```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size
```

```
x_train.head()
```

	Airline	Source	Destination	Date	Month	Year	Dep_Time_Hour	D
10611	4	4	3	18	5	2019	7	
1034	8	2	1	24	4	2019	15	
8123	4	2	1	27	6	2019	2	
4779	4	3	0	1	4	2019	6	
3207	3	3	0	24	5	2019	18	

	Airline	Source	Destination	Date	Month	Year	Dep_Time_
0	-0.410934	-1.658354	2.416648	1.237192	-1.467619	0.0	1.6
1	-1.261305	0.890262	-0.973718	-1.475375	0.250165	0.0	-1.3
2	0.014251	0.040723	-0.295645	-0.531874	1.109057	0.0	-0.6
3	-0.410934	0.890262	-0.973718	-0.178060	0.250165	0.0	0.9
4	-0.410934	-1.658354	2.416648	-1.475375	-1.467619	0.0	0.61

```
y = data['Price']
x = data.drop(columns=['Price'],axis=1)
```

```
### Scaling the Data
```

```
from sklearn.preprocessing import StandardScaler
ss=StandardScaler()
```

```
x_scaled = ss.fit_transform(x)
```

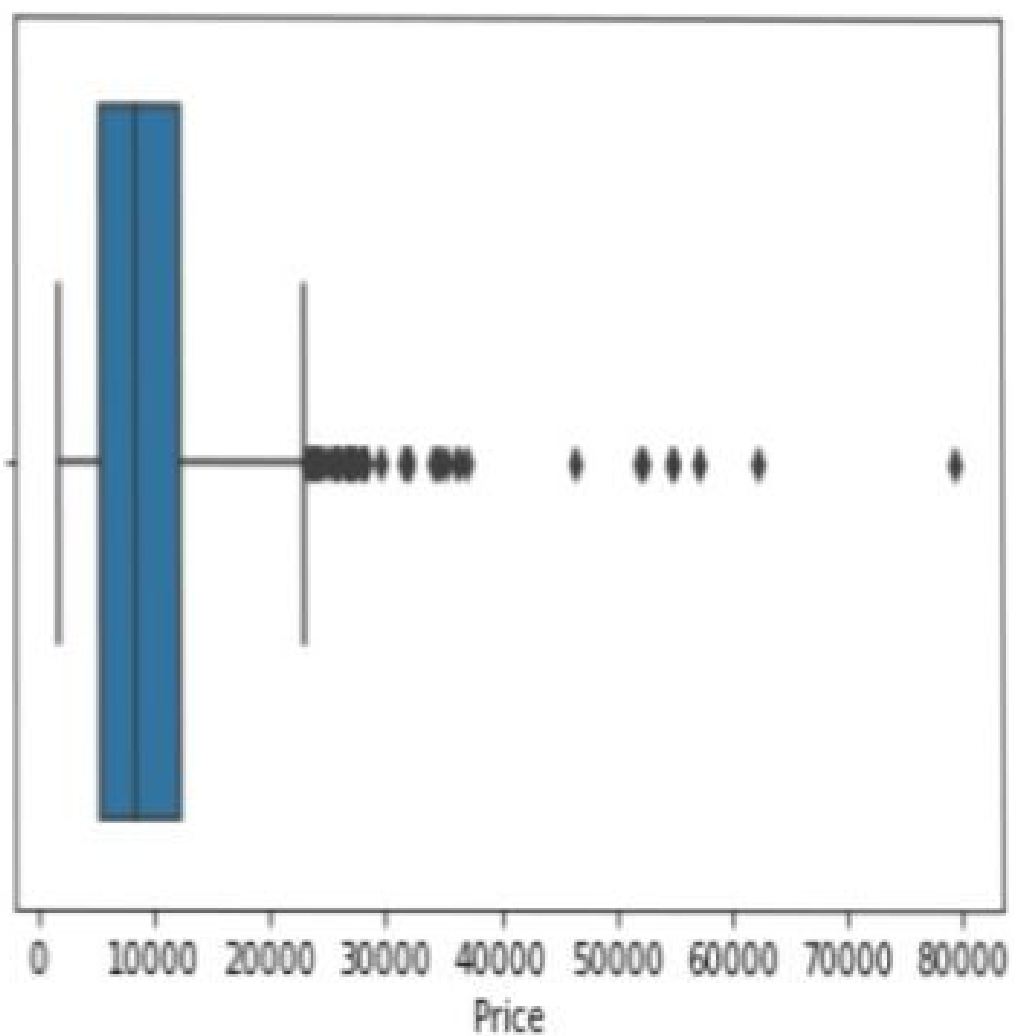
```
x_scaled = pd.DataFrame(x_scaled,columns=x.columns)
x_scaled.head()
```

	Airline	Source	Destination	Date	Month	Year	Dep_Time_
0	-0.410934	-1.658354	2.416648	1.237192	-1.467619	0.0	1.65
1	-1.261305	0.890262	-0.973718	-1.475375	0.250165	0.0	-1.30
2	0.014251	0.040723	-0.295645	-0.531874	1.109057	0.0	-0.60

```
# Detecting the Outliers
import seaborn as sns
sns.boxplot(data['Price'])
```

```
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabo
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
```

```
<AxesSubplot:xlabel='Price'>
```

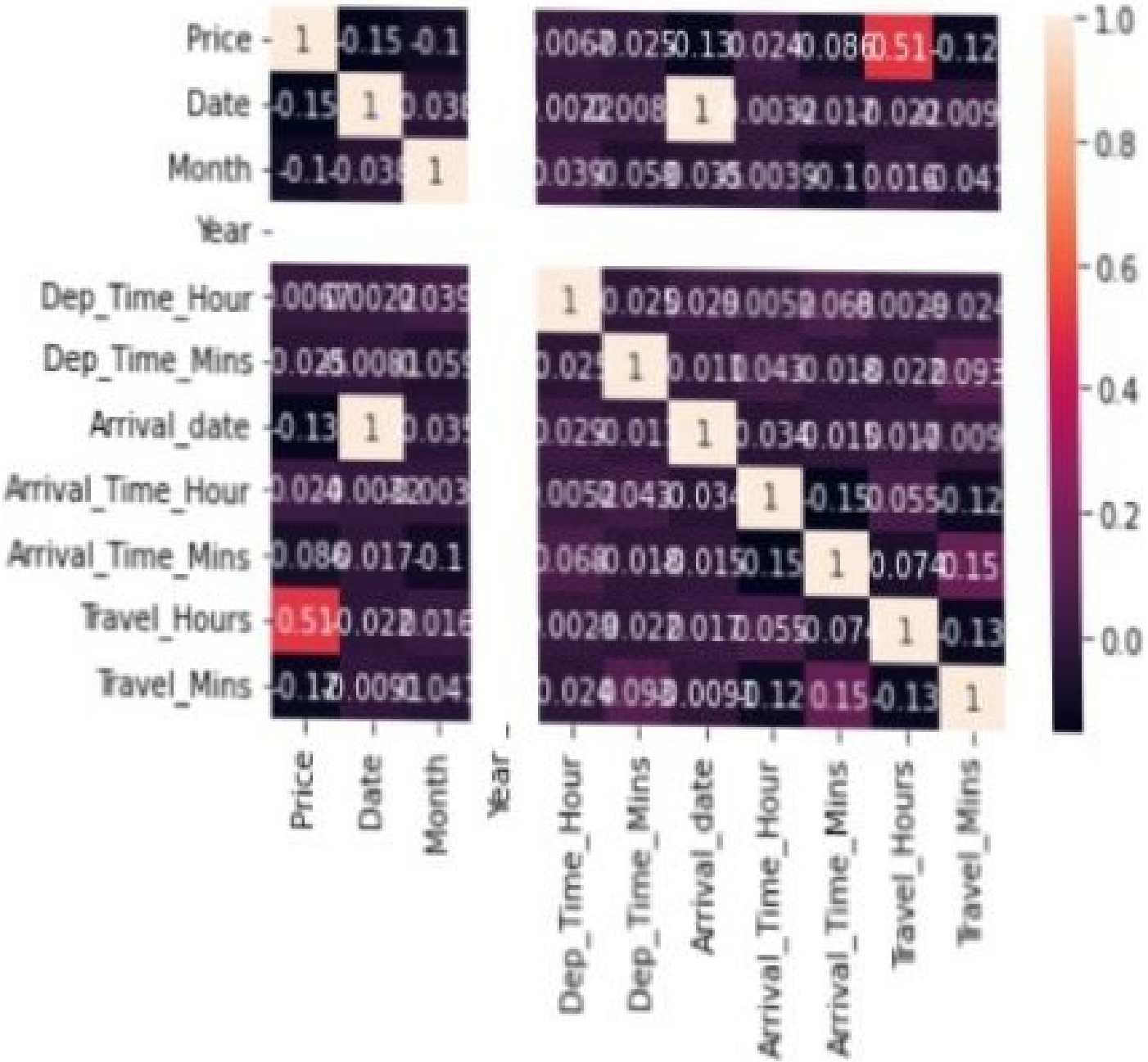


Outlier Detection For 'Price' Column

Sometimes it's best to keep outliers in your data. it captures the valuable information and they can effect on statistical results and detect any errors in your statistical process. Here, we are checking Outliers in the 'Price' column.

```
sns.heatmap(data.corr(),annot=True)
```

```
<AxesSubplot:>
```



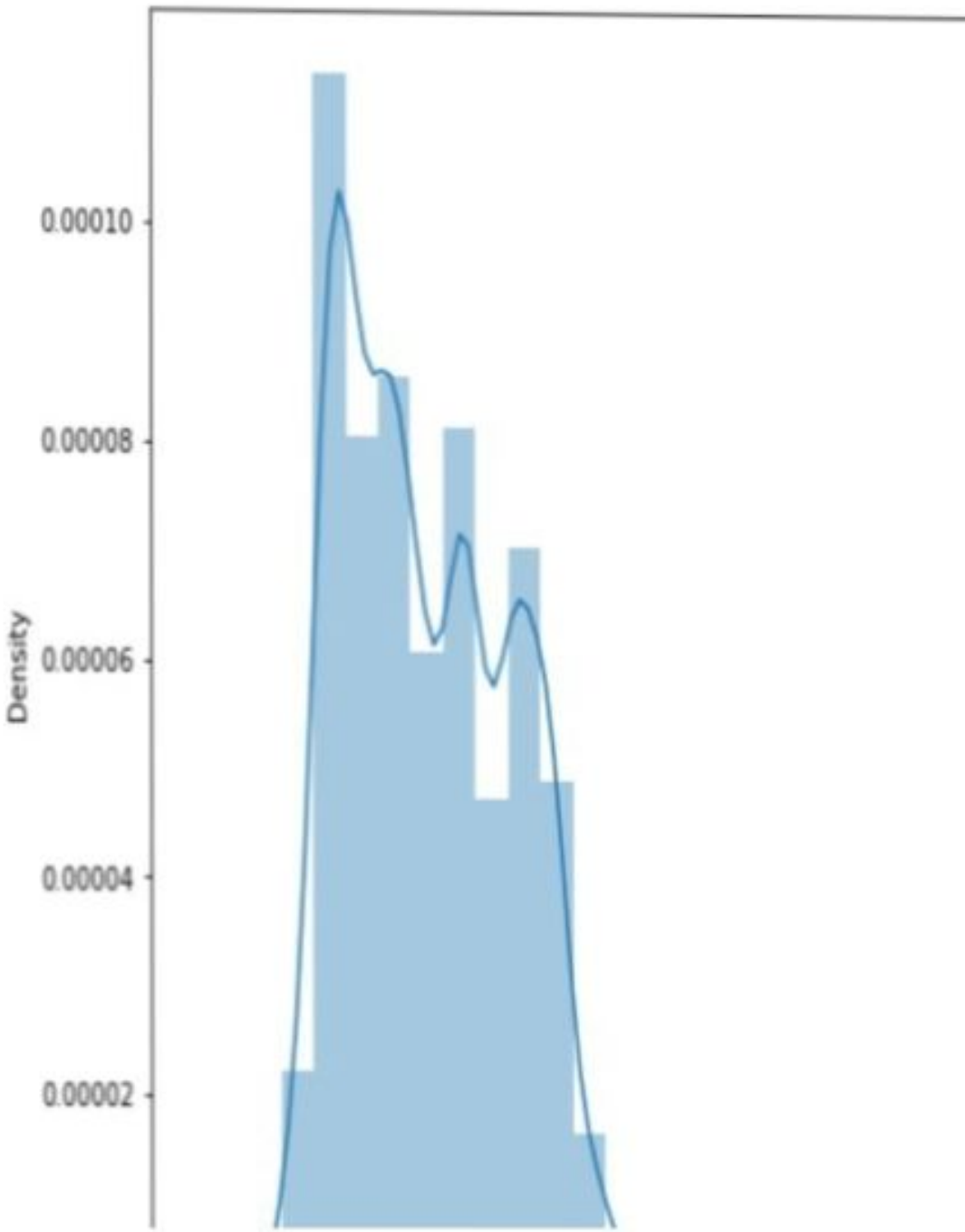
Checking The Correlation Using HeatMap

- Here, I 'm finding the correlation using HeatMap. It visualizes the data in 2-D colored maps making use of color variations. It describes the relationship variables in form of colors instead of numbers it will be plotted on both axes.
- So, by this heatmap we found that correlation between 'Arrival_date' and 'Date'. Remaining all columns don't have the any Correlation.

```
#Distribution of 'PRICE' Column
plt.figure(figsize=(15,8))
sns.distplot(data.Price)
```

```
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seab
nction and will be removed in a future version. Please a
milar flexibility) or `histplot` (an axes-level function
  warnings.warn(msg, FutureWarning)
```

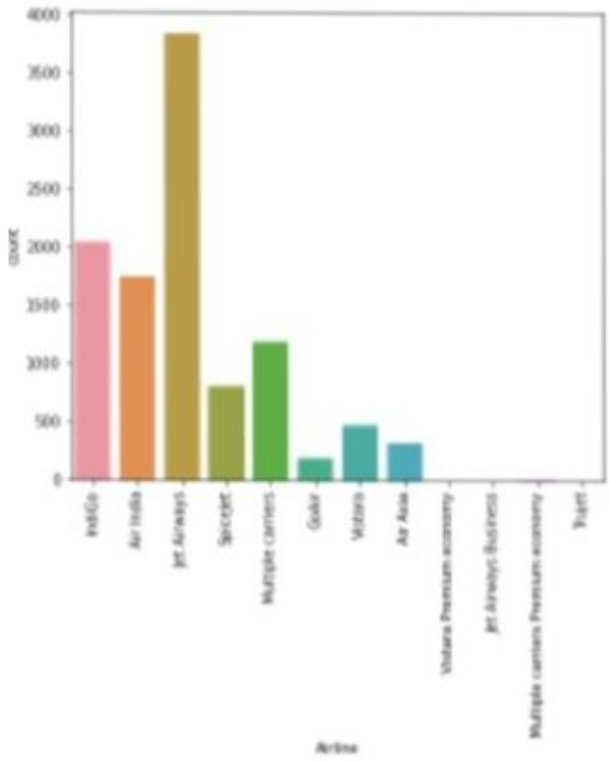
```
<AxesSubplot:xlabel='Price', ylabel='Density'>
```

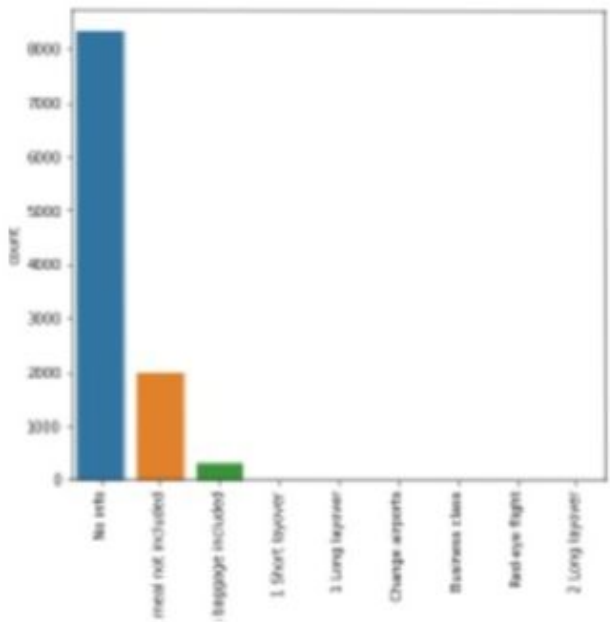


We Now Plot Distribution Plots To Check The Distribution In Numerical Data (Distribution Of 'Price' Column)

- The seaborn.displot() function is used to plot the displot. The displot represents the univariate distribution of data variable as an argument and returns the plot with the density distribution. Here, I used

```
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabo
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabo
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
```





```python
#plotting Countplots for Categorical Data

import seaborn as sns
c=1
plt.figure(figsize=(20,45))

for i in categorical:
    plt.subplot(6,3,c)
    sns.countplot(data[i])
    plt.xticks(rotation=90)
    plt.tight_layout(pad=3.0)
    c=c+1

plt.show()
```

```
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabor
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabor
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabor
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabor
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
  warnings.warn(
C:\Users\SmartBridge-PC\anaconda3\lib\site-packages\seabor
keyword arg: x. From version 0.12, the only valid positio
plicit keyword will result in an error or misinterpretati
```

```
data.describe()
```

	Price
count	10683.000000
mean	9087.064121
std	4611.359167
min	1759.000000
25%	5277.000000
50%	8372.000000
75%	12373.000000
max	79512.000000

```
data.describe()
```

	Price
count	10683.000000
mean	9087.064121
std	4611.359167