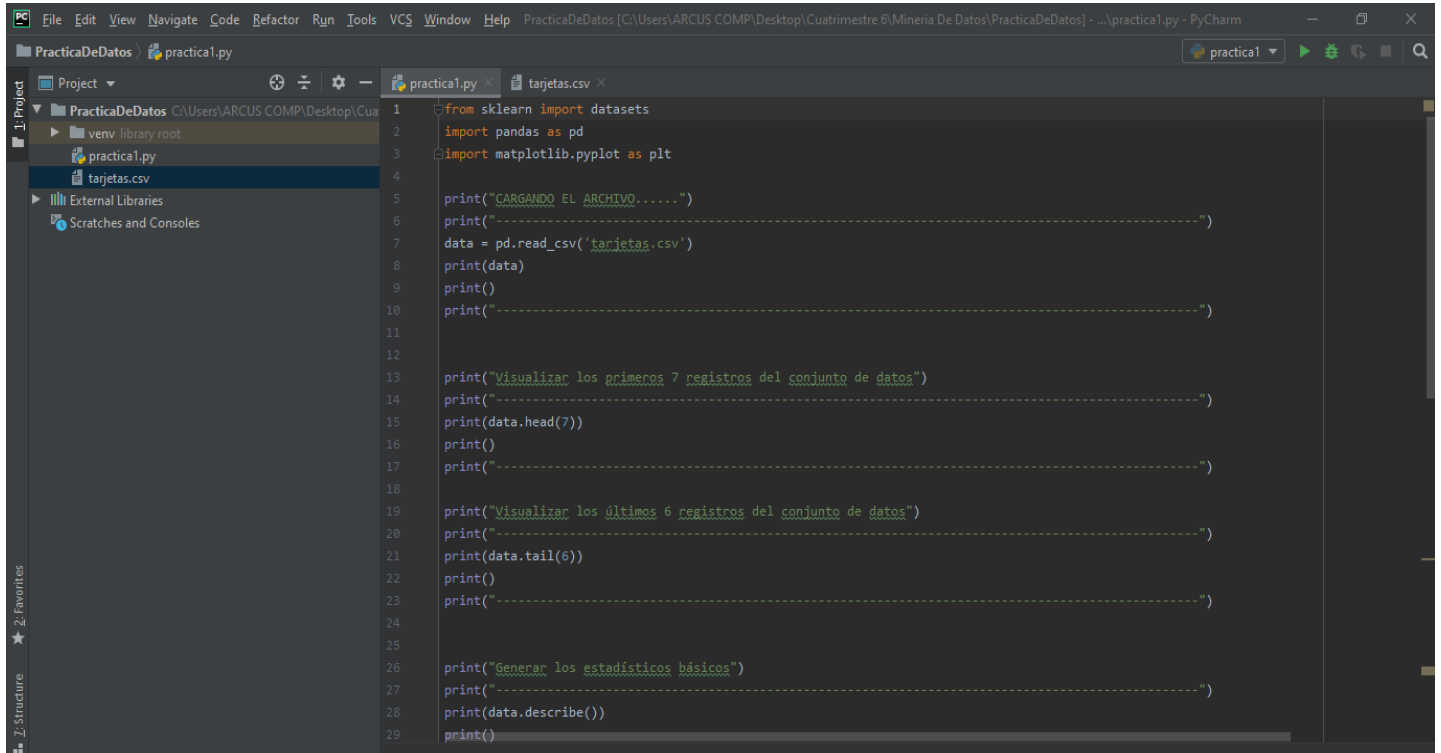
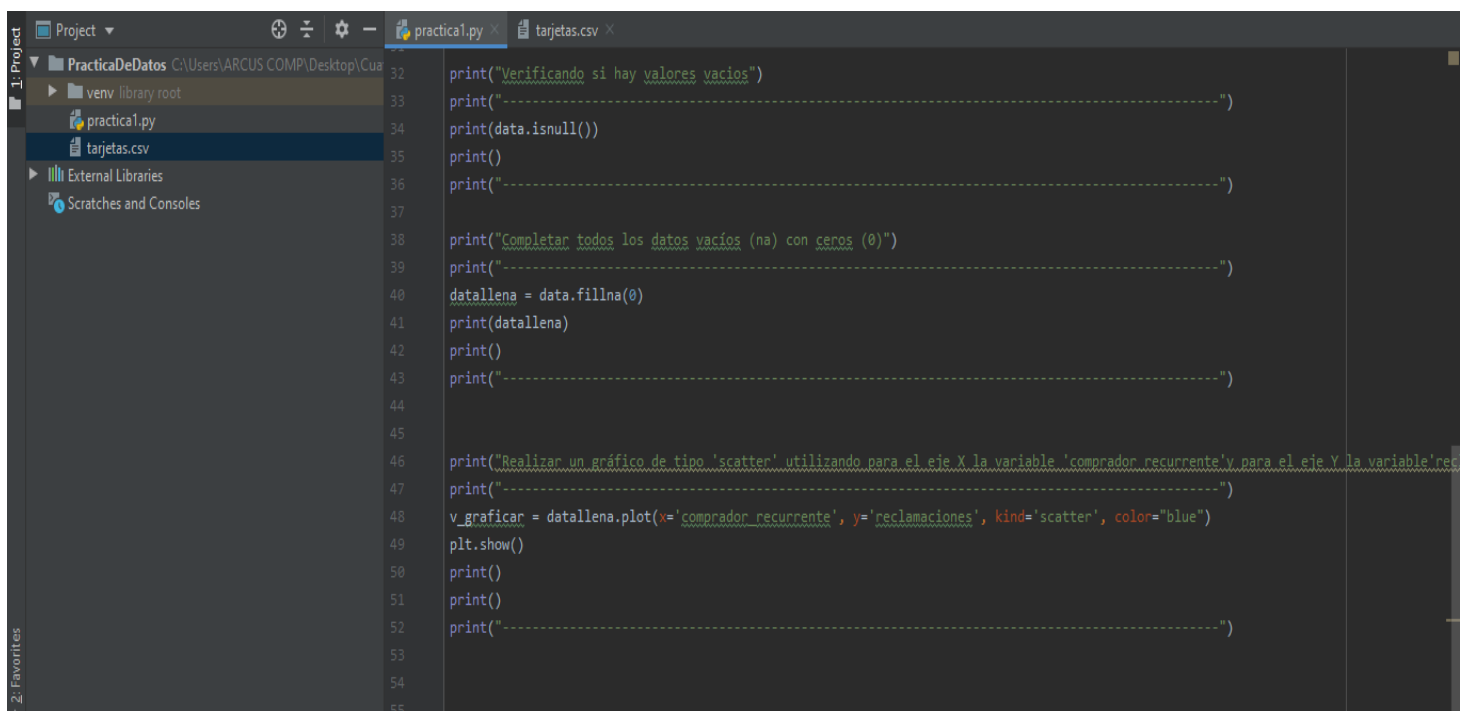


# Minería de datos

## Exploración de Datos



```
1 from sklearn import datasets
2 import pandas as pd
3 import matplotlib.pyplot as plt
4
5 print("CARGANDO EL ARCHIVO.....")
6 print("-----")
7 data = pd.read_csv('tarjetas.csv')
8 print(data)
9 print()
10 print("-----")
11
12
13 print("Visualizar los primeros 7 registros del conjunto de datos")
14 print("-----")
15 print(data.head(7))
16 print()
17 print("-----")
18
19 print("Visualizar los últimos 6 registros del conjunto de datos")
20 print("-----")
21 print(data.tail(6))
22 print()
23 print("-----")
24
25
26 print("Generar los estadísticos básicos")
27 print("-----")
28 print(data.describe())
29 print()
```



```
32 print("Verificando si hay valores vacíos")
33 print("-----")
34 print(data.isnull())
35 print()
36 print("-----")
37
38 print("Completar todos los datos vacíos (na) con ceros (0)")
39 print("-----")
40 dataallena = data.fillna(0)
41 print(dataallena)
42 print()
43 print("-----")
44
45
46 print("Realizar un gráfico de tipo 'scatter' utilizando para el eje X la variable 'comprador recurrente' y para el eje Y la variable 'reclamaciones'")
47 print("-----")
48 v_graficar = dataallena.plot(x='comprador recurrente', y='reclamaciones', kind='scatter', color="blue")
49 plt.show()
50 print()
51 print()
52 print("-----")
53
54
55
```

# Resultados

```
Run: practical x
Visualizar los primeros 7 registros del conjunto de datos
-----
      nombre_tarjeta      fecha  ...  total_compras autorizacion
0  Leilani F. Lowery  03/07/2019  ...      NaN            1
1   Nehru S. Acosta  08/06/2019  ...      6.0            1
2   Rajah G. Trevino 10/22/2019  ...      4.0            1
3    Xena B. Battle  08/17/2019  ...      2.0            1
4    Neve Z. Ramirez 03/17/2019  ...      3.0            1
5    Aiko D. Mccray  05/14/2019  ...      5.0            1
6   Jolie V. Humphrey 08/05/2019  ...      4.0            1

[7 rows x 11 columns]

Visualizar los últimos 6 registros del conjunto de datos
-----
      nombre_tarjeta      fecha  ...  total_compras autorizacion
94   Hayden Q. Parks  08/30/2019  ...      2.0            1
95   Orli Z. Hernandez 10/12/2019  ...      5.0            1
96  Patricia X. Wooten 01/17/2019  ...      NaN            1
97   Carlos H. Wilder  08/29/2019  ...      4.0            1
98   Calista B. Roy   11/18/2018  ...      5.0            1
99   Rose A. Mcfarland 11/12/2018  ...      4.0            1

[6 rows x 11 columns]
```

```
Run: practica1 x
CARGANDO EL ARCHIVO.....
-----
      nombre_tarjeta      fecha ... total_compras autorizacion
0      Leilani F. Lowery 03/07/2019 ...          NaN          1
1      Nehru S. Acosta 08/06/2019 ...          6.0          1
2      Rajah G. Trevino 10/22/2019 ...          4.0          1
3      Xena B. Battle 08/17/2019 ...          2.0          1
4      Neve Z. Ramirez 03/17/2019 ...          3.0          1
..      ... ..
95      Orli Z. Hernandez 10/12/2019 ...          5.0          1
96      Patricia X. Wooten 01/17/2019 ...          NaN          1
97      Carlos H. Wilder 08/29/2019 ...          4.0          1
98      Calista B. Roy 11/18/2018 ...          5.0          1
99      Rose A. Mcfarland 11/12/2018 ...          4.0          1

[100 rows x 11 columns]
```

```
Run: practica1 x
Verificando si hay valores vacios
-----
      nombre_tarjeta fecha hora ... reclamaciones total_compras autorizacion
0      False False False ...          False          True          False
1      False False False ...          False          False          False
2      False False False ...          False          False          False
3      False False False ...          False          False          False
4      False False False ...          False          False          False
..      ... ..
95      False False False ...          False          False          False
96      False False True ...          False          True          False
97      False False False ...          False          False          False
98      False False False ...          False          False          False
99      False False False ...          False          False          False

[100 rows x 11 columns]
```

```
Run: practica1 x
Generar los estadísticos básicos
-----
      hora tipo_tarjeta ... total_compras autorizacion
count  99.000000  100.000000 ...  98.000000  100.0
mean   12.484848   2.500000 ...   3.571429   1.0
std     7.244088   1.123666 ...   1.526400   0.0
min     0.000000   1.000000 ...   1.000000   1.0
25%     6.000000   1.750000 ...   2.000000   1.0
50%    12.000000   2.500000 ...   4.000000   1.0
75%    19.000000   3.250000 ...   5.000000   1.0
max    24.000000   4.000000 ...   6.000000   1.0

[8 rows x 8 columns]
```

```
Completar todos los datos vacíos (na) con ceros (0)
```

	nombre_tarjeta	fecha	...	total_compras	autorizacion
0	Leilani F. Lowery	03/07/2019	...	0.0	1
1	Nehru S. Acosta	08/06/2019	...	6.0	1
2	Rajah G. Trevino	10/22/2019	...	4.0	1
3	Xena B. Battle	08/17/2019	...	2.0	1
4	Neve Z. Ramirez	03/17/2019	...	3.0	1
..	...	...	...	...	...
95	Orli Z. Hernandez	10/12/2019	...	5.0	1
96	Patricia X. Wooten	01/17/2019	...	0.0	1
97	Carlos H. Wilder	08/29/2019	...	4.0	1
98	Calista B. Roy	11/18/2018	...	5.0	1
99	Rose A. Mcfarland	11/12/2018	...	4.0	1

[100 rows x 11 columns]

Realizar un gráfico de tipo 'scatter' utilizando para el eje X la variable 'comprador\_recurrente' y para el eje Y la variable 'reclamaciones'.

Figure 1

