Name: Bhagyashri Thakur 163 A3

PRACTICAL NO.04

Practice Lab Assignment:

1. Perform all the pandas operations in python. Lab Assignment: Read any real life dataset. Store the data into Data Frames. Identify 20 grains for the given dataset. Implement all 20 grains using Pandas methods. The Sample Grains for Sales Dataset as:

- Which was the best month for sales? How much was earned that month?
- Which product sold the most? Why do you think it did?
- Which city sold the most products?
- What Products are most often sold together? Self Study Assignment: Perform advanced Data Manipulation operations

```
import numpy as np
import pandas as pd

all_data=pd.read_csv("/content/1686715083343_all_data.csv")

all_data.head(10)
```

```
   Order ID                     Product  Quantity Ordered  Price Each
\
0  176559.0  Bose SoundSport Headphones               1.0       99.99

1  176560.0                Google Phone               1.0      600.00

2  176560.0            Wired Headphones               1.0       11.99

3  176561.0            Wired Headphones               1.0       11.99

4  176562.0        USB-C Charging Cable               1.0       11.95

5  176563.0  Bose SoundSport Headphones               1.0       99.99

6  176564.0        USB-C Charging Cable               1.0       11.95

7  184074.0      AAA Batteries (4-pack)               1.0        2.99

8  184075.0      AAA Batteries (4-pack)               1.0        2.99

9  184076.0    Lightning Charging Cable               1.0       14.95


         Order Date                         Purchase Address
0  04-07-2019 22:30       682 Chestnut St, Boston, MA 02215
1  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001
```

```
2  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001
3     05/30/19 9:27       333 8th St, Los Angeles, CA 90001
4    04/29/19 13:03  381 Wilson St, San Francisco, CA 94016
5   05-02-2019 7:46        668 Center St, Seattle, WA 98101
6  04-12-2019 10:58         790 Ridge St, Atlanta, GA 30301
7  04-12-2019 10:47     544 1st St, San Francisco, CA 94016
8  04-03-2019 20:20           241 11th St, Austin, TX 73301
9    06/17/19 18:22  752 Chestnut St, Los Angeles, CA 90001
```

Clean up the data!

Drop rows of NAN

```
# Find NAN
nan_df = all_data[all_data.isna().any(axis=1)]
display(nan_df.head())

all_data = all_data.dropna(how='all')
all_data.head(60)
```

```
Empty DataFrame
Columns: [Order ID, Product, Quantity Ordered, Price Each, Order Date,
Purchase Address]
Index: []
```

	Order ID	Product	Quantity Ordered	Price Each
0	176559.0	Bose SoundSport Headphones	1.0	99.99
1	176560.0	Google Phone	1.0	600.00
2	176560.0	Wired Headphones	1.0	11.99
3	176561.0	Wired Headphones	1.0	11.99
4	176562.0	USB-C Charging Cable	1.0	11.95
5	176563.0	Bose SoundSport Headphones	1.0	99.99
6	176564.0	USB-C Charging Cable	1.0	11.95
7	184074.0	AAA Batteries (4-pack)	1.0	2.99
8	184075.0	AAA Batteries (4-pack)	1.0	2.99
9	184076.0	Lightning Charging Cable	1.0	14.95
10	179776.0	Wired Headphones	2.0	11.99

