```
# -*- coding: utf-8 -*-
"""intern_asgmt (1).ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/1ukjnB8phTSyuXgqWwSGJf7kMrh80x0kr

Reg No:2189

Name:K.Keerthi Reddy

Date:11/10/2022
"""

import pandas as pd
import numpy as np

data =pd.read_csv("/content/Enrollments_28092022.csv")
data

data.info()

rows=len(data)
cols=len(data.axes[1])
print("Number of rows:",str(rows))
print("Number of columns:",str(cols))

import matplotlib.pyplot as plt
import statistics as stat

plt.hist(data['DEGREE'])
plt.show()

plt.hist(data['INTERMEDIATE'])
plt.show()

plt.hist(data['SSC'])
plt.show()

cv = lambda x: np.std(x, ddof=1) / np.mean(x) * 100

print("Degree-")
print("Mean=",np.mean(data['DEGREE']))
print("Median=",np.median(data['DEGREE']))
print("Mode=",stat.mode(data['DEGREE']))
print("Range=",max(data['DEGREE'])-min(data['DEGREE']))
print("Co-efficient of Variation=",cv(data['DEGREE']))
data['DEGREE'].describe()

print("Intermediate-")
print("Mean=",np.mean(data['INTERMEDIATE']))
print("Median=",np.median(data['INTERMEDIATE']))
print("Mode=",stat.mode(data['INTERMEDIATE']))
print("Range=",max(data['INTERMEDIATE'])-min(data['INTERMEDIATE']))
print("Co-efficient of Variation=",cv(data['INTERMEDIATE']))
data['INTERMEDIATE'].describe()

print("10th class-")
print("Mean=",np.mean(data['SSC']))
print("Median=",np.median(data['SSC']))
print("Mode=",stat.mode(data['SSC']))
print("Range=",max(data['SSC'])-min(data['SSC']))
print("Co-efficient of Variation=",cv(data['SSC']))
data['SSC'].describe()

data['INTERNSHIP'].value_counts()

courses=['Data Science','Cloud Computing Services (AWS)','MEAN Stack Web Development']
students=[156,90,51]
plt.pie(students, labels = courses,autopct='%1.2f%%')
plt.show

cv = lambda x: np.std(x, ddof=1) / np.mean(x) * 100
```

```
plt.boxplot(data['DEGREE'])
plt.show

plt.boxplot(data['INTERMEDIATE'])
plt.show

plt.boxplot(data['SSC'])
plt.show

#Outliers function
def outlier(a):
  q1 = np.quantile(a, 0.25)

  q3 = np.quantile(a, 0.75)
  med = np.median(a)

  iqr = q3-q1

  upper_bound = q3+(1.5*iqr)
  lower_bound = q1-(1.5*iqr)
  print(iqr, upper_bound, lower_bound)
  print("Inter-Quartile Range:",iqr)
  outliers = a[(a <= lower_bound) | (a >= upper_bound)]
  print('The following are the outliers in the boxplot:\n{}'.format(outliers))

#Degree
outlier(data['DEGREE'])

#Intermediate
outlier(data['INTERMEDIATE'])

#SSC
outlier(data['SSC'])

import scipy.stats as stats

print("Standard Scores of Degree:")
print(stats.zscore(data['DEGREE']))

print("Standard Scores of Intermediate:")
print(stats.zscore(data['INTERMEDIATE']))

print("Standard Scores of 10th class:")
print(stats.zscore(data['SSC']))

def func(b):
  q9 = np.quantile(b, 0.9)
  li=b[b==q9]
  print("No.of students with 90% percentile:",li.count())

#Degree
func(data['DEGREE'])

func(data['INTERMEDIATE'])

func(data['SSC'])
```