```
import pandas as pd
import numpy as np
import matplotlib.pyplot as pyplot
import seaborn as sns
```

```
data = pd.read_csv("/content/Amazon Sales data.csv",encoding = 'latin-1')
```

```
# Show dataframe
data
```

	Region	Country	Item Type	Sales Channel	Order Priority	Order ID	Ship Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit	Order Month	Order Year	Or D Month
0	Australia and Oceania	Tuvalu	0	0	1	669165933	2010-06-27	9925	255.28	159.42	2533654.00	1582243.50	951410.50	5	2010	201
1	Central America and the Caribbean	Grenada	2	1	0	963881480	2012-09-15	2804	205.70	117.11	576782.80	328376.44	248406.36	8	2012	201
2	Europe	Russia	8	0	2	341417157	2014-05-08	1779	651.21	524.96	1158502.59	933903.84	224598.75	5	2014	201
3	Sub-Saharan Africa	Sao Tome and Principe	5	1	0	514321792	2014-07-05	8102	9.33	6.92	75591.66	56065.84	19525.82	6	2014	201
4	Sub-Saharan Africa	Rwanda	8	0	2	115456712	2013-02-06	5062	651.21	524.96	3296425.02	2657347.52	639077.50	2	2013	201
5	Australia and Oceania	Solomon Islands	0	1	0	547995746	2015-02-21	2974	255.28	159.42	759202.72	474115.08	285087.64	2	2015	201
6	Sub-Saharan Africa	Angola	6	0	3	135425221	2011-04-27	4187	668.27	502.54	2798046.49	2104134.98	693911.51	4	2011	201
7	Sub-Saharan Africa	Burkina Faso	11	1	1	871543967	2012-07-27	8082	154.06	90.93	1245112.92	734896.26	510216.66	7	2012	201
8	Sub-Saharan Africa	Republic of the Congo	9	0	3	770463311	2015-08-25	6070	81.73	56.67	496101.10	343986.90	152114.20	7	2015	201
9	Sub-Saharan Africa	Senegal	2	1	1	616607081	2014-05-30	6593	205.70	117.11	1356180.10	772106.23	584073.87	4	2014	201
10	Asia	Kyrgyzstan	11	1	1	814711606	2011-07-12	124	154.06	90.93	19103.44	11275.32	7828.12	6	2011	201
11	Sub-Saharan Africa	Cape Verde	3	0	1	939825713	2014-08-19	4168	109.28	35.84	455479.04	149381.12	306097.92	8	2014	201
12	Asia	Bangladesh	3	1	2	187310731	2017-03-01	8263	109.28	35.84	902980.64	296145.92	606834.72	1	2017	201
13	Central America and the Caribbean	Honduras	6	0	1	522840487	2017-02-13	8974	668.27	502.54	5997054.98	4509793.96	1487261.02	2	2017	201
14	Asia	Mongolia	9	0	0	832401311	2014-02-23	4901	81.73	56.67	400558.73	277739.67	122819.06	2	2014	201
15	Europe	Bulgaria	3	1	3	972292029	2012-06-03	1673	109.28	35.84	182825.44	59960.32	122865.12	4	2012	201
16	Asia	Sri Lanka	4	0	3	419123971	2016-12-18	6952	437.20	263.33	3039414.40	1830670.16	1208744.24	11	2016	201
17	Sub-Saharan Africa	Cameroon	1	0	0	519820964	2015-04-18	5430	47.45	31.79	257653.50	172619.70	85033.80	4	2015	201
18	Asia	Turkmenistan	6	0	2	441619336	2011-01-20	3830	668.27	502.54	2559474.10	1924728.20	634745.90	12	2010	201
19	Australia and Oceania	East Timor	7	1	2	322067916	2012-09-11	5908	421.89	364.69	2492526.12	2154588.52	337937.60	7	2012	201
20	Europe	Norway	0	1	2	819028031	2014-06-28	7450	255.28	159.42	1901836.00	1187679.00	714157.00	5	2014	201
21	Europe	Portugal	0	1	1	860673511	2015-09-03	1273	255.28	159.42	324971.44	202941.66	122029.78	7	2015	201
22	Central America and the Caribbean	Honduras	10	1	2	795490682	2016-07-26	2225	152.58	97.44	339490.50	216804.00	122686.50	6	2016	201
23	Australia and Oceania	New Zealand	5	1	1	142278373	2014-10-04	2187	9.33	6.92	20404.71	15134.04	5270.67	9	2014	201
24	Europe	Moldova	9	1	2	740147912	2016-05-10	5070	81.73	56.67	414371.10	287316.90	127054.20	5	2016	201

25	Europe	France	4	1	1	898523128	2017-06-05	1815	437.20	263.33	793518.00	477943.95	315574.05	5	2017	201
26	Australia and Oceania	Kiribati	5	1	3	347140347	2014-11-10	5398	9.33	6.92	50363.34	37354.16	13009.18	10	2014	201
27	Sub-Saharan Africa	Mali	5	1	2	686048400	2010-05-10	5822	9.33	6.92	54319.26	40288.24	14031.02	5	2010	201
28	Europe	Norway	1	0	0	435608613	2014-07-30	5124	47.45	31.79	243133.80	162891.96	80241.84	7	2014	201
29	Sub-Saharan Africa	The Gambia	6	0	2	886494815	2012-06-09	2370	668.27	502.54	1583799.90	1191019.80	392780.10	5	2012	201
30	Europe	Switzerland	4	0	3	249693334	2012-10-20	8661	437.20	263.33	3786589.20	2280701.13	1505888.07	9	2012	201
31	Sub-Saharan Africa	South Sudan	9	0	0	406502997	2014-01-28	2125	81.73	56.67	173676.25	120423.75	53252.50	12	2013	201
32	Australia and Oceania	Australia	8	1	0	158535134	2015-11-25	2924	651.21	524.96	1904138.04	1534983.04	369155.00	10	2015	201
33	Asia	Myanmar	6	0	1	177713572	2015-03-01	8250	668.27	502.54	5513227.50	4145955.00	1367272.50	1	2015	201
34	Sub-Saharan Africa	Djibouti	10	1	3	756274640	2017-02-25	7327	152.58	97.44	1117953.66	713942.88	404010.78	2	2017	201
35	Central America and the Caribbean	Costa Rica	9	0	2	456767165	2017-05-21	6409	81.73	56.67	523807.57	363198.03	160609.54	5	2017	201
36	Middle East and North Africa	Syria	5	1	2	162052476	2011-12-03	3784	9.33	6.92	35304.72	26185.28	9119.44	11	2011	201
37	Sub-Saharan Africa	The Gambia	7	1	3	825304400	2017-01-23	4767	421.89	364.69	2011149.63	1738477.23	272672.40	1	2017	201
38	Asia	Brunei	8	1	2	320009267	2012-05-08	6708	651.21	524.96	4368316.68	3521431.68	846885.00	4	2012	201
39	Europe	Bulgaria	8	1	3	189965903	2012-02-28	3987	651.21	524.96	2596374.27	2093015.52	503358.75	2	2012	201
40	Sub-Saharan Africa	Niger	9	1	1	699285638	2017-03-28	3015	81.73	56.67	246415.95	170860.05	75555.90	3	2017	201
41	Middle East and North Africa	Azerbaijan	4	1	3	382392299	2010-02-25	7234	437.20	263.33	3162704.80	1904929.22	1257775.58	2	2010	201
42	Sub-Saharan Africa	The Gambia	2	0	1	994022214	2012-06-08	2117	205.70	117.11	435466.90	247921.87	187545.03	6	2012	201
43	Europe	Slovakia	11	1	1	759224212	2012-11-10	171	154.06	90.93	26344.26	15549.03	10795.23	10	2012	201
44	Asia	Myanmar	3	1	1	223359620	2015-11-18	5930	109.28	35.84	648030.40	212531.20	435499.20	11	2015	201
45	Sub-Saharan Africa	Comoros	2	0	1	902102267	2016-04-29	962	205.70	117.11	197883.40	112659.82	85223.58	3	2016	201
46	Europe	Iceland	4	1	0	331438481	2016-12-31	8867	437.20	263.33	3876652.40	2334947.11	1541705.29	12	2016	201
47	Europe	Switzerland	9	1	3	617667090	2011-01-31	273	81.73	56.67	22312.29	15470.91	6841.38	12	2010	201
48	Europe	Macedonia	3	0	0	787399423	2014-11-14	7842	109.28	35.84	856973.76	281057.28	575916.48	10	2014	201
49	Sub-Saharan Africa	Mauritania	8	0	0	837559306	2012-01-13	1266	651.21	524.96	824431.86	664599.36	159832.50	1	2012	201
50	Europe	Albania	3	1	0	385383069	2010-03-18	2269	109.28	35.84	247956.32	81320.96	166635.36	2	2010	201
	Sub-						2013-									

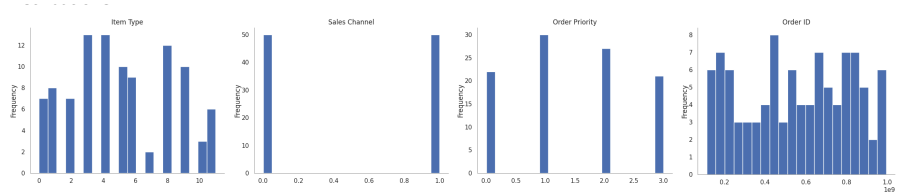
51	Saharan Africa	Lesotho	5	1	2	918419539	09-18	9606	9.33	6.92	89623.98	66473.52	23150.46	8	2013	201
52	Middle East and North Africa	Saudi Arabia	2	1	3	844530045	2013-03-28	4063	205.70	117.11	835759.10	475817.93	359941.17	3	2013	201
53	Sub-Saharan Africa	Sierra Leone	8	0	3	441888415	2012-01-07	3457	651.21	524.96	2251232.97	1814786.72	436446.25	11	2011	201
54	Sub-Saharan Africa	Sao Tome and Principe	5	0	1	508980977	2013-10-24	7637	9.33	6.92	71253.21	52848.04	18405.17	9	2013	201
55	Sub-Saharan Africa	Cote d'Ivoire	3	1	0	114606559	2012-06-27	3482	109.28	35.84	380512.96	124794.88	255718.08	6	2012	201
56	Australia and Oceania	Fiji	3	0	0	647876489	2010-08-01	9905	109.28	35.84	1082418.40	354995.20	727423.20	6	2010	201
57	Europe	Austria	4	0	1	868214595	2015-03-02	2847	437.20	263.33	1244708.40	749700.51	495007.89	2	2015	201
58	Europe	United Kingdom	6	1	2	955357205	2012-02-14	282	668.27	502.54	188452.14	141716.28	46735.86	1	2012	201
59	Sub-Saharan Africa	Djibouti	4	0	1	259353148	2014-04-19	7215	437.20	263.33	3154398.00	1899925.95	1254472.05	4	2014	201
60	Australia and Oceania	Australia	2	0	1	450563752	2013-07-02	682	205.70	117.11	140287.40	79869.02	60418.38	6	2013	201
61	Europe	San Marino	0	1	2	569662845	2013-07-01	4750	255.28	159.42	1212580.00	757245.00	455335.00	6	2013	201
62	Sub-Saharan Africa	Cameroon	8	1	3	177636754	2011-11-15	5518	651.21	524.96	3593376.78	2896729.28	696647.50	11	2011	201
63	Middle East and North Africa	Libya	3	0	1	705784308	2010-11-17	6116	109.28	35.84	668356.48	219197.44	449159.04	10	2010	201
64	Central America and the Caribbean	Haiti	4	0	1	505716836	2013-11-16	1705	437.20	263.33	745426.00	448977.65	296448.35	10	2013	201
65	Sub-Saharan Africa	Rwanda	4	0	1	699358165	2013-11-25	4477	437.20	263.33	1957344.40	1178928.41	778415.99	10	2013	201
66	Sub-Saharan Africa	Gabon	9	0	2	228944623	2012-07-09	8656	81.73	56.67	707454.88	490535.52	216919.36	7	2012	201
67	Central America and the Caribbean	Belize	3	0	3	807025039	2016-09-07	5498	109.28	35.84	600821.44	197048.32	403773.12	7	2016	201
68	Europe	Lithuania	8	0	1	166460740	2010-11-17	8287	651.21	524.96	5396577.27	4350343.52	1046233.75	10	2010	201
69	Sub-Saharan Africa	Madagascar	3	0	2	610425555	2015-05-28	7342	109.28	35.84	802333.76	263137.28	539196.48	4	2015	201
70	Asia	Turkmenistan	8	1	3	462405812	2013-05-20	5010	651.21	524.96	3262562.10	2630049.60	632512.50	4	2013	201
71	Middle East and North Africa	Libya	5	1	2	816200339	2015-09-30	673	9.33	6.92	6279.09	4657.16	1621.93	8	2015	201
72	Sub-Saharan Africa	Democratic Republic of the Congo	1	1	0	585920464	2011-07-15	5741	47.45	31.79	272410.45	182506.39	89904.06	5	2011	201
73	Sub-Saharan Africa	Djibouti	2	1	1	555990016	2017-06-17	8656	205.70	117.11	1780539.20	1013704.16	766835.04	5	2017	201
74	Middle East and North Africa	Pakistan	4	0	2	231145322	2013-08-16	9892	437.20	263.33	4324782.40	2604860.36	1719922.04	7	2013	201
	North						2014-									

75	North America	Mexico	6	0	0	986435210	12-12	6954	668.27	502.54	4647149.58	3494663.16	1152486.42	11	2014	201
76	Australia and Oceania	Federated States of Micronesia	1	1	0	217221009	2014-11-15	9379	47.45	31.79	445033.55	298158.41	146875.14	10	2014	201
77	Asia	Laos	11	0	0	789176547	2011-10-23	3732	154.06	90.93	574951.92	339350.76	235601.16	9	2011	201
78	Europe	Monaco	0	0	1	688288152	2012-06-02	8614	255.28	159.42	2198981.92	1373243.88	825738.04	5	2012	201
79	Australia and Oceania	Samoa	4	1	1	670854651	2013-08-07	9654	437.20	263.33	4220728.80	2542187.82	1678540.98	7	2013	201
80	Europe	Spain	6	0	2	213487374	2012-11-30	4513	668.27	502.54	3015902.51	2267963.02	747939.49	10	2012	201
81	Middle East and North Africa	Lebanon	3	1	2	663110148	2012-10-08	7884	109.28	35.84	861563.52	282562.56	579000.96	9	2012	201
82	Middle East and North Africa	Iran	4	1	1	286959302	2016-12-08	6489	437.20	263.33	2836990.80	1708748.37	1128242.43	11	2016	201
83	Sub-Saharan Africa	Zambia	10	1	2	122583663	2011-01-05	4085	152.58	97.44	623289.30	398042.40	225246.90	1	2011	201
84	Sub-Saharan Africa	Kenya	11	1	2	827844560	2012-04-07	6457	154.06	90.93	994765.42	587135.01	407630.41	3	2012	201
85	North America	Mexico	9	0	2	430915820	2012-03-20	6422	81.73	56.67	524870.06	363934.74	160935.32	2	2012	201
86	Sub-Saharan Africa	Sao Tome and Principe	1	0	0	180283772	2011-01-21	8829	47.45	31.79	418936.05	280673.91	138262.14	1	2011	201
87	Sub-Saharan Africa	The Gambia	0	0	3	494747245	2014-03-20	5559	255.28	159.42	1419101.52	886215.78	532885.74	2	2014	201
88	Middle East and North Africa	Kuwait	5	1	3	513417565	2012-05-18	522	9.33	6.92	4870.26	3612.24	1258.02	4	2012	201
89	Europe	Slovenia	1	0	0	345718562	2016-11-25	4660	47.45	31.79	221117.00	148141.40	72975.60	10	2016	201
90	Sub-Saharan Africa	Sierra Leone	8	0	1	621386563	2016-12-14	948	651.21	524.96	617347.08	497662.08	119685.00	12	2016	201
91	Australia and Oceania	Australia	1	0	1	240470397	2014-07-11	9389	47.45	31.79	445508.05	298476.31	147031.74	7	2014	201
92	Middle East and North Africa	Azerbaijan	8	1	3	423331391	2012-07-24	2021	651.21	524.96	1316095.41	1060944.16	255151.25	6	2012	201
93	Europe	Romania	4	1	1	660643374	2010-12-25	7910	437.20	263.33	3458252.00	2082940.30	1375311.70	11	2010	201
94	Central America and the Caribbean	Nicaragua	1	0	0	963392674	2011-03-21	8156	47.45	31.79	387002.20	259279.24	127722.96	2	2011	201
95	Sub-Saharan Africa	Mali	3	1	3	512878119	2011-09-03	888	109.28	35.84	97040.64	31825.92	65214.72	7	2011	201
96	Asia	Malaysia	5	0	2	810711038	2011-12-28	6267	9.33	6.92	58471.11	43367.64	15103.47	11	2011	201
97	Sub-Saharan Africa	Sierra Leone	11	0	0	728815257	2016-06-29	1485	154.06	90.93	228779.10	135031.05	93748.05	6	2016	201
98	North America	Mexico	9	0	3	559427106	2015-08-08	5767	81.73	56.67	471336.91	326815.89	144521.02	7	2015	201
99	Sub-Saharan Africa	Mozambique	6	0	2	665095412	2012-02-15	5367	668.27	502.54	3586605.09	2697132.18	889472.91	2	2012	201

Next

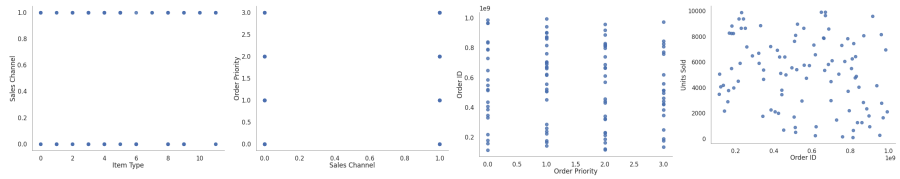
Generate code with `data` | View recommended plots | New interactive sheet

Distributions

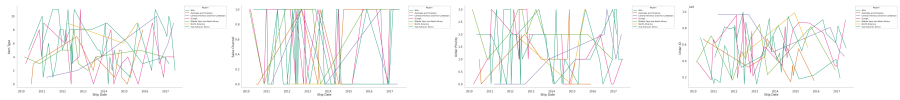


Categorical distributions

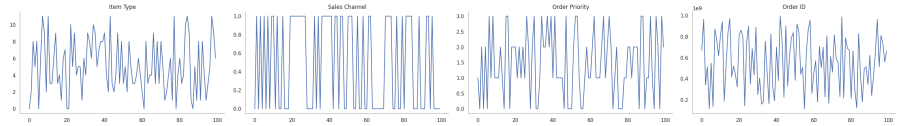
2-d distributions



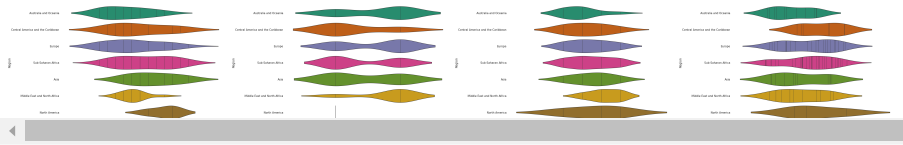
Time series



Values



Faceted distributions



```
data.columns
```

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

```
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
```

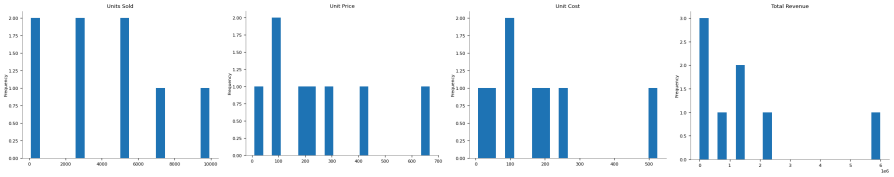
```
# Changing the data type of different column for model training and analysis
data['Order Date'] = pd.to_datetime(data['Order Date'])
data['Ship Date'] = pd.to_datetime(data['Ship Date'])
```

```
data['Region'] = data['Region'].astype(str)
data['Country'] = data['Country'].astype(str)
data['Item Type'] = data['Item Type'].astype(str)
data['Sales Channel'] = data['Sales Channel'].astype(str)
data['Order Priority'] = data['Order Priority'].astype(str)
```

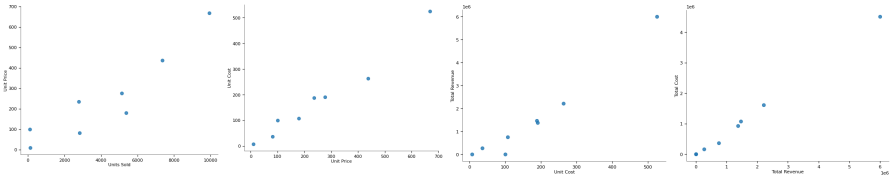
```
data[['Units Sold', 'Unit Price',   'Unit Cost', 'Total Revenue', 'Total Cost', 'Total Profit']].describe()
```

	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit
count	100.000000	100.000000	100.000000	1.000000e+02	1.000000e+02	1.000000e+02
mean	5128.710000	276.761300	191.048000	1.373488e+06	9.318057e+05	4.416820e+05
std	2794.484562	235.592241	188.208181	1.460029e+06	1.083938e+06	4.385379e+05
min	124.000000	9.330000	6.920000	4.870260e+03	3.612240e+03	1.258020e+03
25%	2836.250000	81.730000	35.840000	2.687212e+05	1.688680e+05	1.214436e+05
50%	5382.500000	179.880000	107.275000	7.523144e+05	3.635664e+05	2.907680e+05
75%	7369.000000	437.200000	263.330000	2.212045e+06	1.613870e+06	6.358288e+05
max	9925.000000	668.270000	524.960000	5.997055e+06	4.509794e+06	1.719922e+06

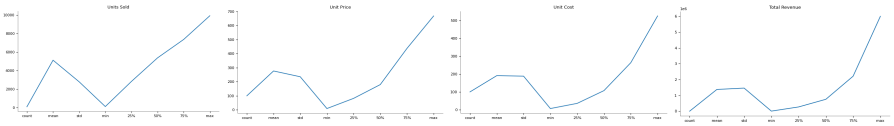
Distributions



2-d distributions



Values




```
data['Order Month'] = data['Order Date'].dt.month
data['Order Year'] = data['Order Date'].dt.year
data['Order Date MonthYear'] = data['Order Date'].dt.strftime('%Y-%m')
data = data.drop(columns=['Order Date'])
```

```
df = data
```

```
df.isnull().sum()
```

	0
Region	0
Country	0
Item Type	0
Sales Channel	0
Order Priority	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
Order Month	0
Order Year	0
Order Date MonthYear	0

dtype: int64

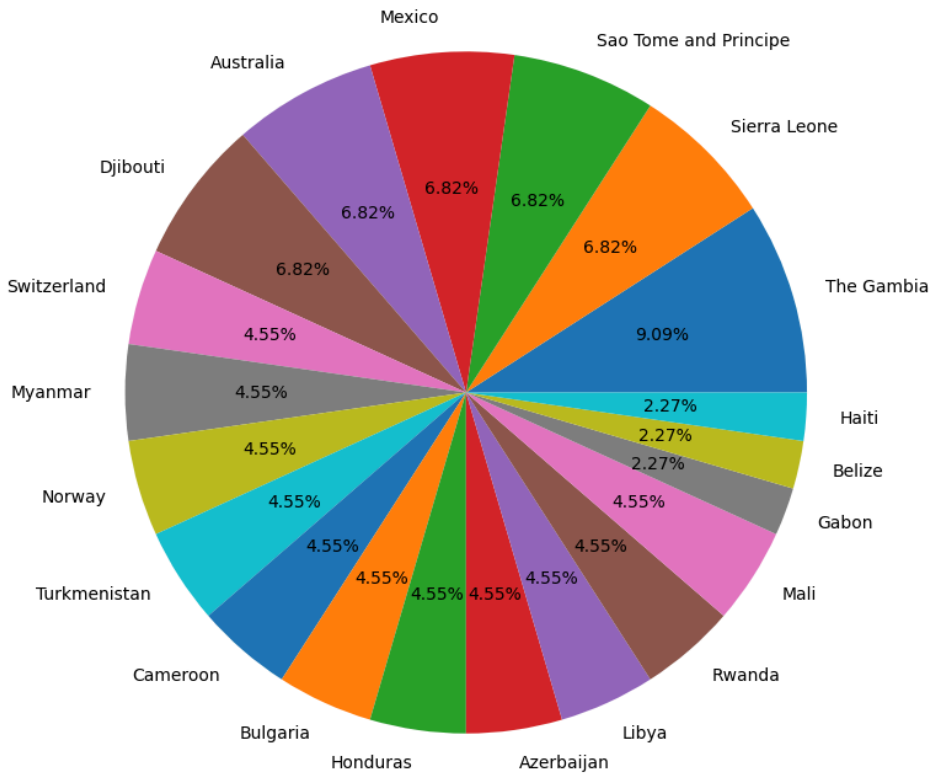
```
pd.set_option('display.max_rows', None)
df['Country'].value_counts()
```

Country	count
The Gambia	4
Sierra Leone	3
Sao Tome and Principe	3
Mexico	3
Australia	3
Djibouti	3
Switzerland	2
Myanmar	2
Norway	2
Turkmenistan	2
Cameroon	2
Bulgaria	2
Honduras	2
Azerbaijan	2
Libya	2
Rwanda	2
Mali	2
Gabon	1
Belize	1
Haiti	1
Lithuania	1
San Marino	1
United Kingdom	1
Austria	1
Fiji	1
Madagascar	1
Cote d'Ivoire	1
Tuvalu	1
Democratic Republic of the Congo	1
Zambia	1
Malaysia	1
Nicaragua	1
Romania	1
Slovenia	1
Kuwait	1
Kenya	1
Iran	1
Pakistan	1
Lebanon	1
Spain	1
Samoa	1
Monaco	1
Laos	1
Saudi Arabia	1
Federated States of Micronesia	1
Slovakia	1
Lesotho	1
Albania	1
Russia	1

Solomon Islands	1
Angola	1
Burkina Faso	1
Republic of the Congo	1
Senegal	1
Kyrgyzstan	1
Cape Verde	1
Bangladesh	1
Mongolia	1
Sri Lanka	1
East Timor	1
Portugal	1
New Zealand	1
Moldova	1
France	1
Kiribati	1
South Sudan	1
Costa Rica	1
Syria	1
Brunei	1
Niger	1
Grenada	1
Comoros	1
Iceland	1
Macedonia	1
Mauritania	1
Mozambique	1

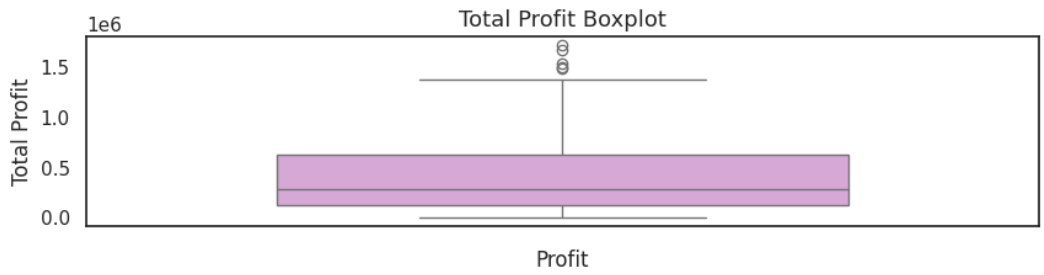
dtype: int64

```
import matplotlib.pyplot as plt
country_names = df.Country.value_counts().index
country_val = df.Country.value_counts().values
# Pie Chart for top 20 country
fig,ax = plt.subplots(figsize=(9,9))
ax.pie(country_val[:20],labels=country_names[:20],autopct='%1.2f%%')
plt.show()
```




```
import matplotlib.pyplot as plt
sns.set(style='white')
fig, ax = plt.subplots(figsize=(10, 2))
sns.boxplot(data['Total Profit'], color="plum", width=.6)

plt.title('Total Profit Boxplot', fontsize=13)
plt.xlabel('Profit')
plt.show()
```




```
def detect_outliers(dataframe, column):
    threshold = 2     ## 2rd standard deviation
    mean = np.mean(column)
    std = np.std(column)
    outliers = []

    for i, value in enumerate(column):
        z_score = (value - mean) / std
        if np.abs(z_score) > threshold:
            outliers.append(i)
            print(dataframe.loc[i])

    return outliers
```

```
outliers = detect_outliers(df, df["Total Profit"])
```

```
Order Year                              2016
Order Date MonthYear                 2016-12
Name: 46, dtype: object
Region                  Middle East and North Africa
Country                                 Pakistan
Item Type                               Cosmetics
Sales Channel                             Offline
Order Priority                                  L
Order ID                                231145322
Ship Date                       2013-08-16 00:00:00
Units Sold                                     9892
Unit Price                                    437.2
Unit Cost                                    263.33
Total Revenue                             4324782.4
Total Cost                               2604860.36
Total Profit                             1719922.04
Order Month                                       7
Order Year                                     2013
Order Date MonthYear                        2013-07
Name: 74, dtype: object
Region                  Australia and Oceania
Country                                 Samoa
Item Type                           Cosmetics
Sales Channel                          Online
Order Priority                              H
Order ID                            670854651
Ship Date                 2013-08-07 00:00:00
Units Sold                                9654
Unit Price                               437.2
Unit Cost                               263.33
Total Revenue                         4220728.8
Total Cost                           2542187.82
Total Profit                         1678540.98
Order Month                                   7
Order Year                                 2013
Order Date MonthYear                    2013-07
Name: 79, dtype: object
Region                                  Europe
Country                                 Romania
Item Type                           Cosmetics
Sales Channel                          Online
Order Priority                              H
Order ID                            660643374
Ship Date                 2010-12-25 00:00:00
Units Sold                                7910
Unit Price                               437.2
Unit Cost                               263.33
Total Revenue                         3458252.0
Total Cost                            2082940.3
Total Profit                          1375311.7
Order Month                                  11
Order Year                                 2010
Order Date MonthYear                    2010-11
Name: 93, dtype: object
```

```
# Print rows where outlier is present for the Total Profit column value
print(outliers)
```

```
[13, 33, 38, 68, 75]
```

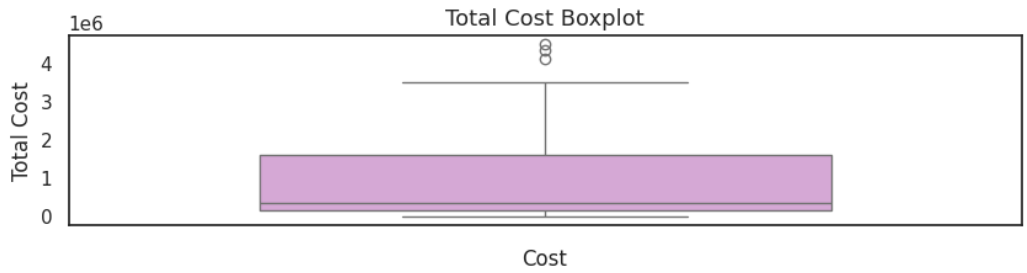
```
list_length = len(outliers)

# Print the number of values in the list
print("The list has", list_length, "outliers in Total Profit column of dataframe data ")
```

```
The list has 5 outliers in Total Profit column of dataframe data
```

```
sns.set(style='white')
fig, ax = plt.subplots(figsize=(10, 2))
sns.boxplot(data['Total Cost'], color="plum", width=.6)

plt.title('Total Cost Boxplot', fontsize=13)
plt.xlabel('Cost')
plt.show()
```




```
def detect_outliers(dataframe, column):
    threshold = 2     ## 3rd standard deviation
    mean = np.mean(column)
    std = np.std(column)
    outliers = []

    for i, value in enumerate(column):
        z_score = (value - mean) / std
        if np.abs(z_score) > threshold:
            outliers.append(i)
            print(dataframe.loc[i])

    return outliers
```

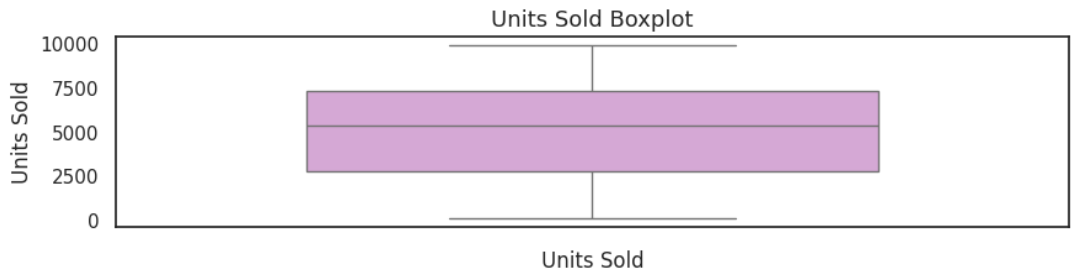
```
outliers = detect_outliers(df, df["Total Cost"])
```

```
Order Priority          1.00
Units Sold           8974.00
Unit Price            668.27
Unit Cost             502.54
Total Revenue     5997054.98
Total Cost        4509793.96
Total Profit      1487261.02
Order Month             2.00
Order Year           2017.00
Name: 13, dtype: float64
Item Type               6.00
Sales Channel           0.00
Order Priority          1.00
Units Sold           8250.00
Unit Price            668.27
Unit Cost             502.54
Total Revenue     5513227.50
Total Cost        4145955.00
Total Profit      1367272.50
Order Month             1.00
Order Year           2015.00
Name: 33, dtype: float64
Item Type               8.00
Sales Channel           1.00
Order Priority          2.00
Units Sold           6708.00
Unit Price            651.21
Unit Cost             524.96
Total Revenue     4368316.68
Total Cost        3521431.68
Total Profit       846885.00
Order Month             4.00
Order Year           2012.00
Name: 38, dtype: float64
Item Type               8.00
Sales Channel           0.00
Order Priority          1.00
Units Sold           8287.00
Unit Price            651.21
Unit Cost             524.96
Total Revenue     5396577.27
Total Cost        4350343.52
Total Profit      1046233.75
Order Month            10.00
Order Year           2010.00
Name: 68, dtype: float64
Item Type               6.00
Sales Channel           0.00
Order Priority          0.00
Units Sold           6954.00
Unit Price            668.27
Unit Cost             502.54
Total Revenue     4647149.58
```

```
Order Month            11.00
Order Year           2014.00
Name: 75, dtype: float64
```

```
sns.set(style='white')
fig, ax = plt.subplots(figsize=(10, 2))
sns.boxplot(data['Units Sold'], color="plum", width=.6)

plt.title('Units Sold Boxplot', fontsize=13)
plt.xlabel('Units Sold')
plt.show()
```




```
# Creating a bar chart for Total Revenue and Order Month
plt.bar(df['Order Month'], df['Total Revenue'])

# Set the chart title and axis labels
plt.title('Number of Orders Purchased by Month and Year')
plt.xticks([1,2,3,4,5,6,7,8,9,10,11,12])
plt.xlabel('Order Month')
plt.ylabel('Total Revenue')

# Rotate the x-axis labels for better readability

# Display the chart
plt.show()
```