FUTURE SALES PREDICTION

Introduction:

Predicting the future sales of a product helps a business manage the manufacturing and advertising cost of the product. There are many more benefits of predicting the future sales of a product. So if you want to learn to predict the future sales of a product with machine learning, this article is for you. In this article, I will take you through the task of future sales prediction with machine learning using Python.

Future Sales Prediction (Case Study)

The dataset given here contains the data about the sales of the product. The dataset is about the advertising cost incurred by the business on various advertising platforms. Below is the description of all the columns in the dataset:

- **TV**: Advertising cost spent in dollars for advertising on TV;
- **Radio**: Advertising cost spent in dollars for advertising on Radio;
- **Newspaper**: Advertising cost spent in dollars for advertising on Newspaper;
- **Sales**: Number of units sold;

So, in the above dataset, the sales of the product depend on the advertisement cost of the product. I hope you now have understood everything about this dataset. Now in the section below, I will take you through the task of future sales prediction with machine learning using Python.

Python libraries

```
import pandas as pd
import numpy as np
from sklearn.model_selection
import train_test_split
from sklearn.linear_model
import LinearRegression
```

Understanding the working of the Prophet library :

Let us now understand the working of the Python Prophet Library using the Dataset of monthly car sales.

This Dataset is a standard univariate time series dataset consisting of both a trend as well as seasonality. The Dataset contains 108 months of data, and a naïve persistence forecast can accomplish a mean absolute error of around 3,235 sales, offering a lower error limit.

Let us begin by loading and summarizing the Dataset

Loading and Summarizing Dataset :

The prophet needs data to be stored in the form of Pandas Data frames. Thus, we will load and summarize the data with the help of the Pandas library.

We can load the data directly from the URL by calling the Pandas **read_csv()** function, summarizing the shape (number of rows and columns) of the data, and looking at the first few rows of the data.

PYTHON SAMPLE CODE :

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

data = pd.read_csv("Sales.csv")
print(data.head())
```

OUTPUT:

```
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
```

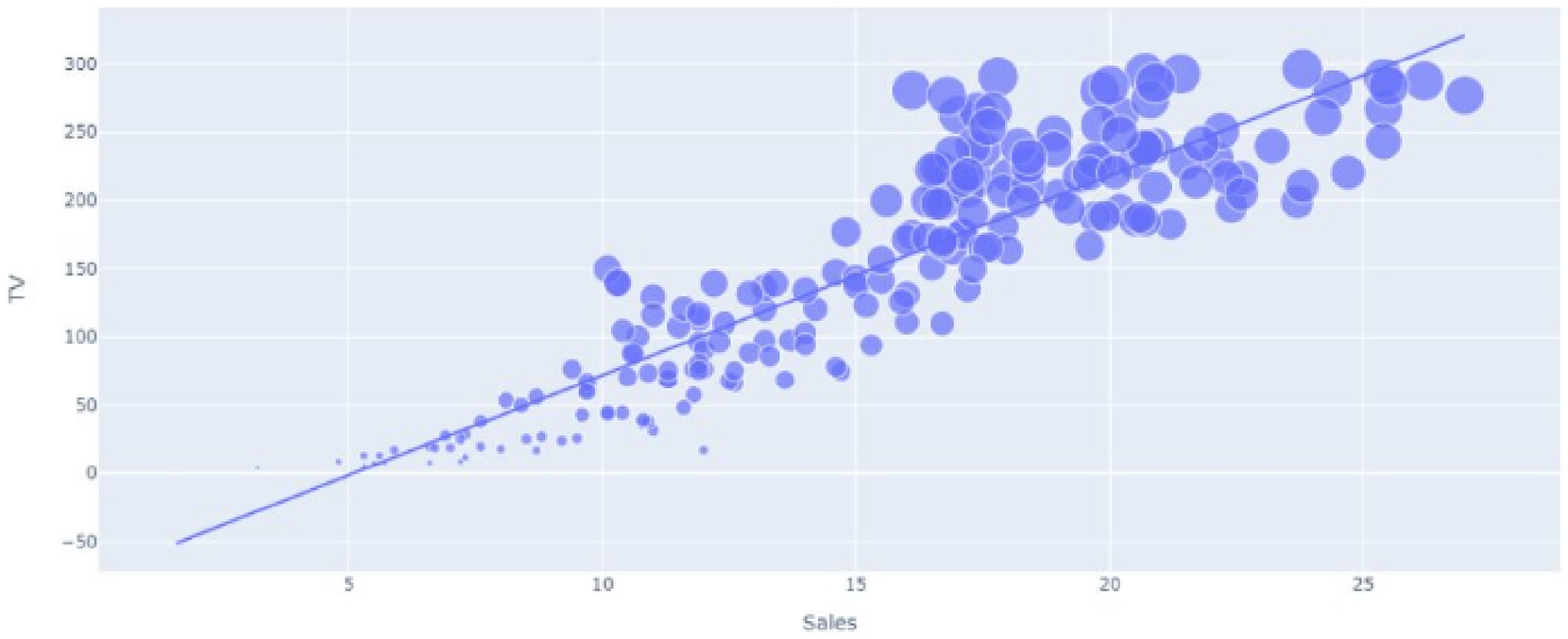
```
print(data.isnull().sum())
```

OUTPUT:

