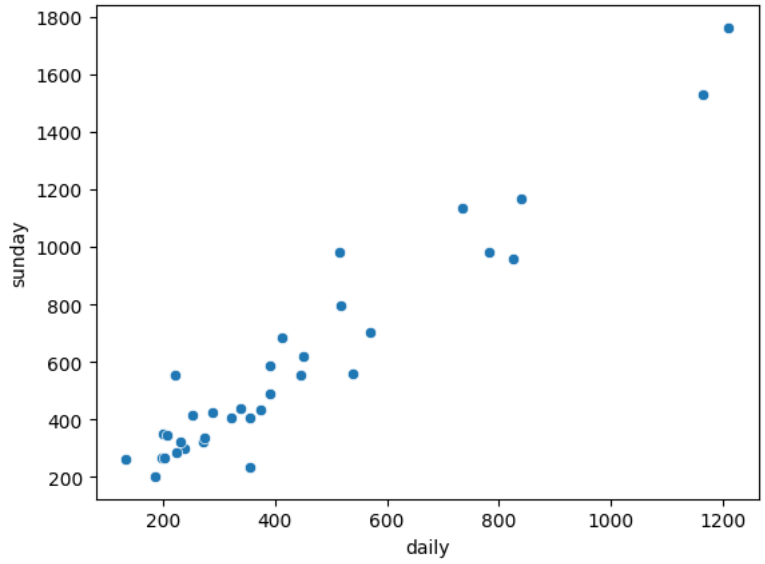
```
import pandas as pd

df = pd.read_csv("https://raw.githubusercontent.com/AmenaNajeeb/Data/master/NewspaperData.csv")

df.head()
```

	Newspaper	daily	sunday
0	Baltimore Sun	391.952	488.506
1	Boston Globe	516.981	798.298
2	Boston Herald	355.628	235.084
3	Charlotte Observer	238.555	299.451
4	Chicago Sun Times	537.780	559.093

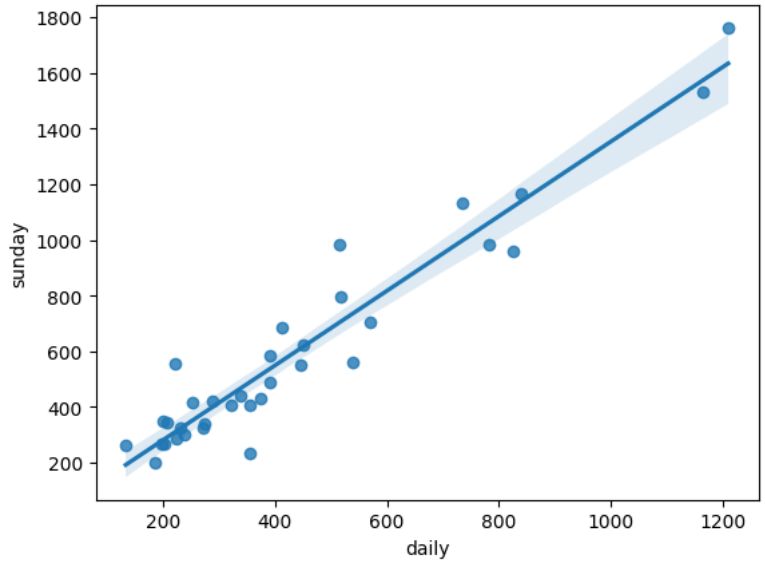
```
import seaborn as sns

sns.scatterplot(x=df["daily"],y=df["sunday"])
```

```
<Axes: xlabel='daily', ylabel='sunday'>
```




```
sns.regplot(x=df["daily"],y=df["sunday"])
```

```
<Axes: xlabel='daily', ylabel='sunday'>
```




```
#statsmodel
import statsmodels.formula.api as smf
```

```
model = smf.ols("sunday~daily",data=df).fit()
```

```
model.params
```

```
Intercept    13.835630
daily         1.339715
dtype: float64
```

```
x = [100] #daily
df2 = pd.DataFrame(x,columns=["daily"])
```

```
df2
```

	daily
0	100

```
model.predict(df2)
```

```
0    147.807106
dtype: float64
```

```
model.summary()
```

OLS Regression Results

Dep. Variable:	sunday	**R-squared:**	0.918
Model:	OLS	**Adj. R-squared:**	0.915
Method:	Least Squares	**F-statistic:**	358.5
Date:	Wed, 17 May 2023	**Prob (F-statistic):**	6.02e-19
Time:	11:29:11	**Log-Likelihood:**	-206.85
No. Observations:	34	**AIC:**	417.7
Df Residuals:	32	**BIC:**	420.8
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	13.8356	35.804	0.386	0.702	-59.095	86.766
daily	1.3397	0.071	18.935	0.000	1.196	1.484

Omnibus:	3.297	**Durbin-Watson:**	2.059
Prob(Omnibus):	0.192	**Jarque-Bera (JB):**	1.990
Skew:	0.396	**Prob(JB):**	0.370
Kurtosis:	3.882	**Cond. No.**	965.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

```
x=df["daily"]
df3=pd.DataFrame(x,columns=["daily"])
data_pred=model.predict(df3)
```

```
from sklearn import metrics
```

```
mse=metrics.mean_squared_error(df["sunday"],data_pred)
from math import sqrt
rmse=sqrt(mse)
print(rmse)
```

```
106.1540958806654
```

✓ 0s completed at 5:19 PM