Statistic

Certainly! Using pandas and numpy, you can perform various statistical calculations on a dataset. Below are examples of how to find mean, median, mode, measures of dispersion (variance, standard deviation, range), measures of skewness and kurtosis, as well as conduct normality tests for a dataset:

import pandas as pd

import numpy as np

from scipy.stats import skew, kurtosis, Shapiro

# Create a sample DataFrame

data = {'value': [2, 4, 5, 7, 8, 9, 12, 15, 18, 22]}

df = pd.DataFrame(data)

# Mean

mean\_value = np.mean(df['value'])

# Median

median\_value = np.median(df['value'])

# Mode

mode\_value = df['value'].mode().iloc[0]

# Variance

variance\_value = np.var(df['value'])

# Standard Deviation

std\_deviation\_value = np.std(df['value'])

# Range

range\_value = np.ptp(df['value'])

# Skewness

skewness\_value = skew(df['value'])

# Kurtosis

kurtosis\_value = kurtosis(df['value'])

# Shapiro-Wilk Test for Normality

stat, p\_value = shapiro(df['value'])

# Check for normality based on the p-value

if p\_value > 0.05:

print("The data appears to be normally distributed.")

else:

print("The data does not appear to be normally distributed.")