```
TV                0
Radio             0
Newspaper         0
Sales             0
dtype: int64
```

```
import plotly.express as px
import plotly.graph_objects as go
figure = px.scatter(data_frame = data, x="Sales",
                    y="TV", size="TV", trendline="ols")
figure.show()
```

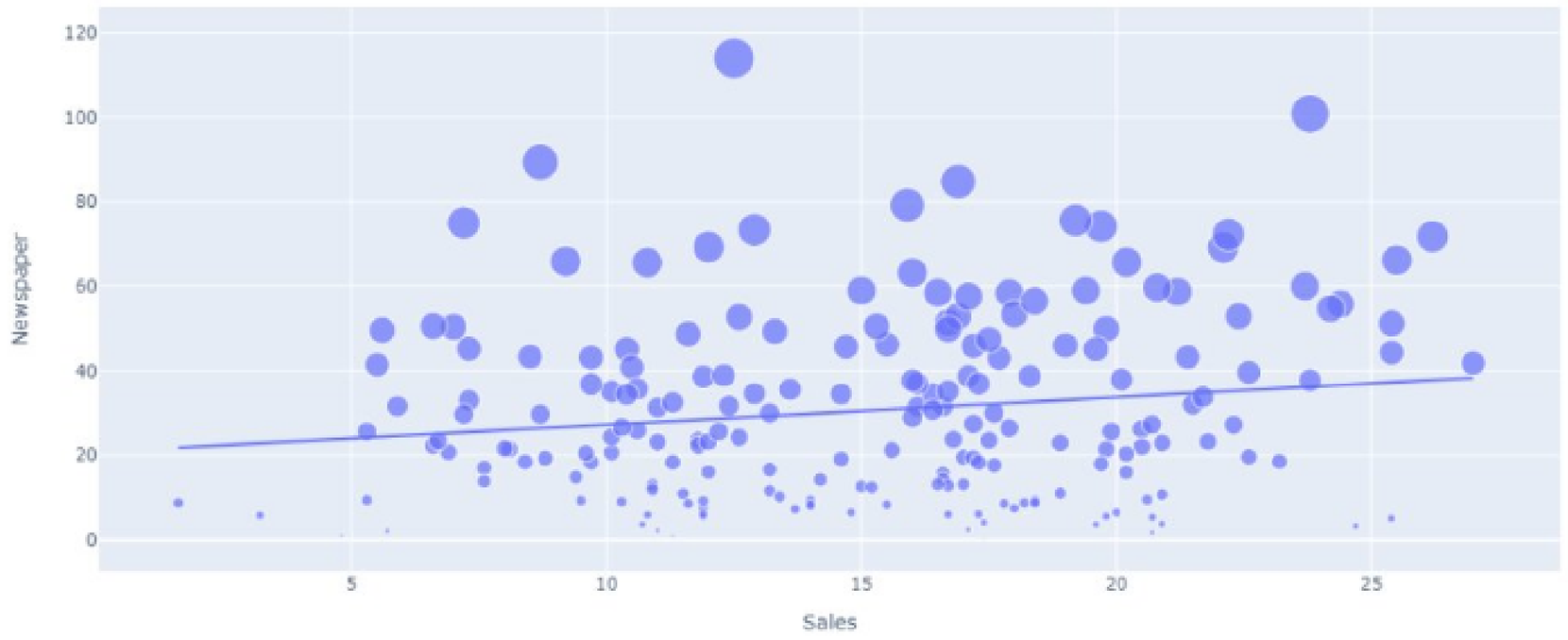
OUTPUT:




```
figure = px.scatter(data_frame = data, x="Sales",
                    y="Newspaper", size="Newspaper", trendline="ols")

figure.show()
```

OUTPUT:




```
correlation = data.corr()
print(correlation["Sales"].sort_values(ascending=False))
```

OUTPUT:

```
Sales        1.000000
TV           0.901208
Radio        0.349631
Newspaper    0.157960
Name: Sales, dtype: float64
```

CONCLUSION :

So ,this is how we can train a machine learning model to predict the future sales of a product. Predicting the future sales of a product helps a business manage the manufacturing and advertising cost of the product. I hope you liked this article on future sales prediction with machine learning. Feel free to ask valuable questions in the comments section below..