

```

{
  "cells": [
    {
      "cell_type": "code",
      "execution_count": 34,
      "id": "e0459162-75ae-44b6-8591-cf88a86da7b1",
      "metadata": {},
      "outputs": [],
      "source": [
        "import pandas as pd\n",
        "import numpy as np\n",
        "import matplotlib.pyplot as plt\n",
        "import seaborn as sns"
      ]
    },
    {
      "cell_type": "code",
      "execution_count": 27,
      "id": "e4e38bb2-997f-4b65-a51f-10490f8c05b9",
      "metadata": {},
      "outputs": [],
      "source": [
        "df = pd.read_csv(r\"C:\\Users\\tamal\\Documents\\Amazon_Project\\Amazon_Sales_data.csv\")"
      ]
    },
    {
      "cell_type": "code",
      "execution_count": 35,
      "id": "8fb7df40-a094-47dc-a3f4-4699c4f867b7",
      "metadata": {},
      "outputs": [
        {
          "data": {
            "text/html": [
              "<div>\n",
              "<style scoped>\n",
              ".dataframe tbody tr th:only-of-type {\n",
              "  vertical-align: middle;\n",
              }\n",
              "\n",
              ".dataframe tbody tr th {\n",
              "  vertical-align: top;\n",
              }\n",
              "\n",
              .dataframe thead th {\n",
              "  text-align: right;\n",
              }\n",
              "</style>\n",
              <table border=\\\"1\\\" class=\\\"dataframe\\\">\n",
              <thead>\n",
              <tr style=\\\"text-align: right;\\\">\n",
              <th></th>\n",
              <th>Region</th>\n",
              <th>Country</th>\n",
              <th>Item Type</th>\n",
              <th>Sales Channel</th>\n",
              <th>Order Priority</th>\n",
              <th>Order Date</th>\n",
              <th>Order ID</th>\n",
              <th>Ship Date</th>\n",
              <th>Units Sold</th>\n",
              <th>Unit Price</th>\n",
              <th>Unit Cost</th>\n",
              <th>Total Revenue</th>\n",
              <th>Total Cost</th>\n",
              <th>Total Profit</th>\n",
              </tr>\n",
              </thead>\n",
              <tbody>\n",
              <tr>\n",
              <th>0</th>\n",
              <td>Australia and Oceania</td>\n",
              <td>Tuvalu</td>\n",
              <td>Baby Food</td>\n",
              <td>Offline</td>\n",
              <td>H</td>\n",
              <td>5/28/2010</td>\n",
              <td>669165933</td>\n",
              <td>6/27/2010</td>\n",
              <td>9925</td>\n",
              <td>255.28</td>\n",
              <td>159.42</td>\n",
              <td>2533654.00</td>\n",
              <td>1582243.50</td>\n",
              <td>951410.50</td>\n",
              </tr>\n",
              <tr>\n",
              <th>1</th>\n",
              <td>Central America and the Caribbean</td>\n",
              <td>Grenada</td>\n",
              <td>Cereal</td>\n",
              <td>Online</td>\n",
              <td>C</td>\n",
              <td>8/22/2012</td>\n",
              <td>963881480</td>\n",
              <td>9/15/2012</td>\n",
              <td>2804</td>\n",
              <td>205.70</td>\n",
              <td>117.11</td>\n",
              <td>576782.80</td>\n",
              <td>328376.44</td>\n",
              <td>248406.36</td>\n",
              </tr>\n",
              <tr>\n",
              <th>2</th>\n",
              <td>Europe</td>\n",

```

```
"      <td>Russia</td>\n",
"      <td>Office Supplies</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>\n",
"      <td>5/2/2014</td>\n",
"      <td>341417157</td>\n",
"      <td>5/8/2014</td>\n",
"      <td>1779</td>\n",
"      <td>651.21</td>\n",
"      <td>524.96</td>\n",
"      <td>1158502.59</td>\n",
"      <td>933903.84</td>\n",
"      <td>224598.75</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>3</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Sao Tome and Principe</td>\n",
"      <td>Fruits</td>\n",
"      <td>Online</td>\n",
"      <td>C</td>\n",
"      <td>6/20/2014</td>\n",
"      <td>514321792</td>\n",
"      <td>7/5/2014</td>\n",
"      <td>8102</td>\n",
"      <td>9.33</td>\n",
"      <td>6.92</td>\n",
"      <td>75591.66</td>\n",
"      <td>56065.84</td>\n",
"      <td>19525.82</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>4</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Rwanda</td>\n",
"      <td>Office Supplies</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>\n",
"      <td>2/1/2013</td>\n",
"      <td>115456712</td>\n",
"      <td>2/6/2013</td>\n",
"      <td>5062</td>\n",
"      <td>651.21</td>\n",
"      <td>524.96</td>\n",
"      <td>3296425.02</td>\n",
"      <td>2657347.52</td>\n",
"      <td>639077.50</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>5</th>\n",
"      <td>Australia and Oceania</td>\n",
"      <td>Solomon Islands</td>\n",
"      <td>Baby Food</td>\n",
"      <td>Online</td>\n",
"      <td>C</td>\n",
"      <td>2/4/2015</td>\n",
"      <td>547995746</td>\n",
"      <td>2/21/2015</td>\n",
"      <td>2974</td>\n",
"      <td>255.28</td>\n",
"      <td>159.42</td>\n",
"      <td>759202.72</td>\n",
"      <td>474115.08</td>\n",
"      <td>285087.64</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>6</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Angola</td>\n",
"      <td>Household</td>\n",
"      <td>Offline</td>\n",
"      <td>M</td>\n",
"      <td>4/23/2011</td>\n",
"      <td>135425221</td>\n",
"      <td>4/27/2011</td>\n",
"      <td>4187</td>\n",
"      <td>668.27</td>\n",
"      <td>502.54</td>\n",
"      <td>2798046.49</td>\n",
"      <td>2104134.98</td>\n",
"      <td>693911.51</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>7</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Burkina Faso</td>\n",
"      <td>Vegetables</td>\n",
"      <td>Online</td>\n",
"      <td>H</td>\n",
"      <td>7/17/2012</td>\n",
"      <td>871543967</td>\n",
"      <td>7/27/2012</td>\n",
"      <td>8082</td>\n",
"      <td>154.06</td>\n",
"      <td>90.93</td>\n",
"      <td>1245112.92</td>\n",
"      <td>734896.26</td>\n",
"      <td>510216.66</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>8</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Republic of the Congo</td>\n",
"      <td>Personal Care</td>\n",
"      <td>Offline</td>\n",
"      <td>M</td>
```

```

"      <td>7/14/2015</td>\n",
"      <td>770463311</td>\n",
"      <td>8/25/2015</td>\n",
"      <td>6070</td>\n",
"      <td>81.73</td>\n",
"      <td>56.67</td>\n",
"      <td>496101.10</td>\n",
"      <td>343986.90</td>\n",
"      <td>152114.20</td>\n",
"    </tr>\n",
"  <tr>\n",
"    <th>9</th>\n",
"    <td>Sub-Saharan Africa</td>\n",
"    <td>Senegal</td>\n",
"    <td>Cereal</td>\n",
"    <td>Online</td>\n",
"    <td>H</td>\n",
"    <td>4/18/2014</td>\n",
"    <td>616607081</td>\n",
"    <td>5/30/2014</td>\n",
"    <td>6593</td>\n",
"    <td>205.70</td>\n",
"    <td>117.11</td>\n",
"    <td>1356180.10</td>\n",
"    <td>772106.23</td>\n",
"    <td>584073.87</td>\n",
"  </tr>\n",
" </tbody>\n",
"</table>\n",
"</div>"
],
"text/plain": [
"      Region          Country      Item Type  \\\n",
"0      Australia and Oceania      Tuvalu      Baby Food  \n",
"1      Central America and the Caribbean      Grenada      Cereal  \n",
"2      Europe      Russia      Office Supplies  \n",
"3      Sub-Saharan Africa      Sao Tome and Principe      Fruits  \n",
"4      Sub-Saharan Africa      Rwanda      Office Supplies  \n",
"5      Australia and Oceania      Solomon Islands      Baby Food  \n",
"6      Sub-Saharan Africa      Angola      Household  \n",
"7      Sub-Saharan Africa      Burkina Faso      Vegetables  \n",
"8      Sub-Saharan Africa      Republic of the Congo      Personal Care  \n",
"9      Sub-Saharan Africa      Senegal      Cereal  \n",
"\n",
"Sales Channel Order Priority Order Date      Order ID      Ship Date      Units Sold  \\\n",
"0      Offline      H      5/28/2010      669165933      6/27/2010      9925  \n",
"1      Online      C      8/22/2012      963881480      9/15/2012      2804  \n",
"2      Offline      L      5/2/2014      341417157      5/8/2014      1779  \n",
"3      Online      C      6/20/2014      514321792      7/5/2014      8102  \n",
"4      Offline      L      2/1/2013      115456712      2/6/2013      5062  \n",
"5      Online      C      2/4/2015      547995746      2/21/2015      2974  \n",
"6      Offline      M      4/23/2011      135425221      4/27/2011      4187  \n",
"7      Online      H      7/17/2012      871543967      7/27/2012      8082  \n",
"8      Offline      M      7/14/2015      770463311      8/25/2015      6070  \n",
"9      Online      H      4/18/2014      616607081      5/30/2014      6593  \n",
"\n",
"Unit Price      Unit Cost      Total Revenue      Total Cost      Total Profit  \n",
"0      255.28      159.42      2533654.00      1582243.50      951410.50  \n",
"1      205.70      117.11      576782.80      328376.44      248406.36  \n",
"2      651.21      524.96      1158502.59      933903.84      224598.75  \n",
"3      9.33      6.92      75591.66      56065.84      19525.82  \n",
"4      651.21      524.96      3296425.02      2657347.52      639077.50  \n",
"5      255.28      159.42      759202.72      474115.08      285087.64  \n",
"6      668.27      502.54      2798046.49      2104134.98      693911.51  \n",
"7      154.06      90.93      1245112.92      734896.26      510216.66  \n",
"8      81.73      56.67      496101.10      343986.90      152114.20  \n",
"9      205.70      117.11      1356180.10      772106.23      584073.87  "
]
},
"execution_count": 35,
"metadata": {},
"output_type": "execute_result"
}
],
"source": [
"df.head(10)"
]
},
{
"cell_type": "code",
"execution_count": 45,
"id": "b8b24caa-a168-46b0-a98b-eeb747aa0bfd",
"metadata": {},
"outputs": [
{
"data": {
"text/html": [
<div>\n",
<style scoped>\n",
.dataframe tbody tr th:only-of-type {\n",
vertical-align: middle;\n",
}\n",
\n",
.dataframe tbody tr th {\n",
vertical-align: top;\n",
}\n",
\n",
.dataframe thead th {\n",
text-align: right;\n",
}\n",
</style>\n",
<table border="1" class="dataframe">\n",
<thead>\n",
<tr style="text-align: right;">\n",
<th></th>\n",
<th>Region</th>\n",

```

```
"      <th>Country</th>\n",
"      <th>Item Type</th>\n",
"      <th>Sales Channel</th>\n",
"      <th>Order Priority</th>\n",
"      <th>Order Date</th>\n",
"      <th>Order ID</th>\n",
"      <th>Ship Date</th>\n",
"      <th>Units Sold</th>\n",
"      <th>Unit Price</th>\n",
"      <th>Unit Cost</th>\n",
"      <th>Total Revenue</th>\n",
"      <th>Total Cost</th>\n",
"      <th>Total Profit</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>95</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Mali</td>\n",
"      <td>Clothes</td>\n",
"      <td>Online</td>\n",
"      <td>M</td>\n",
"      <td>7/26/2011</td>\n",
"      <td>512878119</td>\n",
"      <td>9/3/2011</td>\n",
"      <td>888</td>\n",
"      <td>109.28</td>\n",
"      <td>35.84</td>\n",
"      <td>97040.64</td>\n",
"      <td>31825.92</td>\n",
"      <td>65214.72</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>96</th>\n",
"      <td>Asia</td>\n",
"      <td>Malaysia</td>\n",
"      <td>Fruits</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>\n",
"      <td>11/11/2011</td>\n",
"      <td>810711038</td>\n",
"      <td>12/28/2011</td>\n",
"      <td>6267</td>\n",
"      <td>9.33</td>\n",
"      <td>6.92</td>\n",
"      <td>58471.11</td>\n",
"      <td>43367.64</td>\n",
"      <td>15103.47</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>97</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Sierra Leone</td>\n",
"      <td>Vegetables</td>\n",
"      <td>Offline</td>\n",
"      <td>C</td>\n",
"      <td>6/1/2016</td>\n",
"      <td>728815257</td>\n",
"      <td>6/29/2016</td>\n",
"      <td>1485</td>\n",
"      <td>154.06</td>\n",
"      <td>90.93</td>\n",
"      <td>228779.10</td>\n",
"      <td>135031.05</td>\n",
"      <td>93748.05</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>98</th>\n",
"      <td>North America</td>\n",
"      <td>Mexico</td>\n",
"      <td>Personal Care</td>\n",
"      <td>Offline</td>\n",
"      <td>M</td>\n",
"      <td>7/30/2015</td>\n",
"      <td>559427106</td>\n",
"      <td>8/8/2015</td>\n",
"      <td>5767</td>\n",
"      <td>81.73</td>\n",
"      <td>56.67</td>\n",
"      <td>471336.91</td>\n",
"      <td>326815.89</td>\n",
"      <td>144521.02</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>99</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Mozambique</td>\n",
"      <td>Household</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>\n",
"      <td>2/10/2012</td>\n",
"      <td>665095412</td>\n",
"      <td>2/15/2012</td>\n",
"      <td>5367</td>\n",
"      <td>668.27</td>\n",
"      <td>502.54</td>\n",
"      <td>3586605.09</td>\n",
"      <td>2697132.18</td>\n",
"      <td>889472.91</td>\n",
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>"
```

```
],
"text/plain": [
```

```

"
      Region      Country      Item Type      Sales Channel      \\n",
"95 Sub-Saharan Africa      Mali      Clothes      Online      \\n",
"96      Asia      Malaysia      Fruits      Offline      \\n",
"97 Sub-Saharan Africa      Sierra Leone      Vegetables      Offline      \\n",
"98 North America      Mexico      Personal Care      Offline      \\n",
"99 Sub-Saharan Africa      Mozambique      Household      Offline      \\n",
"\\n",
"      Order Priority      Order Date      Order ID      Ship Date      Units Sold      Unit Price      \\n",
"95      M      7/26/2011      512878119      9/3/2011      888      109.28      \\n",
"96      L      11/11/2011      810711038      12/28/2011      6267      9.33      \\n",
"97      C      6/1/2016      728815257      6/29/2016      1485      154.06      \\n",
"98      M      7/30/2015      559427106      8/8/2015      5767      81.73      \\n",
"99      L      2/10/2012      665095412      2/15/2012      5367      668.27      \\n",
"\\n",
"      Unit Cost      Total Revenue      Total Cost      Total Profit      \\n",
"95      35.84      97040.64      31825.92      65214.72      \\n",
"96      6.92      58471.11      43367.64      15103.47      \\n",
"97      90.93      228779.10      135031.05      93748.05      \\n",
"98      56.67      471336.91      326815.89      144521.02      \\n",
"99      502.54      3586605.09      2697132.18      889472.91      "
]
},
"execution_count": 45,
"metadata": {},
"output_type": "execute_result"
}
],
"source": [
  "df.tail()"
]
},
{
  "cell_type": "code",
  "execution_count": 38,
  "id": "06f408a5-f6bf-463d-8c95-2d6f53dbe8af",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "text/plain": [
          "(100, 14)"
        ]
      },
      "execution_count": 38,
      "metadata": {},
      "output_type": "execute_result"
    }
  ],
  "source": [
    "df.shape"
  ]
},
{
  "cell_type": "code",
  "execution_count": 39,
  "id": "c4322360-a965-453f-b5fb-60995d32f337",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "text/plain": [
          "Index(['Region', 'Country', 'Item Type', 'Sales Channel', 'Order Priority',\\n",
          "      'Order Date', 'Order ID', 'Ship Date', 'Units Sold', 'Unit Price',\\n",
          "      'Unit Cost', 'Total Revenue', 'Total Cost', 'Total Profit'],\\n",
          "      dtype='object')"
        ]
      },
      "execution_count": 39,
      "metadata": {},
      "output_type": "execute_result"
    }
  ],
  "source": [
    "df.columns"
  ]
},
{
  "cell_type": "code",
  "execution_count": 41,
  "id": "fb06d93f-a709-4a3a-ae3f-1804156cb368",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "text/plain": [
          "Region      object\\n",
          "Country      object\\n",
          "Item Type      object\\n",
          "Sales Channel      object\\n",
          "Order Priority      object\\n",
          "Order Date      object\\n",
          "Order ID      int64\\n",
          "Ship Date      object\\n",
          "Units Sold      int64\\n",
          "Unit Price      float64\\n",
          "Unit Cost      float64\\n",
          "Total Revenue      float64\\n",
          "Total Cost      float64\\n",
          "Total Profit      float64\\n",
          "dtype: object"
        ]
      },
      "execution_count": 41,
      "metadata": {},
      "output_type": "execute_result"
    }
  ],

```

```

],
"source": [
  "df.dtypes"
]
},
{
  "cell_type": "code",
  "execution_count": 43,
  "id": "61ba43cc-007f-4e44-aff3-5b9069acc8d7",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "text/plain": [
          "Region          0\\n",
          "Country         0\\n",
          "Item Type       0\\n",
          "Sales Channel   0\\n",
          "Order Priority   0\\n",
          "Order Date      0\\n",
          "Order ID        0\\n",
          "Ship Date       0\\n",
          "Units Sold      0\\n",
          "Unit Price      0\\n",
          "Unit Cost       0\\n",
          "Total Revenue   0\\n",
          "Total Cost      0\\n",
          "Total Profit    0\\n",
          "dtype: int64"
        ]
      },
      "execution_count": 43,
      "metadata": {},
      "output_type": "execute_result"
    }
  ],
  "source": [
    "df.isnull().sum()"
  ]
},
{
  "cell_type": "code",
  "execution_count": 44,
  "id": "26ce581d-0fd2-426d-af71-2a8aef260bf5",
  "metadata": {},
  "outputs": [
    {
      "name": "stdout",
      "output_type": "stream",
      "text": [
        "<class 'pandas.core.frame.DataFrame'>\\n",
        "RangeIndex: 100 entries, 0 to 99\\n",
        "Data columns (total 14 columns):\\n",
        " #   Column          Non-Null Count  Dtype   \\n",
        "---  ---          -

```

#	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)\\n",
        "memory usage: 11.1+ KB\\n"
      ],
      "execution_count": 44,
      "metadata": {},
      "output_type": "stream"
    }
  ],
  "source": [
    "df.info()"
  ]
},
{
  "cell_type": "code",
  "execution_count": 82,
  "id": "85678e79-74b8-4458-82c4-6c309ba4a93d",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "text/html": [
          "<div>\\n",
          "<style scoped>\\n",
          "  .dataframe tbody tr th:only-of-type {\\n",
          "    vertical-align: middle;\\n",
          "  }\\n",
          "\\n",
          "  .dataframe tbody tr th {\\n",
          "    vertical-align: top;\\n",
          "  }\\n",
          "\\n",
          "  .dataframe thead th {\\n",
          "    text-align: right;\\n",
          "  }\\n",
          "</style>\\n",
          "<table border='1' class='dataframe'>\\n",
          "  <thead>\\n",
          "    <tr style='text-align: right;'>\\n",
          "      <th></th>\\n",
          "      <th>Region</th>\\n",

```

```
"      <th>Country</th>\n",
"      <th>Item Type</th>\n",
"      <th>Sales Channel</th>\n",
"      <th>Order Priority</th>\n",
"      <th>Order Date</th>\n",
"      <th>Order ID</th>\n",
"      <th>Ship Date</th>\n",
"      <th>Units Sold</th>\n",
"      <th>Unit Price</th>\n",
"      <th>Unit Cost</th>\n",
"      <th>Total Revenue</th>\n",
"      <th>Total Cost</th>\n",
"      <th>Total Profit</th>\n",
"      <th>Date</th>\n",
"      <th>Month</th>\n",
"      <th>Year</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>0</th>\n",
"      <td>Australia and Oceania</td>\n",
"      <td>Tuvalu</td>\n",
"      <td>Baby Food</td>\n",
"      <td>Offline</td>\n",
"      <td>R</td>\n",
"      <td>2010-05-28</td>\n",
"      <td>669165933</td>\n",
"      <td>6/27/2010</td>\n",
"      <td>9925</td>\n",
"      <td>255.28</td>\n",
"      <td>159.42</td>\n",
"      <td>2533654.00</td>\n",
"      <td>1582243.50</td>\n",
"      <td>951410.50</td>\n",
"      <td>2010-05-28</td>\n",
"      <td>5</td>\n",
"      <td>2010</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>1</th>\n",
"      <td>Central America and the Caribbean</td>\n",
"      <td>Grenada</td>\n",
"      <td>Cereal</td>\n",
"      <td>Online</td>\n",
"      <td>C</td>\n",
"      <td>2012-08-22</td>\n",
"      <td>963881480</td>\n",
"      <td>9/15/2012</td>\n",
"      <td>2804</td>\n",
"      <td>205.70</td>\n",
"      <td>117.11</td>\n",
"      <td>576782.80</td>\n",
"      <td>328376.44</td>\n",
"      <td>248406.36</td>\n",
"      <td>2012-08-22</td>\n",
"      <td>8</td>\n",
"      <td>2012</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>2</th>\n",
"      <td>Europe</td>\n",
"      <td>Russia</td>\n",
"      <td>Office Supplies</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>\n",
"      <td>2014-05-02</td>\n",
"      <td>341417157</td>\n",
"      <td>5/8/2014</td>\n",
"      <td>1779</td>\n",
"      <td>651.21</td>\n",
"      <td>524.96</td>\n",
"      <td>1158502.59</td>\n",
"      <td>933903.84</td>\n",
"      <td>224598.75</td>\n",
"      <td>2014-05-02</td>\n",
"      <td>5</td>\n",
"      <td>2014</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>3</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Sao Tome and Principe</td>\n",
"      <td>Fruits</td>\n",
"      <td>Online</td>\n",
"      <td>C</td>\n",
"      <td>2014-06-20</td>\n",
"      <td>514321792</td>\n",
"      <td>7/5/2014</td>\n",
"      <td>8102</td>\n",
"      <td>9.33</td>\n",
"      <td>6.92</td>\n",
"      <td>75591.66</td>\n",
"      <td>56065.84</td>\n",
"      <td>19525.82</td>\n",
"      <td>2014-06-20</td>\n",
"      <td>6</td>\n",
"      <td>2014</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>4</th>\n",
"      <td>Sub-Saharan Africa</td>\n",
"      <td>Rwanda</td>\n",
"      <td>Office Supplies</td>\n",
"      <td>Offline</td>\n",
"      <td>L</td>
```

```
"      <td>2013-02-01</td>\n",
"      <td>115456712</td>\n",
"      <td>2/6/2013</td>\n",
"      <td>5062</td>\n",
"      <td>651.21</td>\n",
"      <td>524.96</td>\n",
"      <td>3296425.02</td>\n",
"      <td>2657347.52</td>\n",
"      <td>639077.50</td>\n",
"      <td>2013-02-01</td>\n",
"      <td>2</td>\n",
"      <td>2013</td>\n",
"    </tr>\n",
"  <tr>\n",
"    <th>5</th>\n",
"    <td>Australia and Oceania</td>\n",
"    <td>Solomon Islands</td>\n",
"    <td>Baby Food</td>\n",
"    <td>Online</td>\n",
"    <td>C</td>\n",
"    <td>2015-02-04</td>\n",
"    <td>547995746</td>\n",
"    <td>2/21/2015</td>\n",
"    <td>2974</td>\n",
"    <td>255.28</td>\n",
"    <td>159.42</td>\n",
"    <td>759202.72</td>\n",
"    <td>474115.08</td>\n",
"    <td>285087.64</td>\n",
"    <td>2015-02-04</td>\n",
"    <td>2</td>\n",
"    <td>2015</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>6</th>\n",
"    <td>Sub-Saharan Africa</td>\n",
"    <td>Angola</td>\n",
"    <td>Household</td>\n",
"    <td>Offline</td>\n",
"    <td>M</td>\n",
"    <td>2011-04-23</td>\n",
"    <td>135425221</td>\n",
"    <td>4/27/2011</td>\n",
"    <td>4187</td>\n",
"    <td>668.27</td>\n",
"    <td>502.54</td>\n",
"    <td>2798046.49</td>\n",
"    <td>2104134.98</td>\n",
"    <td>693911.51</td>\n",
"    <td>2011-04-23</td>\n",
"    <td>4</td>\n",
"    <td>2011</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>7</th>\n",
"    <td>Sub-Saharan Africa</td>\n",
"    <td>Burkina Faso</td>\n",
"    <td>Vegetables</td>\n",
"    <td>Online</td>\n",
"    <td>H</td>\n",
"    <td>2012-07-17</td>\n",
"    <td>871543967</td>\n",
"    <td>7/27/2012</td>\n",
"    <td>8082</td>\n",
"    <td>154.06</td>\n",
"    <td>90.93</td>\n",
"    <td>1245112.92</td>\n",
"    <td>734896.26</td>\n",
"    <td>510216.66</td>\n",
"    <td>2012-07-17</td>\n",
"    <td>7</td>\n",
"    <td>2012</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>8</th>\n",
"    <td>Sub-Saharan Africa</td>\n",
"    <td>Republic of the Congo</td>\n",
"    <td>Personal Care</td>\n",
"    <td>Offline</td>\n",
"    <td>M</td>\n",
"    <td>2015-07-14</td>\n",
"    <td>770463311</td>\n",
"    <td>8/25/2015</td>\n",
"    <td>6070</td>\n",
"    <td>81.73</td>\n",
"    <td>56.67</td>\n",
"    <td>496101.10</td>\n",
"    <td>343986.90</td>\n",
"    <td>152114.20</td>\n",
"    <td>2015-07-14</td>\n",
"    <td>7</td>\n",
"    <td>2015</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>9</th>\n",
"    <td>Sub-Saharan Africa</td>\n",
"    <td>Senegal</td>\n",
"    <td>Cereal</td>\n",
"    <td>Online</td>\n",
"    <td>H</td>\n",
"    <td>2014-04-18</td>\n",
"    <td>616607081</td>\n",
"    <td>5/30/2014</td>\n",
"    <td>6593</td>\n",
"    <td>205.70</td>\n",
"    <td>117.11</td>
```



```

"      <td>1356180.10</td>\n",
"      <td>772106.23</td>\n",
"      <td>584073.87</td>\n",
"      <td>2014-04-18</td>\n",
"      <td>4</td>\n",
"      <td>2014</td>\n",
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"</div>"
],
"text/plain": [
"      Region          Country      Item Type  \\\n",
"0      Australia and Oceania      Tuvalu      Baby Food  \n",
"1      Central America and the Caribbean      Grenada      Cereal  \n",
"2      Europe      Russia      Office Supplies  \n",
"3      Sub-Saharan Africa      Sao Tome and Principe      Fruits  \n",
"4      Sub-Saharan Africa      Rwanda      Office Supplies  \n",
"5      Australia and Oceania      Solomon Islands      Baby Food  \n",
"6      Sub-Saharan Africa      Angola      Household  \n",
"7      Sub-Saharan Africa      Burkina Faso      Vegetables  \n",
"8      Sub-Saharan Africa      Republic of the Congo      Personal Care  \n",
"9      Sub-Saharan Africa      Senegal      Cereal  \n",
"\n",
"  Sales Channel Order Priority Order Date      Order ID      Ship Date      Units Sold  \\\n",
"0      Offline      H      2010-05-28      669165933      6/27/2010      9925  \n",
"1      Online      C      2012-08-22      963881480      9/15/2012      2804  \n",
"2      Offline      L      2014-05-02      341417157      5/8/2014      1779  \n",
"3      Online      C      2014-06-20      514321792      7/5/2014      8102  \n",
"4      Offline      L      2013-02-01      115456712      2/6/2013      5062  \n",
"5      Online      C      2015-02-04      547995746      2/21/2015      2974  \n",
"6      Offline      M      2011-04-23      135425221      4/27/2011      4187  \n",
"7      Online      H      2012-07-17      871543967      7/27/2012      8082  \n",
"8      Offline      M      2015-07-14      770463311      8/25/2015      6070  \n",
"9      Online      H      2014-04-18      616607081      5/30/2014      6593  \n",
"\n",
"  Unit Price      Unit Cost      Total Revenue      Total Cost      Total Profit      Date  \\\n",
"0      255.28      159.42      2533654.00      1582243.50      951410.50      2010-05-28  \n",
"1      205.70      117.11      576782.80      328376.44      248406.36      2012-08-22  \n",
"2      651.21      524.96      1158502.59      933903.84      224598.75      2014-05-02  \n",
"3      9.33      6.92      75591.66      56065.84      19525.82      2014-06-20  \n",
"4      651.21      524.96      3296425.02      2657347.52      639077.50      2013-02-01  \n",
"5      255.28      159.42      759202.72      474115.08      285087.64      2015-02-04  \n",
"6      668.27      502.54      2798046.49      2104134.98      693911.51      2011-04-23  \n",
"7      154.06      90.93      1245112.92      734896.26      510216.66      2012-07-17  \n",
"8      81.73      56.67      496101.10      343986.90      152114.20      2015-07-14  \n",
"9      205.70      117.11      1356180.10      772106.23      584073.87      2014-04-18  \n",
"\n",
"  Month      Year  \n",
"0      5      2010  \n",
"1      8      2012  \n",
"2      5      2014  \n",
"3      6      2014  \n",
"4      2      2013  \n",
"5      2      2015  \n",
"6      4      2011  \n",
"7      7      2012  \n",
"8      7      2015  \n",
"9      4      2014  "
]
},
"execution_count": 82,
"metadata": {},
"output_type": "execute_result"
}
],
"source": [
"# Convert the 'Order Date' column to datetime format\n",
"df['Order Date'] = pd.to_datetime(df['Order Date'])\n",
"\n",
"# Add the new datetime column to the dataframe\n",
"df['Date'] = df['Order Date']\n",
"\n",
"# Display the first few rows to confirm the changes\n",
"df.head(10)"
]
},
{
"cell_type": "code",
"execution_count": 53,
"id": "284f7d7a-a615-49cc-af5e-8559e1ffeb29",
"metadata": {},
"outputs": [
{
"data": {
"text/html": [
<div>\n",
<style scoped>\n",
.dataframe tbody tr th:only-of-type {\n",
vertical-align: middle;\n",
}\n",
\n",
.dataframe tbody tr th {\n",
vertical-align: top;\n",
}\n",
\n",
\n",
.dataframe thead th {\n",
text-align: right;\n",
}\n",
</style>\n",
<table border="1" class="dataframe">\n",
<thead>\n",
<tr style="text-align: right;">\n",
<th></th>\n",
<th>Order Date</th>\n",
<th>Order ID</th>\n",

```

```
"      <th>Units Sold</th>\n",
"      <th>Unit Price</th>\n",
"      <th>Unit Cost</th>\n",
"      <th>Total Revenue</th>\n",
"      <th>Total Cost</th>\n",
"      <th>Total Profit</th>\n",
"      <th>Date</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>count</th>\n",
"      <td>100</td>\n",
"      <td>1.000000e+02</td>\n",
"      <td>100.000000</td>\n",
"      <td>100.000000</td>\n",
"      <td>100.000000</td>\n",
"      <td>1.000000e+02</td>\n",
"      <td>1.000000e+02</td>\n",
"      <td>1.000000e+02</td>\n",
"      <td>100</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>mean</th>\n",
"      <td>2013-09-16 14:09:36</td>\n",
"      <td>5.550204e+08</td>\n",
"      <td>5128.710000</td>\n",
"      <td>276.761300</td>\n",
"      <td>191.048000</td>\n",
"      <td>1.373488e+06</td>\n",
"      <td>9.318057e+05</td>\n",
"      <td>4.416820e+05</td>\n",
"      <td>2013-09-16 14:09:36</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>min</th>\n",
"      <td>2010-02-02 00:00:00</td>\n",
"      <td>1.146066e+08</td>\n",
"      <td>124.000000</td>\n",
"      <td>9.330000</td>\n",
"      <td>6.920000</td>\n",
"      <td>4.870260e+03</td>\n",
"      <td>3.612240e+03</td>\n",
"      <td>1.258020e+03</td>\n",
"      <td>2010-02-02 00:00:00</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>25%</th>\n",
"      <td>2012-02-14 12:00:00</td>\n",
"      <td>3.389225e+08</td>\n",
"      <td>2836.250000</td>\n",
"      <td>81.730000</td>\n",
"      <td>35.840000</td>\n",
"      <td>2.687212e+05</td>\n",
"      <td>1.688680e+05</td>\n",
"      <td>1.214436e+05</td>\n",
"      <td>2012-02-14 12:00:00</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>50%</th>\n",
"      <td>2013-07-12 12:00:00</td>\n",
"      <td>5.577086e+08</td>\n",
"      <td>5382.500000</td>\n",
"      <td>179.880000</td>\n",
"      <td>107.275000</td>\n",
"      <td>7.523144e+05</td>\n",
"      <td>3.635664e+05</td>\n",
"      <td>2.907680e+05</td>\n",
"      <td>2013-07-12 12:00:00</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>75%</th>\n",
"      <td>2015-04-07 00:00:00</td>\n",
"      <td>7.907551e+08</td>\n",
"      <td>7369.000000</td>\n",
"      <td>437.200000</td>\n",
"      <td>263.330000</td>\n",
"      <td>2.212045e+06</td>\n",
"      <td>1.613870e+06</td>\n",
"      <td>6.358288e+05</td>\n",
"      <td>2015-04-07 00:00:00</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>max</th>\n",
"      <td>2017-05-22 00:00:00</td>\n",
"      <td>9.940222e+08</td>\n",
"      <td>9925.000000</td>\n",
"      <td>668.270000</td>\n",
"      <td>524.960000</td>\n",
"      <td>5.997055e+06</td>\n",
"      <td>4.509794e+06</td>\n",
"      <td>1.719922e+06</td>\n",
"      <td>2017-05-22 00:00:00</td>\n",
"    </tr>\n",
"    <tr>\n",
"      <th>std</th>\n",
"      <td>NaN</td>\n",
"      <td>2.606153e+08</td>\n",
"      <td>2794.484562</td>\n",
"      <td>235.592241</td>\n",
"      <td>188.208181</td>\n",
"      <td>1.460029e+06</td>\n",
"      <td>1.083938e+06</td>\n",
"      <td>4.385379e+05</td>\n",
"      <td>NaN</td>\n",
"    </tr>\n",
"
```

```

" </tbody>\n",
"</table>\n",
"</div>"
],
"text/plain": [
"
      Order Date      Order ID      Units Sold      Unit Price      Unit Cost  \\n",
"count          100      1.000000e+02      100.000000      100.000000      100.000000  \n",
"mean    2013-09-16 14:09:36  5.550204e+08  5128.710000      276.761300      191.048000  \n",
"min      2010-02-02 00:00:00  1.146066e+08      124.000000      9.330000      6.920000  \n",
"25%      2012-02-14 12:00:00  3.389225e+08  2836.250000      81.730000      35.840000  \n",
"50%      2013-07-12 12:00:00  5.577086e+08  5382.500000      179.880000      107.275000  \n",
"75%      2015-04-07 00:00:00  7.907551e+08  7369.000000      437.200000      263.330000  \n",
"max      2017-05-22 00:00:00  9.940222e+08  9925.000000      668.270000      524.960000  \n",
"std              NaN      2.606153e+08      2794.484562      235.592241      188.208181  \n",
"
      Total Revenue      Total Cost      Total Profit              Date  \n",
"count      1.000000e+02      1.000000e+02      1.000000e+02          100  \n",
"mean      1.373488e+06      9.318057e+05      4.416820e+05    2013-09-16 14:09:36  \n",
"min      4.870260e+03      3.612240e+03      1.258020e+03    2010-02-02 00:00:00  \n",
"25%      2.687212e+05      1.688680e+05      1.214436e+05    2012-02-14 12:00:00  \n",
"50%      7.523144e+05      3.635664e+05      2.907680e+05    2013-07-12 12:00:00  \n",
"75%      2.212045e+06      1.613870e+06      6.358288e+05    2015-04-07 00:00:00  \n",
"max      5.997055e+06      4.509794e+06      1.719922e+06    2017-05-22 00:00:00  \n",
"std      1.460029e+06      1.083938e+06      4.385379e+05              NaN  "
]
},
"execution_count": 53,
"metadata": {},
"output_type": "execute_result"
}
],
"source": [
"df.describe()"
]
},
{
"cell_type": "markdown",
"id": "a6c96c34-cd76-4dcl-aafb-6269e7ba3fed",
"metadata": {},
"source": [
"## Monthly Sales Channel Trend"
]
},
{
"cell_type": "code",
"execution_count": 56,
"id": "72635177-cea9-4abb-b4ed-dfbc14340e72",
"metadata": {},
"outputs": [
{
"data": {
"image/png": "iVBORw0KGGoAAAANSUUhEUgAAAk8AAAHHCAYAAACmzLxGAAAOXRFWHRtb2Z0d2FyZQBhYXRwbG90bGl1IHZlcnNpb24zLjkuMSwgaHR0cHM6Ly9tYXRwbG90bGl1Lm9yZy/TGe4hAAA",
"text/plain": [
"<Figure size 640x480 with 1 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
],
"source": [
"df['Month'] = pd.to_datetime(df['Date']).dt.month\n",
"sns.lineplot(x='Month', y='Sales Channel', data=df)\n",
"plt.title('Monthly Sales Channel Trend')\n",
"plt.show()"
]
},
{
"cell_type": "markdown",
"id": "5507ac07-7619-4f4a-83e2-bc7591a7bdde",
"metadata": {},
"source": [
"## Year Wise Sales Channel Trend"
]
},
{
"cell_type": "code",
"execution_count": 57,
"id": "a9bf6bca-de34-42d3-ac49-c783f2b7b0f4",
"metadata": {},
"outputs": [
{
"data": {
"image/png": "iVBORw0KGGoAAAANSUUhEUgAAAk0AAAHHCAYAAAC1OWx7AAAAOXRFWHRtb2Z0d2FyZQBhYXRwbG90bGl1IHZlcnNpb24zLjkuMSwgaHR0cHM6Ly9tYXRwbG90bGl1Lm9yZy/TGe4hAAA",
"text/plain": [
"<Figure size 640x480 with 1 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
],
"source": [
"df['Year'] = pd.to_datetime(df['Date']).dt.year\n",
"sns.lineplot(x='Year', y='Sales Channel', data=df)\n",
"plt.title('Year-Wise Sales Channel Trend')\n",
"plt.show()"
]
},
{
"cell_type": "markdown",
"id": "a3bblfa2-cf94-4823-blab-9d300a3d14d8",
"metadata": {},
"source": [
"## Yearly Month-Wise Sales Channel Trend"
]
}
]

```

```

},
{
  "cell_type": "code",
  "execution_count": 58,
  "id": "d97lcccd-3e93-476f-8f5a-9fc0fe337bd9",
  "metadata": {},
  "outputs": [
    {
      "data": {
        "image/png":
        "text/plain": [
          "<Figure size 640x480 with 1 Axes>"
        ]
      },
      "metadata": {},
      "output_type": "display_data"
    }
  ],
  "source": [
    "sns.lineplot(x='Month', y='Sales Channel', hue='Year', data=df)\n",
    "plt.title('Yearly Month-Wise Sales Channel Trend')\n",
    "plt.show()"
  ]
},
{
  "cell_type": "code",
  "execution_count": 66,
  "id": "d9d62127-eb0c-407d-9fel-561fb331ad34",
  "metadata": {},
  "outputs": [
    {
      "name": "stdout",
      "output_type": "stream",
      "text": [
        "
        Order Date    Order ID    Units Sold    Unit Price    Unit Cost    \\n",
        1.000000    0.068436    0.011993    -0.067132    -0.078461    \n",
        Order ID        0.068436    1.000000    -0.222907    -0.190941    -0.213201    \n",
        Units Sold      0.011993    -0.222907    1.000000    -0.070486    -0.092232    \n",
        Unit Price      -0.067132    -0.190941    -0.070486    1.000000    0.987270    \n",
        Unit Cost       -0.078461    -0.213201    -0.092232    0.987270    1.000000    \n",
        Total Revenue   -0.036936    -0.314688    0.447784    0.752360    0.715623    \n",
        Total Cost      -0.053604    -0.328944    0.374746    0.787905    0.774895    \n",
        Total Profit    0.009523    -0.234638    0.564550    0.557365    0.467214    \n",
        Date            1.000000    0.068436    0.011993    -0.067132    -0.078461    \n",
        Month           0.029971    -0.111219    -0.007995    -0.031917    -0.042016    \n",
        Year            0.990581    0.081752    0.012455    -0.061791    -0.071567    \n",
        \n",
        Total Revenue   Total Cost    Total Profit    Date        Month    \\n",
        Order Date      -0.036936    -0.053604    0.009523    1.000000    0.029971    \n",
        Order ID        -0.314688    -0.328944    -0.234638    0.068436    -0.111219    \n",
        Units Sold      0.447784    0.374746    0.564550    0.011993    -0.007995    \n",
        Unit Price      0.752360    0.787905    0.557365    -0.067132    -0.031917    \n",
        Unit Cost       0.715623    0.774895    0.467214    -0.078461    -0.042016    \n",
        Total Revenue   1.000000    0.983928    0.897327    -0.036936    0.003835    \n",
        Total Cost      0.983928    1.000000    0.804091    -0.053604    -0.015617    \n",
        Total Profit    0.897327    0.804091    1.000000    0.009523    0.051366    \n",
        Date            -0.036936    -0.053604    0.009523    1.000000    0.029971    \n",
        Month           0.003835    -0.015617    0.051366    0.029971    1.000000    \n",
        Year            -0.037128    -0.050899    0.002196    0.990581    -0.106715    \n",
        \n",
        Year    \n",
        Order Date    0.990581    \n",
        Order ID      0.081752    \n",
        Units Sold    0.012455    \n",
        Unit Price    -0.061791    \n",
        Unit Cost     -0.071567    \n",
        Total Revenue -0.037128    \n",
        Total Cost    -0.050899    \n",
        Total Profit  0.002196    \n",
        Date         0.990581    \n",
        Month        -0.106715    \n",
        Year         1.000000    \n",
        1373487.6831\n",
        2131683824748.7039\n",
        3\n"
      ]
    }
  ],
  "source": [
    "# exclude object (string) columns\n",
    "numeric_df = df.select_dtypes(exclude=[object])\n",
    "\n",
    "from pandas import get_dummies\n",
    "\n",
    "categorical_cols = ['Region', 'Country'] # replace with your categorical column names\n",
    "df_encoded = get_dummies(df, columns=categorical_cols)\n",
    "\n",
    "# Exclude non-numeric columns\n",
    "numeric_df = df.select_dtypes(exclude=[object])\n",
    "\n",
    "# Conduct correlation analysis on numeric columns\n",
    "correlation_matrix = numeric_df.corr()\n",
    "\n",
    "# Calculate average sales (Total Revenue)\n",
    "average_sales = df['Total Revenue'].mean()\n",
    "\n",
    "# Calculate sales variance (Total Revenue)\n",
    "sales_variance = df['Total Revenue'].var()\n",
    "\n",
    "# Identify potential outliers using the IQR method\n",
    "Q1 = df['Total Revenue'].quantile(0.25)\n",
    "Q3 = df['Total Revenue'].quantile(0.75)\n",
    "IQR = Q3 - Q1\n",
    "\n",
    "# Define outliers as points outside 1.5 * IQR above Q3 or below Q1\n",
    "outliers = df[(df['Total Revenue'] < (Q1 - 1.5 * IQR)) | (df['Total Revenue'] > (Q3 + 1.5 * IQR))]\n"
  ]
}

```

```

"\n",
"# Display results\n",
"print(correlation_matrix)\n",
"print(average_sales)\n",
"print(sales_variance)\n",
"print(outliers.shape[0])"
],
},
{
"cell_type": "code",
"execution_count": 67,
"id": "80c8e3a6-95e6-4b31-a1ef-032bcc4e40a4",
"metadata": {},
"outputs": [
{
"data": {
"image/png":
"text/plain": [
"<Figure size 1000x600 with 1 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
],
"source": [
"# Creating a line graph for sales trend over time\n",
"\n",
"plt.figure(figsize=(10, 6))\n",
"plt.plot(df['Date'], df['Total Revenue'])\n",
"plt.title('Sales Trend Over Time')\n",
"plt.xlabel('Date')\n",
"plt.ylabel('Total Revenue')\n",
"plt.show()"
]
},
{
"cell_type": "code",
"execution_count": 71,
"id": "af936a0b-deac-4c92-8703-cee29ccabdc",
"metadata": {},
"outputs": [
{
"data": {
"image/png": "iVBORw0KGgoAAAANSUHEUgAAA04AAAJocAYAAABRIrEeAAAAOXRFWHRTb2Z0dFyZQBNYXRwbG90bGliIHZlcnNpb24zLjkuMSwgaHR0cHM6Ly9tYXRwbG90bGliLm9yZy/TGe4hAAA"
"text/plain": [
"<Figure size 1000x600 with 1 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
],
"source": [
"# Creating a bar graph for top-selling products\n",
"\n",
"top_products = df.groupby('Item Type')['Total Revenue'].sum().sort_values(ascending=False).head(10)\n",
"\n",
"plt.figure(figsize=(10, 6))\n",
"sns.barplot(x=top_products.index, y=top_products.values)\n",
"plt.title('Top-Selling Item Types')\n",
"plt.xlabel('Item Type')\n",
"plt.ylabel('Total Revenue')\n",
"plt.xticks(rotation=45) # Rotating x labels for better readability\n",
"plt.show()\n"
]
},
{
"cell_type": "code",
"execution_count": 75,
"id": "321c2216-37de-42ed-a466-4292b84d7f48",
"metadata": {},
"outputs": [
{
"data": {
"image/png":
"text/plain": [
"<Figure size 750x750 with 12 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
],
"source": [
"# Creating a scatter plot for correlation between attributes\n",
"\n",
"sns.pairplot(df[['Total Revenue', 'Units Sold', 'Unit Price']])\n",
"plt.show()"
]
},
{
"cell_type": "code",
"execution_count": 83,
"id": "af952a7c-7d71-4707-9a29-667f667a1365",
"metadata": {},
"outputs": [
{
"data": {
"image/png":
"text/plain": [
"<Figure size 1000x800 with 2 Axes>"
]
},
"metadata": {},
"output_type": "display_data"
}
]
}

```

