In [14]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load data
df = pd.read_csv(r'C:\ProgramData\MySQL\MySQL Server 8.0\Uploads\OnlineRetail_Clean
```

In [8]:
```
print(df.head())              # View first 5 rows
```

```
  InvoiceNo StockCode                          Description  Quantity  \
0    536365    85123A   WHITE HANGING HEART T-LIGHT HOLDER         6
1    536365     71053                  WHITE METAL LANTERN         6
2    536365    84406B       CREAM CUPID HEARTS COAT HANGER         8
3    536365    84029G  KNITTED UNION FLAG HOT WATER BOTTLE         6
4    536365    84029E       RED WOOLLY HOTTIE WHITE HEART.         6

          InvoiceDate  UnitPrice  CustomerID         Country
0  2010-12-01 08:26:00       2.55     17850.0  United Kingdom
1  2010-12-01 08:26:00       3.39     17850.0  United Kingdom
2  2010-12-01 08:26:00       2.75     17850.0  United Kingdom
3  2010-12-01 08:26:00       3.39     17850.0  United Kingdom
4  2010-12-01 08:26:00       3.39     17850.0  United Kingdom
```

In [9]:
```
print(df.info())              # Check data types
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 392692 entries, 0 to 392691
Data columns (total 8 columns):
 #   Column       Non-Null Count   Dtype
---  ------       --------------   -----
 0   InvoiceNo    392692 non-null  int64
 1   StockCode    392692 non-null  object
 2   Description  392692 non-null  object
 3   Quantity     392692 non-null  int64
 4   InvoiceDate  392692 non-null  object
 5   UnitPrice    392692 non-null  float64
 6   CustomerID   392692 non-null  float64
 7   Country      392692 non-null  object
dtypes: float64(2), int64(2), object(4)
memory usage: 24.0+ MB
None
```

In [10]:
```
print(df.describe())
```

```
          InvoiceNo       Quantity      UnitPrice     CustomerID
count  392692.000000  392692.000000  392692.000000  392692.000000
mean   560590.875047      13.119702       3.125914   15287.843865
std     13087.063759     180.492832      22.241836    1713.539549
min    536365.000000       1.000000       0.001000   12346.000000
25%    549234.000000       2.000000       1.250000   13955.000000
50%    561874.000000       6.000000       1.950000   15150.000000
75%    572061.000000      12.000000       3.750000   16791.000000
max    581587.000000   80995.000000    8142.750000   18287.000000
```

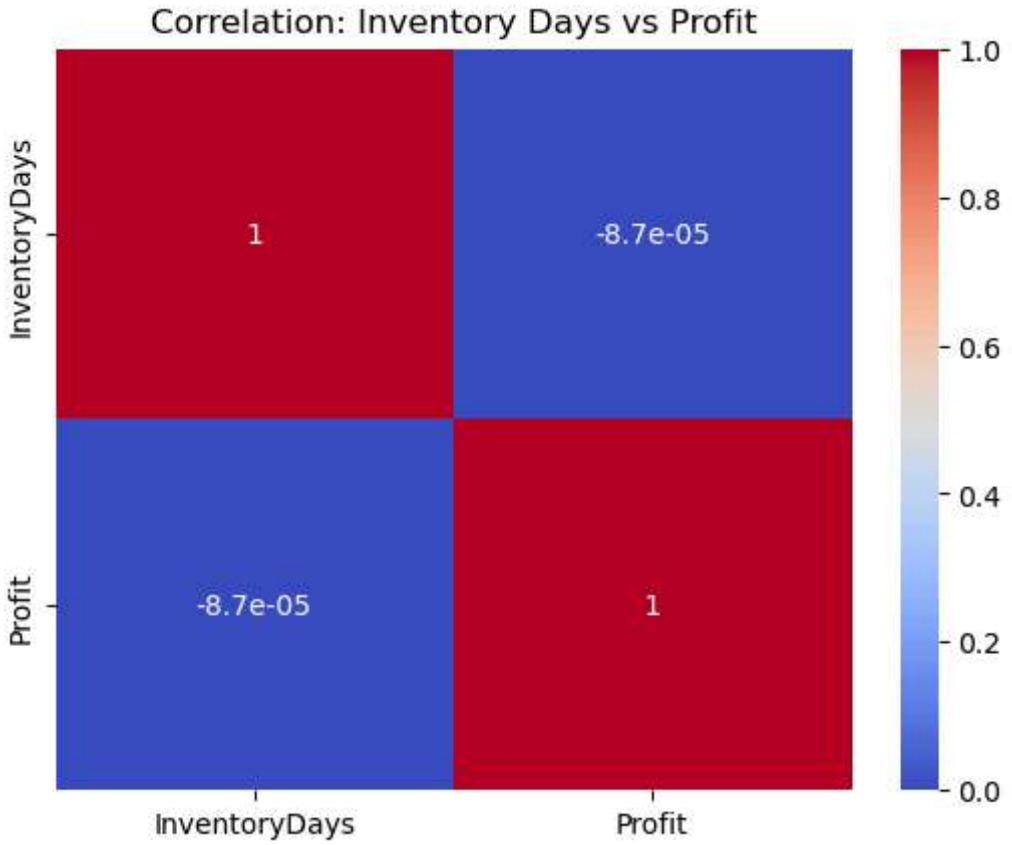
In [15]:
```
df['Revenue'] = df['Quantity'] * df['UnitPrice']
df['Profit'] = df['Revenue'] * 0.25   # 25% assumed margin
```

In [17]:
```
df['InventoryDays'] = np.random.randint(30, 120, size=len(df))
```

In [18]:
```
correlation = df[['InventoryDays', 'Profit']].corr()
print("Correlation Matrix:\n", correlation)
```

```
Correlation Matrix:
                InventoryDays    Profit
InventoryDays       1.000000 -0.000087
Profit             -0.000087  1.000000
```

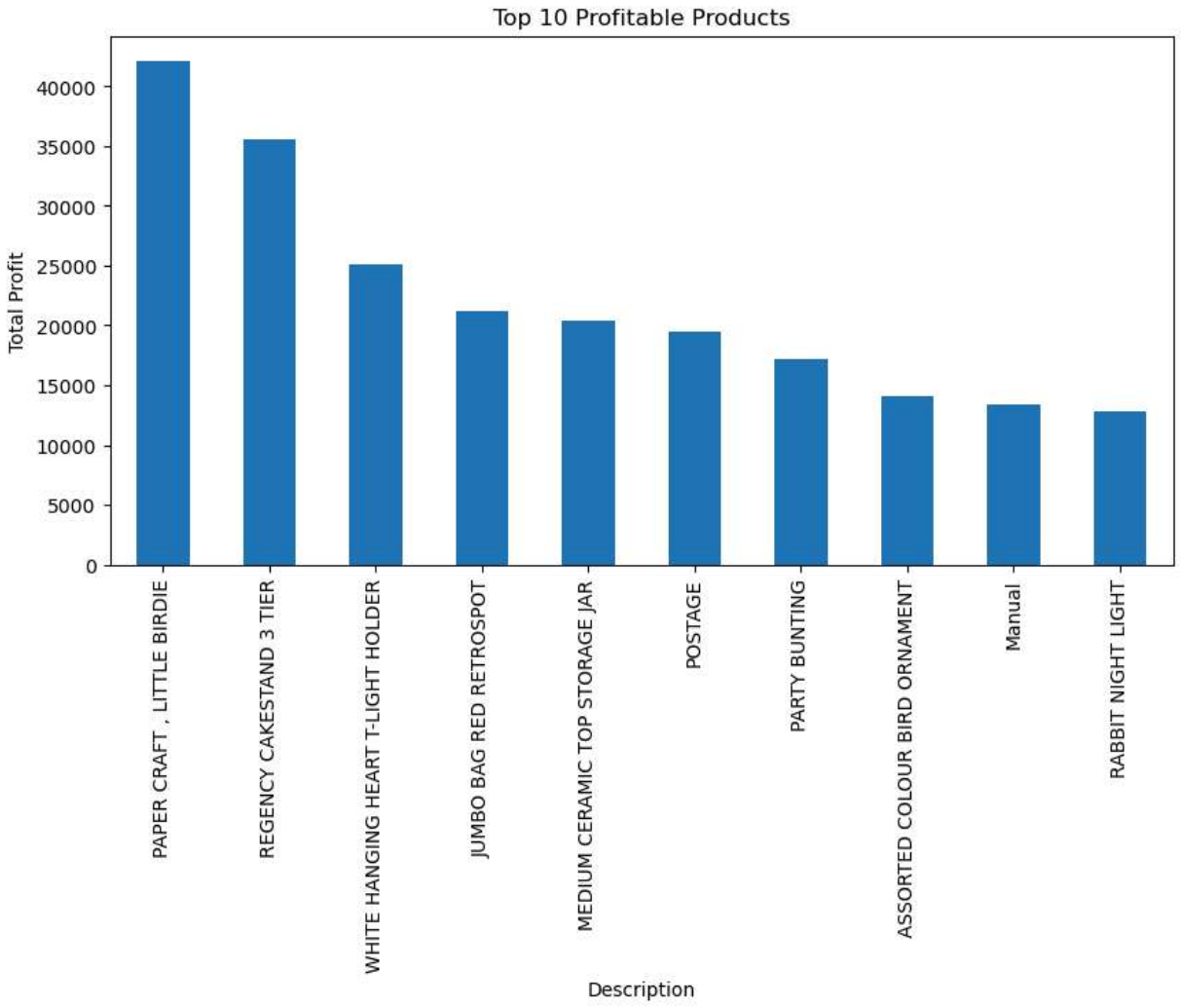
In [19]:
```
sns.heatmap(correlation, annot=True, cmap='coolwarm')
plt.title("Correlation: Inventory Days vs Profit")
plt.show()
```



In [21]:
```
profit_by_product = df.groupby('Description')['Profit'].sum().sort_values(ascending
print(profit_by_product)
```

```
Description
PAPER CRAFT , LITTLE BIRDIE           42117.4000
REGENCY CAKESTAND 3 TIER              35566.1875
WHITE HANGING HEART T-LIGHT HOLDER    25098.0250
JUMBO BAG RED RETROSPOT               21260.1350
MEDIUM CERAMIC TOP STORAGE JAR        20354.1825
POSTAGE                               19450.9900
PARTY BUNTING                         17196.3075
ASSORTED COLOUR BIRD ORNAMENT         14103.2575
Manual                                13354.9825
RABBIT NIGHT LIGHT                    12812.8100
Name: Profit, dtype: float64
```

In [22]:
```
# Bar Plot
profit_by_product.plot(kind='bar', figsize=(10,5), title='Top 10 Profitable Product
plt.ylabel('Total Profit')
plt.show()
```



In [ ]: