In [324...]
```
# importing libraries for analysis
import numpy as np
import pandas as pd
import xlrd
```

In [325...]
```
# importing data from excel
data = pd.read_excel('Frequency of Purchase Analysis Data Question.xlsx',engine = 'o
```

In [326...]
```
data
```

Out[326...]

	Outlet ID	Brand Name	Sales Value	DATE	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Un
0	Outlet 1	Brand 1	395.6906	2018-04-10	NaN	NaN	NaN	NaN	NaN	
1	Outlet 1	Brand 1	395.6906	2018-04-24	NaN	NaN	NaN	NaN	NaN	
2	Outlet 1	Brand 2	724.9328	2018-04-10	NaN	NaN	NaN	NaN	NaN	
3	Outlet 1	Brand 3	150.0466	2018-04-10	NaN	NaN	NaN	NaN	NaN	
4	Outlet 1	Brand 3	300.0932	2018-04-24	NaN	NaN	NaN	NaN	NaN	
...	...	...	...	...	...	...	...	...	...	
25468	Outlet 23559	Brand 3	150.0466	2018-04-05	NaN	NaN	NaN	NaN	NaN	
25469	Outlet 23560	Brand 1	395.6906	2018-04-05	NaN	NaN	NaN	NaN	NaN	
25470	Outlet 23560	Brand 3	300.0932	2018-04-05	NaN	NaN	NaN	NaN	NaN	
25471	Outlet 23561	Brand 1	395.6906	2018-04-05	NaN	NaN	NaN	NaN	NaN	
25472	Outlet 23561	Brand 3	150.0466	2018-04-05	NaN	NaN	NaN	NaN	NaN	

25473 rows × 11 columns

In [327...]
```
data.shape
```

Out[327...]
```
(25473, 11)
```

In [328...]
```
# removing the unnneccesary cells from the data frame
data = data[['Outlet ID','Brand Name','Sales Value','DATE']]
```

In [329...]
```
data
```

Out[329...

	Outlet ID	Brand Name	Sales Value	DATE
0	Outlet 1	Brand 1	395.6906	2018-04-10
1	Outlet 1	Brand 1	395.6906	2018-04-24
2	Outlet 1	Brand 2	724.9328	2018-04-10
3	Outlet 1	Brand 3	150.0466	2018-04-10
4	Outlet 1	Brand 3	300.0932	2018-04-24
...	...	...	...	...
25468	Outlet 23559	Brand 3	150.0466	2018-04-05
25469	Outlet 23560	Brand 1	395.6906	2018-04-05
25470	Outlet 23560	Brand 3	300.0932	2018-04-05
25471	Outlet 23561	Brand 1	395.6906	2018-04-05
25472	Outlet 23561	Brand 3	150.0466	2018-04-05

25473 rows × 4 columns

In [330...

```
# finding number of unique value for different attributes
data.nunique(0)
```

Out[330...

```
Outlet ID      5811
Brand Name       11
Sales Value     334
DATE             25
dtype: int64
```

total number of unique outlets : 5811

total number of brands : 11

In [331...

```
t_brands = 11
t_outlet = data.nunique()['Outlet ID']
t_outlet
```

Out[331...

```
5811
```

In [332...

```
# total number of purchases from each brand
tnp = data.groupby(["Brand Name"]).count()
tnp.drop(columns=['Sales Value','DATE'],inplace=True)
tnp.columns = ['number of times purchased']
tnp['total sales'] = data.groupby(["Brand Name"]).sum()
tnp
```

Out[332...

Brand Name	number of times purchased	total sales
Brand 1	3665	2.291894e+06
Brand 10	2863	4.200337e+06
Brand 11	1	6.197700e+03
Brand 2	3547	2.443848e+06

Brand Name	number of times purchased	total sales
Brand 3	1951	7.526337e+05
Brand 4	546	2.260473e+05
Brand 5	1496	2.766144e+06
Brand 6	1868	1.118381e+06
Brand 7	3301	2.577246e+06
Brand 8	1122	2.657290e+06
Brand 9	5113	7.295189e+06

In [333...
```
tnp.to_excel('Brand_Purchase_Data.xlsx')
```

In [334...
```
onp = data.groupby(["Outlet ID"]).count()
onp.drop(columns=['Sales Value','DATE'],inplace=True)
onp.columns = ['number of times purchased']
onp['total sales'] = data.groupby(["Outlet ID"]).sum()
onp
```

Out[334...

