

# Consulta em Banco de Dados com SQL e Python

In [1]: *# Versão Linguagem Python utilizada*

```
from platform import python_version  
python_version()
```

Out[1]: '3.9.13'

In [17]: *# Importando biblioteca SQLite*

```
import sqlite3  
  
sqlite3.sqlite_version
```

Out[17]: '3.39.3'

In [18]: *# Importando biblioteca Pandas*

```
import pandas as pd  
pd.__version__
```

Out[18]: '2.1.4'

In [3]: *# Conectando ao banco de dados*

```
bd = sqlite3.connect('cap12_dsa.db')
```

In [5]: *# Abrindo um cursor para percorrer os dados no banco de dados*

```
cursor = bd.cursor()  
  
# Query SQL para extrair os nomes das colunas no banco de dados  
  
sql_query = """SELECT name FROM sqlite_master WHERE type = 'table';"""  
cursor.execute(sql_query)
```

Out[5]: <sqlite3.Cursor at 0x20a3a8e65e0>

In [6]: *# Visualizando o resultado*

```
print(cursor.fetchall())
```

```
[('tb_vendas_dsa',)]
```

In [15]: *# Verificando todas as linhas da tabela*

```
query1 = 'SELECT * FROM tb_vendas_dsa'
cursor.execute(query1)
```

*# List comprehension para visualizar os nomes das colunas*

```
nomes_colunas = [description[0] for description in cursor.description]
```

```
print(nomes_colunas)
```

*#tabela*

```
['ID_Pedido', 'ID_Cliente', 'Nome_Produto', 'Valor_Unitario', 'Unidades_Vendidas', 'Custo']
```

In [16]: *# Dados da query1*

```
tabela = cursor.fetchall()
```

*# Carregando os dados como dataframe do Pandas*

```
df = pd.DataFrame(tabela, columns = ['ID_Pedido',
                                     'ID_Cliente',
                                     'Nome_Produto',
                                     'Valor_Unitario',
                                     'Unidades_Vendidas',
                                     'Custo'])
```

```
df
```

Out[16]:

	ID_Pedido	ID_Cliente	Nome_Produto	Valor_Unitario	Unidades_Vendidas	Custo
0	1	63	Produto_38	154.03	7	92.42
1	2	49	Produto_8	171.52	5	102.91
2	3	83	Produto_39	28.97	13	17.38
3	4	37	Produto_2	104.55	4	62.73
4	5	19	Produto_1	77.21	19	46.33
...	...	...	...	...	...	...
495	496	27	Produto_12	155.02	1	93.01
496	497	32	Produto_23	71.04	6	42.62
497	498	80	Produto_1	67.83	13	40.70
498	499	13	Produto_50	187.89	16	112.73
499	500	46	Produto_21	82.81	11	49.69

500 rows × 6 columns

In [39]:

```
# Calculando a média de unidades vendidas

query2 = 'SELECT AVG(Unidades_Vendidas) FROM tb_vendas_dsa'
cursor.execute(query2)

# Retornando resultado da query2

media = cursor.fetchall()

print('A média de unidades vendidas é de:', media)
```

A média de unidades vendidas é de: [(10.506,)]

In [40]:

```
# Calculando a média de unidades vendidas por produto

query3 = 'SELECT Nome_Produto, AVG(Unidades_Vendidas) FROM tb_vendas_dsa GROUP BY Nome_Produto'
cursor.execute(query3)

# Retornando resultado da query3
```

```
media2 = cursor.fetchall()

# Carregando os dados como dataframe do Pandas

media2_produtos = pd.DataFrame(media2, columns = ['Nome_Produto',
                                                  'Média de Unidades Vendidas'])

media2_produtos
```

Out[40]:

	Nome_Produto	Média de Unidades Vendidas
0	Produto_1	12.000000
1	Produto_10	9.500000
2	Produto_11	14.181818
3	Produto_12	8.846154
4	Produto_13	6.000000
5	Produto_14	9.166667
6	Produto_15	9.750000
7	Produto_16	8.250000
8	Produto_17	11.714286
9	Produto_18	13.083333
10	Produto_19	9.727273
11	Produto_2	9.250000
12	Produto_20	7.555556
13	Produto_21	10.285714
14	Produto_22	13.687500
15	Produto_23	10.818182
16	Produto_24	12.272727
17	Produto_25	9.538462
18	Produto_26	9.363636
19	Produto_27	11.100000
20	Produto_28	9.000000
21	Produto_29	9.692308
22	Produto_3	8.909091
23	Produto_30	9.875000

	Nome_Produto	Média de Unidades Vendidas
24	Produto_31	7.900000
25	Produto_32	11.923077
26	Produto_33	12.285714
27	Produto_34	8.100000
28	Produto_35	9.000000
29	Produto_36	9.090909
30	Produto_37	11.000000
31	Produto_38	12.800000
32	Produto_39	12.666667
33	Produto_4	11.153846
34	Produto_40	7.250000
35	Produto_41	11.857143
36	Produto_42	10.272727
37	Produto_43	11.000000
38	Produto_44	7.200000
39	Produto_45	8.875000
40	Produto_46	12.142857
41	Produto_47	10.571429
42	Produto_48	14.000000
43	Produto_49	11.875000
44	Produto_5	10.200000
45	Produto_50	10.545455
46	Produto_6	12.000000
47	Produto_7	13.562500

	Nome_Produto	Média de Unidades Vendidas
48	Produto_8	11.071429
49	Produto_9	7.200000

```
In [47]: # Calculando a média de unidades vendidas por produto,
# quando o valor unitário for maior que 199
query4 = """SELECT Nome_Produto,
                AVG(Unidades_Vendidas)
            FROM
                tb_vendas_dsa
            WHERE
                Valor_Unitario > 199
            GROUP BY
                Nome_Produto"""
cursor.execute(query3)

# Retornando resultado da query4

media3 = cursor.fetchall()

# Carregando os dados como dataframe do Pandas

media3_produtos = pd.DataFrame(media3, columns = ['Nome_Produto',
                                                'Média de Un. Vend. acima de 199'])

media3_produtos
```

Out[47]:

	Nome_Produto	Média de Un. Vend. acima de 199
0	Produto_1	12.000000
1	Produto_10	9.500000
2	Produto_11	14.181818
3	Produto_12	8.846154
4	Produto_13	6.000000
5	Produto_14	9.166667
6	Produto_15	9.750000
7	Produto_16	8.250000
8	Produto_17	11.714286
9	Produto_18	13.083333
10	Produto_19	9.727273
11	Produto_2	9.250000
12	Produto_20	7.555556
13	Produto_21	10.285714
14	Produto_22	13.687500
15	Produto_23	10.818182
16	Produto_24	12.272727
17	Produto_25	9.538462
18	Produto_26	9.363636
19	Produto_27	11.100000
20	Produto_28	9.000000
21	Produto_29	9.692308
22	Produto_3	8.909091
23	Produto_30	9.875000



	Nome_Produto	Média de Un. Vend. acima de 199
24	Produto_31	7.900000
25	Produto_32	11.923077
26	Produto_33	12.285714
27	Produto_34	8.100000
28	Produto_35	9.000000
29	Produto_36	9.090909
30	Produto_37	11.000000
31	Produto_38	12.800000
32	Produto_39	12.666667
33	Produto_4	11.153846
34	Produto_40	7.250000
35	Produto_41	11.857143
36	Produto_42	10.272727
37	Produto_43	11.000000
38	Produto_44	7.200000
39	Produto_45	8.875000
40	Produto_46	12.142857
41	Produto_47	10.571429
42	Produto_48	14.000000
43	Produto_49	11.875000
44	Produto_5	10.200000
45	Produto_50	10.545455
46	Produto_6	12.000000
47	Produto_7	13.562500

	Nome_Produto	Média de Un. Vend. acima de 199
48	Produto_8	11.071429
49	Produto_9	7.200000

```
In [34]: # Verificando os Top 10 produtos mais vendidos
```

```
query5 = """SELECT Nome_Produto,  
                COUNT(Unidades_Vendidas)  
            FROM  
                tb_vendas_dsa  
            GROUP BY  
                Nome_Produto  
            ORDER BY  
                COUNT(Unidades_Vendidas) DESC  
            LIMIT 10"""  
  
cursor.execute(query5)  
  
# Retornando resultado da query5  
  
top10 = cursor.fetchall()  
  
# Carregando os dados como dataframe do Pandas  
  
top10_produtos = pd.DataFrame(top10, columns = ['Nome_Produto',  
                                                'Unidades_Vendidas'])  
  
top10_produtos
```

Nome_Produto	Unidades_Vendidas
--------------	-------------------

Nome_Produto	Unidades_Vendidas
--------------	-------------------

### # Verificando os 10 produtos de maior lucro

### # Verificando os 10 produtos de maior lucro

```
'Custo',  
'Lucro']])
```

```
produtos_custo
```

```
Out[38]:
```

	Nome_Produto	Valor_Unitario	Custo	Lucro
0	Produto_12	197.18	118.31	78.87
1	Produto_20	194.80	116.88	77.92
2	Produto_50	191.30	114.78	76.52
3	Produto_10	191.30	114.78	76.52
4	Produto_21	191.28	114.77	76.51
5	Produto_46	190.45	114.27	76.18
6	Produto_8	171.52	102.91	68.61
7	Produto_34	167.69	100.61	67.08
8	Produto_3	164.55	98.73	65.82
9	Produto_36	161.97	97.18	64.79

```
In [ ]: # Fechando o cursor e encerrando a conexão
```

```
cursor.close()  
con.close()
```