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. Hence the task of future sales prediction with machine learning using **Python** has been achieved in this project.

Future Sales Prediction (Case Study)

The **dataset** chosen 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.

Future Sales Prediction using Python

Import the necessary Python libraries and the dataset:

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegressi
data =
pd.read_csv("https://raw.githubusercontent.com/amankharwal/Website-da
ta/master/advertising.csv")
print(data.head())
```

```
   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
```

Checking for any null values:

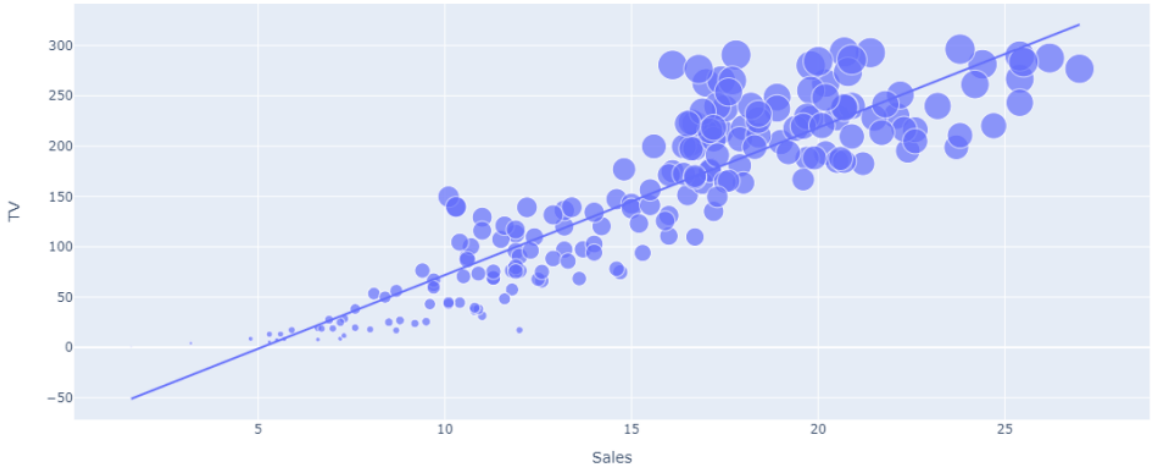
```
print(data.isnull().sum())
```

```
TV           0
Radio        0
Newspaper    0
Sales        0
```

dtype: int64

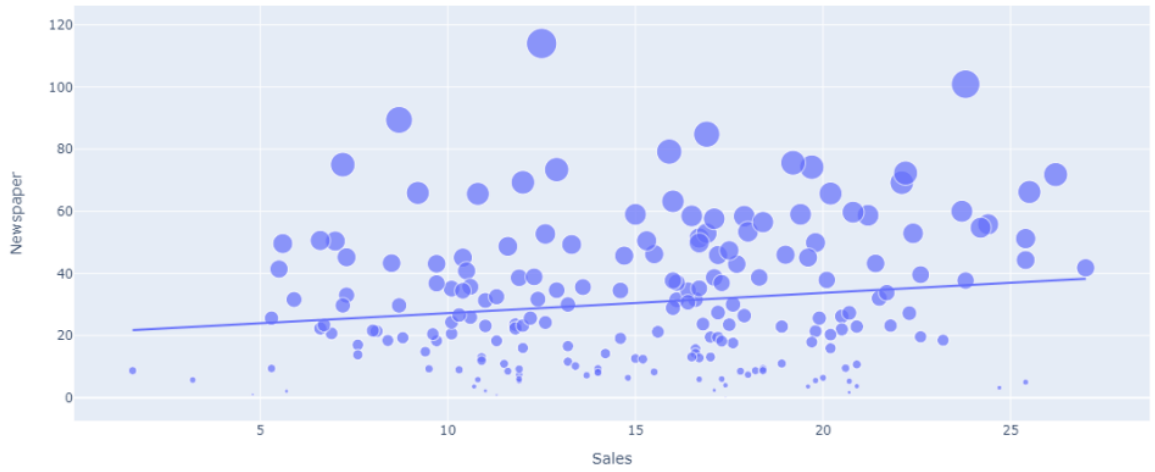
So this dataset doesn't have any null values. Now let's visualize the relationship between the amount spent on advertising on TV and units sold:

```
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()
```



Now let's visualize the relationship between the amount spent on advertising on newspapers and units sold:

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



Now let's visualize the relationship between the amount spent on advertising on radio and units sold:

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

Out of all the amount spent on advertising on various platforms, we can see that the amount spent on advertising the product on TV results in more sales of the product. Now let's have a look at the correlation of all the columns with the sales column:

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

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

Future Sales Prediction Model

Now in this section, we will train a machine learning model to predict the future sales of a product. But before, let's split the data into training and test sets:

```
x = np.array(data.drop(["Sales"], 1))
y = np.array(data["Sales"])
xtrain, xtest, ytrain, ytest = train_test_split(x, y,
test_size=0.2,
random_state=42)
```

Now let's train the model to predict future sales:

```
model = LinearRegression()
model.fit(xtrain, ytrain)
print(model.score(xtest, ytest))
```

Now let's input values into the model according to the features we have used to train it and predict how many units of the product can be sold based on the amount spent on its advertising on various platforms:

```
#features = [[TV, Radio, Newspaper]]
features = np.array([[230.1, 37.8, 69.2]])
print(model.predict(features))
```

```
[21.37254028]
```

Summary

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.