11	179790.0	AAA Batteries (4-pack)	2.0	2.99
12	179815.0	Wired Headphones	2.0	11.99
13	179821.0	AAA Batteries (4-pack)	2.0	2.99
14	179845.0	AAA Batteries (4-pack)	2.0	2.99
15	179871.0	AAA Batteries (4-pack)	2.0	2.99
16	179917.0	AA Batteries (4-pack)	2.0	3.84
17	179918.0	USB-C Charging Cable	2.0	11.95
18	179920.0	AA Batteries (4-pack)	2.0	3.84
19	179943.0	AAA Batteries (4-pack)	2.0	2.99
20	179970.0	AAA Batteries (4-pack)	2.0	2.99
21	180004.0	Lightning Charging Cable	2.0	14.95
22	187091.0	AA Batteries (4-pack)	3.0	3.84
23	187147.0	AAA Batteries (4-pack)	3.0	2.99
24	187170.0	AAA Batteries (4-pack)	3.0	2.99
25	187467.0	AA Batteries (4-pack)	3.0	3.84
26	187475.0	AAA Batteries (4-pack)	3.0	2.99
27	187505.0	AA Batteries (4-pack)	3.0	3.84
28	187567.0	AAA Batteries (4-pack)	3.0	2.99
29	187588.0	AA Batteries (4-pack)	3.0	3.84
30	187656.0	AA Batteries (4-pack)	3.0	3.84
31	187728.0	AAA Batteries (4-pack)	3.0	2.99
32	187740.0	AA Batteries (4-pack)	3.0	3.84
33	187806.0	AAA Batteries (4-pack)	3.0	2.99
34	187857.0	AAA Batteries (4-pack)	3.0	2.99

35	187895.0	AA Batteries (4-pack)	3.0	3.84
37	188069.0	AAA Batteries (4-pack)	3.0	2.99
38	180790.0	AA Batteries (4-pack)	4.0	3.84
39	180941.0	AAA Batteries (4-pack)	4.0	2.99
40	181213.0	AA Batteries (4-pack)	4.0	3.84
41	181642.0	AA Batteries (4-pack)	4.0	3.84
42	181976.0	USB-C Charging Cable	4.0	11.95
43	182448.0	AAA Batteries (4-pack)	4.0	2.99
44	182538.0	AAA Batteries (4-pack)	4.0	2.99
45	182660.0	AAA Batteries (4-pack)	4.0	2.99
46	185349.0	AAA Batteries (4-pack)	5.0	2.99
47	185711.0	AA Batteries (4-pack)	5.0	3.84
48	185915.0	AA Batteries (4-pack)	5.0	3.84
49	186028.0	AAA Batteries (4-pack)	5.0	2.99
50	186331.0	AAA Batteries (4-pack)	5.0	2.99
52	187182.0	AA Batteries (4-pack)	5.0	3.84
53	188953.0	AAA Batteries (4-pack)	5.0	2.99
54	190462.0	AA Batteries (4-pack)	5.0	3.84
55	190838.0	AA Batteries (4-pack)	5.0	3.84
56	192160.0	AAA Batteries (4-pack)	5.0	2.99
57	192932.0	AAA Batteries (4-pack)	5.0	2.99
58	177838.0	AAA Batteries (4-pack)	6.0	2.99
59	179865.0	AAA Batteries (4-pack)	6.0	2.99
60	183636.0	AAA Batteries (4-pack)	6.0	2.99

61	177160.0	AAA Batteries (4-pack)	7.0	2.99

	Order Date	Purchase Address
0	04-07-2019 22:30	682 Chestnut St, Boston, MA 02215
1	04-12-2019 14:38	669 Spruce St, Los Angeles, CA 90001
2	04-12-2019 14:38	669 Spruce St, Los Angeles, CA 90001
3	05/30/19 9:27	333 8th St, Los Angeles, CA 90001
4	04/29/19 13:03	381 Wilson St, San Francisco, CA 94016
5	05-02-2019 7:46	668 Center St, Seattle, WA 98101
6	04-12-2019 10:58	790 Ridge St, Atlanta, GA 30301
7	04-12-2019 10:47	544 1st St, San Francisco, CA 94016
8	04-03-2019 20:20	241 11th St, Austin, TX 73301
9	06/17/19 18:22	752 Chestnut St, Los Angeles, CA 90001
10	04/21/19 21:38	978 Wilson St, Seattle, WA 98101
11	04-10-2019 9:58	475 Adams St, Boston, MA 02215
12	04/28/19 15:42	91 9th St, Seattle, WA 98101
13	04/22/19 16:19	975 Main St, Dallas, TX 75001
14	08-11-2019 13:08	541 Adams St, Los Angeles, CA 90001
15	04/24/19 9:17	362 Lake St, San Francisco, CA 94016
16	04/24/19 22:19	152 Spruce St, Los Angeles, CA 90001
17	04/14/19 19:27	216 7th St, Dallas, TX 75001
18	04-11-2019 15:36	598 14th St, Seattle, WA 98101
19	08/14/19 11:08	344 River St, Portland, OR 97035
20	04/18/19 18:44	668 Lake St, Los Angeles, CA 90001
21	08-06-2019 19:38	307 Lincoln St, New York City, NY 10001
22	04-04-2019 13:50	303 Ridge St, San Francisco, CA 94016
23	04/23/19 17:42	929 Willow St, Los Angeles, CA 90001
24	08-09-2019 23:11	810 Ridge St, Los Angeles, CA 90001
25	04/13/19 11:50	747 Center St, New York City, NY 10001
26	04-02-2019 14:05	859 Wilson St, Los Angeles, CA 90001
27	04/14/19 11:40	752 Ridge St, Atlanta, GA 30301
28	04/16/19 20:44	855 Highland St, New York City, NY 10001
29	04/16/19 20:44	80 Main St, Boston, MA 02215
30	04/23/19 16:44	376 6th St, Portland, OR 97035
31	04-02-2019 14:29	430 Park St, San Francisco, CA 94016
32	09/30/19 1:58	77 South St, Atlanta, GA 30301
33	04/19/19 18:54	159 1st St, Los Angeles, CA 90001
34	09/23/19 20:40	384 Hickory St, San Francisco, CA 94016
35	09-01-2019 20:54	647 14th St, Austin, TX 73301
37	10/21/19 14:07	472 Chestnut St, Atlanta, GA 30301
38	04-03-2019 17:01	626 7th St, Boston, MA 02215
39	04/21/19 19:39	725 2nd St, San Francisco, CA 94016
40	10/28/19 14:51	919 5th St, New York City, NY 10001
41	10-01-2019 16:02	695 12th St, San Francisco, CA 94016
42	04-02-2019 11:02	561 2nd St, Atlanta, GA 30301
43	04-07-2019 8:52	456 Dogwood St, Atlanta, GA 30301
44	04/23/19 22:06	516 6th St, Atlanta, GA 30301
45	11-05-2019 22:47	291 Chestnut St, Portland, OR 97035

```
46     04/26/19 7:26           675 Center St, Boston, MA 02215
47  11-10-2019 11:54       481 South St, Los Angeles, CA 90001
48    11/14/19 17:15      749 Sunset St, Los Angeles, CA 90001
49  11-12-2019 13:54  414 Chestnut St, San Francisco, CA 94016
50  04-09-2019 11:35         553 2nd St, Los Angeles, CA 90001
52  04-06-2019 13:05      34 South St, New York City, NY 10001
53     04/30/19 9:32        20 Maple St, Los Angeles, CA 90001
54  04-04-2019 12:28        273 Jackson St, Portland, OR 97035
55    04/13/19 12:59             787 Main St, Boston, MA 02215
56     04/22/19 0:03      498 Washington St, Atlanta, GA 30301
57  12-08-2019 16:51         553 Dogwood St, Seattle, WA 98101
58     12/29/19 0:16      872 Pine St, San Francisco, CA 94016
59    12/14/19 11:24            83 Willow St, Austin, TX 73301
60    04/30/19 10:28             116 11th St, Boston, MA 02215
61    04/26/19 11:37              689 7th St, Dallas, TX 75001
```

Get rid of text in order date column

```
all_data = all_data[all_data['Order Date'].str[0:2]!='Or']
all_data.head(60)
```

```
    Order ID                    Product  Quantity Ordered  Price Each
\
0   176559.0  Bose SoundSport Headphones               1.0       99.99

1   176560.0                Google Phone               1.0      600.00

2   176560.0             Wired Headphones               1.0       11.99

3   176561.0             Wired Headphones               1.0       11.99

4   176562.0         USB-C Charging Cable               1.0       11.95

5   176563.0  Bose SoundSport Headphones               1.0       99.99

6   176564.0         USB-C Charging Cable               1.0       11.95

7   184074.0       AAA Batteries (4-pack)               1.0        2.99

8   184075.0       AAA Batteries (4-pack)               1.0        2.99

9   184076.0    Lightning Charging Cable               1.0       14.95

10  179776.0             Wired Headphones               2.0       11.99

11  179790.0       AAA Batteries (4-pack)               2.0        2.99

12  179815.0             Wired Headphones               2.0       11.99
```

13	179821.0	AAA Batteries (4-pack)	2.0	2.99
14	179845.0	AAA Batteries (4-pack)	2.0	2.99
15	179871.0	AAA Batteries (4-pack)	2.0	2.99
16	179917.0	AA Batteries (4-pack)	2.0	3.84
17	179918.0	USB-C Charging Cable	2.0	11.95
18	179920.0	AA Batteries (4-pack)	2.0	3.84
19	179943.0	AAA Batteries (4-pack)	2.0	2.99
20	179970.0	AAA Batteries (4-pack)	2.0	2.99
21	180004.0	Lightning Charging Cable	2.0	14.95
22	187091.0	AA Batteries (4-pack)	3.0	3.84
23	187147.0	AAA Batteries (4-pack)	3.0	2.99
24	187170.0	AAA Batteries (4-pack)	3.0	2.99
25	187467.0	AA Batteries (4-pack)	3.0	3.84
26	187475.0	AAA Batteries (4-pack)	3.0	2.99
27	187505.0	AA Batteries (4-pack)	3.0	3.84
28	187567.0	AAA Batteries (4-pack)	3.0	2.99
29	187588.0	AA Batteries (4-pack)	3.0	3.84
30	187656.0	AA Batteries (4-pack)	3.0	3.84
31	187728.0	AAA Batteries (4-pack)	3.0	2.99
32	187740.0	AA Batteries (4-pack)	3.0	3.84
33	187806.0	AAA Batteries (4-pack)	3.0	2.99
34	187857.0	AAA Batteries (4-pack)	3.0	2.99
35	187895.0	AA Batteries (4-pack)	3.0	3.84
37	188069.0	AAA Batteries (4-pack)	3.0	2.99

38	180790.0	AA Batteries (4-pack)	4.0	3.84
39	180941.0	AAA Batteries (4-pack)	4.0	2.99
40	181213.0	AA Batteries (4-pack)	4.0	3.84
41	181642.0	AA Batteries (4-pack)	4.0	3.84
42	181976.0	USB-C Charging Cable	4.0	11.95
43	182448.0	AAA Batteries (4-pack)	4.0	2.99
44	182538.0	AAA Batteries (4-pack)	4.0	2.99
45	182660.0	AAA Batteries (4-pack)	4.0	2.99
46	185349.0	AAA Batteries (4-pack)	5.0	2.99
47	185711.0	AA Batteries (4-pack)	5.0	3.84
48	185915.0	AA Batteries (4-pack)	5.0	3.84
49	186028.0	AAA Batteries (4-pack)	5.0	2.99
50	186331.0	AAA Batteries (4-pack)	5.0	2.99
52	187182.0	AA Batteries (4-pack)	5.0	3.84
53	188953.0	AAA Batteries (4-pack)	5.0	2.99
54	190462.0	AA Batteries (4-pack)	5.0	3.84
55	190838.0	AA Batteries (4-pack)	5.0	3.84
56	192160.0	AAA Batteries (4-pack)	5.0	2.99
57	192932.0	AAA Batteries (4-pack)	5.0	2.99
58	177838.0	AAA Batteries (4-pack)	6.0	2.99
59	179865.0	AAA Batteries (4-pack)	6.0	2.99
60	183636.0	AAA Batteries (4-pack)	6.0	2.99
61	177160.0	AAA Batteries (4-pack)	7.0	2.99

Order Date Purchase Address

0	04-07-2019 22:30	682 Chestnut St, Boston, MA 02215
1	04-12-2019 14:38	669 Spruce St, Los Angeles, CA 90001
2	04-12-2019 14:38	669 Spruce St, Los Angeles, CA 90001
3	05/30/19 9:27	333 8th St, Los Angeles, CA 90001
4	04/29/19 13:03	381 Wilson St, San Francisco, CA 94016
5	05-02-2019 7:46	668 Center St, Seattle, WA 98101
6	04-12-2019 10:58	790 Ridge St, Atlanta, GA 30301
7	04-12-2019 10:47	544 1st St, San Francisco, CA 94016
8	04-03-2019 20:20	241 11th St, Austin, TX 73301
9	06/17/19 18:22	752 Chestnut St, Los Angeles, CA 90001
10	04/21/19 21:38	978 Wilson St, Seattle, WA 98101
11	04-10-2019 9:58	475 Adams St, Boston, MA 02215
12	04/28/19 15:42	91 9th St, Seattle, WA 98101
13	04/22/19 16:19	975 Main St, Dallas, TX 75001
14	08-11-2019 13:08	541 Adams St, Los Angeles, CA 90001
15	04/24/19 9:17	362 Lake St, San Francisco, CA 94016
16	04/24/19 22:19	152 Spruce St, Los Angeles, CA 90001
17	04/14/19 19:27	216 7th St, Dallas, TX 75001
18	04-11-2019 15:36	598 14th St, Seattle, WA 98101
19	08/14/19 11:08	344 River St, Portland, OR 97035
20	04/18/19 18:44	668 Lake St, Los Angeles, CA 90001
21	08-06-2019 19:38	307 Lincoln St, New York City, NY 10001
22	04-04-2019 13:50	303 Ridge St, San Francisco, CA 94016
23	04/23/19 17:42	929 Willow St, Los Angeles, CA 90001
24	08-09-2019 23:11	810 Ridge St, Los Angeles, CA 90001
25	04/13/19 11:50	747 Center St, New York City, NY 10001
26	04-02-2019 14:05	859 Wilson St, Los Angeles, CA 90001
27	04/14/19 11:40	752 Ridge St, Atlanta, GA 30301
28	04/16/19 20:44	855 Highland St, New York City, NY 10001
29	04/16/19 20:44	80 Main St, Boston, MA 02215
30	04/23/19 16:44	376 6th St, Portland, OR 97035
31	04-02-2019 14:29	430 Park St, San Francisco, CA 94016
32	09/30/19 1:58	77 South St, Atlanta, GA 30301
33	04/19/19 18:54	159 1st St, Los Angeles, CA 90001
34	09/23/19 20:40	384 Hickory St, San Francisco, CA 94016
35	09-01-2019 20:54	647 14th St, Austin, TX 73301
37	10/21/19 14:07	472 Chestnut St, Atlanta, GA 30301
38	04-03-2019 17:01	626 7th St, Boston, MA 02215
39	04/21/19 19:39	725 2nd St, San Francisco, CA 94016
40	10/28/19 14:51	919 5th St, New York City, NY 10001
41	10-01-2019 16:02	695 12th St, San Francisco, CA 94016
42	04-02-2019 11:02	561 2nd St, Atlanta, GA 30301
43	04-07-2019 8:52	456 Dogwood St, Atlanta, GA 30301
44	04/23/19 22:06	516 6th St, Atlanta, GA 30301
45	11-05-2019 22:47	291 Chestnut St, Portland, OR 97035
46	04/26/19 7:26	675 Center St, Boston, MA 02215
47	11-10-2019 11:54	481 South St, Los Angeles, CA 90001
48	11/14/19 17:15	749 Sunset St, Los Angeles, CA 90001
49	11-12-2019 13:54	414 Chestnut St, San Francisco, CA 94016
50	04-09-2019 11:35	553 2nd St, Los Angeles, CA 90001

```
52   04-06-2019 13:05      34 South St, New York City, NY 10001
53      04/30/19 9:32        20 Maple St, Los Angeles, CA 90001
54   04-04-2019 12:28        273 Jackson St, Portland, OR 97035
55     04/13/19 12:59             787 Main St, Boston, MA 02215
56      04/22/19 0:03      498 Washington St, Atlanta, GA 30301
57   12-08-2019 16:51         553 Dogwood St, Seattle, WA 98101
58      12/29/19 0:16      872 Pine St, San Francisco, CA 94016
59     12/14/19 11:24            83 Willow St, Austin, TX 73301
60     04/30/19 10:28            116 11th St, Boston, MA 02215
61     04/26/19 11:37             689 7th St, Dallas, TX 75001
```

Make columns correct type

```
all_data['Quantity Ordered'] = pd.to_numeric(all_data['Quantity
Ordered'])
all_data['Price Each'] = pd.to_numeric(all_data['Price Each'])
```

Augment data with additional columns

Add month column¶

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

```
   Order ID                     Product  Quantity Ordered  Price Each
\
0  176559.0  Bose SoundSport Headphones               1.0       99.99

1  176560.0                Google Phone               1.0      600.00

2  176560.0             Wired Headphones               1.0       11.99

3  176561.0             Wired Headphones               1.0       11.99

4  176562.0         USB-C Charging Cable               1.0       11.95


         Order Date                       Purchase Address  Month
0  04-07-2019 22:30       682 Chestnut St, Boston, MA 02215      4
1  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
2  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
3     05/30/19 9:27       333 8th St, Los Angeles, CA 90001      5
4    04/29/19 13:03  381 Wilson St, San Francisco, CA 94016      4
```

Add month column (alternative method)

```
all_data['Month 2'] = pd.to_datetime(all_data['Order Date']).dt.month
all_data.head()
```

```
   Order ID                     Product  Quantity Ordered  Price Each
\
0  176559.0  Bose SoundSport Headphones               1.0       99.99

1  176560.0                Google Phone               1.0      600.00

2  176560.0             Wired Headphones               1.0       11.99

3  176561.0             Wired Headphones               1.0       11.99

4  176562.0         USB-C Charging Cable               1.0       11.95


         Order Date                         Purchase Address  Month
Month 2
0  04-07-2019 22:30       682 Chestnut St, Boston, MA 02215      4
4
1  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
4
2  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
4
3     05/30/19 9:27       333 8th St, Los Angeles, CA 90001      5
5
4    04/29/19 13:03  381 Wilson St, San Francisco, CA 94016      4
4
```

Add city column

```
def get_city(address):
    return address.split(",")[1].strip(" ")

def get_state(address):
    return address.split(",")[2].split(" ")[1]

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

```
   Order ID                     Product  Quantity Ordered  Price Each
\
0  176559.0  Bose SoundSport Headphones               1.0       99.99

1  176560.0                Google Phone               1.0      600.00

2  176560.0             Wired Headphones               1.0       11.99
```

```
3  176561.0            Wired Headphones                1.0       11.99

4  176562.0        USB-C Charging Cable                1.0       11.95


         Order Date                        Purchase Address Month
Month 2 \
0  04-07-2019 22:30       682 Chestnut St, Boston, MA 02215      4
4
1  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
4
2  04-12-2019 14:38    669 Spruce St, Los Angeles, CA 90001      4
4
3     05/30/19 9:27       333 8th St, Los Angeles, CA 90001      5
5
4    04/29/19 13:03  381 Wilson St, San Francisco, CA 94016      4
4

                  City
0         Boston  (MA)
1    Los Angeles  (CA)
2    Los Angeles  (CA)
3    Los Angeles  (CA)
4  San Francisco  (CA)
```

Data Exploration!

Question 1: What was the best month for sales? How much was earned that month?

```
all_data['Sales'] = all_data['Quantity Ordered'].astype('int64')*
all_data['Price Each'].astype('float')

all_data.groupby(['Month']).sum().max
```

```
<ipython-input-30-9826895e33f5>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.sum is deprecated. In a future
version, numeric_only will default to False. Either specify
numeric_only or select only columns which should be valid for the
function.
  all_data.groupby(['Month']).sum().max

<bound method NDFrame._add_numeric_operations.<locals>.max of
Order ID  Quantity Ordered  Price Each  Month 2    Sales
Month
4      7335546.0             123.0      885.80      160  1210.76
5       353124.0               2.0      111.98       10   111.98
6       184076.0               1.0       14.95        6    14.95
8       726962.0               9.0       23.92       32    50.83
```

```
9      2378802.0              17.0      591.44       90   616.62
10      550924.0              11.0       10.67       30    39.69
11      740314.0              19.0       13.66       44    65.31
12      550635.0              17.0        8.97       36    50.83>
```

Question 2: What city sold the most product?

```
city_max=all_data.groupby(['City']).sum()
print(max(city_max))
```

```
Sales

<ipython-input-31-79b556d70b46>:1: FutureWarning: The default value of
numeric_only in DataFrameGroupBy.sum is deprecated. In a future
version, numeric_only will default to False. Either specify
numeric_only or select only columns which should be valid for the
function.
  city_max=all_data.groupby(['City']).sum()
```

Question 4: What products are most often sold together?

```
df = all_data[all_data['Order ID'].duplicated(keep=False)]

# Referenced:
https://stackoverflow.com/questions/27298178/concatenate-strings-from-
several-rows-using-pandas-groupby
df['Grouped'] = df.groupby('Order ID')['Product'].transform(lambda x:
','.join(x))
df2 = df[['Order ID', 'Grouped']].drop_duplicates()
print(df['Grouped'])
```

```
3                              Google Phone,Wired Headphones
4                              Google Phone,Wired Headphones
18                          Google Phone,USB-C Charging Cable
19                          Google Phone,USB-C Charging Cable
30         Bose SoundSport Headphones,Bose SoundSport Hea...
                                    ...
186792           34in Ultrawide Monitor,AA Batteries (4-pack)
186803                Wired Headphones,AAA Batteries (4-pack)
186804                Wired Headphones,AAA Batteries (4-pack)
186841                      Google Phone,USB-C Charging Cable
186842                      Google Phone,USB-C Charging Cable
Name: Grouped, Length: 14649, dtype: object

C:\Users\ZEALIN~1\AppData\Local\Temp/ipykernel_14640/4070466232.py:4:
SettingWithCopyWarning:
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/stable/user_guide/indexing.html#
returning-a-view-versus-a-copy
  df['Grouped'] = df.groupby('Order ID')['Product'].transform(lambda
x: ','.join(x))

from itertools import combinations
from collections import Counter

count = Counter()

for row in df2['Grouped']:
    row_list = row.split(',')
    count.update(Counter(combinations(row_list, 2)))

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

('iPhone', 'Lightning Charging Cable') 1005
('Google Phone', 'USB-C Charging Cable') 987
('iPhone', 'Wired Headphones') 447
('Google Phone', 'Wired Headphones') 414
('Vareebadd Phone', 'USB-C Charging Cable') 361
('iPhone', 'Apple Airpods Headphones') 360
('Google Phone', 'Bose SoundSport Headphones') 220
('USB-C Charging Cable', 'Wired Headphones') 160
('Vareebadd Phone', 'Wired Headphones') 143
('Lightning Charging Cable', 'Wired Headphones') 92
```

What product sold the most? Why do you think it sold the most?

```
product_group = all_data.groupby('Product')
quantity_ordered = product_group.sum()['Quantity Ordered']

print(quantity_ordered)

Product
20in Monitor                   4129
27in 4K Gaming Monitor         6244
27in FHD Monitor               7550
34in Ultrawide Monitor         6199
AA Batteries (4-pack)         27635
AAA Batteries (4-pack)        31017
Apple Airpods Headphones      15661
Bose SoundSport Headphones    13457
Flatscreen TV                  4819
Google Phone                   5532
LG Dryer                        646
LG Washing Machine              666
Lightning Charging Cable      23217
Macbook Pro Laptop             4728
```

```
ThinkPad Laptop                4130
USB-C Charging Cable          23975
Vareebadd Phone                2068
Wired Headphones              20557
iPhone                         6849
Name: Quantity Ordered, dtype: int64
```

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

print(prices)
```

```
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
```