Outlet ID	number of times purchased	total sales
Outlet 1	6	2139.5366
Outlet 10	4	1786.7488
Outlet 100	3	7468.2568
Outlet 1000	9	9349.3012
Outlet 1001	5	3872.9116
...	...	...
Outlet 995	6	3699.2062
Outlet 996	9	6205.8504
Outlet 997	6	7610.4926
Outlet 998	6	3769.2204
Outlet 999	9	13469.7246

5811 rows × 2 columns

In [335...
```
onp.to_excel('Outlet_Purchase_Data.xlsx')
```

total number of brands = 11
the above data shows total number of times certain brands were bought and number of times product was purchase from certain outlet

All brands

In [336...

```
# finding the outlets purchased for different number of times
outlet_count = data.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = data.groupby(['Outlet ID']).sum()
outlet_count
```

Out[336...

Outlet ID	number of times purchased	total sales
Outlet 1	6	2139.5366
Outlet 10	4	1786.7488
Outlet 100	3	7468.2568
Outlet 1000	9	9349.3012
Outlet 1001	5	3872.9116
...	...	...
Outlet 995	6	3699.2062
Outlet 996	9	6205.8504
Outlet 997	6	7610.4926
Outlet 998	6	3769.2204
Outlet 999	9	13469.7246

5811 rows × 2 columns

In [337...

```
# finding the number of outlets with respect to number of times purchase and its tot
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[337...

number of times purchased	number of outlets	total sales
1	818	8.108168e+05
2	717	9.274539e+05
3	1210	3.478157e+06
4	744	2.955618e+06
5	489	2.305511e+06
6	468	3.085521e+06
7	344	2.464057e+06
8	235	1.989405e+06
9	786	8.318669e+06

In [338...

```
# frequncy of purchase of brand1
fp = len(data)/t_outlet
fp
```

Out[338...
```
4.383582860092927
```

In [339...
```
#total number of sales value
ts = data['Sales Value'].sum()
ts
```

Out[339...
```
26335207.9194
```

In [340...
```
frequency_dict = {}
frequency_dict['Brand Name'] = []
frequency_dict['Frequency of Purchase'] = []
frequency_dict['Frequency of Purchase without'] = []
frequency_dict['Total Sales'] = []
```

Considering for only Brand 1

In [341...
```
brand1 = data[data['Brand Name']=='Brand 1']
```

In [342...
```
brand1
brand1.to_excel('test.xlsx')
```

In [343...
```
len(brand1)
```

Out[343...
```
3665
```

In [344...
```
# finding the outlets purchased for different number of times
outlet_count = brand1.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand1.groupby(['Outlet ID']).sum()
outlet_count
```

Out[344...

Outlet ID	number of times purchased	total sales
Outlet 1	2	791.3812
Outlet 10	3	1636.7022
Outlet 1002	2	1007.1970
Outlet 1003	2	1007.1970
Outlet 1004	1	1223.0128
...	...	...
Outlet 993	1	1187.0718
Outlet 994	1	1223.0128
Outlet 995	2	1007.1970
Outlet 996	1	395.6906
Outlet 998	2	1007.1970

2133 rows × 2 columns

In [345…
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[345…

number of times purchased	number of outlets	total sales
1	1121	720219.7180
2	644	803817.6916
3	259	479013.3278
4	78	193059.0908
5	19	51672.9700
6	12	44111.2100

In [346…
```
# frequncy of purchase of brand1
fp = len(brand1)/t_outlet
fp
```

Out[346… `0.630700395801067`

In [347…
```
#total number of sales value
ts = brand1['Sales Value'].sum()
ts
```

Out[347… `2291894.0082`

In [348…
```
wfp = len(brand1)/len(outlet_count)
wfp
```

Out[348… `1.7182372245663384`

In [349…
```
frequency_dict['Brand Name'].append('Brand 1')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

considering for brand2 alone

In [350…
```
brand2 = data[data['Brand Name']=='Brand 2']
brand2
```

Out[350…

	Outlet ID	Brand Name	Sales Value	DATE
2	Outlet 1	Brand 2	724.9328	2018-04-10
7	Outlet 2	Brand 2	362.4664	2018-04-17
8	Outlet 2	Brand 2	538.2660	2018-04-17

	Outlet ID	Brand Name	Sales Value	DATE
9	Outlet 2	Brand 2	585.8100	2018-04-17
12	Outlet 3	Brand 2	276.0948	2018-04-24
...	...	...	...	...
25449	Outlet 23552	Brand 2	878.7150	2018-04-12
25456	Outlet 23554	Brand 2	292.9050	2018-04-12
25460	Outlet 23556	Brand 2	585.8100	2018-04-12
25463	Outlet 23557	Brand 2	1464.5250	2018-04-19
25465	Outlet 23558	Brand 2	292.9050	2018-04-05

3547 rows × 4 columns

In [351...
```
# finding the outlets purchased for different number of times
outlet_count = brand2.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand2.groupby(['Outlet ID']).sum()
outlet_count
```

Out[351...

Outlet ID	number of times purchased	total sales
Outlet 1	1	724.9328
Outlet 1037	1	585.8100
Outlet 1039	2	944.6540
Outlet 104	1	362.4664
Outlet 1040	2	784.4194
...	...	...
Outlet 890	1	585.8100
Outlet 9	2	661.0314
Outlet 926	1	292.9050
Outlet 930	1	292.9050
Outlet 931	1	292.9050

2329 rows × 2 columns

In [352...
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[352...

number of times purchased	number of outlets	total sales
1	1436	1.039915e+06

number of times purchased	number of outlets	total sales
2	637	8.438762e+05
3	197	3.874004e+05
4	49	1.326351e+05
5	10	4.002214e+04

In [353...
```
# frequncy of purchase of brand2
fp = len(brand2)/t_outlet
fp
```

Out[353... 0.6103940801927379

In [354...
```
#total number of sales value
ts = brand2['Sales Value'].sum()
ts
```

Out[354... 2443848.4244

In [355...
```
wfp = len(brand2)/len(outlet_count)
wfp
```

Out[355... 1.5229712322885358

In [356...
```
frequency_dict['Brand Name'].append('Brand 2')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 3

In [357...
```
brand3 = data[data['Brand Name']=='Brand 3']
brand3
```

Out[357...

	Outlet ID	Brand Name	Sales Value	DATE
3	Outlet 1	Brand 3	150.0466	2018-04-10
4	Outlet 1	Brand 3	300.0932	2018-04-24
13	Outlet 3	Brand 3	150.0466	2018-04-24
14	Outlet 3	Brand 3	450.1398	2018-04-17
22	Outlet 5	Brand 3	150.0466	2018-04-10
...	...	...	...	...
25464	Outlet 23557	Brand 3	150.0466	2018-04-19
25466	Outlet 23558	Brand 3	300.0932	2018-04-05
25468	Outlet 23559	Brand 3	150.0466	2018-04-05
25470	Outlet 23560	Brand 3	300.0932	2018-04-05

	Outlet ID	Brand Name	Sales Value	DATE
25472	Outlet 23561	Brand 3	150.0466	2018-04-05

1951 rows × 4 columns

In [358…
```
# finding the outlets purchased for different number of times
outlet_count = brand3.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand3.groupby(['Outlet ID']).sum()
outlet_count
```

Out[358…

Outlet ID	number of times purchased	total sales
Outlet 1	2	450.1398
Outlet 10	1	150.0466
Outlet 100	1	600.1864
Outlet 1000	2	600.1864
Outlet 1007	1	300.0932
...	...	...
Outlet 982	1	450.1398
Outlet 99	1	300.0932
Outlet 992	2	900.2796
Outlet 996	1	150.0466
Outlet 999	1	750.2330

1762 rows × 2 columns

In [359…
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[359…

number of times purchased	number of outlets	total sales
1	1588	612790.3144
2	159	127689.6566
3	15	12153.7746

In [360…
```
# frequncy of purchase of brand2
fp = len(brand3)/t_outlet
fp
```

Out[360…
```
0.33574255721906726
```

In [361...
```
#total number of sales value
ts = brand3['Sales Value'].sum()
ts
```

Out[361...
```
752633.7456000001
```

In [362...
```
wfp = len(brand3)/len(outlet_count)
wfp
```

Out[362...
```
1.1072644721906924
```

In [363...
```
frequency_dict['Brand Name'].append('Brand 3')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 4

