```
pip install ucimlrepo
```

```
Requirement already satisfied: ucimlrepo in /usr/local/lib/python3.10/dist-packages (0.0.6)
```

```
from ucimlrepo import fetch_ucirepo

# fetch dataset
automobile = fetch_ucirepo(id=10)

# data (as pandas dataframes)
X = automobile.data.features
y = automobile.data.targets

# metadata
print(automobile.metadata)

# variable information
print(automobile.variables)
```

```
{'uci_id': 10, 'name': 'Automobile', 'repository_url': 'https://archive.ics.uci.edu/dataset/10/automobile', 'data_url': 'https://archive.ics.uci.edu/static/public/10/data.csv', 'abstr
                  name     role         type demographic  \
0                price  Feature   Continuous        None
1          highway-mpg  Feature   Continuous        None
2             city-mpg  Feature   Continuous        None
3             peak-rpm  Feature   Continuous        None
4           horsepower  Feature   Continuous        None
5    compression-ratio  Feature   Continuous        None
6               stroke  Feature   Continuous        None
7                 bore  Feature   Continuous        None
8          fuel-system  Feature  Categorical        None
9          engine-size  Feature   Continuous        None
10   num-of-cylinders  Feature      Integer        None
11        engine-type  Feature  Categorical        None
12        curb-weight  Feature   Continuous        None
13             height  Feature   Continuous        None
14              width  Feature   Continuous        None
15             length  Feature   Continuous        None
16         wheel-base  Feature   Continuous        None
17    engine-location  Feature       Binary        None
18       drive-wheels  Feature  Categorical        None
19         body-style  Feature  Categorical        None
20       num-of-doors  Feature      Integer        None
21         aspiration  Feature       Binary        None
22          fuel-type  Feature       Binary        None
23               make  Feature  Categorical        None
24  normalized-losses  Feature   Continuous        None
25          symboling   Target      Integer        None

                                    description units missing_values
0                     continuous from 5118 to 45400  None            yes
1                          continuous from 16 to 54  None             no
2                          continuous from 13 to 49  None             no
3                        continuous from 4150 to 6600  None            yes
4                           continuous from 48 to 288  None            yes
5                             continuous from 7 to 23  None             no
6                        continuous from 2.07 to 4.17  None            yes
7                        continuous from 2.54 to 3.94  None            yes
8        1bbl, 2bbl, 4bbl, idi, mfi, mpfi, spdi, spfi  None             no
9                           continuous from 61 to 326  None             no
10         eight, five, four, six, three, twelve, two  None             no
```

```
11            dohc, dohcv, l, ohc, ohcf, ohcv, rotor  None            no
12                      continuous from 1488 to 4066  None            no
13                      continuous from 47.8 to 59.8  None            no
14                      continuous from 60.3 to 72.3  None            no
15                    continuous from 141.1 to 208.1  None            no
16                       continuous from 86.6 120.9   None            no
17                                     front, rear    None            no
18                                   4wd, fwd, rwd    None            no
19     hardtop, wagon, sedan, hatchback, convertible  None            no
20                                       four, two    None           yes
21                                     std, turbo     None            no
22                                     diesel, gas    None            no
23  alfa-romero, audi, bmw, chevrolet, dodge, hond...  None            no
24                      continuous from 65 to 256     None           yes
25                         -3, -2, -1, 0, 1, 2, 3     None            no
```

```
!pip install hvplot
```

```
Requirement already satisfied: hvplot in /usr/local/lib/python3.10/dist-packages (0.9.2)
Requirement already satisfied: bokeh>=1.0.0 in /usr/local/lib/python3.10/dist-packages (from hvplot) (3.3.4)
Requirement already satisfied: colorcet>=2 in /usr/local/lib/python3.10/dist-packages (from hvplot) (3.1.0)
Requirement already satisfied: holoviews>=1.11.0 in /usr/local/lib/python3.10/dist-packages (from hvplot) (1.17.1)
Requirement already satisfied: pandas in /usr/local/lib/python3.10/dist-packages (from hvplot) (2.0.3)
Requirement already satisfied: numpy>=1.15 in /usr/local/lib/python3.10/dist-packages (from hvplot) (1.25.2)
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from hvplot) (24.0)
Requirement already satisfied: panel>=0.11.0 in /usr/local/lib/python3.10/dist-packages (from hvplot) (1.3.8)
Requirement already satisfied: param<3.0,>=1.12.0 in /usr/local/lib/python3.10/dist-packages (from hvplot) (2.1.0)
Requirement already satisfied: Jinja2>=2.9 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (3.1.3)
Requirement already satisfied: contourpy>=1 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (1.2.1)
Requirement already satisfied: pillow>=7.1.0 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (9.4.0)
Requirement already satisfied: PyYAML>=3.10 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (6.0.1)
Requirement already satisfied: tornado>=5.1 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (6.3.3)
Requirement already satisfied: xyzservices>=2021.09.1 in /usr/local/lib/python3.10/dist-packages (from bokeh>=1.0.0->hvplot) (2024.4.0)
Requirement already satisfied: pyviz-comms>=0.7.4 in /usr/local/lib/python3.10/dist-packages (from holoviews>=1.11.0->hvplot) (3.0.2)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.10/dist-packages (from pandas->hvplot) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas->hvplot) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas->hvplot) (2024.1)
Requirement already satisfied: markdown in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (3.6)
Requirement already satisfied: markdown-it-py in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (3.0.0)
Requirement already satisfied: linkify-it-py in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (2.0.3)
Requirement already satisfied: mdit-py-plugins in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (0.4.0)
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (2.31.0)
Requirement already satisfied: tqdm>=4.48.0 in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (4.66.2)
Requirement already satisfied: bleach in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (6.1.0)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.10/dist-packages (from panel>=0.11.0->hvplot) (4.11.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2>=2.9->bokeh>=1.0.0->hvplot) (2.1.5)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.8.2->pandas->hvplot) (1.16.0)
Requirement already satisfied: webencodings in /usr/local/lib/python3.10/dist-packages (from bleach->panel>=0.11.0->hvplot) (0.5.1)
Requirement already satisfied: uc-micro-py in /usr/local/lib/python3.10/dist-packages (from linkify-it-py->panel>=0.11.0->hvplot) (1.0.3)
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py->panel>=0.11.0->hvplot) (0.1.2)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->panel>=0.11.0->hvplot) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->panel>=0.11.0->hvplot) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->panel>=0.11.0->hvplot) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->panel>=0.11.0->hvplot) (2024.2.2)
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import hvplot.pandas

from sklearn.model_selection import train_test_split
from sklearn.impute import SimpleImputer
from sklearn import metrics
from sklearn. linear_model import LinearRegression

%matplotlib inline


x1 = pd.DataFrame(X)
y1 = pd.DataFrame(y)
df = pd.concat([x1, y1], axis= 1)
df
```

