```
#import the necessary laibraries
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

```
#create the dataframe
df=pd.read_excel('F:\Python\customer_retention_dataset.xlsx')
df
```

	Gender	Age	City	Pin Code	Duration of Shoping	Frequecy	Access	Device	Screen Size	OS	...	User satisfaction cannot exist without trust	Offering a wide variety	Complete and relevant product information	Monetary savings	Convenience of patronizing the online retailer	Sense of adventure	e-tailer enhances	Gratification shopping	Certain roles	Gett va mo sp
0	0	3	Delhi	110009	5	4	4	3	5	1	...	5	3	3	2	3	4	4	5	4	
1	1	2	Delhi	110030	5	5	2	1	2	3	...	5	5	5	5	3	3	3	5	5	
2	1	2	Greater Noida	201308	4	5	3	1	4	2	...	4	5	5	5	4	4	4	3	3	
3	0	2	Karnal	132001	4	1	3	1	4	3	...	4	4	4	5	4	4	5	4	3	
4	1	2	Bangalore	530068	3	2	2	1	2	3	...	5	4	4	5	4	3	1	5	1	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
264	1	2	Solan	173212	2	1	3	1	4	2	...	4	4	3	2	3	2	4	3	4	
265	1	3	Ghaziabad	201008	2	4	3	1	5	2	...	5	5	5	5	5	5	5	5	5	
266	1	4	Bangalore	560010	3	1	3	2	5	1	...	4	3	5	4	4	2	3	4	4	
267	1	1	Solan	173229	3	1	2	1	4	2	...	5	5	4	4	4	4	3	3	4	
268	1	4	Ghaziabad	201009	3	4	3	1	4	2	...	5	5	5	5	5	5	5	5	5	

269 rows × 47 columns

```
df.shape
```

```
(269, 47)
```

```
#check how many rows in our dataset
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 269 entries, 0 to 268
Data columns (total 47 columns):
 #   Column                                  Non-Null Count  Dtype
---  ------                                  --------------  -----
 0   Gender                                  269 non-null    int64
 1   Age                                     269 non-null    int64
 2   City                                    269 non-null    object
 3   Pin Code                                269 non-null    int64
 4   Duration of Shoping                     269 non-null    int64
 5   Frequecy                                269 non-null    int64
 6   Access                                  269 non-null    int64
 7   Device                                  269 non-null    int64
 8   Screen Size                             269 non-null    int64
 9   OS                                      269 non-null    int64
 10  Browser                                 269 non-null    int64
 11  Channel                                 269 non-null    int64
 12  How do you reach                        269 non-null    int64
 13  Purchase decision                       269 non-null    int64
 14  Payment Option                          269 non-null    int64
 15  How do you abandon                      269 non-null    int64
 16  Why did you abandon                     269 non-null    int64
 17  Content                                 269 non-null    int64
 18  Information for product comparison      269 non-null    int64
 19  Complete for purchase decision.         269 non-null    int64
 20  All relevant information                269 non-null    int64
 21  Ease of navigation                      269 non-null    int64
 22  Loading and processing speed            269 non-null    int64
 23  User friendly Interface                 269 non-null    int64
 24  Convenient Payment methods              269 non-null    int64
 25  Trust that the online retail store      269 non-null    int64
 26  Empathy towards the customers           269 non-null    int64
 27  Guarantee the privacy of the customer   269 non-null    int64
 28  Responsiveness                          269 non-null    int64
 29  Monetary benefit and discounts          269 non-null    int64
 30  Enjoyment                               269 non-null    int64
```

```
#check data types of each columns
df.dtypes
```

```
Gender                                                 int64
Age                                                    int64
City                                                  object
Pin Code                                               int64
Duration of Shoping                                    int64
Frequecy                                               int64
Access                                                 int64
Device                                                 int64
Screen Size                                            int64
OS                                                     int64
Browser                                                int64
Channel                                                int64
How do you reach                                       int64
Purchase decision                                      int64
Payment Option                                         int64
How do you abandon                                     int64
Why did you abandon                                    int64
Content                                                int64
Information for product comparison                     int64
Complete for purchase decision.                        int64
All relevant information                               int64
Ease of navigation                                     int64
Loading and processing speed                           int64
User friendly Interface                                int64
Convenient Payment methods                             int64
Trust that the online retail store                     int64
Empathy towards the customers                          int64
Guarantee the privacy of the customer                  int64
Responsiveness                                         int64
Monetary benefit and discounts                         int64
Enjoyment                                              int64
Convenient and flexible                                int64
Return and replacement policy                          int64
Loyalty programs                                       int64
Displaying quality Information                         int64
User derive satisfaction                               int64
Net Benefit derived                                    int64
User satisfaction cannot exist without trust           int64
Offering a wide variety                                int64
```