In [364...
```
brand4 = data[data['Brand Name']=='Brand 4']
brand4
```

Out[364...

	Outlet ID	Brand Name	Sales Value	DATE
5	Outlet 1	Brand 4	173.0828	2018-04-24
15	Outlet 3	Brand 4	220.9098	2018-04-24
18	Outlet 4	Brand 4	220.9098	2018-04-10
24	Outlet 5	Brand 4	220.9098	2018-04-10
30	Outlet 7	Brand 4	220.9098	2018-04-10
...	...	...	...	...
24716	Outlet 23376	Brand 4	2700.3860	2018-04-26
24733	Outlet 23379	Brand 4	441.8196	2018-04-28
24738	Outlet 23380	Brand 4	441.8196	2018-04-11
25439	Outlet 23548	Brand 4	270.0386	2018-04-20
25440	Outlet 23548	Brand 4	346.1656	2018-04-20

546 rows × 4 columns

In [365...
```
# finding the outlets purchased for different number of times
outlet_count = brand4.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand4.groupby(['Outlet ID']).sum()
outlet_count
```

Out[365...

Outlet ID	number of times purchased	total sales
Outlet 1	1	173.0828

Outlet ID	number of times purchased	total sales
Outlet 103	1	220.9098
Outlet 1037	1	270.0386
Outlet 104	1	220.9098
Outlet 106	1	220.9098
...	...	...
Outlet 91	1	220.9098
Outlet 92	1	220.9098
Outlet 93	1	220.9098
Outlet 95	1	220.9098
Outlet 96	1	220.9098

406 rows × 2 columns

In [366...
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[366...

number of times purchased	number of outlets	total sales
1	303	146858.4352
2	73	50729.2782
3	24	21639.5950
4	5	5074.0768
5	1	1745.9402

In [367...
```
# frequncy of purchase of brand2
fp = len(brand4)/t_outlet
fp
```

Out[367... 0.09395973154362416

In [368...
```
#total number of sales value
ts = brand4['Sales Value'].sum()
ts
```

