In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

In [2]:
```
ecom_data=pd.read_csv('Ecommerce Customers')
```

In [5]:
```
#ecom_data.inf()
#ecom_data.head()
ecom_data.corr()
```

Out[5]:

	Avg. Session Length	Time on App	Time on Website	Length of Membership	Yearly Amount Spent
Avg. Session Length	1.000000	-0.027826	-0.034987	0.060247	0.355088
Time on App	-0.027826	1.000000	0.082388	0.029143	0.499328
Time on Website	-0.034987	0.082388	1.000000	-0.047582	-0.002641
Length of Membership	0.060247	0.029143	-0.047582	1.000000	0.809084
Yearly Amount Spent	0.355088	0.499328	-0.002641	0.809084	1.000000

In [6]:
```
ecom_data.columns
```

Out[6]:
```
Index(['Email', 'Address', 'Avatar', 'Avg. Session Length', 'Time on App',
       'Time on Website', 'Length of Membership', 'Yearly Amount Spent'],
      dtype='object')
```

In [9]:
```
x=ecom_data[['Avg. Session Length', 'Time on App',
       'Time on Website', 'Length of Membership']]
y=ecom_data['Yearly Amount Spent']
```

In [10]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.4,random_state=101)
```

In [11]:
```
from sklearn.linear_model import LinearRegression
```

In [12]:
```
lm=LinearRegression()

lm.fit(x_train,y_train)
```

Out[12]:
```
LinearRegression()
```

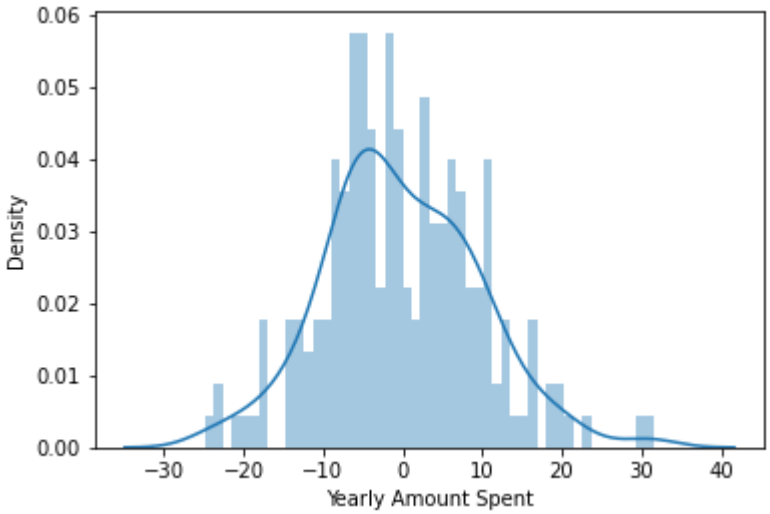
In [14]:
```
prediction = lm.predict(x_test)
```

In [16]:
```
#sns.scatterplot(y_test, prediction)
sns.distplot(y_test- prediction,bins=50)
```

```
C:\Users\admin\anaconda3\lib\site-packages\seaborn\distributions.py:2551: FutureWarning:
`distplot` is a deprecated function and will be removed in a future version. Please adap
```

```
t your code to use either `displot` (a figure-level function with similar flexibility) o
r `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

Out[16]: <AxesSubplot:xlabel='Yearly Amount Spent', ylabel='Density'>



In [17]:
```
lm.coef_
```

Out[17]:
```
array([25.69154034, 37.89259966,  0.56058149, 61.64859402])
```

In [24]:
```
coeff_df=pd.DataFrame(lm.coef_, x_train.columns,columns=['coefficent'])
coeff_df
#x_test.columns
```

Out[24]:

	coefficent
Avg. Session Length	25.691540
Time on App	37.892600
Time on Website	0.560581
Length of Membership	61.648594

In [ ]:
```

```