In [49]:
```
# Standard data science imports
import numpy as np
import pandas as pd
from pandas import Series, DataFrame
from numpy import arange

# Visualization libraries
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

# Scikit-learn
import sklearn
from sklearn import datasets
from sklearn import preprocessing
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.metrics import classification_report
from sklearn.metrics import mean_squared_error
from sklearn.metrics import r2_score
```

In [2]:
```
from sklearn.linear_model import Ridge
from sklearn.linear_model import RidgeCV
from sklearn.model_selection import RepeatedKFold
```

In [3]:
```
# Warning Code
import warnings
warnings.filterwarnings('ignore')
```

In [4]:
```
churn_df = pd.read_csv('churn_clean.csv', index_col=0)
```

In [5]:
```
# Examine dataset
churn_df.columns
```

Out[5]:
```
Index(['Customer_id', 'Interaction', 'UID', 'City', 'State', 'County', 'Zip',
       'Lat', 'Lng', 'Population', 'Area', 'TimeZone', 'Job', 'Children',
       'Age', 'Income', 'Marital', 'Gender', 'Churn', 'Outage_sec_perweek',
       'Email', 'Contacts', 'Yearly_equip_failure', 'Techie', 'Contract',
       'Port_modem', 'Tablet', 'InternetService', 'Phone', 'Multiple',
       'OnlineSecurity', 'OnlineBackup', 'DeviceProtection', 'TechSupport',
       'StreamingTV', 'StreamingMovies', 'PaperlessBilling', 'PaymentMethod',
       'Tenure', 'MonthlyCharge', 'Bandwidth_GB_Year', 'Item1', 'Item2',
       'Item3', 'Item4', 'Item5', 'Item6', 'Item7', 'Item8'],
      dtype='object')
```

In [6]:
```
churn_df.shape
```

Out[6]:
```
(10000, 49)
```

In [7]:
```
churn_df.head()
```

Out[7]:

CaseOrder	Customer_id	Interaction	UID	City	State	C
1	K409198	aa90260b-4141-4a24-8e36-b04ce1f4f77b	e885b299883d4f9fb18e39c75155d990	Point Baker	AK	Pri W
2	S120509	fb76459f-c047-4a9d-8af9-e0f7d4ac2524	f2de8bef964785f41a2959829830fb8a	West Branch	MI	Og
3	K191035	344d114c-3736-4be5-98f7-c72c281e2d35	f1784cfa9f6d92ae816197eb175d3c71	Yamhill	OR	Y
4	D90850	abfa2b40-2d43-4994-b15a-989b8c79e311	dc8a365077241bb5cd5ccd305136b05e	Del Mar	CA	
5	K662701	68a861fd-0d20-4e51-a587-8a90407ee574	aabb64a116e83fdc4befc1fbab1663f9	Needville	TX	

5 rows × 49 columns

In [8]:
```
churn_df.info
```

Out[8]:
```
<bound method DataFrame.info of          Customer_id
Interaction  \
CaseOrder
1             K409198  aa90260b-4141-4a24-8e36-b04ce1f4f77b
2             S120509  fb76459f-c047-4a9d-8af9-e0f7d4ac2524
3             K191035  344d114c-3736-4be5-98f7-c72c281e2d35
4              D90850  abfa2b40-2d43-4994-b15a-989b8c79e311
5             K662701  68a861fd-0d20-4e51-a587-8a90407ee574
...               ...                                   ...
9996          M324793  45deb5a2-ae04-4518-bf0b-c82db8dbe4a4
9997          D861732  6e96b921-0c09-4993-bbda-a1ac6411061a
9998          I243405  e8307ddf-9a01-4fff-bc59-4742e03fd24f
9999          I641617  3775ccfc-0052-4107-81ae-9657f81ecdf3
10000          T38070  9de5fb6e-bd33-4995-aec8-f01d0172a499

                                      UID          City State  \
CaseOrder
1           e885b299883d4f9fb18e39c75155d990   Point Baker    AK
2           f2de8bef964785f41a2959829830fb8a   West Branch    MI
3           f1784cfa9f6d92ae816197eb175d3c71       Yamhill    OR
4           dc8a365077241bb5cd5ccd305136b05e       Del Mar    CA
5           aabb64a116e83fdc4befc1fbab1663f9     Needville    TX
...                                      ...           ...   ...
9996        9499fb4de537af195d16d046b79fd20a   Mount Holly    VT
9997        c09a841117fa81b5c8e19afec2760104   Clarksville    TN
9998        9c41f212d1e04dca84445019bbc9b41c      Mobeetie    TX
9999        3e1f269b40c235a1038863ecf6b7a0df    Carrollton    GA
10000       0ea683a03a3cd544aefe8388aab16176  Clarkesville    GA

                          County    Zip       Lat        Lng  Population  ...
\
CaseOrder                                                                  ...
1           Prince of Wales-Hyder  99927  56.25100 -133.37571          38  ...
2                         Ogemaw  48661  44.32893  -84.24080       10446  ...
3                        Yamhill  97148  45.35589 -123.24657        3735  ...
4                      San Diego  92014  32.96687 -117.24798       13863  ...
5                      Fort Bend  77461  29.38012  -95.80673       11352  ...
...                          ...    ...       ...        ...         ...  ...
9996                     Rutland   5758  43.43391  -72.78734         640  ...
9997                   Montgomery  37042  36.56907  -87.41694       77168  ...
9998                      Wheeler  79061  35.52039 -100.44180         406  ...
9999                      Carroll  30117  33.58016  -85.13241       35575  ...
10000                   Habersham  30523  34.70783  -83.53648       12230  ...

            MonthlyCharge Bandwidth_GB_Year Item1  Item2  Item3  Item4 Item5  \
CaseOrder
1              172.455519        904.536110     5      5      5      3     4
2              242.632554        800.982766     3      4      3      3     4
3              159.947583       2054.706961     4      4      2      4     4
4              119.956840       2164.579412     4      4      4      2     5
5              149.948316        271.493436     4      4      4      3     4
...                   ...               ...   ...    ...    ...    ...   ...
9996           159.979400       6511.252601     3      2      3      3     4
9997           207.481100       5695.951810     4      5      5      4     4
9998           169.974100       4159.305799     4      4      4      4     4
9999           252.624000       6468.456752     4      4      6      4     3
10000          217.484000       5857.586167     2      2      3      3     3
```

```
          Item6 Item7  Item8
CaseOrder
1             4     3      4
2             3     4      4
3             3     3      3
4             4     3      3
5             4     4      5
...         ...   ...    ...
9996          3     2      3
9997          5     2      5
9998          4     4      5
9999          3     5      4
10000         3     4      1

[10000 rows x 49 columns]>
```

In [9]:
```
churn_df.describe()
```

Out[9]:

	Zip	Lat	Lng	Population	Children	Age	
count	10000.000000	10000.000000	10000.000000	10000.000000	10000.0000	10000.000000	100(
mean	49153.319600	38.757567	-90.782536	9756.562400	2.0877	53.078400	398(
std	27532.196108	5.437389	15.156142	14432.698671	2.1472	20.698882	281(
min	601.000000	17.966120	-171.688150	0.000000	0.0000	18.000000	3(
25%	26292.500000	35.341828	-97.082813	738.000000	0.0000	35.000000	192(
50%	48869.500000	39.395800	-87.918800	2910.500000	1.0000	53.000000	331(
75%	71866.500000	42.106908	-80.088745	13168.000000	3.0000	71.000000	532(
max	99929.000000	70.640660	-65.667850	111850.000000	10.0000	89.000000	2589(

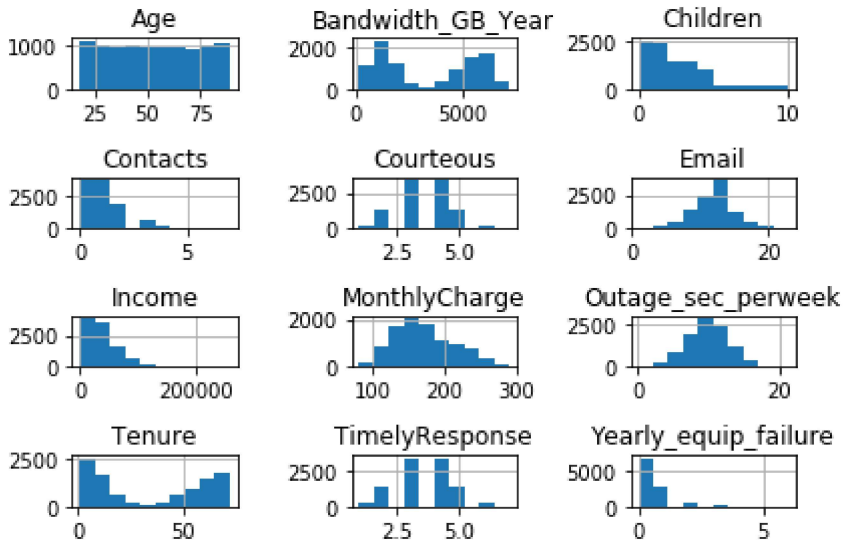
8 rows × 22 columns

In [10]:
```
churn_df.dtypes
```

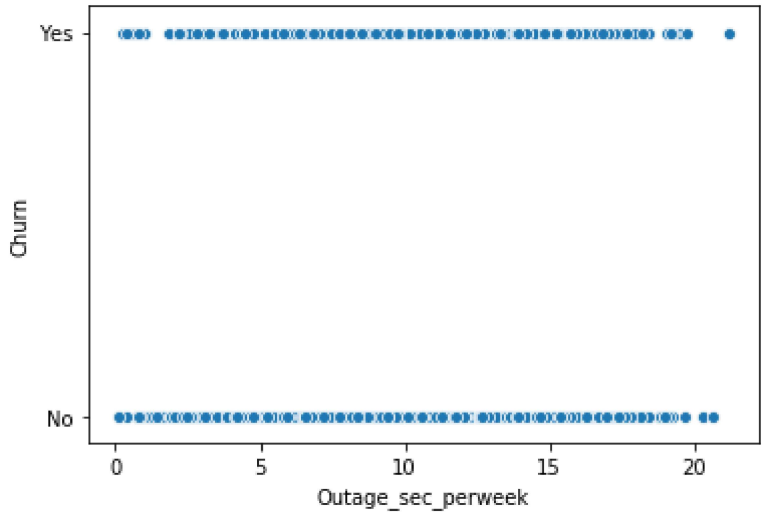
Out[10]:
```
Customer_id             object
Interaction             object
UID                     object
City                    object
State                   object
County                  object
Zip                      int64
Lat                    float64
Lng                    float64
Population               int64
Area                    object
TimeZone                object
Job                     object
Children                 int64
Age                      int64
Income                 float64
Marital                 object
Gender                  object
Churn                   object
Outage_sec_perweek     float64
Email                    int64
Contacts                 int64
Yearly_equip_failure     int64
Techie                  object
Contract                object
Port_modem              object
Tablet                  object
InternetService         object
Phone                   object
Multiple                object
OnlineSecurity          object
OnlineBackup            object
DeviceProtection        object
TechSupport             object
StreamingTV             object
StreamingMovies         object
PaperlessBilling        object
PaymentMethod           object
Tenure                 float64
MonthlyCharge          float64
Bandwidth_GB_Year      float64
Item1                    int64
Item2                    int64
Item3                    int64
Item4                    int64
Item5                    int64
Item6                    int64
Item7                    int64
Item8                    int64
dtype: object
```

In [11]:
```
# Rename survey items
churn_df.rename(columns = {'Item1':'TimelyResponse',
 'Item2':'Fixes',
 'Item3':'Replacements',
 'Item4':'Reliability',
 'Item5':'Options',
 'Item6':'Respectful',
 'Item7':'Courteous',
 'Item8':'Listening'},
 inplace=True)
```

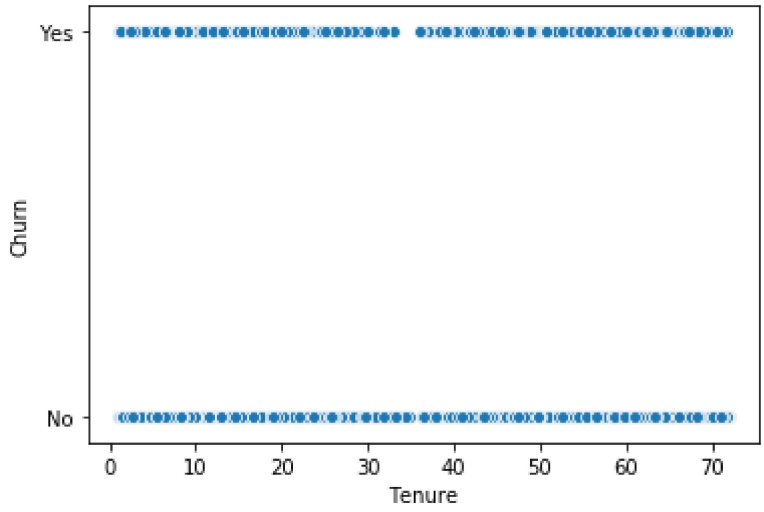
In [12]:
```
# Create histograms
churn_df[['Children', 'Age', 'Income', 'Outage_sec_perweek', 'Email',
 'Contacts', 'Yearly_equip_failure', 'Tenure', 'MonthlyCharge',
 'Bandwidth_GB_Year', 'TimelyResponse', 'Courteous']].hist()
plt.savefig('churn_pyplot.jpg')
plt.tight_layout()
```



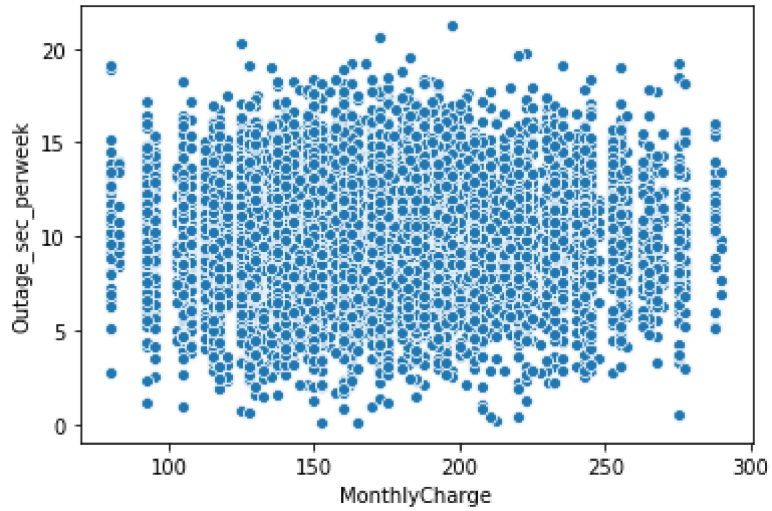
In [13]:
```
# Crerate scatterplot Outage_sec_perweek
sns.scatterplot(x=churn_df['Outage_sec_perweek'], y=churn_df['Churn'])
plt.show()
```



In [14]:
```
# Crerate scatterplot Tenure
sns.scatterplot(x=churn_df['Tenure'], y=churn_df['Churn'])
plt.show()
```



In [15]:
```
# Crerate scatterplot MonthlyCharge to Outage_sec_perweek
sns.scatterplot(x=churn_df['MonthlyCharge'], y=churn_df['Outage_sec_perweek'])
plt.show()
```



In [16]:
```
# Scatter matrix of nuerical variables
churn_numeric = churn_df[['Children', 'Age', 'Income', 'Outage_sec_perweek',
 'Email', 'Contacts','Yearly_equip_failure', 'Tenure',
 'MonthlyCharge', 'Bandwidth_GB_Year', 'Replacements',
 'Reliability', 'Options', 'Respectful', 'Courteous',
 'Listening']]
pd.plotting.scatter_matrix(churn_numeric, figsize = [15, 15])
```

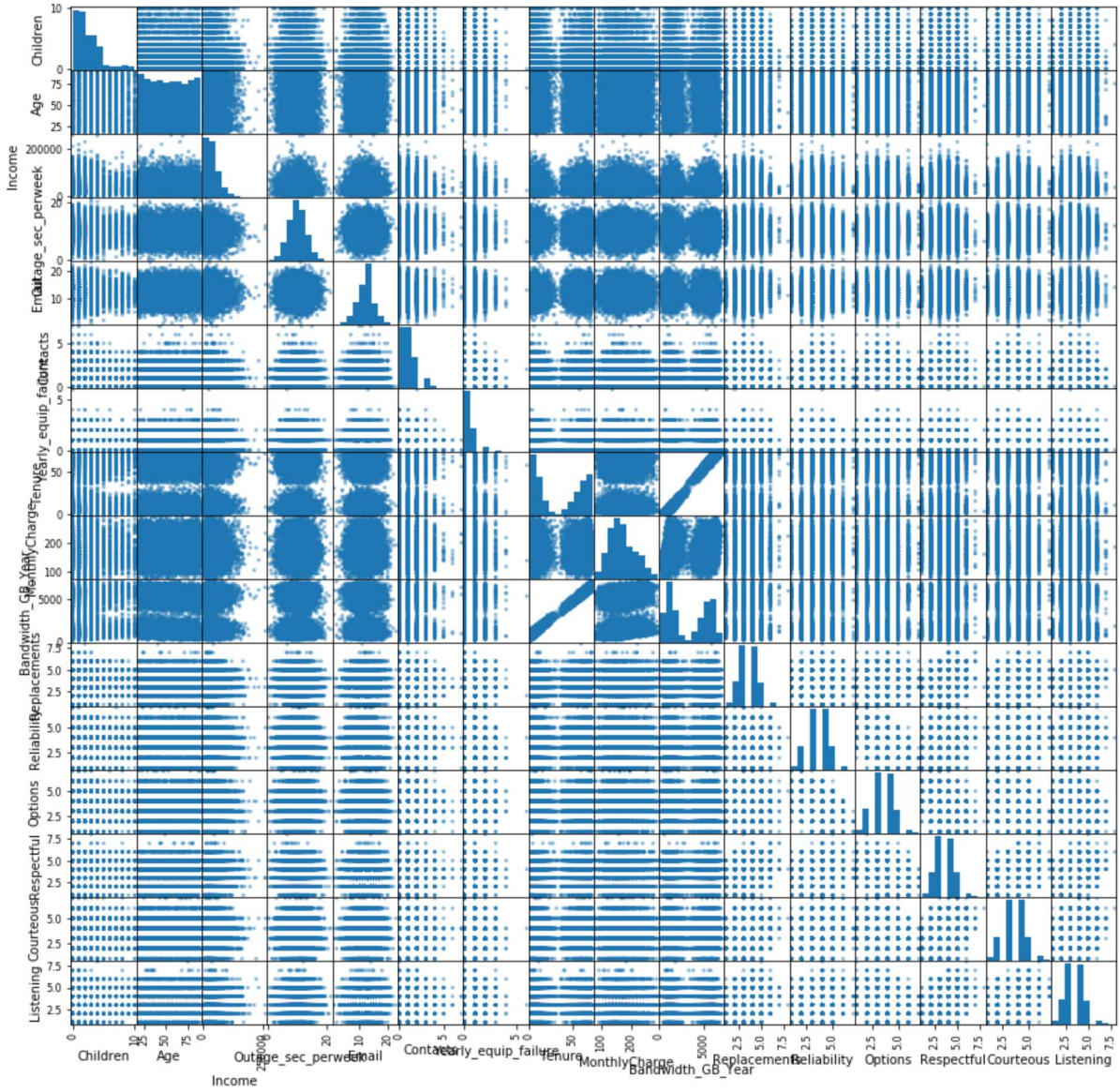
Out[16]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD38A0F08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD38EAD08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3922308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3958C88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3991A48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD39CAB88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3A02C48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3A58F08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3A58FC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3A97208>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3AE90C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3B22148>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3B5A2C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3B923C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3BCC488>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3C04608
       >],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3C3D748>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3C74748>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3CAF8C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3CE9A08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3D21A88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3D5BC08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3D93CC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3DCAD48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3E02E88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3E3E048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3E77108>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD38B2688>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD386DA48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD37F3108>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3789148>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD36B48C8
       >],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD36FA288>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3697088>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD352CF88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD365FEC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD353CE48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3790BC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD388C6C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD350BD88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3548E48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD35C4508>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3EC4048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD349E188>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD346E048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD2B8C348>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD342ACC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD32B6208
       >],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3341508>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD336D788>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD33EF8C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3386948>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD32EEA88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3333B88>,
```

```
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3EF0C08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3F29D08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3F62E88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3FB8E88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD3FD7048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD40111C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4049308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4081388>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD40BA508>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD40F3648
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4129648>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD41657C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD419E908>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD41D7988>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD420EB08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4249BC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4280C48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD42B9DC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD42F2F08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4347E88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD43670C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD439F248>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD43D8388>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD44114C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4449608>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD4483708
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD56DC848>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5714988>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD574CB08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5785C48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD57BED88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD57F5E88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD582FFC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD586E148>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD58A52C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD58F9EC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5916508>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD594E688>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD59867C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD59C0908>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD59FAA88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5A31B48
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5A6AC88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5AA4E08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5ADDF48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5B170C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5B50248>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5B88348>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5BC2488>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5BFA5C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5C4EE88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5C6D888>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5CA6988>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5CDEB08>,
```

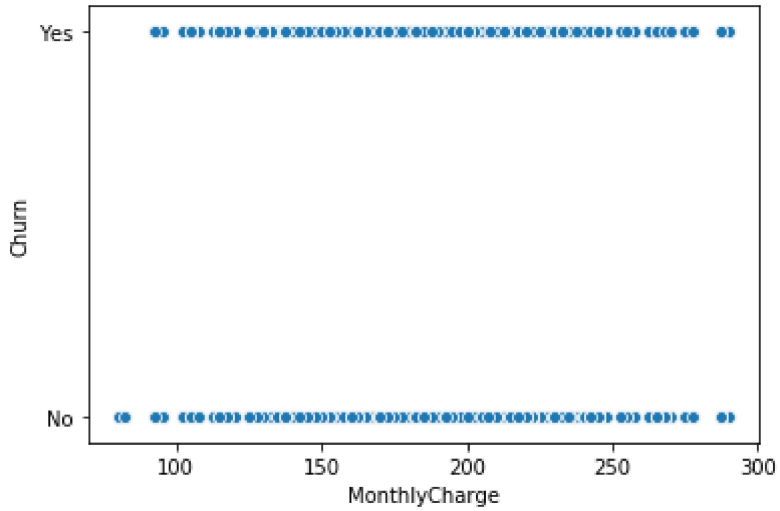
```
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5D17C48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5D50D88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5D87F08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5DC1FC8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5DFD148>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5E362C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5E6D408>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5EA7548>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5EE0648>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5F18788>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5F528C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5F8AA48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5FDEE88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD5FFAC88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6034DC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD606EF48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD60A90C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD60E3208>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD611C308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6154448
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD618A588>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD61C4708>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD61FD848>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6234988>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD626FAC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD62A7C08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD62E0D48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6319EC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6357048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6390148>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD63C62C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6400408>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6438548>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD64716C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD64A8708>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD64E3848
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD651C9C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6553B08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD658CC48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD65C4D88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD65FBEC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6639048>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6675188>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD66AD308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD66E5448>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD671E548>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD67576C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD678E808>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD67C7948>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6800AC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6838B88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6871CC8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD68AAE48>,
```

```
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD68E2F88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD691E108>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6958208>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6992388>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD69CB4C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6A01608>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6A3B788>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6A74848>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6AAC988>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6AE5B08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6B1DC48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6B56D88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6B8FE88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6BC8FC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6C04148
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6C3C2C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6C76408>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6CAF548>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6CE7688>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6D1F7C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6D59908>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6D92A88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6DE2EC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6E02CC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6E3BE08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6E75F88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6EB0108>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6EE9248>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6F22348>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6F59488>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6F935C8
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD6FCC748>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7005888>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD703D988>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7077B08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD70B0C48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD70E7D88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7138E88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7159FC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7196148>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD71CD2C8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7207408>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7240548>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7279688>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD72AF708>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD72EA848>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD73239C8
    >],
   [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD735AB08>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7393C48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD73CBD48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7404E88>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD743EFC8>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD747C188>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD74B32C8>,
```

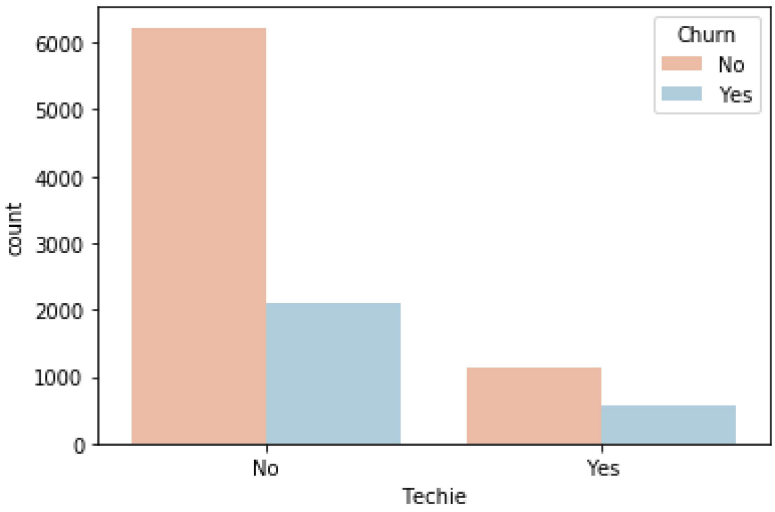
```
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD74EB3C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7525548>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD755E688>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD75937C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD75CD948>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7608A08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD763FB48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7679C88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD76B2E08
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD76EAF48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD77250C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD775F208>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7799348>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD77D0488>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7809608>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7843748>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD787C848>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD78B49C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD78EBB08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7924C48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD795EDC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7996E88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD79CEFC8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7A0A188>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7A432C8
>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7A7D408>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7AB7508>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7AEF648>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7B27788>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7B608C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7B98A48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7BD0B48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7C07C88>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7C41E08>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7C7BF48>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7CB80C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7CF21C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7D29308>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7D62448>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7D9A5C8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x000001EDD7DD3708
>]],
      dtype=object)
```



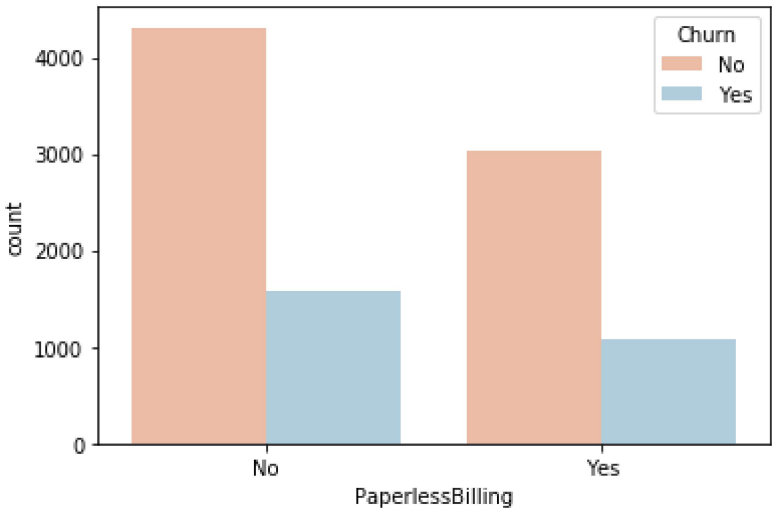
In [17]:
```
# Scatterplot MonthlyCharge to churn
sns.scatterplot(x = churn_df['MonthlyCharge'], y = churn_df['Churn'])
plt.show()
```



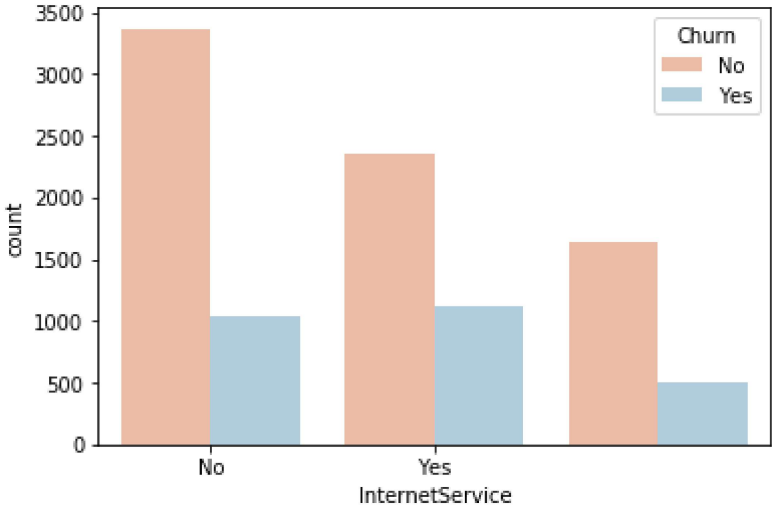
In [18]:
```
# Binary dataset
plt.figure()
sns.countplot(x='Techie', hue='Churn', data=churn_df, palette='RdBu')
plt.xticks([0,1], ['No', 'Yes'])
plt.show()
```



In [19]:
```
plt.figure()
sns.countplot(x='PaperlessBilling', hue='Churn', data=churn_df, palette='RdBu'
)
plt.xticks([0,1], ['No', 'Yes'])
plt.show()
```

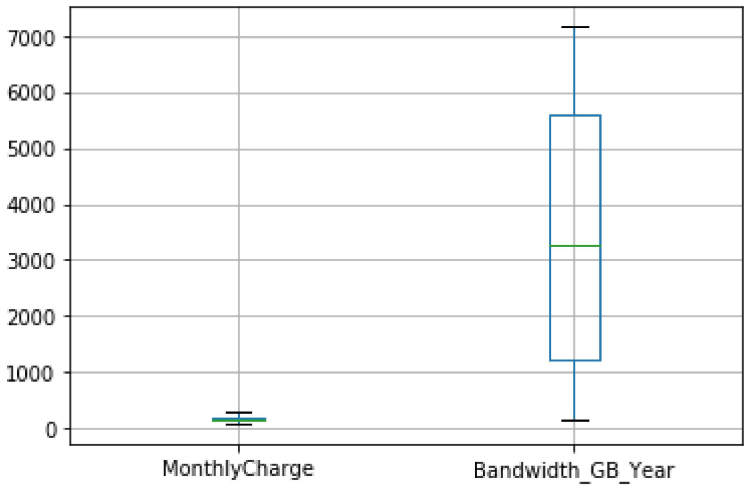


In [20]:
```
plt.figure()
sns.countplot(x='InternetService', hue='Churn', data=churn_df, palette='RdBu')
plt.xticks([0,1], ['No', 'Yes'])
plt.show()
```

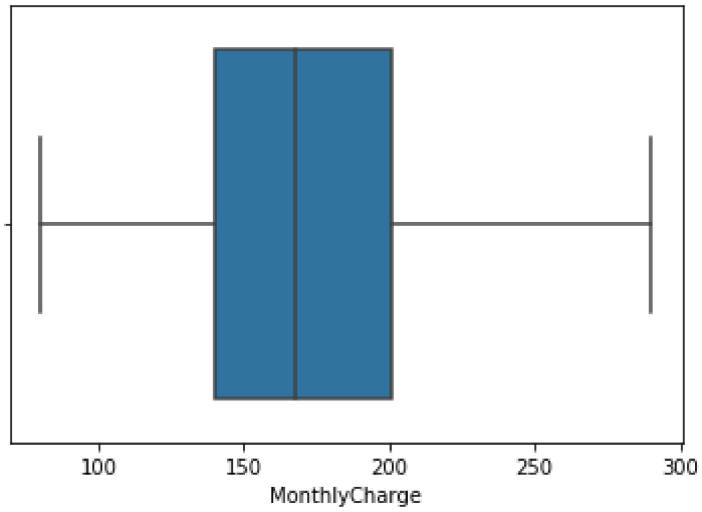


In [21]:
```
# Boxplots MonthlyCharge and Bandwidth_GB_Year
churn_df.boxplot(column=['MonthlyCharge','Bandwidth_GB_Year'])
```

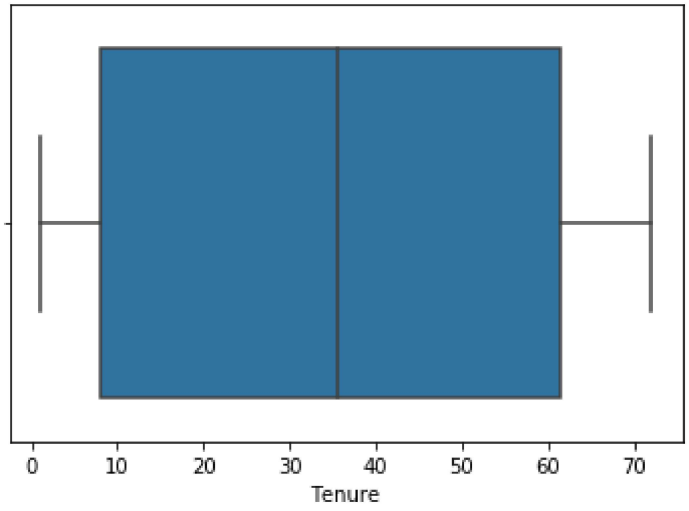
Out[21]: <matplotlib.axes._subplots.AxesSubplot at 0x1eddd32e548>



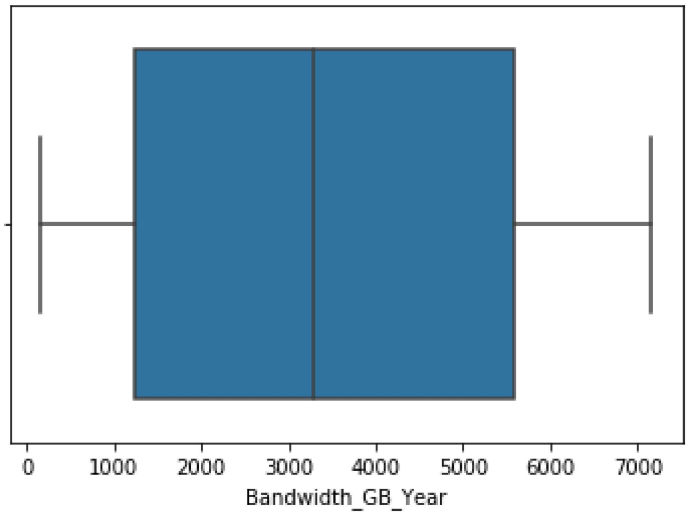
In [22]:
```
sns.boxplot('MonthlyCharge', data = churn_df)
plt.show()
```



In [23]:
```
sns.boxplot('Tenure', data = churn_df)
plt.show()
```



In [24]:
```
sns.boxplot('Bandwidth_GB_Year', data = churn_df)
plt.show()
```



In [25]:
```
# Find null values
data_null = churn_df.isnull().sum()
print(data_null)
```

```
Customer_id             0
Interaction             0
UID                     0
City                    0
State                   0
County                  0
Zip                     0
Lat                     0
Lng                     0
Population              0
Area                    0
TimeZone                0
Job                     0
Children                0
Age                     0
Income                  0
Marital                 0
Gender                  0
Churn                   0
Outage_sec_perweek      0
Email                   0
Contacts                0
Yearly_equip_failure    0
Techie                  0
Contract                0
Port_modem              0
Tablet                  0
InternetService         0
Phone                   0
Multiple                0
OnlineSecurity          0
OnlineBackup            0
DeviceProtection        0
TechSupport             0
StreamingTV             0
StreamingMovies         0
PaperlessBilling        0
PaymentMethod           0
Tenure                  0
MonthlyCharge           0
Bandwidth_GB_Year       0
TimelyResponse          0
Fixes                   0
Replacements            0
Reliability             0
Options                 0
Respectful              0
Courteous               0
Listening               0
dtype: int64
```

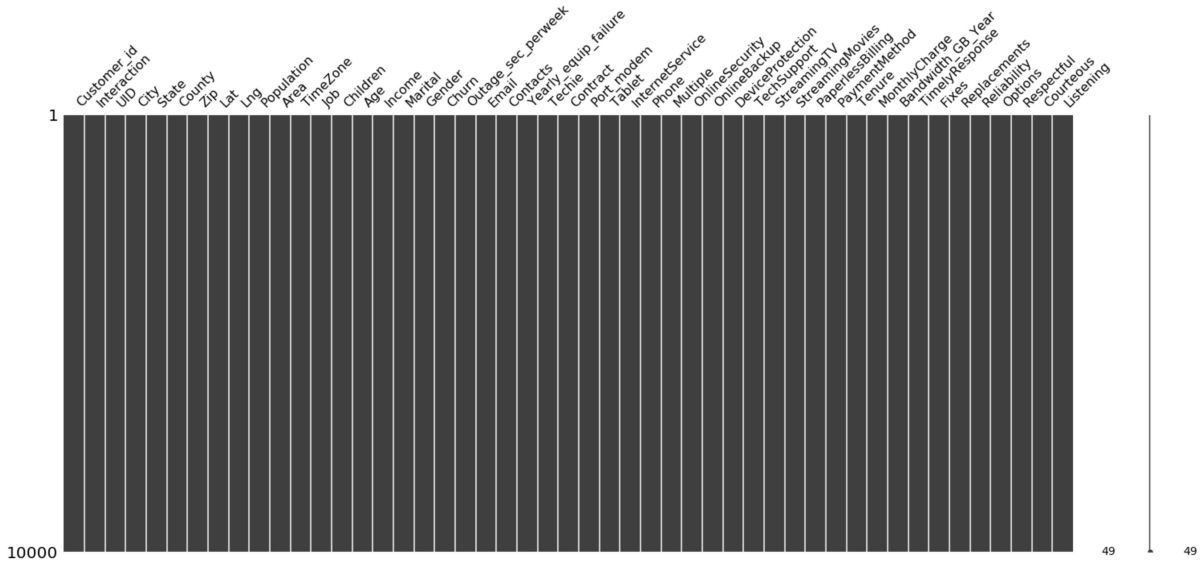
In [26]:
```
# Look for missing data

# Install library
!pip install missingno

# Importing libraries
import missingno as msno

# Visualize missing values
msno.matrix(churn_df);
```

```
Requirement already satisfied: missingno in c:\users\coope\anaconda3\lib\site
-packages (0.5.0)
Requirement already satisfied: matplotlib in c:\users\coope\anaconda3\lib\sit
e-packages (from missingno) (3.1.1)
Requirement already satisfied: numpy in c:\users\coope\anaconda3\lib\site-pac
kages (from missingno) (1.16.5)
Requirement already satisfied: scipy in c:\users\coope\anaconda3\lib\site-pac
kages (from missingno) (1.3.1)
Requirement already satisfied: seaborn in c:\users\coope\anaconda3\lib\site-p
ackages (from missingno) (0.9.0)
Requirement already satisfied: cycler>=0.10 in c:\users\coope\anaconda3\lib\s
ite-packages (from matplotlib->missingno) (0.10.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\coope\anaconda3
\lib\site-packages (from matplotlib->missingno) (1.1.0)
Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.1 in
c:\users\coope\anaconda3\lib\site-packages (from matplotlib->missingno) (2.4.
2)
Requirement already satisfied: python-dateutil>=2.1 in c:\users\coope\anacond
a3\lib\site-packages (from matplotlib->missingno) (2.8.0)
Requirement already satisfied: pandas>=0.15.2 in c:\users\coope\anaconda3\lib
\site-packages (from seaborn->missingno) (0.25.1)
Requirement already satisfied: six in c:\users\coope\anaconda3\lib\site-packa
ges (from cycler>=0.10->matplotlib->missingno) (1.12.0)
Requirement already satisfied: setuptools in c:\users\coope\anaconda3\lib\sit
e-packages (from kiwisolver>=1.0.1->matplotlib->missingno) (41.4.0)
Requirement already satisfied: pytz>=2017.2 in c:\users\coope\anaconda3\lib\s
ite-packages (from pandas>=0.15.2->seaborn->missingno) (2019.3)
```



In [27]:
```
# Dummy variables for qualitative binary data
churn_df['DummyGender'] = [1 if v == 'Male' else 0 for v in churn_df['Gender'
]]
churn_df['DummyChurn'] = [1 if v == 'Yes' else 0 for v in churn_df['Churn']] #
## If the customer left (churned) they get a '1'
churn_df['DummyTechie'] = [1 if v == 'Yes' else 0 for v in churn_df['Techie']]
churn_df['DummyContract'] = [1 if v == 'Two Year' else 0 for v in churn_df['Co
ntract']]
churn_df['DummyPort_modem'] = [1 if v == 'Yes' else 0 for v in churn_df['Port_
modem']]
churn_df['DummyTablet'] = [1 if v == 'Yes' else 0 for v in churn_df['Tablet']]
churn_df['DummyInternetService'] = [1 if v == 'Fiber Optic' else 0 for v in ch
urn_df['InternetService']]
churn_df['DummyPhone'] = [1 if v == 'Yes' else 0 for v in churn_df['Phone']]
churn_df['DummyMultiple'] = [1 if v == 'Yes' else 0 for v in churn_df['Multipl
e']]
churn_df['DummyOnlineSecurity'] = [1 if v == 'Yes' else 0 for v in churn_df['O
nlineSecurity']]
churn_df['DummyOnlineBackup'] = [1 if v == 'Yes' else 0 for v in churn_df['Onl
ineBackup']]
churn_df['DummyDeviceProtection'] = [1 if v == 'Yes' else 0 for v in churn_df[
'DeviceProtection']]
churn_df['DummyTechSupport'] = [1 if v == 'Yes' else 0 for v in churn_df['Tech
Support']]
churn_df['DummyStreamingTV'] = [1 if v == 'Yes' else 0 for v in churn_df['Stre
amingTV']]
churn_df['StreamingMovies'] = [1 if v == 'Yes' else 0 for v in churn_df['Strea
mingMovies']]
churn_df['DummyPaperlessBilling'] = [1 if v == 'Yes' else 0 for v in churn_df[
'PaperlessBilling']]
```

In [28]:
```
 # Drop original data for dummy data
churn_df = churn_df.drop(columns=['Gender', 'Churn', 'Techie', 'Contract', 'Po
rt_modem', 'Tablet',
 'InternetService', 'Phone', 'Multiple', 'OnlineSecurity',
 'OnlineBackup', 'DeviceProtection', 'TechSupport',
 'StreamingTV', 'StreamingMovies', 'PaperlessBilling'])
```

In [29]:
```
churn_df.head()
```

Out[29]:

CaseOrder	Customer_id	Interaction	UID	City	State	C
1	K409198	aa90260b-4141-4a24-8e36-b04ce1f4f77b	e885b299883d4f9fb18e39c75155d990	Point Baker	AK	Pri W
2	S120509	fb76459f-c047-4a9d-8af9-e0f7d4ac2524	f2de8bef964785f41a2959829830fb8a	West Branch	MI	Og
3	K191035	344d114c-3736-4be5-98f7-c72c281e2d35	f1784cfa9f6d92ae816197eb175d3c71	Yamhill	OR	Y
4	D90850	abfa2b40-2d43-4994-b15a-989b8c79e311	dc8a365077241bb5cd5ccd305136b05e	Del Mar	CA	
5	K662701	68a861fd-0d20-4e51-a587-8a90407ee574	aabb64a116e83fdc4befc1fbab1663f9	Needville	TX	

5 rows × 48 columns

In [30]:
```
# Drop less meaningful data
churn_df = churn_df.drop(columns=['Customer_id', 'Interaction', 'UID',
 'City', 'State', 'County', 'Zip', 'Lat', 'Lng',
 'Area', 'TimeZone', 'Job', 'Marital', 'PaymentMethod'])
churn_df.head()
```

Out[30]:

CaseOrder	Population	Children	Age	Income	Outage_sec_perweek	Email	Contacts	Yearly_e
1	38	0	68	28561.99	7.978323	10	0	
2	10446	1	27	21704.77	11.699080	12	0	
3	3735	4	50	9609.57	10.752800	9	0	
4	13863	1	48	18925.23	14.913540	15	2	
5	11352	0	83	40074.19	8.147417	16	2	

5 rows × 34 columns

In [31]:
```
# Move DummyChurn to end
churn_df = churn_df[['Children', 'Age', 'Income', 'Outage_sec_perweek', 'Emai
l', 'Contacts',
 'Yearly_equip_failure', 'Tenure', 'MonthlyCharge', 'Bandwidth_GB_Year',
 'TimelyResponse', 'Fixes', 'Replacements',
 'Reliability', 'Options', 'Respectful', 'Courteous', 'Listening',
 'DummyGender', 'DummyTechie', 'DummyContract',
 'DummyPort_modem', 'DummyTablet', 'DummyInternetService', 'DummyPhone',
 'DummyMultiple', 'DummyOnlineSecurity', 'DummyOnlineBackup',
 'DummyDeviceProtection', 'DummyTechSupport', 'DummyStreamingTV',
 'DummyPaperlessBilling', 'DummyChurn',]]
```

In [32]:
```
churn_df.head()
```

Out[32]:

CaseOrder	Children	Age	Income	Outage_sec_perweek	Email	Contacts	Yearly_equip_failure
1	0	68	28561.99	7.978323	10	0	1
2	1	27	21704.77	11.699080	12	0	1
3	4	50	9609.57	10.752800	9	0	1
4	1	48	18925.23	14.913540	15	2	0
5	0	83	40074.19	8.147417	16	2	1

5 rows × 33 columns

In [33]:
```
# List features
features = (list(churn_df.columns[:-1]))
print('Features for analysis include: \n', features)
```

```
Features for analysis include:
 ['Children', 'Age', 'Income', 'Outage_sec_perweek', 'Email', 'Contacts', 'Ye
arly_equip_failure', 'Tenure', 'MonthlyCharge', 'Bandwidth_GB_Year', 'TimelyR
esponse', 'Fixes', 'Replacements', 'Reliability', 'Options', 'Respectful', 'C
ourteous', 'Listening', 'DummyGender', 'DummyTechie', 'DummyContract', 'Dummy
Port_modem', 'DummyTablet', 'DummyInternetService', 'DummyPhone', 'DummyMulti
ple', 'DummyOnlineSecurity', 'DummyOnlineBackup', 'DummyDeviceProtection', 'D
ummyTechSupport', 'DummyStreamingTV', 'DummyPaperlessBilling']
```

In [34]:
```
#Extract cleaned dataset
churn_df.to_csv('churn_prepared.csv')
```

In [35]:
```
churn_df = pd.read_csv('churn_prepared.csv')
```

In [36]:
```
X = churn_df.drop('DummyChurn', axis=1).values
y = churn_df['DummyChurn'].values
```

In [42]:
```
churn_df = churn_df[['Outage_sec_perweek', 'MonthlyCharge', 'Bandwidth_GB_Yea
r', 'DummyChurn']]
```

In [43]:
```
target_column = churn_df['DummyChurn']
predictors = list(set(list(churn_df.columns))-set(target_column))
churn_df[predictors] = churn_df[predictors]/churn_df[predictors].max()
```

In [44]:
```
X = churn_df.iloc[:, 1:-1].values
y = churn_df.iloc[:,-1].values
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.30, rand
om_state=40)
```

In [45]:
```
rr = Ridge(alpha=0.01)
rr.fit(X_train, y_train)
pred_train_rr = rr.predict(X_train)

pred_test_rr = rr.predict(X_test)
```

In [50]:
```
print(np.sqrt(mean_squared_error(y_test, pred_test_rr)))
print(r2_score(y_test, pred_test_rr))
```

```
0.3510387173061011
0.3637303509289348
```

In [51]:
```
pd.DataFrame(X_train).to_csv('X_train_set.csv')
pd.DataFrame(X_test).to_csv('X_test_set.csv')
pd.DataFrame(y_train).to_csv('y_train_set.csv')
pd.DataFrame(y_test).to_csv('y_test_set.csv')
```

In [ ]:
```

```