# **Practical – 4**

**Aim: Practical 4: Write a program to compute summary statistics such as mean, median, mode, standard deviation and variance of the given different types of data.**

* Code:

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

data\_set = pd.read\_csv("/content/drive/MyDrive/temp/prac\_4.csv")

df = pd.DataFrame(data\_set)

# ! getting only int or flot data cols

num\_data = df.select\_dtypes(include=["int64", "float64"])

stats = {

"Mean": num\_data.mean(),

"Median": num\_data.median(),

"Mode": num\_data.mode().iloc[0],

"Standard Deviation": num\_data.std(),

"Variance": num\_data.var()

}

state\_table = pd.DataFrame(stats)

print("Result\n")

print(state\_table)

fig, axes = plt.subplots(2, 2, figsize=(10, 8))

temp\_state\_table = state\_table.drop(index = "Release Year", errors = "ignore")

temp\_state\_table2 = state\_table.drop(index = "Active Players (millions)", errors = "ignore")

temp\_table = num\_data.drop(columns = "Release Year")

# ! Hist chart

axes[0, 0].hist(num\_data, bins=5, edgecolor='black', alpha=0.7)

axes[0, 0].set\_title("Histogram of Numerical Data")

axes[0, 0].set\_xlabel("Value")

axes[0, 0].set\_ylabel("Frequency")

# ! Box chart

temp\_table.boxplot(ax=axes[0, 1])

axes[0, 1].set\_title("Box Plot of Numerical Data")

axes[0, 1].set\_xlabel("Columns")

axes[0, 1].set\_ylabel("Value")

# ! Bar Plot graph

temp\_state\_table[["Mean", "Median", "Mode"]].plot(kind="bar", ax=axes[1, 0])

axes[1, 0].set\_title("Comparison of Mean, Median, Mode")

axes[1, 0].set\_ylabel("Value")

# ! Bar Plot

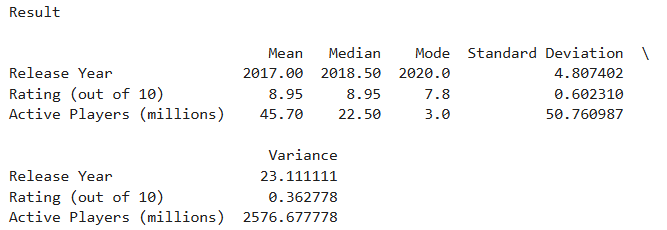
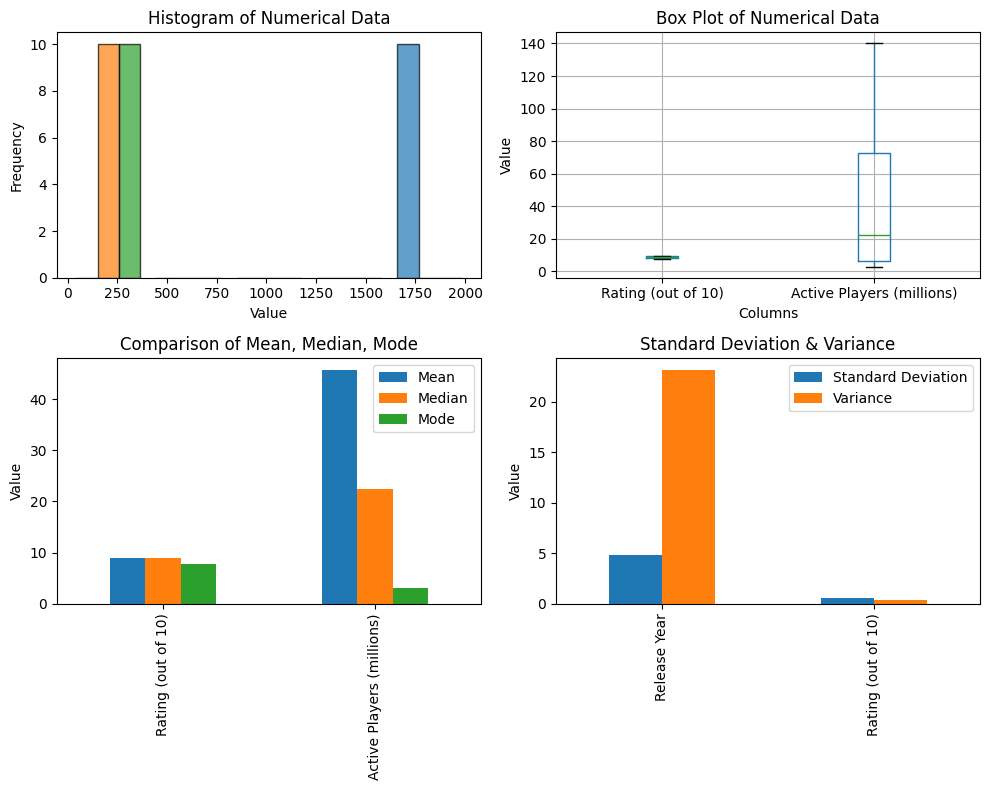
temp\_state\_table2[["Standard Deviation", "Variance"]].plot(kind="bar", ax=axes[1, 1])

axes[1, 1].set\_title("Standard Deviation & Variance")

axes[1, 1].set\_ylabel("Value")

plt.tight\_layout()

plt.show()

* Output:

Faculty Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_