In [14]:

```
import numpy as np
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt
from sklearn import preprocessing ,svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.linear_model import Ridge,Lasso
from sklearn.preprocessing import StandardScaler
dv=pd.read_csv(r"C:\Users\magam\Downloads\Advertising.csv")
dv.head(10)
```

Out[14]:

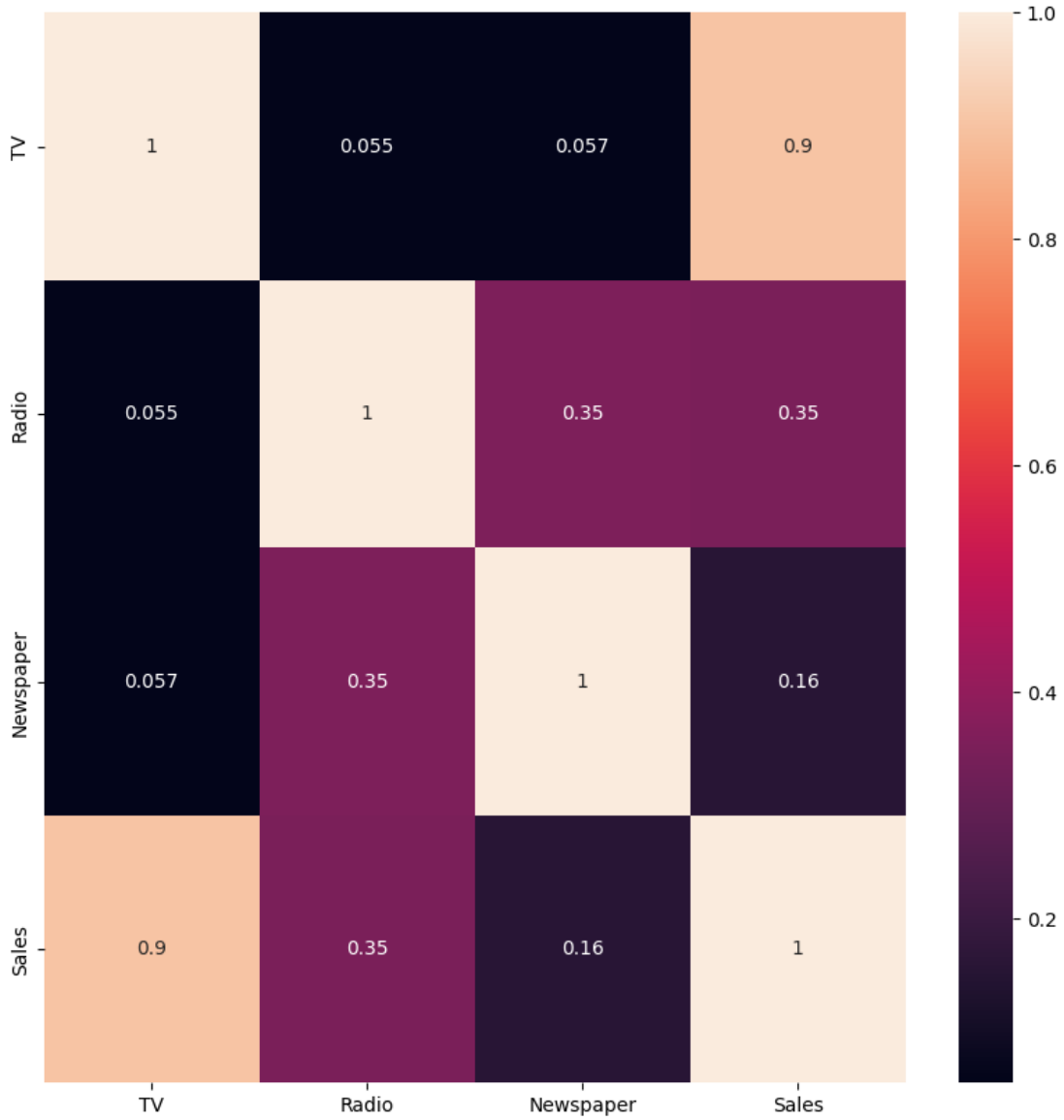
	TV	Radio	Newspaper	Sales
0	230.1	37.8	69.2	22.1
1	44.5	39.3	45.1	10.4
2	17.2	45.9	69.3	12.0
3	151.5	41.3	58.5	16.5
4	180.8	10.8	58.4	17.9
5	8.7	48.9	75.0	7.2
6	57.5	32.8	23.5	11.8
7	120.2	19.6	11.6	13.2
8	8.6	2.1	1.0	4.8
9	199.8	2.6	21.2	15.6

In [15]:

```
plt.figure(figsize = (10, 10))
sb.heatmap(dv.corr(), annot = True)
```

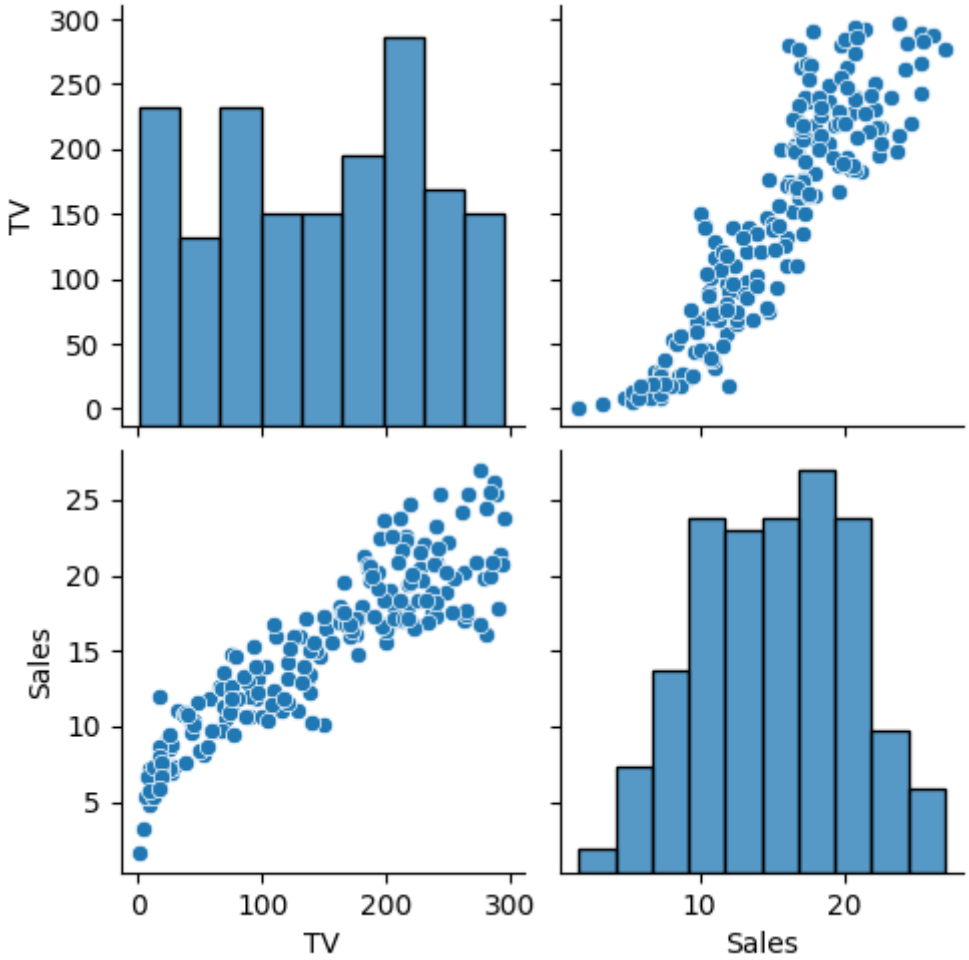
Out[15]:

<Axes: >



In [16]:

```
dv.drop(columns = ["Radio", "Newspaper"], inplace = True)
sb.pairplot(dv)
dv.Sales = np.log(dv.Sales)
```



In [17]:

```
features = dv.columns[0:2]
target = dv.columns[-1]
#X and y values
x = dv[features].values
y = dv[target].values
#splot
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_s
print("The dimension of X_train is {}".format(x_train.shape))
print("The dimension of X_test is {}".format(x_test.shape))
#Scale features
scaler = StandardScaler()
x_train = scaler.fit_transform(x_train)
x_test = scaler.transform(x_test)
```

```
The dimension of X_train is (140, 2)
The dimension of X_test is (60, 2)
```

In [18]:

```
#Model
lr = LinearRegression()
#Fit model
lr.fit(x_train, y_train)
#predict
#prediction = lr.predict(X_test)
#actual
actual = y_test
train_score_lr = lr.score(x_train, y_train)
test_score_lr = lr.score(x_test, y_test)
print("\nLinear Regression Model:\n")
print("The train score for lr model is {}".format(train_score_lr))
print("The test score for lr model is {}".format(test_score_lr))
```

```
Linear Regression Model:

The train score for lr model is 1.0
The test score for lr model is 1.0
```

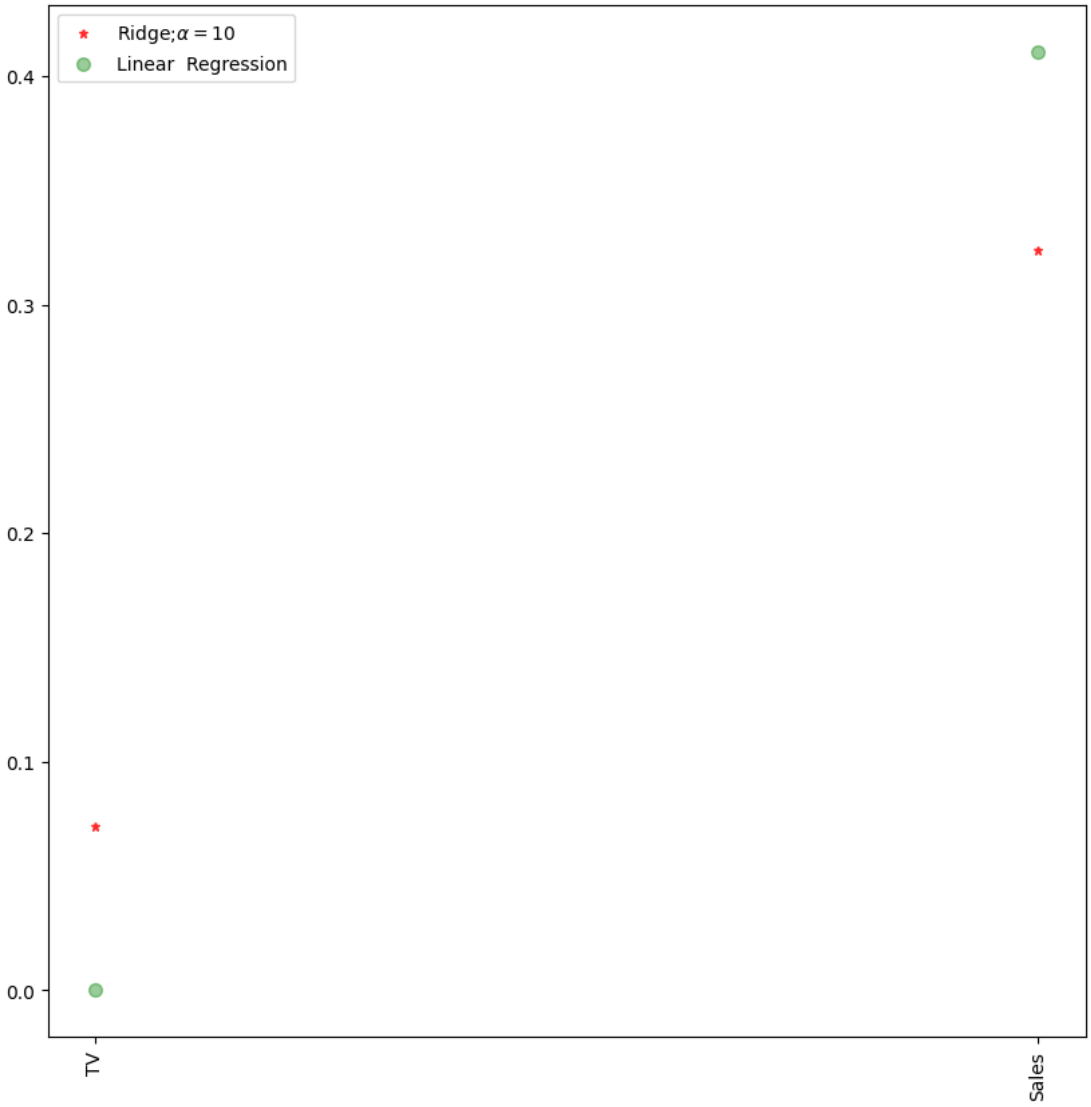
In [19]:

```
#Ridge Regression Model
ridgeReg = Ridge(alpha=10)
ridgeReg.fit(x_train,y_train)
#train and test scorefor ridge regression
train_score_ridge = ridgeReg.score(x_train, y_train)
test_score_ridge = ridgeReg.score(x_test, y_test)
print("\nRidge Model:\n")
print("The train score for ridge model is {}".format(train_score_ridge))
print("The test score for ridge model is {}".format(test_score_ridge))
```

```
Ridge Model:

The train score for ridge model is 0.990287139194161
The test score for ridge model is 0.9844266285141221
```

In [20]:

```
plt.figure(figsize=(10,10))
plt.plot(features,ridgeReg.coef_,alpha=0.7,linestyle='none',marker='*',markersize
plt.plot(features,lr.coef_,alpha=0.4,linestyle='none',marker='o',markersize=7,col
plt.xticks(rotation=90)
plt.legend()
plt.show()
```



Lasso regression

In [21]:

```
print("\nLasso Model: \n")
lasso = Lasso(alpha = 10)
lasso.fit(x_train,y_train)
train_score_ls =lasso.score(x_train,y_train)
test_score_ls =lasso.score(x_test,y_test)
print("The train score for ls model is {}".format(train_score_ls))
print("The test score for ls model is {}".format(test_score_ls))
```

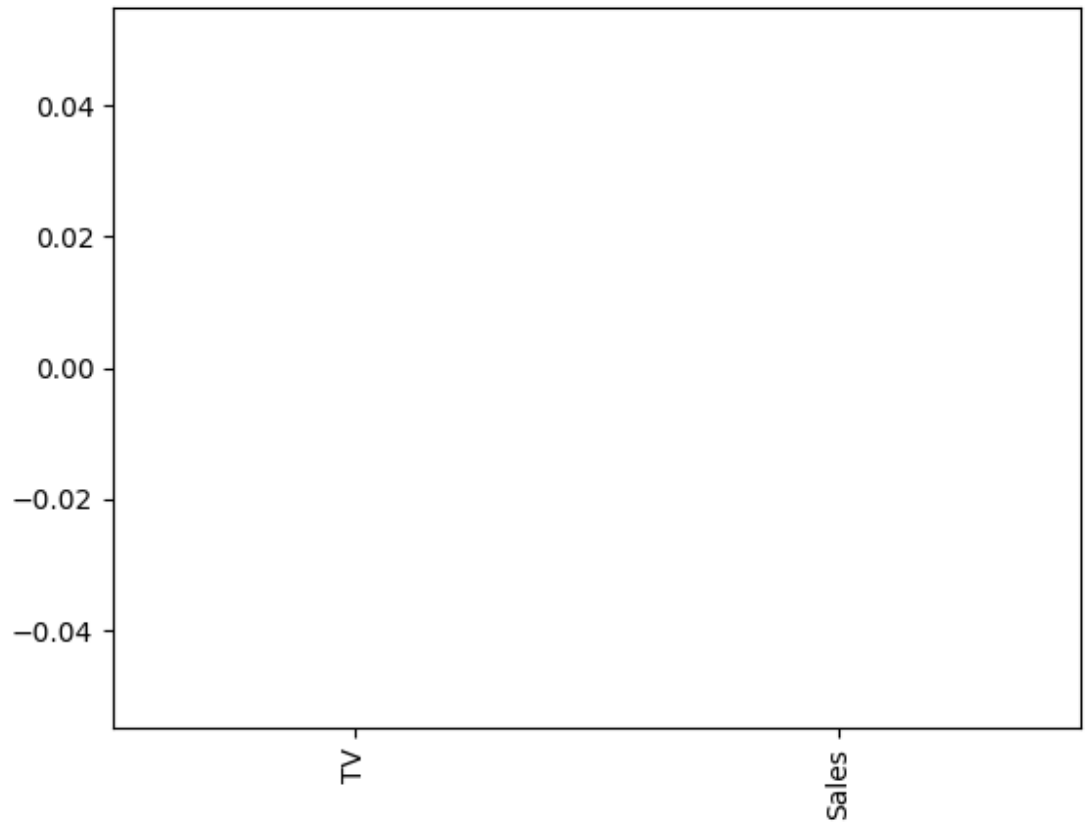
```
Lasso Model:

The train score for ls model is 0.0
The test score for ls model is -0.0042092253233847465
```

In [26]:

```
pd.Series(lasso.coef_, features).sort_values(ascending = True).plot(kind ="bar")
```

Out[26]:

```
<Axes: >
```



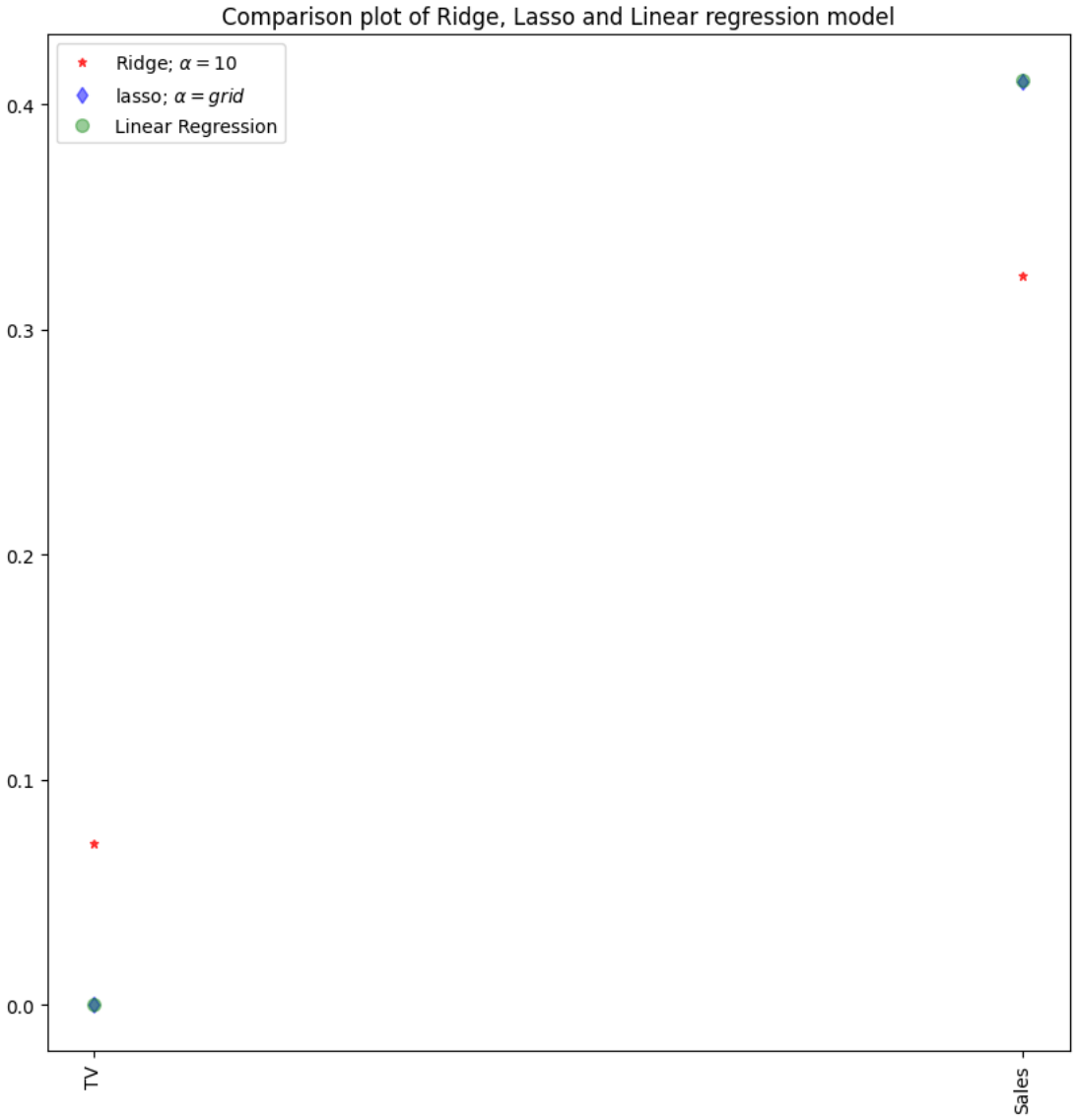
In [28]:

```
#Using the linear cv model
from sklearn.linear_model import LassoCV
#Lasso Cross validation
lasso_cv = LassoCV(alphas = [0.0001, 0.001,0.01, 0.1, 1, 10], random_state=0).fit
#score
print(lasso_cv.score(x_train, y_train))
print(lasso_cv.score(x_test, y_test))
```

```
0.9999999343798134
0.9999999152638072
```

In [30]:

```
plt.figure(figsize = (10, 10))
#add plot for ridge regression
plt.plot(features,ridgeReg.coef_,alpha=0.7,linestyle='none',marker='*',markersize
#add plot for lasso regression
plt.plot(lasso_cv.coef_,alpha=0.5,linestyle='none',marker='d',markersize=6,color=
#add plot for linear model
plt.plot(features,lr.coef_,alpha=0.4,linestyle='none',marker='o',markersize=7,col
#rotate axis
plt.xticks(rotation = 90)
plt.legend()
plt.title("Comparison plot of Ridge, Lasso and Linear regression model")
plt.show()
```



In [31]:

```
from sklearn.linear_model import RidgeCV
#Ridge Cross validation
ridge_cv = RidgeCV(alphas = [0.0001, 0.001,0.01, 0.1, 1, 10]).fit(x_train, y_trai
#score
print("The train score for ridge model is {}".format(ridge_cv.score(x_train, y_tr
print("The train score for ridge model is {}".format(ridge_cv.score(x_test, y_tes
```

```
The train score for ridge model is 0.999999999997627
The train score for ridge model is 0.9999999999962467
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [ ]:

```
1
```

In [24]:

```
import numpy as np
import pandas as pd
import seaborn as sb
import matplotlib.pyplot as plt
from sklearn import preprocessing ,svm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.linear_model import Ridge,Lasso
from sklearn.preprocessing import StandardScaler
dv=pd.read_csv(r"C:\Users\magam\Downloads\fiat500_VehicleSelection_Dataset.csv")
dv.head(10)
```

Out[24]:

	ID	model	engine_power	age_in_days	km	previous_owners	lat	lo
0	1	lounge	51	882	25000	1	44.907242	8.61156
1	2	pop	51	1186	32500	1	45.666359	12.24189
2	3	sport	74	4658	142228	1	45.503300	11.41784
3	4	lounge	51	2739	160000	1	40.633171	17.63460
4	5	pop	73	3074	106880	1	41.903221	12.49565
5	6	pop	74	3623	70225	1	45.000702	7.68227
6	7	lounge	51	731	11600	1	44.907242	8.61156
7	8	lounge	51	1521	49076	1	41.903221	12.49565
8	9	sport	73	4049	76000	1	45.548000	11.54947
9	10	sport	51	3653	89000	1	45.438301	10.99170

In [14]:

```
1 dv.info()
```

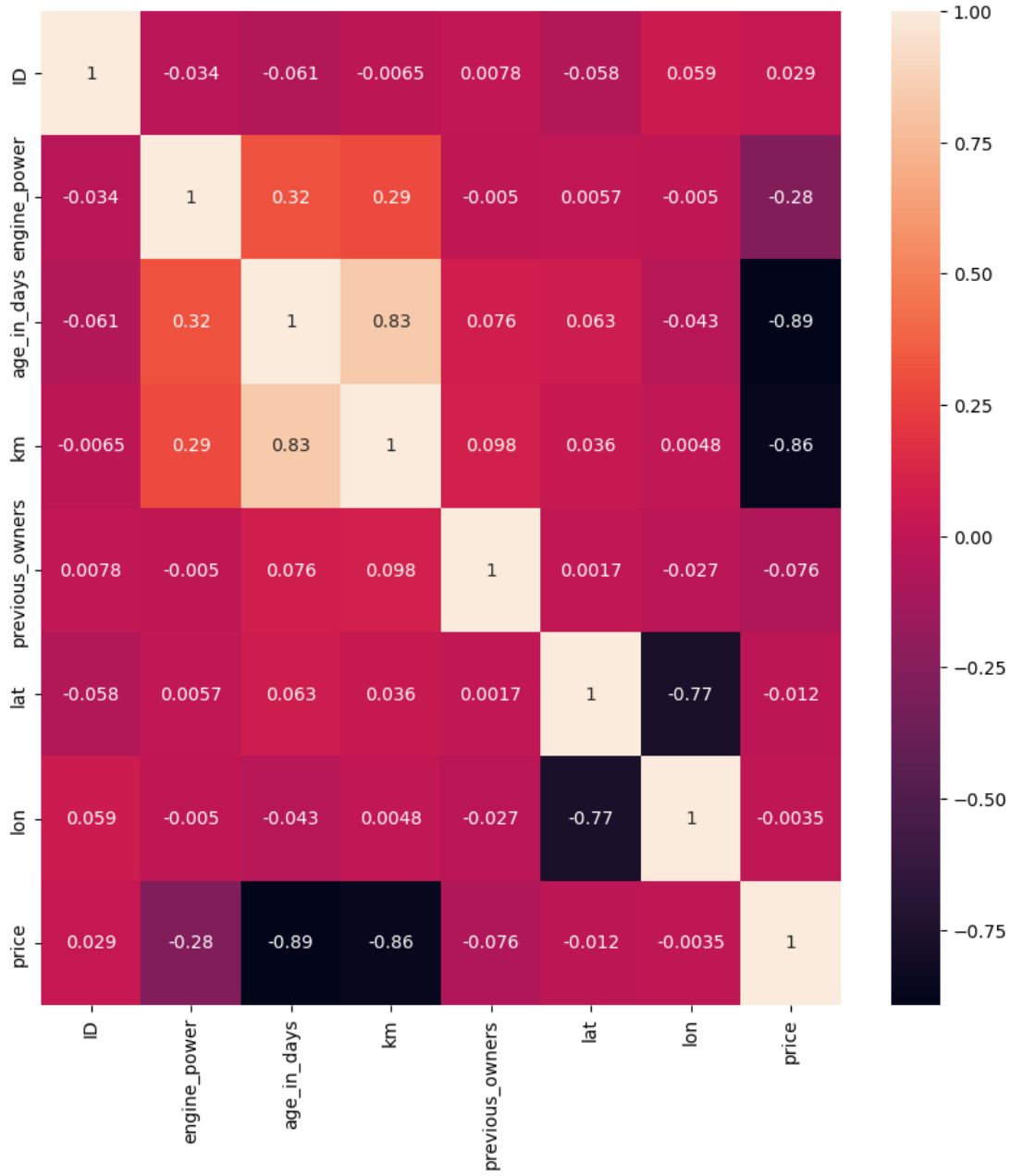
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1538 entries, 0 to 1537
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   ID               1538 non-null   int64  
 1   model            1538 non-null   object 
 2   engine_power     1538 non-null   int64  
 3   age_in_days      1538 non-null   int64  
 4   km               1538 non-null   int64  
 5   previous_owners  1538 non-null   int64  
 6   lat              1538 non-null   float64
 7   lon              1538 non-null   float64
 8   price            1538 non-null   int64  
dtypes: float64(2), int64(6), object(1)
memory usage: 108.3+ KB
```

In [15]:

```
dat=dv
dat.drop(columns=["model"], inplace = True)
plt.figure(figsize = (10, 10))
sb.heatmap(dv.corr(), annot = True)
```

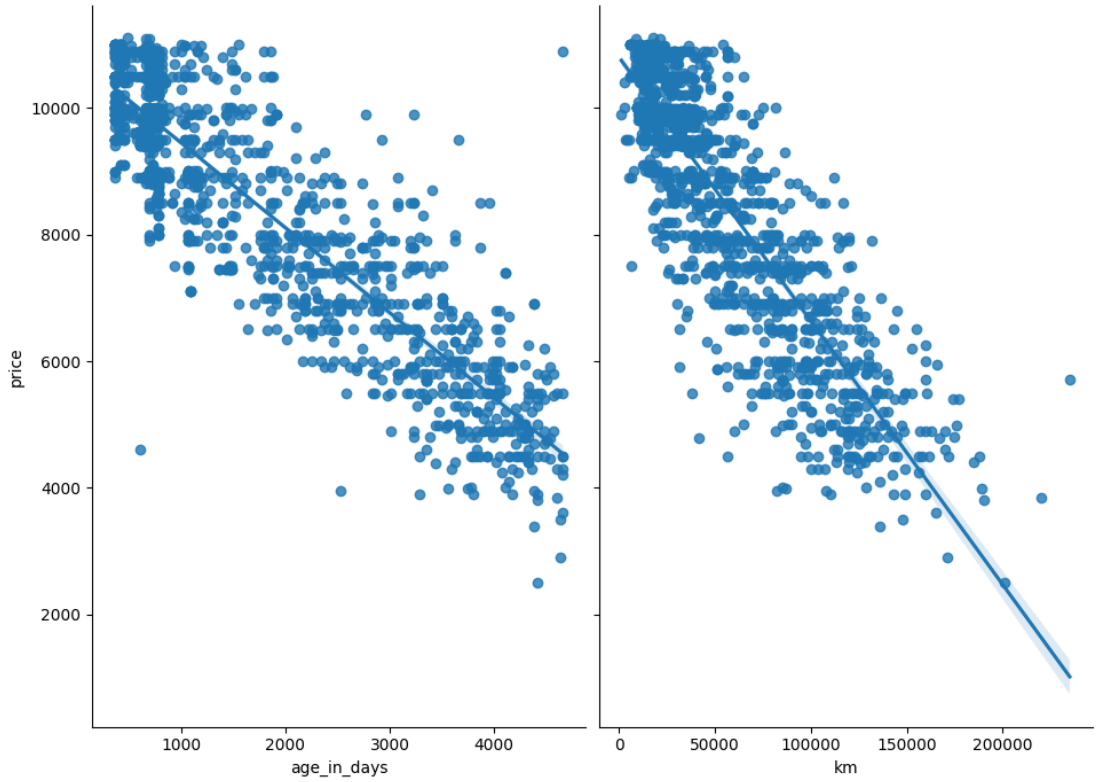
Out[15]:

<Axes: >



In [25]:

```
sb.pairplot(dv,x_vars=['age_in_days','km'],y_vars='price',height=7,aspect=0.7,kin
#sb.pairplot(dv)
dv.price = np.log(dv.price)
```



In [26]:

```
dv.columns
```

Out[26]:

```
Index(['ID', 'model', 'engine_power', 'age_in_days', 'km', 'previous_ow
ners',
       'lat', 'lon', 'price'],
      dtype='object')
```

In [27]:

```
features = dv.columns[3:5]
target = dv.columns[-1]
#X and y values
x = dv[features].values
y = dv[target].values
#splot
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_s
print("The dimension of X_train is {}".format(x_train.shape))
print("The dimension of X_test is {}".format(x_test.shape))
#Scale features
scaler = StandardScaler()
x_train = scaler.fit_transform(x_train)
x_test = scaler.transform(x_test)
```

```
The dimension of X_train is (1076, 2)
The dimension of X_test is (462, 2)
```

In [28]:

```
#Model
lr = LinearRegression()
#Fit model
lr.fit(x_train, y_train)
#actual
actual = y_test
train_score_lr = lr.score(x_train, y_train)
test_score_lr = lr.score(x_test, y_test)
print("\nLinear Regression Model:\n")
print("The train score for lr model is {}".format(train_score_lr))
print("The test score for lr model is {}".format(test_score_lr))
```

```
Linear Regression Model:

The train score for lr model is 0.8407959504511326
The test score for lr model is 0.8396374461789744
```

Ridge Regression

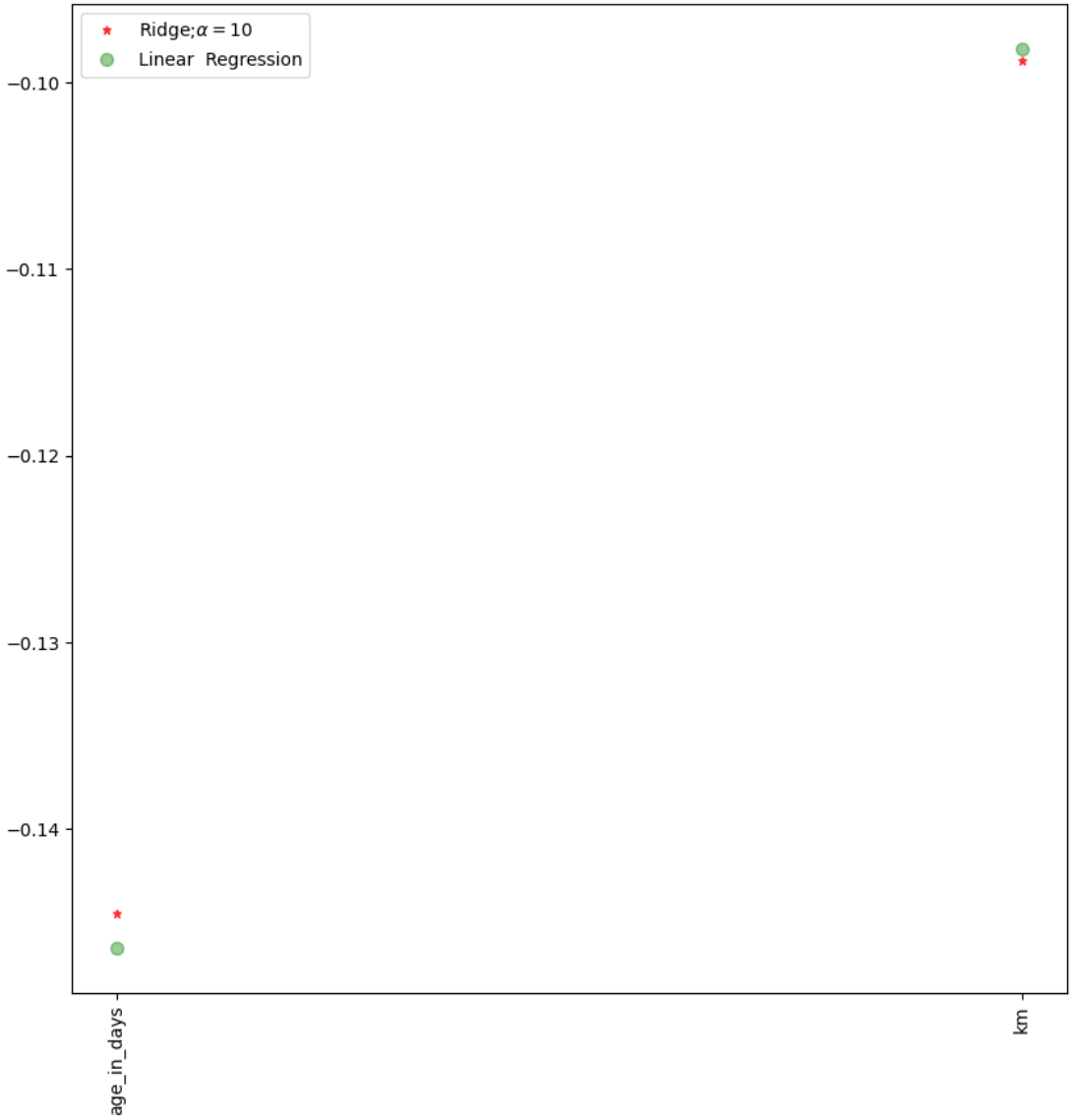
In [29]:

```
#Ridge Regression Model
ridgeReg = Ridge(alpha=10)
ridgeReg.fit(x_train,y_train)
#train and test scorefor ridge regression
train_score_ridge = ridgeReg.score(x_train, y_train)
test_score_ridge = ridgeReg.score(x_test, y_test)
print("\nRidge Model:\n")
print("The train score for ridge model is {}".format(train_score_ridge))
print("The test score for ridge model is {}".format(test_score_ridge))
```

```
Ridge Model:

The train score for ridge model is 0.8407665046629869
The test score for ridge model is 0.8395909167380576
```

In [30]:

```
plt.figure(figsize=(10,10))
plt.plot(features,ridgeReg.coef_,alpha=0.7,linestyle='none',marker='*',markersize
plt.plot(features,lr.coef_,alpha=0.4,linestyle='none',marker='o',markersize=7,col
plt.xticks(rotation=90)
plt.legend()
plt.show()
```



Lasso Regression

In [31]:

```
print("\nLasso Model: \n")
lasso = Lasso(alpha = 10)
lasso.fit(x_train,y_train)
train_score_ls =lasso.score(x_train,y_train)
test_score_ls =lasso.score(x_test,y_test)
print("The train score for ls model is {}".format(train_score_ls))
print("The test score for ls model is {}".format(test_score_ls))
```

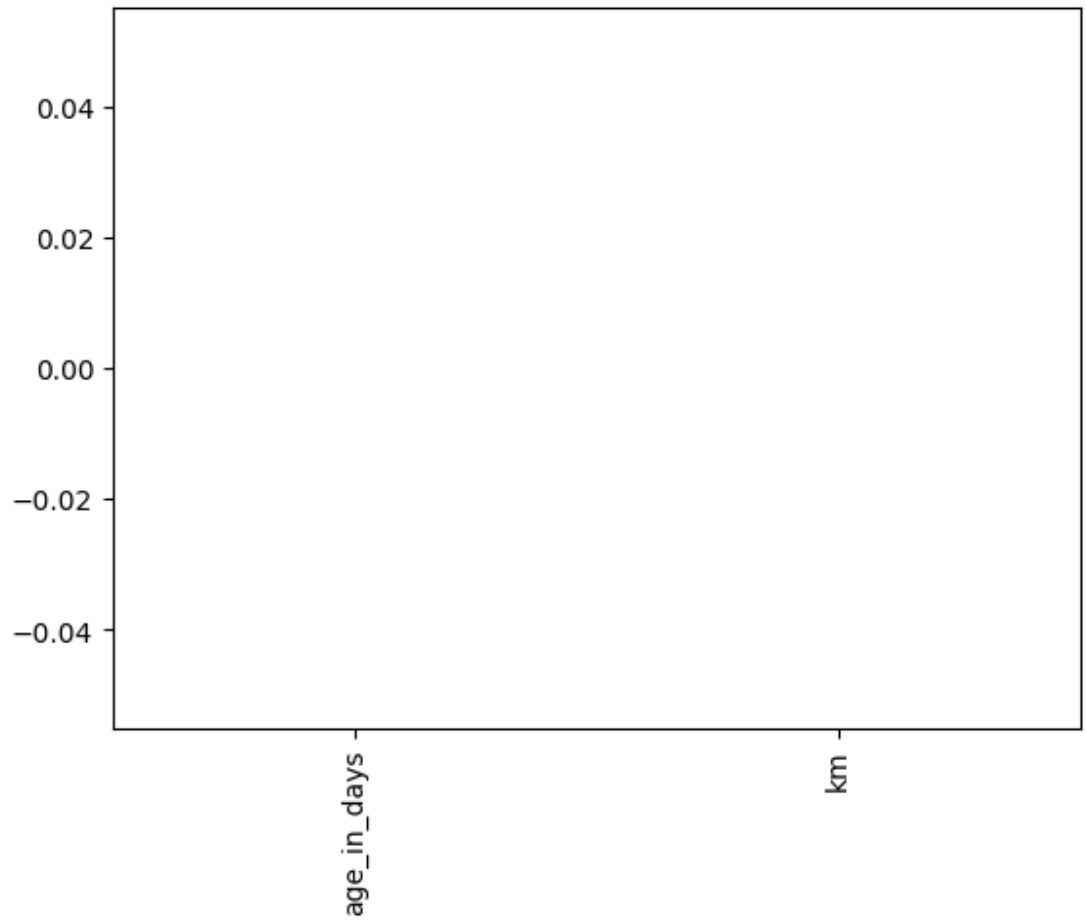
```
Lasso Model:

The train score for ls model is 0.0
The test score for ls model is -0.0004974348027177999
```

In [32]:

```
pd.Series(lasso.coef_, features).sort_values(ascending = True).plot(kind ="bar")
```

Out[32]:

```
<Axes: >
```



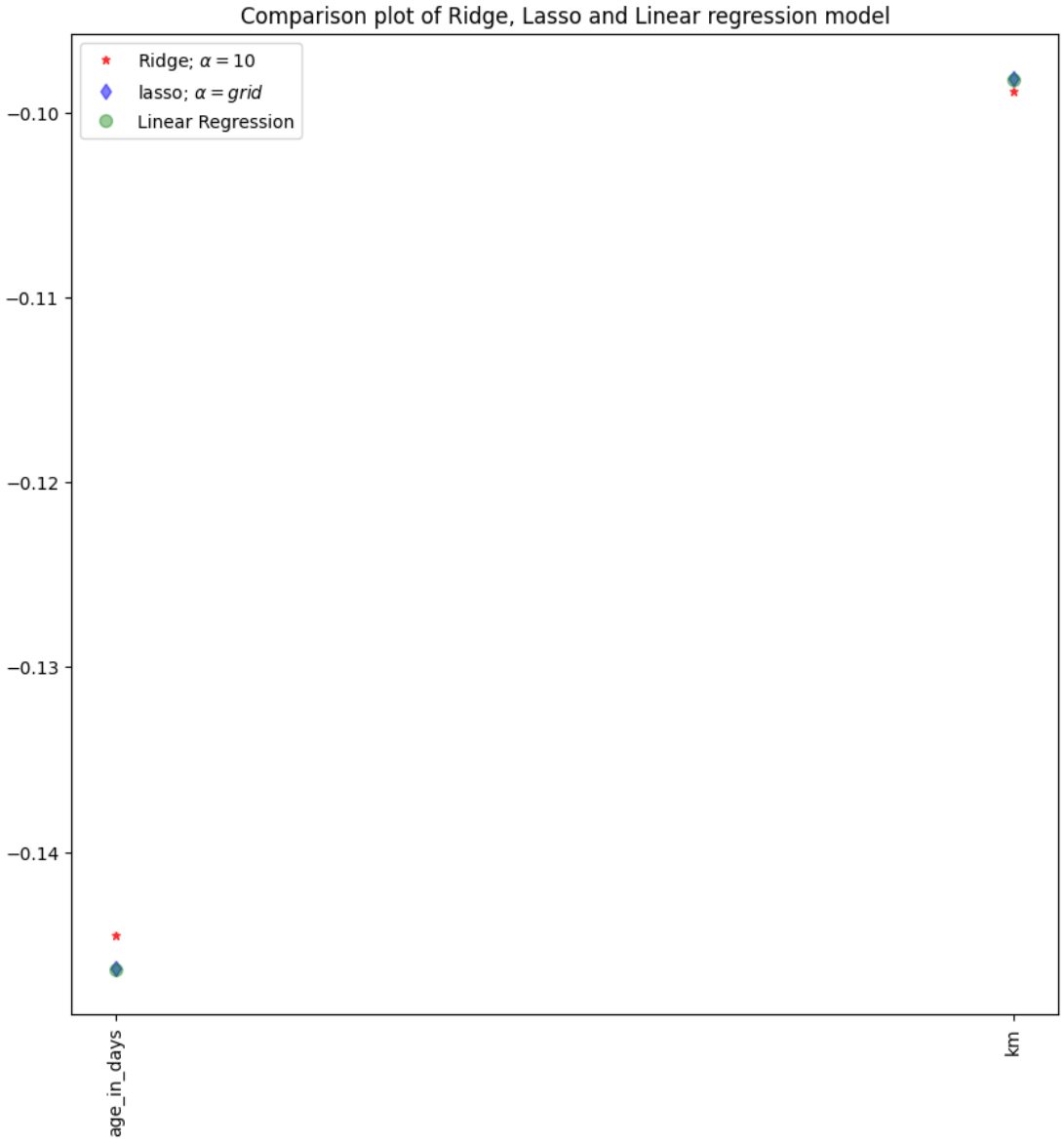
In [33]:

```
#Using the linear cv model
from sklearn.linear_model import LassoCV
#Lasso Cross validation
lasso_cv = LassoCV(alphas = [0.0001, 0.001,0.01, 0.1, 1, 10], random_state=0).fit
#score
print(lasso_cv.score(x_train, y_train))
print(lasso_cv.score(x_test, y_test))
```

```
0.8407957935001775
0.8396376947295177
```

In [34]:

```
plt.figure(figsize = (10, 10))
#add plot for ridge regression
plt.plot(features,ridgeReg.coef_,alpha=0.7,linestyle='none',marker='*',markersize
#add plot for lasso regression
plt.plot(lasso_cv.coef_,alpha=0.5,linestyle='none',marker='d',markersize=6,color=
#add plot for linear model
plt.plot(features,lr.coef_,alpha=0.4,linestyle='none',marker='o',markersize=7,col
#rotate axis
plt.xticks(rotation = 90)
plt.legend()
plt.title("Comparison plot of Ridge, Lasso and Linear regression model")
plt.show()
```



In [35]:

```
from sklearn.linear_model import RidgeCV
#Ridge Cross validation
ridge_cv = RidgeCV(alphas = [0.0001, 0.001,0.01, 0.1, 1, 10]).fit(x_train, y_trai
#score
print("The train score for ridge model is {}".format(ridge_cv.score(x_train, y_tr
print("The train score for ridge model is {}".format(ridge_cv.score(x_test, y_tes
```

```
The train score for ridge model is 0.8407665046629889
The train score for ridge model is 0.8395909167380597
```

In [36]:

```
from sklearn.linear_model import ElasticNet
regr=ElasticNet()
regr.fit(x,y)
print(regr.coef_)
print(regr.intercept_)
```

```
[-1.14013910e-04 -2.51291433e-06]
9.34861331834344
```

In [37]:

```
y_pred_elastc=regr.predict(x_train)
mean_squared_error=np.mean((y_pred_elastc-y_train)**2)
print(mean_squared_error)
```

```
0.1679531699267824
```

In [ ]:

```

```

In [ ]:

```

```

In [ ]:

```

```

In [ ]:

```

```