Out[368... 226047.3254

In [369...
```
wfp = len(brand4)/len(outlet_count)
wfp
```

Out[369... 1.3448275862068966

In [370…
```
frequency_dict['Brand Name'].append('Brand 4')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 5

In [371…
```
brand5 = data[data['Brand Name']=='Brand 5']
brand5
```

Out[371…

	Outlet ID	Brand Name	Sales Value	DATE
10	Outlet 3	Brand 5	350.2408	2018-04-17
19	Outlet 5	Brand 5	350.2408	2018-04-24
47	Outlet 11	Brand 5	350.2408	2018-04-17
62	Outlet 15	Brand 5	350.2408	2018-04-03
68	Outlet 17	Brand 5	350.2408	2018-04-03
…	…	…	…	…
25324	Outlet 23513	Brand 5	764.1000	2018-04-21
25333	Outlet 23514	Brand 5	764.1000	2018-04-07
25334	Outlet 23514	Brand 5	764.1000	2018-04-21
25347	Outlet 23516	Brand 5	764.1000	2018-04-25
25448	Outlet 23551	Brand 5	2292.3000	2018-04-19

1496 rows × 4 columns

In [372…
```
# finding the outlets purchased for different number of times
outlet_count = brand5.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand5.groupby(['Outlet ID']).sum()
outlet_count
```

Out[372…

Outlet ID	number of times purchased	total sales
Outlet 1002	1	1528.2000
Outlet 1003	1	700.4816
Outlet 1004	1	764.1000
Outlet 1023	1	764.1000
Outlet 1025	1	764.1000
…	…	…
Outlet 971	1	1528.2000
Outlet 973	1	1528.2000
Outlet 991	1	1528.2000

Outlet ID	number of times purchased	total sales
Outlet 993	1	764.1000
Outlet 994	1	764.1000

1300 rows × 2 columns

In [373...
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[373...

number of times purchased	number of outlets	total sales
1	1112	2.105609e+06
2	180	6.353515e+05
3	8	2.518383e+04

In [374...
```
# frequncy of purchase of brand2
fp = len(brand5)/t_outlet
fp
```

Out[374...
```
0.25744278093271383
```

In [375...
```
#total number of sales value
ts = brand5['Sales Value'].sum()
ts
```

Out[375...
```
2766143.88
```

In [376...
```
wfp = len(brand5)/len(outlet_count)
wfp
```

Out[376...
```
1.1507692307692308
```

In [377...
```
frequency_dict['Brand Name'].append('Brand 5')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 6

In [378...
```
brand6 = data[data['Brand Name']=='Brand 6']
brand6
```

Out[378...

	Outlet ID	Brand Name	Sales Value	DATE
83	Outlet 20	Brand 6	276.6608	2018-04-17

	Outlet ID	Brand Name	Sales Value	DATE
128	Outlet 30	Brand 6	366.2020	2018-04-17
134	Outlet 32	Brand 6	1383.3040	2018-04-17
141	Outlet 34	Brand 6	679.2000	2018-04-17
142	Outlet 34	Brand 6	1106.6432	2018-04-17
...	...	...	...	...
25220	Outlet 23487	Brand 6	553.3216	2018-04-18
25258	Outlet 23496	Brand 6	366.2020	2018-04-18
25266	Outlet 23497	Brand 6	1098.6060	2018-04-17
25315	Outlet 23512	Brand 6	549.3030	2018-04-07
25446	Outlet 23551	Brand 6	829.9824	2018-04-19

1868 rows × 4 columns

In [379...

```
# finding the outlets purchased for different number of times
outlet_count = brand6.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand6.groupby(['Outlet ID']).sum()
outlet_count
```

Out[379...

Outlet ID	number of times purchased	total sales
Outlet 1002	2	1379.2854
Outlet 1003	2	1379.2854
Outlet 1004	1	829.9824
Outlet 1023	2	1655.9462
Outlet 1027	2	1562.3864
...	...	...
Outlet 991	2	1379.2854
Outlet 993	1	732.4040
Outlet 994	1	829.9824
Outlet 995	2	1285.7256
Outlet 998	2	1379.2854

1369 rows × 2 columns

In [380...

```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[380...

	number of outlets	total sales

number of times purchased	number of outlets	total sales
number of times purchased		
1	912	582795.8802
2	421	480188.1174
3	30	44579.4052
4	6	10818.0712

In [381...
```
# frequncy of purchase of brand2
fp = len(brand6)/t_outlet
fp
```

Out[381...
```
0.3214593013250731
```

In [382...
```
#total number of sales value
ts = brand6['Sales Value'].sum()
ts
```

Out[382...
```
1118381.474
```

In [383...
```
wfp = len(brand6)/len(outlet_count)
wfp
```

Out[383...
```
1.364499634769905
```

In [384...
```
frequency_dict['Brand Name'].append('Brand 6')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 7

In [385...
```
brand7 = data[data['Brand Name']=='Brand 7']
brand7
```

Out[385...

