Case Study 2 – Data Visualization

Problem Statement:

You work in XYZ Company as a Python developer. The company officials want you to build a Python program.

Tasks To Be Performed:

1. Load cars data as dataframe using pandas and create a bar plot between number of cylinders and frequency of cars with that many number of cylinders

- Set xlabel as Number of cylinders
- Set ylabel as Frequency of cars
- Draw a bar plot

2. Write code to load data from cars and print a bar graph of count of columns with null values.
3. Use the 'mpg' (Miles Per Gallon column) and draw a histogram:

- Set xlabel: Miles per gallon
- Set ylabel: Frequency
- Set title as Miles Per Gallon Histogram
- Use mpg column to generate a histogram

4. Draw a boxplot on the card dataframes hp column:

- Set xlabel: Car Horsepower
- Set title as Boxplot for car horsepower
- Use hp column to generate a boxplot

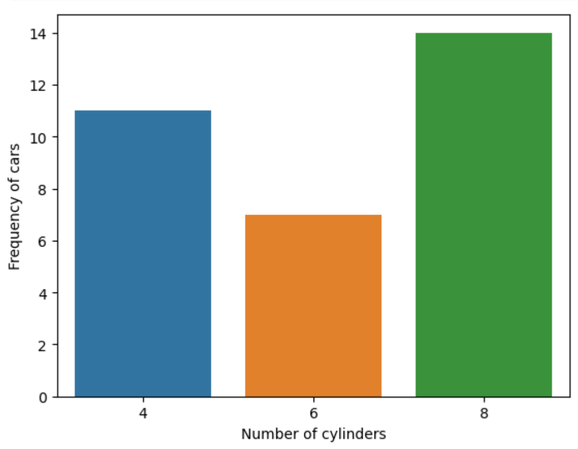
In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns

df = pd.read_csv(r"csv files\cars-6.csv")
df.head()
```

Out[1]:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2

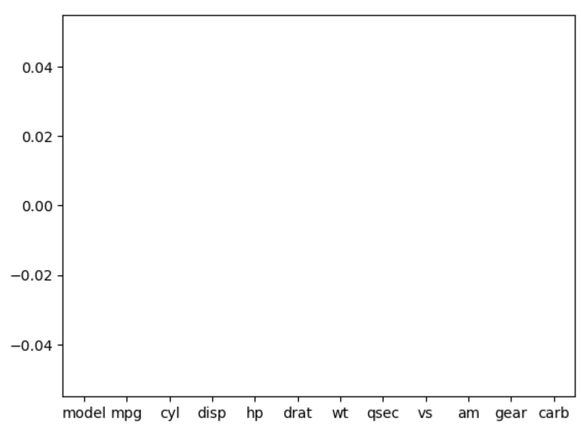
In [10]:
```
sns.barplot(y=df['cyl'].value_counts(),x=df['cyl'].value_counts().keys())
plt.xlabel('Number of cylinders')
plt.ylabel('Frequency of cars')
plt.show()
```



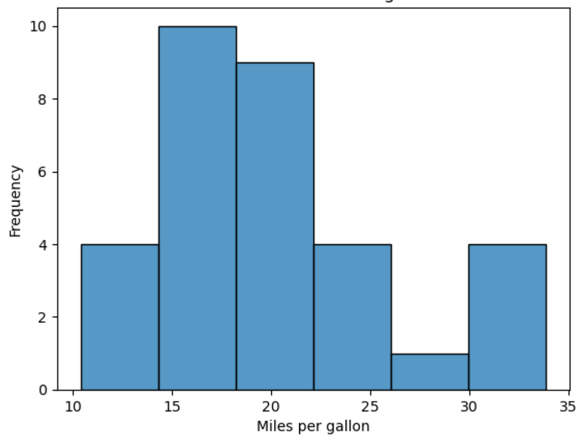
In [15]:
```
df.isna().sum()
```

Out[15]:
```
model    0
mpg      0
cyl      0
disp     0
hp       0
drat     0
wt       0
qsec     0
vs       0
am       0
gear     0
carb     0
dtype: int64
```

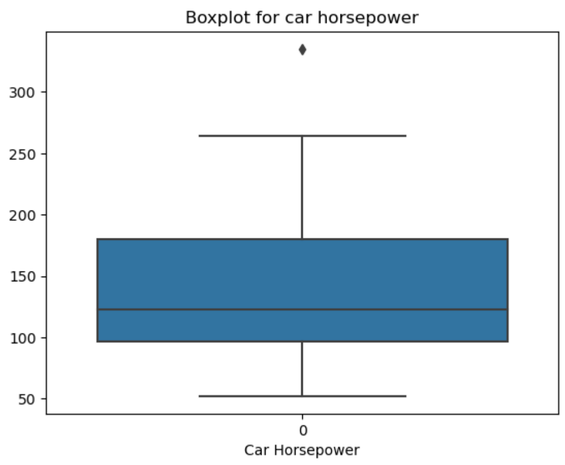
In [20]:
```
sns.barplot(x=df.isna().sum().keys(), y=df.isna().sum())
plt.show()
```



In [16]:
```
sns.histplot(df['mpg'])
plt.xlabel('Miles per gallon')
plt.ylabel('Frequency')
plt.title('Miles Per Gallon Histogram')
plt.show()
```



In [19]:
```
sns.boxplot(df['hp'])
plt.xlabel('Car Horsepower')
plt.title('Boxplot for car horsepower')
plt.show()
```


Boxplot for car horsepower
300
250
200
150
100
50
0
Car Horsepower

In [ ]: