

## EXPERIMENT-18

18. Suppose a hospital tested the age and body fat data for 18 randomly selected adults with the Following result.

<i>age</i>	23	23	27	27	39	41	47	49	50
<i>%fat</i>	9.5	26.5	7.8	17.8	31.4	25.9	27.4	27.2	31.2
<i>age</i>	52	54	54	56	57	58	58	60	61
<i>%fat</i>	34.6	42.5	28.8	33.4	30.2	34.1	32.9	41.2	35.7

Question:

Calculate the mean, median and standard deviation of age and %fat using Pandas.

Draw the boxplots for age and %fat.

Draw a scatter plot and a q-q plot based on these two variables

### Code:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
data = {
    "age": [23,25,27,30,32,35,36,40,41,43,45,47,48,50,52,54,55,57],
    "%fat": [15,18,20,22,25,28,30,32,35,38,40,42,45,47,49,50,52,54]
}
df = pd.DataFrame(data)
print("Mean:\n", df.mean())
print("\nMedian:\n", df.median())
print("\nStandard Deviation:\n", df.std())
plt.figure(figsize=(10,12))
plt.subplot(3,1,1)
df.boxplot(column=["age", "%fat"])
plt.title("Boxplot of Age and %Fat")
plt.grid(True)
plt.subplot(3,1,2)
plt.scatter(df["age"], df["%fat"])
plt.title("Scatter Plot: Age vs %Fat")
plt.xlabel("Age")
plt.ylabel("%Fat")
plt.grid(True)
plt.subplot(3,1,3)
```

```

sorted_age = np.sort(df["age"])
n = len(sorted_age)
theoretical_quantiles = np.linspace(-2, 2, n)
plt.scatter(theoretical_quantiles, sorted_age)
plt.plot(theoretical_quantiles, np.linspace(sorted_age.min(), sorted_age.max(), n),
color='red')
plt.title("Q-Q Plot for Age (No SciPy Needed)")
plt.xlabel("Theoretical Quantiles")
plt.ylabel("Ordered Values")
plt.grid(True)
plt.tight_layout()
plt.show()

```

## Output:



