In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

In [5]:
```
import warnings
warnings.filterwarnings('ignore')
```

In [7]:
```
bp = pd.read_excel(r'C:\Users\kripal\Forsage\QVI_transaction_data.xlsx')
```

In [9]:
```
bp
```

Out[9]:

	DATE	STORE_NBR	LYLTY_CARD_NBR	TXN_ID	PROD_NBR	PROD_NAME	PI
0	43390	1	1000	1	5	Natural Chip Compny SeaSalt175g	
1	43599	1	1307	348	66	CCs Nacho Cheese 175g	
2	43605	1	1343	383	61	Smiths Crinkle Cut Chips Chicken 170g	
3	43329	2	2373	974	69	Smiths Chip Thinly S/Cream&Onion 175g	
4	43330	2	2426	1038	108	Kettle Tortilla ChpsHny&Jlpno Chili 150g	
...	...	...	...	...	...	...	
264831	43533	272	272319	270088	89	Kettle Sweet Chilli And Sour Cream 175g	
264832	43325	272	272358	270154	74	Tostitos Splash Of Lime 175g	
264833	43410	272	272379	270187	51	Doritos Mexicana 170g	
264834	43461	272	272379	270188	42	Doritos Corn Chip Mexican Jalapeno 150g	
264835	43365	272	272380	270189	74	Tostitos Splash Of Lime 175g	

264836 rows × 8 columns

In [13]:
```
bp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 264836 entries, 0 to 264835
Data columns (total 8 columns):
 #   Column          Non-Null Count   Dtype  
---  ------          --------------   -----  
 0   DATE            264836 non-null  int64  
 1   STORE_NBR       264836 non-null  int64  
 2   LYLTY_CARD_NBR  264836 non-null  int64  
 3   TXN_ID          264836 non-null  int64  
 4   PROD_NBR        264836 non-null  int64  
 5   PROD_NAME       264836 non-null  object 
 6   PROD_QTY        264836 non-null  int64  
 7   TOT_SALES       264836 non-null  float64
dtypes: float64(1), int64(6), object(1)
memory usage: 16.2+ MB
```

In [15]:
```
bp.isnull().sum()
```

Out[15]:
```
DATE              0
STORE_NBR         0
LYLTY_CARD_NBR    0
TXN_ID            0
PROD_NBR          0
PROD_NAME         0
PROD_QTY          0
TOT_SALES         0
dtype: int64
```

In [17]:
```
bp.dtypes
```

Out[17]:
```
DATE                int64
STORE_NBR           int64
LYLTY_CARD_NBR      int64
TXN_ID              int64
PROD_NBR            int64
PROD_NAME          object
PROD_QTY            int64
TOT_SALES         float64
dtype: object
```

In [19]:
```
bp.columns
```

Out[19]:
```
Index(['DATE', 'STORE_NBR', 'LYLTY_CARD_NBR', 'TXN_ID', 'PROD_NBR',
       'PROD_NAME', 'PROD_QTY', 'TOT_SALES'],
      dtype='object')
```

In [21]:
```
bp.describe()
```

Out[21]:

	DATE	STORE_NBR	LYLTY_CARD_NBR	TXN_ID	PROD_NBR	
count	264836.000000	264836.00000	2.648360e+05	2.648360e+05	264836.000000	26
mean	43464.036260	135.08011	1.355495e+05	1.351583e+05	56.583157	
std	105.389282	76.78418	8.057998e+04	7.813303e+04	32.826638	
min	43282.000000	1.00000	1.000000e+03	1.000000e+00	1.000000	
25%	43373.000000	70.00000	7.002100e+04	6.760150e+04	28.000000	
50%	43464.000000	130.00000	1.303575e+05	1.351375e+05	56.000000	
75%	43555.000000	203.00000	2.030942e+05	2.027012e+05	85.000000	
max	43646.000000	272.00000	2.373711e+06	2.415841e+06	114.000000	

In [33]:
```
print(bp['PROD_NAME'].unique())
```

```
['Natural Chip        Compny SeaSalt175g' 'CCs Nacho Cheese    175g'
 'Smiths Crinkle Cut  Chips Chicken 170g'
 'Smiths Chip Thinly  S/Cream&Onion 175g'
 'Kettle Tortilla ChpsHny&Jlpno Chili 150g'
 'Old El Paso Salsa   Dip Tomato Mild 300g'
 'Smiths Crinkle Chips Salt & Vinegar 330g'
 'Grain Waves         Sweet Chilli 210g'
 'Doritos Corn Chip Mexican Jalapeno 150g'
 'Grain Waves Sour    Cream&Chives 210G'
 'Kettle Sensations   Siracha Lime 150g' 'Twisties Cheese     270g'
 'WW Crinkle Cut      Chicken 175g' 'Thins Chips Light&  Tangy 175g'
 'CCs Original 175g' 'Burger Rings 220g'
 'NCC Sour Cream &    Garden Chives 175g'
 'Doritos Corn Chip Southern Chicken 150g' 'Cheezels Cheese Box 125g'
 'Smiths Crinkle      Original 330g'
 'Infzns Crn Crnchers Tangy Gcamole 110g'
 'Kettle Sea Salt     And Vinegar 175g'
 'Smiths Chip Thinly  Cut Original 175g' 'Kettle Original 175g'
 'Red Rock Deli Thai  Chilli&Lime 150g' 'Pringles Sthrn FriedChicken 134g'
 'Pringles Sweet&Spcy BBQ 134g' 'Red Rock Deli SR    Salsa & Mzzrlla 150g'
 'Thins Chips         Originl saltd 175g'
 'Red Rock Deli Sp    Salt & Truffle 150G'
 'Smiths Thinly       Swt Chli&S/Cream175G' 'Kettle Chilli 175g'
 'Doritos Mexicana    170g' 'Smiths Crinkle Cut  French OnionDip 150g'
 'Natural ChipCo      Hony Soy Chckn175g'
 'Dorito Corn Chp     Supreme 380g' 'Twisties Chicken270g'
 'Smiths Thinly Cut   Roast Chicken 175g'
 'Smiths Crinkle Cut  Tomato Salsa 150g'
 'Kettle Mozzarella   Basil & Pesto 175g'
 'Infuzions Thai SweetChili PotatoMix 110g'
 'Kettle Sensations   Camembert & Fig 150g'
 'Smith Crinkle Cut   Mac N Cheese 150g'
 'Kettle Honey Soy    Chicken 175g' 'Thins Chips Seasonedchicken 175g'
 'Smiths Crinkle Cut  Salt & Vinegar 170g'
 'Infuzions BBQ Rib   Prawn Crackers 110g'
 'GrnWves Plus Btroot & Chilli Jam 180g'
 'Tyrrells Crisps     Lightly Salted 165g'
 'Kettle Sweet Chilli And Sour Cream 175g'
 'Doritos Salsa       Medium 300g' 'Kettle 135g Swt Pot Sea Salt'
 'Pringles SourCream  Onion 134g' 'Doritos Corn Chips  Original 170g'
 'Twisties Cheese     Burger 250g'
 'Old El Paso Salsa   Dip Chnky Tom Ht300g'
 'Cobs Popd Swt/Chlli &Sr/Cream Chips 110g'
 'Woolworths Mild     Salsa 300g'
 'Natural Chip Co     Tmato Hrb&Spce 175g'
 'Smiths Crinkle Cut  Chips Original 170g'
 'Cobs Popd Sea Salt  Chips 110g'
 'Smiths Crinkle Cut  Chips Chs&Onion170g'
 'French Fries Potato Chips 175g'
 'Old El Paso Salsa   Dip Tomato Med 300g'
 'Doritos Corn Chips  Cheese Supreme 170g'
 'Pringles Original   Crisps 134g' 'RRD Chilli&         Coconut 150g'
 'WW Original Corn    Chips 200g' 'Thins Potato Chips  Hot & Spicy 175g'
 'Cobs Popd Sour Crm  &Chives Chips 110g'
 'Smiths Crnkle Chip  Orgnl Big Bag 380g'
 'Doritos Corn Chips  Nacho Cheese 170g'
 'Kettle Sensations   BBQ&Maple 150g' 'WW D/Style Chip     Sea Salt 200g'
 'Pringles Chicken    Salt Crips 134g' 'WW Original Stacked Chips 160g'
 'Smiths Chip Thinly  CutSalt/Vinegr175g' 'Cheezels Cheese 330g'
 'Tostitos Lightly    Salted 175g' 'Thins Chips Salt &  Vinegar 175g'
```

```
'Smiths Crinkle Cut  Chips Barbecue 170g' 'Cheetos Puffs 165g'
'RRD Sweet Chilli &  Sour Cream 165g' 'WW Crinkle Cut      Original 175g'
'Tostitos Splash Of  Lime 175g' 'Woolworths Medium   Salsa 300g'
'Kettle Tortilla ChpsBtroot&Ricotta 150g' 'CCs Tasty Cheese    175g'
'Woolworths Cheese   Rings 190g' 'Tostitos Smoked     Chipotle 175g'
'Pringles Barbeque   134g' 'WW Supreme Cheese   Corn Chips 200g'
'Pringles Mystery    Flavour 134g'
'Tyrrells Crisps     Ched & Chives 165g'
'Snbts Whlgrn Crisps Cheddr&Mstrd 90g' 'Cheetos Chs & Bacon Balls 190g'
'Pringles Slt Vingar 134g' 'Infuzions SourCream&Herbs Veg Strws 110g'
'Kettle Tortilla ChpsFeta&Garlic 150g'
'Infuzions Mango     Chutny Papadums 70g'
'RRD Steak &         Chimuchurri 150g' 'RRD Honey Soy       Chicken 165g'
'Sunbites Whlegrn    Crisps Frch/Onin 90g' 'RRD Salt & Vinegar  165g'
'Doritos Cheese      Supreme 330g' 'Smiths Crinkle Cut  Snag&Sauce 150g'
'WW Sour Cream &OnionStacked Chips 160g' 'RRD Lime & Pepper   165g'
'Natural ChipCo Sea  Salt & Vinegr 175g'
'Red Rock Deli Chikn&Garlic Aioli 150g'
'RRD SR Slow Rst     Pork Belly 150g' 'RRD Pc Sea Salt     165g'
'Smith Crinkle Cut   Bolognese 150g' 'Doritos Salsa Mild  300g']
```

In [51]:
```
bp['DATE'] = pd.to_numeric(bp['DATE'], errors='coerce')
bp['DATE'] = pd.to_datetime(bp['DATE'], origin='1899-12-30', unit='D')
```

In [53]:
```
bp
```

Out[53]:

	DATE	STORE_NBR	LYLTY_CARD_NBR	TXN_ID	PROD_NBR	PROD_NAME	PI
0	2018-10-17	1	1000	1	5	Natural Chip Compny SeaSalt175g	
1	2019-05-14	1	1307	348	66	CCs Nacho Cheese 175g	
2	2019-05-20	1	1343	383	61	Smiths Crinkle Cut Chips Chicken 170g	
3	2018-08-17	2	2373	974	69	Smiths Chip Thinly S/Cream&Onion 175g	
4	2018-08-18	2	2426	1038	108	Kettle Tortilla ChpsHny&Jlpno Chili 150g	
...	...	...	...	...	...	...	
264831	2019-03-09	272	272319	270088	89	Kettle Sweet Chilli And Sour Cream 175g	
264832	2018-08-13	272	272358	270154	74	Tostitos Splash Of Lime 175g	
264833	2018-11-06	272	272379	270187	51	Doritos Mexicana 170g	
264834	2018-12-27	272	272379	270188	42	Doritos Corn Chip Mexican Jalapeno 150g	
264835	2018-09-22	272	272380	270189	74	Tostitos Splash Of Lime 175g	

264836 rows × 8 columns

In [57]:
```
bp['DATE'].head()
```

Out[57]:
```
0   2018-10-17
1   2019-05-14
2   2019-05-20
3   2018-08-17
4   2018-08-18
Name: DATE, dtype: datetime64[ns]
```

In [63]:
```
bp['DATE'].dtype
```

Out[63]:
```
dtype('<M8[ns]')
```

In [75]:
```
# Extract pack size
bp['PACK_SIZE'] = bp['PROD_NAME'].str.extract(r'(\d+)g').astype(float)
```

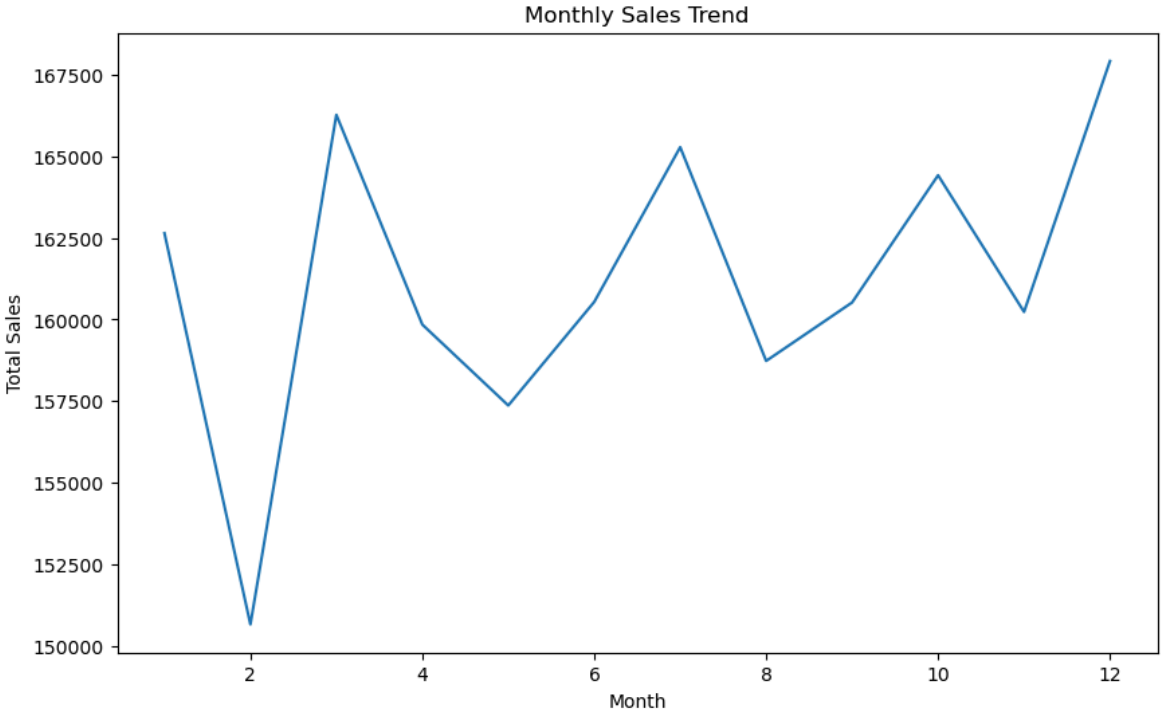
```
# Extract brand name
bp['BRAND'] = bp['PROD_NAME'].str.split().str[0]

# Extract transaction month and day of the week
bp['MONTH'] = bp['DATE'].dt.month
bp['DAY_OF_WEEK'] = bp['DATE'].dt.day_name()
```

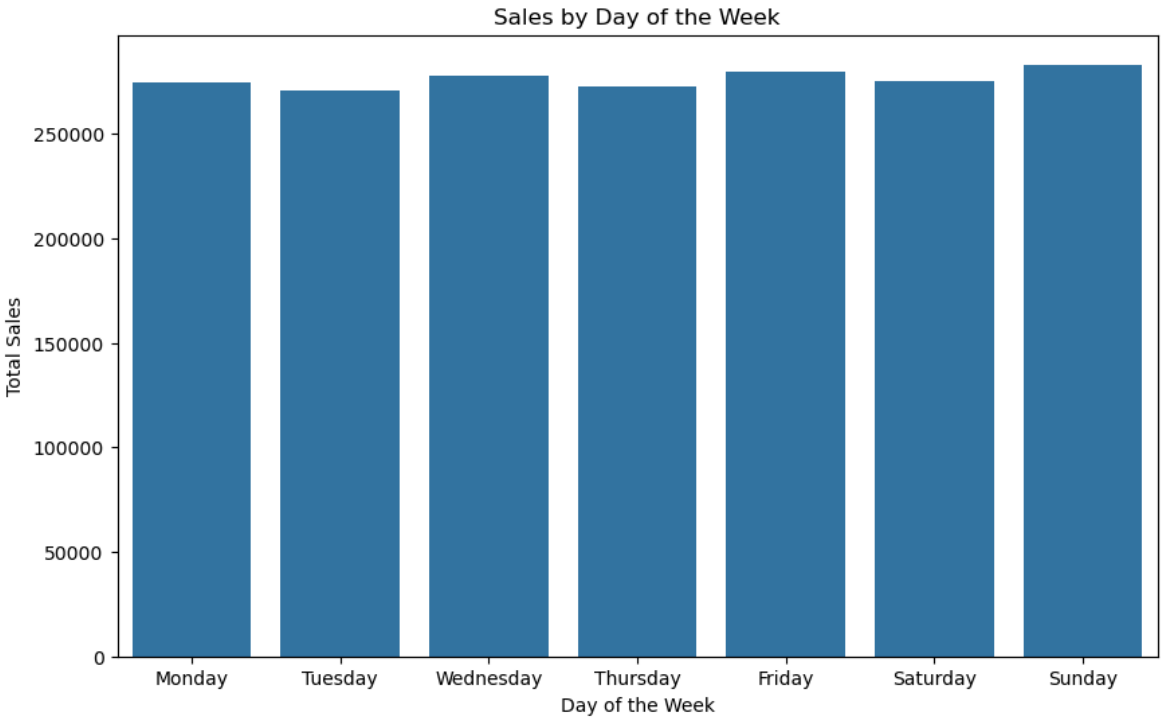
In [77]:
```
# Total spend by customer
customer_spend = bp.groupby('LYLTY_CARD_NBR')['TOT_SALES'].sum().reset_index()

# Sales trends by month
monthly_sales = bp.groupby('MONTH')['TOT_SALES'].sum().reset_index()

# Sales trends by day of the week
weekday_sales = bp.groupby('DAY_OF_WEEK')['TOT_SALES'].sum().reset_index()

# Top-selling brands
top_brands = bp['BRAND'].value_counts().head(10)

print("Top-Selling Brands:")
print(top_brands)
```

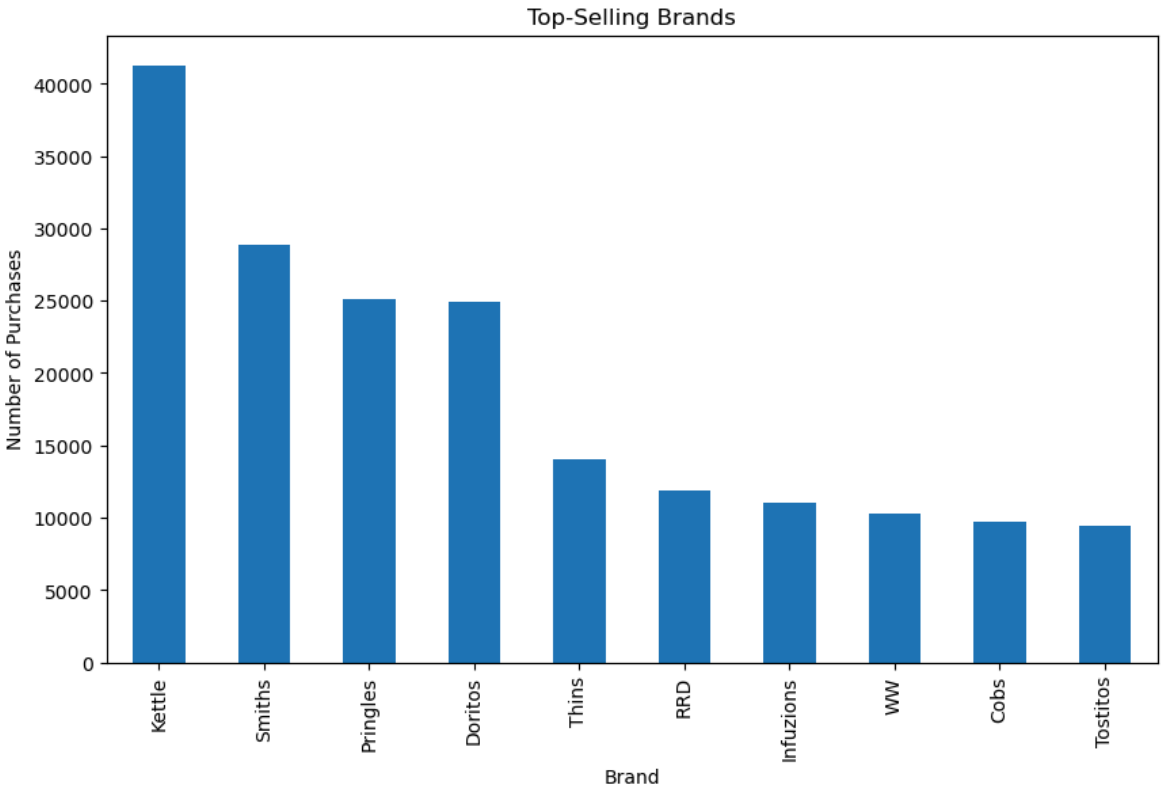
```
Top-Selling Brands:
BRAND
Kettle       41288
Smiths       28860
Pringles     25102
Doritos      24962
Thins        14075
RRD          11894
Infuzions    11057
WW           10320
Cobs          9693
Tostitos      9471
Name: count, dtype: int64
```

In [79]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Monthly sales trend
plt.figure(figsize=(10, 6))
sns.lineplot(data=monthly_sales, x='MONTH', y='TOT_SALES')
plt.title('Monthly Sales Trend')
plt.xlabel('Month')
plt.ylabel('Total Sales')
plt.show()

# Sales by day of the week
plt.figure(figsize=(10, 6))
sns.barplot(data=weekday_sales, x='DAY_OF_WEEK', y='TOT_SALES', order=['Monday',
plt.title('Sales by Day of the Week')
plt.xlabel('Day of the Week')
plt.ylabel('Total Sales')
plt.show()

# Top-selling brands
plt.figure(figsize=(10, 6))
top_brands.plot(kind='bar')
plt.title('Top-Selling Brands')
plt.xlabel('Brand')
```

```
plt.ylabel('Number of Purchases')
plt.show()
```






Top-Selling Brands
Number of Purchases
0
5000
10000
15000
20000
25000
30000
35000
40000
Kettle
Smiths
Pringles
Doritos
Thins
RRD
Infuzions
WW
Cobs
Tostitos
Brand

In [ ]: