Practical 1 Find the below data set and perform the following operations:-

Dataset name: -mtcars_DataDescription

1. Read the dataset.

2. Find the head of the dataset.

3. Find the Datatype of Dataset (each column).

4. From the given dataset 'mtcars.csv', plot a histogram to check the frequency distribution of the variable 'mpg' (Miles per gallon).

5. Find the highest frequency of interval.

6. Which can be inferred from scatter plot of 'mpg' (Miles per gallon) vs 'wt' (Weight of car) from the dataset mtcars.csv.

Step 1: Read the dataset

import pandas as pd

 # Assuming the dataset 'mtcars.csv' is in the current directory
mtcars = pd.read_csv('mtcars.csv')

Step 2: Find the head of the dataset

To examine the first few rows of the dataset:

print(mtcars.head())

Step 3: Find the Datatype of Dataset (each column)

To determine the datatype of each column:

print(mtcars.dtypes)

Step 4: Plot a histogram of the 'mpg' variable

To visualize the frequency distribution of the 'mpg' (Miles per gallon) variable using a histogram:

```python
import matplotlib.pyplot as plt
import pandas as pd

mtcars = pd.read_csv('mtcars.csv')
plt.hist(mtcars['mpg'], bins=10, edgecolor='black')
plt.xlabel('Miles per gallon (mpg)')
plt.ylabel('Frequency')
plt.title('Histogram of Miles per gallon (mpg)')
plt.show()
```

Step 5: Find the highest frequency of interval

```python
import matplotlib.pyplot as plt
import numpy as np

# Generate some example data
data = np.random.normal(loc=0, scale=1, size=1000)

# Create a histogram
plt.hist(data, bins=10, edgecolor='black')
plt.xlabel('Values')
plt.ylabel('Frequency')
plt.title('Histogram of Data')
plt.grid(True)

# Display the plot
plt.show()
```

 # Get the histogram data
```
counts, bins, _ = plt.hist(data, bins=10, edgecolor='black')
```

 # Find the bin with the maximum frequency
```
max_frequency = np.max(counts)
max_index = np.argmax(counts)
max_interval = (bins[max_index], bins[max_index + 1])

print(f"The interval with the highest frequency is {max_interval} with frequency {max_frequency}.")
```

Step 6: Inference from scatter plot of 'mpg' vs 'wt'

To infer from the scatter plot of 'mpg' (Miles per gallon) vs 'wt' (Weight of car), we would typically look for a trend or pattern that shows how the mileage (mpg) changes with the weight (wt) of the car. Common inferences could include:

- *__Negative correlation:__ If the scatter plot shows a downward trend (as weight increases, mpg decreases), it suggests that heavier cars tend to have lower mileage.*
- *__No correlation:__ If there's no clear pattern (scatter points are randomly distributed), then there might not be a strong relationship between weight and mileage.*
- *__Positive correlation:__ In some cases, a positive trend (as weight increases, mpg increases) could imply that heavier cars are more fuel-efficient, but this is less common for standard combustion engine cars.*

Here's how you could create a scatter plot to observe this relationship:

```
plt.scatter(mtcars['wt'], mtcars['mpg'], color='blue')
plt.xlabel('Weight of car (wt)')
plt.ylabel('Miles per gallon (mpg)')
plt.title('Scatter plot of Miles per gallon vs Weight of car')
plt.show()
```

After generating this scatter plot, observe the direction of the relationship between 'mpg' and 'wt' to make your inference.