  C.CodCli != 4
"""

```

```

df = pd.read_sql_query(query, conn)
df

```

	SIGLA	NOMEPRO
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 DISTINCT
    L.UFLoc
FROM
    Local as L
INNER JOIN Viagem as V
    ON L.NomeLoc = V.Destino
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.CPF 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

```
query = """
SELECT DISTINCT
    P.Sigla, P.NomePro
FROM
    Consultor as C NATURAL JOIN Equipe as E NATURAL JOIN Projeto as P
WHERE
    C.DataNascCon >= 1987
"""
df = pd.read_sql_query(query, conn)
df
```

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

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

```
query = """
SELECT DISTINCT
    C.CPF, C.NomeCon
FROM
    Consultor as C
NATURAL JOIN Viagem as V
JOIN Local as L
WHERE (L.NomeLoc == V.Origem OR L.NomeLoc == V.Destino) AND C.UFNascCon == L.UFLoc
"""
df = pd.read_sql_query(query, conn)
df
```

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

▼ 7. Para que locais não houve nenhuma viagem?

```
query = """
SELECT
    Lo.UFLoc
FROM
    Local as Lo

EXCEPT

SELECT DISTINCT
    L.UFLoc
FROM
    Viagem as V
INNER JOIN Local as L
    ON L.NomeLoc == V.Destino
"""

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

	UFLOC
0	NY
1	PR
2	SF
3	WA

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

```
query = """
SELECT DISTINCT
    P.Sigla, P.NomePro
FROM
    Projeto as P
NATURAL JOIN Viagem as V
NATURAL JOIN Local as L
WHERE L.UFLoc == "CA" AND V.Destino = L.NomeLoc
"""

df = pd.read_sql_query(query, conn)
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?

```
query = """
SELECT
    C.CPF, C.NomeCon
FROM Consultor as C
NATURAL JOIN
(SELECT
    V.Sigla, V.CPF
FROM
    Viagem as V INNER JOIN Consultor as C

EXCEPT

SELECT
    E.Sigla, C.CPF
FROM
    Equipe as E INNER JOIN Consultor as C ON E.CPF = C.CPF
    INNER JOIN Projeto as P ON E.Sigla = P.Sigla)
"""
```

```
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 DISTINCT
    C.CPF, C.NomeCon
FROM
    Consultor as C

EXCEPT

SELECT DISTINCT
    C.CPF, C.NomeCon
FROM
    (Projeto as P INNER JOIN Equipe as E ON P.Sigla = E.Sigla)
    INNER JOIN Consultor as C ON C.CPF = E.CPF
WHERE P.DataInicio > 2019 AND P.DataInicio < 2020
```

"""

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

	CPF	NOMECON
0	3074973994	Julieta Maciel Sousa
1	13414263351	Samuel Araujo Azevedo
2	17787063122	Larissa Fernandes Araujo
3	29762035631	Caio Dias Ribeiro
4	31266012516	Leonor Goncalves Rocha
5	34646193892	Rodrigo Araujo Cardoso
6	62465487841	Lucas Fernandes Costa
7	85726311353	Brenda Oliveira Goncalves