

Programa:

```
import pandas as pd
from sklearn.model_selection import train_test_split

# Leer el archivo csv y convertirlo en un DataFrame
df = pd.read_csv('paleteria.csv')
print('Dataset:')
print(df)

# Separa las características del DataFrame y las imprime
X = df.drop(['y'], axis=1)
print('\nCaracterísticas:')
print(X)

# Separa las etiquetas del DataFrame y las imprime
y = df['y']
print('\nEtiquetas:')
print(y)

# Separa el dataset en un subconjunto de entrenamiento (50%)
# y un subconjunto de prueba (50%) y los imprime como salida
X_train, X_test, y_train, y_test = train_test_split(X, y
, test_size=0.5)
print('\nConjunto de entrenamiento:')
print(X_train)
print(y_train)
print('\nConjunto de prueba:')
print(X_test)
print(y_test)
```

Salida:

Dataset:

	x	y
0	3	5
1	7	11
2	11	21
3	15	16

4 18 16  
5 27 28  
6 29 27  
7 30 25  
8 30 35  
9 31 30  
10 31 40  
11 32 32  
12 33 34  
13 33 32  
14 34 34  
15 36 37  
16 36 38  
17 36 34  
18 37 36  
19 38 38  
20 39 37  
21 39 36  
22 39 45  
23 40 39  
24 41 41  
25 42 40  
26 42 44  
27 43 37  
28 44 44  
29 45 46  
30 46 46  
31 47 49  
32 50 51

Características:

x

0 3

1 7

2 11

3 15

4 18

5 27

6 29

7 30

8 30

9 31

10 31

11 32

12 33

13 33

14 34

15 36

16 36

17 36

18 37

19 38

20 39

21 39

22 39

23 40

24 41

25 42

26 42

27 43

28 44

29 45

30 46

31 47

32 50

Etiquetas:

0 5

1 11

2 21

3 16

4 16

5 28

6 27

7 25

8 35

9 30

10 40

11 32

12 34

13 32

14 34

15 37

16 38

17 34

18 36

19 38

20 37

21 36

22 45

23 39

24 41

25 40

26 44

27 37

28 44

29 46

30 46

31 49

32 51

Name: y, dtype: int64

Conjunto de entrenamiento:

x

3 15

7 30

6 29

2 11

19 38

27 43

8 30

32 50

10 31

14 34

26 42

21 39

28 44

20 39

0 3

30 46

3 16

7 25

6 27

2 21

19 38

27 37

8 35

32 51

10 40

14 34

26 44

21 36

28 44

20 37

0 5

30 46

Name: y, dtype: int64

Conjunto de prueba:

x

31 47

12 33

18 37

22 39

16 36

13 33

1 7

5 27

23 40

17 36

24 41

11 32

25 42

4 18

15 36

9 31

29 45

31 49

12 34

18 36

22 45

16 38

13 32

1 11

5 28

23 39

17 34

24 41

11 32

25 40

4 16

15 37

9 30

29 46

Name: y, dtype: int64