Análise de Dados com Linguagem Python

Projeto 1

Instalando e Carregando os Pacotes

In [1]:
```
# Versão da Linguagem Python
from platform import python_version
print('Versão da Linguagem Python Usada Neste Jupyter Notebook:', python_version())
```
```
Versão da Linguagem Python Usada Neste Jupyter Notebook: 3.9.12
```

In [2]:
```
# Instala o pacote iPython-sql
# https://pypi.org/project/ipython-sql/
!pip install -q ipython-sql
```

In [3]:
```
# Importando os pacotes
import pandas as pd
import sqlite3
```

In [4]:
```
# Versões dos pacotes usados neste jupyter notebook
%reload_ext watermark
%watermark -a "Guilherme Paleari" --iversions
```
```
Author: Guilherme Paleari

sqlite3: 2.6.0
pandas : 1.2.4
```

In [5]:
```
# Criamos a conexão a um banco de dados SQLite
cnn = sqlite3.connect('database/dbprojeto1.db')
```

In [6]:
```
# Carregamos a extensão SQL
%load_ext sql
```

In [7]:
```
# Definimos o banco de dados
%sql sqlite:///database/dbprojeto1.db
```

Com os dados em mãos dos pacientes que desenvolveram diabetes ou não, irei gerar uma amostra de dados com os pacientes com mais de 50 anos e para cada um deles indicar em uma nova coluna se o paciente está normal (BMI menor que 30) ou obeso (BMI maior ou igual a 30). Então gerar um novo arquivo CSV e encaminhar para o Cientista de Dados ou dar continuidade ao projeto.

In [8]:
```
# Carregando os dados do formato .csv para o Dataset chamado df
df = pd.read_csv('dataset/diabetes.csv')
```

In [9]:
```
# Tipo do dataframe
type(df)
```
Out[9]:
```
pandas.core.frame.DataFrame
```

In [10]:
```
# Nesse Dataset podemos ver que tem 768 linhas e 9 colunas.
df.shape
```
Out[10]:
```
(768, 9)
```

In [11]:
```
# Uma prévia dos dados com as 5 primeiras linhas.
df.head()
```
Out[11]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

In [12]:
```
# Copia o dataframe (diabetes) para dentro do banco de dados como uma tabela
df.to_sql('diabetes', cnn)
```

Executando as consultas SQL usando diretamente Linguagem SQL dentro do Jupyter Notebook

In [13]:
```
%%sql

SELECT COUNT(*) FROM diabetes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[13]:

COUNT(*)
768

In [14]:
```
%%sql

SELECT Age, Glucose, Outcome FROM diabetes WHERE Glucose > 190
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[14]:

Age	Glucose	Outcome
53	197	1
41	196	1
41	194	1
57	196	1
31	197	0
24	193	0
34	191	0
59	194	1
29	196	1
25	193	1
39	197	1
67	194	0
55	195	1
28	198	1
62	197	1
22	199	1
31	195	1

In [15]:
```
# Listando todas as colunas do Dataset
df.columns
```
Out[15]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

Criando nova tabela para selecionar os pacientes maiores de 50 anos.

In [16]:
```
%%sql

CREATE TABLE pacientes (Pregnancies INT,
                        Glucose INT,
                        BloodPressure INT,
                        SkinThickness INT,
                        Insulin INT,
                        BMI DECIMAL(8, 2),
                        DiabetesPedigreeFunction DECIMAL(8, 2),
                        Age INT,
                        Outcome INT);
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[16]:
```
[]
```

In [17]:
```
%%sql

SELECT * FROM pacientes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[17]:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome

Inserindo os pacientes através de filtro com SQL.

In [18]:
```
%%sql

INSERT INTO pacientes(Pregnancies,
                      Glucose,
                      BloodPressure,
                      SkinThickness,
                      Insulin,
                      BMI,
                      DiabetesPedigreeFunction,
                      Age,
                      Outcome)
SELECT Pregnancies,
       Glucose,
       BloodPressure,
       SkinThickness,
       Insulin,
       BMI,
       DiabetesPedigreeFunction,
       Age,
       Outcome
FROM diabetes WHERE Age > 50;
```
```
 * sqlite:///database/dbprojeto1.db
81 rows affected.
```
Out[18]:
```
[]
```

In [19]:
```
%%sql

SELECT * FROM pacientes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[19]:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
2	197	70	45	543	30.5	0.158	53	1
8	125	96	0	0	0	0.232	54	1
10	139	80	0	0	27.1	1.441	57	0
1	189	60	23	846	30.1	0.398	59	1
5	166	72	19	175	25.8	0.587	51	1
11	143	94	33	146	36.6	0.254	51	1
13	145	82	19	110	22.2	0.245	57	0
5	109	75	26	0	36	0.546	60	0
4	111	72	47	207	37.1	1.39	56	1
9	171	110	24	240	45.4	0.721	54	1
8	176	90	34	300	33.7	0.467	58	1
2	109	92	0	0	42.7	0.845	54	0
4	134	72	0	0	23.8	0.277	60	1
4	146	92	0	0	31.2	0.539	61	1
5	132	80	0	0	26.8	0.186	69	0
0	105	84	0	0	27.9	0.741	62	1
3	128	78	0	0	21.1	0.268	55	0
5	147	78	0	0	33.7	0.218	65	0
8	181	68	36	495	30.1	0.615	60	1
6	103	72	32	190	37.7	0.324	55	0
8	196	76	29	280	37.5	0.605	57	1
5	162	104	0	0	37.7	0.151	52	1
7	179	95	31	0	34.2	0.164	60	0
2	158	90	0	0	31.6	0.805	66	1
7	142	60	33	190	28.8	0.687	61	0
7	181	84	21	192	35.9	0.586	51	1
11	155	76	28	150	33.3	1.353	51	1
3	142	80	15	0	32.4	0.2	63	0
13	106	70	0	0	34.2	0.251	52	0
5	114	74	0	0	24.9	0.744	57	0
2	108	80	0	0	27	0.259	52	1
7	136	74	26	135	26	0.647	51	0
0	161	50	0	0	21.9	0.254	65	0
8	112	72	0	0	23.6	0.84	58	0
6	194	78	0	0	23.5	0.129	59	1
8	95	72	0	0	36.8	0.485	57	0
5	158	70	0	0	29.8	0.207	63	0
5	103	108	37	0	39.2	0.305	65	0
4	146	78	0	0	38.5	0.52	67	1
12	140	82	43	325	39.2	0.528	58	1
5	144	82	26	285	32	0.452	58	1
6	137	61	0	0	24.2	0.151	55	0
2	119	0	0	0	19.6	0.832	72	0
1	135	54	0	0	26.7	0.687	62	0
10	148	84	48	237	37.6	1.001	51	1
9	134	74	33	60	25.9	0.46	81	0
0	137	84	27	0	27.3	0.231	59	0
4	132	86	31	0	28	0.419	63	0
0	173	78	32	265	46.5	1.159	58	0
8	194	80	0	0	26.1	0.551	67	0
6	166	74	0	0	26.6	0.304	66	0
7	195	70	33	145	25.1	0.163	55	1
8	120	78	0	0	25	0.409	64	0
9	91	68	0	0	24.2	0.2	58	0
9	145	88	34	165	30.3	0.771	53	1
7	125	86	0	0	37.6	0.304	51	0
6	129	90	7	326	19.6	0.582	60	0
0	57	60	0	0	21.7	0.735	67	0
10	90	85	32	0	34.9	0.825	56	1
5	187	76	27	207	43.6	1.034	53	1
6	114	88	0	0	27.8	0.247	66	0
8	110	76	0	0	27.8	0.237	58	0
6	125	76	0	0	33.8	0.121	54	1
2	197	70	99	0	34.7	0.575	62	1
12	121	78	17	0	26.5	0.259	62	0
8	124	76	24	600	28.7	0.687	52	1
3	176	86	27	156	33.3	1.154	52	1
7	150	78	29	126	35.2	0.692	54	1
11	127	106	0	0	39	0.19	51	0
10	162	84	0	0	27.7	0.182	54	0
4	145	82	18	0	32.5	0.235	70	1
8	91	82	0	0	35.6	0.587	68	0
9	156	86	0	0	24.8	0.23	53	1
5	136	82	0	0	0	0.64	69	0
1	168	88	29	0	35	0.905	52	1
10	94	72	18	0	23.1	0.595	56	0
5	97	76	27	0	35.6	0.378	52	1
2	105	75	0	0	23.3	0.56	53	0
0	123	72	0	0	36.3	0.258	52	1
6	190	92	0	0	35.5	0.278	66	1
10	101	76	48	180	32.9	0.171	63	0

Criando uma coluna para classificar os pacientes com 50 anos.

In [20]:
```
%%sql

ALTER TABLE pacientes
ADD Perfil VARCHAR(10);
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[20]:
```
[]
```

In [21]:
```
%%sql

SELECT * FROM pacientes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[21]:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Perfil
2	197	70	45	543	30.5	0.158	53	1	None
8	125	96	0	0	0	0.232	54	1	None
10	139	80	0	0	27.1	1.441	57	0	None
1	189	60	23	846	30.1	0.398	59	1	None
5	166	72	19	175	25.8	0.587	51	1	None
11	143	94	33	146	36.6	0.254	51	1	None
13	145	82	19	110	22.2	0.245	57	0	None
5	109	75	26	0	36	0.546	60	0	None
4	111	72	47	207	37.1	1.39	56	1	None
9	171	110	24	240	45.4	0.721	54	1	None
8	176	90	34	300	33.7	0.467	58	1	None
2	109	92	0	0	42.7	0.845	54	0	None
4	134	72	0	0	23.8	0.277	60	1	None
4	146	92	0	0	31.2	0.539	61	1	None
5	132	80	0	0	26.8	0.186	69	0	None
0	105	84	0	0	27.9	0.741	62	1	None
3	128	78	0	0	21.1	0.268	55	0	None
5	147	78	0	0	33.7	0.218	65	0	None
8	181	68	36	495	30.1	0.615	60	1	None
6	103	72	32	190	37.7	0.324	55	0	None
8	196	76	29	280	37.5	0.605	57	1	None
5	162	104	0	0	37.7	0.151	52	1	None
7	179	95	31	0	34.2	0.164	60	0	None
2	158	90	0	0	31.6	0.805	66	1	None
7	142	60	33	190	28.8	0.687	61	0	None
7	181	84	21	192	35.9	0.586	51	1	None
11	155	76	28	150	33.3	1.353	51	1	None
3	142	80	15	0	32.4	0.2	63	0	None
13	106	70	0	0	34.2	0.251	52	0	None
5	114	74	0	0	24.9	0.744	57	0	None
2	108	80	0	0	27	0.259	52	1	None
7	136	74	26	135	26	0.647	51	0	None
0	161	50	0	0	21.9	0.254	65	0	None
8	112	72	0	0	23.6	0.84	58	0	None
6	194	78	0	0	23.5	0.129	59	1	None
8	95	72	0	0	36.8	0.485	57	0	None
5	158	70	0	0	29.8	0.207	63	0	None
5	103	108	37	0	39.2	0.305	65	0	None
4	146	78	0	0	38.5	0.52	67	1	None
12	140	82	43	325	39.2	0.528	58	1	None
5	144	82	26	285	32	0.452	58	1	None
6	137	61	0	0	24.2	0.151	55	0	None
2	119	0	0	0	19.6	0.832	72	0	None
1	135	54	0	0	26.7	0.687	62	0	None
10	148	84	48	237	37.6	1.001	51	1	None
9	134	74	33	60	25.9	0.46	81	0	None
0	137	84	27	0	27.3	0.231	59	0	None
4	132	86	31	0	28	0.419	63	0	None
0	173	78	32	265	46.5	1.159	58	0	None
8	194	80	0	0	26.1	0.551	67	0	None
6	166	74	0	0	26.6	0.304	66	0	None
7	195	70	33	145	25.1	0.163	55	1	None
8	120	78	0	0	25	0.409	64	0	None
9	91	68	0	0	24.2	0.2	58	0	None
9	145	88	34	165	30.3	0.771	53	1	None
7	125	86	0	0	37.6	0.304	51	0	None
6	129	90	7	326	19.6	0.582	60	0	None
0	57	60	0	0	21.7	0.735	67	0	None
10	90	85	32	0	34.9	0.825	56	1	None
5	187	76	27	207	43.6	1.034	53	1	None
6	114	88	0	0	27.8	0.247	66	0	None
8	110	76	0	0	27.8	0.237	58	0	None
6	125	76	0	0	33.8	0.121	54	1	None
2	197	70	99	0	34.7	0.575	62	1	None
12	121	78	17	0	26.5	0.259	62	0	None
8	124	76	24	600	28.7	0.687	52	1	None
3	176	86	27	156	33.3	1.154	52	1	None
7	150	78	29	126	35.2	0.692	54	1	None
11	127	106	0	0	39	0.19	51	0	None
10	162	84	0	0	27.7	0.182	54	0	None
4	145	82	18	0	32.5	0.235	70	1	None
8	91	82	0	0	35.6	0.587	68	0	None
9	156	86	0	0	24.8	0.23	53	1	None
5	136	82	0	0	0	0.64	69	0	None
1	168	88	29	0	35	0.905	52	1	None
10	94	72	18	0	23.1	0.595	56	0	None
5	97	76	27	0	35.6	0.378	52	1	None
2	105	75	0	0	23.3	0.56	53	0	None
0	123	72	0	0	36.3	0.258	52	1	None
6	190	92	0	0	35.5	0.278	66	1	None
10	101	76	48	180	32.9	0.171	63	0	None

In [22]:
```
%%sql

UPDATE pacientes
SET Perfil = 'Normal'
WHERE BMI < 30;
```
```
 * sqlite:///database/dbprojeto1.db
38 rows affected.
```
Out[22]:
```
[]
```

In [23]:
```
%%sql

SELECT * FROM pacientes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[23]:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Perfil
2	197	70	45	543	30.5	0.158	53	1	None
8	125	96	0	0	0	0.232	54	1	Normal
10	139	80	0	0	27.1	1.441	57	0	Normal
1	189	60	23	846	30.1	0.398	59	1	None
5	166	72	19	175	25.8	0.587	51	1	Normal
11	143	94	33	146	36.6	0.254	51	1	None
13	145	82	19	110	22.2	0.245	57	0	Normal
5	109	75	26	0	36	0.546	60	0	None
4	111	72	47	207	37.1	1.39	56	1	None
9	171	110	24	240	45.4	0.721	54	1	None
8	176	90	34	300	33.7	0.467	58	1	None
2	109	92	0	0	42.7	0.845	54	0	None
4	134	72	0	0	23.8	0.277	60	1	Normal
4	146	92	0	0	31.2	0.539	61	1	None
5	132	80	0	0	26.8	0.186	69	0	Normal
0	105	84	0	0	27.9	0.741	62	1	Normal
3	128	78	0	0	21.1	0.268	55	0	Normal
5	147	78	0	0	33.7	0.218	65	0	None
8	181	68	36	495	30.1	0.615	60	1	None
6	103	72	32	190	37.7	0.324	55	0	None
8	196	76	29	280	37.5	0.605	57	1	None
5	162	104	0	0	37.7	0.151	52	1	None
7	179	95	31	0	34.2	0.164	60	0	None
2	158	90	0	0	31.6	0.805	66	1	None
7	142	60	33	190	28.8	0.687	61	0	Normal
7	181	84	21	192	35.9	0.586	51	1	None
11	155	76	28	150	33.3	1.353	51	1	None
3	142	80	15	0	32.4	0.2	63	0	None
13	106	70	0	0	34.2	0.251	52	0	None
5	114	74	0	0	24.9	0.744	57	0	Normal
2	108	80	0	0	27	0.259	52	1	Normal
7	136	74	26	135	26	0.647	51	0	Normal
0	161	50	0	0	21.9	0.254	65	0	Normal
8	112	72	0	0	23.6	0.84	58	0	Normal
6	194	78	0	0	23.5	0.129	59	1	Normal
8	95	72	0	0	36.8	0.485	57	0	None
5	158	70	0	0	29.8	0.207	63	0	Normal
5	103	108	37	0	39.2	0.305	65	0	None
4	146	78	0	0	38.5	0.52	67	1	None
12	140	82	43	325	39.2	0.528	58	1	None
5	144	82	26	285	32	0.452	58	1	None
6	137	61	0	0	24.2	0.151	55	0	Normal
2	119	0	0	0	19.6	0.832	72	0	Normal
1	135	54	0	0	26.7	0.687	62	0	Normal
10	148	84	48	237	37.6	1.001	51	1	None
9	134	74	33	60	25.9	0.46	81	0	Normal
0	137	84	27	0	27.3	0.231	59	0	Normal
4	132	86	31	0	28	0.419	63	0	Normal
0	173	78	32	265	46.5	1.159	58	0	None
8	194	80	0	0	26.1	0.551	67	0	Normal
6	166	74	0	0	26.6	0.304	66	0	Normal
7	195	70	33	145	25.1	0.163	55	1	Normal
8	120	78	0	0	25	0.409	64	0	Normal
9	91	68	0	0	24.2	0.2	58	0	Normal
9	145	88	34	165	30.3	0.771	53	1	None
7	125	86	0	0	37.6	0.304	51	0	None
6	129	90	7	326	19.6	0.582	60	0	Normal
0	57	60	0	0	21.7	0.735	67	0	Normal
10	90	85	32	0	34.9	0.825	56	1	None
5	187	76	27	207	43.6	1.034	53	1	None
6	114	88	0	0	27.8	0.247	66	0	Normal
8	110	76	0	0	27.8	0.237	58	0	Normal
6	125	76	0	0	33.8	0.121	54	1	None
2	197	70	99	0	34.7	0.575	62	1	None
12	121	78	17	0	26.5	0.259	62	0	Normal
8	124	76	24	600	28.7	0.687	52	1	Normal
3	176	86	27	156	33.3	1.154	52	1	None
7	150	78	29	126	35.2	0.692	54	1	None
11	127	106	0	0	39	0.19	51	0	None
10	162	84	0	0	27.7	0.182	54	0	Normal
4	145	82	18	0	32.5	0.235	70	1	None
8	91	82	0	0	35.6	0.587	68	0	None
9	156	86	0	0	24.8	0.23	53	1	Normal
5	136	82	0	0	0	0.64	69	0	Normal
1	168	88	29	0	35	0.905	52	1	None
10	94	72	18	0	23.1	0.595	56	0	Normal
5	97	76	27	0	35.6	0.378	52	1	None
2	105	75	0	0	23.3	0.56	53	0	Normal
0	123	72	0	0	36.3	0.258	52	1	None
6	190	92	0	0	35.5	0.278	66	1	None
10	101	76	48	180	32.9	0.171	63	0	None

In [24]:
```
%%sql

UPDATE pacientes
SET Perfil = 'Obeso'
WHERE BMI >= 30;
```
```
 * sqlite:///database/dbprojeto1.db
43 rows affected.
```
Out[24]:
```
[]
```

In [25]:
```
%%sql

SELECT * FROM pacientes
```
```
 * sqlite:///database/dbprojeto1.db
Done.
```
Out[25]:

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Perfil
2	197	70	45	543	30.5	0.158	53	1	Obeso
8	125	96	0	0	0	0.232	54	1	Normal
10	139	80	0	0	27.1	1.441	57	0	Normal
1	189	60	23	846	30.1	0.398	59	1	Obeso
5	166	72	19	175	25.8	0.587	51	1	Normal
11	143	94	33	146	36.6	0.254	51	1	Obeso
13	145	82	19	110	22.2	0.245	57	0	Normal
5	109	75	26	0	36	0.546	60	0	Obeso
4	111	72	47	207	37.1	1.39	56	1	Obeso
9	171	110	24	240	45.4	0.721	54	1	Obeso
8	176	90	34	300	33.7	0.467	58	1	Obeso
2	109	92	0	0	42.7	0.845	54	0	Obeso
4	134	72	0	0	23.8	0.277	60	1	Normal
4	146	92	0	0	31.2	0.539	61	1	Obeso
5	132	80	0	0	26.8	0.186	69	0	Normal
0	105	84	0	0	27.9	0.741	62	1	Normal
3	128	78	0	0	21.1	0.268	55	0	Normal
5	147	78	0	0	33.7	0.218	65	0	Obeso
8	181	68	36	495	30.1	0.615	60	1	Obeso
6	103	72	32	190	37.7	0.324	55	0	Obeso
8	196	76	29	280	37.5	0.605	57	1	Obeso
5	162	104	0	0	37.7	0.151	52	1	Obeso
7	179	95	31	0	34.2	0.164	60	0	Obeso
2	158	90	0	0	31.6	0.805	66	1	Obeso
7	142	60	33	190	28.8	0.687	61	0	Normal
7	181	84	21	192	35.9	0.586	51	1	Obeso
11	155	76	28	150	33.3	1.353	51	1	Obeso
3	142	80	15	0	32.4	0.2	63	0	Obeso
13	106	70	0	0	34.2	0.251	52	0	Obeso
5	114	74	0	0	24.9	0.744	57	0	Normal
2	108	80	0	0	27	0.259	52	1	Normal
7	136	74	26	135	26	0.647	51	0	Normal
0	161	50	0	0	21.9	0.254	65	0	Normal
8	112	72	0	0	23.6	0.84	58	0	Normal
6	194	78	0	0	23.5	0.129	59	1	Normal
8	95	72	0	0	36.8	0.485	57	0	Obeso
5	158	70	0	0	29.8	0.207	63	0	Normal
5	103	108	37	0	39.2	0.305	65	0	Obeso
4	146	78	0	0	38.5	0.52	67	1	Obeso
12	140	82	43	325	39.2	0.528	58	1	Obeso
5	144	82	26	285	32	0.452	58	1	Obeso
6	137	61	0	0	24.2	0.151	55	0	Normal
2	119	0	0	0	19.6	0.832	72	0	Normal
1	135	54	0	0	26.7	0.687	62	0	Normal
10	148	84	48	237	37.6	1.001	51	1	Obeso
9	134	74	33	60	25.9	0.46	81	0	Normal
0	137	84	27	0	27.3	0.231	59	0	Normal
4	132	86	31	0	28	0.419	63	0	Normal
0	173	78	32	265	46.5	1.159	58	0	Obeso
8	194	80	0	0	26.1	0.551	67	0	Normal
6	166	74	0	0	26.6	0.304	66	0	Normal
7	195	70	33	145	25.1	0.163	55	1	Normal
8	120	78	0	0	25	0.409	64	0	Normal
9	91	68	0	0	24.2	0.2	58	0	Normal
9	145	88	34	165	30.3	0.771	53	1	Obeso
7	125	86	0	0	37.6	0.304	51	0	Obeso
6	129	90	7	326	19.6	0.582	60	0	Normal
0	57	60	0	0	21.7	0.735	67	0	Normal
10	90	85	32	0	34.9	0.825	56	1	Obeso
5	187	76	27	207	43.6	1.034	53	1	Obeso
6	114	88	0	0	27.8	0.247	66	0	Normal
8	110	76	0	0	27.8	0.237	58	0	Normal
6	125	76	0	0	33.8	0.121	54	1	Obeso
2	197	70	99	0	34.7	0.575	62	1	Obeso
12	121	78	17	0	26.5	0.259	62	0	Normal
8	124	76	24	600	28.7	0.687	52	1	Normal
3	176	86	27	156	33.3	1.154	52	1	Obeso
7	150	78	29	126	35.2	0.692	54	1	Obeso
11	127	106	0	0	39	0.19	51	0	Obeso
10	162	84	0	0	27.7	0.182	54	0	Normal
4	145	82	18	0	32.5	0.235	70	1	Obeso
8	91	82	0	0	35.6	0.587	68	0	Obeso
9	156	86	0	0	24.8	0.23	53	1	Normal
5	136	82	0	0	0	0.64	69	0	Normal
1	168	88	29	0	35	0.905	52	1	Obeso
10	94	72	18	0	23.1	0.595	56	0	Normal
5	97	76	27	0	35.6	0.378	52	1	Obeso
2	105	75	0	0	23.3	0.56	53	0	Normal
0	123	72	0	0	36.3	0.258	52	1	Obeso
6	190	92	0	0	35.5	0.278	66	1	Obeso
10	101	76	48	180	32.9	0.171	63	0	Obeso

Carregando os Dados no Pandas e Salvando o CSV

In [26]:
```
# Criando uma query que vai consultar a tabela "pacientes" e gravando o resultado da query em um cursor
# Cursor é o resultado de uma consulta SQL
query = cnn.execute("SELECT * FROM pacientes")
query
```
Out[26]:
```
<sqlite3.Cursor at 0x2c84332b2d0>
```

In [27]:
```
# Fazendo um Loop com List Comprehension (Programação Python)
# Saindo do SQL e voltando para o Python
# Para cada coluna no resultado da minha query, busque cada uma das colunas e grave no objeto cols
cols = [coluna[0] for coluna in query.description]
cols
```
Out[27]:
```
['Pregnancies',
 'Glucose',
 'BloodPressure',
 'SkinThickness',
 'Insulin',
 'BMI',
 'DiabetesPedigreeFunction',
 'Age',
 'Outcome',
 'Perfil']
```

In [28]:
```
# Gerando o dataframe
# Aqui estou indo la no banco de dados e trazendo os resultados que me interessa e gravando num Dataframe do Pandas
resultado = pd.DataFrame.from_records(data = query.fetchall(), columns = cols)
```

In [29]:
```
# Shape (linhas e colunas)
resultado.shape
```
Out[29]:
```
(81, 10)
```

In [30]:
```
# Visualiza
resultado.head()
```
Out[30]:

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Perfil
0	2	197	70	45	543	30.5	0.158	53	1	Obeso
1	8	125	96	0	0	0.0	0.232	54	1	Normal
2	10	139	80	0	0	27.1	1.441	57	0	Normal
3	1	189	60	23	846	30.1	0.398	59	1	Obeso
4	5	166	72	19	175	25.8	0.587	51	1	Normal

In [31]:
```
# Salva em CSV
resultado.to_csv('dataset/pacientes.csv', index = False)
```

Fim