

In []:

```

class Univariate():

    def QuanQual(self, dataset):
        quan=[]
        qual=[]
        for C_N in dataset.columns:
            if(dataset[C_N].dtype=="O"):
                qual.append(C_N)
            else:
                quan.append(C_N)
        return qual, quan

    def AnalysisTable(self, dataset, Quan):
        import pandas as pd
        import numpy as np
        UniTable=pd.DataFrame(columns=Quan[1:], index=["Mean", "Median", "Mode", "25th-Percentile", "50th-Percentile", "75th-Percentile", "90th-Percentile", "99th-Percentile", "100th-Percentile", "IQR", "IQR*1.5", "Lesser Boundary", "Greater Boundary", "Min value", "Max value", "Outlier"])
        dataset["salary"]=dataset["salary"].fillna(0)
        for C_N in Quan[1:]:
            UniTable[C_N]["Mean"]=dataset[C_N].mean()
            UniTable[C_N]["Median"]=dataset[C_N].median()
            UniTable[C_N]["Mode"]=dataset[C_N].mode()[0]
            UniTable[C_N]["25th-Percentile"]=np.percentile(dataset[C_N], 25)
            UniTable[C_N]["50th-Percentile"]=np.percentile(dataset[C_N], 50)
            UniTable[C_N]["75th-Percentile"]=np.percentile(dataset[C_N], 75)
            UniTable[C_N]["90th-Percentile"]=np.percentile(dataset[C_N], 90)
            UniTable[C_N]["99th-Percentile"]=np.percentile(dataset[C_N], 99)
            UniTable[C_N]["100th-Percentile"]=np.percentile(dataset[C_N], 100)
            UniTable[C_N]["IQR"]=UniTable[C_N]["75th-Percentile"]-UniTable[C_N]["25th-Percentile"]
            UniTable[C_N]["IQR*1.5"]=(UniTable[C_N]["IQR"])*1.5
            UniTable[C_N]["Lesser Boundary"]=UniTable[C_N]["25th-Percentile"]-UniTable[C_N]["IQR*1.5"]
            UniTable[C_N]["Greater Boundary"]=UniTable[C_N]["75th-Percentile"]+UniTable[C_N]["IQR*1.5"]
            UniTable[C_N]["Min value"]=dataset[C_N].min()
            UniTable[C_N]["Max value"]=dataset[C_N].max()
            if(UniTable[C_N]["Lesser Boundary"]>UniTable[C_N]["Min value"] or UniTable[C_N]["Greater Boundary"]>UniTable[C_N]["Max value"]):
                UniTable[C_N]["Outlier"]="Present"
            else:
                UniTable[C_N]["Outlier"]="Not Present"
        return UniTable

    def frequency(self, dataset, c_n):
        import pandas as pd
        FreqTable=pd.DataFrame(columns=["UniqueValues", "Frequency", "Relative-Freq", "Cumulative-Freq"])
        FreqTable["UniqueValues"]=dataset[c_n].value_counts().sort_index().index
        FreqTable["Frequency"]=dataset[c_n].value_counts().sort_index().values
        FreqTable["Relative-Freq"]=(FreqTable["Frequency"]/(len(FreqTable)))*100
        FreqTable["Cumulative-Freq"]=FreqTable["Relative-Freq"].cumsum()
        return FreqTable

```