```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt


df = pd.read_csv('/content/QVI_data.csv')


df.head()
```

	LYLTY_CARD_NBR	DATE	STORE_NBR	TXN_ID	PROD_NBR	PROD_NAME	PROD_QTY	T
0	1000	2018-10-17	1	1	5	Natural Chip Compny SeaSalt175g	2.0	
1	1002	2018-09-16	1	2	58	Red Rock Deli Chikn&Garlic Aioli 150g	1.0	
2	1003	2019-03-07	1	3	52	Grain Waves Sour Cream&Chives 210G	1.0	
3	1003	2019-03-08	1	4	106	Natural ChipCo Hony Soy Chckn175g	1.0	
4	1004	2018-11-02	1	5	96	WW Original Stacked Chips 160g	1.0	

To undo cell deletion use Ctrl+M Z or the 'Undo' option in the 'Edit' menu

```
df['DATE'].min(),df['DATE'].max()
```

```
('2018-07-01', '2019-06-30')
```

So this dataset consists of all the transactions that have taken place over the course of one year.(1st July 2018 - 30th June 2019)

```
df.dtypes
```

```
LYLTY_CARD_NBR        int64
DATE                 object
STORE_NBR             int64
TXN_ID                int64
```

```
TXN_ID              int64
PROD_NBR            int64
PROD_NAME          object
```

```
# The datatype of DATE is object so we have to convert that into datetime
df['Date'] = pd.to_datetime(df['DATE'])
```

```
BRAND               object
```

```
df['Date']
```

```
0       2018-10-17
1       2018-09-16
2       2019-03-07
3       2019-03-08
4       2018-11-02
           ...
97476   2019-01-06
97477   2019-03-06
97478   2019-03-28
97479   2019-06-14
97480   2018-07-09
Name: Date, Length: 97481, dtype: datetime64[ns]
```

```
df['Date'].dtype
```

```
dtype('<M8[ns]')
```

```
df['Month'] = df['Date'].dt.month # Extracting the month
```

To undo cell deletion use Ctrl+M Z or the 'Undo' option in the 'Edit' menu

```
12    8357
7     8315
8     8255
1     8252
3     8248
5     8247
10    8192
9     8093
6     8065
4     8050
11    7974
2     7433
Name: Month, dtype: int64
```

First we convert the Date from object to datetime and then extract the month from the date. Here 1 implies January 2 implied Feb and so on till 12(December). This shows us the total number of transactions that have taken place every month and the maximum transactions happened in January which is 8357.

```
df.groupby('Month')['TOT_SALES'].sum()
```

```
Month
1     61235.60
2     55541.50
```

```
3     61252.90
4     59702.10
5     58420.45
6     59875.00
7     61689.20
8     58907.95
9     60570.80
10    60915.70
11    58977.90
12    62148.50
Name: TOT_SALES, dtype: float64
```

```
len(df['LYLTY_CARD_NBR'].unique()) # The total number of customers are 27,419
```

```
27419
```

```
df.groupby('LYLTY_CARD_NBR')['TXN_ID'].count()  # This shows the number of transac
```

```
LYLTY_CARD_NBR
1000       1
1002       1
1003       2
1004       1
1005       1
          ..
100168     6
100169     7
100170    10
100171     7
```



```
df['LYLTY_CARD_NBR'].mean()
```

```
52128.87668366143
```

```
df.groupby('LYLTY_CARD_NBR')['TXN_ID'].count().mean() # The mean number of transac
```

```
3.5552354206936796
```

This shows the us the total number of sales every month. The maximum of sales happened in December. The maxium number of transactions happened in January but the total saleswas the highest in December.

```
df[df['STORE_NBR']==77] # There are 543 rows so 543 transactions have happened in S
```

	LYLTY_CARD_NBR	DATE	STORE_NBR	TXN_ID	PROD_NBR	PROD_NAME	PROD
73365	77000	2019-	77	74911	18	Cheetos Chs &	

		03-28				Bacon Balls 190g
73366	77000	2019-04-13	77	74912	69	Smiths Chip Thinly S/Cream&Onion 175g
73367	77000	2018-09-26	77	74910	36	Kettle Chilli 175g
73368	77001	2019-02-27	77	74913	7	Smiths Crinkle Original 330g
73369	77001	2019-01-21	77	74914	9	Kettle Tortilla ChpsBtroot&Ricotta 150g
...	...	...	...	...	...	...
73903	77499	2018-10-03	77	75460	30	Doritos Corn Chips Cheese Supreme 170g
73904	77500	2019-03-12	77	75461	81	Pringles Original Crisps 134g
73905	77501	2018-10-14	77	75462	40	Thins Chips Seasonedchicken 175g
73906	77502	2018-10-02	77	75463	94	Burger Rings 220g
		2019			69	Smiths Chip Thinly S/Cream&Onion 175g

To undo cell deletion use Ctrl+M Z or the 'Undo' option in the 'Edit' menu

```
df[df['STORE_NBR']==77]['TOT_SALES'].sum() # Total number of sales which have taker
```

```
2928.7
```

Similarly we can do the same type of analysis for Stores 86,88

To undo cell deletion use Ctrl+M Z or the 'Undo' option in the 'Edit' menu