In [40]:
```
import matplotlib .pyplot as plt
import seaborn as sns
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from xgboost import XGBRegressor
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error


df = pd.read_excel(r"D:\\Downloads\\online+retail\\Online Retail.xlsx")
df.head(5)
df.tail(5)
```

Out[40]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID
541904	581587	22613	PACK OF 20 SPACEBOY NAPKINS	12	2011-12-09 12:50:00	0.85	12680.0
541905	581587	22899	CHILDREN'S APRON DOLLY GIRL	6	2011-12-09 12:50:00	2.10	12680.0
541906	581587	23254	CHILDRENS CUTLERY DOLLY GIRL	4	2011-12-09 12:50:00	4.15	12680.0
541907	581587	23255	CHILDRENS CUTLERY CIRCUS PARADE	4	2011-12-09 12:50:00	4.15	12680.0
541908	581587	22138	BAKING SET 9 PIECE RETROSPOT	3	2011-12-09 12:50:00	4.95	12680.0

In [4]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 541909 entries, 0 to 541908
Data columns (total 8 columns):
 #   Column       Non-Null Count   Dtype
---  ------       --------------   -----
 0   InvoiceNo    541909 non-null  object
 1   StockCode    541909 non-null  object
 2   Description  540455 non-null  object
 3   Quantity     541909 non-null  int64
 4   InvoiceDate  541909 non-null  datetime64[ns]
 5   UnitPrice    541909 non-null  float64
 6   CustomerID   406829 non-null  float64
 7   Country      541909 non-null  object
dtypes: datetime64[ns](1), float64(2), int64(1), object(4)
memory usage: 33.1+ MB
```

In [6]:
```
df.isnull().sum()
```

Out[6]:
```
InvoiceNo          0
StockCode          0
Description      1454
Quantity           0
InvoiceDate        0
UnitPrice          0
CustomerID     135080
Country            0
dtype: int64
```

In [8]:
```
df.shape[0]
```

Out[8]:
```
541909
```

In [10]:
```
df.dropna(subset=['CustomerID'], inplace=True)
```

In [12]:
```
df.isnull().sum()
```

Out[12]:
```
InvoiceNo      0
StockCode      0
Description    0
Quantity       0
InvoiceDate    0
UnitPrice      0
CustomerID     0
Country        0
dtype: int64
```

In [18]:
```
df = df[df['Quantity'] > 0]
df
```

Out[18]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0
1	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0
...	...	...	...	...	...	...	...
541904	581587	22613	PACK OF 20 SPACEBOY NAPKINS	12	2011-12-09 12:50:00	0.85	12680.0
541905	581587	22899	CHILDREN'S APRON DOLLY GIRL	6	2011-12-09 12:50:00	2.10	12680.0
541906	581587	23254	CHILDRENS CUTLERY DOLLY GIRL	4	2011-12-09 12:50:00	4.15	12680.0
541907	581587	23255	CHILDRENS CUTLERY CIRCUS PARADE	4	2011-12-09 12:50:00	4.15	12680.0
541908	581587	22138	BAKING SET 9 PIECE RETROSPOT	3	2011-12-09 12:50:00	4.95	12680.0

397884 rows × 8 columns

In [19]:
```
df = df[df['UnitPrice'] > 0]
df
```

Out[19]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0
1	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850.0
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	2010-12-01 08:26:00	3.39	17850.0
...	...	...	...	...	...	...	...
541904	581587	22613	PACK OF 20 SPACEBOY NAPKINS	12	2011-12-09 12:50:00	0.85	12680.0
541905	581587	22899	CHILDREN'S APRON DOLLY GIRL	6	2011-12-09 12:50:00	2.10	12680.0
541906	581587	23254	CHILDRENS CUTLERY DOLLY GIRL	4	2011-12-09 12:50:00	4.15	12680.0
541907	581587	23255	CHILDRENS CUTLERY CIRCUS PARADE	4	2011-12-09 12:50:00	4.15	12680.0
541908	581587	22138	BAKING SET 9 PIECE RETROSPOT	3	2011-12-09 12:50:00	4.95	12680.0

397884 rows × 8 columns

In [15]:
```
df['InvoiceDate'] = pd.to_datetime(df['InvoiceDate'])
df['InvoiceDate']
```

Out[15]:
```
0        2010-12-01 08:26:00
1        2010-12-01 08:26:00
2        2010-12-01 08:26:00
3        2010-12-01 08:26:00
4        2010-12-01 08:26:00
                 ...
541904   2011-12-09 12:50:00
541905   2011-12-09 12:50:00
541906   2011-12-09 12:50:00
541907   2011-12-09 12:50:00
541908   2011-12-09 12:50:00
Name: InvoiceDate, Length: 397884, dtype: datetime64[ns]
```

feature engineering

In [20]:
```
snapshot_date = df['InvoiceDate'].max() + pd.Timedelta(days=1)
snapshot_date
```

Out[20]:
```
Timestamp('2011-12-10 12:50:00')
```

In [21]:
```
df['TotalPrice'] = df['Quantity'] * df['UnitPrice']
df['TotalPrice']
```

Out[21]:
```
0         15.30
1         20.34
2         22.00
3         20.34
4         20.34
          ...
541904    10.20
541905    12.60
541906    16.60
541907    16.60
541908    14.85
Name: TotalPrice, Length: 397884, dtype: float64
```

In [23]:
```
df.head(3)
```

Out[23]:

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Cou
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850.0	Un Kingc
1	536365	71053	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850.0	Un Kingc
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850.0	Un Kingc

In [24]:
```
rfm = df.groupby('CustomerID').agg({
    'InvoiceDate': lambda x: (snapshot_date - x.max()).days,
    'InvoiceNo': 'count',
    'TotalPrice': 'sum'
}).reset_index()
rfm
```

Out[24]:

	CustomerID	InvoiceDate	InvoiceNo	TotalPrice
0	12346.0	326	1	77183.60
1	12347.0	2	182	4310.00
2	12348.0	75	31	1797.24
3	12349.0	19	73	1757.55
4	12350.0	310	17	334.40
...	...	...	...	...
4333	18280.0	278	10	180.60
4334	18281.0	181	7	80.82
4335	18282.0	8	12	178.05
4336	18283.0	4	756	2094.88
4337	18287.0	43	70	1837.28

4338 rows × 4 columns

In [25]:
```
rfm.columns = ['CustomerID', 'Recency', 'Frequency', 'Monetary']
rfm.columns
```

Out[25]:
```
Index(['CustomerID', 'Recency', 'Frequency', 'Monetary'], dtype='object')
```

In [26]:
```
rfm['AOV'] = rfm['Monetary'] / rfm['Frequency']
rfm['AOV']
```

Out[26]:
```
0       77183.600000
1          23.681319
2          57.975484
3          24.076027
4          19.670588
              ...
4333       18.060000
4334       11.545714
4335       14.837500
4336        2.771005
4337       26.246857
Name: AOV, Length: 4338, dtype: float64
```

In [48]:
```
#  Prepare Features and Targe
```

In [30]:
```
X = rfm[['Recency', 'Frequency', 'AOV']]
y = rfm['Monetary']
X ,y
```

Out[30]:
```
(      Recency  Frequency           AOV
 0         326          1  77183.600000
 1           2        182     23.681319
 2          75         31     57.975484
 3          19         73     24.076027
 4         310         17     19.670588
 ...       ...        ...           ...
 4333      278         10     18.060000
 4334      181          7     11.545714
 4335        8         12     14.837500
 4336        4        756      2.771005
 4337       43         70     26.246857

 [4338 rows x 3 columns],
 0       77183.60
 1        4310.00
 2        1797.24
 3        1757.55
 4         334.40
            ...
 4333      180.60
 4334       80.82
 4335      178.05
 4336     2094.88
 4337     1837.28
 Name: Monetary, Length: 4338, dtype: float64)
```

Train-Test Split

In [35]:
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_sta
X_train, X_test, y_train, y_test
```

Out[35]:
```
(      Recency  Frequency        AOV
227        212         86  31.207209
964        120          5  50.420000
2045         3         68  18.922059
1025       219         38  16.068158
4240        13        140   5.249571
...        ...        ...        ...
3444       157         22  29.182727
466         31        215  19.382419
3092       364          3  57.900000
3772        29        105  27.857048
860         18         42  27.597857

[3470 rows x 3 columns],
      Recency  Frequency         AOV
3975       53         13   10.149231
1448      207         31   15.412581
2662       17         65   16.226615
17          8         85   15.448235
1634      111          4  358.000000
...       ...        ...         ...
3466       36         25   33.643200
2643        8         60   15.920667
416         4        154    9.972532
1616      186          5   20.650000
2478      234          7   21.471429

[868 rows x 3 columns],
227     2683.82
964      252.10
2045    1286.70
1025     610.59
4240     734.94
         ...
3444     642.02
466     4167.22
3092     173.70
3772    2924.99
860     1159.11
Name: Monetary, Length: 3470, dtype: float64,
3975     131.94
1448     477.79
2662    1054.73
17      1313.10
1634    1432.00
         ...
3466     841.08
2643     955.24
416     1535.77
1616     103.25
2478     150.30
Name: Monetary, Length: 868, dtype: float64)
```

Model Training (XGBoost)

In [38]:
```
xgb_model = XGBRegressor()
xgb_model.fit(X_train, y_train)
y_pred_xgb = xgb_model.predict(X_test)
y_pred_xgb
```

Out[38]:
```
array([ 1.46377472e+02,  4.93421753e+02,  1.06320459e+03,  1.38473877e+03,
        1.89829700e+03,  1.14612354e+03,  4.04853516e+02,  2.33366318e+02,
        9.11333069e+02,  5.69715942e+02,  2.38526099e+03,  5.38625391e+04,
        1.07255481e+03,  1.06218681e+02,  1.97317249e+03,  2.21340576e+03,
        5.06940308e+02,  1.79440808e+03,  1.20605640e+03,  1.53019974e+02,
        1.12093682e+04,  3.84862964e+03,  2.93449707e+02,  2.28129736e+03,
        2.40683032e+03,  3.34879700e+02,  3.20638843e+03,  4.29376068e+02,
        2.66751001e+03,  4.15355713e+02,  4.25988831e+02,  1.08996582e+03,
        5.67015625e+02,  3.36070648e+02,  1.89729202e+02,  1.99156689e+03,
        1.69977051e+04,  1.20017725e+03,  1.11379492e+03,  2.34253931e+03,
        4.54045117e+03,  3.65623840e+02,  8.54273376e+02,  1.93969925e+02,
        1.52360315e+03,  3.57456451e+02,  2.72105133e+02,  1.22224121e+03,
        2.14848373e+02,  2.01050037e+03,  6.01108521e+02,  1.25954980e+03,
        5.06988617e+02,  2.22686676e+02,  1.79482996e+03,  2.25732468e+02,
        1.06366284e+03,  1.01866748e+04,  1.28342371e+03,  1.96419348e+03,
        3.95147552e+02,  2.09975754e+02,  1.09743274e+03,  5.22247437e+02,
        2.22209692e+03,  3.73118561e+02,  1.74175830e+03,  1.11839414e+04,
        3.27362622e+03,  6.96266663e+02,  4.44466766e+02,  8.73438477e+03,
        3.45217438e+02,  1.94604843e+02,  7.26584534e+02,  5.36620239e+02,
        4.51156836e+03,  1.42935266e+03,  7.44186584e+02,  2.93405182e+02,
        3.90834015e+02,  2.86101044e+02,  5.24014160e+02,  4.15115784e+02,
        3.89631256e+02,  6.93132891e+04,  1.39540787e+02,  3.12035553e+02,
        5.28877930e+02,  1.57160925e+03,  2.00684402e+02,  1.96231030e+03,
        7.93905334e+02,  5.90467407e+02,  1.38485828e+03,  3.26352356e+02,
        6.00026489e+02,  2.39220508e+03,  6.41759766e+02,  2.29907275e+03,
        4.36574609e+03,  1.33891553e+03,  2.16928024e+02,  3.21662256e+03,
        3.79261993e+02,  1.47358813e+03,  7.32383484e+02,  6.43398132e+02,
        1.69103796e+03,  1.00144739e+03,  3.71476257e+02,  8.08250000e+02,
        3.23461121e+02,  3.15170990e+02,  1.09745044e+03,  2.14338669e+02,
        6.94352478e+02,  6.60589453e+03,  2.82578931e+03,  7.66515747e+02,
        5.99504028e+02,  5.60152161e+02,  3.95327539e+03,  1.83044861e+03,
        3.36303802e+02,  1.57429224e+03,  1.33535538e+02,  8.14698975e+02,
        1.99733179e+03,  8.41999329e+02,  4.26215668e+02,  1.48260388e+03,
        1.37613449e+02,  6.35460938e+02,  1.26560059e+03,  1.37639282e+03,
        2.47303833e+02,  6.45903076e+02,  1.00124023e+03,  2.02167346e+03,
        2.63623535e+03,  8.07185364e+02,  9.52630859e+02,  7.68837891e+02,
        3.23307556e+02,  7.44186584e+02,  5.39098816e+02,  1.49782312e+03,
        3.40530664e+03,  5.03637744e+03,  1.25454932e+03,  2.07020410e+03,
        2.86551544e+02,  2.63793457e+02,  2.70180005e+03,  7.55380325e+01,
        8.65244507e+02,  2.26931177e+03,  5.54551453e+02,  5.12270947e+03,
        4.22859924e+02,  6.62345215e+02,  8.00548279e+02,  1.16791428e+02,
        9.93204468e+02,  2.13315649e+03,  4.90492859e+02,  5.30031250e+02,
        4.72333350e+03,  2.20780334e+02,  8.95535583e+02,  4.95368115e+03,
        1.04152722e+03,  1.43489612e+03,  1.89455002e+02,  4.06297089e+02,
        2.99674591e+02,  1.99038855e+03,  2.52193497e+02,  3.01091064e+02,
        2.91404205e+02,  1.93588272e+02,  2.88538391e+02,  7.75409119e+02,
        1.73129004e+03,  5.00072314e+03,  1.35360123e+02,  1.70297836e+02,
        1.27923134e+02,  2.04892899e+02,  1.47739548e+02,  1.53750031e+02,
        3.40899963e+02,  1.44309021e+02,  6.96751938e+01,  2.63371704e+02,
        9.96018616e+02,  7.08572815e+02,  1.91019714e+03,  6.71254395e+02,
        7.60976746e+02,  3.11929199e+03,  1.19830835e+03,  2.05763428e+02,
        9.93070755e+01,  6.20124316e+03,  1.09798073e+02,  7.39660767e+02,
        4.13749573e+02,  1.07435791e+03,  8.93507344e+04,  4.51782013e+02,
        2.18582520e+03,  1.56490771e+03,  3.81028076e+02,  2.08281082e+02,
        7.94889709e+02,  1.39975494e+02,  1.93362634e+03,  3.98824727e+04,
        1.09888745e+03,  1.41048645e+03,  8.89982422e+02,  5.47988464e+02,
```

```
 6.32837280e+02,  7.90741748e+03,  4.03356445e+02,  5.41960571e+02,
 1.63192163e+03,  1.56832739e+03,  4.46899445e+02,  8.31041199e+02,
 1.30665100e+03,  6.28829688e+04,  1.33319763e+02,  1.53137238e+02,
 1.34065796e+03,  7.34981323e+02,  1.64721878e+02,  1.87855823e+03,
 9.55023376e+02,  1.47784883e+04,  3.17455261e+02,  1.08629785e+03,
 1.41548523e+02,  1.14062366e+03,  6.30207520e+02,  6.45893677e+02,
 5.83146240e+02,  5.42544434e+02,  6.68073547e+02,  7.04259094e+02,
 4.55706024e+02,  4.25116455e+02,  4.40820312e+02,  8.11035095e+02,
 3.09650909e+02,  6.21295593e+02,  1.52054382e+02,  2.09657013e+02,
 1.25912338e+02,  3.92193628e+03,  9.48955322e+02,  6.35211975e+02,
 2.42919510e+02,  1.08699806e+02,  1.14175636e+02,  3.00929810e+03,
 1.56949548e+03,  7.57021667e+02,  2.55054626e+02,  2.78331421e+03,
 1.92522473e+03,  2.27886279e+03,  1.11340149e+03,  8.04957642e+02,
 2.16204956e+02,  1.20007056e+03,  1.12656555e+03,  3.93220752e+03,
 4.30056702e+02,  1.14835339e+03,  3.93161523e+03,  5.34093445e+02,
 8.55989502e+02,  6.55970764e+02,  3.10582275e+02,  7.64321045e+02,
 2.27223892e+02,  2.70726294e+03,  2.43270096e+02,  1.60450714e+02,
 3.91668213e+02,  5.65305481e+02,  1.83939685e+03,  3.08152069e+02,
 8.08338989e+02,  3.22892548e+02,  6.07459229e+02,  5.39973389e+02,
 7.73388428e+02,  5.90377075e+02,  3.29329468e+02,  2.46458557e+02,
 1.07338940e+03,  8.50226517e+01,  3.96904395e+03,  8.96962036e+02,
 4.43570648e+02,  6.26015015e+02,  3.65729126e+02,  6.20792358e+02,
 2.38556534e+02,  3.86691956e+02,  1.14032361e+03,  7.87843604e+03,
 1.68654999e+02,  5.93335632e+02,  3.87997932e+01,  3.99845551e+02,
 2.18719238e+02,  5.37428833e+02,  1.78933090e+02,  5.44272803e+03,
 1.63994824e+03,  1.87629395e+03,  1.67345413e+02,  4.82045361e+03,
 3.28696930e+02,  2.83850317e+03,  7.24733337e+02,  9.18585156e+03,
 2.91652271e+03,  1.56238983e+02,  4.73058685e+02,  1.10502673e+03,
 1.92695511e+02,  1.40154209e+04,  2.98725525e+02,  8.83937866e+02,
 1.75234746e+04,  3.21177271e+03,  9.42202854e+00,  8.29298218e+02,
 4.22705200e+02,  8.22006836e+02,  6.57400146e+02,  5.65914062e+02,
 1.44420276e+03,  4.33975128e+02,  7.38145691e+02,  9.12824219e+02,
 1.38563428e+03,  8.22559586e+01,  1.65584326e+03,  3.81212622e+03,
 2.81022754e+03,  1.62413925e+02,  6.97716003e+02,  1.86266260e+03,
 2.47644196e+02,  3.24301422e+02,  3.66699365e+03,  1.26791748e+04,
 9.78167053e+02,  2.50319971e+03,  2.39535278e+03,  6.24304565e+02,
 1.29403711e+03,  2.35802319e+03,  8.44057846e+01,  3.80428009e+01,
 4.24022888e+02,  1.70713965e+04,  2.83135669e+03,  3.57644440e+02,
 1.00135297e+03,  1.91916702e+02,  2.55473862e+02,  6.06052063e+02,
 7.41475525e+02,  1.31922516e+02,  4.83749688e+04,  3.55759521e+02,
 3.17403259e+02,  2.39563583e+02,  4.82040771e+02,  8.48919128e+02,
 6.27815186e+02,  2.50562408e+02,  7.92821198e+01,  1.60457443e+02,
 4.63313916e+03,  6.57235168e+02,  1.49147797e+02,  1.59762720e+03,
-7.05842743e+01,  2.58425629e+02,  4.38336792e+02,  3.25041412e+02,
 8.83937866e+02,  4.36735938e+03,  1.35076062e+03,  4.36147034e+02,
 1.77499402e+03,  1.66266449e+02,  7.84097351e+02,  3.03365753e+02,
 5.36836975e+02,  1.52915771e+02,  1.89301123e+03,  1.40072852e+03,
 9.10921387e+02,  2.96010895e+02,  6.38174988e+02,  3.48653369e+03,
 1.29735168e+02,  1.04102551e+03,  7.20705078e+02,  6.38437805e+02,
 2.79350250e+02,  2.32864624e+02,  4.87198389e+03,  2.28451935e+02,
 6.26844849e+02,  4.57715674e+03,  2.51774475e+02,  2.30132202e+03,
 4.12394453e+04,  1.58332983e+03,  4.63389435e+02,  9.09033508e+02,
 4.59943817e+02,  2.52691284e+03,  6.57983337e+02,  1.94897446e+02,
 3.06991882e+02,  8.53292969e+03,  6.05830139e+02,  1.16918242e+04,
 4.86741162e+03,  3.50426880e+02,  2.84751434e+02,  3.76014023e+01,
 1.21682861e+03,  2.34399734e+02,  4.95700439e+02,  7.37873840e+02,
```

```
1.18199854e+03,  1.32988632e+02,  1.95015930e+03,  4.59150977e+03,
2.10490646e+02,  6.39852881e+03,  6.56362488e+02,  1.40461655e+02,
3.47083954e+02,  3.10682037e+02,  2.10910352e+04,  7.69677185e+02,
1.83031448e+02,  1.10334558e+03,  2.08900253e+02,  6.22961121e+02,
3.79040918e+03,  9.98761230e+02,  8.89605530e+02,  5.35403198e+02,
1.54879600e+04,  8.01049683e+02,  4.13130432e+02,  2.10197363e+03,
4.85942497e+01,  5.71201294e+02,  5.02083252e+02,  2.54383936e+03,
3.94139282e+02,  6.00540527e+02,  1.12183716e+03,  1.84717041e+03,
6.61284521e+03,  1.13458594e+03,  8.02525940e+02,  4.86023224e+02,
3.39822021e+03,  4.38381592e+03,  3.26279846e+02,  3.69724579e+02,
1.70419189e+02,  3.99323438e+03,  1.58979834e+03,  1.44575403e+03,
3.40634430e+02,  3.83406769e+02,  5.07424805e+02,  3.83152374e+02,
6.55843933e+02,  2.48230850e+02,  1.97402435e+02,  1.84830774e+03,
1.59275513e+02,  4.13171045e+03,  2.00210217e+03,  3.67340088e+03,
2.18907007e+03,  5.29928040e+02,  3.89458643e+03,  1.57682617e+03,
2.04585251e+02,  1.04738477e+03,  2.51665405e+02,  4.89452832e+03,
1.42420670e+02,  8.19442825e+01,  4.83264496e+02,  5.47963074e+02,
1.28654346e+03,  9.87569763e+02,  4.49194580e+03,  1.23741611e+04,
2.45849030e+02,  7.23278503e+02,  5.06383514e+02,  8.69941650e+02,
8.01937714e+01,  1.68380688e+03,  6.26604919e+02,  1.35960358e+02,
2.96527026e+03,  1.31558884e+02,  2.71379126e+03,  8.91107422e+02,
1.94640283e+03,  1.27797882e+02,  6.83674683e+02,  5.80524109e+02,
1.57819543e+03,  1.39539990e+03,  4.34940977e+04,  1.32988632e+02,
1.53526538e+03,  3.57766553e+03,  4.29456360e+02,  1.54584244e+02,
2.28912939e+03,  1.95385803e+02,  1.62565720e+02,  2.78373375e+01,
2.54704453e+02,  6.00549683e+02,  2.49288116e+02,  4.25315094e+02,
2.93233545e+03,  1.44575403e+03,  4.60390663e+01,  1.63192902e+02,
6.93995972e+02,  1.68518713e+03,  2.57673281e+04,  6.57400146e+02,
6.24003125e+03,  3.21653778e+02,  8.08011780e+01,  1.61327588e+03,
3.01850189e+02,  1.11838086e+03,  8.58914160e+03,  1.39694861e+03,
4.61495117e+02,  3.11929199e+03,  6.20554321e+02,  2.61160431e+02,
4.13880280e+02,  5.79773743e+02,  2.15737207e+03,  3.09146881e+02,
1.85404065e+03,  8.95535583e+02,  3.21447021e+02,  2.23696716e+02,
9.43686035e+03,  4.48068457e+03,  2.50373804e+03,  1.03951324e+02,
6.37912109e+02,  2.19797876e+03,  1.93602188e+05,  7.30734558e+02,
2.18400635e+03,  7.31759399e+02,  2.72139557e+02,  1.41893311e+02,
7.17326050e+02,  2.92941223e+02,  5.09947174e+02,  5.00948212e+02,
1.28441455e+03,  4.83509827e+02,  2.40924576e+02,  7.05120697e+01,
1.61517395e+02,  7.67594055e+02,  2.96146698e+02,  2.16125781e+03,
1.37784753e+03,  4.87946106e+02,  4.39707123e+02,  4.61897168e+03,
2.36499243e+03,  1.57632458e+03,  1.64522327e+03,  4.21354736e+02,
1.01957520e+02,  3.79876434e+02,  1.59662537e+02,  2.32028465e+01,
4.47555664e+03,  4.52784973e+02,  6.49938538e+02,  2.15679053e+03,
3.06962964e+03,  3.98882355e+02,  4.03546777e+03,  1.08858252e+03,
1.33792798e+03,  1.68205139e+02,  1.99052353e+02,  2.16956274e+03,
1.87578445e+02,  3.07476105e+02,  7.25600586e+02,  9.37810791e+02,
3.46536591e+02,  4.29208130e+02,  8.18787109e+02,  1.97091284e+03,
1.63516739e+02,  4.17342102e+02,  4.65006836e+03,  1.85715851e+02,
2.98249298e+02,  2.53264697e+03,  9.04287842e+02,  1.09915000e+04,
2.73845044e+03,  3.21025269e+02,  6.21126648e+02,  1.32591339e+02,
2.36203442e+03,  2.50797607e+02,  5.32013916e+02,  4.77127533e+02,
3.06697296e+02,  4.35341980e+02,  1.01876178e+03,  2.90857178e+02,
1.52132004e+02,  1.98528284e+03,  2.26097485e+03,  5.49152283e+02,
7.88666077e+02,  1.57635315e+03,  1.66951794e+03,  3.54445068e+02,
4.23366602e+03,  1.47135635e+02,  1.66074265e+02,  4.87524353e+02,
1.82837622e+03,  1.98868005e+03,  3.72177887e+02,  7.05697388e+02,
```

```
  1.48479590e+03,  1.26466408e+02,  6.06082268e+01,  6.84839050e+02,
  5.64623584e+03,  1.27795728e+03,  9.84226624e+02,  5.74809265e+02,
  3.94646265e+03,  5.27425098e+00,  2.65662811e+02,  5.25179736e+03,
  2.66017944e+03,  1.16765173e+03,  2.68876984e+02,  1.00511664e+03,
  2.71971680e+02,  2.63751129e+02,  2.11302780e+02,  3.10352588e+03,
  1.54296985e+03,  9.19453125e+03,  2.51384497e+03,  1.63809891e+02,
  4.71956573e+02,  2.00999048e+03,  4.56951660e+02,  1.14387170e+03,
  1.71243683e+02,  5.62699768e+02,  6.64127832e+03,  2.82042505e+03,
  8.09326401e+01,  3.77513000e+02,  1.26611462e+03,  1.78146088e+02,
  3.99826562e+03,  4.97550568e+02,  7.81608826e+02,  1.43667920e+03,
  3.99701538e+02,  1.83888153e+02,  1.69997421e+02,  1.29338440e+03,
  4.97314392e+02,  6.31153857e+03,  1.44151523e+04,  4.86266016e+03,
  1.22338892e+03,  4.85350891e+02,  3.50805695e+02,  6.55774719e+02,
  4.58934326e+03,  9.46007202e+02,  1.17638159e+03,  3.25084839e+01,
  5.39536084e+03,  2.02167346e+03,  6.76302307e+02,  1.25339038e+03,
  3.13420593e+02,  3.51074341e+02,  1.45231781e+02,  3.56205933e+02,
  1.40154209e+04,  1.65870972e+03,  1.49685144e+03,  5.35985400e+03,
  7.20805176e+02,  1.99309863e+03,  6.94623535e+02,  3.87034027e+02,
  9.84507080e+02,  1.61512952e+03,  3.21025269e+02,  6.82212031e+04,
  1.23610146e+02,  2.02178116e+02,  1.05973840e+03,  1.25261182e+03,
  9.77428436e+01,  3.17168549e+02,  9.71716736e+02,  4.10990021e+02,
  6.85038574e+02,  3.75125214e+02,  2.61846832e+02,  8.70412109e+02,
  8.21345337e+02,  1.08763184e+03,  4.76269043e+03,  2.96911670e+03,
  8.04582672e+01,  1.98794678e+02,  2.78218964e+02,  4.39143143e+01,
  3.20404083e+02,  1.23981567e+03,  3.31146759e+02,  6.90336487e+02,
  9.77994019e+02,  3.61520477e+02,  1.15271106e+03,  5.13868701e+03,
  3.83728027e+02,  3.24776904e+03,  7.71511780e+02,  6.03473633e+02,
  1.77897864e+03,  1.49930530e+03,  1.25045105e+03,  8.53197250e+01,
  1.04813147e+03,  2.85962463e+02,  3.51227448e+02,  4.22859924e+02,
  7.93878662e+02,  1.27280383e+03,  3.34456299e+02,  1.63027191e+02,
  1.29914478e+03,  3.55572900e+03,  8.86560425e+02,  7.38263379e+03,
  1.06107117e+03,  6.13099023e+03,  1.43494400e+02,  5.34093445e+02,
  4.89535645e+02,  1.25017975e+02,  2.06182910e+03,  6.15357539e+04,
  2.79938062e+03,  1.05963379e+03,  2.33119781e+02,  1.49377710e+03,
  5.94903442e+02,  1.51838501e+03,  3.62016821e+03,  7.30625671e+02,
  2.02439026e+02,  2.10734848e+02,  4.34859070e+02,  4.71643097e+02,
  1.01206873e+03,  5.06771515e+02,  1.54296985e+03,  7.10356522e+01,
  3.95532532e+02,  2.42590027e+02,  1.37613449e+02,  6.98739136e+02,
  4.69793555e+03,  2.08186157e+02,  8.35801514e+02,  2.06213965e+03,
  1.26881335e+03,  2.84993927e+02,  1.74235062e+02,  4.77192078e+02,
  4.38743896e+03,  5.67156787e+03,  1.09173950e+03,  4.30532806e+02,
  2.63623535e+03,  3.78089404e+03,  6.97768616e+02,  3.02529633e+02,
  7.91867615e+02,  1.62418237e+03,  3.29695343e+02,  7.36438232e+02,
  2.88165845e+03,  3.53395691e+02,  1.35410999e+03,  2.31289246e+02,
  3.11045654e+03,  2.07202332e+02,  3.55921448e+02,  1.04619102e+02,
  6.18235596e+02,  3.07652252e+02,  1.32783584e+04,  1.20714326e+04,
  6.81204590e+02,  2.13648120e+03,  7.32793335e+02,  4.91939819e+02,
  2.73055176e+02,  7.95610352e+03,  1.19811292e+03,  7.74874390e+02,
  9.49739075e+02,  1.42695178e+03,  8.23191833e+01,  1.62371124e+02],
dtype=float32)
```

Model Evaluation (XGBoost)

In [39]:
```
mae_xgb = mean_absolute_error(y_test, y_pred_xgb)
rmse_xgb = np.sqrt(mean_squared_error(y_test, y_pred_xgb))
print("XGBoost MAE:", mae_xgb)
print("XGBoost RMSE:", rmse_xgb)
```

```
XGBoost MAE: 526.9712263089546
XGBoost RMSE: 4162.316623172757
```

In [41]:
```
rf_model = RandomForestRegressor()
rf_model.fit(X_train, y_train)
y_pred_rf = rf_model.predict(X_test)
y_pred_rf
```

Out[41]:
```
array([1.34325900e+02, 4.81247900e+02, 1.06096260e+03, 1.30939170e+03,
       1.17127350e+03, 1.45111180e+03, 3.84103300e+02, 2.22796900e+02,
       9.96827000e+02, 5.56067700e+02, 2.55753820e+03, 4.09709200e+04,
       1.12213320e+03, 1.15578000e+02, 2.04684060e+03, 2.05652050e+03,
       4.94882000e+02, 1.87316030e+03, 1.45183200e+03, 1.57406400e+02,
       1.08153259e+04, 3.51269560e+03, 2.89319800e+02, 2.39948230e+03,
       2.49067430e+03, 3.25832000e+02, 3.28005480e+03, 4.44799200e+02,
       2.67540220e+03, 4.26431970e+02, 4.28668500e+02, 1.02611080e+03,
       5.39406200e+02, 3.45906900e+02, 1.82779500e+02, 1.99392920e+03,
       1.96246666e+04, 1.21310670e+03, 1.15178410e+03, 2.34914300e+03,
       4.84145040e+03, 3.35013700e+02, 8.85424900e+02, 1.83092500e+02,
       1.42799530e+03, 3.83019500e+02, 3.37400900e+02, 8.94206900e+02,
       2.16941000e+02, 2.09117390e+03, 5.65445500e+02, 1.26709370e+03,
       4.64704500e+02, 2.30294900e+02, 1.79839250e+03, 2.02558600e+02,
       1.05109230e+03, 8.17106650e+03, 1.29406970e+03, 2.06383800e+03,
       3.86716200e+02, 1.92143900e+02, 1.08476920e+03, 5.75416300e+02,
       2.28111550e+03, 3.61188800e+02, 1.78472770e+03, 7.88347700e+03,
       3.43490760e+03, 7.02176900e+02, 4.64332700e+02, 8.62268940e+03,
       3.55987600e+02, 1.86516600e+02, 6.65708800e+02, 5.43243000e+02,
       5.21839050e+03, 1.40484640e+03, 6.97228900e+02, 2.99991000e+02,
       3.95838600e+02, 2.88705600e+02, 5.21785300e+02, 4.21376900e+02,
       3.54604600e+02, 3.65095290e+04, 1.23677200e+02, 3.35910500e+02,
       4.92020000e+02, 1.53679610e+03, 2.05380500e+02, 2.22965700e+03,
       7.84860300e+02, 5.45915300e+02, 1.34660910e+03, 3.54583700e+02,
       5.61557800e+02, 6.73959030e+03, 5.96971900e+02, 2.25886770e+03,
       4.47641250e+03, 1.31766770e+03, 2.13094800e+02, 3.09289830e+03,
       3.87309200e+02, 1.52927650e+03, 7.16925500e+02, 6.22110500e+02,
       1.69511100e+03, 1.02911100e+03, 3.75307000e+02, 7.73757000e+02,
       3.35039200e+02, 3.37670700e+02, 1.17180800e+03, 2.00425700e+02,
       7.23634900e+02, 6.69321740e+03, 2.97713020e+03, 7.64297300e+02,
       5.90338000e+02, 5.67183500e+02, 3.63590770e+03, 1.85640640e+03,
       3.44434600e+02, 1.52591130e+03, 1.34994000e+02, 7.99530800e+02,
       2.09486510e+03, 8.48786300e+02, 4.25949800e+02, 1.47663360e+03,
       1.38405000e+02, 6.04464000e+02, 1.32008720e+03, 1.36705110e+03,
       2.42201200e+02, 6.48273300e+02, 1.05172450e+03, 2.06430500e+03,
       2.63692740e+03, 8.20524200e+02, 9.24948300e+02, 7.73461900e+02,
       3.60188200e+02, 7.05345600e+02, 4.73225500e+02, 1.58024010e+03,
       4.29680480e+03, 6.95180840e+03, 1.30901270e+03, 1.96257770e+03,
       3.21136500e+02, 2.66528400e+02, 2.69053570e+03, 5.53190000e+01,
       9.35459100e+02, 2.30287330e+03, 5.69047400e+02, 5.17986400e+03,
       3.71118300e+02, 6.54817400e+02, 8.06588200e+02, 1.04343500e+02,
       1.01684020e+03, 2.27604270e+03, 4.96216800e+02, 5.14235600e+02,
       4.52757860e+03, 2.08089200e+02, 9.08657500e+02, 4.57909740e+03,
       9.84307100e+02, 1.49322430e+03, 1.66560200e+02, 4.08350000e+02,
       2.85760500e+02, 1.90487870e+03, 2.70339400e+02, 2.83187500e+02,
       3.10789600e+02, 1.77839300e+02, 2.81712400e+02, 7.49592400e+02,
       1.68603610e+03, 4.86307490e+03, 1.23339500e+02, 2.25861100e+02,
       1.77134800e+02, 1.92451600e+02, 1.53390400e+02, 1.54466100e+02,
       3.09204200e+02, 1.42195900e+02, 6.24024000e+01, 2.61672600e+02,
       1.06426600e+03, 6.98874500e+02, 1.97352210e+03, 7.06195900e+02,
       7.55819270e+02, 3.15906720e+03, 1.15252010e+03, 1.88853200e+02,
       9.03429000e+01, 5.94565450e+03, 1.14200100e+02, 6.92172700e+02,
       3.86432400e+02, 1.06198540e+03, 7.83108059e+04, 4.41828000e+02,
       2.21909870e+03, 1.57264430e+03, 3.43696200e+02, 2.12487700e+02,
       8.20535600e+02, 1.35132100e+02, 1.78756340e+03, 7.73274528e+04,
       1.08616430e+03, 1.42476070e+03, 8.80239500e+02, 5.66787200e+02,
```

```
6.10697200e+02, 7.75691710e+03, 4.17917600e+02, 5.48241000e+02,
1.58459120e+03, 1.51653610e+03, 5.11557800e+02, 8.49400100e+02,
1.33137260e+03, 6.38973651e+04, 1.24839600e+02, 1.93794800e+02,
1.36053800e+03, 7.44449100e+02, 1.74795200e+02, 2.05645340e+03,
9.18126600e+02, 1.30138829e+04, 3.97554800e+02, 1.02659010e+03,
1.41963800e+02, 1.07051500e+03, 6.25178200e+02, 6.51250300e+02,
5.92151200e+02, 5.20196600e+02, 6.83679800e+02, 6.91519200e+02,
4.70293500e+02, 4.02958400e+02, 4.11586400e+02, 9.53470600e+02,
2.76511300e+02, 5.52399400e+02, 1.58834900e+02, 2.31941700e+02,
1.35572200e+02, 3.90945500e+03, 9.84863300e+02, 6.30394700e+02,
2.27072100e+02, 1.10084100e+02, 1.15167900e+02, 2.78275470e+03,
1.56986220e+03, 7.01693800e+02, 2.73575400e+02, 2.68589640e+03,
1.85586760e+03, 2.13763820e+03, 1.14259340e+03, 8.03077550e+02,
2.16420600e+02, 1.12612640e+03, 1.12090830e+03, 4.12269260e+03,
4.12744900e+02, 1.21393820e+03, 3.89381670e+03, 5.01015500e+02,
9.28839500e+02, 6.59690200e+02, 3.27087100e+02, 5.33925800e+02,
2.25627800e+02, 2.54800550e+03, 2.35353800e+02, 1.70653200e+02,
3.95192300e+02, 5.32236900e+02, 1.88113490e+03, 2.98737900e+02,
8.33932800e+02, 3.62169650e+02, 6.37934000e+02, 4.96545400e+02,
7.77910100e+02, 6.30838700e+02, 3.16904700e+02, 2.33815500e+02,
1.07038060e+03, 1.17125300e+02, 4.28400840e+03, 9.43139200e+02,
4.16067100e+02, 5.85123500e+02, 3.17221500e+02, 5.77318600e+02,
2.52041900e+02, 3.86867800e+02, 1.13413840e+03, 7.62201500e+03,
1.68471200e+02, 6.25539900e+02, 2.88640000e+01, 3.56261500e+02,
2.07987600e+02, 5.26686400e+02, 1.95829100e+02, 5.61598460e+03,
1.52048220e+03, 1.94494070e+03, 1.72964700e+02, 4.88301740e+03,
3.38195800e+02, 2.80271020e+03, 6.81749300e+02, 8.71514670e+03,
2.91513820e+03, 1.79803800e+02, 4.90097800e+02, 1.04396350e+03,
1.96312800e+02, 1.35049030e+04, 2.55639700e+02, 9.64480100e+02,
1.67883944e+04, 3.17604590e+03, 4.08151000e+01, 8.06808300e+02,
4.48631200e+02, 7.99984700e+02, 6.56677900e+02, 5.90708300e+02,
1.40391790e+03, 4.46703800e+02, 3.67287800e+02, 9.86518600e+02,
1.43929640e+03, 8.91435000e+01, 1.70619300e+03, 3.85112070e+03,
2.59859040e+03, 1.59206200e+02, 6.98287100e+02, 1.67191140e+03,
2.18614000e+02, 3.33109400e+02, 3.67932960e+03, 1.38733410e+04,
1.11476840e+03, 2.50516680e+03, 2.58337050e+03, 6.23511500e+02,
1.25395580e+03, 2.41313030e+03, 8.88652000e+01, 4.27380000e+01,
4.00872700e+02, 1.67614839e+04, 2.64197560e+03, 3.90511200e+02,
1.00372380e+03, 2.00441900e+02, 2.57070900e+02, 5.90333700e+02,
6.94209500e+02, 1.24548300e+02, 4.20649094e+04, 3.98510300e+02,
3.06794200e+02, 2.51248700e+02, 4.48118300e+02, 7.75593400e+02,
6.16980300e+02, 2.88541000e+02, 7.18064000e+01, 1.51906500e+02,
1.98408512e+04, 6.77807200e+02, 1.76292400e+02, 1.55645690e+03,
7.60726000e+01, 2.53491400e+02, 4.28503000e+02, 4.02717400e+02,
9.54357600e+02, 3.78851610e+03, 1.68356410e+03, 4.49309600e+02,
1.64318900e+03, 1.55147100e+02, 7.54056900e+02, 3.15714700e+02,
5.86349300e+02, 1.35305100e+02, 2.03755570e+03, 1.43293800e+03,
9.66817300e+02, 3.11962000e+02, 6.63819400e+02, 3.07107510e+03,
1.26957700e+02, 1.17514430e+03, 6.76945300e+02, 6.18398300e+02,
2.89350900e+02, 2.26967800e+02, 4.78879280e+03, 3.10238500e+02,
6.43690500e+02, 4.63755130e+03, 2.36734300e+02, 2.64168960e+03,
2.60701278e+04, 1.62617450e+03, 4.63358000e+02, 1.04312810e+03,
6.90975000e+02, 2.60560250e+03, 6.59081500e+02, 2.09744300e+02,
2.97922300e+02, 7.82209210e+03, 6.21646200e+02, 1.07960569e+04,
4.75077430e+03, 3.81346200e+02, 2.78088300e+02, 4.46460000e+01,
1.13877300e+03, 2.04701500e+02, 4.87767600e+02, 7.42065200e+02,
```

```
1.17944760e+03, 1.33966800e+02, 2.00040100e+03, 6.37634060e+03,
2.31672700e+02, 6.43811200e+03, 6.75118000e+02, 1.59481800e+02,
3.34971800e+02, 3.19733400e+02, 3.11225640e+03, 9.59503500e+02,
1.85794900e+02, 1.01653620e+03, 1.99734200e+02, 5.96190800e+02,
3.69236410e+03, 1.05399060e+03, 9.42117900e+02, 5.48701900e+02,
1.07751632e+04, 8.02431100e+02, 4.31919300e+02, 2.13205600e+03,
5.98964000e+01, 6.11875700e+02, 4.90874600e+02, 2.24319510e+03,
4.13412700e+02, 6.30352500e+02, 1.17092070e+03, 1.88932310e+03,
7.07024720e+03, 1.10062980e+03, 8.01660500e+02, 4.52717100e+02,
3.40121920e+03, 4.35833430e+03, 3.21568100e+02, 3.76995930e+02,
1.88906600e+02, 3.69083680e+03, 1.60973350e+03, 1.35376460e+03,
3.45081800e+02, 3.69502600e+02, 4.91326300e+02, 3.91580500e+02,
6.22132400e+02, 2.57460400e+02, 2.14263700e+02, 1.86233090e+03,
1.64327300e+02, 4.07836270e+03, 1.96193040e+03, 3.60057890e+03,
2.23844260e+03, 6.46221300e+02, 3.87326570e+03, 1.65694840e+03,
1.99566800e+02, 1.02075680e+03, 2.75802200e+02, 4.56239250e+03,
1.49737100e+02, 7.65413000e+01, 4.56761800e+02, 5.74694600e+02,
1.23194100e+03, 1.02978940e+03, 4.45812400e+03, 1.24598783e+04,
2.30099200e+02, 7.01782400e+02, 5.25037700e+02, 8.31527300e+02,
8.25300000e+01, 1.69968200e+03, 5.91668600e+02, 1.26923500e+02,
2.95715271e+03, 1.21881100e+02, 2.52736360e+03, 8.64170600e+02,
2.11674760e+03, 1.23795500e+02, 6.54807300e+02, 5.70993900e+02,
1.46677390e+03, 1.35127220e+03, 5.40862136e+04, 1.38011800e+02,
1.50393550e+03, 3.73913470e+03, 3.93095000e+02, 1.82630900e+02,
2.31878600e+03, 1.77263400e+02, 1.50086300e+02, 4.94691000e+01,
2.57792400e+02, 5.86751000e+02, 2.31824900e+02, 4.36629600e+02,
2.75869880e+03, 1.41221400e+03, 6.41199000e+01, 1.49805900e+02,
7.00220900e+02, 1.90768120e+03, 2.26757752e+04, 6.56312000e+02,
