Importing the necessary libraries

In [12]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import os
```

Merging all the csv files into one

In [34]:
```
df =pd.read_csv('Sales_Data/Sales_April_2019.csv')
df.head()
```

Out[34]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001
1	NaN	NaN	NaN	NaN	NaN	NaN
2	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215
3	176560	Google Phone	1	600	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001
4	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001

In [35]:
```
# Listing down number of files in Sales data folder
files=[file for file in os.listdir('./Sales_Data')]

for file in files:
    print(file)
```

```
Sales_April_2019.csv
Sales_August_2019.csv
Sales_December_2019.csv
Sales_February_2019.csv
Sales_January_2019.csv
Sales_July_2019.csv
Sales_June_2019.csv
Sales_March_2019.csv
Sales_May_2019.csv
Sales_November_2019.csv
Sales_October_2019.csv
Sales_September_2019.csv
```

In [36]:
```
#Creating the empty data frame
all_months_data=pd.DataFrame()

# Reading all the files
files=[file for file in os.listdir('./Sales_Data')]

# Concanating data from each file to all_months_data
for file in files:
    df=pd.read_csv("./Sales_Data/"+file)
    all_months_data=pd.concat([all_months_data, df])

# Exporting to csv file
all_months_data.to_csv("all_data.csv", index=False)
```

Data Cleaning

In [38]:
```
# Find out the null values in data
nan_df=all_months_data[all_months_data.isna().any(axis=1)]
nan_df
```

Out[38]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address
1	NaN	NaN	NaN	NaN	NaN	NaN
356	NaN	NaN	NaN	NaN	NaN	NaN
735	NaN	NaN	NaN	NaN	NaN	NaN
1433	NaN	NaN	NaN	NaN	NaN	NaN
1553	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...
10012	NaN	NaN	NaN	NaN	NaN	NaN
10274	NaN	NaN	NaN	NaN	NaN	NaN
10878	NaN	NaN	NaN	NaN	NaN	NaN
11384	NaN	NaN	NaN	NaN	NaN	NaN
11662	NaN	NaN	NaN	NaN	NaN	NaN

545 rows × 6 columns

In [39]:
```
# Drop Null values: As all the values in row are null
all_months_data.dropna(inplace= True)
```

In [40]:
```
# Drop repeated headers
index_name=all_months_data.loc[all_months_data['Order ID']=='Order ID'].index
index_name
```

Out[40]:
```
Int64Index([  519,  1149,  1155,  2878,  2893,  3036,  3209,  3618,  4138,
             4645,
            ...
             8644,  9325,  9502,  9615,  9954, 10000, 10387, 11399, 11468,
            11574],
           dtype='int64', length=355)
```

In [41]:
```
all_months_data.drop(index_name, inplace= True)
```

Converting the data to correct data type for computing

In [46]:
```
all_months_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 182735 entries, 0 to 11685
Data columns (total 6 columns):
 #   Column            Non-Null Count   Dtype 
---  ------            --------------   ----- 
 0   Order ID          182735 non-null  object
 1   Product           182735 non-null  object
 2   Quantity Ordered  182735 non-null  object
 3   Price Each        182735 non-null  object
 4   Order Date        182735 non-null  object
 5   Purchase Address  182735 non-null  object
dtypes: object(6)
memory usage: 9.8+ MB
```

In [47]:
```
all_months_data['Quantity Ordered']= pd.to_numeric(all_months_data['Quantity O
rdered'])
all_months_data['Price Each']= pd.to_numeric(all_months_data['Price Each'])
```

In [48]:
```
# Exporting cleaned data into csv files
all_months_data.to_csv("all_data01.csv", index=False)
```

In [76]:
```
# Creating new data frame with cleaned data
all_data= pd.read_csv("all_data01.csv")
all_data.shape
```

Out[76]:
```
(182735, 6)
```

In [77]:
```
all_data.head()
```

Out[77]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001

Agument data with additional column

Task 1: Add additional column with month data

In [78]:
```
all_data['Month']=all_data['Order Date'].str[0:2]
all_data['Month']=all_data['Month'].astype('int32')
all_data.head()
```

Out[78]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4

Task 2: Add additional column for sales

In [79]:
```
all_data['Sales']= all_data['Quantity Ordered'] * all_data['Price Each']
all_data.head()
```

Out[79]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4	23.90
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4	99.99
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	600.00
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	11.99
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4	11.99

Task 3: Add column for city

In [80]:
```
#all_data.drop(columns='City')
def getCity(address):
    return address.split(',')[1]
all_data['City']=all_data['Purchase Address'].apply(lambda x: getCity(x))

all_data.head()
```

Out[80]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4	23.90	Dallas
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4	99.99	Boston
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles

Task 4: Add column for state as city can have duplicates (Same city name in different states)

In [81]:
```
#all_data.drop(columns='State')
def getState(address):
    return address.split(',')[2].split(' ')[1]

all_data['State']=all_data['Purchase Address'].apply(lambda x: getState(x))
all_data.head()
```

Out[81]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4	23.90	Dallas	TX
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4	99.99	Boston	MA
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles	CA
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles	CA
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles	CA

In [84]:
```
# Combining state and city column

all_data['City']= all_data['Purchase Address'].apply(lambda x: f"{getCity(x)}
 ({getState(x)})")
all_data.head()
```

Out[84]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4	23.90	Dallas (TX)	TX
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4	99.99	Boston (MA)	MA
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA

Q1: What is the best month for sales? How much was earned that month?

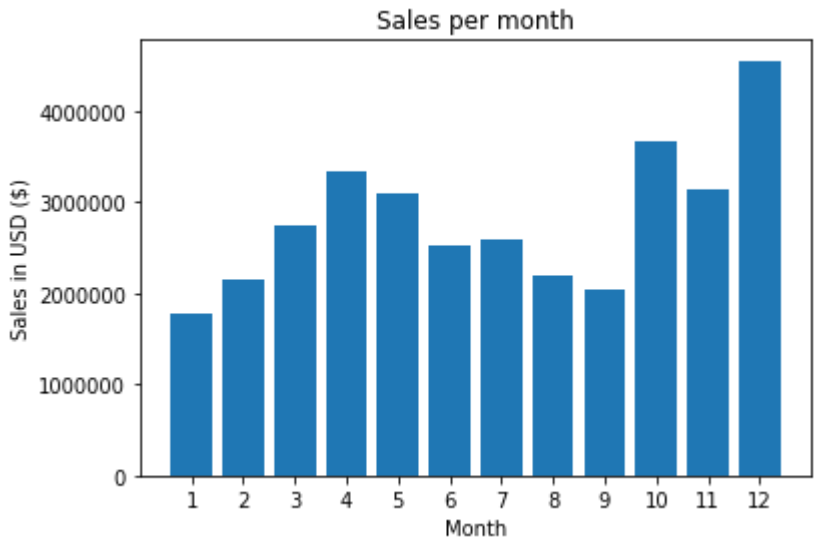
In [54]:
```
result=all_data.groupby('Month').sum()
result
```

Out[54]:

Month	Order ID	Quantity Ordered	Price Each	Sales
1	1391148846	10670	1.776156e+06	1.786511e+06
2	1832181724	13162	2.145209e+06	2.158127e+06
3	2520234917	16697	2.740393e+06	2.755969e+06
4	3333217539	20226	3.313522e+06	3.336376e+06
5	3289290354	18344	3.084756e+06	3.101881e+06
6	2876409942	14964	2.508863e+06	2.524465e+06
7	3224356195	15781	2.572765e+06	2.587445e+06
8	2840777365	13185	2.178184e+06	2.191698e+06
9	2886380303	12827	2.038020e+06	2.050361e+06
10	5374745444	22356	3.658885e+06	3.679254e+06
11	4964034464	19479	3.130969e+06	3.149785e+06
12	7593831188	27764	4.533152e+06	4.557905e+06

In [55]:
```
months=range(1, 13)

plt.bar(months, result['Sales'])
plt.xticks(months)
plt.ylabel('Sales in USD ($)')
plt.xlabel('Month')
plt.title('Sales per month')
plt.show()
```



Q2: What city had the highest number of sales?

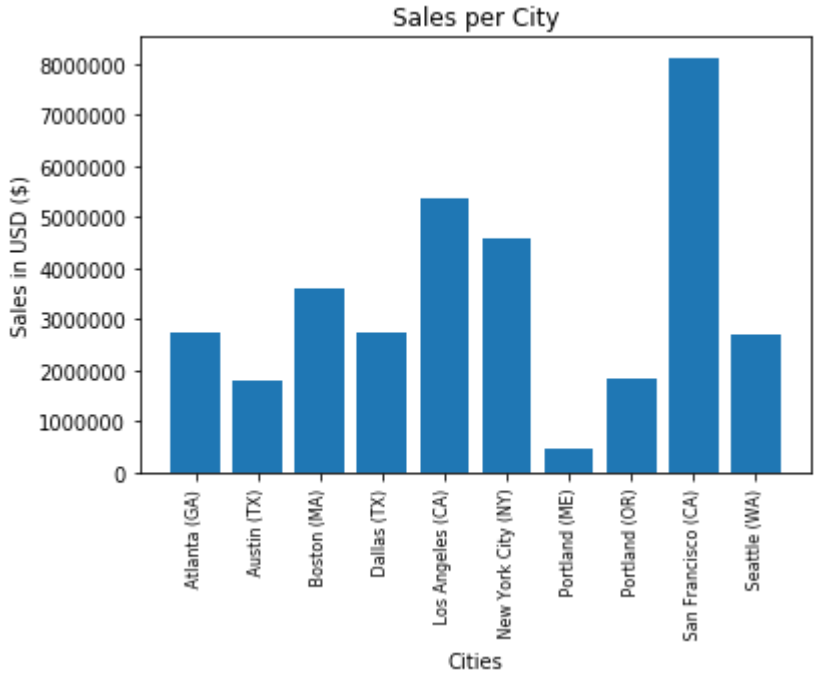
In [85]:
```
result=all_data.groupby('City').sum()
result
```

Out[85]:

City	Order ID	Quantity Ordered	Price Each	Month	Sales
Atlanta (GA)	3366679952	16304	2.726247e+06	103160	2.741642e+06
Austin (TX)	2238557673	10933	1.777232e+06	68570	1.786746e+06
Boston (MA)	4517043947	22123	3.580075e+06	138632	3.604081e+06
Dallas (TX)	3360280471	16453	2.702619e+06	102976	2.717794e+06
Los Angeles (CA)	6698824762	32722	5.323916e+06	204933	5.354040e+06
New York City (NY)	5643048641	27470	4.553287e+06	172980	4.581659e+06
Portland (ME)	554542995	2696	4.417017e+05	16889	4.441105e+05
Portland (OR)	2269245678	11110	1.822878e+06	69540	1.832539e+06
San Francisco (CA)	10127798551	49363	8.074046e+06	310174	8.124121e+06
Seattle (WA)	3350585611	16281	2.678871e+06	103213	2.693049e+06

In [92]:
```
cities=[city for city, df in all_data.groupby('City')]

plt.bar(cities, result['Sales'])
plt.xticks(cities, rotation='vertical', size=8)
plt.ylabel('Sales in USD ($)')
plt.xlabel('Cities')
plt.title('Sales per City')
plt.show()
```



Question 3: At what time should we display advertisements to likelihood of customer's buying product?

In [101]:
```
all_data.head()
```

Out[101]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
0	176558	USB-C Charging Cable	2	11.95	04/19/19 08:46	917 1st St, Dallas, TX 75001	4	23.90	Dallas (TX)	TX
1	176559	Bose SoundSport Headphones	1	99.99	04/07/19 22:30	682 Chestnut St, Boston, MA 02215	4	99.99	Boston (MA)	MA
2	176560	Google Phone	1	600.00	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA
3	176560	Wired Headphones	1	11.99	04/12/19 14:38	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA
4	176561	Wired Headphones	1	11.99	04/30/19 09:27	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA

In [103]:
```
# Conveting Order date string into python date time format for easy calculation and formatting
all_data['Order Date']= pd.to_datetime(all_data['Order Date'])
all_data.head()
```

Out[103]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
0	176558	USB-C Charging Cable	2	11.95	2019-04-19 08:46:00	917 1st St, Dallas, TX 75001	4	23.90	Dallas (TX)	TX
1	176559	Bose SoundSport Headphones	1	99.99	2019-04-07 22:30:00	682 Chestnut St, Boston, MA 02215	4	99.99	Boston (MA)	MA
2	176560	Google Phone	1	600.00	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA
3	176560	Wired Headphones	1	11.99	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA
4	176561	Wired Headphones	1	11.99	2019-04-30 09:27:00	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA

In [104]:
```
all_data['Hour'] =all_data['Order Date'].dt.hour
all_data['Min'] =all_data['Order Date'].dt.minute
all_data.head()
```

Out[104]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State	Ho
0	176558	USB-C Charging Cable	2	11.95	2019-04-19 08:46:00	917 1st St, Dallas, TX 75001	4	23.90	Dallas (TX)	TX	
1	176559	Bose SoundSport Headphones	1	99.99	2019-04-07 22:30:00	682 Chestnut St, Boston, MA 02215	4	99.99	Boston (MA)	MA	
2	176560	Google Phone	1	600.00	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA	
3	176560	Wired Headphones	1	11.99	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA	
4	176561	Wired Headphones	1	11.99	2019-04-30 09:27:00	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA	

In [105]:
```
result=all_data.groupby('Hour').count()
result
```

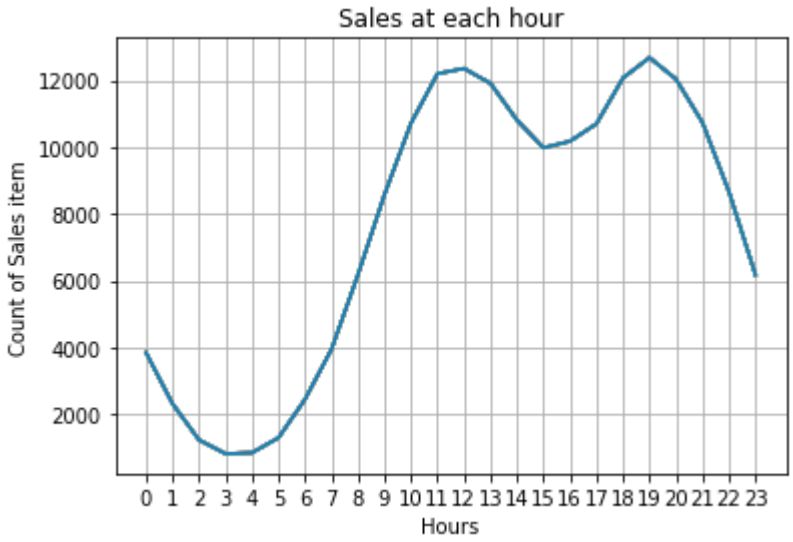
Out[105]:

Hour	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State	Min
0	3844	3844	3844	3844	3844	3844	3844	3844	3844	3844	3844
1	2313	2313	2313	2313	2313	2313	2313	2313	2313	2313	2313
2	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227	1227
3	816	816	816	816	816	816	816	816	816	816	816
4	843	843	843	843	843	843	843	843	843	843	843
5	1293	1293	1293	1293	1293	1293	1293	1293	1293	1293	1293
6	2445	2445	2445	2445	2445	2445	2445	2445	2445	2445	2445
7	3945	3945	3945	3945	3945	3945	3945	3945	3945	3945	3945
8	6184	6184	6184	6184	6184	6184	6184	6184	6184	6184	6184
9	8583	8583	8583	8583	8583	8583	8583	8583	8583	8583	8583
10	10721	10721	10721	10721	10721	10721	10721	10721	10721	10721	10721
11	12202	12202	12202	12202	12202	12202	12202	12202	12202	12202	12202
12	12360	12360	12360	12360	12360	12360	12360	12360	12360	12360	12360
13	11913	11913	11913	11913	11913	11913	11913	11913	11913	11913	11913
14	10810	10810	10810	10810	10810	10810	10810	10810	10810	10810	10810
15	9981	9981	9981	9981	9981	9981	9981	9981	9981	9981	9981
16	10179	10179	10179	10179	10179	10179	10179	10179	10179	10179	10179
17	10702	10702	10702	10702	10702	10702	10702	10702	10702	10702	10702
18	12074	12074	12074	12074	12074	12074	12074	12074	12074	12074	12074
19	12685	12685	12685	12685	12685	12685	12685	12685	12685	12685	12685
20	12040	12040	12040	12040	12040	12040	12040	12040	12040	12040	12040
21	10738	10738	10738	10738	10738	10738	10738	10738	10738	10738	10738
22	8663	8663	8663	8663	8663	8663	8663	8663	8663	8663	8663
23	6174	6174	6174	6174	6174	6174	6174	6174	6174	6174	6174

In [110]:
```
hours=[hour for hour, df in all_data.groupby('Hour')]

plt.plot(hours, result)
plt.xticks(hours)
plt.grid()
plt.xlabel("Hours")
plt.ylabel("Count of Sales item")
plt.title("Sales at each hour")
plt.show()

# Recommondation is to display add  around 11am (11) or 7pm (19)
```



Q4: What products are most often sold together?

In [111]:
```
all_data.head()
# Need to see all the duplicate order Id to analyse this problem.
```

Out[111]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State	Ho
0	176558	USB-C Charging Cable	2	11.95	2019-04-19 08:46:00	917 1st St, Dallas, TX 75001	4	23.90	Dallas (TX)	TX	
1	176559	Bose SoundSport Headphones	1	99.99	2019-04-07 22:30:00	682 Chestnut St, Boston, MA 02215	4	99.99	Boston (MA)	MA	
2	176560	Google Phone	1	600.00	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA	
3	176560	Wired Headphones	1	11.99	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA	
4	176561	Wired Headphones	1	11.99	2019-04-30 09:27:00	333 8th St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA	

In [115]:
```
# Creating new data frame with duplicated order Id
df =all_data[all_data['Order ID'].duplicated(keep=False)]
df.head(20)
```

Out[115]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
2	176560	Google Phone	1	600.00	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA
3	176560	Wired Headphones	1	11.99	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA
17	176574	Google Phone	1	600.00	2019-04-03 19:42:00	20 Hill St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA
18	176574	USB-C Charging Cable	1	11.95	2019-04-03 19:42:00	20 Hill St, Los Angeles, CA 90001	4	11.95	Los Angeles (CA)	CA
29	176585	Bose SoundSport Headphones	1	99.99	2019-04-07 11:31:00	823 Highland St, Boston, MA 02215	4	99.99	Boston (MA)	MA
30	176585	Bose SoundSport Headphones	1	99.99	2019-04-07 11:31:00	823 Highland St, Boston, MA 02215	4	99.99	Boston (MA)	MA
31	176586	AAA Batteries (4-pack)	2	2.99	2019-04-10 17:00:00	365 Center St, San Francisco, CA 94016	4	5.98	San Francisco (CA)	CA
32	176586	Google Phone	1	600.00	2019-04-10 17:00:00	365 Center St, San Francisco, CA 94016	4	600.00	San Francisco (CA)	CA
118	176672	Lightning Charging Cable	1	14.95	2019-04-12 11:07:00	778 Maple St, New York City, NY 10001	4	14.95	New York City (NY)	NY
119	176672	USB-C Charging Cable	1	11.95	2019-04-12 11:07:00	778 Maple St, New York City, NY 10001	4	11.95	New York City (NY)	NY
128	176681	Apple Airpods Headphones	1	150.00	2019-04-20 10:39:00	331 Cherry St, Seattle, WA 98101	4	150.00	Seattle (WA)	WA

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State
129	176681	ThinkPad Laptop	1	999.99	2019-04-20 10:39:00	331 Cherry St, Seattle, WA 98101	4	999.99	Seattle (WA)	WA
137	176689	Bose SoundSport Headphones	1	99.99	2019-04-24 17:15:00	659 Lincoln St, New York City, NY 10001	4	99.99	New York City (NY)	NY
138	176689	AAA Batteries (4-pack)	2	2.99	2019-04-24 17:15:00	659 Lincoln St, New York City, NY 10001	4	5.98	New York City (NY)	NY
187	176739	34in Ultrawide Monitor	1	379.99	2019-04-05 17:38:00	730 6th St, Austin, TX 73301	4	379.99	Austin (TX)	TX
188	176739	Google Phone	1	600.00	2019-04-05 17:38:00	730 6th St, Austin, TX 73301	4	600.00	Austin (TX)	TX
222	176774	Lightning Charging Cable	1	14.95	2019-04-25 15:06:00	372 Church St, Los Angeles, CA 90001	4	14.95	Los Angeles (CA)	CA
223	176774	USB-C Charging Cable	1	11.95	2019-04-25 15:06:00	372 Church St, Los Angeles, CA 90001	4	11.95	Los Angeles (CA)	CA
230	176781	iPhone	1	700.00	2019-04-03 07:37:00	976 Hickory St, Dallas, TX 75001	4	700.00	Dallas (TX)	TX
231	176781	Lightning Charging Cable	1	14.95	2019-04-03 07:37:00	976 Hickory St, Dallas, TX 75001	4	14.95	Dallas (TX)	TX

In [127]:
```
df['Grouped']=df.groupby('Order ID')['Product'].transform(lambda x: ','.join(x
))
df.head()
```

```
C:\ProgramData\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  """Entry point for launching an IPython kernel.
```

Out[127]:

	Order ID	Product	Quantity Ordered	Price Each	Order Date	Purchase Address	Month	Sales	City	State	H
2	176560	Google Phone	1	600.00	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA	
3	176560	Wired Headphones	1	11.99	2019-04-12 14:38:00	669 Spruce St, Los Angeles, CA 90001	4	11.99	Los Angeles (CA)	CA	
17	176574	Google Phone	1	600.00	2019-04-03 19:42:00	20 Hill St, Los Angeles, CA 90001	4	600.00	Los Angeles (CA)	CA	
18	176574	USB-C Charging Cable	1	11.95	2019-04-03 19:42:00	20 Hill St, Los Angeles, CA 90001	4	11.95	Los Angeles (CA)	CA	
29	176585	Bose SoundSport Headphones	1	99.99	2019-04-07 11:31:00	823 Highland St, Boston, MA 02215	4	99.99	Boston (MA)	MA	

In [129]:
```
# Removing duplicate Order Id
df=df[['Order ID', 'Grouped']].drop_duplicates()
df.head(100)
```

Out[129]:

	Order ID	Grouped
2	176560	Google Phone,Wired Headphones
17	176574	Google Phone,USB-C Charging Cable
29	176585	Bose SoundSport Headphones,Bose SoundSport Hea...
31	176586	AAA Batteries (4-pack),Google Phone
118	176672	Lightning Charging Cable,USB-C Charging Cable
...	...	...
2802	179311	Wired Headphones,USB-C Charging Cable
2819	179328	Wired Headphones,AA Batteries (4-pack)
2832	179340	Macbook Pro Laptop,Apple Airpods Headphones
2850	179357	Google Phone,USB-C Charging Cable
2871	179377	USB-C Charging Cable,AA Batteries (4-pack)

100 rows × 2 columns

In [131]:
```
from itertools import combinations
from collections import Counter

count=Counter()

for row in df['Grouped']:
    row_list=row.split(',')
    count.update(Counter(combinations(row_list, 2)))

print(count)
```

```
Counter({('iPhone', 'Lightning Charging Cable'): 973, ('Google Phone', 'USB-C
Charging Cable'): 954, ('iPhone', 'Wired Headphones'): 437, ('Google Phone',
'Wired Headphones'): 401, ('Vareebadd Phone', 'USB-C Charging Cable'): 348,
('iPhone', 'Apple Airpods Headphones'): 343, ('Google Phone', 'Bose SoundSpor
t Headphones'): 219, ('USB-C Charging Cable', 'Wired Headphones'): 156, ('Var
eebadd Phone', 'Wired Headphones'): 141, ('Lightning Charging Cable', 'Wired
Headphones'): 92, ('USB-C Charging Cable', 'Bose SoundSport Headphones'): 76,
('Lightning Charging Cable', 'Apple Airpods Headphones'): 76, ('Vareebadd Pho
ne', 'Bose SoundSport Headphones'): 74, ('Apple Airpods Headphones', 'Wired H
eadphones'): 69, ('Lightning Charging Cable', 'USB-C Charging Cable'): 58,
('Lightning Charging Cable', 'AA Batteries (4-pack)'): 53, ('Lightning Chargi
ng Cable', 'Lightning Charging Cable'): 53, ('Bose SoundSport Headphones', 'W
ired Headphones'): 51, ('AA Batteries (4-pack)', 'Lightning Charging Cable'):
49, ('AAA Batteries (4-pack)', 'USB-C Charging Cable'): 48, ('AA Batteries (4
-pack)', 'AAA Batteries (4-pack)'): 47, ('AAA Batteries (4-pack)', 'AAA Batte
ries (4-pack)'): 46, ('USB-C Charging Cable', 'USB-C Charging Cable'): 46,
('USB-C Charging Cable', 'AAA Batteries (4-pack)'): 44, ('Apple Airpods Headp
hones', 'AAA Batteries (4-pack)'): 44, ('AAA Batteries (4-pack)', 'Lightning
Charging Cable'): 44, ('AAA Batteries (4-pack)', 'Wired Headphones'): 42, ('W
ired Headphones', 'USB-C Charging Cable'): 42, ('AA Batteries (4-pack)', 'Wir
ed Headphones'): 42, ('Wired Headphones', 'AAA Batteries (4-pack)'): 42, ('US
B-C Charging Cable', 'Lightning Charging Cable'): 41, ('AA Batteries (4-pac
k)', 'Apple Airpods Headphones'): 39, ('AAA Batteries (4-pack)', 'AA Batterie
s (4-pack)'): 39, ('Wired Headphones', 'AA Batteries (4-pack)'): 39, ('Bose S
oundSport Headphones', 'AAA Batteries (4-pack)'): 37, ('AA Batteries (4-pac
k)', 'USB-C Charging Cable'): 37, ('Lightning Charging Cable', 'Bose SoundSpo
rt Headphones'): 37, ('USB-C Charging Cable', 'AA Batteries (4-pack)'): 36,
('Wired Headphones', 'Lightning Charging Cable'): 36, ('Apple Airpods Headpho
nes', 'Lightning Charging Cable'): 35, ('Lightning Charging Cable', 'AAA Batt
eries (4-pack)'): 35, ('Wired Headphones', 'Wired Headphones'): 35, ('AA Batt
eries (4-pack)', 'AA Batteries (4-pack)'): 35, ('USB-C Charging Cable', 'Appl
e Airpods Headphones'): 34, ('AAA Batteries (4-pack)', 'Apple Airpods Headpho
nes'): 33, ('Apple Airpods Headphones', 'Bose SoundSport Headphones'): 32,
('USB-C Charging Cable', '27in FHD Monitor'): 31, ('Bose SoundSport Headphone
s', 'Lightning Charging Cable'): 31, ('Wired Headphones', 'Apple Airpods Head
phones'): 29, ('Apple Airpods Headphones', 'AA Batteries (4-pack)'): 29, ('Ap
ple Airpods Headphones', 'USB-C Charging Cable'): 28, ('Bose SoundSport Headp
hones', 'Bose SoundSport Headphones'): 27, ('Bose SoundSport Headphones', 'AA
Batteries (4-pack)'): 26, ('Bose SoundSport Headphones', 'USB-C Charging Cabl
e'): 25, ('AA Batteries (4-pack)', 'Bose SoundSport Headphones'): 25, ('Apple
Airpods Headphones', 'Apple Airpods Headphones'): 23, ('Wired Headphones', 'B
ose SoundSport Headphones'): 21, ('AAA Batteries (4-pack)', '27in FHD Monito
r'): 20, ('27in FHD Monitor', 'AAA Batteries (4-pack)'): 20, ('34in Ultrawide
Monitor', 'AA Batteries (4-pack)'): 19, ('AAA Batteries (4-pack)', 'Bose Soun
dSport Headphones'): 18, ('27in FHD Monitor', 'Lightning Charging Cable'): 1
8, ('Lightning Charging Cable', '27in FHD Monitor'): 18, ('Wired Headphones',
'27in 4K Gaming Monitor'): 18, ('Bose SoundSport Headphones', 'Apple Airpods
Headphones'): 18, ('AA Batteries (4-pack)', 'iPhone'): 17, ('Wired Headphone
s', '34in Ultrawide Monitor'): 17, ('34in Ultrawide Monitor', 'Lightning Char
ging Cable'): 17, ('ThinkPad Laptop', 'AAA Batteries (4-pack)'): 16, ('iPhon
e', 'AAA Batteries (4-pack)'): 16, ('Lightning Charging Cable', '27in 4K Gami
ng Monitor'): 16, ('Lightning Charging Cable', 'Google Phone'): 16, ('27in 4K
Gaming Monitor', 'Lightning Charging Cable'): 16, ('34in Ultrawide Monitor',
'USB-C Charging Cable'): 15, ('27in FHD Monitor', 'AA Batteries (4-pack)'): 1
5, ('Wired Headphones', 'iPhone'): 15, ('AAA Batteries (4-pack)', '27in 4K Ga
ming Monitor'): 15, ('iPhone', 'USB-C Charging Cable'): 15, ('20in Monitor',
'USB-C Charging Cable'): 15, ('Lightning Charging Cable', '20in Monitor'): 1
```

5, ('Lightning Charging Cable', '34in Ultrawide Monitor'): 15, ('Apple Airpod
s Headphones', 'Google Phone'): 14, ('USB-C Charging Cable', 'iPhone'): 14,
('Bose SoundSport Headphones', '27in FHD Monitor'): 14, ('27in 4K Gaming Moni
tor', 'AAA Batteries (4-pack)'): 14, ('iPhone', 'AA Batteries (4-pack)'): 14,
('Google Phone', 'AA Batteries (4-pack)'): 13, ('AA Batteries (4-pack)', '34i
n Ultrawide Monitor'): 13, ('AA Batteries (4-pack)', '27in 4K Gaming Monito
r'): 13, ('AA Batteries (4-pack)', 'Flatscreen TV'): 12, ('Apple Airpods Head
phones', '27in FHD Monitor'): 12, ('27in FHD Monitor', 'Bose SoundSport Headp
hones'): 12, ('27in FHD Monitor', 'USB-C Charging Cable'): 12, ('Google Phon
e', 'Lightning Charging Cable'): 12, ('AAA Batteries (4-pack)', '34in Ultrawi
de Monitor'): 12, ('AAA Batteries (4-pack)', 'iPhone'): 12, ('Apple Airpods H
eadphones', 'iPhone'): 12, ('Wired Headphones', '27in FHD Monitor'): 12, ('20
in Monitor', 'Wired Headphones'): 12, ('27in FHD Monitor', 'Apple Airpods Hea
dphones'): 12, ('USB-C Charging Cable', 'Google Phone'): 12, ('27in 4K Gaming
Monitor', 'AA Batteries (4-pack)'): 12, ('34in Ultrawide Monitor', 'AAA Batte
ries (4-pack)'): 12, ('Wired Headphones', 'Macbook Pro Laptop'): 12, ('AAA Ba
tteries (4-pack)', 'Google Phone'): 11, ('USB-C Charging Cable', 'ThinkPad La
ptop'): 11, ('34in Ultrawide Monitor', 'Wired Headphones'): 11, ('Apple Airpo
ds Headphones', '27in 4K Gaming Monitor'): 11, ('AA Batteries (4-pack)', '27i
n FHD Monitor'): 11, ('Google Phone', 'AAA Batteries (4-pack)'): 11, ('USB-C
Charging Cable', '34in Ultrawide Monitor'): 11, ('27in 4K Gaming Monitor', 'U
SB-C Charging Cable'): 11, ('Macbook Pro Laptop', 'Lightning Charging Cabl
e'): 11, ('AA Batteries (4-pack)', 'Google Phone'): 11, ('AAA Batteries (4-pa
ck)', 'ThinkPad Laptop'): 11, ('Lightning Charging Cable', 'Flatscreen TV'):
11, ('Macbook Pro Laptop', 'Bose SoundSport Headphones'): 11, ('AAA Batteries
(4-pack)', 'Macbook Pro Laptop'): 10, ('Wired Headphones', 'ThinkPad Lapto
p'): 10, ('USB-C Charging Cable', '20in Monitor'): 10, ('USB-C Charging Cabl
e', '27in 4K Gaming Monitor'): 10, ('27in 4K Gaming Monitor', 'Apple Airpods
Headphones'): 10, ('Bose SoundSport Headphones', '34in Ultrawide Monitor'): 1
0, ('USB-C Charging Cable', 'Flatscreen TV'): 10, ('27in FHD Monitor', 'Wired
Headphones'): 10, ('Apple Airpods Headphones', 'Macbook Pro Laptop'): 10, ('A
A Batteries (4-pack)', '20in Monitor'): 10, ('AAA Batteries (4-pack)', 'Flats
creen TV'): 10, ('Lightning Charging Cable', 'iPhone'): 10, ('Macbook Pro Lap
top', 'USB-C Charging Cable'): 10, ('Lightning Charging Cable', 'Macbook Pro
Laptop'): 10, ('Flatscreen TV', 'AAA Batteries (4-pack)'): 10, ('Bose SoundSp
ort Headphones', '27in 4K Gaming Monitor'): 10, ('Apple Airpods Headphones',
'ThinkPad Laptop'): 9, ('Wired Headphones', 'Google Phone'): 9, ('Flatscreen
TV', 'Lightning Charging Cable'): 9, ('20in Monitor', 'Bose SoundSport Headph
ones'): 9, ('20in Monitor', 'Lightning Charging Cable'): 9, ('Macbook Pro Lap
top', 'AA Batteries (4-pack)'): 9, ('ThinkPad Laptop', 'USB-C Charging Cabl
e'): 9, ('ThinkPad Laptop', 'Lightning Charging Cable'): 9, ('27in FHD Monito
r', '27in FHD Monitor'): 9, ('AA Batteries (4-pack)', 'ThinkPad Laptop'): 9,
('27in 4K Gaming Monitor', 'Wired Headphones'): 9, ('Lightning Charging Cabl
e', 'ThinkPad Laptop'): 9, ('Bose SoundSport Headphones', 'Flatscreen TV'):
9, ('AA Batteries (4-pack)', 'Macbook Pro Laptop'): 8, ('Flatscreen TV', 'AA
Batteries (4-pack)'): 8, ('27in 4K Gaming Monitor', 'Bose SoundSport Headphon
es'): 8, ('ThinkPad Laptop', 'Bose SoundSport Headphones'): 8, ('Vareebadd Ph
one', 'AA Batteries (4-pack)'): 8, ('Apple Airpods Headphones', 'Flatscreen T
V'): 8, ('ThinkPad Laptop', 'AA Batteries (4-pack)'): 8, ('USB-C Charging Cab
le', 'Macbook Pro Laptop'): 8, ('34in Ultrawide Monitor', 'Apple Airpods Head
phones'): 8, ('Bose SoundSport Headphones', 'Google Phone'): 8, ('Macbook Pro
Laptop', 'Wired Headphones'): 7, ('Google Phone', '27in FHD Monitor'): 7, ('M
acbook Pro Laptop', 'AAA Batteries (4-pack)'): 7, ('34in Ultrawide Monitor',
'iPhone'): 7, ('34in Ultrawide Monitor', '34in Ultrawide Monitor'): 7, ('Flat
screen TV', 'USB-C Charging Cable'): 7, ('AAA Batteries (4-pack)', '20in Moni
tor'): 7, ('Bose SoundSport Headphones', 'iPhone'): 7, ('ThinkPad Laptop', 'A
pple Airpods Headphones'): 7, ('Google Phone', 'Apple Airpods Headphones'):

```
7, ('Macbook Pro Laptop', '27in 4K Gaming Monitor'): 7, ('iPhone', '27in 4K G
aming Monitor'): 6, ('20in Monitor', 'Apple Airpods Headphones'): 6, ('Macboo
k Pro Laptop', 'Apple Airpods Headphones'): 6, ('Flatscreen TV', 'Flatscreen
TV'): 6, ('iPhone', '34in Ultrawide Monitor'): 6, ('Wired Headphones', '20in
Monitor'): 6, ('Vareebadd Phone', 'Apple Airpods Headphones'): 6, ('USB-C Cha
rging Cable', 'Vareebadd Phone'): 6, ('27in 4K Gaming Monitor', '34in Ultrawi
de Monitor'): 6, ('Bose SoundSport Headphones', '20in Monitor'): 6, ('iPhon
e', 'Flatscreen TV'): 6, ('Apple Airpods Headphones', '20in Monitor'): 6, ('W
ired Headphones', 'Vareebadd Phone'): 6, ('34in Ultrawide Monitor', 'Bose Sou
ndSport Headphones'): 6, ('Google Phone', 'iPhone'): 6, ('Wired Headphones',
'Flatscreen TV'): 6, ('Apple Airpods Headphones', '34in Ultrawide Monitor'):
5, ('iPhone', 'Bose SoundSport Headphones'): 5, ('27in 4K Gaming Monitor', 'M
acbook Pro Laptop'): 5, ('27in 4K Gaming Monitor', '27in 4K Gaming Monitor'):
5, ('Flatscreen TV', '34in Ultrawide Monitor'): 5, ('Apple Airpods Headphone
s', 'Vareebadd Phone'): 5, ('27in 4K Gaming Monitor', 'Google Phone'): 5, ('2
7in FHD Monitor', '34in Ultrawide Monitor'): 5, ('Flatscreen TV', 'Apple Airp
ods Headphones'): 5, ('27in FHD Monitor', 'Macbook Pro Laptop'): 5, ('20in Mo
nitor', 'AA Batteries (4-pack)'): 5, ('34in Ultrawide Monitor', '27in FHD Mon
itor'): 5, ('Macbook Pro Laptop', '34in Ultrawide Monitor'): 4, ('iPhone', 'V
areebadd Phone'): 4, ('Bose SoundSport Headphones', 'ThinkPad Laptop'): 4,
('20in Monitor', 'Macbook Pro Laptop'): 4, ('Vareebadd Phone', '34in Ultrawid
e Monitor'): 4, ('Flatscreen TV', 'Wired Headphones'): 4, ('LG Dryer', 'AA Ba
tteries (4-pack)'): 4, ('27in FHD Monitor', '27in 4K Gaming Monitor'): 4, ('T
hinkPad Laptop', 'Flatscreen TV'): 4, ('Flatscreen TV', 'iPhone'): 4, ('27in
4K Gaming Monitor', 'ThinkPad Laptop'): 4, ('Vareebadd Phone', 'Google Phon
e'): 4, ('Macbook Pro Laptop', 'Google Phone'): 4, ('27in 4K Gaming Monitor',
'27in FHD Monitor'): 4, ('Lightning Charging Cable', 'LG Washing Machine'):
4, ('27in FHD Monitor', 'ThinkPad Laptop'): 4, ('ThinkPad Laptop', 'Wired Hea
dphones'): 4, ('Bose SoundSport Headphones', 'Macbook Pro Laptop'): 4, ('AAA
Batteries (4-pack)', 'Vareebadd Phone'): 4, ('LG Washing Machine', 'AAA Batte
ries (4-pack)'): 4, ('Macbook Pro Laptop', 'ThinkPad Laptop'): 3, ('34in Ultr
awide Monitor', 'Macbook Pro Laptop'): 3, ('Lightning Charging Cable', 'Varee
badd Phone'): 3, ('Flatscreen TV', '27in FHD Monitor'): 3, ('Google Phone',
'ThinkPad Laptop'): 3, ('20in Monitor', '20in Monitor'): 3, ('ThinkPad Lapto
p', 'iPhone'): 3, ('Vareebadd Phone', 'Flatscreen TV'): 3, ('Flatscreen TV',
'Macbook Pro Laptop'): 3, ('34in Ultrawide Monitor', 'ThinkPad Laptop'): 3,
('Macbook Pro Laptop', 'iPhone'): 3, ('Vareebadd Phone', 'iPhone'): 3, ('Wire
d Headphones', 'LG Washing Machine'): 3, ('Google Phone', '34in Ultrawide Mon
itor'): 3, ('Flatscreen TV', 'Bose SoundSport Headphones'): 3, ('AA Batteries
(4-pack)', 'Vareebadd Phone'): 3, ('27in FHD Monitor', '20in Monitor'): 3,
('iPhone', 'Google Phone'): 3, ('iPhone', 'ThinkPad Laptop'): 3, ('27in 4K Ga
ming Monitor', 'iPhone'): 3, ('Google Phone', 'Google Phone'): 3, ('Flatscree
n TV', 'Google Phone'): 3, ('Google Phone', 'Macbook Pro Laptop'): 3, ('27in
4K Gaming Monitor', 'Flatscreen TV'): 3, ('Apple Airpods Headphones', 'LG Dry
er'): 3, ('20in Monitor', 'AAA Batteries (4-pack)'): 3, ('iPhone', 'Macbook P
ro Laptop'): 3, ('34in Ultrawide Monitor', 'Google Phone'): 2, ('Macbook Pro
Laptop', '20in Monitor'): 2, ('Lightning Charging Cable', 'LG Dryer'): 2, ('F
latscreen TV', '27in 4K Gaming Monitor'): 2, ('34in Ultrawide Monitor', 'Flat
screen TV'): 2, ('Macbook Pro Laptop', 'LG Washing Machine'): 2, ('Macbook Pr
o Laptop', '27in FHD Monitor'): 2, ('ThinkPad Laptop', 'ThinkPad Laptop'): 2,
('Bose SoundSport Headphones', 'Vareebadd Phone'): 2, ('Vareebadd Phone', 'Th
inkPad Laptop'): 2, ('20in Monitor', 'ThinkPad Laptop'): 2, ('iPhone', 'iPhon
e'): 2, ('27in FHD Monitor', 'LG Dryer'): 2, ('Vareebadd Phone', '27in 4K Gam
ing Monitor'): 2, ('27in 4K Gaming Monitor', '20in Monitor'): 2, ('LG Washing
Machine', 'Lightning Charging Cable'): 2, ('LG Washing Machine', 'Bose SoundS
port Headphones'): 2, ('AA Batteries (4-pack)', 'LG Dryer'): 2, ('Vareebadd P
hone', 'AAA Batteries (4-pack)'): 2, ('20in Monitor', 'Google Phone'): 2, ('T
```

```
hinkPad Laptop', 'Google Phone'): 2, ('Flatscreen TV', 'ThinkPad Laptop'): 2,
('ThinkPad Laptop', '27in FHD Monitor'): 2, ('Macbook Pro Laptop', 'Macbook P
ro Laptop'): 2, ('27in FHD Monitor', 'Flatscreen TV'): 2, ('Google Phone', '2
0in Monitor'): 2, ('27in 4K Gaming Monitor', 'Vareebadd Phone'): 1, ('ThinkPa
d Laptop', 'Macbook Pro Laptop'): 1, ('27in FHD Monitor', 'iPhone'): 1, ('Var
eebadd Phone', 'Lightning Charging Cable'): 1, ('Google Phone', 'Vareebadd Ph
one'): 1, ('20in Monitor', 'iPhone'): 1, ('LG Dryer', 'Vareebadd Phone'): 1,
('Macbook Pro Laptop', 'Flatscreen TV'): 1, ('20in Monitor', '27in FHD Monito
r'): 1, ('Google Phone', 'Flatscreen TV'): 1, ('LG Washing Machine', 'Google
Phone'): 1, ('LG Washing Machine', 'Wired Headphones'): 1, ('LG Dryer', 'Flat
screen TV'): 1, ('27in FHD Monitor', 'LG Washing Machine'): 1, ('20in Monito
r', '34in Ultrawide Monitor'): 1, ('34in Ultrawide Monitor', '20in Monitor'):
1, ('34in Ultrawide Monitor', 'LG Washing Machine'): 1, ('Google Phone', '27i
n 4K Gaming Monitor'): 1, ('LG Washing Machine', 'iPhone'): 1, ('LG Dryer',
'Wired Headphones'): 1, ('27in FHD Monitor', 'Vareebadd Phone'): 1, ('LG Wash
ing Machine', '27in 4K Gaming Monitor'): 1, ('LG Washing Machine', 'Apple Air
pods Headphones'): 1, ('27in 4K Gaming Monitor', 'LG Dryer'): 1, ('20in Monit
or', 'LG Washing Machine'): 1, ('LG Dryer', 'Google Phone'): 1, ('Vareebadd P
hone', '27in FHD Monitor'): 1, ('ThinkPad Laptop', '27in 4K Gaming Monitor'):
1, ('20in Monitor', 'Flatscreen TV'): 1, ('USB-C Charging Cable', 'LG Drye
r'): 1, ('LG Washing Machine', '20in Monitor'): 1, ('Flatscreen TV', '20in Mo
nitor'): 1, ('27in FHD Monitor', 'Google Phone'): 1, ('LG Dryer', 'AAA Batter
ies (4-pack)'): 1, ('ThinkPad Laptop', '34in Ultrawide Monitor'): 1, ('iPhon
e', 'LG Washing Machine'): 1, ('AAA Batteries (4-pack)', 'LG Dryer'): 1, ('LG
Dryer', '27in 4K Gaming Monitor'): 1, ('iPhone', '20in Monitor'): 1, ('LG Dry
er', 'Lightning Charging Cable'): 1, ('ThinkPad Laptop', 'LG Dryer'): 1, ('LG
Washing Machine', 'AA Batteries (4-pack)'): 1})
```

In [132]:
```
# Most common 10 values
count.most_common(10)
```

Out[132]:
```
[(('iPhone', 'Lightning Charging Cable'), 973),
 (('Google Phone', 'USB-C Charging Cable'), 954),
 (('iPhone', 'Wired Headphones'), 437),
 (('Google Phone', 'Wired Headphones'), 401),
 (('Vareebadd Phone', 'USB-C Charging Cable'), 348),
 (('iPhone', 'Apple Airpods Headphones'), 343),
 (('Google Phone', 'Bose SoundSport Headphones'), 219),
 (('USB-C Charging Cable', 'Wired Headphones'), 156),
 (('Vareebadd Phone', 'Wired Headphones'), 141),
 (('Lightning Charging Cable', 'Wired Headphones'), 92)]
```

In [160]:
```
# Most common 3 items purchased together
for row in df['Grouped']:
    row_list=row.split(',')
    count.update(Counter(combinations(row_list, 3)))

for key, value in count.most_common(10):
    print (key, value)
```

```
('iPhone', 'Lightning Charging Cable') 1946
('Google Phone', 'USB-C Charging Cable') 1908
('iPhone', 'Wired Headphones') 874
('Google Phone', 'Wired Headphones') 802
('Vareebadd Phone', 'USB-C Charging Cable') 696
('iPhone', 'Apple Airpods Headphones') 686
('Google Phone', 'Bose SoundSport Headphones') 438
('USB-C Charging Cable', 'Wired Headphones') 312
('Vareebadd Phone', 'Wired Headphones') 282
('Lightning Charging Cable', 'Wired Headphones') 184
```

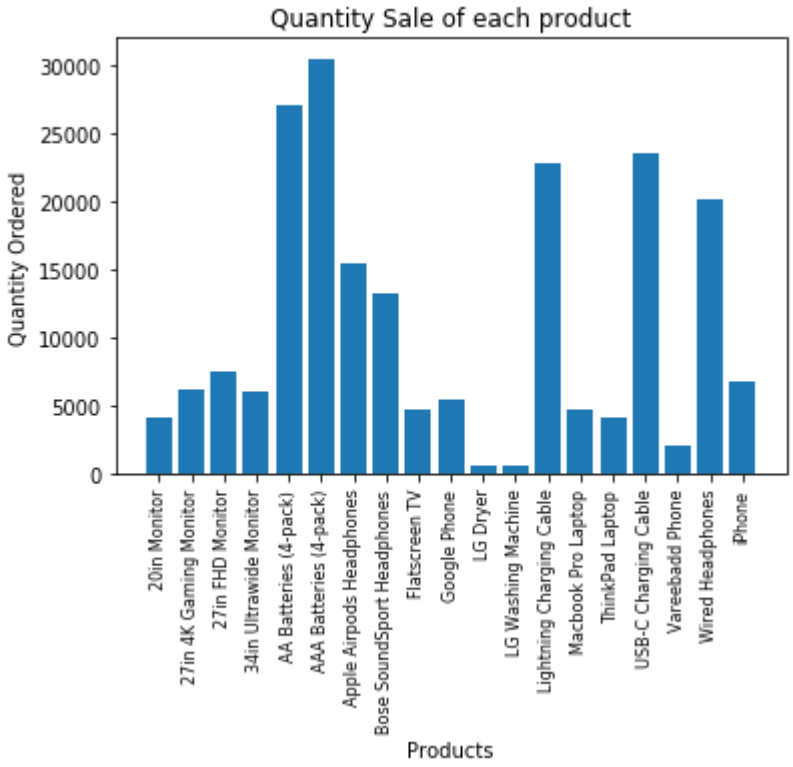
Q5: Which product sold most? why do you think that sold most?

In [142]:
```
product_group=all_data.groupby('Product')
quantity_ordered=product_group.sum()['Quantity Ordered']
product=[product for product, df in product_group]

quantity_ordered
```

Out[142]:
```
Product
20in Monitor                  4058
27in 4K Gaming Monitor        6134
27in FHD Monitor              7429
34in Ultrawide Monitor        6076
AA Batteries (4-pack)        27148
AAA Batteries (4-pack)       30487
Apple Airpods Headphones     15383
Bose SoundSport Headphones   13236
Flatscreen TV                 4724
Google Phone                  5440
LG Dryer                       640
LG Washing Machine             649
Lightning Charging Cable     22841
Macbook Pro Laptop            4645
ThinkPad Laptop               4054
USB-C Charging Cable         23555
Vareebadd Phone               2023
Wired Headphones             20201
iPhone                        6732
Name: Quantity Ordered, dtype: int64
```

In [146]:
```
plt.bar(product, quantity_ordered)
plt.xticks(product, rotation=90, size=8)
plt.xlabel("Products")
plt.ylabel("Quantity Ordered")
plt.title("Quantity Sale of each product")
plt.show()
```



In [149]:
```
prices= all_data.groupby('Product').mean()['Price Each']
prices
```

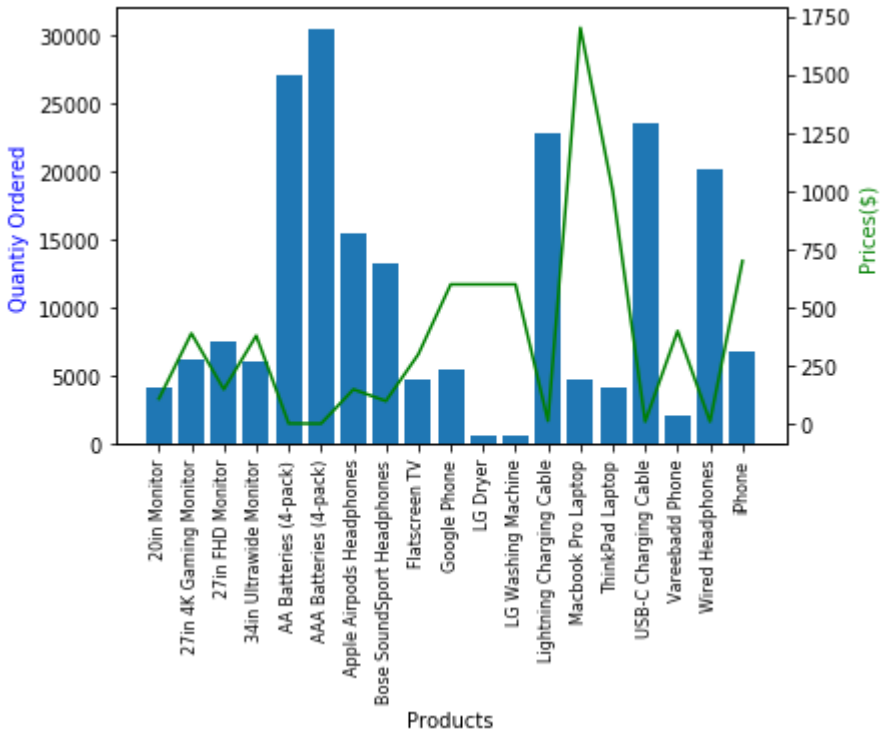
Out[149]:
```
Product
20in Monitor                  109.99
27in 4K Gaming Monitor        389.99
27in FHD Monitor              149.99
34in Ultrawide Monitor        379.99
AA Batteries (4-pack)           3.84
AAA Batteries (4-pack)          2.99
Apple Airpods Headphones      150.00
Bose SoundSport Headphones     99.99
Flatscreen TV                 300.00
Google Phone                  600.00
LG Dryer                      600.00
LG Washing Machine            600.00
Lightning Charging Cable       14.95
Macbook Pro Laptop           1700.00
ThinkPad Laptop               999.99
USB-C Charging Cable           11.95
Vareebadd Phone               400.00
Wired Headphones               11.99
iPhone                        700.00
Name: Price Each, dtype: float64
```

In [159]:
```
fig, ax1=plt.subplots()

ax2=ax1.twinx()
ax1.bar(product, quantity_ordered)
ax2.plot(product, prices, 'g-')

ax1.set_xlabel("Products")
ax1.set_xticklabels(product, rotation='vertical', size=8)
ax1.set_ylabel('Quantiy Ordered', color='b')
ax2.set_ylabel('Prices($)', color='g')

plt.show()
```



Product with low cost is sold most.