	price	highway-mpg	city-mpg	peak-rpm	horsepower	compression-ratio	stroke	bore	fuel-system	engine-size	...	wheel-base	engine-location	dr wh
0	13495.0	27	21	5000.0	111.0	9.0	2.68	3.47	mpfi	130	...	88.6	front	
1	16500.0	27	21	5000.0	111.0	9.0	2.68	3.47	mpfi	130	...	88.6	front	
2	16500.0	26	19	5000.0	154.0	9.0	3.47	2.68	mpfi	152	...	94.5	front	
3	13950.0	30	24	5500.0	102.0	10.0	3.40	3.19	mpfi	109	...	99.8	front	
4	17450.0	22	18	5500.0	115.0	8.0	3.40	3.19	mpfi	136	...	99.4	front	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
200	16845.0	28	23	5400.0	114.0	9.5	3.15	3.78	mpfi	141	...	109.1	front	
201	19045.0	25	19	5300.0	160.0	8.7	3.15	3.78	mpfi	141	...	109.1	front	
202	21485.0	23	18	5500.0	134.0	8.8	2.87	3.58	mpfi	173	...	109.1	front	
203	22470.0	27	26	4800.0	106.0	23.0	3.40	3.01	idi	145		109.1	front	

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 205 entries, 0 to 204
Data columns (total 26 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   price              201 non-null    float64
 1   highway-mpg        205 non-null    int64
 2   city-mpg           205 non-null    int64
 3   peak-rpm           203 non-null    float64
 4   horsepower         203 non-null    float64
 5   compression-ratio  205 non-null    float64
 6   stroke             201 non-null    float64
 7   bore               201 non-null    float64
 8   fuel-system        205 non-null    object
 9   engine-size        205 non-null    int64
 10  num-of-cylinders   205 non-null    int64
```

```
 11  engine-type        205 non-null    object
 12  curb-weight        205 non-null    int64
 13  height             205 non-null    float64
 14  width              205 non-null    float64
 15  length             205 non-null    float64
 16  wheel-base         205 non-null    float64
 17  engine-location    205 non-null    object
 18  drive-wheels       205 non-null    object
 19  body-style         205 non-null    object
 20  num-of-doors       203 non-null    float64
 21  aspiration         205 non-null    object
 22  fuel-type          205 non-null    object
 23  make               205 non-null    object
 24  normalized-losses  164 non-null    float64
 25  symboling          205 non-null    int64
dtypes: float64(12), int64(6), object(8)
memory usage: 41.8+ KB
```

```
df.fillna(method='ffill', inplace=True)
df.fillna(method='bfill', inplace=True)
```

```
df
```

	price	highway-mpg	city-mpg	peak-rpm	horsepower	compression-ratio	stroke	bore	fuel-system	engine-size	...	wheel-base	engine-location	dr wh
0	13495.0	27	21	5000.0	111.0	9.0	2.68	3.47	mpfi	130	...	88.6	front	
1	16500.0	27	21	5000.0	111.0	9.0	2.68	3.47	mpfi	130	...	88.6	front	
2	16500.0	26	19	5000.0	154.0	9.0	3.47	2.68	mpfi	152	...	94.5	front	
3	13950.0	30	24	5500.0	102.0	10.0	3.40	3.19	mpfi	109	...	99.8	front	
4	17450.0	22	18	5500.0	115.0	8.0	3.40	3.19	mpfi	136	...	99.4	front	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	
200	16845.0	28	23	5400.0	114.0	9.5	3.15	3.78	mpfi	141	...	109.1	front	
201	19045.0	25	19	5300.0	160.0	8.7	3.15	3.78	mpfi	141	...	109.1	front	
202	21485.0	23	18	5500.0	134.0	8.8	2.87	3.58	mpfi	173	...	109.1	front	
203	22470.0	27	26	4800.0	106.0	23.0	3.40	3.01	idi	145		109.1	front	

```
df.describe()
```

	price	highway-mpg	city-mpg	peak-rpm	horsepower	compression-ratio	stroke	bore	engine-size	cy
count	205.000000	205.000000	205.000000	205.000000	205.000000	205.000000	205.000000	205.000000	205.000000	20
mean	13312.712195	30.751220	25.219512	5131.463415	106.048780	10.142537	3.253366	3.324878	126.907317	
std	8102.472461	6.886443	6.542142	480.933330	43.468803	3.972040	0.313937	0.273049	41.642693	
min	5118.000000	16.000000	13.000000	4150.000000	48.000000	7.000000	2.070000	2.540000	61.000000	
25%	7775.000000	25.000000	19.000000	4800.000000	70.000000	8.600000	3.110000	3.130000	97.000000	
50%	10295.000000	30.000000	24.000000	5200.000000	95.000000	9.000000	3.290000	3.310000	120.000000	
75%	16503.000000	34.000000	30.000000	5500.000000	120.000000	9.400000	3.410000	3.580000	141.000000	
max	45400.000000	54.000000	49.000000	6600.000000	288.000000	23.000000	4.170000	3.940000	326.000000	1

