

In [11]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn import preprocessing, svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

In []:

```
df=pd.read_csv(r"C:\Users\konduru jeevana priy\Downloads\bottle.csv.zip")
df
```

In []:

```
df=df[['Salnty', 'T_degC']]
```

In []:

```
df.columns=['Sal', 'Temp']
```

In []:

```
df.head(10)
```

In []:

```
sns.lmplot(x="Sal", y="Temp", data=df, order=2, ci=None)
```

In []:

```
df.describe()
```

In []:

```
df.info()
```

In []:

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

In []:

```
x=np.array(df['Sal']).reshape(-1,1)
y=np.array(df['Temp']).reshape(-1,1)
```

In []:

```
df.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
regr=LinearRegression()
regr.fit(x_train,y_train)
print(regr.score(x_test,y_test))
```

In []:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_pred,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

In []:

```
df500=df[:][:500]
sns.lmplot(x="Sal",y="Temp",data=df500,order=1,ci=None)
x=np.array(df500['Sal']).reshape(-1,1)
y=np.array(df500['Temp']).reshape(-1,1)
df500.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
```

In []:

```
regr=LinearRegression()
regr.fit(x_train,y_train)
print("Regression:",regr.score(x_test,y_test))
```

In []:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_test,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

In []:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_test,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
model=LinearRegression()
model.fit(x_train,y_train)
y_pred=model.predict(x_test)
r2=r2_score(y_test,y_pred)
print("R2 Score:",r2)
```

In []:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn import preprocessing, svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

In []:

```
df=pd.read_csv(r"C:\Users\konduru jeevana priy\Desktop\fiat500_VehicleSelection_Dataset.
df
```

In []:

```
df=df[['age_in_days', 'km']]
df.columns=['age', 'km']
```

In []:

```
df.head(10)
```

In []:

```
sns.lmplot(x="age", y="km", data=df, order=2, ci=None)
```

In []:

```
df.describe()
```

In []:

```
df.info()
```

In []:

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

In []:

```
x=np.array(df['age']).reshape(-1,1)
y=np.array(df['km']).reshape(-1,1)
```

In []:

```
df.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
regr=LinearRegression()
regr.fit(x_train,y_train)
print(regr.score(x_test,y_test))
```

In []:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_pred,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

In []:

```
df500=df[:][:500]
sns.lmplot(x="age",y="km",data=df500,order=1,ci=None)
x=np.array(df500['age']).reshape(-1,1)
y=np.array(df500['km']).reshape(-1,1)
df500.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
```

In []:

```
regr=LinearRegression()
regr.fit(x_train,y_train)
print("Regression:",regr.score(x_test,y_test))
```

In []:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_test,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

In []:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn import preprocessing,svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
```

In []:

```
df=pd.read_csv(r"C:\Users\konduru jeevana priy\Downloads\data.csv")
df
```

In []:

```
df=df[['sqft_living','sqft_lot']]
df.head(10)
```

In []:

```
sns.lmplot(x='sqft_living',y='sqft_lot',data=df,order=2,ci=None)
```

In [2]:

```
df.describe()
```

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[2], line 1
----> 1 df.describe()
```

NameError: name 'df' is not defined

In [3]:

```
df.info()
```

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[3], line 1
----> 1 df.info()
```

NameError: name 'df' is not defined

In [4]:

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

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[4], line 1
----> 1 df.fillna(method='ffill',inplace=True)
```

NameError: name 'df' is not defined

In [5]:

```
x=np.array(df['sqft_living']).reshape(-1,1)
y=np.array(df['sqft_lot']).reshape(-1,1)
```

-
NameError Traceback (most recent call last)

Cell In[5], line 1
----> 1 x=np.array(df['sqft_living']).reshape(-1,1)
 2 y=np.array(df['sqft_lot']).reshape(-1,1)

NameError: name 'df' is not defined

In [6]:

```
df.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
regr=LinearRegression()
regr.fit(x_train,y_train)
print(regr.score(x_test,y_test))
```

-
NameError Traceback (most recent call last)

Cell In[6], line 1
----> 1 df.dropna(inplace=True)
 2 x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
 3 regr=LinearRegression()

NameError: name 'df' is not defined

In [7]:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_pred,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

-
NameError Traceback (most recent call last)

Cell In[7], line 1
----> 1 y_pred=regr.predict(x_test)
 2 plt.scatter(x_test,y_pred,color='b')
 3 plt.plot(x_test,y_pred,color='k')

NameError: name 'regr' is not defined

In [8]:

```
df500=df[:][:500]
sns.lmplot(x="sqft_living",y="sqft_lot",data=df500,order=1,ci=None)
x=np.array(df500['sqft_living']).reshape(-1,1)
y=np.array(df500['sqft_lot']).reshape(-1,1)
df500.dropna(inplace=True)
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25)
```

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[8], line 1
----> 1 df500=df[:][:500]
      2 sns.lmplot(x="sqft_living",y="sqft_lot",data=df500,order=1,ci=Non
e)
      3 x=np.array(df500['sqft_living']).reshape(-1,1)
```

NameError: name 'df' is not defined

In [9]:

```
regr=LinearRegression()
regr.fit(x_train,y_train)
print("Regression:",regr.score(x_test,y_test))
```

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[9], line 2
      1 regr=LinearRegression()
----> 2 regr.fit(x_train,y_train)
      3 print("Regression:",regr.score(x_test,y_test))
```

NameError: name 'x_train' is not defined

In [10]:

```
y_pred=regr.predict(x_test)
plt.scatter(x_test,y_test,color='b')
plt.plot(x_test,y_pred,color='k')
plt.show()
```

```
-----
-
NameError                                Traceback (most recent call las
t)
Cell In[10], line 1
----> 1 y_pred=regr.predict(x_test)
      2 plt.scatter(x_test,y_test,color='b')
      3 plt.plot(x_test,y_pred,color='k')
```

NameError: name 'x_test' is not defined

In []:

In []:

In []:

In []:

In []:

In []: