

```

In [1]:
#Python Assignment 1

In [9]:
import pandas as pd

In [10]:
data =
pd.read_csv("D:\KSR\Python\Assignments\Assignment_1\Ele_Store1.csv")

In [23]:
data.shape

Out[23]:
(15000, 17)

In [24]:
#1)Query the list of columns?

In [9]:
data.columns

Out[9]:
Index(['OrderID', 'OrderDate', 'UnitCost', 'Price', 'OrderQty',
       'CostofSales',
       'Sales', 'Profit', 'Channel',
       'Manufacturer', 'ProductSubCategory', 'ProductCategory',
       'Region',
       'City', 'Country'],
      dtype='object')

```

```

In [10]:
#2)Query the top 5 rows?

```

```

In [25]:
data.head(5)

```

Out[25]:

	OrderID	OrderDate	UnitCost	Price	OrderQty	CostofSales	Sales	Profit	Channel
0	7077	13-09-2017	76.094968	304.00	9	684.854710	2714.7200	2029.865290	Store
1	117	14-09-2017	7.491753	12.99	4	29.967011	50.1414	20.174389	Store
2	7018	15-09-2017	10.122338	159.99	9	91.101039	1395.1128	1304.011761	Store

	OrderID	OrderDate	UnitCost	Price	OrderQty	CostofSales	Sales	Profit	Chann
3	140	16-09-2017	0.576153	25.69	18	10.370759	462.4200	452.049241	Store
4	491	17-09-2017	108.508777	304.00	9	976.578991	2614.4000	1637.821009	Onlin

In [12]:

#3) Query the bottom 5 rows?

In [26]:

data.tail(5)

Out[26]:

	OrderID	OrderDate	UnitCost	Price	OrderQty	CostofSales	Sales	Profit	Chann
14995	22348	03-10-2058	16.240455	25.0	12	194.885455	300.0	105.114545	Store
14996	22349	04-10-2058	55.387212	999.0	10	553.872121	9990.0	9436.127879	Resel
14997	22350	05-10-2058	363.634986	588.0	13	4727.254816	7408.8	2681.545184	Store
14998	22351	06-10-2058	113.237461	279.0	12	1358.849536	3348.0	1989.150464	Onlin
14999	22352	07-10-2058	10.936943	15.9	20	218.738862	318.0	99.261138	Store

In [14]:

#4) How many total number of orders placed?

In [36]:

```
data[['OrderID']].count()  
Out[36]:  
OrderID    15000  
dtype: int64  
In [30]:  
#5) What are the unique channels present and how many are there?  
In [99]:
```

```
data.groupby(['Channel'])['Channel'].count().reset_index(name =  
'TotalCount')
```

```
Out[99]:
```

	Channel	TotalCount
0	Catalog	884
1	Online	2351
2	Reseller	1784
3	Store	9981

```
In [38]:
```

```
#6) How many cities are there?
```

```
In [40]:
```

```
data[['City']].count()
```

```
Out[40]:
```

```
City    15000
```

```
dtype: int64
```

```
In [41]:
```

```
#7) Query the unique Countries and count of the countries?
```

```
In [60]:
```

```
data.groupby(['Country'])[['Country']].count().rename(columns =  
{"Country": "TotalCount"})
```

```
Out[60]:
```

	TotalCount
Country	
Armenia	44
Australia	149
Bhutan	48
Canada	359
China	1673
Denmark	42
France	781

	TotalCount
Country	
Germany	1118
Greece	25
India	152
Iran	89
Ireland	37
Italy	111
Japan	298
Kyrgyzstan	46
Malta	41
Pakistan	93
Poland	37
Portugal	38
Romania	39
Russia	166
Singapore	49
Slovenia	34
South Korea	72
Spain	32
Sweden	30
Switzerland	35
Syria	86
Taiwan	50
Thailand	82
Turkmenistan	119
United Kingdom	477
United States	8510
the Netherlands	38

In [61]:

#8) Query the unique Regions and count of the regions?

In [71]:

```
data.groupby(['Region'])[['Region']].count().rename(columns = {"Region": "TotalCount"})
```

Out[71]:

	TotalCount
Region	
Asia	3050
Europe	3081
North America	8869

In [72]:

#9) Query the total cities count by Country and order it from highest count to lowest count?

In [75]:

```
data.groupby("Country")[['City']].count().rename(columns={"City": "TotalCount"}).sort_values('TotalCount', ascending = False)
```

Out[75]:

	TotalCount
Country	
United States	8510
China	1673
Germany	1118
France	781
United Kingdom	477
Canada	359
Japan	298
Russia	166
India	152
Australia	149
Turkmenistan	119
Italy	111
Pakistan	93
Iran	89
Syria	86
Thailand	82
South Korea	72
Taiwan	50
Singapore	49
Bhutan	48
Kyrgyzstan	46

	TotalCount
Country	
Armenia	44
Denmark	42
Malta	41
Romania	39
the Netherlands	38
Portugal	38
Ireland	37
Poland	37
Switzerland	35
Slovenia	34
Spain	32
Sweden	30
Greece	25

In [76]:

#10) How many products are there? Display count of products

In [78]:

```
data['ProductName'].nunique()
```

Out[78]:

1638

In [79]:

#11) Query the total sales and total profit?

In [82]:

```
data.agg({'Sales':'sum','Profit':'sum'}).astype(int)
```

Out[82]:

Sales 55391759

Profit 31587437

dtype: int64

In [83]:

#12) Query the total sales by region and country?

In [85]:

```
data.groupby(['Region','Country'])[['Sales']].sum().rename(columns = {"Sales":"TotalSales"}).astype(int)
```

Out[85]:

		TotalSales
--	--	------------

Region	Country	
Asia	Armenia	172934
	Australia	632642
	Bhutan	161291
	China	7557973
	India	598404
	Iran	367953
	Japan	958334
	Kyrgyzstan	127114
	Pakistan	296570
	Singapore	99223
	South Korea	250601
	Syria	329671
	Taiwan	189009
	Thailand	258248
	Turkmenistan	412370
Europe	Denmark	79145
	France	2628400
	Germany	4314956
	Greece	68527
	Ireland	92123
	Italy	316881
	Malta	112295
	Poland	87829
	Portugal	126381
	Romania	102278
	Russia	434240
	Slovenia	79881
	Spain	143433
	Sweden	78106
	Switzerland	99362
	United Kingdom	1324267
	the Netherlands	107746
North America	Canada	1148316

		TotalSales
Region	Country	
	United States	31635240

In [86]:

#13) Query the total order qty by promotion name?

In [91]:

```
data.groupby('PromotionName')['OrderQty'].sum().reset_index(name='TotalOrderQty')
```

Out[91]:

	PromotionName	TotalOrderQty
0	Asian Holiday Promotion	17004
1	Asian Spring Promotion	10980
2	Asian Summer Promotion	9768
3	European Back-to-Scholl Promotion	7363
4	European Holiday Promotion	12328
5	European Spring Promotion	13205
6	No Discount	80062
7	North America Back-to-School Promotion	39393
8	North America Holiday Promotion	30392
9	North America Spring Promotion	30631

In [100]:

#14) Query the top 7 countries with sales from highest to lowest order?

In [103]:

```
data.groupby('Country')[['Sales']].sum().sort_values('Sales', ascending=False).astype(int).head(7).rename(columns={"Sales": "TotalSales"})
```

Out[103]:

	TotalSales
Country	
United States	31635240
China	7557973
Germany	4314956
France	2628400
United Kingdom	1324267
Canada	1148316

	TotalSales
Country	
Japan	958334

In [104]:

#15) How much the total sales happened in Asia region?

In [112]:

```
data[data.Region =='Asia'][['Sales']].sum().astype(int)
```

Out[112]:

```
Sales    12412345
```

```
dtype: int64
```

In [113]:

#16) How much total sales happened in Asia region and India?

In [149]:

```
data[(data.Region =='Asia') & (data.Country ==  
'India')][['Sales']].sum().astype(int)
```

Out[149]:

```
Sales    598404
```

```
dtype: int64
```

In [128]:

#17) How much profit got in Europe region and Denmark country?

In [129]:

```
data[(data.Region == 'Europe') & (data.Country ==  
'Denmark')][['Profit']].sum().astype(int)
```

Out[129]:

```
Profit    45194
```

```
dtype: int64
```

In [130]:

#18) Query the total order qty by Manufacturer and Product Name?

In [133]:

```
data.groupby(['Manufacturer', 'ProductName'])[['OrderQty']].count().re  
name(columns = {"OrderQty": "TotalOrderQty"})
```

Out[133]:

		TotalOrderQty
Manufacturer	ProductName	
A. Datum Corporation	A. Datum Advanced Digital Camera M300 Azure	7
	A. Datum Advanced Digital Camera M300 Black	10

		TotalOrderQty
Manufacturer	ProductName	
	A. Datum Advanced Digital Camera M300 Green	7
	A. Datum Advanced Digital Camera M300 Grey	12
	A. Datum Advanced Digital Camera M300 Orange	10
...
Wide World Importers	WWI Wireless Transmitter and Bluetooth Headphones X250 Black	4
	WWI Wireless Transmitter and Bluetooth Headphones X250 Blue	6
	WWI Wireless Transmitter and Bluetooth Headphones X250 Red	6
	WWI Wireless Transmitter and Bluetooth Headphones X250 Silver	4
	WWI Wireless Transmitter and Bluetooth Headphones X250 White	5

1638 rows × 1 columns

In [134]:

#19) Query the top 7 products with orderqty and sales and show it descending order based on total orderqty?

In [155]:

```
data.groupby('ProductName').agg({'OrderQty':'sum','Sales':'sum'}) \
    .head(7).rename(columns = {"OrderQty":"TotalOrderQty",
    "Sales":"TotalSales"}) \
        .sort_values('TotalOrderQty', ascending = False) \
            .head(7)
```

Out[155]:

	TotalOrderQty	TotalSales
ProductName		
A. Datum Advanced Digital Camera M300 Grey	207	38208.950
A. Datum Advanced Digital Camera M300 Black	142	26701.025
A. Datum Advanced Digital Camera M300 Orange	115	21385.325
A. Datum Advanced Digital Camera M300 Silver	103	19340.100
A. Datum Advanced Digital Camera M300 Green	96	17935.775

	TotalOrderQty	TotalSales
ProductName		
A. Datum Advanced Digital Camera M300 Pink	89	16616.275
A. Datum Advanced Digital Camera M300 Azure	77	14252.485

In [156]:

#20) Query the total order quantity greater than 10000 by city?

In [158]:

```
data.groupby(['City'])[['OrderQty']].sum() \
    .rename(columns = {"OrderQty": "TotalOrderQty"}) \
    .query('TotalOrderQty > 10000')
```

Out[158]:

	TotalOrderQty
City	
Beijing	30547
Berlin	13841
Bethesda	20498
North Harford	24806
Seattle	12174

In [159]:

#21)Query the total sales less than 1000 by product, display it in ascending order?

In [161]:

```
data.groupby(['ProductName'])[['Sales']].sum() \
    .rename(columns = {"Sales": "TotalSales"}) \
    .query('TotalSales > 1000') \
    .sort_values("TotalSales", ascending = True) \
    .astype(int)
```

Out[161]:

	TotalSales
ProductName	
Contoso DVD 48 DVD Storage Binder M50 Red	1011
SV 22xDVD X680 Black	1014
SV DVD Player M110 Silver	1026
Contoso Optical Wheel OEM PS/2 Mouse E60 White	1030
Contoso Phone with Memory Dialing-single line E88 Grey	1032
...	...

	TotalSales
ProductName	
Proseware Projector 1080p DLP86 Silver	358806
Proseware Projector 1080p LCD86 Black	360567
Contoso Projector 1080p X980 White	399903
Proseware Projector 1080p LCD86 White	417575
Contoso Projector 1080p X980 Black	445115

1588 rows × 1 columns

In [162]:

#22) Query the total profit greater than 200000 by product, display it in descending order?

In [164]:

```
data.groupby(['ProductName'])[['Profit']].sum() \
    .rename(columns = {"Profit":"TotalProfit"}) \
    .query('TotalProfit > 200000') \
    .sort_values("TotalProfit", ascending = False) \
    .astype(int)
```

Out[164]:

	TotalProfit
ProductName	
Contoso Projector 1080p X980 White	320513
Proseware Projector 1080p LCD86 White	315379
Contoso Projector 1080p X980 Black	252654
Proseware Projector 1080p LCD86 Black	232556
Proseware Projector 1080p DLP86 Silver	228823
Proseware Projector 1080p DLP86 Black	208705

In [165]:

#23) How much total sales happened in China and Beijing?

In [183]:

```
data[(data.Country == 'China') & (data.City == \
'Beijing)][['Sales']].sum().astype(int)
```

Out[183]:

Sales 6596953

dtype: int64

In [184]:

#24) How much total sales happened in Asian Holiday Promotion?

```

In [189]:
data[data.PromotionName == 'Asian Holiday
Promotion'][['Sales']].sum().astype(int)
Out[189]:
Sales    3615193
dtype: int64
In [190]:
#25) How much total profit value by Contoso, Ltd Manufacturer?
In [191]:
data[data.Manufacturer == 'Contoso, Ltd
'][['Profit']].sum().astype(int)
Out[191]:
Profit    0
dtype: int64
In [192]:
#26) Query the total sales and total order qty by product category
and order it by OrderQty from highest to lowest?
In [193]:
data.groupby(['ProductCategory']).agg({'Sales': 'sum', 'OrderQty': 'sum'})
) \
    .rename(columns =
{"Sales": "TotalSales", 'OrderQty': 'TotalOrderQty'}) \
    .sort_values("TotalOrderQty", ascending = False) \
    .astype(int)
Out[193]:

```

	TotalSales	TotalOrderQty
ProductCategory		
Cell phones	5841155	96074
Computers	21306971	69425
Cameras and camcorders	17048931	43893
TV and Video	9107839	22714
Music, Movies and Audio Books	1057480	9953
Audio	1029381	9067

```

In [219]:
#27) Query the total sales value in between 1057 and 26700 by Region?
In [224]:
data[data['Sales'].between(1057, 26700)] \

```

```
.groupby('Region')['Sales']\n    .sum()
```

Out[224]:

Region

```
Asia           1.076776e+07\nEurope        9.331408e+06\nNorth America 2.929052e+07\nName: Sales, dtype: float64
```

In [225]:

#28) How many orders placed in Pittsfield city?

In [11]:

```
data.query("City == 'Pittsfield'")['OrderID'].count()
```

Out[11]:

```
OrderID      31\ndtype: int64
```

In [12]:

#29) Query the total sales and total profit by region, country and city?

In [15]:

```
data.groupby(['Region', 'Country', 'City']).agg({'Sales':'sum',\n    'Profit':'sum'})\n    .rename(columns =\n        {"Sales":"TotalSales", "Profit":"TotalProfit"}).astype(int)
```

Out[15]:

			TotalSales	TotalProfit
Region	Country	City		
Asia	Armenia	Yerevan	172934	111970
	Australia	Canberra	211239	93840
		Sydney	421402	250393
	Bhutan	Thimphu	161291	88175
	China	Beijing	6596953	3810006
...
North America	United States	Waukesha	195396	95561
		Wheat Ridge	129569	57945
		Winchester	94264	52615
		Worcester	183950	108701

			TotalSales	TotalProfit
Region	Country	City		
		Yakima	100843	70798

263 rows × 2 columns

In [16]:

#30) Query the total unit cost and total price by productcategory, productssubcategory and Manufacturer?

In [18]:

```
data.groupby(['ProductCategory', 'ProductSubCategory', 'Manufacturer']).agg({"UnitCost": "sum", "Price": "sum"}). \
    .rename(columns = {"UnitCost": "TotalUnitCost", "Price": "TotalPrice"}).astype(int)
```

Out[18]:

			TotalUnitCost	TotalPrice
ProductCategory	ProductSubCategory	Manufacturer		
Audio	Bluetooth Headphones	Northwind Traders	4694	10698
		Wide World Importers	5348	12740
	MP4&MP3	Contoso, Ltd	15180	37464
	Recording Pen	Wide World Importers	10165	22585
Cameras and camcorders	Camcorders	Fabrikam, Inc.	286756	705377
	Cameras & Camcorders Accessories	Contoso, Ltd	12221	26301
	Digital Cameras	A. Datum Corporation	88363	210914
	Digital SLR Cameras	A. Datum Corporation	55606	148662
		Contoso, Ltd	61069	155672
		Fabrikam, Inc.	64966	174022
Cell phones	Cell phones Accessories	Contoso, Ltd	1526	3530
	Home & Office Phones	Contoso, Ltd	12044	27389

			TotalUnitCost	TotalPrice
ProductCategory	ProductSubCategory	Manufacturer		
Computers	Smart phones & PDAs	The Phone Company	103969	239147
	Touch Screen Phones	Contoso, Ltd	12851	27104
		The Phone Company	55433	130579
	Computers Accessories	Contoso, Ltd	16153	35657
		Southridge Video	12789	24305
	Desktops	Adventure Works	43576	89568
		Wide World Importers	31821	64659
	Laptops	Adventure Works	93609	224875
		Fabrikam, Inc.	52102	127405
		Proseware, Inc.	26831	59963
		Wide World Importers	45528	119908
Music, Movies and Audio Books	Monitors	Adventure Works	23132	54600
		Proseware, Inc.	17530	42299
		Wide World Importers	22386	58944
	Printers, Scanners & Fax	Proseware, Inc.	65123	155502
	Projectors & Screens	Contoso, Ltd	78241	198399
		Proseware, Inc.	99812	259541
		Wide World Importers	62842	146677
	Movie DVD	Contoso, Ltd	16690	44897
		Southridge Video	17559	42566

			TotalUnitCost	TotalPrice
ProductCategory	ProductSubCategory	Manufacturer		
TV and Video	Car Video	Southridge Video	84685	179049
	Home Theater System	Contoso, Ltd	97944	205953
		Litware, Inc.	87195	187306
	Televisions	Adventure Works	71821	174949
	VCD & DVD	Southridge Video	8797	18504

In []: