Data Visualization with Matplotlib Assignment

Problem Statement:

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

Tasks to be performed:

In [ ]:
```
#1. Load cars data as dataframe using pandas and create a bar plot between number c
# 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.
```

In [47]:
```
import os
os.chdir("C:\\Users\\veena\\OneDrive\\Desktop")
```

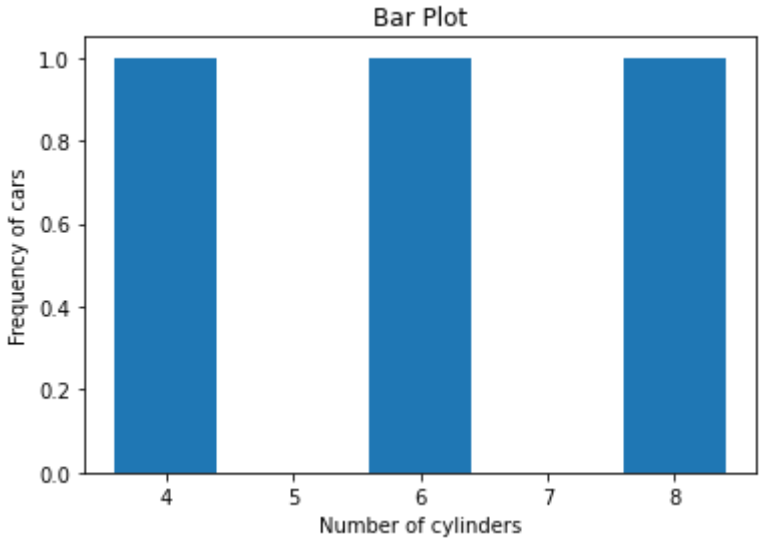
In [48]:
```
import pandas as pd
import matplotlib as plt
%matplotlib inline
dataframe=pd.read_csv('cars-3.csv')
dataframe.head()
```

Out[48]:

	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 [52]:
```
x=list(dataframe.cyl)
```

In [54]:
```
import matplotlib.pyplot  as plt
%matplotlib inline
plt.bar(x,height=1)
plt.title('Bar Plot')
plt.xlabel('Number of cylinders')
plt.ylabel('Frequency of cars')
#plt.bar()
plt.show()
```



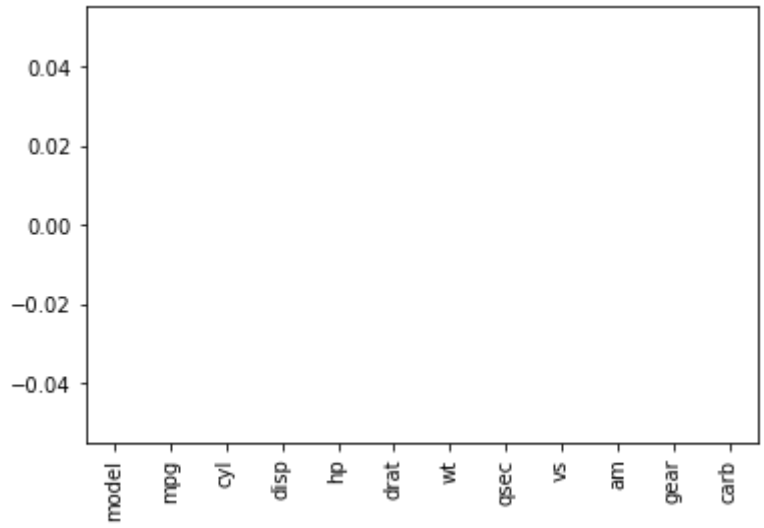
In [ ]:
```
#2. Write code to load data from cars and print a bar graph of count of columns wi
```

In [4]:
```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
dataframe=pd.read_csv('cars-3.csv')
dataframe.head()
```

Out[4]:

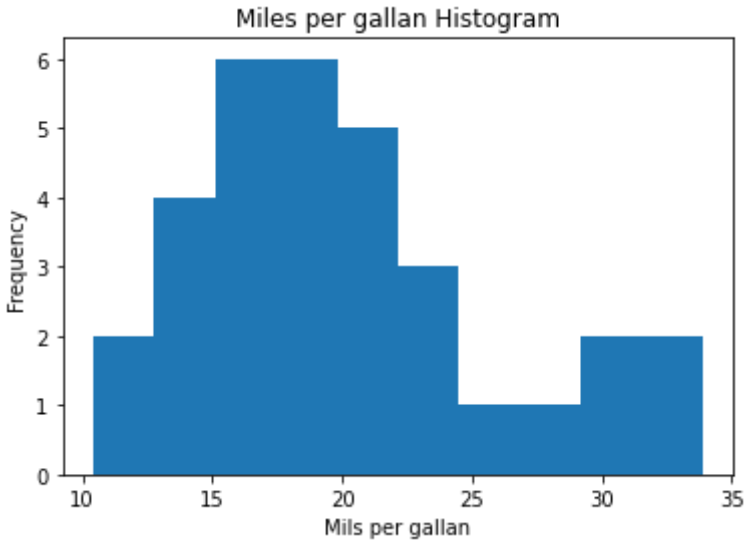
	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7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

In [5]:
```
dataframe.isnull().sum().plot.bar()
plt.show()
```



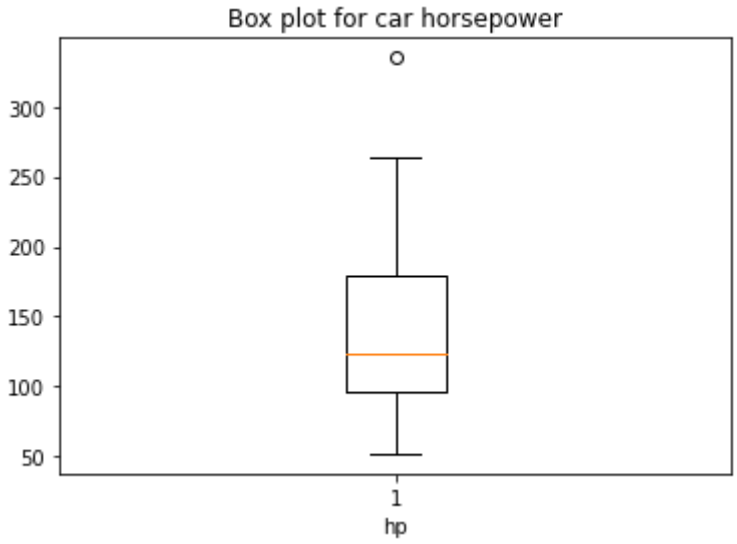
In [ ]:
```
#3. Use the 'mpg' (Miles Per Gallon column) and draw a histogram
#i. Set xlabel: Miles per gallon
#ii. Set ylabel: Frequency
#iii. Set title as Miles Per Gallon Histogram
#iv. Use mpg column to generate a histogram
```

In [9]:
```
plt.hist(dataframe['mpg'])
plt.title('Miles per gallan Histogram')
plt.xlabel('Mils per gallan')
plt.ylabel('Frequency')
plt.show()
```



In [12]:
```
#4. Draw a boxplot on the card dataframes hp column
#i. Set xlabel: Car Horsepower
#ii. Set title as Boxplot for car horsepower
#iii. Use hp column to generate a boxplot
```

In [14]:
```
plt.boxplot(dataframe.hp)
plt.title('Box plot for car horsepower')
plt.xlabel('hp')
plt.show()
```



Module 6: Data Visualization Assignment

Problem Statement: Consider yourself to be Sam who is a data scientist. He has been approached by a telecom company to build some aesthetic graphs to make better sense of the customer data. Tasks to be performed:

In [ ]:
```
#1. Sam has to build a bar-plot for the 'Contract' column
#a. Set the x-axis label to be 'Contract Type of customer'
#b. Set the y-axis label to be 'Count'
#c. Set the title of the plot to be 'Distribution of Contract'
#d. Assign 'orange' color to all the bars
```

In [26]:
```
import os
os.chdir("C:\\Users\\veena\\OneDrive\\Desktop")
```

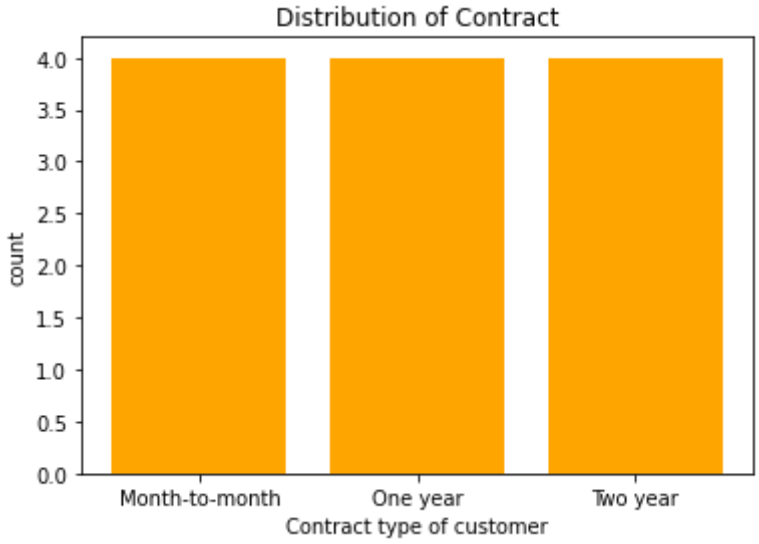
In [29]:
```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
df=pd.read_csv('customer_churn.csv')
df.head()
```

Out[29]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service
1	5575-GNVDE	Male	0	No	No	34	Yes	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No
3	7795-CFOCW	Male	0	No	No	45	No	No phone service
4	9237-HQITU	Female	0	No	No	2	Yes	No

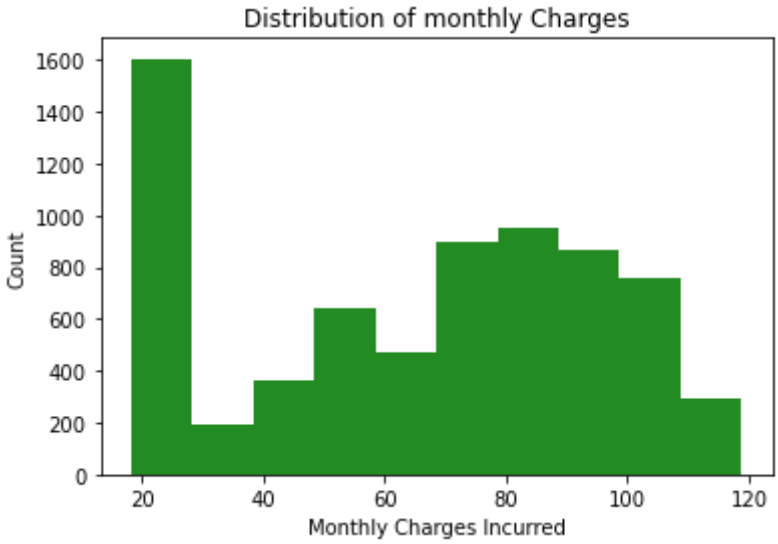
5 rows × 21 columns

In [57]:
```
x=list(df.Contract)
plt.bar(x,height=4,color='orange')
plt.title('Distribution of Contract')
plt.xlabel('Contract type of customer')
plt.ylabel('count')
plt.show()
```



In [ ]:
```
#2. Sam has to build a histogram for the 'MonthlyCharges' column
#a. Set the x-axis label to be 'Monthly Charges Incurred'
#b. Set the y-axis label to be 'Count'
#c. Set the title of the plot to be 'Distribution of Monthly Charges'
#d. Assign 'forestgreen' color to the bins
```

In [60]:
```
plt.hist([df.MonthlyCharges],color='forestgreen')
plt.title('Distribution of monthly Charges')
plt.xlabel('Monthly Charges Incurred')
plt.ylabel('Count')
plt.show()
```



In [ ]:
```
#3. Sam has to build a scatter-plot between 'TotalCharges' & 'tenure'. 'TotalCharge
#the y-axis and 'tenure' should be on the x-axis
#a. Set the x-axis label to be 'Tenure of the customer'
#b. Set the y-axis label to be 'Total chargesIncurred'
#c. Set the title of the plot to be 'Total Charges vs Tenure'
#d. Assign 'indigo' color to the points
```