6.01140340e+03, 3.25610400e+02, 9.09680000e+01, 1.60983760e+03,
3.09980400e+02, 1.10794080e+03, 9.93966030e+03, 1.44565020e+03,
4.75060600e+02, 3.11354520e+03, 5.96186200e+02, 2.47851200e+02,
4.53486100e+02, 5.18490400e+02, 2.39008390e+03, 3.15053600e+02,
1.80922840e+03, 9.07108100e+02, 3.14989800e+02, 2.25583300e+02,
8.04900070e+03, 4.37040690e+03, 2.50077900e+03, 9.83796000e+01,
5.64264100e+02, 2.12770880e+03, 1.12673437e+05, 7.07516200e+02,
2.20722870e+03, 6.65718500e+02, 2.49305200e+02, 1.59343500e+02,
7.41454600e+02, 2.93745000e+02, 4.80567300e+02, 5.18195100e+02,
1.21436570e+03, 4.88500300e+02, 2.42870700e+02, 8.62870000e+01,
1.57829500e+02, 7.41458100e+02, 2.83940400e+02, 1.41658990e+03,
1.36727970e+03, 4.59856900e+02, 4.20110200e+02, 4.18369340e+03,
2.27432080e+03, 1.44323110e+03, 1.63120100e+03, 4.09687000e+02,
1.07012900e+02, 3.27560000e+02, 2.07354100e+02, 4.07535000e+01,
4.04862880e+03, 4.20987000e+02, 6.29515700e+02, 1.88908290e+03,
3.11945860e+03, 3.90396900e+02, 4.10893910e+03, 1.02394140e+03,
1.27836600e+03, 1.64114800e+02, 2.19215300e+02, 2.12547700e+03,
1.68562500e+02, 2.97276200e+02, 7.28283700e+02, 9.17764500e+02,
3.54586600e+02, 4.42775700e+02, 8.41332300e+02, 1.93469470e+03,
1.65566800e+02, 4.12354200e+02, 4.73496470e+03, 1.86192300e+02,
3.17054800e+02, 2.53744200e+03, 8.25058100e+02, 1.08574977e+04,
3.22392060e+03, 3.14958500e+02, 5.98155700e+02, 1.15292900e+02,
2.32318290e+03, 2.27619100e+02, 3.97871800e+02, 4.72683600e+02,
2.80237800e+02, 4.14189900e+02, 9.94079900e+02, 3.01575000e+02,
1.45360600e+02, 1.98424940e+03, 2.28865600e+03, 5.76324500e+02,
8.77098000e+02, 1.49058180e+03, 1.97942070e+03, 3.69062700e+02,
4.11430460e+03, 1.56447400e+02, 1.65654800e+02, 4.58068200e+02,
1.88112970e+03, 1.91916280e+03, 3.75548200e+02, 7.07397000e+02,
```

```
1.55440050e+03, 1.14979400e+02, 7.58189000e+01, 7.41937200e+02,
5.69085880e+03, 1.21507560e+03, 9.78079400e+02, 6.25665400e+02,
3.95256190e+03, 1.29000000e+01, 2.59146800e+02, 5.66517340e+03,
2.73112010e+03, 1.18295800e+03, 2.20566700e+02, 1.01966940e+03,
2.86193100e+02, 2.76013600e+02, 2.37366400e+02, 3.00087230e+03,
1.51109710e+03, 7.74909730e+03, 2.83004670e+03, 1.59786300e+02,
5.73188000e+02, 2.07335890e+03, 5.09800300e+02, 1.07309010e+03,
1.66804500e+02, 5.58411100e+02, 6.82067350e+03, 2.68739240e+03,
9.14337000e+01, 3.47308500e+02, 1.29233440e+03, 2.12603800e+02,
4.07934420e+03, 4.99955400e+02, 8.57698000e+02, 1.44146180e+03,
3.87778200e+02, 1.74246600e+02, 1.89451200e+02, 1.40990580e+03,
4.85858200e+02, 7.55811150e+03, 2.27132029e+04, 4.71532820e+03,
1.18166940e+03, 5.41520900e+02, 3.35533900e+02, 6.10517100e+02,
4.61763760e+03, 8.97193500e+02, 1.22924510e+03, 4.10520000e+01,
5.80694320e+03, 2.14230200e+03, 7.10755900e+02, 1.32206630e+03,
2.99897200e+02, 3.44891700e+02, 1.34101700e+02, 3.66873200e+02,
1.38419634e+04, 1.65816960e+03, 1.26315090e+03, 5.38580460e+03,
6.89772200e+02, 2.10342830e+03, 7.14921400e+02, 3.67072500e+02,
9.89203200e+02, 1.84865840e+03, 3.14832600e+02, 1.84255842e+04,
1.34489300e+02, 2.00296800e+02, 1.00245610e+03, 1.32399090e+03,
1.14060800e+02, 3.32418500e+02, 9.53230100e+02, 3.82214400e+02,
6.88671900e+02, 3.52283900e+02, 2.23012900e+02, 8.88891300e+02,
8.85558300e+02, 1.09665090e+03, 4.01076750e+03, 3.10054960e+03,
8.08353000e+01, 2.23710600e+02, 2.89818400e+02, 4.58965000e+01,
3.09925100e+02, 1.25672090e+03, 3.24886500e+02, 7.03452100e+02,
9.87570100e+02, 3.52060500e+02, 1.19650710e+03, 4.90110120e+03,
3.64600900e+02, 3.23025830e+03, 7.71050020e+02, 5.74087900e+02,
2.18209510e+03, 1.47216580e+03, 1.43953470e+03, 9.26524000e+01,
1.06423590e+03, 3.05137700e+02, 3.19528600e+02, 3.70973100e+02,
8.74436200e+02, 1.24520650e+03, 3.50230200e+02, 1.57151900e+02,
1.27962890e+03, 3.36882200e+03, 8.25268900e+02, 6.65573980e+03,
9.07698400e+02, 4.98140250e+03, 1.36011700e+02, 5.15914700e+02,
4.92964000e+02, 1.13780400e+02, 2.21288160e+03, 4.92941000e+02,
2.76170370e+03, 1.08103190e+03, 2.33802400e+02, 1.65293990e+03,
6.32950800e+02, 1.53347720e+03, 3.65506880e+03, 6.96048300e+02,
1.87804600e+02, 2.02703000e+02, 3.97580920e+02, 4.96146200e+02,
1.00988070e+03, 5.30370500e+02, 1.55087460e+03, 1.14458900e+02,
3.71871900e+02, 2.32295200e+02, 1.41467400e+02, 6.87877900e+02,
3.69177910e+03, 2.17994000e+02, 8.40590200e+02, 2.00907600e+03,
1.28112670e+03, 2.92466200e+02, 1.76266900e+02, 5.74825000e+02,
4.53091780e+03, 6.30375120e+03, 1.07821610e+03, 4.28705700e+02,
2.64339400e+03, 3.59681410e+03, 6.85998200e+02, 3.00719100e+02,
7.79568400e+02, 1.62486240e+03, 3.27411300e+02, 6.91676000e+02,
2.98116260e+03, 3.45157800e+02, 1.46481230e+03, 2.19527000e+02,
3.37153110e+03, 1.08384500e+02, 3.65729200e+02, 1.12570500e+02,
6.25287000e+02, 3.02130800e+02, 1.53392852e+04, 3.46914822e+04,
6.75048500e+02, 2.17726210e+03, 7.70521200e+02, 4.65488300e+02,
2.69053800e+02, 7.56236590e+03, 1.15003640e+03, 8.58428800e+02,
9.52355200e+02, 1.57591750e+03, 1.01809300e+02, 1.48874600e+02])
```

In [42]:
```
mae_rf = mean_absolute_error(y_test, y_pred_rf)
rmse_rf = np.sqrt(mean_squared_error(y_test, y_pred_rf))
print("Random Forest MAE:", mae_rf)
print("Random Forest RMSE:", rmse_rf)
```

```
Random Forest MAE: 478.39716129032274
Random Forest RMSE: 5596.062709058427
```

Predict LTV for all customers

In [43]:
```
rfm['Predicted_LTV'] = xgb_model.predict(X)
rfm['Segment'] = pd.qcut(rfm['Predicted_LTV'], 4, labels=['Low', 'Medium', 'High',
rfm['Segment']
```

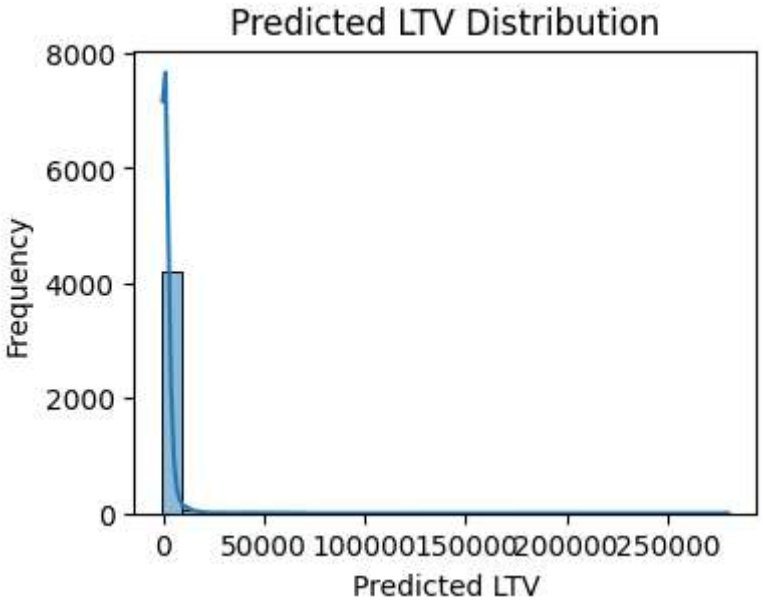
Out[43]:
```
0          VIP
1          VIP
2          VIP
3          VIP
4       Medium
         ...
4333       Low
4334       Low
4335       Low
4336       VIP
4337       VIP
Name: Segment, Length: 4338, dtype: category
Categories (4, object): ['Low' < 'Medium' < 'High' < 'VIP']
```

Save final output

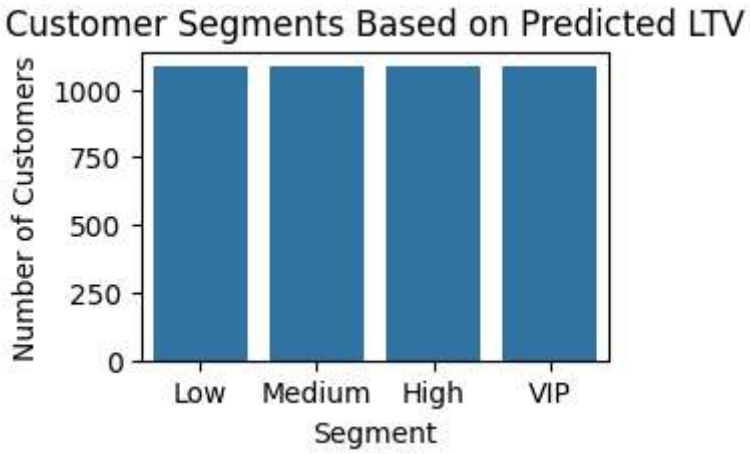
In [45]:
```
rfm[['CustomerID', 'Predicted_LTV', 'Segment']].to_csv("ltv_predictions.csv", index
```

Visualization

In [60]:
```
plt.figure(figsize=(4, 3))
sns.histplot(rfm['Predicted_LTV'], bins=30, kde=True)
plt.title("Predicted LTV Distribution")
plt.xlabel("Predicted LTV")
plt.ylabel("Frequency")
plt.show()
```



In [57]:
```
plt.figure(figsize=(3, 2))
sns.countplot(x='Segment', data=rfm, order=['Low', 'Medium', 'High', 'VIP'])
plt.title("Customer Segments Based on Predicted LTV")
plt.xlabel("Segment")
plt.ylabel("Number of Customers")
plt.show()
```



In [ ]: