

```
import pandas as pd
```

```
data = pd.read_csv("https://github.com/SantyPaul19/Practica1/raw/refs/heads/main/diabetes.csv", encoding='latin-1')
```

```
data
```

	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
...	...	...	...	...	...	...	...	...	...
763	10	101	76	48	180	32.9	0.171	63	0
764	2	122	70	27	0	36.8	0.340	27	0
765	5	121	72	23	112	26.2	0.245	30	0
766	1	126	60	0	0	30.1	0.349	47	1
767	1	93	70	31	0	30.4	0.315	23	0

768 rows × 9 columns

Próximos pasos:

[Generar código con data](#)
[Ver gráficos recomendados](#)
[New interactive sheet](#)

```
import random
```

```
# Generar provincias al azar
```


```
def crear_lista_random_de_valores(valores):
```

```
    return [random.choice(valores) for _ in range(768)]
```

```
ciudades = ["Azuay", "Bolívar", "Cañar", "Carchi", "Chimborazo", "Cotopaxi", "El Oro",
            "Esmeraldas", "Galápagos", "Guayas", "Imbabura", "Loja", "Los Ríos", "Manabí",
            "Morona Santiago", "Napó", "Orellana", "Pastaza", "Pichincha", "Santa Elena",
            "Santo Domingo de los Tsáchilas", "Sucumbíos", "Tungurahua", "Zamora Chinchipe"]
ciudades_random = crear_lista_random_de_valores(ciudades)
```

```
data['Provincia'] = ciudades_random
```

```
data
```



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Pr
0	6	148	72	35	0	33.6	0.627	50	1	
1	1	85	66	29	0	26.6	0.351	31	0	Chir
2	8	183	64	0	0	23.3	0.672	32	1	Tun
3	1	89	66	23	94	28.1	0.167	21	0	
4	0	137	40	35	168	43.1	2.288	33	1	Chir
...	...	...	...	...	...	...	...	...	...	...
763	10	101	76	48	180	32.9	0.171	63	0	
764	2	122	70	27	0	36.8	0.340	27	0	C
765	5	121	72	23	112	26.2	0.245	30	0	
766	1	126	60	0	0	30.1	0.349	47	1	Esn
767	1	93	70	31	0	30.4	0.315	23	0	Ir

768 rows × 10 columns

Próximos pasos:

[Generar código con data](#)[Ver gráficos recomendados](#)[New interactive sheet](#)

```
!pip install faker
```



Collecting faker

Downloading faker-37.1.0-py3-none-any.whl.metadata (15 kB)

Requirement already satisfied: tzdata in /usr/local/lib/python3.11/dist-packages (from faker) (2025.2)

Downloading faker-37.1.0-py3-none-any.whl (1.9 MB)

1.9/1.9 MB 20.8 MB/s eta 0:00:00

Installing collected packages: faker

Successfully installed faker-37.1.0


```
from faker import Faker
fake = Faker()
```



```
lista = []
```

```
for _ in range(768):
    lista.append({'nombre': fake.name(), 'apellido' : fake.last_name(), 'email' : fake.email()})
```

```
df = pd.DataFrame(lista)
```

```
df
```




	nombre	apellido	email	
0	Loretta Scott	Jones	tylerparks@example.org	
1	Elizabeth Hinton	Valdez	timothy86@example.com	
2	George Wood	Davis	cmorton@example.org	
3	Austin Cole	Johnson	ingramlisa@example.org	
4	Carol Martin	Johnson	galvancameron@example.org	
...	...	...	...	
763	Bobby Fox	Day	jaredgardner@example.org	
764	Gabriela Zimmerman	Baker	iadams@example.net	
765	Rachel Turner	Nguyen	vguerrero@example.net	
766	Diane Wyatt	Stone	keithrebecca@example.net	
767	Amanda Green	Morris	jeremyanderson@example.com	

768 rows × 3 columns

Próximos pasos: [Generar código con df](#) [Ver gráficos recomendados](#) [New interactive sheet](#)

```
df_concatenated = pd.concat([data, df], axis=1)
```

df\_concatenated



	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome	Pr
0	6	148	72	35	0	33.6	0.627	50	1	
1	1	85	66	29	0	26.6	0.351	31	0	Chir
2	8	183	64	0	0	23.3	0.672	32	1	Tun
3	1	89	66	23	94	28.1	0.167	21	0	
4	0	137	40	35	168	43.1	2.288	33	1	Chir
...	...	...	...	...	...	...	...	...	...	...
763	10	101	76	48	180	32.9	0.171	63	0	
764	2	122	70	27	0	36.8	0.340	27	0	C
765	5	121	72	23	112	26.2	0.245	30	0	
766	1	126	60	0	0	30.1	0.349	47	1	Esn
767	1	93	70	31	0	30.4	0.315	23	0	Ir

768 rows × 13 columns

Próximos pasos: [Generar código con df\\_concatenated](#) [Ver gráficos recomendados](#) [New interactive sheet](#)

```
import os
from pathlib import Path

archivo_csv = Path('/content/misdata.csv')
```

```

if archivo_csv.exists():
    print(f"El archivo CSV '{archivo_csv}' existe.")
    os.remove(archivo_csv)
    df_concatenated.to_csv("misdata.csv", index=False)
    print(f"El archivo CSV '{archivo_csv}' se actualizo.")
else:
    df_concatenated.to_csv("misdata.csv", index=False)
    print(f"El archivo CSV '{archivo_csv}' fue creado.")

```

```

↗ El archivo CSV '/content/misdata.csv' existe.
  El archivo CSV '/content/misdata.csv' se actualizo.

```

```
data1 = pd.read_csv("misdata.csv")
```

```
!pip install pygwalker
```

```

↗ Requirement already satisfied: Send2Trash>=1.8.0 in /usr/local/lib/python3.11/dist-packages (from notebook>=4.4.
Requirement already satisfied: terminado>=0.8.3 in /usr/local/lib/python3.11/dist-packages (from notebook>=4.4.1
Requirement already satisfied: prometheus-client in /usr/local/lib/python3.11/dist-packages (from notebook>=4.4.
Requirement already satisfied: nbclassic>=0.4.7 in /usr/local/lib/python3.11/dist-packages (from notebook>=4.4.1
Requirement already satisfied: platformdirs>=2.5 in /usr/local/lib/python3.11/dist-packages (from jupyter-core>=
Requirement already satisfied: notebook-shim>=0.2.3 in /usr/local/lib/python3.11/dist-packages (from nbclassic>=
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.11/dist-packages (from nbconvert>=5->not
Requirement already satisfied: bleach!=5.0.0 in /usr/local/lib/python3.11/dist-packages (from bleach[css]!=5.0.0
Requirement already satisfied: defusedxml in /usr/local/lib/python3.11/dist-packages (from nbconvert>=5->noteboo
Requirement already satisfied: jupyterlab-pygments in /usr/local/lib/python3.11/dist-packages (from nbconvert>=5
Requirement already satisfied: mistune<4,>=2.0.3 in /usr/local/lib/python3.11/dist-packages (from nbconvert>=5->
Requirement already satisfied: nbclient>=0.5.0 in /usr/local/lib/python3.11/dist-packages (from nbconvert>=5->no
Requirement already satisfied: pandocfilters>=1.4.1 in /usr/local/lib/python3.11/dist-packages (from nbconvert>=
Requirement already satisfied: fastjsonschema>=2.15 in /usr/local/lib/python3.11/dist-packages (from nbformat->n
Requirement already satisfied: jsonschema>=2.6 in /usr/local/lib/python3.11/dist-packages (from nbformat->notebo
Requirement already satisfied: argon2-cffi-bindings in /usr/local/lib/python3.11/dist-packages (from argon2-cffi
Requirement already satisfied: webencodings in /usr/local/lib/python3.11/dist-packages (from bleach!=5.0.0->blea
Requirement already satisfied: tinycss2<1.5,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from bleach[css]
Requirement already satisfied: attrs>=22.2.0 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=2.6->n
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in /usr/local/lib/python3.11/dist-packages (
Requirement already satisfied: referencing>=0.28.4 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=
Requirement already satisfied: rpds-py>=0.7.1 in /usr/local/lib/python3.11/dist-packages (from jsonschema>=2.6->
Requirement already satisfied: jupyter-server<3,>=1.8 in /usr/local/lib/python3.11/dist-packages (from notebook-
Requirement already satisfied: cffi>=1.0.1 in /usr/local/lib/python3.11/dist-packages (from argon2-cffi-bindings
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.11/dist-packages (from beautifulsoup4->nb
Requirement already satisfied: pycparser in /usr/local/lib/python3.11/dist-packages (from cffi>=1.0.1->argon2-cf
Requirement already satisfied: anyio>=3.1.0 in /usr/local/lib/python3.11/dist-packages (from jupyter-server<3,>=
Requirement already satisfied: websocket-client in /usr/local/lib/python3.11/dist-packages (from jupyter-server<
Requirement already satisfied: sniffio>=1.1 in /usr/local/lib/python3.11/dist-packages (from anyio>=3.1.0->jupy
Downloading pygwalker-0.4.9.15-py3-none-any.whl (4.5 MB)
4.5/4.5 MB 41.4 MB/s eta 0:00:00
Downloading gw_dsl_parser-0.1.49.1-py3-none-any.whl (956 kB)
956.2/956.2 kB 36.5 MB/s eta 0:00:00
Downloading kanaries_track-0.0.5-py3-none-any.whl (8.6 kB)
Downloading segment_analytics_python-2.2.3-py2.py3-none-any.whl (24 kB)
Downloading ipylab-1.0.0-py3-none-any.whl (100 kB)
100.2/100.2 kB 8.1 MB/s eta 0:00:00
Downloading anywidget-0.9.18-py3-none-any.whl (220 kB)
220.7/220.7 kB 17.8 MB/s eta 0:00:00
Downloading appdirs-1.4.4-py2.py3-none-any.whl (9.6 kB)
Downloading arrow-1.3.0-py3-none-any.whl (66 kB)
66.4/66.4 kB 4.6 MB/s eta 0:00:00
Downloading astor-0.8.1-py2.py3-none-any.whl (27 kB)
Downloading quickjs-1.19.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.manylinux_2_28_x86_64.whl (2.2
2.2/2.2 MB 75.7 MB/s eta 0:00:00
Downloading backoff-2.2.1-py3-none-any.whl (15 kB)
Downloading dateutils-0.6.12-py2.py3-none-any.whl (5.7 kB)
Downloading jedi-0.19.2-py2.py3-none-any.whl (1.6 MB)
1.6/1.6 MB 66.8 MB/s eta 0:00:00
Downloading monotonic-1.6-py2.py3-none-any.whl (8.2 kB)
Downloading psygnal-0.12.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (765 kB)
765.5/765.5 kB 42.9 MB/s eta 0:00:00
Downloading types_python_dateutil-2.9.0.20241206-py3-none-any.whl (14 kB)
Downloading wasmtime-32.0.0-py3-none-manylinux1_x86_64.whl (7.6 MB)
7.6/7.6 MB 105.1 MB/s eta 0:00:00
Installing collected packages: quickjs, monotonic, appdirs, wasmtime, types-python-dateutil, psygnal, jedi, back
Successfully installed anywidget-0.9.18 appdirs-1.4.4 arrow-1.3.0 astor-0.8.1 backoff-2.2.1 dateutils-0.6.12 gw-

```

```
import pygwalker as pyg
```

Haz doble clic (o ingresa) para editar

```
pyg.walk(data1)
```



DataVisualizationChat

Chart 1+ New

What visualization your want to draw from the datasetAsk >

Field List

- # Outcome ::
- [Icon] Provincia ::
- [Icon] nombre ::
- [Icon] apellido ::
- [Icon] email ::
- [Icon] Measure names ::

---

- # Pregnancies ::
- # Glucose ::
- # BloodPressure ::
- # SkinThickness ::
- # Insulin ::
- # BMI ::
- # DiabetesPedigreeFunction ::
- # Age ::
- # Row count ::
- # Measure values ::

Filters

Color

Drop Field Here

Opacity

Drop Field Here

Size

Drop Field Here

Shape

Drop Field Here

Details

Drop Field Here

X-AxisProvincia

Y-AxisGlucose sum

sum(Glucose)

Provincia

```
<pygwalker.api.pygwalker.PygWalker at 0x7dbb47e3a850>
```

```
pyg.walk(data1)
```



Chart 1 + New

What visualization your want to draw from the dataset

Ask



Field List

- # Outcome
- Provincia
- nombre
- apellido
- email
- Measure names
- Pregnancies

Filters

Age  
notIn: []

Color

Drop Field Here

Opacity

X-Axis

Provincia

Y-Axis

Glucose sum  
Row count  
Age sum

