```
import numpy as np
import pandas as pd
import matplotlib as mt
import matplotlib.pyplot as plt
%matplotlib inline
warnings.filterwarnings('ignore')
```

Evaluating warnings: ignoring UTF-8: not a valid region tag

```
df1 = pd.read_csv('Datasets/code_violations.csv')
df2 = pd.read_csv('Datasets/property_violation.csv')
code_violations = pd.concat([df1, df2], ignore_index=True)
code_violations.shape
```

(54428, 12)

Step 1. Data cleaning and transformation

### 1. a. Clean the dataset
#### Drop unwanted columns
#### Drop of these columns and characters

```
[In]: [0]: [0]: [0
```