🌸 **Handling imbalanced dataset in ML** 🌳

credit card fraud dataset

1. Analysis dataset 😊

☐ original data

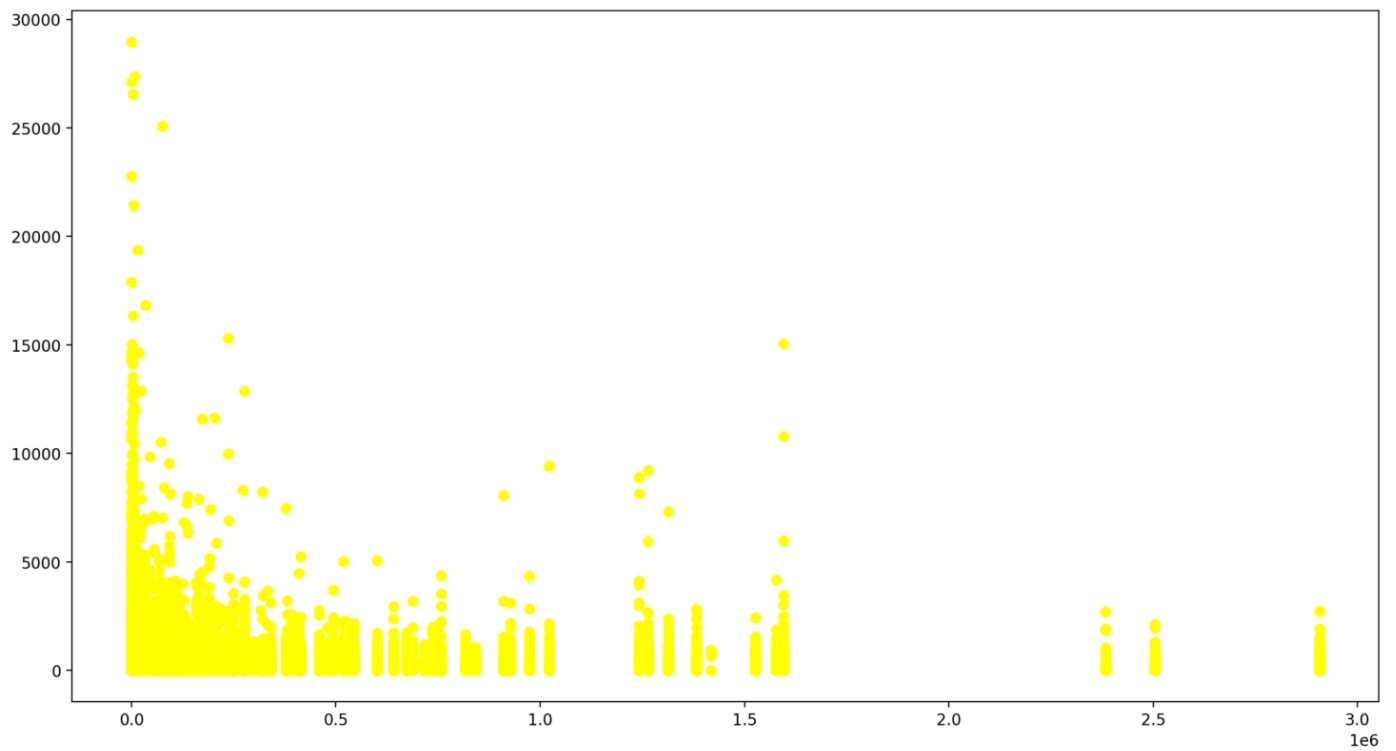### 1.1 original dataframe

	Unnamed: 0	trans_date_trans_time	cc_num	merchant	cate
0	0	2020-06-21 12:14:25	2291163933867244	fraud_Kirlin and Sons	per
1	1	2020-06-21 12:14:33	3573030041201292	fraud_Sporer-Keebler	per
2	2	2020-06-21 12:14:53	3598215285024754	fraud_Swaniawski, Nitzsche and Welch	hea
3	3	2020-06-21 12:15:15	3591919803438423	fraud_Haley Group	mis
4	4	2020-06-21 12:15:17	3526826139003047	fraud_Johnston-Casper	trav

```
class 0 = no fraud data 99.47899852839083
```

```
class 1 = fraud data 0.5210014716091717
```

☐ scatterplot

1.2 scatterplot

☐ boxplot

☐ Decision Tree Classifier

2. confusion matrix 🔢

evaluation:

| option ⌄ |

`select next option`

3. credit card fraud transaction geomap 🚓 :

click here