	Outlet ID	Brand Name	Sales Value	DATE
74	Outlet 19	Brand 7	814.7004	2018-04-03
84	Outlet 20	Brand 7	438.6500	2018-04-17
85	Outlet 20	Brand 7	814.7004	2018-04-17
87	Outlet 21	Brand 7	362.0136	2018-04-24
94	Outlet 23	Brand 7	227.4188	2018-04-03
...	...	...	...	...
25353	Outlet 23519	Brand 7	227.4188	2018-04-09
25386	Outlet 23532	Brand 7	988.0662	2018-04-28
25432	Outlet 23547	Brand 7	219.3250	2018-04-13
25433	Outlet 23547	Brand 7	329.3554	2018-04-13

	Outlet ID	Brand Name	Sales Value	DATE
25447	Outlet 23551	Brand 7	724.0272	2018-04-19

3301 rows × 4 columns

In [386…
```
# finding the outlets purchased for different number of times
outlet_count = brand7.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand7.groupby(['Outlet ID']).sum()
outlet_count
```

Out[386…

Outlet ID	number of times purchased	total sales
Outlet 1000	1	1086.0408
Outlet 1001	2	1332.1376
Outlet 1002	1	909.6752
Outlet 1003	2	1633.7024
Outlet 1004	1	909.6752
...	...	...
Outlet 995	2	1406.2836
Outlet 996	3	1271.6888
Outlet 997	3	2528.9446
Outlet 998	2	1382.7380
Outlet 999	2	2325.3544

1981 rows × 2 columns

In [387…
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[387…

number of times purchased	number of outlets	total sales
1	1078	857363.7820
2	615	957120.1500
3	187	438362.3588
4	82	239555.4814
5	11	42544.3522
6	7	38073.6880
7	1	4225.8692

In [388…
```
# frequncy of purchase of brand2
fp = len(brand7)/t_outlet
fp
```

Out[388…
```
0.5680605747719841
```

In [389…
```
#total number of sales value
ts = brand7['Sales Value'].sum()
ts
```

Out[389…
```
2577245.6816
```

In [390…
```
wfp = len(brand7)/len(outlet_count)
wfp
```

Out[390…
```
1.6663301362948006
```

In [391…
```
frequency_dict['Brand Name'].append('Brand 7')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 8

In [392…
```
brand8 = data[data['Brand Name']=='Brand 8']
brand8
```

Out[392…

	Outlet ID	Brand Name	Sales Value	DATE
78	Outlet 19	Brand 8	2689.632	2018-04-03
79	Outlet 19	Brand 8	2689.632	2018-04-17
96	Outlet 23	Brand 8	448.272	2018-04-03
108	Outlet 25	Brand 8	448.272	2018-04-03
126	Outlet 29	Brand 8	2689.632	2018-04-17
…	…	…	…	…
25223	Outlet 23487	Brand 8	1344.816	2018-04-18
25270	Outlet 23497	Brand 8	2689.632	2018-04-17
25319	Outlet 23512	Brand 8	5379.264	2018-04-07
25325	Outlet 23513	Brand 8	5379.264	2018-04-21
25335	Outlet 23514	Brand 8	2689.632	2018-04-21

1122 rows × 4 columns

In [393…
```
# finding the outlets purchased for different number of times
outlet_count = brand8.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand8.groupby(['Outlet ID']).sum()
outlet_count
```

Out[393...

Outlet ID	number of times purchased	total sales
Outlet 1002	1	2689.632
Outlet 1003	1	2689.632
Outlet 1023	1	2689.632
Outlet 1025	1	2689.632
Outlet 1026	1	896.544
...	...	...
Outlet 94	1	1793.088
Outlet 962	1	5379.264
Outlet 98	1	2689.632
Outlet 991	1	2241.360
Outlet 996	1	1344.816

1012 rows × 2 columns

In [394...

```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[394...

number of times purchased	number of outlets	total sales
1	907	2.170102e+06
2	100	4.450496e+05
3	5	4.213757e+04

In [395...

```
# frequncy of purchase of brand2
fp = len(brand8)/t_outlet
fp
```

Out[395...

```
0.19308208569953536
```

In [396...

```
#total number of sales value
ts = brand8['Sales Value'].sum()
ts
```

Out[396...

```
2657289.6846000003
```

In [397...

```
wfp = len(brand8)/len(outlet_count)
wfp
```

Out[397...

```
1.108695652173913
```

In [398…
```
frequency_dict['Brand Name'].append('Brand 8')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 9

In [399…
```
brand9 = data[data['Brand Name']=='Brand 9']
brand9
```

Out[399…

	Outlet ID	Brand Name	Sales Value	DATE
88	Outlet 21	Brand 9	179.3088	2018-04-24
89	Outlet 21	Brand 9	631.9956	2018-04-24
92	Outlet 22	Brand 9	179.3088	2018-04-10
93	Outlet 22	Brand 9	462.0824	2018-04-10
99	Outlet 23	Brand 9	924.1648	2018-04-24
…	…	…	…	…
25346	Outlet 23515	Brand 9	2949.3128	2018-04-25
25358	Outlet 23520	Brand 9	2106.6520	2018-04-30
25388	Outlet 23532	Brand 9	3581.3084	2018-04-28
25392	Outlet 23533	Brand 9	1263.9912	2018-04-11
25435	Outlet 23547	Brand 9	1895.9868	2018-04-13

5113 rows × 4 columns

In [400…
```
# finding the outlets purchased for different number of times
outlet_count = brand9.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand9.groupby(['Outlet ID']).sum()
outlet_count
```

Out[400…

Outlet ID	number of times purchased	total sales
Outlet 100	1	2527.9824
Outlet 1000	4	3322.9860
Outlet 1001	2	1094.0780
Outlet 101	1	2527.9824
Outlet 102	2	811.3044
…	…	…
Outlet 992	4	5069.4356
Outlet 994	1	1053.3260
Outlet 996	2	1596.9124
Outlet 997	2	2188.1560

Outlet ID	number of times purchased	total sales
Outlet 999	4	4607.3532

3049 rows × 2 columns

In [401…
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[401…

number of times purchased	number of outlets	total sales
1	1666	2.573965e+06
2	948	2.804138e+06
3	270	1.100688e+06
4	116	5.001336e+05
5	28	1.678628e+05
6	15	9.802645e+04
7	2	1.763226e+04
8	3	2.589982e+04
9	1	6.843393e+03

In [402…
```
# frequncy of purchase of brand2
fp = len(brand9)/t_outlet
fp
```

Out[402… `0.8798829805541215`

In [403…
```
#total number of sales value
ts = brand9['Sales Value'].sum()
ts
```

Out[403… `7295188.8292`

In [404…
```
wfp = len(brand9)/len(outlet_count)
wfp
```

Out[404… `1.6769432600852738`

In [405…
```
frequency_dict['Brand Name'].append('Brand 9')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 10

In [406…
```
brand10 = data[data['Brand Name']=='Brand 10']
brand10
```

Out[406…

	Outlet ID	Brand Name	Sales Value	DATE
98	Outlet 23	Brand 10	4340.0880	2018-04-24
103	Outlet 24	Brand 10	1446.6960	2018-04-10
114	Outlet 26	Brand 10	1446.6960	2018-04-10
123	Outlet 28	Brand 10	1446.6960	2018-04-24
147	Outlet 35	Brand 10	289.3392	2018-04-10
…	…	…	…	…
25357	Outlet 23520	Brand 10	1446.6960	2018-04-30
25387	Outlet 23532	Brand 10	2893.3920	2018-04-28
25391	Outlet 23533	Brand 10	1446.6960	2018-04-11
25434	Outlet 23547	Brand 10	1446.6960	2018-04-13
25445	Outlet 23550	Brand 10	482.2320	2018-04-17

2863 rows × 4 columns

In [407…
```
# finding the outlets purchased for different number of times
outlet_count = brand10.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand10.groupby(['Outlet ID']).sum()
outlet_count
```

Out[407…

Outlet ID	number of times purchased	total sales
Outlet 100	1	4340.088
Outlet 1000	2	4340.088
Outlet 1001	1	1446.696
Outlet 101	1	4340.088
Outlet 1025	1	2893.392
…	…	…
Outlet 992	2	5786.784
Outlet 994	1	1446.696
Outlet 996	1	1446.696
Outlet 997	1	2893.392
Outlet 999	2	5786.784

2257 rows × 2 columns

In [408…
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
```

```
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[408...

number of times purchased	number of outlets	total sales
1	1714	2.578398e+06
2	486	1.377351e+06
3	51	2.077455e+05
4	6	3.684252e+04

In [409...

```
# frequncy of purchase of brand2
fp = len(brand10)/t_outlet
fp
```

Out[409... 0.49268628463259334

In [410...

```
#total number of sales value
ts = brand10['Sales Value'].sum()
ts
```

Out[410... 4200337.1663999995

In [411...

```
wfp = len(brand10)/len(outlet_count)
wfp
```

Out[411... 1.2684980062029243

In [412...

```
frequency_dict['Brand Name'].append('Brand 10')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

Brand 11

In [413...

```
brand11 = data[data['Brand Name']=='Brand 11']
brand11
```

Out[413...

	Outlet ID	Brand Name	Sales Value	DATE
11234	Outlet 2578	Brand 11	6197.7	2018-04-27

In [414...

```
# finding the outlets purchased for different number of times
outlet_count = brand11.groupby(['Outlet ID']).count()
outlet_count.drop(columns=['Sales Value','DATE'],inplace=True)
outlet_count.columns = ['number of times purchased']
outlet_count['total sales'] = brand11.groupby(['Outlet ID']).sum()
outlet_count
```

Out[414...

Outlet ID	number of times purchased	total sales

Outlet ID	number of times purchased	total sales
Outlet 2578	1	6197.7

In [415...]
```
f_count = outlet_count.groupby(['number of times purchased']).count()
f_count.columns = ['number of outlets']
f_count['total sales'] = outlet_count.groupby(['number of times purchased']).sum()
f_count
```

Out[415...]

number of times purchased	number of outlets	total sales
1	1	6197.7

In [416...]
```
# frequncy of purchase of brand2
fp = len(brand11)/t_outlet
fp
```

Out[416...]
```
0.00017208742040956807
```

In [417...]
```
#total number of sales value
ts = brand11['Sales Value'].sum()
ts
```

Out[417...]
```
6197.7
```

In [418...]
```
wfp = len(brand11)/len(outlet_count)
wfp
```

Out[418...]
```
1.0
```

In [419...]
```
frequency_dict['Brand Name'].append('Brand 11')
frequency_dict['Frequency of Purchase'].append(fp)
frequency_dict['Frequency of Purchase without'].append(wfp)
frequency_dict['Total Sales'].append(ts)
```

In [420...]
```
frequency_data = pd.DataFrame(frequency_dict)
frequency_data.sort_values('Frequency of Purchase')
```

Out[420...]

	Brand Name	Frequency of Purchase	Frequency of Purchase without	Total Sales
10	Brand 11	0.000172	1.000000	6.197700e+03
3	Brand 4	0.093960	1.344828	2.260473e+05
7	Brand 8	0.193082	1.108696	2.657290e+06
4	Brand 5	0.257443	1.150769	2.766144e+06
5	Brand 6	0.321459	1.364500	1.118381e+06
2	Brand 3	0.335743	1.107264	7.526337e+05

	Brand Name	Frequency of Purchase	Frequency of Purchase without	Total Sales
9	Brand 10	0.492686	1.268498	4.200337e+06
6	Brand 7	0.568061	1.666330	2.577246e+06
1	Brand 2	0.610394	1.522971	2.443848e+06
0	Brand 1	0.630700	1.718237	2.291894e+06
8	Brand 9	0.879883	1.676943	7.295189e+06

In [422…

```
frequency_data.to_excel('frequency of purchase.xlsx')
```

In [ ]: