

```
[ ] import numpy as np
import pandas as pd
```

```
[ ] import os
for dirname, _, filenames in os.walk('data'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

```
[ ] data = pd.read_csv('data/cars.csv')
```

```
[ ] df.head()
```



	Car Make	Car Model	Year	Engine Size (L)
0	Porsche	911	2022	3.0
1	Lamborghini	Huracan	2021	5.2
2	Ferrari	488 GTB	2022	3.9
3	Audi	R8	2022	5.2
4	McLaren	720S	2021	3.9

```
[ ] plt.hist(df['Price (in USD)'], color='blue')
plt.title('Range of prices')
plt.xlabel('Horsepower')
plt.ylabel('Torque')
```



```
Text(0, 0.5, 'Torque')
```



```
[ ] data.shape
```



```
(1007, 8)
```

```
[ ] data.info()
```



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1007 entries, 0 to 1006
Data columns (total 8 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Car Make    1007         object
```

```
[ ] data.info()
```



```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1007 entries, 0 to 1006
Data columns (total 8 columns):
 #   Column                                  Non-Null Count  Dtype  
---  -
 0   Car Make                               1007  object 
 1   Car Model                              1007  object 
 2   Year                                   1007  object 
 3   Engine Size (L)                        997    float64
 4   Horsepower                             1007  float64
 5   Torque (lb-ft)                         1007  float64
 6   0-60 MPH Time (seconds)                1007  float64
 7   Price (in USD)                         1007  float64
dtypes: int64(1), object(7)
memory usage: 63.1+ KB
```

```
[ ] data['Car Make'] = data['Car Make'].str.lower()
data['Car Model'] = data['Car Model'].str.lower()
data['Year'] = data['Year'].astype(int)
data['Engine Size (L)'] = data['Engine Size (L)'].str.replace(' ', '')
data['Horsepower'] = data['Horsepower'].str.replace(' ', '')
data['Torque (lb-ft)'] = data['Torque (lb-ft)'].str.replace(' ', '')
data['0-60 MPH Time (seconds)'] = data['0-60 MPH Time (seconds)'].str.replace(' ', '')
```

```
[ ] data['Price (in USD)'] = data['Price (in USD)'].str.replace(' ', '')
data['Price (in USD)'] = data['Price (in USD)'].str.replace(' ', '')
```

```
[ ] data.head()
```

```
[ ] data.head()
```



	Car Make	Car Model	Year	Engine Size (L)
0	Porsche	911	2022	
1	Lamborghini	Huracan	2021	5.2
2	Ferrari	488 GTB	2022	3.9
3	Audi	R8	2022	5.2
4	McLaren	720S	2021	

```
[ ] data.isna().sum()
```



	0
Car Make	0
Car Model	0
Year	0
Engine Size (L)	0
Horsepower	0



```
[ ] data.isna().sum()
```

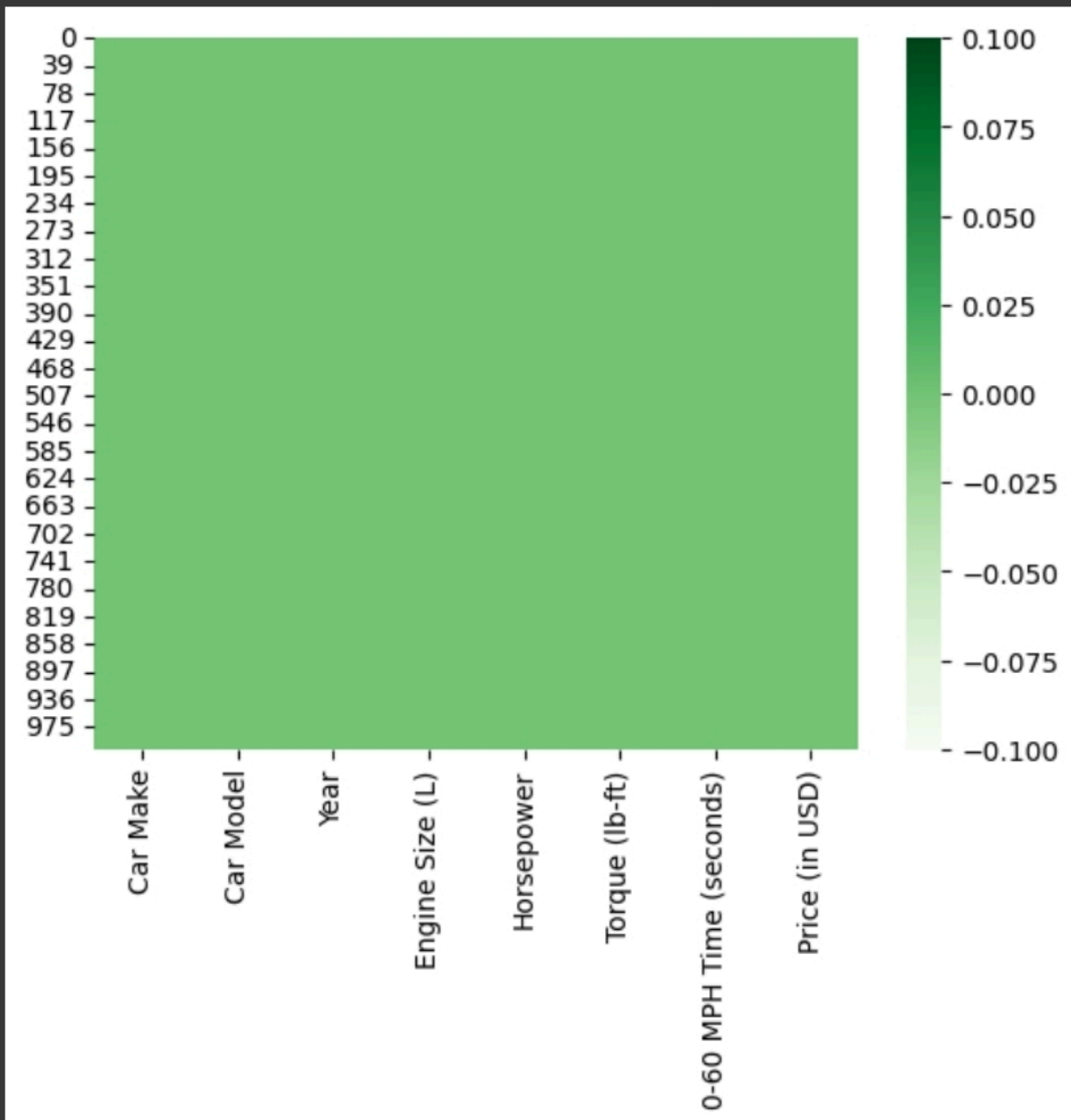


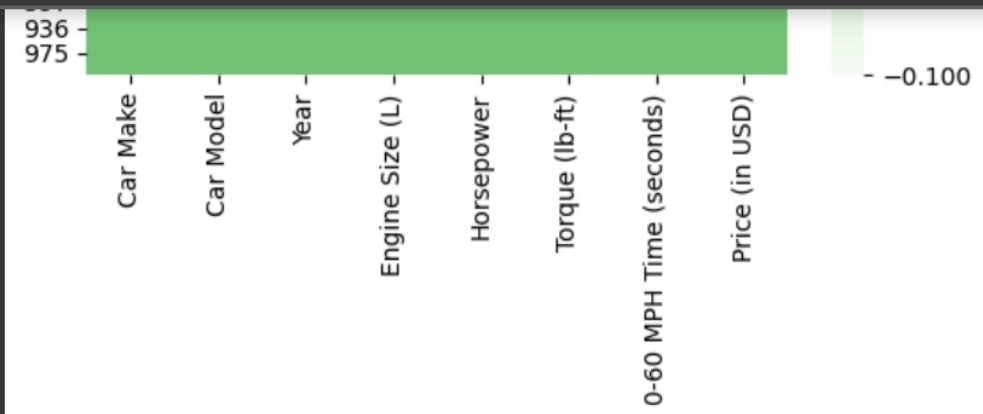
	0
Car Make	0
Car Model	0
Year	0
Engine Size (L)	0
Horsepower	0
Torque (lb-ft)	0
0-60 MPH Time (seconds)	0
Price (in USD)	0

dtype: int64



```
sns.heatmap(data.isna(), cmap='Gre  
plt.show()
```





```
[ ] sns.countplot(data=data, x='Year')  
plt.show()
```

