NORMAL DISTRIBUTION PROPERTIES

from google.colab import files

data\_to\_load = files.upload()

import pandas as pd

import csv

import statistics

import plotly.figure\_factory as ff

import plotly.graph\_objects as go

import random

df = pd.read\_csv("StudentsPerformance.csv")

ws = df["writing score"].tolist()

wme = statistics.mean(ws)

wmd = statistics.median(ws)

wstd = statistics.stdev(ws)

wmo = statistics.mode(ws)

fstds, fstde = wme-wstd, wme+wstd

sstds, sstde = wme-(2\*wstd), wme+(2\*wstd)

tstds, tstde = wme-(3\*wstd), wme+(3\*wstd)

fig=ff.create\_distplot([ws], ["writing score"], show\_hist=False)

fig.add\_trace(go.Scatter(x=[wme, wme], y=[0, 0.17], mode="lines", name="Mean"))

fig.add\_trace(go.Scatter(x=[fstds, fstds], y=[0, 0.17], mode="lines", name="Standard Deviation 1"))

fig.add\_trace(go.Scatter(x=[fstde, fstde], y=[0, 0.17], mode="lines", name="Standard Deviation 1"))

fig.add\_trace(go.Scatter(x=[sstds, sstds], y=[0, 0.17], mode="lines", name="Standard Deviation 2"))

fig.add\_trace(go.Scatter(x=[sstde, sstde], y=[0, 0.17], mode="lines", name="Standard Deviation 2"))

fig.show()

l1=[result for result in ws if result > fstds and result < fstde]

l2=[result for result in ws if result > sstds and result < sstde]

l3=[result for result in ws if result > tstds and result < tstde]

print("Mean is {}".format(wme))

print("Median is {}".format(wmd))

print("Mode is {}".format(wmo))

print("Standard Deviation is {}".format(wstd))

print("{}% of data is in the first standard deviation".format(len(l1)\*100.0/len(ws)))

print("{}% of data is in the second standard deviation".format(len(l2)\*100.0/len(ws)))

print("{}% of data is in the third standard deviation".format(len(l3)\*100.0/len(ws)))

