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

```
data = pd.read_csv('/content/telecom_customer_churn.csv')
```

```
data
```

	Customer ID	Gender	Age	Married	Number of Dependents	City	Zip Code	Latitude	Longitude
0	0002-ORFBO	Female	37	Yes	0	Frazier Park	93225	34.827662	-118.999073
1	0003-MKNFE	Male	46	No	0	Glendale	91206	34.162515	-118.203869
2	0004-TLHLJ	Male	50	No	0	Costa Mesa	92627	33.645672	-117.922613
3	0011-IGKFF	Male	78	Yes	0	Martinez	94553	38.014457	-122.115432
4	0013-EXCHZ	Female	75	Yes	0	Camarillo	93010	34.227846	-119.079903
...	...	...	...	...	...	...	...	...	...
7038	9987-LUTYD	Female	20	No	0	La Mesa	91941	32.759327	-116.997260
7039	9992-RRAMN	Male	40	Yes	0	Riverbank	95367	37.734971	-120.954271
7040	9992-UJOEL	Male	22	No	0	Elk	95432	39.108252	-123.645121
7041	9993-LHIEB	Male	21	Yes	0	Solana Beach	92075	33.001813	-117.263628
7042	9995-HOTOH	Male	36	Yes	0	Sierra City	96125	39.600599	-120.636358

7043 rows × 38 columns
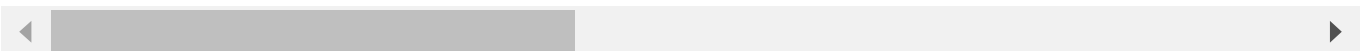
```
data.describe()
```

	Age	Number of Dependents	Zip Code	Latitude	Longitude	Number of Referrals	Ter
count	7043.000000	7043.000000	7043.000000	7043.000000	7043.000000	7043.000000	7043
mean	46.509726	0.468692	93486.070567	36.197455	-119.756684	1.951867	32
std	16.750352	0.962802	1856.767505	2.468929	2.154425	3.001199	24
min	19.000000	0.000000	90001.000000	32.555828	-124.301372	0.000000	1
25%	32.000000	0.000000	92101.000000	33.990646	-121.788090	0.000000	9
50%	46.000000	0.000000	93518.000000	36.205465	-119.595293	0.000000	29
75%	60.000000	0.000000	95329.000000	38.161321	-117.969795	3.000000	55
max	80.000000	9.000000	96150.000000	41.962127	-114.192901	11.000000	72

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 38 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   Customer ID                       7043 non-null   object
 1   Gender                            7043 non-null   object
 2   Age                               7043 non-null   int64
 3   Married                           7043 non-null   object
 4   Number of Dependents              7043 non-null   int64
 5   City                              7043 non-null   object
 6   Zip Code                          7043 non-null   int64
 7   Latitude                          7043 non-null   float64
 8   Longitude                         7043 non-null   float64
 9   Number of Referrals               7043 non-null   int64
 10  Tenure in Months                  7043 non-null   int64
 11  Offer                             3166 non-null   object
 12  Phone Service                     7043 non-null   object
 13  Avg Monthly Long Distance Charges 6361 non-null   float64
 14  Multiple Lines                    6361 non-null   object
 15  Internet Service                  7043 non-null   object
 16  Internet Type                     5517 non-null   object
 17  Avg Monthly GB Download           5517 non-null   float64
 18  Online Security                   5517 non-null   object
 19  Online Backup                     5517 non-null   object
 20  Device Protection Plan            5517 non-null   object
 21  Premium Tech Support              5517 non-null   object
 22  Streaming TV                      5517 non-null   object
 23  Streaming Movies                  5517 non-null   object
 24  Streaming Music                   5517 non-null   object
```

```
25  Unlimited Data                 5517 non-null   object
26  Contract                       7043 non-null   object
27  Paperless Billing              7043 non-null   object
28  Payment Method                 7043 non-null   object
29  Monthly Charge                 7043 non-null   float64
30  Total Charges                  7043 non-null   float64
31  Total Refunds                  7043 non-null   float64
32  Total Extra Data Charges       7043 non-null   int64
33  Total Long Distance Charges    7043 non-null   float64
34  Total Revenue                  7043 non-null   float64
35  Customer Status                7043 non-null   object
36  Churn Category                 1869 non-null   object
37  Churn Reason                   1869 non-null   object
dtypes: float64(9), int64(6), object(23)
memory usage: 2.0+ MB
```

```
data.isnull().sum()
```

	0
Customer ID	0
Gender	0
Age	0
Married	0
Number of Dependents	0
City	0
Zip Code	0
Latitude	0
Longitude	0
Number of Referrals	0
Tenure in Months	0
Offer	3877
Phone Service	0
Avg Monthly Long Distance Charges	682
Multiple Lines	682
Internet Service	0
Internet Type	1526
Avg Monthly GB Download	1526
Online Security	1526
Online Backup	1526
Device Protection Plan	1526
Premium Tech Support	1526
Streaming TV	1526
Streaming Movies	1526
Streaming Music	1526
Unlimited Data	1526
Contract	0
Paperless Billing	0
Payment Method	0
Monthly Charge	0

Total Charges	0
Total Refunds	0
Total Extra Data Charges	0
Total Long Distance Charges	0
Total Revenue	0
Customer Status	0
Churn Category	5174
Churn Reason	5174

dtype: int64

```
data.dropna(inplace=True)
```

```
data.isnull().sum()
```

	0
Customer ID	0
Gender	0
Age	0
Married	0
Number of Dependents	0
City	0
Zip Code	0
Latitude	0
Longitude	0
Number of Referrals	0
Tenure in Months	0
Offer	0
Phone Service	0
Avg Monthly Long Distance Charges	0
Multiple Lines	0
Internet Service	0
Internet Type	0
Avg Monthly GB Download	0
Online Security	0
Online Backup	0
Device Protection Plan	0
Premium Tech Support	0
Streaming TV	0
Streaming Movies	0
Streaming Music	0
Unlimited Data	0
Contract	0
Paperless Billing	0
Payment Method	0
Monthly Charge	0

Total Charges	0
Total Refunds	0
Total Extra Data Charges	0
Total Long Distance Charges	0
Total Revenue	0
Customer Status	0
Churn Category	0
Churn Reason	0

dtype: int64

data

	Customer ID	Gender	Age	Married	Number of Dependents	City	Zip Code	Latitude	Longitude
2	0004-TLHLJ	Male	50	No	0	Costa Mesa	92627	33.645672	-117.922613
3	0011-IGKFF	Male	78	Yes	0	Martinez	94553	38.014457	-122.115432
54	0094-OIFMO	Female	78	No	1	North Hollywood	91605	34.207295	-118.400022
67	0112-QWPNC	Male	61	Yes	0	Valyermo	93563	34.395830	-117.734568
77	0125-LZQXK	Male	42	No	0	Big Oak Flat	95305	37.818589	-120.256995
...	...	...	...	...	...	...	...	...	...
7007	9940-RHLFB	Female	53	No	0	Sunset Beach	90742	33.719221	-118.073596
7010	9944-HKVVB	Female	55	No	0	King City	93930	36.220761	-120.980777
7012	9947-OTFQU	Male	65	No	0	Covina	91722	34.097345	-117.906736
7036	9985-MWVIX	Female	53	No	0	Hume	93628	36.807595	-118.901544
7039	9992-RRAMN	Male	40	Yes	0	Riverbank	95367	37.734971	-120.954271

681 rows × 38 columns

```
data.columns
```

```
Index(['Customer ID', 'Gender', 'Age', 'Married', 'Number of Dependents',
       'City', 'Zip Code', 'Latitude', 'Longitude', 'Number of Referrals',
       'Tenure in Months', 'Offer', 'Phone Service',
       'Avg Monthly Long Distance Charges', 'Multiple Lines',
       'Internet Service', 'Internet Type', 'Avg Monthly GB Download',
       'Online Security', 'Online Backup', 'Device Protection Plan',
```

```
                'Premium Tech Support', 'Streaming TV', 'Streaming Movies',
                'Streaming Music', 'Unlimited Data', 'Contract', 'Paperless Billing',
                'Payment Method', 'Monthly Charge', 'Total Charges', 'Total Refunds',
                'Total Extra Data Charges', 'Total Long Distance Charges',
                'Total Revenue', 'Customer Status', 'Churn Category', 'Churn Reason'],
            dtype='object')
```
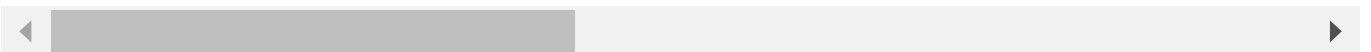
```
services = ['Phone Service', 'Multiple Lines', 'Internet Service', 'Online Security', 'Onlir
            'Device Protection Plan', 'Premium Tech Support', 'Streaming TV', 'Streaming Mo\
data['Total Services'] = data[services].apply(lambda x: x.str.contains('Yes').sum(), axis=1)
```

```
data['High Data User'] = data['Avg Monthly GB Download'].apply(lambda x: 1 if x > data['Avg
```

```
data['Referral Status'] = data['Number of Referrals'].apply(lambda x: 1 if x > 0 else 0)
```
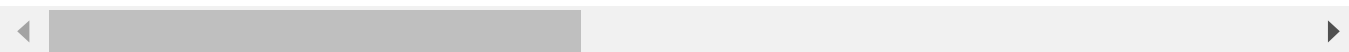
```
data.columns
```

```
Index(['Customer ID', 'Gender', 'Age', 'Married', 'Number of Dependents',
       'City', 'Zip Code', 'Latitude', 'Longitude', 'Number of Referrals',
       'Tenure in Months', 'Offer', 'Phone Service',
       'Avg Monthly Long Distance Charges', 'Multiple Lines',
       'Internet Service', 'Internet Type', 'Avg Monthly GB Download',
       'Online Security', 'Online Backup', 'Device Protection Plan',
       'Premium Tech Support', 'Streaming TV', 'Streaming Movies',
       'Streaming Music', 'Unlimited Data', 'Contract', 'Paperless Billing',
       'Payment Method', 'Monthly Charge', 'Total Charges', 'Total Refunds',
       'Total Extra Data Charges', 'Total Long Distance Charges',
       'Total Revenue', 'Customer Status', 'Churn Category', 'Churn Reason',
       'Total Services', 'High Data User', 'Referral Status'],
      dtype='object')
```

```
data
```

	Customer ID	Gender	Age	Married	Number of Dependents	City	Zip Code	Latitude	Longitude
2	0004-TLHLJ	Male	50	No	0	Costa Mesa	92627	33.645672	-117.922613
3	0011-IGKFF	Male	78	Yes	0	Martinez	94553	38.014457	-122.115432
54	0094-OIFMO	Female	78	No	1	North Hollywood	91605	34.207295	-118.400022
67	0112-QWPNC	Male	61	Yes	0	Valyermo	93563	34.395830	-117.734568
77	0125-LZQXK	Male	42	No	0	Big Oak Flat	95305	37.818589	-120.256995
...	...	...	...	...	...	...	...	...	...
7007	9940-RHLFB	Female	53	No	0	Sunset Beach	90742	33.719221	-118.073596
7010	9944-HKVVB	Female	55	No	0	King City	93930	36.220761	-120.980777
7012	9947-OTFQU	Male	65	No	0	Covina	91722	34.097345	-117.906736
7036	9985-MWVIX	Female	53	No	0	Hume	93628	36.807595	-118.901544
7039	9992-RRAMN	Male	40	Yes	0	Riverbank	95367	37.734971	-120.954271

681 rows × 41 columns

```python
from urllib.request import DataHandler
from sklearn.preprocessing import MinMaxScaler, StandardScaler

numerical_columns = ['Age', 'Number of Dependents', 'Latitude', 'Longitude',
                'Number of Referrals', 'Tenure in Months', 'Avg Monthly Long Distance (
                'Avg Monthly GB Download', 'Monthly Charge', 'Total Charges',
                'Total Refunds', 'Total Extra Data Charges', 'Total Long Distance Charg
```

```
                    'Total Revenue']

scaler = MinMaxScaler()

data[numerical_columns] = scaler.fit_transform(data[numerical_columns])

data.head()
```
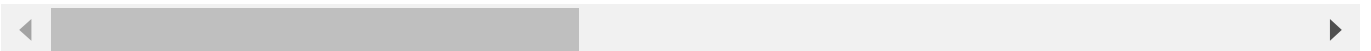
	Customer ID	Gender	Age	Married	Number of Dependents	City	Zip Code	Latitude	Longitud
2	0004-TLHLJ	Male	0.508197	No	0.00	Costa Mesa	92627	0.115863	0.71717
3	0011-IGKFF	Male	0.967213	Yes	0.00	Martinez	94553	0.580316	0.24119
54	0094-OIFMO	Female	0.967213	No	0.25	North Hollywood	91605	0.175570	0.66297
67	0112-QWPNC	Male	0.688525	Yes	0.00	Valyermo	93563	0.195614	0.73851
77	0125-LZQXK	Male	0.377049	No	0.00	Big Oak Flat	95305	0.559493	0.45216

5 rows × 41 columns

```
data.isnull().sum()
```

	0
Customer ID	0
Gender	0
Age	0
Married	0
Number of Dependents	0
City	0
Zip Code	0
Latitude	0
Longitude	0
Number of Referrals	0
Tenure in Months	0
Offer	0
Phone Service	0
Avg Monthly Long Distance Charges	0
Multiple Lines	0
Internet Service	0
Internet Type	0
Avg Monthly GB Download	0
Online Security	0
Online Backup	0
Device Protection Plan	0
Premium Tech Support	0
Streaming TV	0
Streaming Movies	0
Streaming Music	0
Unlimited Data	0
Contract	0