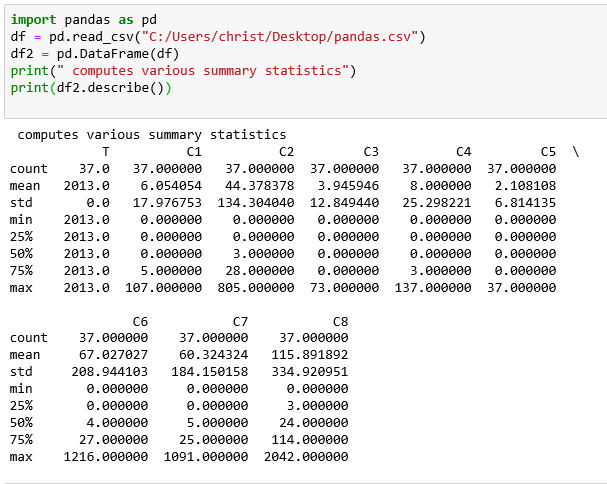
Statistical Analysis of Data Using Pandas



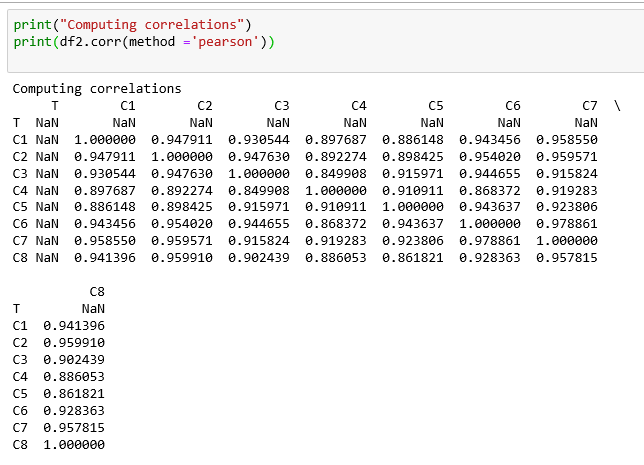
**INFERENCE:**

From the descriptive statistics we know that the count of each column is 37.

From mean we understood that C8 is the most repeated crime in 2013, Since C8 has the highest mean, 115.89, of all crimes.

From standard deviation we can infer that C8 has the greater spread in data with respect to mean.

C8 has the maximum value of all the crime data’s.

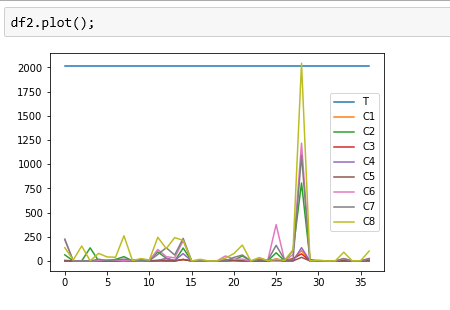


**INFERENCE:**

Here correlation of every variables w.r.t every other variables is positive. This means that all the data’s are positively correlated.

Every data to itself gives 1 since correlation of each variable w.r.t the same variable is always 1.

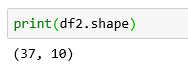
The correlation of C8 to C2 to is maximum from the correlation table.



**INFERENCE:**

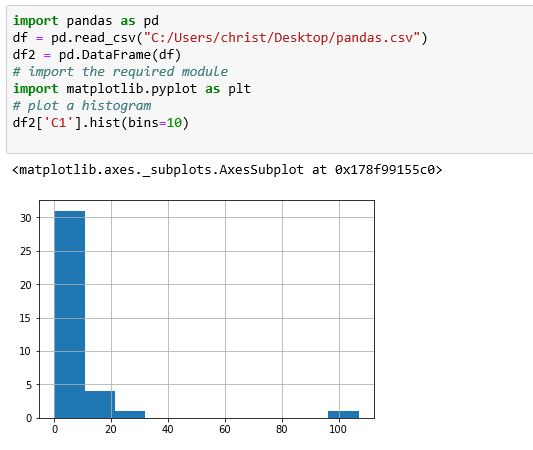
From the plot we can infer that C8 is the crime which is most common in 2013.

C1, C3, C4, C5 has minimum crime rate compared to other values.



**INFERENCE:**

This shows that there are 37 rows and 10 columns in the dataset given for analysis.



**INFERENCE:**

This is the histogram of the crime C1. There are 10 bins in this histogram.

This is the graphical display of the data C1 i.e. visualisation of crime 1.

**CODE:**

import pandas as pd

df = pd.read\_csv("C:/Users/christ/Desktop/pandas.csv")

df2 = pd.DataFrame(df)

print(" computes various summary statistics")

print(df2.describe())

df2.plot();

print("Computing correlations")

print(df2.corr(method ='pearson'))

import pandas as pd

df = pd.read\_csv("C:/Users/christ/Desktop/pandas.csv")

df2 = pd.DataFrame(df)

# import the required module

import matplotlib.pyplot as plt

# plot a histogram

df2['C1'].hist(bins=10)