```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import statsmodels.formula.api as smf
from statsmodels.graphics.regressionplots import influence_plot

data = pd.read_csv('Downloads/NewspaperData.csv')
data
```

	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
5	Chicago Tribune	733.775	1133.249
6	Cincinnati Enquirer	198.832	348.744
7	Denver Post	252.624	417.779
8	Des Moines Register	206.204	344.522
9	Hartford Courant	231.177	323.084
10	Houston Chronicle	449.755	620.752
11	Kansas City Star	288.571	423.305
12	Los Angeles Daily News	185.736	202.614
13	Los Angeles Times	1164.388	1531.527
14	Miami Herald	444.581	553.479
15	Minneapolis Star Tribune	412.871	685.975
16	New Orleans Times-Picayune	273.280	324.241

```
data.head(6)
```

	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
5	Chicago Tribune	733.775	1133.249

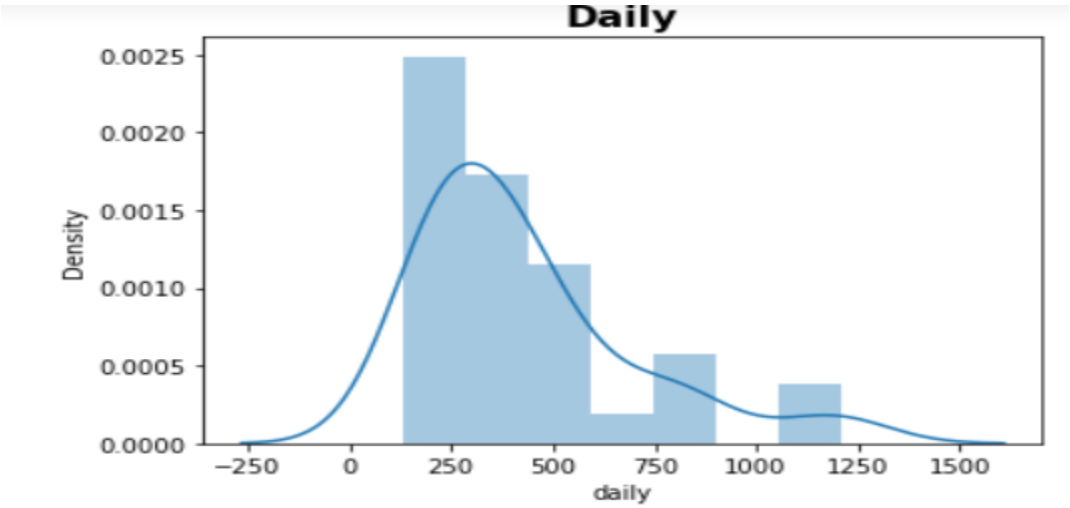
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 34 entries, 0 to 33
Data columns (total 3 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Newspaper  34 non-null     object
 1   daily      34 non-null     float64
 2   sunday     34 non-null     float64
dtypes: float64(2), object(1)
memory usage: 944.0+ bytes
```

data.corr()

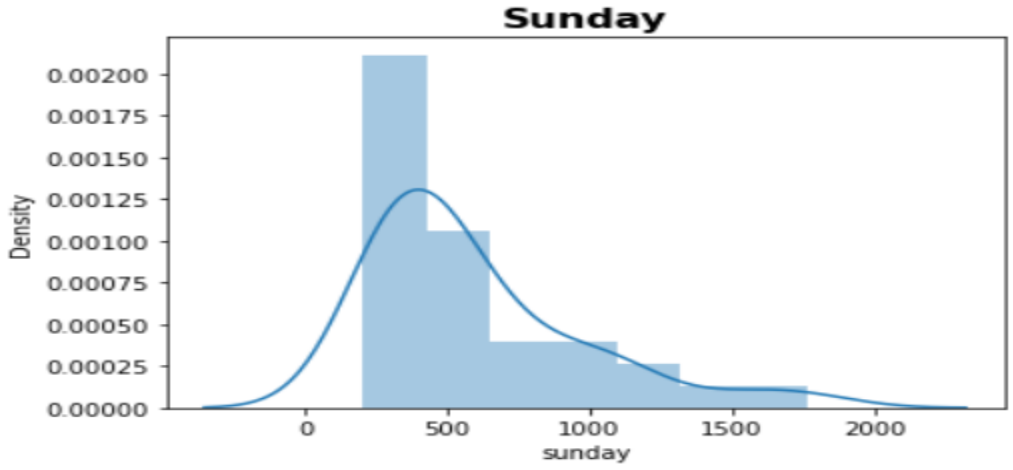
	daily	sunday
daily	1.000000	0.958154
sunday	0.958154	1.000000

```
sns.distplot(data['daily'])
plt.title('Daily',fontsize=16,fontweight='bold')
fig = plt.figure(figsize=(16,8))
```



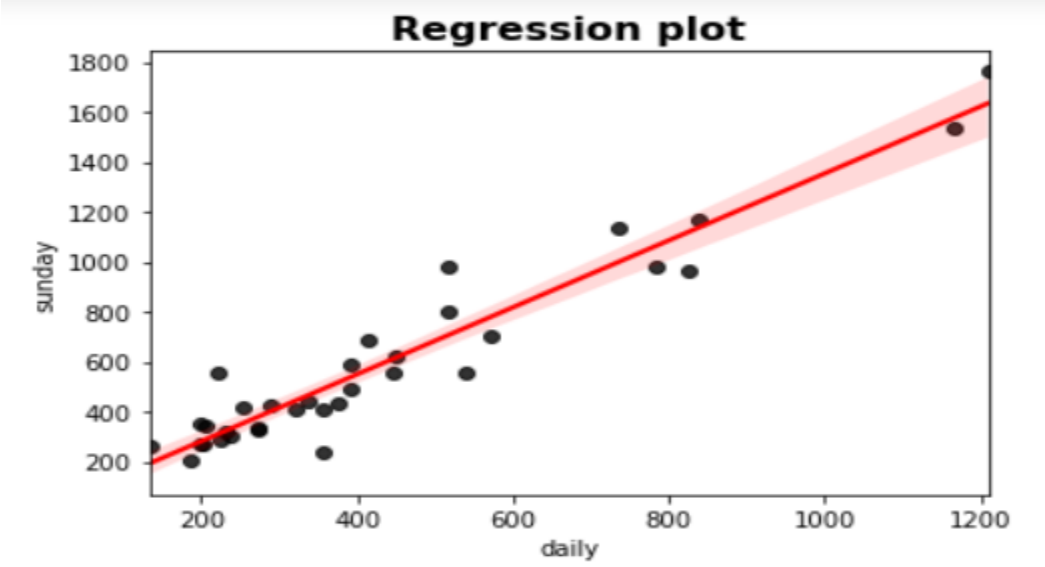

```
<Figure size 1152x576 with 0 Axes>
```

```
sns.distplot(data['sunday'])
plt.title('Sunday',fontsize=16,fontweight='bold')
fig=plt.figure(figsize=(16,8))
```




```
<Figure size 1152x576 with 0 Axes>
```

```
sns.regplot(data['daily'],data['sunday'],line_kws={'color':'red'},scatter_kws={'color':'k'})
plt.title('Regression plot',fontsize=16,fontweight='bold')
fig=plt.figure(figsize=(16,8))
```




```
<Figure size 1152x576 with 0 Axes>
```

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

```
model.params
```

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

```
print(model.tvalues,'\n',model.pvalues)
```

```
Intercept     0.386427
daily        18.934840
dtype: float64
 Intercept    7.017382e-01
daily        6.016802e-19
dtype: float64
```

```
print(model.rsquared,model.rsquared_adj)
```

```
0.9180596895873294 0.9154990548869335
```

```
a = pd.Series([200,300])
a
```

```
0    200
1    300
dtype: int64
```

x = pd.DataFrame(a,columns=['daily'])
x

	daily
0	200
1	300

model.predict(x)

```
0    281.778581
1    415.750057
dtype: float64
```