

Parte 1 Trabajar en un archivo llamado cuaderno1.ipynb Crear un conjunto de datos usando la librería <https://faker.readthedocs.io/en/master/>  
El número de registros que se debe generar es de 10000 El conjunto de datos tendrá las siguientes características: Para cada registro usar las características: fake.name(), fake.last\_name(), fake.email(), fake.address() Pasar los registros a un dataframe de pandas Agregar dos características más: Edad, comprendidas entre 30 y 40 años Provincia, entre las opciones (Loja, Pichincha, Guayas, Azuay, Manabí) Pasar el dataframe a un archivo csv

Parte 2 Trabajar en un archivo llamado cuaderno2.ipynb Leer el archivo previo generado Visualizar la información a través de PyGWalker (instalar previo - <https://github.com/Kanaries/pygwalker>) Un gráfico que muestre el número de registros por cada edad Un gráfico que muestre el número de registros por cada provincia

```
1 !pip install pandas
2 !pip install pygwalker
3 !pip install faker
```

```
1 import pandas as pd
2 from faker import Faker
3 fake = Faker()
```

```
1 lista=[]
```

```
1 for x in range(0,10000):
2     lista.append({'nombre': fake.name(), "apellido":fake.last_name(), "email":fake.email(), "direccion":fake.address()})
3     #fake.last_name(), fake.email(), fake.address()
```

```
1 lista
```

```
1 data = pd.DataFrame(lista)
```

```
1 data
```

	nombre	apellido	email	direccion
0	Janice Peterson	Nguyen	ronnie36@example.com	26630 Renee Field Suite 396\nAntoniovview, WA 7...
1	Kenneth Armstrong	Johnson	veronicawells@example.com	085 Willie Pines Suite 840\nPort Glenda, MN 06818
2	Mr. Darin Hamilton MD	Davis	castanedawarren@example.org	45113 Carlson Squares Suite 461\nPort Adrienne...
3	Brandon King	Weber	jacksondaniel@example.net	2364 Davis Divide Apt. 076\nSteeleview, AS 34296
4	Daniel Bryan	Flowers	eric58@example.com	0397 Hunt Crossing\nWest Madelineside, NV 94151
...	...	...	...	...
9995	Dr. Kimberly Stokes	Strickland	owenskaren@example.net	0385 Patricia Landing\nScottport, KS 64819
9996	Jonathan Martin	Jones	cathy33@example.net	1250 Wagner Mount\nSouth Sandra, AK 80204
9997	Jodi Martinez	Wilkinson	theresarichards@example.net	70371 Vicki Ramp\nNew Douglasland, KS 82361
9998	Christopher Walker	Barnett	mullenrebecca@example.com	49803 Joshua Springs Apt. 063\nPatricktown, NE...
9999	Jeffrey Williams	Kirk	ryan63@example.com	852 David Crossroad\nEast Wendymouth, GA 04814

10000 rows × 4 columns

Next steps:

Generate code with data

View recommended plots

```
1 data.to_csv("misdatacuaderno1.csv", index=False)# convertir en data.csv
```

```
1 data = pd.read_csv("misdatacuaderno1.csv") ##leemos la data
```

```
1 data
```

	nombre	apellido	email	direccion
0	Janice Peterson	Nguyen	ronnie36@example.com	26630 Renee Field Suite 396\nAntoniovview, WA 7...
1	Kenneth Armstrong	Johnson	veronicawells@example.com	085 Willie Pines Suite 840\nPort Glenda, MN 06818
2	Mr. Darin Hamilton MD	Davis	castanedawarren@example.org	45113 Carlson Squares Suite 461\nPort Adrienne...
3	Brandon King	Weber	jacksondaniel@example.net	2364 Davis Divide Apt. 076\nSteeleview, AS 34296
4	Daniel Bryan	Flowers	eric58@example.com	0397 Hunt Crossing\nWest Madelineside, NV 94151
...	...	...	...	...
9995	Dr. Kimberly Stokes	Strickland	owenskaren@example.net	0385 Patricia Landing\nScottport, KS 64819
9996	Jonathan Martin	Jones	cathy33@example.net	1250 Wagner Mount\nSouth Sandra, AK 80204
9997	Jodi Martinez	Wilkinson	theresarichards@example.net	70371 Vicki Ramp\nNew Douglasland, KS 82361
9998	Christopher Walker	Barnett	mullenrebecca@example.com	49803 Joshua Springs Apt. 063\nPatricktown, NE...
9999	Jeffrey Williams	Kirk	ryan63@example.com	852 David Crossroad\nEast Wendymouth, GA 04814

10000 rows × 4 columns

Next steps: [Generate code with data](#) [View recommended plots](#)

```
1 import random as rd

1 prov=['LOJA','PICHINCHA','GUAYAS','AZUAY','MANABI']
2 #edad = str(rd.randint(30, 40))

1 def obtener_edad(d):
2     edad=str(rd.randint(30,40))
3     edad=edad+'años'
4     return edad

1 def obtener_provincia(p):
2     provincia=rd.choice(prov)
3     return provincia

1 data['edad']=data['nombre'].apply(obtener_edad)
2 data['provincia']=data['nombre'].apply(obtener_provincia)
```

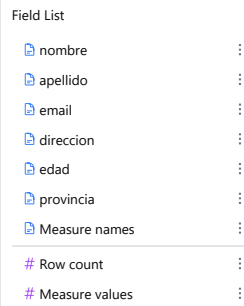
1 data

	nombre	apellido	email	direccion	edad	provincia	
0	Janice Peterson	Nguyen	ronnie36@example.com	26630 Renee Field Suite 396\nAntoniovie, WA 7...	38años	LOJA	
1	Kenneth Armstrong	Johnson	veronicawells@example.com	085 Willie Pines Suite 840\nPort Glenda, MN 06818	39años	GUAYAS	
2	Mr. Darin Hamilton MD	Davis	castanedawarren@example.org	45113 Carlson Squares Suite 461\nPort Adrienne...	34años	AZUAY	
3	Brandon King	Weber	jacksondaniel@example.net	2364 Davis Divide Apt. 076\nSteeleview, AS 34296	34años	MANABI	
4	Daniel Bryan	Flowers	eric58@example.com	0397 Hunt Crossing\nWest Madelineside, NV 94151	38años	LOJA	
...	...	...	...	...	...	...	
9995	Dr. Kimberly Stokes	Strickland	owenskaren@example.net	0385 Patricia Landing\nScottport, KS 64819	37años	AZUAY	
9996	Jonathan Martin	Jones	cathy33@example.net	1250 Wagner Mount\nSouth Sandra, AK 80204	31años	GUAYAS	

Next steps: [Generate code with data](#) [View recommended plots](#)

```
1 import pygwalker as pyg

1 pyg.walk(data)
```

Ask 

Color

---

 provincia

Size

### Details

Y-Axis

Row count

