```
import os
```

```
os.chdir("/content/drive/Shared/Jupyter/projects/Sales-Projects")
```

```
import pandas as pd
import numpy as np
import plotly.express as px
```

```
df = pd.read_csv('Amazon sales data.csv')
```

```
df.head(20)
```

	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	5/28/2010	669381691	6/27/2010	9625	255.28	159.42	2455941.50	1534053.50	921888.00
1	Central America and the Caribbean	Grenada	Cereal	Online	C	8/22/2012	963081093	9/12/2012	2804	205.70	117.11	576792.80	328376.44	248416.36
2	Europe	Russia	Office Supplies	Offline	L	5/2/2014	341417157	5/8/2014	1779	651.21	524.96	1158502.59	933903.84	224598.75
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	6/20/2014	514313782	7/9/2014	8102	9.33	6.92	75591.66	56093.84	19525.82
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2/13/2013	115467517	2/16/2013	5962	651.21	524.96	3883015.02	3129417.52	753597.50
5	Australia and Oceania	Solomon Islands	Baby Food	Offline	C	2/4/2015	547987948	2/21/2015	2974	255.28	159.42	759202.72	474118.08	285084.64
6	Sub-Saharan Africa	Angola	Household	Online	M	4/23/2011	135425221	4/27/2011	4147	668.27	502.54	2770884.38	2084134.38	686750.11
7	Sub-Saharan Africa	Burkina Faso	Vegetables	Online	M	7/17/2012	972245967	7/27/2012	8092	154.06	90.93	1246912.52	735696.36	511216.86
8	Sub-Saharan Africa	Republic of the Congo	Personal Care	Online	M	7/14/2015	770401201	8/25/2015	6070	81.73	56.67	495943.10	343956.90	151986.20
9	Sub-Saharan Africa	Senegal	Cereal	Online	H	4/18/2014	616807091	5/30/2014	8383	205.70	117.11	1724180.10	981728.21	742451.89
10	Asia	Kyrgyzstan	Vegetables	Online	L	8/24/2011	614711698	9/10/2011	134	154.06	90.93	20647.04	12184.62	8462.42
11	Europe	Bulgaria	Clothes	Online	M	3/13/2012	509657013	3/19/2012	4430	109.28	35.84	484079.04	158811.12	325267.92
12	Asia	Bangladesh	Clothes	Online	H	1/15/2017	187150731	3/1/2017	9263	109.28	35.84	923058.64	331991.92	591066.72
13	Central America and the Caribbean	Honduras	Household	Offline	H	2/8/2017	523584687	2/13/2017	8974	668.27	502.54	5997558.98	4509790.96	1487768.02
14	Asia	Mongolia	Personal Care	Online	C	2/18/2014	832401311	2/23/2014	4901	81.73	56.67	400593.13	277758.67	122834.46
15	Europe	Belgium	Clothes	Online	C	4/17/2013	117192921	5/12/2013	1670	109.28	35.84	182557.60	59853.12	122704.48
16	Asia	Sri Lanka	Cosmetics	Online	M	11/19/2016	419112691	12/18/2016	5952	437.20	263.33	2601641.40	1567679.16	1034962.24
17	Sub-Saharan Africa	Cameroon	Beverages	Online	C	4/15/2011	538629840	5/8/2011	9240	47.45	31.79	438318.00	293796.00	144522.00
18	Europe	Luxembourg	Clothes	Online	L	11/23/2010	441631928	12/20/2010	3930	109.28	35.84	429470.40	140851.20	288619.20
19	Australia and Oceania	East Timor	Meat	Online	L	7/31/2012	323879239	8/11/2012	5898	421.89	364.69	2488291.22	2150690.62	337600.60

```
df['Order Date'] = df['Order Date'].str.replace('-','/')
df['Ship Date'] = df['Ship Date'].str.replace('-','/')
```

```
df.head(5)
```

```
df = df.apply(str)
df['Units Sold'] = df['Units Sold'].astype(int)
df['Unit Cost'] = df['Unit Cost'].astype(float)
df['Unit Price'] = df['Unit Price'].astype(float)
df['Total Revenue'] = df['Total Revenue'].str.strip()
```

```
['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']
```

```
df.head(20)
```

	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	5/28/2010	669381691	6/27/2010	9625	255.28	159.42	2455941.50	1534053.50	921888.00
1	Central America and the Caribbean	Grenada	Cereal	Online	C	8/22/2012	963081093	9/12/2012	2804	205.70	117.11	576792.80	328376.44	248416.36
2	Europe	Russia	Office Supplies	Offline	L	5/2/2014	341417157	5/8/2014	1779	651.21	524.96	1158502.59	933903.84	224598.75
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	6/20/2014	514313782	7/9/2014	8102	9.33	6.92	75591.66	56093.84	19525.82
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2/13/2013	115467517	2/16/2013	5962	651.21	524.96	3883015.02	3129417.52	753597.50
5	Australia and Oceania	Solomon Islands	Baby Food	Offline	C	2/4/2015	547987948	2/21/2015	2974	255.28	159.42	759202.72	474118.08	285084.64
6	Sub-Saharan Africa	Angola	Household	Online	M	4/23/2011	135425221	4/27/2011	4147	668.27	502.54	2770884.38	2084134.38	686750.11
7	Sub-Saharan Africa	Burkina Faso	Vegetables	Online	M	7/17/2012	972245967	7/27/2012	8092	154.06	90.93	1246912.52	735696.36	511216.86
8	Sub-Saharan Africa	Republic of the Congo	Personal Care	Online	M	7/14/2015	770401201	8/25/2015	6070	81.73	56.67	495943.10	343956.90	151986.20
9	Sub-Saharan Africa	Senegal	Cereal	Online	H	4/18/2014	616807091	5/30/2014	8383	205.70	117.11	1724180.10	981728.21	742451.89
10	Asia	Kyrgyzstan	Vegetables	Online	L	8/24/2011	614711698	9/10/2011	134	154.06	90.93	20647.04	12184.62	8462.42
11	Europe	Bulgaria	Clothes	Online	M	3/13/2012	509657013	3/19/2012	4430	109.28	35.84	484079.04	158811.12	325267.92
12	Asia	Sri Lanka	Cosmetics	Online	M	11/19/2016	419112691	12/18/2016	5952	437.20	263.33	2601641.40	1567679.16	1034962.24
13	Sub-Saharan Africa	Cameroon	Beverages	Online	C	4/15/2011	538629840	5/8/2011	9240	47.45	31.79	438318.00	293796.00	144522.00
14	Asia	Turkmenistan	Household	Online	L	12/28/2010	441631928	1/20/2011	3930	668.27	502.54	2626074.11	1975187.22	650886.89
15	Australia and Oceania	East Timor	Meat	Online	L	7/31/2012	323879239	8/11/2012	5898	421.89	364.69	2488291.22	2150690.62	337600.60

```
df['Order Date'] = pd.to_datetime(df['Order Date'])
df.head()
```

	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	Year	Month
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	2010-05-28	669381691	6/27/2010	9625	255.28	159.42	2455941.50	1534053.50	921888.00	2010	05
1	Central America and the Caribbean	Grenada	Cereal	Online	C	2012-08-22	963081093	9/12/2012	2804	205.70	117.11	576792.80	328376.44	248416.36	2012	08
2	Europe	Russia	Office Supplies	Offline	L	2014-05-02	341417157	5/8/2014	1779	651.21	524.96	1158502.59	933903.84	224598.75	2014	05
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	2014-06-20	514313782	7/9/2014	8102	9.33	6.92	75591.66	56093.84	19525.82	2014	06
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2013-02-13	115467517	2/16/2013	5962	651.21	524.96	3883015.02	3129417.52	753597.50	2013	02

```
df['Year'] = df['Order Date'].astype(int)
df['Year'] = df['Order Date'].apply(lambda x x.split('-')[0])
df['Month'] = df['Order Date'].apply(lambda x x.split('-')[1])
```

```
df.head()
```

	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	Year	Month
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	2010-05-28	669381691	6/27/2010	9625	255.28	159.42	2455941.50	1534053.50	921888.00	2010	05
1	Central America and the Caribbean	Grenada	Cereal	Online	C	2012-08-22	963081093	9/12/2012	2804	205.70	117.11	576792.80	328376.44	248416.36	2012	08
2	Europe	Russia	Office Supplies	Offline	L	2014-05-02	341417157	5/8/2014	1779	651.21	524.96	1158502.59	933903.84	224598.75	2014	05
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	2014-06-20	514313782	7/9/2014	8102	9.33	6.92	75591.66	56093.84	19525.82	2014	06
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2013-02-13	115467517	2/16/2013	5962	651.21	524.96	3883015.02	3129417.52	753597.50	2013	02

```
df = df.drop('Order ID',axis=1)
```

```
df.head()
```

	Region	Country	Item Type	Sales Channel	Order Priority	Order Date	Ship Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit	Year	Month
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	2010-05-28	6/27/2010	9625	255.28	159.42	2455941.50	1534053.50	921888.00	2010	05
1	Central America and the Caribbean	Grenada	Cereal	Online	C	2012-08-22	9/12/2012	2804	205.70	117.11	576792.80	328376.44	248416.36	2012	08
2	Europe	Russia	Office Supplies	Offline	L	2014-05-02	5/8/2014	1779	651.21	524.96	1158502.59	933903.84	224598.75	2014	05
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	2014-06-20	7/9/2014	8102	9.33	6.92	75591.66	56093.84	19525.82	2014	06
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2013-02-13	2/16/2013	5962	651.21	524.96	3883015.02	3129417.52	753597.50	2013	02

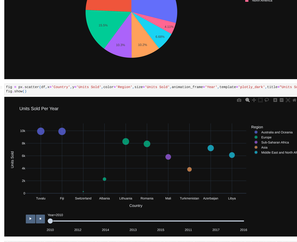
```
df.drop(['Order ID','Ship Date'],axis=1)
```

```
df.head()
```

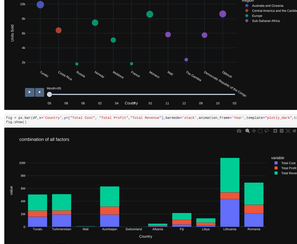
	Region	Country	Item Type	Sales Channel	Order Priority	Order Date	Units Sold	Unit Price	Unit Cost	Total Revenue	Total Cost	Total Profit	Year	Month
0	Australia and Oceania	Tuvalu	Baby Food	Offline	H	2010-05-28	9625	255.28	159.42	2455941.50	1534053.50	921888.00	2010	05
1	Central America and the Caribbean	Grenada	Cereal	Online	C	2012-08-22	2804	205.70	117.11	576792.80	328376.44	248416.36	2012	08
2	Europe	Russia	Office Supplies	Offline	L	2014-05-02	1779	651.21	524.96	1158502.59	933903.84	224598.75	2014	05
3	Sub-Saharan Africa	San Tome and Principe	Fruits	Online	C	2014-06-20	8102	9.33	6.92	75591.66	56093.84	19525.82	2014	06
4	Sub-Saharan Africa	Rwanda	Office Supplies	Online	L	2013-02-13	5962	651.21	524.96	3883015.02	3129417.52	753597.50	2013	02

```
import plotly.express as px
```

```
fig = px.pie(df,values='Total Revenue',names='Item Type',template='plotly_dark',title='Total revenue per item')
fig.show()
```




```
fig = px.pie(df,values='Total Revenue',names='Region',template='plotly_dark',title='Total revenue in Regions')
fig.show()
```




```
fig = px.scatter(df,x='Country',y='Units Sold',color='Region',size='Units Sold',animation_frame='Year',template='plotly_dark',title='Units Sold Per
fig.show()
```