### DATA CLEANING STARTS

```
In [1]: pip install pandoc
       Requirement already satisfied: pandoc in c:\users\ritika shukla\appdata\local\programs\python\python311\lib\site
       -packages (2.4)
       Requirement already satisfied: plumbum in c:\users\ritika shukla\appdata\local\programs\python\python311\lib\sit
       e-packages (from pandoc) (1.8.3)
       Requirement already satisfied: ply in c:\users\ritika shukla\appdata\local\programs\python\python311\lib\site-pa
       ckages (from pandoc) (3.11)
       Requirement already satisfied: pywin32 in c:\users\ritika shukla\appdata\local\programs\python\python311\lib\sit
       e-packages (from plumbum->pandoc) (306)
       Note: you may need to restart the kernel to use updated packages.
       [notice] A new release of pip is available: 24.0 -> 24.2
       [notice] To update, run: python.exe -m pip install --upgrade pip
In [2]: import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns
        import numpy as np
        import pandoc
In [3]: df=pd.read_csv(r'C:\Users\ritika shukla\Downloads\qua\purchase.csv')
        df.head()
Out[3]:
           LYLTY_CARD_NBR
                                          LIFESTAGE PREMIUM_CUSTOMER
        0
                             YOUNG SINGLES/COUPLES
                                                                 Premium
                             YOUNG SINGLES/COUPLES
        1
                       1002
                                                               Mainstream
        2
                       1003
                                     YOUNG FAMILIES
                                                                  Budget
        3
                       1004
                             OLDER SINGLES/COUPLES
                                                               Mainstream
        4
                       1005 MIDAGE SINGLES/COUPLES
                                                               Mainstream
In [4]: df.tail()
                                             LIFESTAGE PREMIUM CUSTOMER
              LYLTY_CARD_NBR
Out[4]:
        72632
                        2370651 MIDAGE SINGLES/COUPLES
                                                                  Mainstream
        72633
                        2370701
                                         YOUNG FAMILIES
                                                                   Mainstream
        72634
                        2370751
                                         YOUNG FAMILIES
                                                                    Premium
                                         OLDER FAMILIES
        72635
                        2370961
                                                                      Budget
                        2373711 YOUNG SINGLES/COUPLES
        72636
                                                                  Mainstream
In [5]: df.info()
       <class 'pandas.core.frame.DataFrame'>
       RangeIndex: 72637 entries, 0 to 72636
       Data columns (total 3 columns):
       # Column
                              Non-Null Count Dtype
       - - -
       0 LYLTY CARD NBR
                              72637 non-null int64
        1
           LIFESTAGE
                              72637 non-null object
           PREMIUM CUSTOMER 72637 non-null object
       dtypes: int64(1), object(2)
       memory usage: 1.7+ MB
In [6]: df1=pd.read_excel(r'C:\Users\ritika shukla\Downloads\qua\transaction.xlsx')
        df1.head()
        df1.info()
```

```
0
            DATE
                              264836 non-null
                                               datetime64[ns]
        1
             STORE NBR
                              264836 non-null
                                               int64
            LYLTY CARD NBR
        2
                             264836 non-null
                                               int64
        3
            TXN ID
                              264836 non-null
                                               int64
            PROD_NBR
                              264836 non-null int64
        4
            PROD NAME
                              264836 non-null
        5
                                               object
        6
            PROD QTY
                              264836 non-null
                                               int64
            TOT SALES
                              264836 non-null float64
       dtypes: datetime64[ns](1), float64(1), int64(5), object(1)
       memory usage: 16.2+ MB
In [7]: df merged=df.merge(df1[['DATE','STORE NBR','LYLTY CARD NBR', 'PROD NAME','PROD QTY','TOT SALES']])
         df_merged.head()
           LYLTY_CARD_NBR
                                    LIFESTAGE PREMIUM_CUSTOMER DATE STORE_NBR
                                                                                         PROD_NAME PROD_QTY TOT_SALES
                                                                                           Natural Chip
                                        YOUNG
                                                                      2018-
         0
                                                             Premium
                                                                                              Compny
                                                                                                               2
                                                                                                                          6.0
                        1000
                              SINGLES/COUPLES
                                                                      10-17
                                                                                           SeaSalt175a
                                                                                         Red Rock Deli
                                        YOUNG
                                                                      2018-
                                                                                          Chikn&Garlic
                                                                                                                          2.7
         1
                                                          Mainstream
                              SINGLES/COUPLES
                                                                      09-16
                                                                                             Aioli 150g
                                                                                          Grain Waves
                                                                      2019-
                                                                                                Sour
         2
                        1003
                               YOUNG FAMILIES
                                                              Budget
                                                                                                               1
                                                                                                                          36
                                                                                        Cream&Chives
                                                                      03-07
                                                                                                210G
                                                                                               Natural
                                                                                          ChipCo Hony
                                                                      2019-
         3
                               YOUNG FAMILIES
                        1003
                                                              Budget
                                                                                                               1
                                                                                                                          3.0
                                                                      03-08
                                                                                                 Sov
                                                                                           Chckn175g
                                                                                           WW Original
                                        OI DER
                                                                      2018-
         4
                                                          Mainstream
                                                                                         Stacked Chips
                                                                                                                          1.9
                              SINGLES/COUPLES
                                                                      11-02
                                                                                                 160g
In [8]: df_merged.info()
       <class 'pandas.core.frame.DataFrame'>
       RangeIndex: 264836 entries, 0 to 264835
       Data columns (total 8 columns):
        #
            Column
                                Non-Null Count
                                                  Dtype
                                _____
        0
            LYLTY CARD NBR
                                264836 non-null
                                                int64
                                264836 non-null
            LIFESTAGE
        1
                                                  object
        2
            PREMIUM CUSTOMER
                                264836 non-null
                                                  object
                                264836 non-null
        3
            DATE
                                                  datetime64[ns]
        4
            STORE NBR
                                264836 non-null
                                                  int64
        5
            PROD NAME
                                264836 non-null
                                                  object
        6
            PROD QTY
                                264836 non-null
                                                  int64
            TOT SALES
                                264836 non-null float64
       dtypes: datetime64[ns](1), float64(1), int64(3), object(3)
       memory usage: 16.2+ MB
In [9]: df merged['PROD NAME']=df merged['PROD NAME'].str.lower()
         df_merged.head()
Out[9]:
           LYLTY_CARD_NBR
                                    LIFESTAGE PREMIUM_CUSTOMER DATE STORE_NBR PROD_NAME PROD_QTY TOT_SALES
                                                                                           natural chip
                                        YOUNG
                                                                      2018-
                                                                                                              2
         0
                                                                                                                         6.0
                        1000
                                                             Premium
                                                                                              compny
                              SINGLES/COUPLES
                                                                      10-17
                                                                                           seasalt175g
                                                                                          red rock deli
                                        YOUNG
                                                                      2018-
         1
                                                                                                                         27
                                                          Mainstream
                                                                                      1
                                                                                          chikn&garlic
                                                                                                              1
                              SINGLES/COUPLES
                                                                      09-16
                                                                                            aioli 150g
                                                                                          grain waves
                                                                      2019-
                                                                                                sour
         2
                        1003
                               YOUNG FAMILIES
                                                              Budget
                                                                                                               1
                                                                                                                         3.6
                                                                                         cream&chives
                                                                      03-07
                                                                                                210g
                                                                                         natural chipco
                                                                      2019-
         3
                        1003
                               YOUNG FAMILIES
                                                                                                                         3.0
                                                              Budget
                                                                                            hony soy
                                                                      03-08
                                                                                           chckn175g
                                                                                           ww original
                                        OLDER
                                                                      2018-
                                                                                                                         1.9
         4
                                                          Mainstream
                                                                                         stacked chips
                                                                                                              1
                              SINGLES/COUPLES
                                                                      11-02
                                                                                                160a
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 264836 entries, 0 to 264835
Data columns (total 8 columns):

Non-Null Count

-----

Dtype

#

In [10]: regex='chip'

Column

```
df_merged=df_merged[df_merged['PROD_NAME'].str.contains(regex)==True]

Out[10]:

LYLTY_CARD_NBR

LIFESTAGE PREMIUM_CUSTOMER DATE STORE_NBR PROD_NAME PROD_QTY TOT_SALE

O 1000 YOUNG Premium 2018- 1 natural chip compny 2 6
```

onc[Tel:		LILIT_CARD_NBR	LIFESTAGE	PREMIUM_CUSTOMER	DATE	STORE_NBR	PROD_NAME	PROD_Q11	TOT_SALE
	0	1000	YOUNG SINGLES/COUPLES	Premium	2018- 10-17	1	natural chip compny seasalt175g	2	6
	3	1003	YOUNG FAMILIES	Budget	2019- 03-08	1	natural chipco hony soy chckn175g	1	3
	4	1004	OLDER SINGLES/COUPLES	Mainstream	2018- 11-02	1	ww original stacked chips 160g	1	1
	14	1011	OLDER SINGLES/COUPLES	Mainstream	2018- 12-19	1	smiths crinkle cut chips barbecue 170g	1	2
	17	1013	RETIREES	Budget	2019- 03-04	1	doritos corn chip southern chicken 150g	1	3
						***	•••		
	264814	2330251	RETIREES	Budget	2018- 11-29	77	smiths crnkle chip orgnl big bag 380g	2	11
	264818	2330291	OLDER SINGLES/COUPLES	Mainstream	2019- 06-18	77	ww d/style chip sea salt 200g	1	1
	264819	2330311	YOUNG SINGLES/COUPLES	Budget	2018- 11-09	77	tostitos smoked chipotle 175g	1	4
	264828	2370581	OLDER SINGLES/COUPLES	Budget	2018- 12-17	88	doritos corn chip southern chicken 150g	2	7
	264835	2373711	YOUNG SINGLES/COUPLES	Mainstream	2018- 12-14	88	smiths crinkle chips salt & vinegar 330g	2	11

74570 rows × 8 columns

In [11]: df\_merged['pos']=df\_merged['PROD\_NAME'].str.find(' ')
 df\_merged['BRAND']=df\_merged.apply(lambda x: x['PROD\_NAME'][0:x['pos']], axis=1)
 df\_merged.head()

Out[11]:	LYLTY_CARD_NBR		LIFESTAGE PREMIUM_CUSTO		DATE	STORE_NBR	PROD_NAME	PROD_QTY	TOT_SALES p
	0	1000	YOUNG SINGLES/COUPLES	Premium	2018- 10-17	1	natural chip compny seasalt175g	2	6.0
	3	1003	YOUNG FAMILIES	Budget	2019- 03-08	1	natural chipco hony soy chckn175g	1	3.0
	4	1004	OLDER SINGLES/COUPLES	Mainstream	2018- 11-02	1	ww original stacked chips 160g	1	1.9
	14	1011	OLDER SINGLES/COUPLES	Mainstream	2018- 12-19	1	smiths crinkle cut chips barbecue 170g	1	2.9
	17	1013	RETIREES	Budget	2019- 03-04	1	doritos corn chip southern chicken 150g	1	3.9
į	4								•

In [12]: df\_merged.drop(['LYLTY\_CARD\_NBR','pos'], axis='columns', inplace=True)
 df\_merged.head()

```
Out[12]:
                       LIFESTAGE PREMIUM_CUSTOMER
                                                            DATE STORE_NBR
                                                                                       PROD_NAME PROD_QTY TOT_SALES
                                                                                                                               BRAND
                           YOUNG
                                                            2018-
                                                                                  natural chip compny
           0
                                                 Premium
                                                                              1
                                                                                                               2
                                                                                                                                natural
                SINGLES/COUPLES
                                                            10-17
                                                                                         seasalt175g
                                                            2019-
                                                                                   natural chipco hony
                  YOUNG FAMILIES
           3
                                                   Budget
                                                                              1
                                                                                                                          3.0
                                                                                                                                natural
                                                            03-08
                                                                                       soy chckn175g
                           OLDER
                                                            2018-
                                                                                   ww original stacked
           4
                                               Mainstream
                                                                              1
                                                                                                               1
                                                                                                                          1.9
                                                                                                                                   ww
                SINGLES/COUPLES
                                                            11-02
                                                                                          chips 160g
                           OLDER
                                                            2018-
                                                                                     smiths crinkle cut
                                               Mainstream
           14
                                                                                                                          2.9
                                                                                                                                smiths
                SINGLES/COUPLES
                                                            12-19
                                                                                  chips barbecue 170g
                                                                                     doritos corn chip
                                                            2019-
          17
                         RETIREES
                                                   Budget
                                                                              1
                                                                                     southern chicken
                                                                                                               1
                                                                                                                          3.9
                                                                                                                                doritos
                                                            03-04
                                                                                               150g
In [13]: df_merged['pos']=df_merged['PROD_NAME'].str.find('chip')
          df_merged['new']=df_merged.apply(lambda x: x['PROD_NAME'][x['pos']: ], axis=1)
          df merged.head()
                      LIFESTAGE PREMIUM_CUSTOMER DATE STORE_NBR PROD_NAME PROD_QTY TOT_SALES
                                                                                                                     BRAND
                                                                                                                             pos
                                                                                natural chip
                         YOUNG
                                                         2018-
                                                                                                     2
                                                                                                                 6.0
                                                                                                                                8
                                                Premium
                                                                                                                      natural
                                                                                                                                       con
                                                                                   compny
               SINGLES/COUPLES
                                                          10-17
                                                                                seasalt175g
                                                                                                                                   seasalt:
                                                                              natural chipco
                                                                                                                                   chipco I
                                                         2019
                YOUNG FAMILIES
                                                                                                                 3.0
                                                                                                                                8
           3
                                                                           1
                                                                                  hony sov
                                                                                                     1
                                                                                                                      natural
                                                 Budget
                                                         03-08
                                                                                 chckn175g
                                                                                                                                    chckn<sup>3</sup>
                                                                                ww original
                                                         2018-
                          OI DER
                                             Mainstream
                                                                              stacked chips
                                                                                                     1
                                                                                                                 1.9
                                                                                                                               20
                                                                                                                                    chips
                                                                                                                         ww
               SINGLES/COUPLES
                                                          11-02
                                                                                      160g
                                                                              smiths crinkle
                          OLDER
                                                         2018-
                                                                                  cut chips
                                                                                                                 29
                                                                                                                               20
                                             Mainstream
                                                                           1
                                                                                                     1
                                                                                                                       smiths
                                                                                                                                      barbo
               SINGLES/COUPLES
                                                         12-19
                                                                                  barbecue
                                                                                      170g
                                                                                doritos com
                                                         2019-
                                                                                                                                      sout
                       RETIREES
          17
                                                                                                                               13
                                                                              chip southern
                                                                                                     1
                                                                                                                 3.9
                                                                                                                       doritos
                                                 Budget
                                                         03-04
                                                                                                                                       chi
                                                                               chicken 150a
In [14]: df_merged.drop(['PROD_NAME', 'pos'], axis='columns', inplace=True)
          df merged.head()
Out[14]:
                          LIFESTAGE PREMIUM_CUSTOMER
                                                                 DATE STORE_NBR
                                                                                     PROD_QTY
                                                                                                  TOT_SALES
                                                                                                               BRAND
                                                                                                                                  new
                                                              2018-10-
                                                                                                                           chip compny
           0
                                                                                               2
                                                    Premium
                                                                                                           6.0
                                                                                                                natural
                   SINGLES/COUPLES
                                                                                                                           seasalt175g
                                                                    17
                                                              2019-03-
                                                                                                                        chipco hony soy
           3
                     YOUNG FAMILIES
                                                      Budget
                                                                                                           3.0
                                                                                                                natural
                                                                    80
                                                                                                                            chckn175g
                              OLDER
                                                              2018-11-
           4
                                                  Mainstream
                                                                                                           1.9
                                                                                                                             chips 160g
                                                                                                                   ww
                   SINGLES/COUPLES
                                                                    02
                              OLDER
                                                              2018-12-
                                                                                                                         chips barbecue
                                                  Mainstream
           14
                                                                                                                smiths
                   SINGLES/COUPLES
                                                                    19
                                                                                                                                 170g
                                                              2019-03-
                                                                                                                          chip southern
          17
                           RETIREES
                                                      Budget
                                                                                                           3.9
                                                                                                                doritos
                                                                    04
                                                                                                                           chicken 150g
In [15]: df_merged['pos']=df_merged['new'].str.find(' ')
          df_merged['PRODUCTS']=df merged.apply(lambda x: x['new'][0:x['pos']], axis=1)
          df_merged.drop(['new','pos'], axis='columns', inplace=True)
          df merged.head()
                             LIFESTAGE PREMIUM CUSTOMER
                                                                    DATE
                                                                           STORE NBR
                                                                                        PROD QTY
                                                                                                     TOT SALES
                                                                                                                  BRAND
                                                                                                                          PRODUCTS
                                 YOUNG
                                                                  2018-10-
           0
                                                                                                  2
                                                      Premium
                                                                                      1
                                                                                                              6.0
                                                                                                                   natural
                                                                                                                                  chip
                      SINGLES/COUPLES
                                                                       17
                                                                  2019-03-
                       YOUNG FAMILIES
           3
                                                        Budget
                                                                                      1
                                                                                                              3.0
                                                                                                                   natural
                                                                                                                                chipco
                                                                       08
                                 OLDER
                                                                  2018-11-
                                                                                                  1
           4
                                                    Mainstream
                                                                                      1
                                                                                                              19
                                                                                                                                 chips
                                                                                                                      ww
                      SINGLES/COUPLES
                                 OLDER
                                                                  2018-12-
                                                    Mainstream
                                                                                                                    smiths
                                                                                                              2.9
                                                                                                                                 chips
           14
                                                                                      1
                      SINGLES/COUPLES
                                                                       19
                                                                  2019-03-
           17
                              RETIREES
                                                        Budget
                                                                                      1
                                                                                                              3.9
                                                                                                                   doritos
                                                                                                                                  chip
                                                                       04
```

In [16]: df\_merged['PRODUCTS']=df\_merged['PRODUCTS'].str.replace('chip','chips') df\_merged.head() LIFESTAGE PREMIUM\_CUSTOMER Out[16]: DATE STORE\_NBR PROD\_QTY TOT\_SALES **BRAND PRODUCTS** YOUNG 2018-10-0 Premium 2 1 6.0 natural chips SINGLES/COUPLES 17 2019-03-3 YOUNG FAMILIES Budget 3.0 natural chipsco 08 **OLDER** 2018-11-1 4 Mainstream 1 1.9 ww chipss SINGLES/COUPLES 02 **OLDER** 2018-12-Mainstream smiths chipss 14 29 SINGLES/COUPLES 19 2019-03chips 17 **RETIREES** Budget 1 1 doritos 39 In [17]: df\_merged=df\_merged[df\_merged['PRODUCTS']=='chips'] df merged.head() Out[17]: LIFESTAGE PREMIUM\_CUSTOMER DATE STORE\_NBR PROD\_QTY TOT\_SALES **BRAND PRODUCTS** YOUNG 2018-10-0 Premium 2 6.0 natural chips 1 SINGLES/COUPLES 17 2019-03-**RETIREES** 17 Budget 3.9 doritos chips 1 YOUNG 2019-02-30 Premium 1 1 natural chips 3.0 SINGLES/COUPLES 20 2018-11-31 YOUNG FAMILIES Budget 2 6.0 smiths chips 21 2018-10-YOUNG FAMILIES 40 Mainstream 1 1 3.0 natural chips 30 In [18]: df merged.tail() Out[18]:

	LIFESTAGE	PREMIUM_CUSTOMER	DATE	STORE_NBR	PROD_QTY	TOT_SALES	BRAND	PRODUCTS
264771	YOUNG SINGLES/COUPLES	Premium	2018-12- 27	272	2	7.8	doritos	chips
264811	RETIREES	Mainstream	2019-06- 20	77	2	6.0	natural	chips
264814	RETIREES	Budget	2018-11- 29	77	2	11.8	smiths	chips
264818	OLDER SINGLES/COUPLES	Mainstream	2019-06- 18	77	1	1.9	ww	chips
264828	OLDER SINGLES/COUPLES	Budget	2018-12- 17	88	2	7.8	doritos	chips

In [19]: df\_merged['month']=df\_merged['DATE'].dt.month df merged['year']=df merged['DATE'].dt.year df merged.head()

> C:\Users\ritika shukla\AppData\Local\Temp\ipykernel\_13888\191690799.py:1: 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#retu rning-a-view-versus-a-copy

df merged['month']=df merged['DATE'].dt.month

C:\Users\ritika shukla\AppData\Local\Temp\ipykernel 13888\191690799.py:2: 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#retu rning-a-view-versus-a-copy

df merged['year']=df merged['DATE'].dt.year

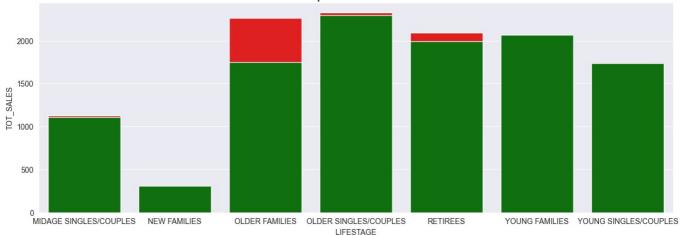
9]:		LIFESTAGE	PREMIUM_CUSTOMER	DATE	STORE_NBR	PROD_QTY	TOT_SALES	BRAND	PRODUCTS	month	year
	0	YOUNG SINGLES/COUPLES	Premium	2018- 10-17	1	2	6.0	natural	chips	10	2018
	17	RETIREES	Budget	2019- 03-04	1	1	3.9	doritos	chips	3	2019
	30	YOUNG SINGLES/COUPLES	Premium	2019- 02-20	1	1	3.0	natural	chips	2	2019
	31	YOUNG FAMILIES	Budget	2018- 11-21	1	2	6.0	smiths	chips	11	2018
	40	YOUNG FAMILIES	Mainstream	2018- 10-30	1	1	3.0	natural	chips	10	2018

## DATA CLEANING DONE

# **VISUALIZATION**

```
In [20]: df_merged.head()
                     LIFESTAGE PREMIUM_CUSTOMER DATE STORE_NBR PROD_QTY TOT_SALES BRAND PRODUCTS month
                                                                                                                                   vear
                         YOUNG
                                                         2018-
                                               Premium
                                                                                       2
                                                                                                  6.0
                                                                                                       natural
                                                                                                                     chips
                                                                                                                                10 2018
              SINGLES/COUPLES
                                                         10-17
                                                         2019-
          17
                       RETIREES
                                                 Budget
                                                                                                  3.9
                                                                                                        doritos
                                                                                                                     chips
                                                                                                                                3 2019
                                                         03-04
                         YOUNG
                                                         2019-
                                                                                       1
                                                                                                       natural
                                                                                                                                2 2019
          30
                                               Premium
                                                                                                  3.0
                                                                                                                     chips
               SINGLES/COUPLES
                                                         02-20
                                                         2018-
                YOUNG FAMILIES
                                                                                                                               11 2018
          31
                                                 Budget
                                                                                                  6.0
                                                                                                        smiths
                                                                                                                     chips
                                                         11-21
                                                         2018-
                YOUNG FAMILIES
                                             Mainstream
                                                                                                  3.0
                                                                                                       natural
                                                                                                                     chips
                                                                                                                               10 2018
                                                         10-30
In [21]: df 2018=df merged[df merged['year']==2018]
          df_2019=df_merged[df_merged['year']==2019]
          df_2018 = df_2018['TOT_SALES'].sum()
          df_2019=df_2019['TOT_SALES'].sum()
          print("Total Sales 2018 :", int(df_2018))
print("Total Sales 2019 :", int(df_2019))
         Total Sales 2018 : 67453
         Total Sales 2019 : 64695
In [22]: df_2018=df_merged[df_merged['year']==2018]
          df 2019=df merged[df merged['year']==2019]
          df 2018 qty=df 2018['PROD QTY'].sum()
          df_2019_qty=df_2019['PROD_QTY'].sum()
          df 2018=df 2018['TOT SALES'].sum()
          df 2019=df 2019['TOT SALES'].sum()
          print("Total Prod 2018 :", df_2018_qty)
print("Total Prod 2019 :", df_2019_qty)
          print("Avg Prod sold 2018 :", (df 2018) / (df 2018 qty))
          print("Avg Prod sold 2019 :", (df_2019) / (df_2019_qty))
         Total Prod 2018 : 18068
         Total Prod 2019 : 17373
         Avg Prod sold 2018 : 3.7333296435687404
         Avg Prod sold 2019 : 3.7239049099176875
In [30]: sns.set({'figure.figsize': (15,5)})
          df_2018=df_merged[df_merged['year']==2018]
          df_2019=df_merged[df_merged['year']==2019]
          df 2018=df 2018.groupby(['LIFESTAGE'])[['TOT SALES']].sum()
          df_2019=df_2019.groupby(['LIFESTAGE'])[['TOT_SALES']].sum()
          sns.barplot(x='LIFESTAGE',y='TOT_SALES', data=df_2018, color='red')
sns.barplot(x='LIFESTAGE',y='TOT_SALES', data=df_2019, color='green')
          plt.title("Sales Comparision 2018 v/s 2019", fontsize=20)
          plt.show()
```

#### Sales Comparision 2018 v/s 2019



```
In [31]: sns.set({'figure.figsize': (15,5)})

df_2018=df_merged[df_merged['year']==2018]
    df_2019=df_merged[df_merged['year']==2019]
    df_2018=df_2018.groupby(['month'])[['TOT_SALES']].sum()
    df_2019=df_2019.groupby(['month'])[['TOT_SALES']].sum()

sns.lineplot(x='month', y='TOT_SALES', data=df_2018, color='red')
    plt.title("MONTH_WISE_SALES_2018", fontsize=20)
    plt.show()
```

# 1950 1950 1850

month

10

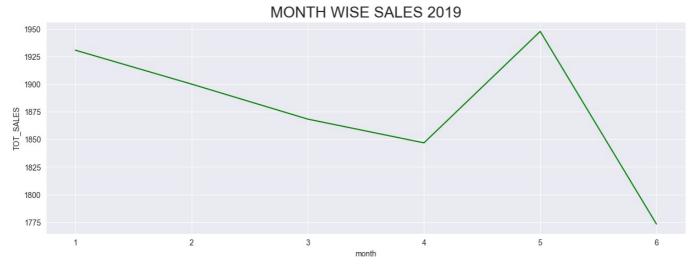
MONTH WISE SALES 2018

```
In [32]: sns.set({'figure.figsize': (15,5)})

df_2019=df_2019.groupby(['month'])[['TOT_SALES']].sum()
sns.lineplot(x='month', y='TOT_SALES', data=df_2019, color='green')
plt.title("MONTH WISE SALES 2019", fontsize=20)

plt.show()
```

8

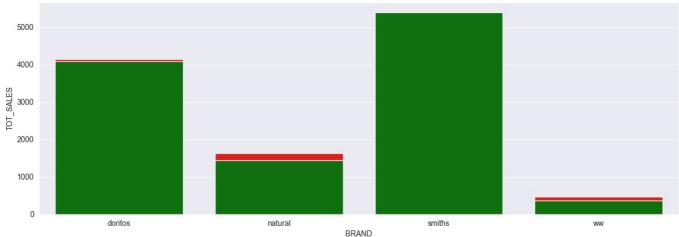


In [33]: sns.set({'figure.figsize': (15,5)})

```
df_2018_BRAND=df_merged[df_merged['year']==2018]
df_2019_BRAND=df_merged[df_merged['year']==2019]

df_2018_BRAND=df_2018_BRAND.groupby(['BRAND'])[['TOT_SALES']].sum()
df_2019_BRAND=df_2019_BRAND.groupby(['BRAND'])[['TOT_SALES']].sum()

sns.barplot(x='BRAND', y='TOT_SALES', data=df_2018_BRAND, color='red')
sns.barplot(x='BRAND', y='TOT_SALES', data=df_2019_BRAND, color='green')
plt.show()
```



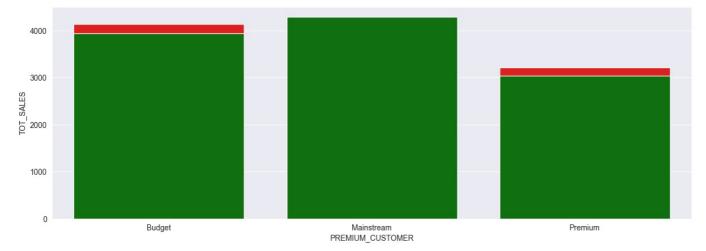
```
In [34]: sns.set({'figure.figsize': (15,5)})

df_2018_customer=df_merged[df_merged['year']==2018]
    df_2019_customer=df_merged[df_merged['year']==2019]

df_2018_customer=df_2018_customer.groupby(['PREMIUM_CUSTOMER'])[['TOT_SALES']].sum()
    df_2019_customer=df_2019_customer.groupby(['PREMIUM_CUSTOMER'])[['TOT_SALES']].sum()

sns.barplot(x='PREMIUM_CUSTOMER', y='TOT_SALES', data=df_2018_customer, color='red')
    sns.barplot(x='PREMIUM_CUSTOMER', y='TOT_SALES', data=df_2019_customer, color='green')

plt.show()
```



```
In [35]: sns.set({'figure.figsize': (15,5)})

df_2018_store = df_merged[df_merged['year']==2018]

df_2019_store = df_merged[df_merged['year']==2019]

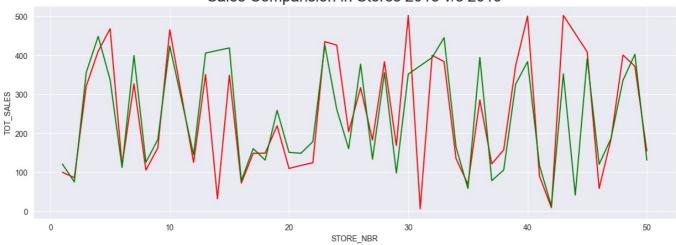
df_2018_store=df_2018_store.groupby(['STORE_NBR'])[['TOT_SALES']].sum()

df_2019_store=df_2019_store.groupby(['STORE_NBR'])[['TOT_SALES']].sum()

sns.lineplot(x='STORE_NBR', y='TOT_SALES', data=df_2018_store, color='red')
sns.lineplot(x='STORE_NBR', y='TOT_SALES', data=df_2019_store, color='green')

plt.title("Sales_Comparision_in_Stores_2018_v/s_2019", fontsize=20)
plt.show()
```

# Sales Comparision in Stores 2018 v/s 2019



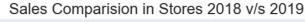
```
In [36]: sns.set({'figure.figsize': (15,5)})

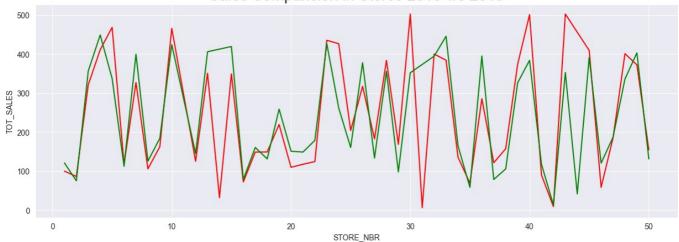
df_merged=df_merged[df_merged['STORE_NBR']<=50]
    df_2018_store = df_merged[df_merged['year']==2018]
    df_2019_store = df_merged[df_merged['year']==2019]

df_2018_store=df_2018_store.groupby(['STORE_NBR'])[['TOT_SALES']].sum()
    df_2019_store=df_2019_store.groupby(['STORE_NBR'])[['TOT_SALES']].sum()

sns.lineplot(x='STORE_NBR', y='TOT_SALES', data=df_2018_store, color='red')
    sns.lineplot(x='STORE_NBR', y='TOT_SALES', data=df_2019_store, color='green')

plt.title("Sales Comparision in Stores 2018 v/s 2019", fontsize=20)
    plt.show()</pre>
```







Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js