

```
In [2]: import pandas as pd

In [3]: data = pd.read_csv(r"D:\Unified_Internship\My_projectWork\Project\Amazon Sales_dataa.csv.csv")

In [3]: data.head()

Out[3]:
```

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order ID | Ship Date | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|---|-----------------------------------|-----------------------|-----------------|---------------|----------------|--------------|-----------|------------|------------|------------|-----------|---------------|------------|--------------|
| 0 | Australia and Oceania | Tuvalu | Baby Food | Offline | | H 2010-05-28 | 669165933 | 2010-06-27 | 9925 | 255.28 | 159.42 | 2533654.00 | 1582243.50 | 951410.50 |
| 1 | Central America and the Caribbean | Grenada | Cereal | Online | | C 2012-08-22 | 963881480 | 2012-08-15 | 2804 | 205.70 | 117.11 | 576782.80 | 328376.44 | 248406.36 |
| 2 | Europe | Russia | Office Supplies | Offline | | L 2014-05-02 | 341417157 | 2014-05-08 | 1779 | 651.21 | 524.96 | 1158502.59 | 933903.84 | 224598.75 |
| 3 | Sub-Saharan Africa | Sao Tome and Principe | Fruits | Online | | C 2014-06-20 | 514321792 | 2014-07-05 | 8102 | 9.33 | 6.92 | 75591.66 | 56065.84 | 19525.82 |
| 4 | Sub-Saharan Africa | Rwanda | Office Supplies | Offline | | L 2013-02-01 | 115456712 | 2013-02-06 | 5062 | 651.21 | 524.96 | 3296425.02 | 2657347.52 | 639077.50 |

```
In [4]: data.tail()

Out[4]:
```

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order ID | Ship Date | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|----|--------------------|--------------|---------------|---------------|----------------|--------------|-----------|------------|------------|------------|-----------|---------------|------------|--------------|
| 95 | Sub-Saharan Africa | Mali | Clothes | Online | | M 2011-07-28 | 812878119 | 2011-09-03 | 888 | 109.28 | 35.84 | 97040.64 | 31825.92 | 65214.72 |
| 96 | Asia | Malaysia | Fruits | Offline | | L 2011-11-11 | 810711038 | 2011-12-28 | 6267 | 9.33 | 6.92 | 58471.11 | 43367.64 | 15103.47 |
| 97 | Sub-Saharan Africa | Sierra Leone | Vegetables | Offline | | C 2016-06-01 | 728815257 | 2016-06-29 | 1485 | 154.06 | 90.93 | 228779.10 | 135031.05 | 93748.05 |
| 98 | North America | Mexico | Personal Care | Offline | | M 2015-07-30 | 559427106 | 2015-08-08 | 5767 | 81.73 | 56.67 | 471336.91 | 326815.89 | 144521.02 |
| 99 | Sub-Saharan Africa | Mozambique | Household | Offline | | L 2012-02-10 | 665095412 | 2012-02-15 | 5367 | 668.27 | 502.54 | 3586605.09 | 2697132.18 | 889472.91 |

```
In [5]: data.shape

Out[5]: (100, 14)

In [6]: data.columns

Out[6]: Index(['Region', 'Country', 'Item Type', 'Sales Channel', 'Order Priority', 'Order Date', 'Order ID', 'Ship Date', 'Units Sold', 'Unit Price', 'Unit Cost', 'Total Revenue', 'Total Cost', 'Total Profit'], dtype='object')

In [7]: data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 14 columns):
# Column Non-Null Count Dtype
---  --
0 Region 100 non-null object
1 Country 100 non-null object
2 Item Type 100 non-null object
3 Sales Channel 100 non-null object
4 Order Priority 100 non-null object
5 Order Date 100 non-null object
6 Order ID 100 non-null int64
7 Ship Date 100 non-null object
8 Units Sold 100 non-null int64
9 Unit Price 100 non-null float64
10 Unit Cost 100 non-null float64
11 Total Revenue 100 non-null float64
12 Total Cost 100 non-null float64
13 Total Profit 100 non-null float64
dtypes: float64(5), int64(2), object(7)
memory usage: 11.1+ KB

In [8]: data.isnull().sum()

Out[8]:
Region      0
Country     0
Item Type   0
Sales Channel 0
Order Priority 0
Order Date  0
Order ID    0
Ship Date   0
Units Sold  0
Unit Price  0
Unit Cost   0
Total Revenue 0
Total Cost  0
Total Profit 0
dtype: int64

In [9]: data.describe()

Out[9]:
```

| | Order ID | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|-------|--------------|-------------|------------|------------|---------------|--------------|--------------|
| count | 1.000000e+02 | 100.000000 | 100.000000 | 100.000000 | 1.000000e+02 | 1.000000e+02 | 1.000000e+02 |
| mean | 5.550204e+08 | 5128.710000 | 276.761300 | 191.048000 | 1.373488e+06 | 9.318057e+05 | 4.416820e+05 |
| std | 2.606153e+08 | 2794.484562 | 235.592241 | 188.206181 | 1.460029e+06 | 1.083938e+06 | 4.385379e+05 |
| min | 1.146066e+08 | 124.000000 | 9.330000 | 6.920000 | 4.870260e+03 | 3.612240e+03 | 1.258020e+03 |
| 25% | 3.389225e+08 | 2836.250000 | 81.730000 | 35.840000 | 2.687212e+05 | 1.688680e+05 | 1.214436e+05 |
| 50% | 5.577086e+08 | 5382.500000 | 179.880000 | 107.275000 | 7.523144e+05 | 3.635664e+05 | 2.907680e+05 |
| 75% | 7.907515e+08 | 7369.000000 | 437.200000 | 263.330000 | 2.212045e+06 | 1.613870e+06 | 6.358288e+05 |
| max | 9.940222e+08 | 9925.000000 | 668.270000 | 524.960000 | 5.897955e+06 | 4.509794e+06 | 1.719822e+06 |

```
In [10]: data['Item Type'].unique()

Out[10]: array(['Baby Food', 'Cereal', 'Office Supplies', 'Fruits', 'Household', 'Vegetables', 'Personal Care', 'Clothes', 'Cosmetics', 'Beverages', 'Meat', 'Snacks'], dtype=object)

In [11]: data['Order Year'] = pd.datetimeIndex(data['Order Date']).year
data['Order Quarter'] = pd.datetimeIndex(data['Order Date']).quarter
data['Order Month'] = pd.datetimeIndex(data['Order Date']).month

In [12]: data1 = data[['Region', 'Country', 'Item Type', 'Sales Channel', 'Order Priority', 'Order Date', 'Order Year', 'Order Quarter', 'Order Month', 'Order ID', 'Ship Date', 'Units Sold', 'Unit Price', 'Unit Cost', 'Total Revenue', 'Total Cost', 'Total Profit']]

In [13]: data1

Out[13]:
```

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order Year | Order Quarter | Order Month | Order ID | Ship Date | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|-----|-----------------------------------|-----------------------|-----------------|---------------|----------------|--------------|------------|---------------|-------------|--------------|------------|------------|------------|-----------|---------------|------------|--------------|
| 0 | Australia and Oceania | Tuvalu | Baby Food | Offline | | H 2010-05-28 | 2010 | | 2 | 5 669165933 | 2010-06-27 | 9925 | 255.28 | 159.42 | 2533654.00 | 1582243.50 | 951410.50 |
| 1 | Central America and the Caribbean | Grenada | Cereal | Online | | C 2012-08-22 | 2012 | | 3 | 8 963881480 | 2012-08-15 | 2804 | 205.70 | 117.11 | 576782.80 | 328376.44 | 248406.36 |
| 2 | Europe | Russia | Office Supplies | Offline | | L 2014-05-02 | 2014 | | 2 | 5 341417157 | 2014-05-08 | 1779 | 651.21 | 524.96 | 1158502.59 | 933903.84 | 224598.75 |
| 3 | Sub-Saharan Africa | Sao Tome and Principe | Fruits | Online | | C 2014-06-20 | 2014 | | 2 | 6 514321792 | 2014-07-05 | 8102 | 9.33 | 6.92 | 75591.66 | 56065.84 | 19525.82 |
| 4 | Sub-Saharan Africa | Rwanda | Office Supplies | Offline | | L 2013-02-01 | 2013 | | 1 | 2 115456712 | 2013-02-06 | 5062 | 651.21 | 524.96 | 3296425.02 | 2657347.52 | 639077.50 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 95 | Sub-Saharan Africa | Mali | Clothes | Online | | M 2011-07-28 | 2011 | | 3 | 7 512878119 | 2011-09-03 | 888 | 109.28 | 35.84 | 97040.64 | 31825.92 | 65214.72 |
| 96 | Asia | Malaysia | Fruits | Offline | | L 2011-11-11 | 2011 | | 4 | 11 810711038 | 2011-12-28 | 6267 | 9.33 | 6.92 | 58471.11 | 43367.64 | 15103.47 |
| 97 | Sub-Saharan Africa | Sierra Leone | Vegetables | Offline | | C 2016-06-01 | 2016 | | 2 | 6 728815257 | 2016-06-29 | 1485 | 154.06 | 90.93 | 228779.10 | 135031.05 | 93748.05 |
| 98 | North America | Mexico | Personal Care | Offline | | M 2015-07-30 | 2015 | | 3 | 7 559427106 | 2015-08-08 | 5767 | 81.73 | 56.67 | 471336.91 | 326815.89 | 144521.02 |
| 99 | Sub-Saharan Africa | Mozambique | Household | Offline | | L 2012-02-10 | 2012 | | 1 | 2 665095412 | 2012-02-15 | 5367 | 668.27 | 502.54 | 3586605.09 | 2697132.18 | 889472.91 |

100 rows × 17 columns

```
In [14]: sort_data1 = data1.sort_values(by = ['Units Sold'],ascending = False)

In [15]: items_sold = sort_data1[['Region','Item Type','Order Year','Order Quarter','Order Month','Units Sold']]

In [16]: items_sold.head(5)

Out[16]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold |
|----|------------------------------|-----------|------------|---------------|-------------|------------|
| 0 | Australia and Oceania | Baby Food | 2010 | 2 | 5 | 9925 |
| 56 | Australia and Oceania | Clothes | 2010 | 2 | 6 | 9905 |
| 74 | Middle East and North Africa | Cosmetics | 2013 | 3 | 7 | 9892 |
| 79 | Australia and Oceania | Cosmetics | 2013 | 3 | 7 | 9654 |
| 51 | Sub-Saharan Africa | Fruits | 2013 | 3 | 8 | 9606 |

```
In [17]: items_sold.tail(5)

Out[17]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold |
|----|------------------------------|---------------|------------|---------------|-------------|------------|
| 88 | Middle East and North Africa | Fruits | 2012 | 2 | 4 | 522 |
| 58 | Europe | Household | 2012 | 1 | 1 | 282 |
| 47 | Europe | Personal Care | 2010 | 4 | 12 | 273 |
| 43 | Europe | Vegetables | 2012 | 4 | 10 | 171 |
| 10 | Asia | Vegetables | 2011 | 2 | 6 | 124 |

```
In [18]: sort_data2 = data1.sort_values(by = ['Total Revenue'],ascending=False)

In [19]: items_revenue = sort_data2[['Region','Item Type','Order Year','Order Quarter','Order Month','Units Sold','Total Revenue']]

In [20]: items_revenue.head(5)

Out[20]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold | Total Revenue |
|----|-----------------------------------|-----------------|------------|---------------|-------------|------------|---------------|
| 13 | Central America and the Caribbean | Household | 2017 | 1 | 2 | 8974 | 5997054.98 |
| 33 | Asia | Household | 2015 | 1 | 1 | 8250 | 5513227.50 |
| 68 | Europe | Office Supplies | 2010 | 4 | 10 | 8287 | 5396577.27 |
| 75 | North America | Household | 2014 | 4 | 11 | 6954 | 4647149.58 |
| 38 | Asia | Office Supplies | 2012 | 2 | 4 | 6708 | 4368316.68 |

```
In [21]: items_revenue.tail(5)

Out[21]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold | Total Revenue |
|----|------------------------------|---------------|------------|---------------|-------------|------------|---------------|
| 47 | Europe | Personal Care | 2010 | 4 | 12 | 273 | 22312.29 |
| 23 | Australia and Oceania | Fruits | 2014 | 3 | 9 | 2187 | 20404.71 |
| 71 | Asia | Vegetables | 2011 | 2 | 6 | 124 | 19103.44 |
| 17 | Middle East and North Africa | Fruits | 2015 | 3 | 8 | 673 | 6279.09 |
| 88 | Middle East and North Africa | Fruits | 2012 | 2 | 4 | 522 | 4870.26 |

```
In [22]: sort_data3 = data1.sort_values(by = ['Total Profit'],ascending=False)

In [23]: items_profit = sort_data3[['Region','Item Type','Order Year','Order Quarter','Order Month','Units Sold','Total Revenue','Total Profit']]

In [24]: items_profit.head(5)

Out[24]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold | Total Revenue | Total Profit |
|----|-----------------------------------|-----------|------------|---------------|-------------|------------|---------------|--------------|
| 74 | Middle East and North Africa | Cosmetics | 2013 | 3 | 7 | 9892 | 4324782.40 | 1719922.04 |
| 79 | Australia and Oceania | Cosmetics | 2013 | 3 | 7 | 9654 | 4220728.80 | 1678540.98 |
| 46 | Europe | Cosmetics | 2016 | 4 | 12 | 8867 | 3876652.40 | 1541705.29 |
| 30 | Europe | Cosmetics | 2012 | 3 | 9 | 8661 | 3786589.20 | 1505888.07 |
| 13 | Central America and the Caribbean | Household | 2017 | 1 | 2 | 8974 | 5997054.98 | 1487261.02 |

```
In [25]: items_profit.tail(5)

Out[25]:
```

| | Region | Item Type | Order Year | Order Quarter | Order Month | Units Sold | Total Revenue | Total Profit |
|----|------------------------------|---------------|------------|---------------|-------------|------------|---------------|--------------|
| 10 | Asia | Vegetables | 2011 | 2 | 6 | 124 | 19103.44 | 7828.12 |
| 47 | Europe | Personal Care | 2010 | 4 | 12 | 273 | 22312.29 | 6841.38 |
| 23 | Australia and Oceania | Fruits | 2014 | 3 | 9 | 2187 | 20404.71 | 5270.67 |
| 71 | Middle East and North Africa | Fruits | 2015 | 3 | 8 | 673 | 6279.09 | 1621.93 |
| 88 | Middle East and North Africa | Fruits | 2012 | 2 | 4 | 522 | 4870.26 | 1258.02 |

```
In [26]: Year_wise = data1.sort_values(by = ['Order Month'])

In [27]: Year_2010 = Year_wise[Year_wise['Order Year'] == 2010]

In [28]: Year_2010

Out[28]:
```

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order Year | Order Quarter | Order Month | Order ID | Ship Date | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|----|------------------------------|--------------|-----------------|---------------|----------------|--------------|------------|---------------|-------------|--------------|------------|------------|------------|-----------|---------------|------------|--------------|
| 50 | Europe | Albania | Clothes | Online | | C 2010-02-05 | 2010 | | 1 | 2 385383069 | 2010-03-18 | 2269 | 109.28 | 35.84 | 247956.32 | 81320.96 | 166635.36 |
| 41 | Middle East and North Africa | Azerbaijan | Cosmetics | Online | | M 2010-02-06 | 2010 | | 1 | 2 382389299 | 2010-02-25 | 7234 | 437.20 | 263.33 | 3162704.80 | 1904929.22 | 1257775.58 |
| 0 | Australia and Oceania | Tuvalu | Baby Food | Offline | | H 2010-05-28 | 2010 | | 2 | 5 669165933 | 2010-06-27 | 9925 | 255.28 | 159.42 | 2533654.00 | 1582243.50 | 951410.50 |
| 27 | Sub-Saharan Africa | Mali | Fruits | Online | | L 2010-05-07 | 2010 | | 2 | 5 666084800 | 2010-05-10 | 5822 | 9.33 | 6.92 | 54319.26 | 40288.24 | 14031.02 |
| 56 | Australia and Oceania | Fiji | Clothes | Offline | | C 2010-10-24 | 2010 | | 2 | 6 648764849 | 2010-08-03 | 9905 | 109.28 | 35.84 | 1082418.40 | 354995.20 | 727423.20 |
| 68 | Europe | Lithuania | Office Supplies | Offline | | H 2010-10-24 | 2010 | | 4 | 10 166460740 | 2010-11-17 | 8287 | 651.21 | 524.96 | 5396577.27 | 4350343.52 | 1046233.75 |
| 63 | Middle East and North Africa | Libya | Clothes | Offline | | H 2010-10-30 | 2010 | | 4 | 10 705784308 | 2010-11-17 | 6116 | 109.28 | 35.84 | 668356.48 | 219197.44 | 449159.04 |
| 93 | Europe | Romania | Cosmetics | Online | | H 2010-11-26 | 2010 | | 4 | 11 660643374 | 2010-12-25 | 7910 | 437.20 | 263.33 | 3458252.00 | 2082940.30 | 1375311.70 |
| 47 | Europe | Switzerland | Personal Care | Online | | M 2010-08-30 | 2010 | | 4 | 12 617667900 | 2011-01-31 | 273 | 81.73 | 56.67 | 22312.29 | 15470.91 | 6841.38 |
| 18 | Asia | Turkmenistan | Household | Offline | | L 2010-12-30 | 2010 | | 4 | 12 441619336 | 2011-01-20 | 3830 | 668.27 | 502.54 | 2559471.04 | 1924728.20 | 634745.90 |

```
In [29]: Year_2011 = Year_wise[Year_wise['Order Year'] == 2011]

In [30]: Year_2011

Out[30]:
```

| | Region | Country | Item Type | Sales Channel | Order Priority | Order Date | Order Year | Order Quarter | Order Month | Order ID | Ship Date | Units Sold | Unit Price | Unit Cost | Total Revenue | Total Cost | Total Profit |
|----|-----------------------------------|----------------------------------|-----------------|---------------|----------------|--------------|------------|---------------|-------------|--------------|------------|------------|------------|-----------|---------------|------------|--------------|
| 86 | Sub-Saharan Africa | Sao Tome and Principe | Beverages | Offline | | C 2011-01-16 | 2011 | | 1 | 1 180283772 | 2011-01-21 | 8829 | 47.45 | 31.79 | 418936.05 | 280673.91 | 138262.14 |
| 83 | Sub-Saharan Africa | Zambia | Snacks | Offline | | L 2011-01-04 | 2011 | | 1 | 1 122583663 | 2011-01-05 | 4085 | 152.58 | 97.44 | 632389.30 | 398042.40 | 23246.90 |
| 94 | Central America and the Caribbean | Nicaragua | Beverages | Offline | | C 2011-03-21 | 2011 | | 1 | 2 963392674 | 2011-03-21 | 8156 | 47.45 | 31.79 | 387002.20 | 259279.24 | 127722.96 |
| 6 | Sub-Saharan Africa | Angola | Household | Offline | | M 2011-04-23 | 2011 | | 2 | 4 135425221 | 2011-04-27 | 4187 | 668.27 | 502.54 | 2788046.49 | 2104134.98 | 693911.51 |
| 72 | Sub-Saharan Africa | Democratic Republic of the Congo | Beverages | Online | | C 2011-05-25 | 2011 | | 2 | 5 585920464 | 2011-07-15 | 5741 | 47.45 | 31.79 | 272410.45 | 182506.39 | 89904.06 |
| 10 | Asia | Kyrgyzstan | Vegetables | Online | | H 2011-07-12 | 2011 | | 2 | 6 814711606 | 2011-07-12 | 124 | 154.06 | 90.93 | 19103.44 | 11275.32 | 7828.12 |
| 95 | Sub-Saharan Africa | Mali | Clothes | Online | | M 2011-07-26 | 2011 | | 3 | 7 512878119 | 2011-09-03 | 888 | 109.28 | 35.84 | 97040.64 | 31825.92 | 65214.72 |
| 77 | Asia | Laos | Vegetables | Offline | | C 2011-09-15 | 2011 | | 3 | 9 789176547 | 2011-10-23 | 3732 | 154.06 | 90.93 | 574951.92 | 339350.76 | 235601.16 |
| 96 | Asia | Malaysia | Fruits | Offline | | L 2011-11-11 | 2011 | | 4 | 11 810711038 | 2011-12-28 | 6267 | 9.33 | 6.92 | 58471.11 | 43367.64 | 15103.47 |
| 53 | Sub-Saharan Africa | Sierra Leone | Office Supplies | Offline | | M 2011-11-11 | 2011 | | 4 | 11 443888415 | 2012-0 | | | | | | |