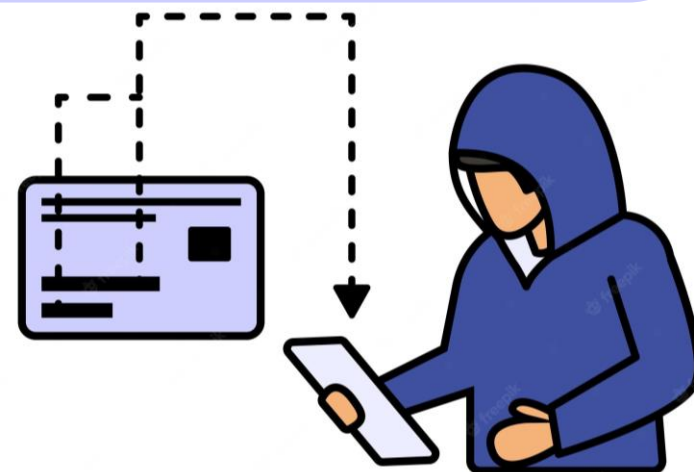




# DETECCIÓN DE FRAUDE EN TARJETAS DE CRÉDITO



FECHA  
LUGAR

SEP 2013  
EUROPA

2 DÍAS

- 284,807 TRANSACCIONES
- 492 FRAUDES 0,172% 

DATOS

V1 – V28  
TRANSFORMACIÓN  
PCA

TIME

AMOUNT

# TRATAMIENTO INICIAL

## ▼ Descripción del dataset

El conjunto de datos trabajado fue tomado de [Kaggle](#).

Característica	Descripción
V1 - V28	Características numéricas que son el resultado de la transformación PCA
Tiempo	Segundos transcurridos entre cada transacción y la primera transacción
Importe	Importe de la transacción
Clase	Fraude (1) o de otro tipo (0)

✓  
0 s

```
[55] # Tamaño de la data  
df.shape
```

```
(284807, 31)
```

✓  
0 s

```
[56] # Corroborar los nombres de las columnas  
df.columns
```

```
Index(['Time', 'V1', 'V2', 'V3', 'V4', 'V5', 'V6', 'V7', 'V8', 'V9', 'V10',  
      'V11', 'V12', 'V13', 'V14', 'V15', 'V16', 'V17', 'V18', 'V19', 'V20',  
      'V21', 'V22', 'V23', 'V24', 'V25', 'V26', 'V27', 'V28', 'Amount',  
      'Class'],  
      dtype='object')
```

✓  
0 s

```
[57] # Chequear duplicados  
df.duplicated().sum()
```

```
1081
```

✓  
1 s

```
[58] # Remover duplicados y volver a chequear  
df = df.drop_duplicates()  
df.duplicated().sum()
```

```
0
```

# TIPOS DE DATOS Y FALTANTES

	Columna	Únicos	Nulos	% nulos	Tipos de datos:	
0	Time	124592	0	0.0	Time	float64
1	V1	275663	0	0.0	V1	float64
2	V2	275663	0	0.0	V2	float64
3	V3	275663	0	0.0	V3	float64
4	V4	275663	0	0.0	V4	float64
5	V5	275663	0	0.0	V5	float64
6	V6	275663	0	0.0	V6	float64
7	V7	275663	0	0.0	V7	float64
8	V8	275663	0	0.0	V8	float64
9	V9	275663	0	0.0	V9	float64
10	V10	275663	0	0.0	V10	float64
11	V11	275663	0	0.0	V11	float64
12	V12	275663	0	0.0	V12	float64
13	V13	275663	0	0.0	V13	float64
14	V14	275663	0	0.0	V14	float64
15	V15	275663	0	0.0	V15	float64
16	V16	275663	0	0.0	V16	float64
17	V17	275663	0	0.0	V17	float64
18	V18	275663	0	0.0	V18	float64
19	V19	275663	0	0.0	V19	float64
20	V20	275663	0	0.0	V20	float64
21	V21	275663	0	0.0	V21	float64
22	V22	275663	0	0.0	V22	float64
23	V23	275663	0	0.0	V23	float64
24	V24	275663	0	0.0	V24	float64
25	V25	275663	0	0.0	V25	float64
26	V26	275663	0	0.0	V26	float64
27	V27	275663	0	0.0	V27	float64
28	V28	275663	0	0.0	V28	float64
29	Amount	32767	0	0.0	Amount	float64
30	Class	2	0	0.0	Class	int64
						dtype: object

# BALANCE DE LOS DATOS

0 – No fraude  
1 – Fraude

```
✓ [61] # Verificamos coherencia en la variable Clase  
0s df.Class.unique()  
  
array([0, 1])  
  
✓ [62] # Balanceo de los datos  
0s df.Class.value_counts(normalize = True)  
  
0    0.998333  
1    0.001667  
Name: Class, dtype: float64
```

# ANÁLISIS EXPLORATORIO