In [1]:
```
import os
os.chdir("C:\\Users\\veena\\OneDrive\\Desktop")
```

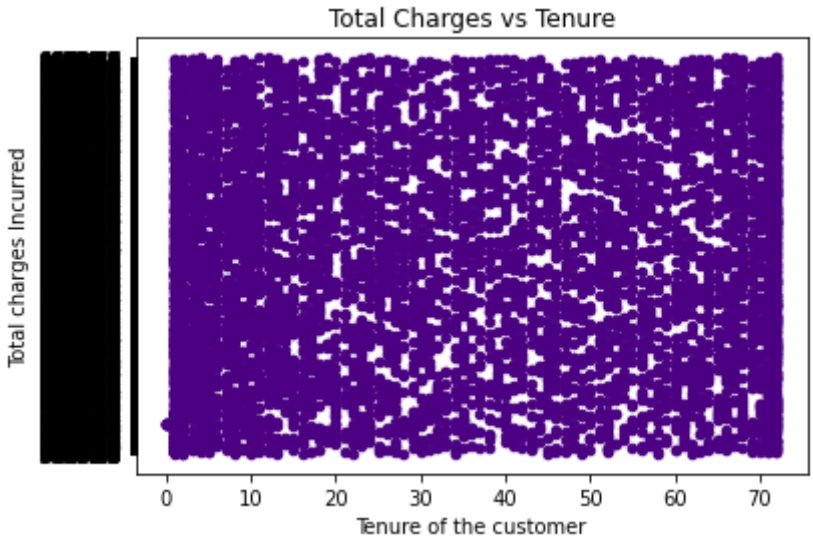
In [2]:
```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
df=pd.read_csv('customer_churn.csv')
df.head()
```

Out[2]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service
1	5575-GNVDE	Male	0	No	No	34	Yes	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No
3	7795-CFOCW	Male	0	No	No	45	No	No phone service
4	9237-HQITU	Female	0	No	No	2	Yes	No

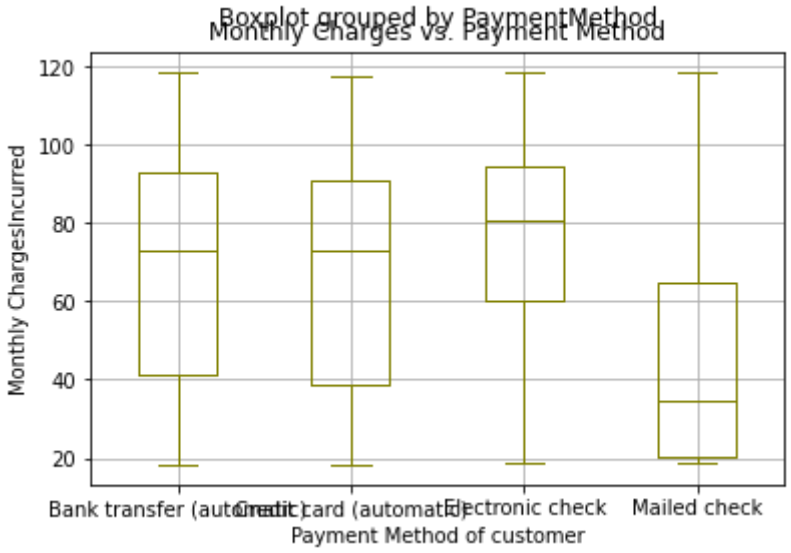
5 rows × 21 columns

In [4]:
```
df.plot.scatter(x='tenure', y='TotalCharges', color='indigo')
plt.title('Total Charges vs Tenure')
plt.xlabel('Tenure of the customer')
plt.ylabel('Total charges Incurred')
plt.show()
```



In [ ]:
```
#4. Sam has to build a box-plot between 'MonthlyCharges' & 'PaymentMethod'.
#'MonthlyCharges' should be on the y-axis and 'PaymentMethod' should be on the x-ax
#a. Set the x-axis label to be 'Payment Method of customer'
#b. Set the y-axis label to be 'Monthly ChargesIncurred'
#c. Set the title of plot to be 'Monthly Charges vs. Payment Method'
#d. Assign 'olive' color to the box-plots
```

In [81]:
```
df.boxplot(by='PaymentMethod',column='MonthlyCharges',color='olive')
plt.title('Monthly Charges vs. Payment Method')
plt.xlabel('Payment Method of customer')
plt.ylabel('Monthly ChargesIncurred')
plt.show()
```



Module 6: Assignment 1 - Data Visualization

Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to visualize the mtcars.csv dataset with various plots. Dataset Link Tasks to be performed:

1. Start off by importing the cars.csv file in the jupyter notebook.
2. Generate a line plot graph for the column 'model' and 'hp'. a. Map the 'model' column on the x-axis. b. Map the 'hp' column on the y-axis. c. Provide the x-axis label as Models of the cars. d. Provide the y-axis label as Horse-Power of Cars. e. Set the title as Model Names vs HorsePower.

In [82]:
```
import os
os.chdir("C:\\Users\\veena\\OneDrive\\Desktop")
```

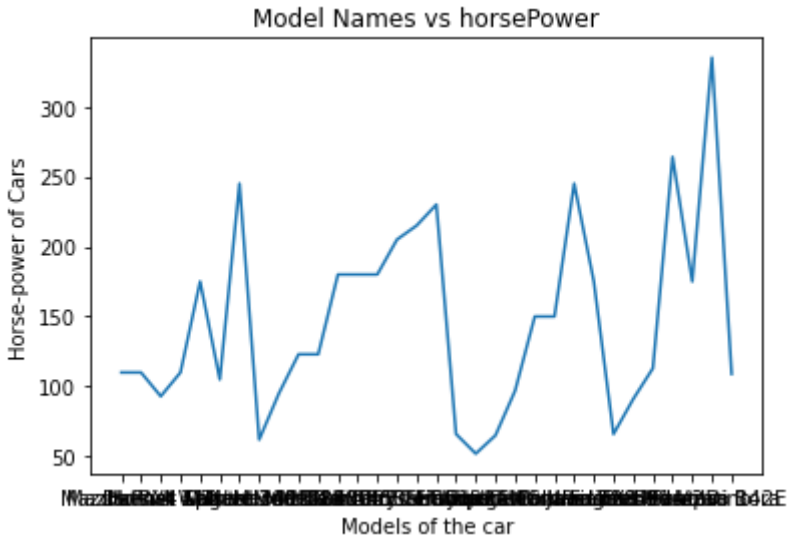
In [84]:
```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
df1=pd.read_csv('cars-3.csv')
df1.head()
```

Out[84]:

	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 [89]:
```
x=list(df1.model)
y=list(df1.hp)
plt.title('Model Names vs horsePower')
plt.xlabel('Models of the car')
plt.ylabel('Horse-power of Cars')
```

```
plt.plot(x,y)
plt.show()
```

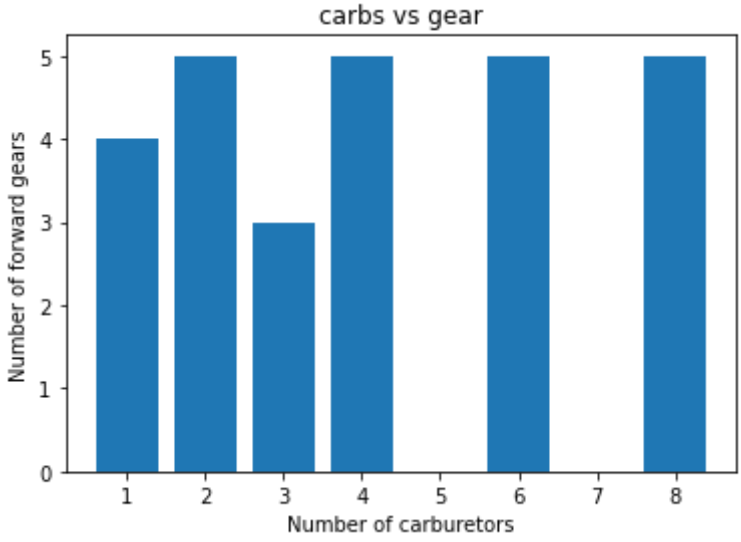


Module 6: Assignment 2 - Data Visualization

Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to visualize the mtcars.csv dataset with various plots. Tasks to be performed: Dataset Link

1. Generate a bar plot graph for the columns 'carbs' and 'gear'. a. Map the 'carbs' onto the x-axis. b. Map the 'gear' onto the y-axis. c. Provide the x-axis label as Number of carburetors. d. Provide the y-axis label as Number of forward gears. e. Set the title as carbs vs gear

In [91]:
```
x=list(df1.carb)
y=list(df1.gear)
plt.bar(x,y)
plt.title('carbs vs gear')
plt.xlabel('Number of carburetors')
plt.ylabel('Number of forward gears')
plt.show()
```

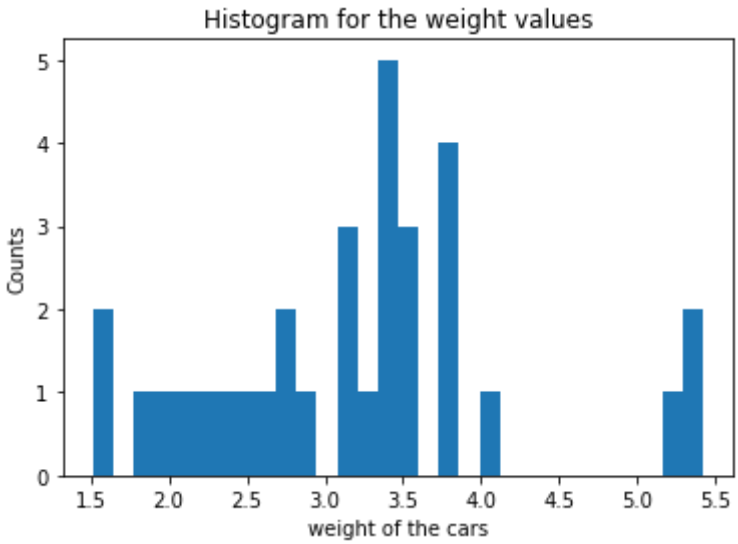


Module 6: Assignment 3 - Data Visualization

Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to visualize the mtcars.csv dataset with various plots. Tasks to be performed: Dataset Link

1. Plot a histogram for the column 'wt'. a. Map the 'wt' onto the x-axis. b. Provide the x-axis label as 'weight of the cars'. c. Provide the y-axis label as 'Count' d. Set the number of bins as 30. e. Set the title as 'Histogram for the weight values'

In [93]:
```
plt.hist(df1.wt,bins=30)
plt.xlabel('weight of the cars')
plt.ylabel('Counts')
plt.title('Histogram for the weight values')
plt.show()
```

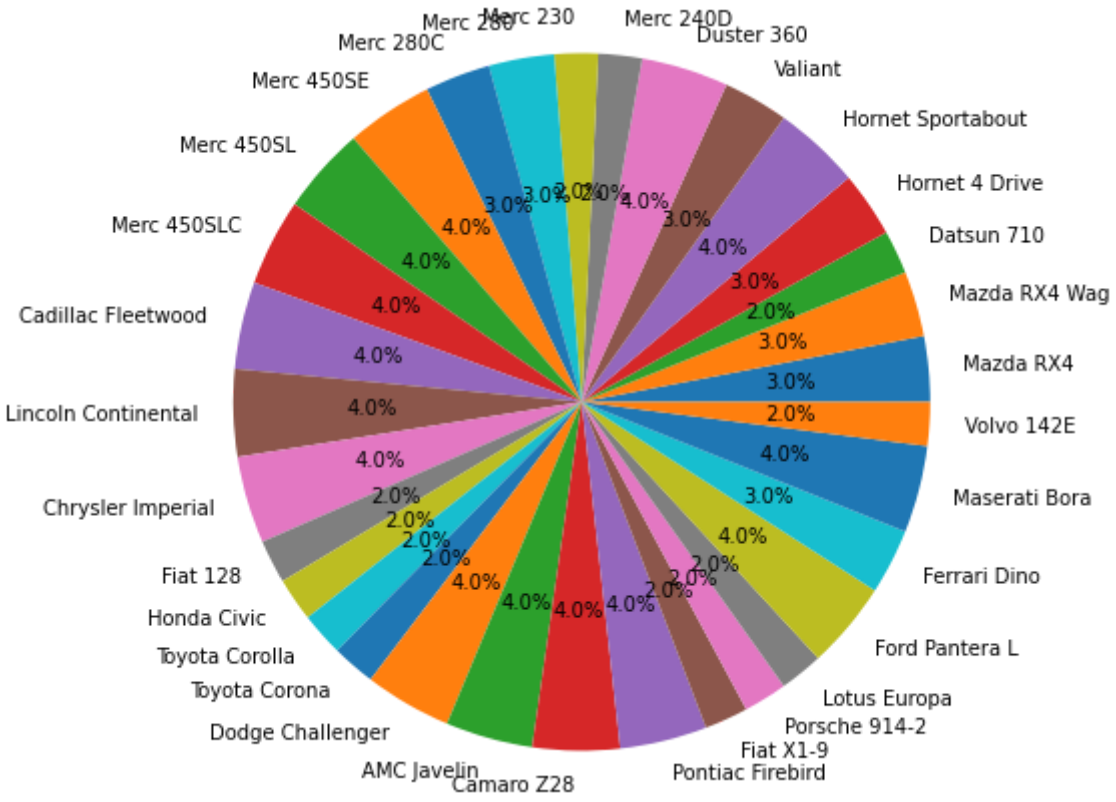


Module 6: Assignment 4 - Data Visualization

Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to visualize the mtcars.csv dataset with various plots. Tasks to be performed: Dataset Link

1. Plot a pie chart for columns: 'cyl' and 'model' form the mtcars.csv data frame.

In [113…
```
plt.axis('equal')
plt.pie(df1.cyl,labels=df1.model, radius=2,autopct='%0.1f%%')
plt.show()
```



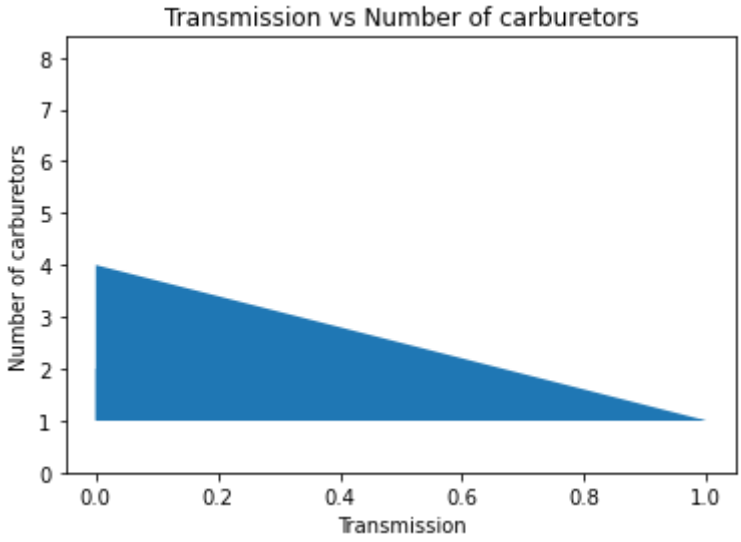
Module 6: Assignment 5 - Data Visualization

Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to visualize the mtcars.csv dataset with various plots. Tasks to be performed: Dataset Link

1. Plot the area chart for the colmns: 'am' and 'carb'. a. Set the 'am' on the x-axis. b. Set the 'carb' on the y-axis. c. Provide the x-axis label as Transmission. d. Provide the y-axis labe as Number of carburetors. e. Provide the title as Transmission vs Number of carburetors.

In [118…

```
plt.stackplot(df1.am, df1.carb)
plt.xlabel('Transmission')
plt.ylabel('Number of carburetors')
plt.title('Transmission vs Number of carburetors')
plt.show()
```


Transmission vs Number of carburetors
Number of carburetors
Transmission
0
1
2
3
4
5
6
7
8
0.0
0.2
0.4
0.6
0.8
1.0

In [ ]: