In [1]:

```
# importing the requried libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:

```
# loading the dataset
data=pd.read_csv(r"C:\Users\Fundemics\Downloads\auto-mpg.csv")
data
```

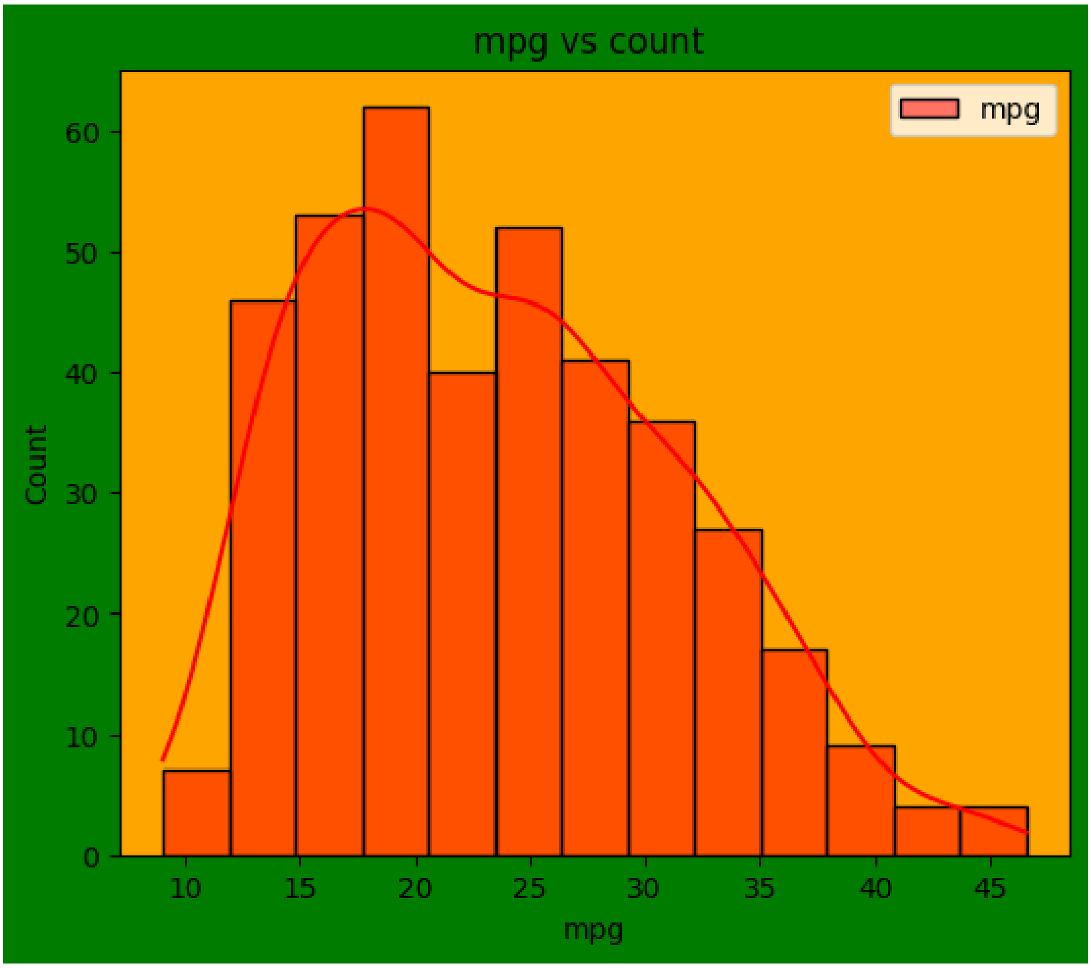
Out[2]:

	mpg	cylinders	displacement	horsepower	weight	accele
0	18.0	8	307.0	130	3504	
1	15.0	8	350.0	165	3693	
2	18.0	8	318.0	150	3436	
3	16.0	8	304.0	150	3433	
4	17.0	8	302.0	140	3449	
...	...	...	...	...	...	
393	27.0	4	140.0	86	2790	
394	44.0	4	97.0	52	2130	
395	32.0	4	135.0	84	2295	
396	28.0	4	120.0	79	2625	
397	31.0	4	119.0	82	2720	

398 rows × 9 columns

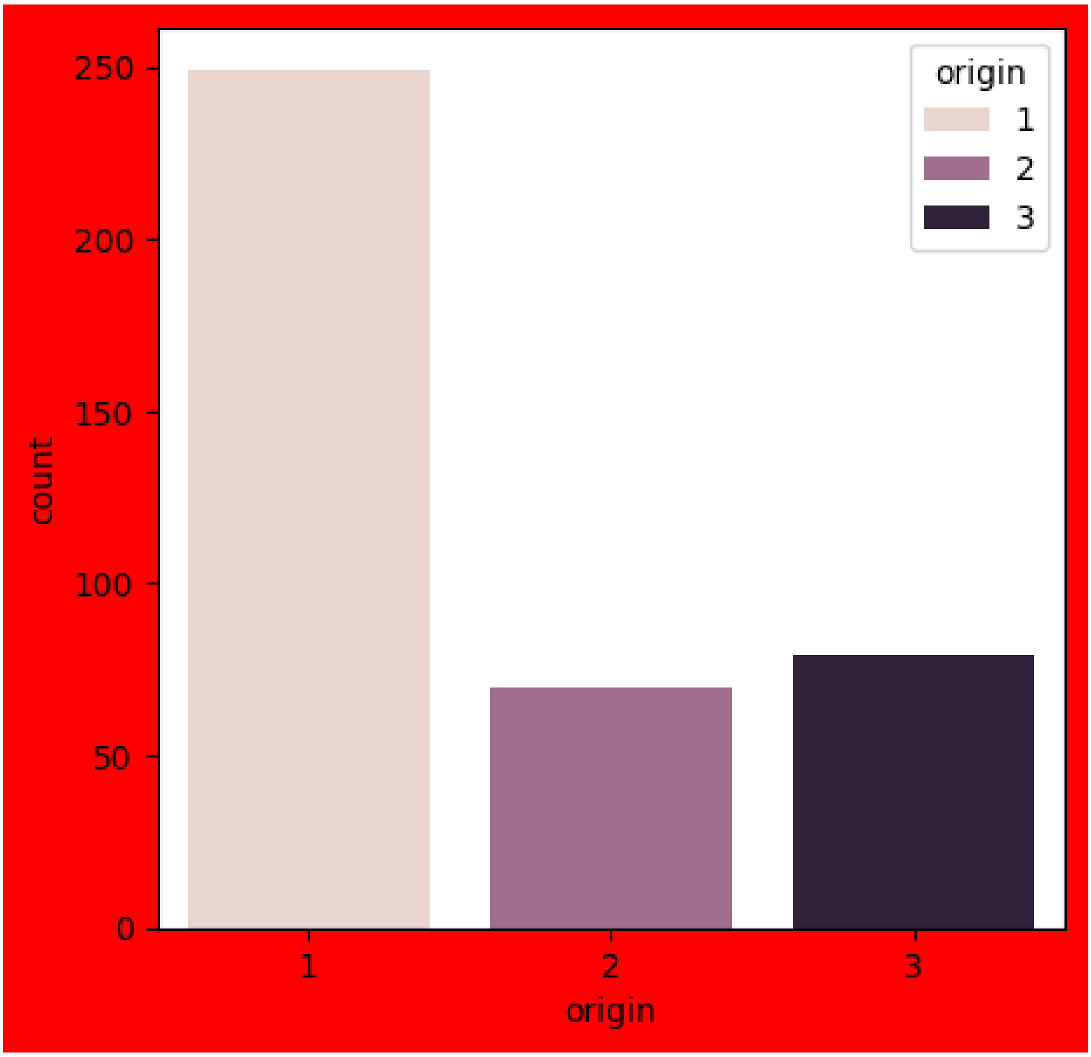
In [3]:

```
'''1.what is the distribution of MPG in the dataset'''
h=data.mpg.values.tolist()
fig=plt.figure(figsize=(6,5),facecolor='green')
axes=plt.axes(facecolor='orange')
sns.histplot(h,color='red',kde=True,label='mpg')
plt.xlabel('mpg')
plt.title('mpg vs count ')
plt.legend()
plt.show()
```



In [4]:

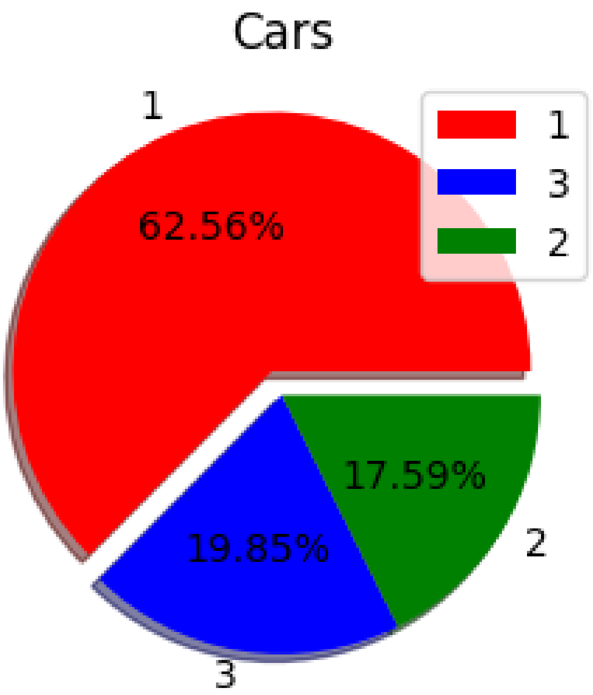
```
'''2.how many cars are from each origin '''
plt.figure(figsize=(5,5),facecolor='red')
sns.countplot(x=data['origin'],hue=data['origin'])
plt.show()
```



In [40]:

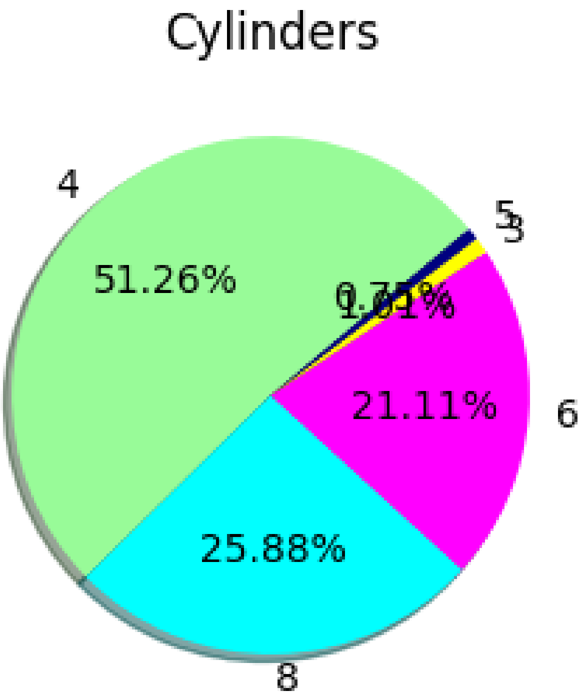
```
'''3.what is the proportion of cars from each origin '''
plt.figure(figsize=(4,3))

plt.pie(x=data.origin.value_counts().tolist(),colors=['red','blu
       ,autopct='%1.2f%%',explode=[0.1,0,0],shadow=True)
plt.legend()
plt.title('Cars')
plt.show()
```



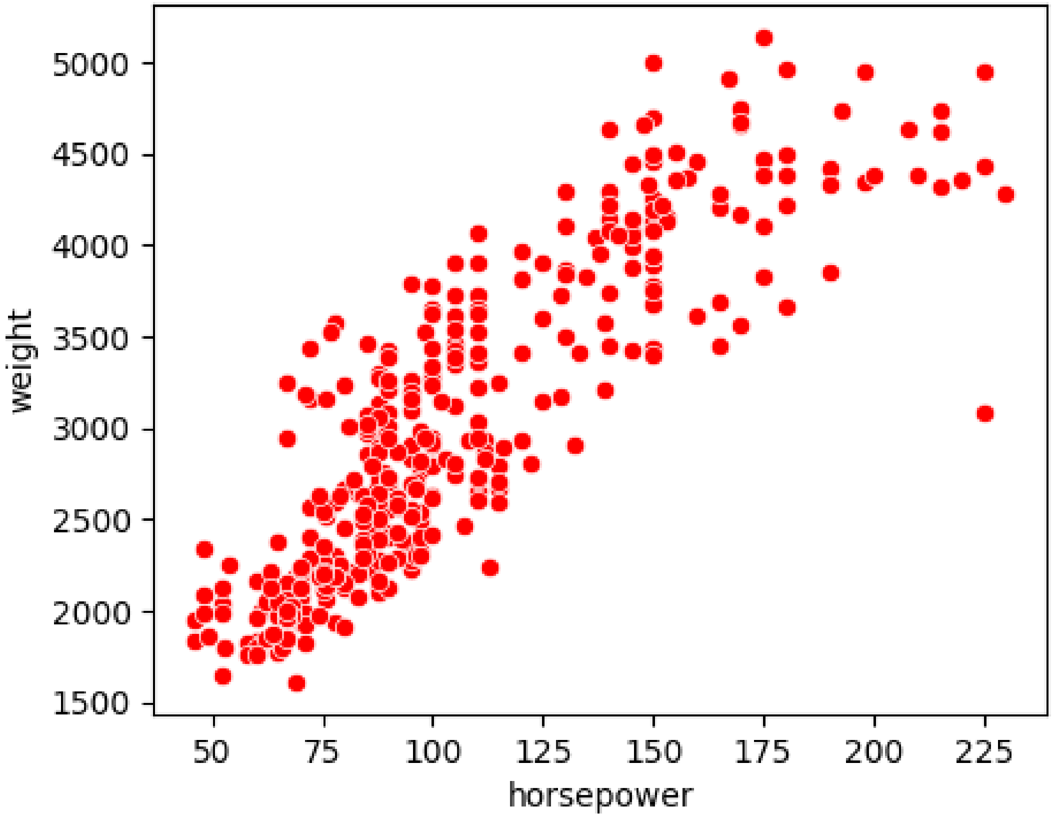
In [6]:

```
'''4.what is the proportion of cars from each cylinder'''
plt.figure(figsize=(10,3),facecolor='white')
plt.axes(facecolor='orange')
plt.pie(data.cylinders.value_counts().tolist(),labels=data.cylin
plt.title('Cylinders')
plt.show()
```



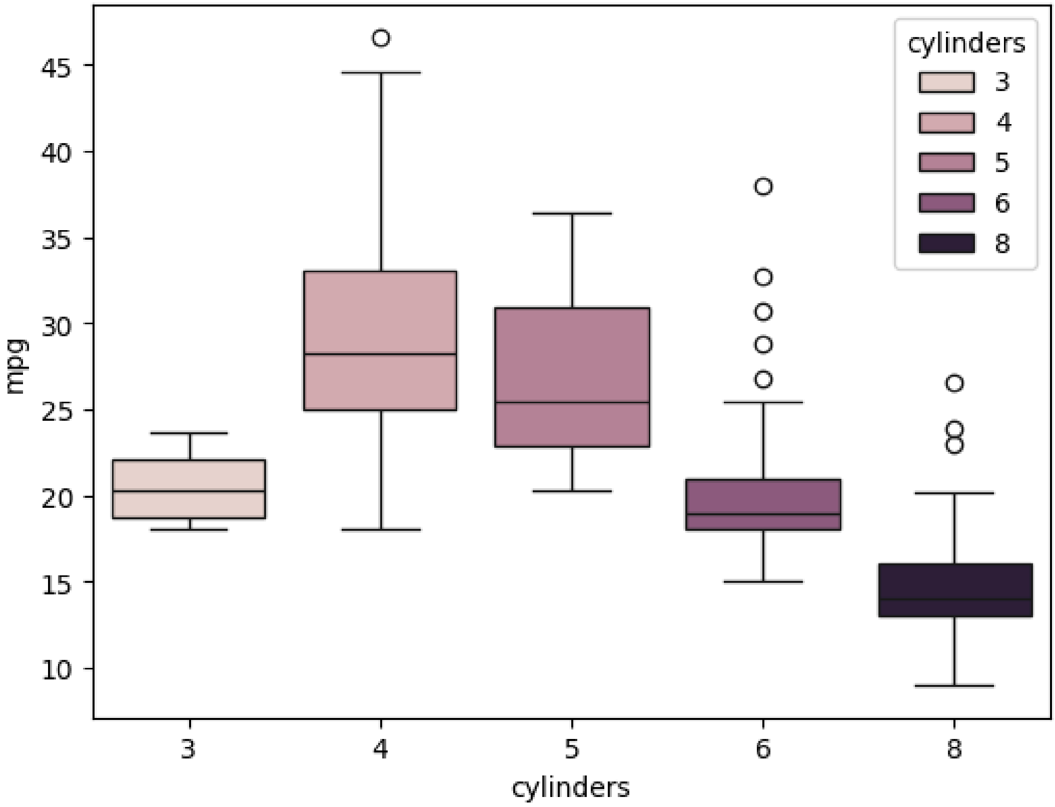
In [7]:

```
'''5.'''
plt.figure(figsize=(5,4))
data['horsepower']=pd.to_numeric(data['horsepower'],errors='coer
data[['horsepower','weight']].corr()
sns.scatterplot(x=data['horsepower'],y=data['weight'],color='red
plt.show()
```



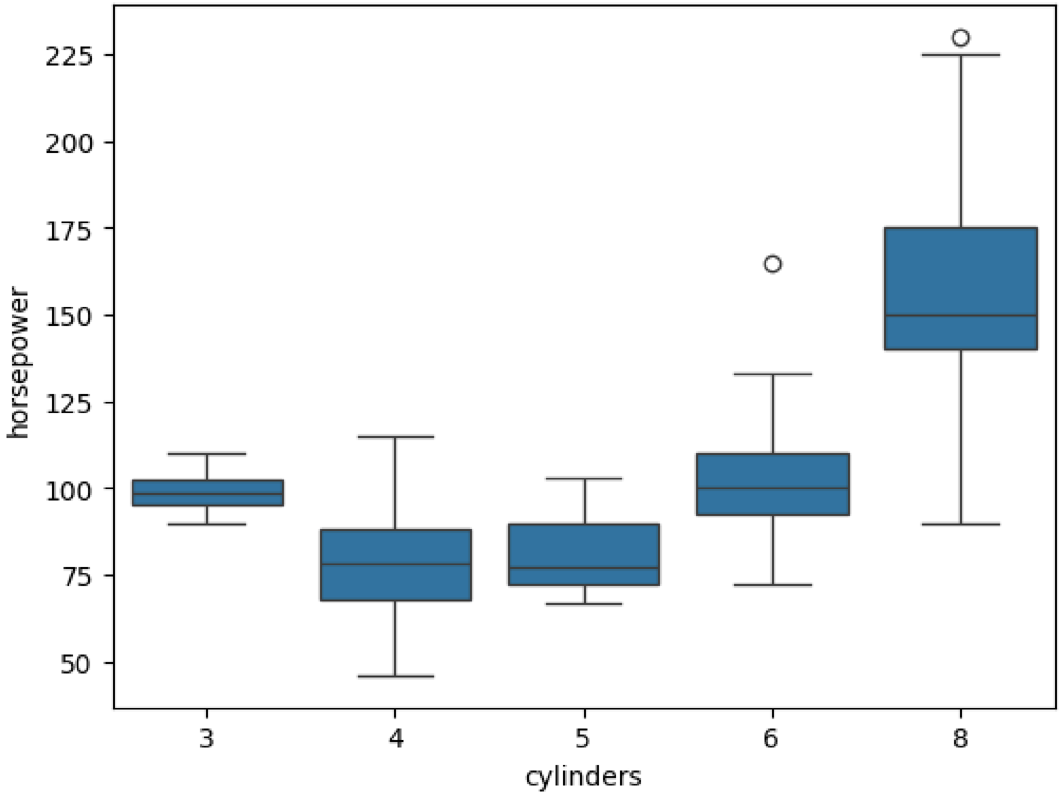
In [35]:

```
'''6.How does MPG vary across different numbers of cylinders?(box
sns.boxplot(data=data,x='cylinders',y='mpg',hue='cylinders')
plt.show()
```



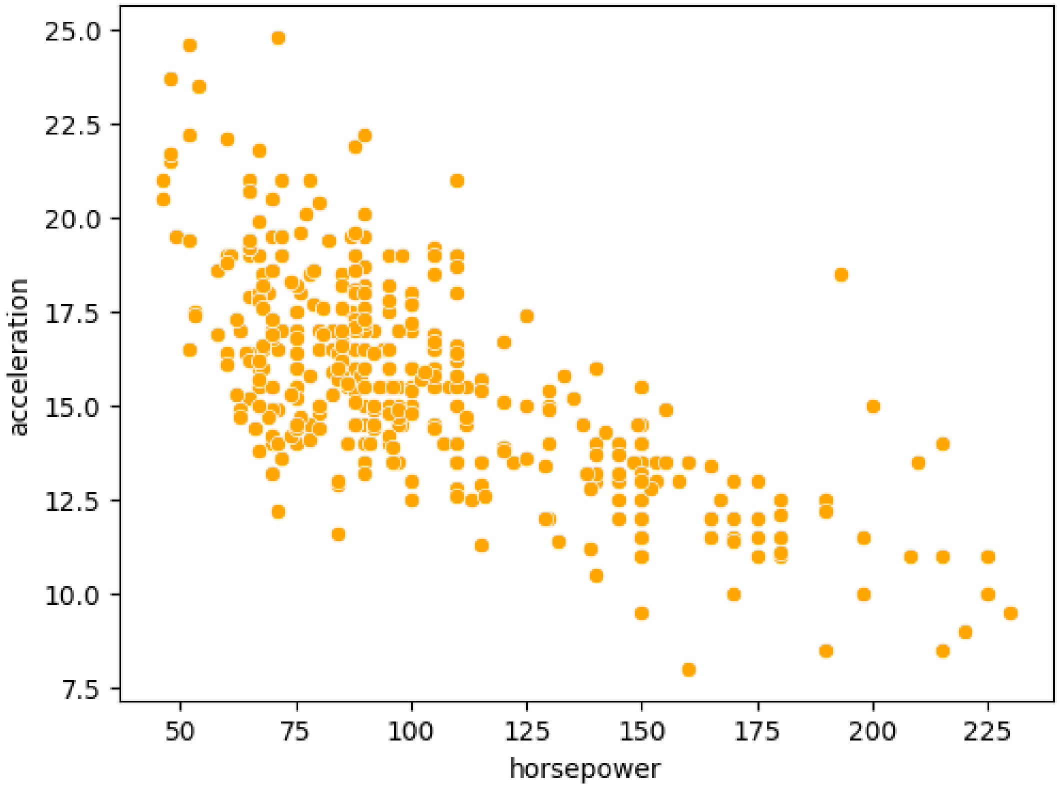
In [18]:

```
'''7.How does horsepower vary across different numbers of cylind
sns.boxplot(data=data,x='cylinders',y='horsepower')
plt.show()
```



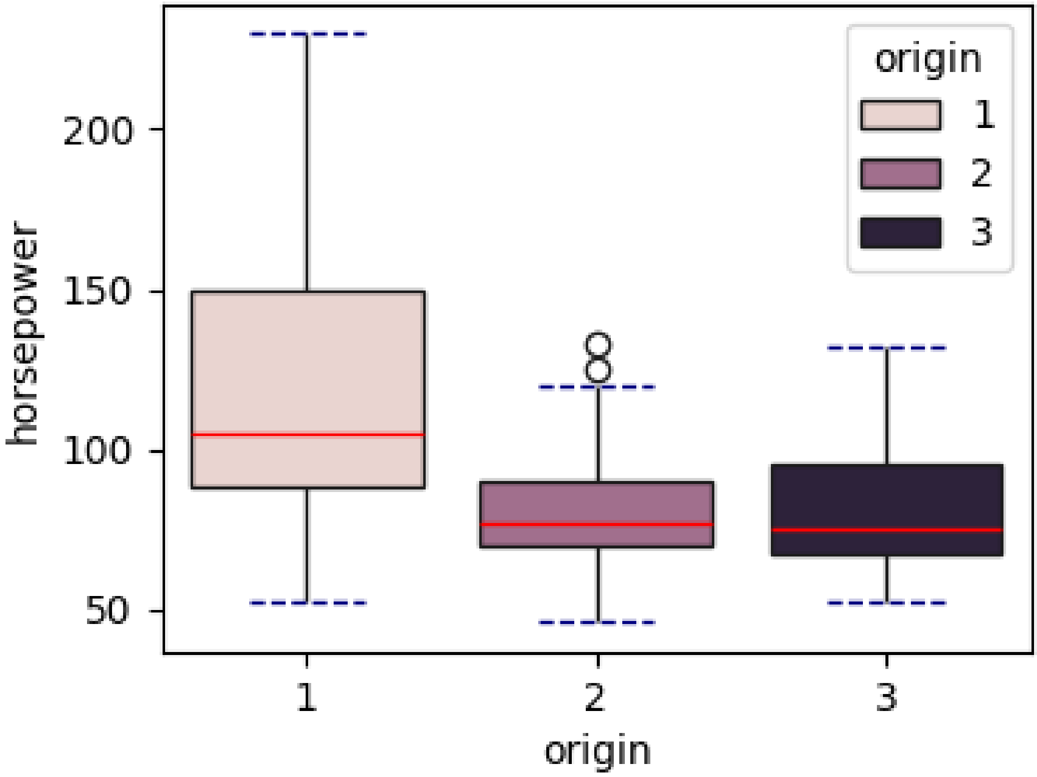
In [19]:

```
'''8.What are the relationships among horsepower,and acceleratio
sns.scatterplot(data=data,x='horsepower',y='acceleration',color=
plt.show()
```



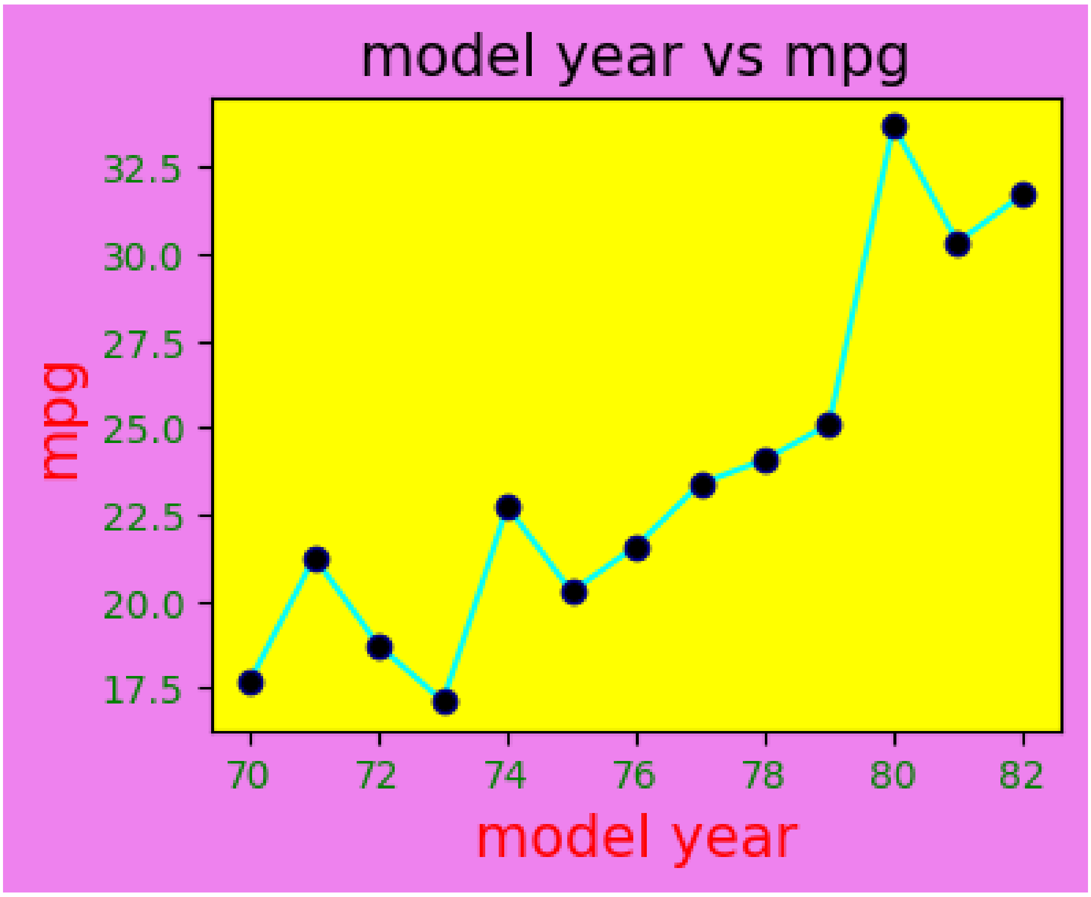
In [11]:

```
'''9.How does horsepower vary by origin?  (boxplot)'''
plt.figure(figsize=(4,3))
sns.boxplot(data=data,x='origin',y='horsepower',hue='origin',med
plt.show()
```



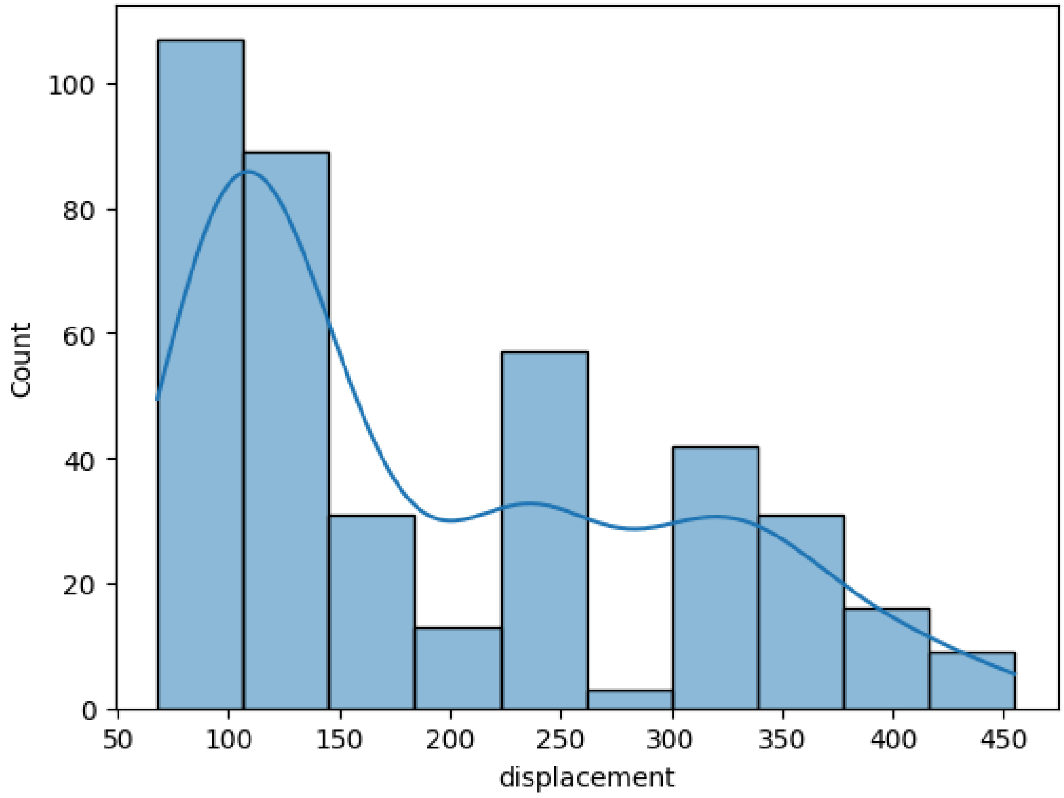
In [12]:

```
'''10.What is trend of MPG across different model years?(using l
year_mpg = data.groupby('model year')['mpg'].mean()
plt.figure(figsize=(4,3),facecolor='violet')
plt.axes(facecolor='yellow')
plt.plot(year_mpg.index, year_mpg.values,color='cyan',marker='o'
plt.xlabel('model year',color='red',size=15)
plt.ylabel('mpg',color='red',size=15)
plt.title('model year vs mpg',color='black',size=15)
plt.xticks(color='green')
plt.yticks(color='green')
plt.show()
```



In [13]:

```
'''11.What is the distribution of displacement in the dataset?(u
sns.histplot(data=data,x='displacement',kde=True,)
plt.show()
```



In [14]:

```
'''12.How does the weiht affects the mpg? (using correlation)'''
data[['weight','mpg']].corr()
```

Out[14]:

	weight	mpg
weight	1.000000	-0.831741
mpg	-0.831741	1.000000

In [15]:

```
'''13.What relationships can be inferred between MPG,horsepower a
data[['horsepower','mpg','weight']].corr()
```

Out[15]:

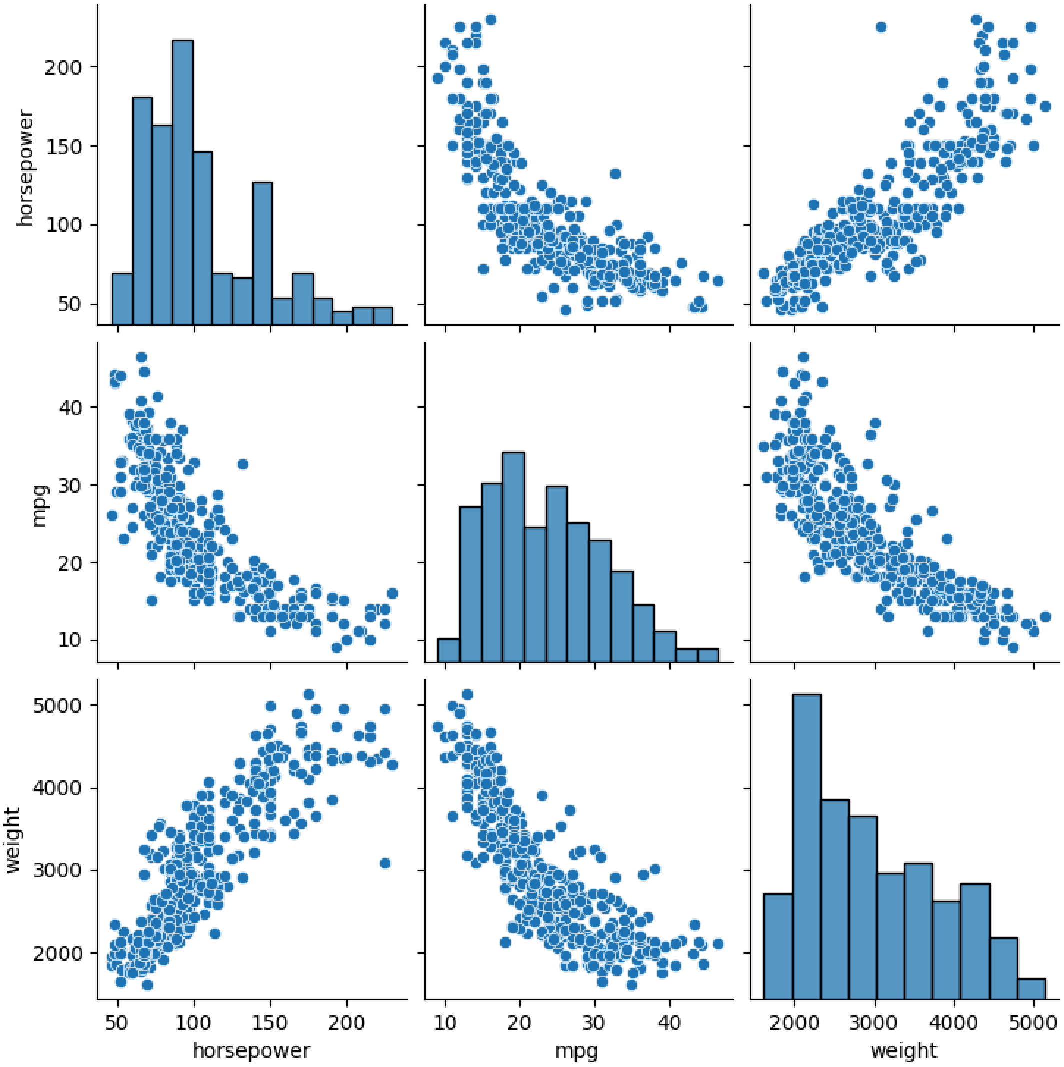
	horsepower	mpg	weight
horsepower	1.000000	-0.778427	0.864538
mpg	-0.778427	1.000000	-0.831741
weight	0.864538	-0.831741	1.000000

In [16]:

```
sns.pairplot(data[['horsepower','mpg','weight']])
```

Out[16]:

```
<seaborn.axisgrid.PairGrid at 0x1fc000433d0>
```

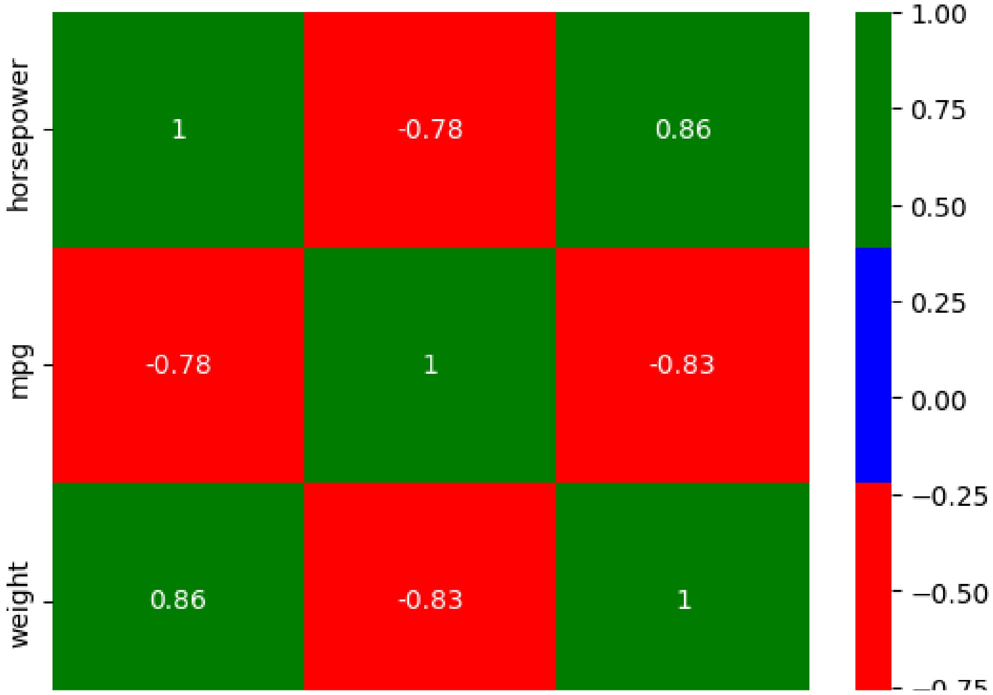


In [41]:

```
sns.heatmap(data[['horsepower','mpg','weight']].corr(),annot=Tru
```

Out[41]:

```
<Axes: >
```


horsepower
mpg
weight