```
#check null values in data set
df.isnull().sum()
```

```
Gender                                                0
Age                                                   0
Pin Code                                              0
Duration of Shoping                                   0
Frequecy                                              0
Access                                                0
Device                                                0
Screen Size                                           0
OS                                                    0
Browser                                               0
Channel                                               0
How do you reach                                      0
Purchase decision                                     0
Payment Option                                        0
How do you abandon                                    0
Why did you abandon                                   0
Content                                               0
Information for product comparison                    0
Complete for purchase decision.                       0
All relevant information                              0
Ease of navigation                                    0
Loading and processing speed                          0
User friendly Interface                               0
Convenient Payment methods                            0
Trust that the online retail store                    0
Empathy towards the customers                         0
Guarantee the privacy of the customer                 0
Responsiveness                                        0
Monetary benefit and discounts                        0
Enjoyment                                             0
Convenient and flexible                               0
Return and replacement policy                         0
Loyalty programs                                      0
Displaying quality Information                        0
User derive satisfaction                              0
Net Benefit derived                                   0
User satisfaction cannot exist without trust          0
Offering a wide variety                               0
Complete and relevant product information             0
Monetary savings                                      0
Convenience of patronizing the online retailer        0
```

```
#Remove the city column from datafrme due to data type is object
df.drop('City',axis=1,inplace=True)

#Check the description of dataframe
df.describe()
```

	Gender	Age	Pin Code	Duration of Shoping	Frequecy	Access	Device	Screen Size	OS	Browser	...	User satisfaction cannot exist without trust	Offering a wide variety	Complete and relevant product information	Monetary savings	Convenience of patronizing the online retailer	Sense of adventure	e-tailer enhances	Grati sl
count	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	...	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269.000000	269
mean	0.669145	2.959108	220465.747212	3.524164	2.672862	3.260223	1.676580	4.282528	1.776952	1.275093	...	4.182156	4.148699	4.349442	4.263941	3.914498	3.553903	3.223048	3
std	0.471398	1.066012	140524.341051	1.436586	1.651788	1.135887	0.843904	0.923426	0.797892	0.645429	...	1.072162	0.842110	0.755953	1.000485	0.693879	1.065869	1.219581	1
min	0.000000	1.000000	110008.000000	1.000000	1.000000	2.000000	1.000000	2.000000	1.000000	1.000000	...	1.000000	2.000000	2.000000	2.000000	3.000000	1.000000	1.000000	1
25%	0.000000	2.000000	122018.000000	3.000000	1.000000	2.000000	1.000000	4.000000	1.000000	1.000000	...	4.000000	4.000000	4.000000	4.000000	3.000000	3.000000	3.000000	3
50%	1.000000	3.000000	201303.000000	4.000000	2.000000	3.000000	1.000000	4.000000	2.000000	1.000000	...	4.000000	4.000000	5.000000	5.000000	4.000000	4.000000	3.000000	3
75%	1.000000	4.000000	201310.000000	5.000000	4.000000	5.000000	2.000000	5.000000	2.000000	1.000000	...	5.000000	5.000000	5.000000	5.000000	4.000000	4.000000	4.000000	4
max	1.000000	5.000000	560037.000000	5.000000	5.000000	5.000000	4.000000	5.000000	3.000000	4.000000	...	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5

8 rows × 46 columns

```
#check how many columns in our data set
df.columns
```

```
Index(['Gender', 'Age', 'Pin Code', 'Duration of Shoping', 'Frequecy ',
       'Access', 'Device', 'Screen Size', 'OS', 'Browser', 'Channel',
       'How do you reach', 'Purchase decision', 'Payment Option',
       'How do you abandon', 'Why did you abandon', 'Content',
       'Information for product comparison', 'Complete for purchase decision.',
       'All relevant information', 'Ease of navigation',
       'Loading and processing speed', 'User friendly Interface',
       'Convenient Payment methods', 'Trust that the online retail store',
       'Empathy towards the customers',
       'Guarantee the privacy of the customer', 'Responsiveness',
       'Monetary benefit and discounts', 'Enjoyment',
       'Convenient and flexible', 'Return and replacement policy',
       'Loyalty programs', 'Displaying quality Information',
       'User derive satisfaction', 'Net Benefit derived',
       'User satisfaction cannot exist without trust',
       'Offering a wide variety', 'Complete and relevant product information',
       'Monetary savings', 'Convenience of patronizing the online retailer',
       'Sense of adventure', 'e-tailer enhances', 'Gratification shopping',
       'Certain roles', 'Getting value for money spent'],
      dtype='object')
```

```
#to visualize the data
sns.heatmap(df.isnull())
```