```
df.columns
```

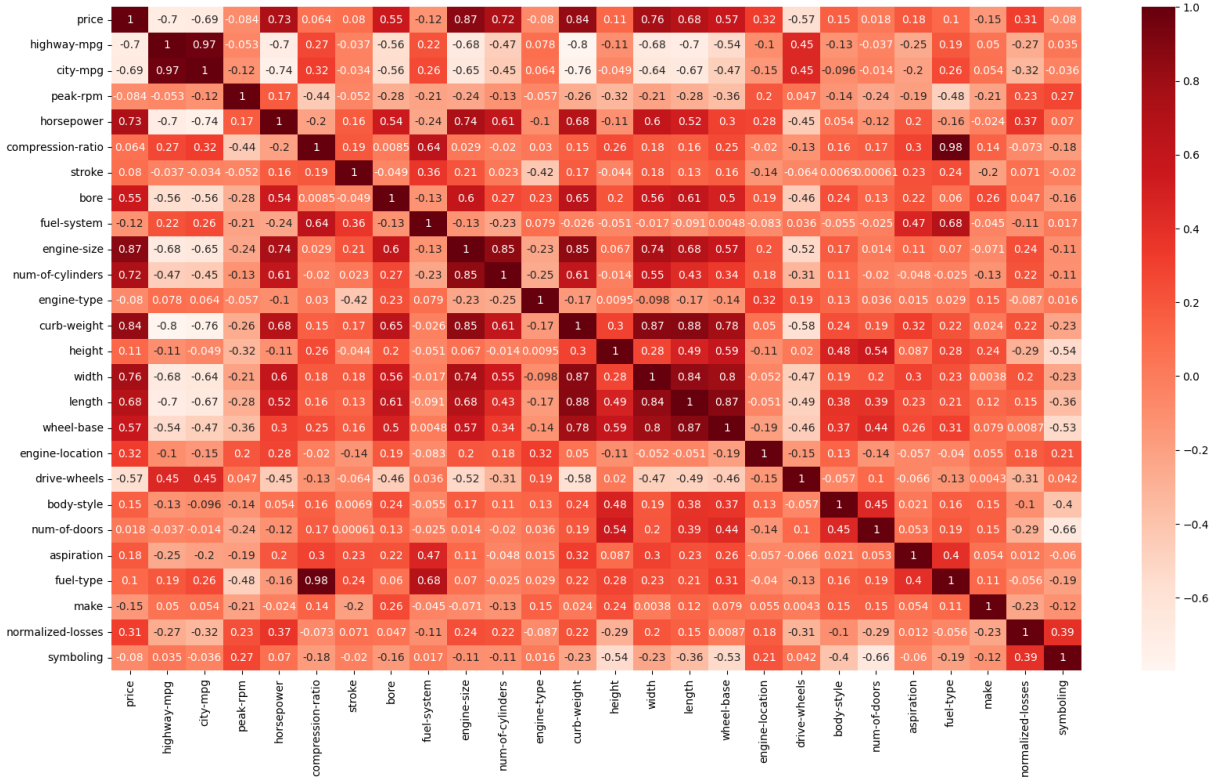
```
Index(['price', 'highway-mpg', 'city-mpg', 'peak-rpm', 'horsepower',
       'compression-ratio', 'stroke', 'bore', 'fuel-system', 'engine-size',
       'num-of-cylinders', 'engine-type', 'curb-weight', 'height', 'width',
       'length', 'wheel-base', 'engine-location', 'drive-wheels', 'body-style',
       'num-of-doors', 'aspiration', 'fuel-type', 'make', 'normalized-losses',
       'symboling'],
      dtype='object')
```

```
def objecttonumerical(dFrame, List):
    if dFrame[List].dtypes == 'object':
        cat_values = dFrame[List].unique()
        range_values = range(1, len(cat_values) + 1)
        map = dict(zip(cat_values, range_values))
        print(f"{List}:", map)
        dFrame[List] = dFrame[List].map(map)
    return dFrame

for i in df.select_dtypes(include=['object']).columns:
      objecttonumerical(df, i)
```

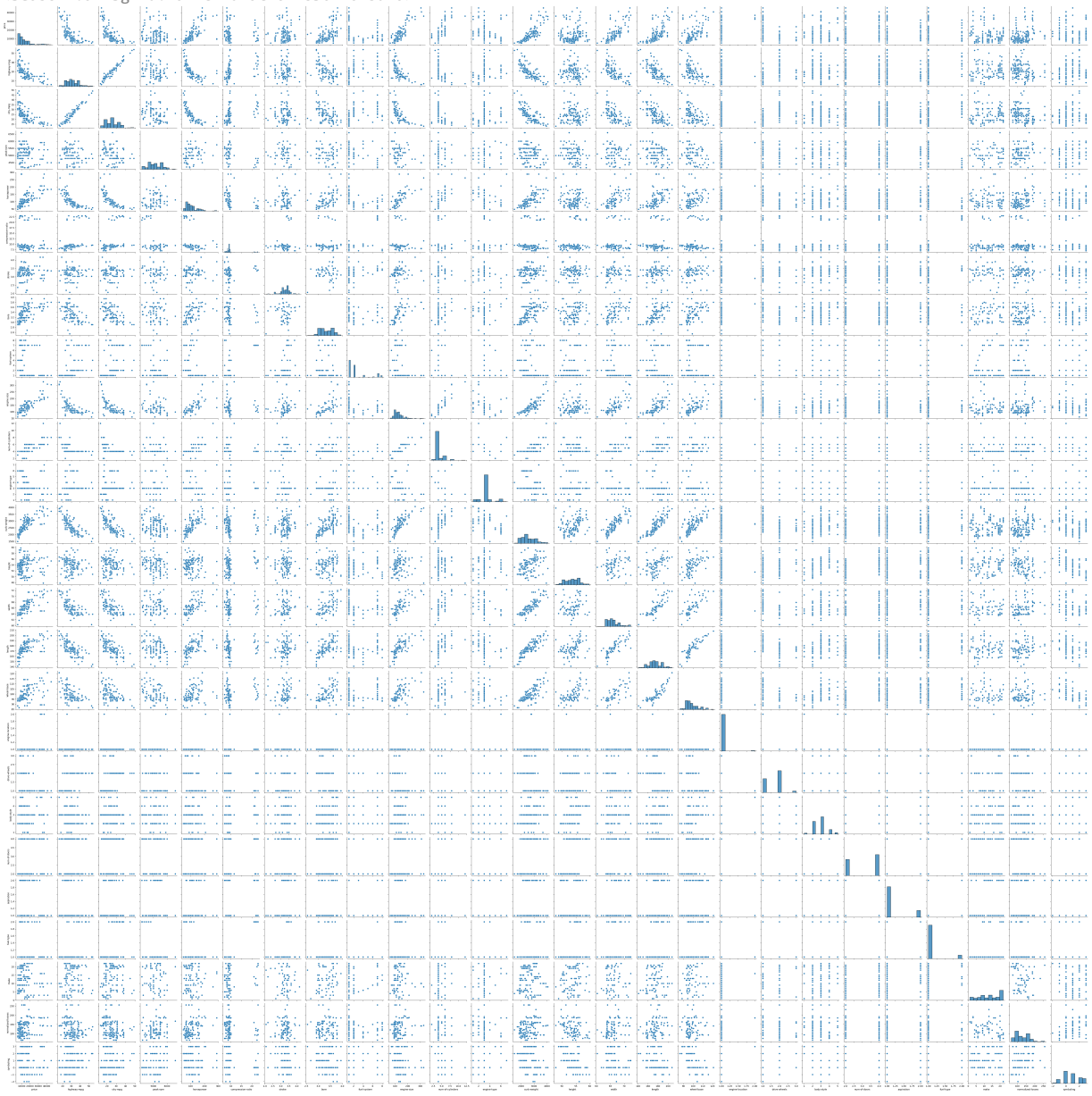
```
plt.figure(figsize=(20, 11))
sns.heatmap(df.corr(), annot=True, cmap='Reds' )
```

```
<Axes: >
```




```
sns.pairplot(df)
```

```
<seaborn.axisgrid.PairGrid at 0x78504901bdf0>
```



```
X=df.drop('price', axis=1)

y=df['highway-mpg']

print("X=",X.shape,"\ny=", y.shape)
```

```
X= (205, 25)
y= (205,)
```

```
print("X=", X.shape, "\ny=", y.shape)
```

```
X= (205, 25)
y= (205,)
```

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=101)
```

```
X_train.shape
```

```
(143, 25)
```

```
X_test.shape
```

```
(62, 25)
```

```
model = LinearRegression()
```

```
model.fit(X_train, y_train)
```

```
▾ LinearRegression
LinearRegression()
```

```
pd.DataFrame(model.coef_, X.columns, columns=['Coedicients'])
```

	Coedicients
highway-mpg	1.000000e+00
city-mpg	-1.742799e-15
peak-rpm	-2.029357e-17
horsepower	3.374986e-16
compression-ratio	-7.763704e-16
stroke	-5.037228e-17
bore	7.226780e-17
fuel-system	-3.664579e-17
engine-size	-7.252997e-16
num-of-cylinders	-1.855127e-16
engine-type	7.640085e-16
curb-weight	3.213432e-17
height	-1.042719e-16
width	-2.165628e-16
length	-3.447523e-16
wheel-base	-6.243697e-16
engine-location	1.083722e-16
drive-wheels	-8.005516e-17
body-style	1.498610e-17
num-of-doors	-3.766393e-16
aspiration	1.184186e-17
fuel-type	-1.919084e-15
make	2.828787e-15
normalized-losses	-3.533726e-17
symboling	3.559951e-16

```
y_pred = model.predict(X_test)
```

```
MAE= metrics.mean_absolute_error(y_test, y_pred)
MSE=metrics.mean_squared_error(y_test, y_pred)
RMSE= np.sqrt(MSE)
```

```
MAE
```

```
2.7390276426881282e-14
```

```
MSE
```

```
9.75948175388266e-28
```

```
RMSE
```

```
3.124016925991705e-14
```

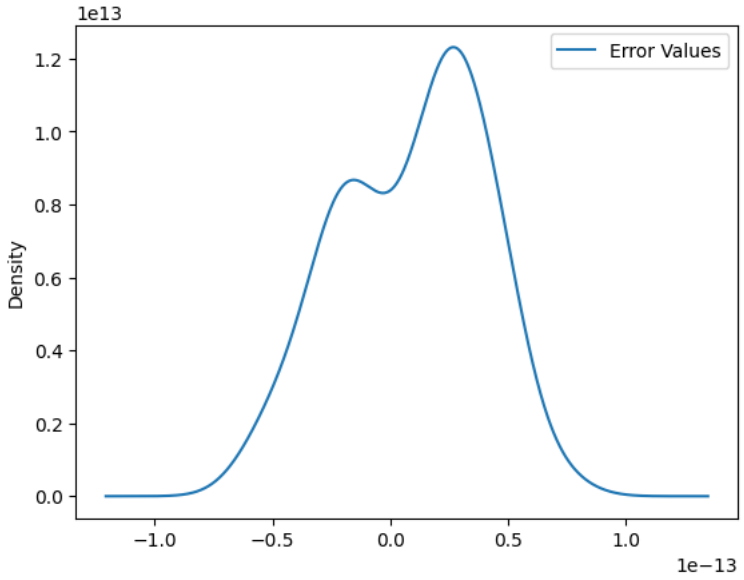
```
df['highway-mpg'].mean()
```

```
30.75121951219512
```

```
test_residual= y_test - y_pred
```

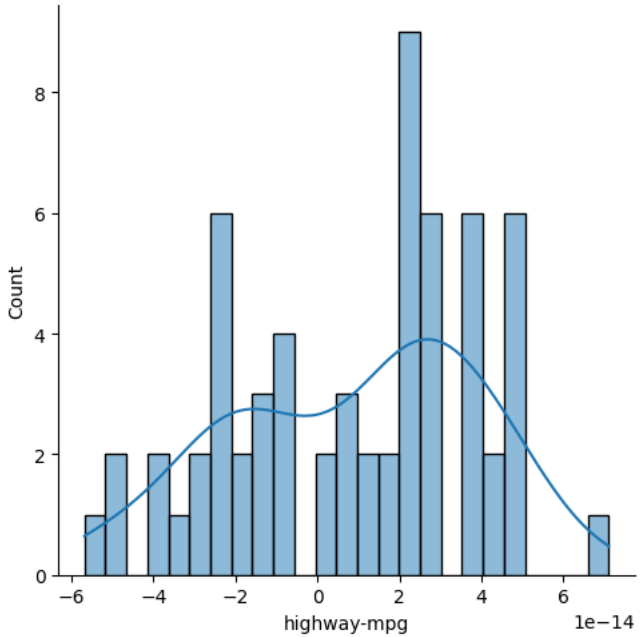
```
pd.DataFrame({'Error Values': (test_residual)}).plot.kde()
```

```
<Axes: ylabel='Density'>
```




```
sns.displot(test_residual, bins=25, kde=True)
```

```
<seaborn.axisgrid.FacetGrid at 0x78502f0c0220>
```




```
sns.scatterplot(x=y_test, y=test_residual)

plt.axhline(y=0, color='r', ls='--')
```

```
<matplotlib.lines.Line2D at 0x78501eb43ca0>
```