In [110…

```python
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn import preprocessing
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.metrics import classification_report
%matplotlib inline
import pylab
from pylab import rcParams
import statsmodels.api as sm
import statistics
from scipy import stats
import sklearn
from scipy.stats import chisquare
from scipy.stats import chi2_contingency
import warnings
warnings.filterwarnings('ignore')
import matplotlib as mpl
COLOR = 'white'
mpl.rcParams['text.color'] = COLOR
mpl.rcParams['axes.labelcolor'] = COLOR
mpl.rcParams['xtick.color'] = COLOR
mpl.rcParams['ytick.color'] = COLOR

medical_df = pd.read_csv('C:/Users/MichaelRupert/Downloads/e9d8sm5uf8df75k650df/medical_raw_data.csv');

#Check for duplicaes
Is_dups_bool = medical_df.duplicated()
print(Is_dups_bool.value_counts())

#Remove Unwanted columns
new_Med_DF  =  medical_df.iloc[0:9999,6:53]
print(new_Med_DF)


new_Med_DF.rename(columns = {'Income' : 'Household_Income',
                             'TotalCharge': 'Daily_Average_Charges', 'Additional_Charges' : 'Average_Daily_
                             'Item1':'Timely_Admission','Item2':'Timely_Treatment','Item3':'Timely_Visits',
                             'Item4':'Reliability','Item5':'Options','Item6':'Hours_Treatment',
                             'Item7':'Courteous_Staff','Item8':'Active_Listening'},inplace = True)
print(new_Med_DF.head(1))
```

```
False    10000
dtype: int64
     State        County    Zip       Lat        Lng  Population      Area  \
0       AL        Morgan  35621  34.34960  -86.72508        2951  Suburban
1       FL       Jackson  32446  30.84513  -85.22907       11303     Urban
2       SD     Minnehaha  57110  43.54321  -96.63772       17125  Suburban
3       MN        Waseca  56072  43.89744  -93.51479        2162  Suburban
4       VA  King William  23181  37.59894  -76.88958        5287     Rural
...    ...           ...    ...       ...        ...         ...       ...
9994    FL  Indian River  32948  27.88942  -80.73347        7908     Urban
9995    NC        Warren  27563  36.42886  -78.23716        4762     Urban
9996    NJ      Atlantic   8340  39.43609  -74.87302        1251     Urban
9997    TN    Montgomery  37171  36.36655  -87.29988         532     Rural
9998    SD    Pennington  57775  44.10354 -102.01593         271     Rural

              Timezone                                    Job  Children  ...  \
0      America/Chicago      Psychologist, sport and exercise       1.0  ...
1      America/Chicago          Community development worker       3.0  ...
2      America/Chicago               Chief Executive Officer       3.0  ...
3      America/Chicago                   Early years teacher       0.0  ...
4     America/New_York           Health promotion specialist       NaN  ...
...                ...                                    ...       ...  ...
9994  America/New_York                      Technical author       6.0  ...
9995  America/New_York                 Programmer, multimedia       NaN  ...
9996  America/New_York         Restaurant manager, fast food       4.0  ...
9997   America/Chicago              Psychologist, occupational       3.0  ...
9998    America/Denver  Outdoor activities/education manager       3.0  ...

      TotalCharge Additional_charges Item1  Item2 Item3 Item4 Item5  Item6  \
0     3191.048774       17939.403420     3      3     2     2     4      3
1     4214.905346       17612.998120     3      4     3     4     4      4
2     2177.586768       17505.192460     2      4     4     4     3      4
3     2465.118965       12993.437350     3      5     5     3     4      5
4     1885.655137        3716.525786     2      1     3     3     5      3
...           ...                ...   ...    ...   ...   ...   ...    ...
9994  5857.563763       12045.862390     2      3     2     1     5      3
9995  6651.241294        8927.642189     3      2     2     3     4      3
9996  7851.522660       28507.147340     3      3     4     2     5      3
9997  7725.953391       15281.214660     3      3     3     4     4      2
9998  8462.831883        7781.678412     5      5     3     4     4      3

      Item7  Item8
0         3      4
1         3      3
2         3      3
3         5      5
4         4      3
...     ...    ...
9994      4      2
9995      4      2
9996      4      4
9997      3      2
9998      4      3

[9999 rows x 47 columns]
  State  County    Zip      Lat       Lng  Population      Area  \
0    AL  Morgan  35621  34.3496 -86.72508        2951  Suburban

          Timezone                               Job  Children  ...  \
0  America/Chicago  Psychologist, sport and exercise       1.0  ...

   Daily_Average_Charges Additional_charges Timely_Admission  \
0            3191.048774        17939.40342                3

   Timely_Treatment Timely_Visits Reliability Options  Hours_Treatment  \
0                 3             2           2       4                3

   Courteous_Staff  Active_Listening
0                3                 4

[1 rows x 47 columns]
```

In [111…

```python
print(new_Med_DF.head(5))
new_Med_DF.index = np.arange(1,len(new_Med_DF)+1)
print(new_Med_DF.head(5))
```

```
  State        County    Zip       Lat       Lng  Population      Area  \
0    AL        Morgan  35621  34.34960 -86.72508        2951  Suburban
1    FL       Jackson  32446  30.84513 -85.22907       11303     Urban
2    SD     Minnehaha  57110  43.54321 -96.63772       17125  Suburban
3    MN        Waseca  56072  43.89744 -93.51479        2162  Suburban
4    VA  King William  23181  37.59894 -76.88958        5287     Rural

           Timezone                               Job  Children  ...  \
0   America/Chicago  Psychologist, sport and exercise       1.0  ...
1   America/Chicago      Community development worker       3.0  ...
2   America/Chicago           Chief Executive Officer       3.0  ...
3   America/Chicago               Early years teacher       0.0  ...
4  America/New_York       Health promotion specialist       NaN  ...

   Daily_Average_Charges Additional_charges Timely_Admission  \
0            3191.048774       17939.403420                3
1            4214.905346       17612.998120                3
2            2177.586768       17505.192460                2
3            2465.118965       12993.437350                3
4            1885.655137        3716.525786                2

   Timely_Treatment Timely_Visits Reliability Options  Hours_Treatment  \
0                 3             2           2       4                3
1                 4             3           4       4                4
2                 4             4           4       3                4
3                 5             5           3       4                5
4                 1             3           3       5                3

   Courteous_Staff  Active_Listening
0                3                 4
1                3                 3
2                3                 3
3                5                 5
4                4                 3

[5 rows x 47 columns]
  State        County    Zip       Lat       Lng  Population      Area  \
1    AL        Morgan  35621  34.34960 -86.72508        2951  Suburban
2    FL       Jackson  32446  30.84513 -85.22907       11303     Urban
3    SD     Minnehaha  57110  43.54321 -96.63772       17125  Suburban
4    MN        Waseca  56072  43.89744 -93.51479        2162  Suburban
5    VA  King William  23181  37.59894 -76.88958        5287     Rural

           Timezone                               Job  Children  ...  \
1   America/Chicago  Psychologist, sport and exercise       1.0  ...
2   America/Chicago      Community development worker       3.0  ...
3   America/Chicago           Chief Executive Officer       3.0  ...
4   America/Chicago               Early years teacher       0.0  ...
5  America/New_York       Health promotion specialist       NaN  ...

   Daily_Average_Charges Additional_charges Timely_Admission  \
1            3191.048774       17939.403420                3
2            4214.905346       17612.998120                3
3            2177.586768       17505.192460                2
4            2465.118965       12993.437350                3
5            1885.655137        3716.525786                2

   Timely_Treatment Timely_Visits Reliability Options  Hours_Treatment  \
1                 3             2           2       4                3
2                 4             3           4       4                4
3                 4             4           4       3                4
4                 5             5           3       4                5
5                 1             3           3       5                3

   Courteous_Staff  Active_Listening
1                3                 4
2                3                 3
3                3                 3
4                5                 5
5                4                 3

[5 rows x 47 columns]
```

In [112…

```python
my_dict = {"Rural": 1, "Suburban": 2, "Urban": 3}
my_dict1 ={
  "America/Puerto_Rico":-4,
  "America/Detroit":-5,
  "America/Indiana/Indianapolis":-5,
  "America/Indiana/Marengo": -5,
  "America/Indiana/Vincennes": -5,
  "America/Indiana/Vevay":-5,
  "America/Indiana/Winamac":-5,
  "America/Kentucky/Louisville":-5,
  "America/New_York":-5,
  "America/Toronto":-5,
  "America/Chicago":-6,
  "America/Indiana/Knox":-6,
  "America/Indiana/Tell_City": -6,
  "America/Menominee": -6,
  "America/North_Dakota/Beulah": -6,
  "America/North_Dakota/New_Salem": -6,
  "America/Boise":-7,
  "America/Denver":-7,
  "America/Phoenix":-7,
  "America/Los_Angeles": -8,
  "America/Anchorage": -9,
  "America/Nome":-9,
  "America/Sitka":-9,
  "America/Yakutat":-9,
  "America/Adak":-10,
  "Pacific/Honolulu":-10}
my_dict2 = {
  "No Schooling Completed": 0,
  "Nursery School to 8th Grade": 8,
  "9th Grade to 12th Grade, No Diploma": 11,
  "GED or Alternative Credential": 12,
  "Regular High School Diploma": 12,
  "Some College, Less than 1 Year": 13,
  "Some College, 1 or More Years, No Degree": 14,
  "Associate's Degree": 15,
  "Bachelor's Degree": 16,
  "Master's Degree": 18,
  "Professional School Degree": 20,
  "Doctorate Degree": 24
}
my_dict3={"Yes": 1,"No": 0}
my_dict4={"Male": 1,"Female": 2,"Prefer not to answer": 0}
my_dict5={"Low":1,"Medium":2,"High":3}
my_dict6={"Emergency Admission":1,   "Elective Admission":2,    "Observation Admission":3}
my_dict7={"Blood Work":1,  "Intravenous":2,  "CT Scan":3,   "MRI":4}
my_dict8 ={"Divorced":1,  "Married":2,      "Widowed":3,       "Never Married":4,  "Separated":5 }
my_dict9={"Full Time": 1,  "Retired":2,    "Unemployed":3, "Student":4,    "Part Time":5}
my_dict10 = {'AL' :  1, 'FL': 2, 'SD' : 3, 'MN' : 4,'VA' : 5, 'OK': 6, 'OH': 7, 'MS' : 8, 'WI': 9, 'IA':10, 'CA
 'NE':15, 'PA':16, 'AR':17, 'WV':18, 'KS':19, 'MA':20, 'KY':21, 'NY':22, 'VT':23, 'DC':24, 'IL':25, 'ND':26, 'S
 'NM':29, 'NH':30, 'GA':31, 'NC':32, 'MD':33, 'TN':34, 'WA': 35,'TX':36, 'CO':37, 'NJ':38, 'LA':39,'OR':40, 'AZ
 'ID':42, 'UT':43, 'RI': 44, 'MT':45, 'PR':46, 'NV':47, 'CT':48, 'HI':49, 'WY':50 ,'DE' : 51}

new_Med_DF = new_Med_DF.replace({"Area": my_dict,"Timezone": my_dict1,"Education":my_dict2,"ReAdmis": my_dict3,
                    "Initial_admin": my_dict6,"HighBlood": my_dict3,
                    "Stroke": my_dict3,"Complication_risk":my_dict5,"Arthritis":my_dict3,"Hyperlipidemia": my_
                    "Allergic_rhinitis": my_dict3,"Reflux_esophagitis": my_dict3,"Asthma": my_dict3,"Services'
print(new_Med_DF.iloc[:,9:30])
```

```
      Children   Age  Education  Employment  Household_Income  Marital  \
1          1.0  53.0         13           1          86575.93        1
2          3.0  51.0         14           1          46805.99        2
3          3.0  53.0         14           2          14370.14        3
4          0.0  78.0         12           2          39741.49        2
5          NaN  22.0         12           1           1209.56        3
...        ...   ...        ...         ...               ...      ...
9995       6.0  69.0         12           1          39797.05        5
9996       NaN  25.0         16           4          45967.61        3
9997       4.0  87.0         12           1          14983.02        3
9998       3.0   NaN         12           1          65917.81        5
9999       3.0  43.0         16           1          29702.32        1

      Gender  ReAdmis  VitD_levels  Doc_visits  ...  VitD_supp  Soft_drink  \
1          1        0    17.802330           6  ...          0         NaN
2          2        0    18.994640           4  ...          1         0.0
3          2        0    17.415889           4  ...          0         0.0
4          1        0    17.420079           4  ...          0         0.0
5          2        0    16.870524           5  ...          2         1.0
...      ...      ...          ...         ...  ...        ...         ...
9995       1        0    17.229222           4  ...          2         0.0
9996       1        0    16.481612           4  ...          1         0.0
9997       1        1    18.451601           5  ...          0         0.0
9998       2        1    15.752751           4  ...          0         1.0
9999       1        1    21.956305           5  ...          1         0.0

      Initial_admin  HighBlood  Stroke  Complication_risk  Overweight  \
1                 1          1       0                  2         0.0
2                 1          1       0                  3         1.0
3                 2          1       0                  2         1.0
4                 2          0       1                  2         0.0
5                 2          0       0                  1         0.0
...             ...        ...     ...                ...         ...
9995              1          0       0                  2         0.0
9996              1          1       0                  2         NaN
9997              2          1       0                  2         1.0
9998              2          1       0                  3         1.0
9999              1          0       0                  2         1.0

      Arthritis  Diabetes  Hyperlipidemia
1             1         1               0
2             0         0               0
3             0         1               0
4             1         0               0
5             0         0               1
...         ...       ...             ...
9995          1         0               0
9996          0         0               0
9997          1         1               0
9998          0         0               0
9999          0         0               0

[9999 rows x 21 columns]
```

In [113…

```python
# Check to see how man zeros in population
zeros = (new_Med_DF["Population"] == 0).sum()
print(zeros)
print(new_Med_DF[35:45])
```

```
109
    State                County    Zip       Lat       Lng  Population  Area  \
36     23                Orange   5043  43.80814 -72.20823        1009     1
37     20             Middlesex   2468  42.32731 -71.23076        5688     2
38     20             Middlesex   2138  42.38004 -71.13463       38835     1
39     15               Douglas  68164  41.29905 -96.10955       29143     1
40     16            Clearfield  16878  40.94827 -78.32867        1816     2
41     21                  Bath  40360  38.15445 -83.76738        6875     2
42     15               Douglas  68132  41.26466 -96.00236       14569     3
43     24  District of Columbia  20228  38.88631 -77.03016           0     1
44     16            Lackawanna  18504  41.42553 -75.69992       20531     1
45      1                  Bibb  35074  33.22242 -87.12425           0     2

    Timezone                                        Job  Children  \
36        -5                             Radio producer       0.0
37        -5               Designer, television/film set       1.0
38        -5                  Conference centre manager       1.0
39        -6               Advertising account executive       3.0
40        -5                  Civil Service fast streamer       3.0
41        -5              Training and development officer       3.0
42        -6                              Buyer, retail       NaN
43        -5                             Event organiser       0.0
44        -5                IT technical support officer       2.0
45        -6  Historic buildings inspector/conservation officer       0.0

    ...  Daily_Average_Charges  Additional_charges  Timely_Admission  \
36  ...            2805.845854        10729.606930                 4
37  ...            2040.063839         4204.826360                 4
38  ...            2775.809508         9646.930063                 5
39  ...            3246.774111         9711.930249                 2
40  ...            2285.954413        11658.486190                 5
41  ...            3319.270971        13347.990770                 3
42  ...            3360.320065        14748.528000                 3
43  ...            3523.303828         9441.679490                 4
44  ...            3574.732199         4620.845473                 3
45  ...            4783.979430         7361.105691                 3

    Timely_Treatment  Timely_Visits  Reliability  Options  Hours_Treatment  \
36                 4              4            3        4                5
37                 4              5            5        1                3
38                 5              4            4        3                5
39                 1              2            6        2                4
40                 3              2            4        4                4
41                 4              2            3        4                3
42                 3              4            4        4                3
43                 3              4            4        2                4
44                 2              4            4        3                3
45                 5              3            1        6                3

    Courteous_Staff  Active_Listening
36                3                 4
37                5                 5
38                5                 6
39                4                 3
40                3                 4
41                4                 4
42                5                 3
43                4                 6
44                2                 5
45                3                 2

[10 rows x 47 columns]
```

In [114…

```python
# Replace the zeros with NaN in population

new_Med_DF["Population"] =  new_Med_DF["Population"].replace(0,np.NAN)
print(new_Med_DF[35:45])
```

```
    State                County    Zip       Lat       Lng  Population  Area  \
36     23                Orange   5043  43.80814 -72.20823      1009.0     1
37     20             Middlesex   2468  42.32731 -71.23076      5688.0     2
38     20             Middlesex   2138  42.38004 -71.13463     38835.0     1
39     15               Douglas  68164  41.29905 -96.10955     29143.0     1
40     16            Clearfield  16878  40.94827 -78.32867      1816.0     2
41     21                  Bath  40360  38.15445 -83.76738      6875.0     2
42     15               Douglas  68132  41.26466 -96.00236     14569.0     3
43     24  District of Columbia  20228  38.88631 -77.03016         NaN     1
44     16            Lackawanna  18504  41.42553 -75.69992     20531.0     1
45      1                  Bibb  35074  33.22242 -87.12425         NaN     2

    Timezone                                        Job  Children  \
36        -5                             Radio producer       0.0
37        -5               Designer, television/film set       1.0
38        -5                  Conference centre manager       1.0
39        -6               Advertising account executive       3.0
40        -5                  Civil Service fast streamer       3.0
41        -5              Training and development officer       3.0
42        -6                              Buyer, retail       NaN
43        -5                             Event organiser       0.0
44        -5                IT technical support officer       2.0
45        -6  Historic buildings inspector/conservation officer       0.0

    ...  Daily_Average_Charges  Additional_charges  Timely_Admission  \
36  ...            2805.845854        10729.606930                 4
37  ...            2040.063839         4204.826360                 4
38  ...            2775.809508         9646.930063                 5
39  ...            3246.774111         9711.930249                 2
40  ...            2285.954413        11658.486190                 5
41  ...            3319.270971        13347.990770                 3
42  ...            3360.320065        14748.528000                 3
43  ...            3523.303828         9441.679490                 4
44  ...            3574.732199         4620.845473                 3
45  ...            4783.979430         7361.105691                 3

    Timely_Treatment  Timely_Visits  Reliability  Options  Hours_Treatment  \
36                 4              4            3        4                5
37                 4              5            5        1                3
38                 5              4            4        3                5
39                 1              2            6        2                4
40                 3              2            4        4                4
41                 4              2            3        4                3
42                 3              4            4        4                3
43                 3              4            4        2                4
44                 2              4            4        3                3
45                 5              3            1        6                3

    Courteous_Staff  Active_Listening
36                3                 4
37                5                 5
38                5                 6
39                4                 3
40                3                 4
41                4                 4
42                5                 3
43                4                 6
44                2                 5
45                3                 2

[10 rows x 47 columns]
```

In [115…

```python
nan_values = new_Med_DF.isnull()
nan_columns = nan_values.any()
columns_with_nan = new_Med_DF.columns[nan_columns].tolist()
print(columns_with_nan)
```

```
['Population', 'Children', 'Age', 'Household_Income', 'Soft_drink', 'Overweight', 'Anxiety', 'Initial_days']
```

In [116…

```python
print(new_Med_DF[["Population", "Children", "Age", "Household_Income", "Soft_drink", "Overweight", "Anxiety", '
```

```
Population          10071.892214
Children                2.097423
Age                    53.295676
Household_Income    40481.492257
Soft_drink              0.258099
Overweight              0.709105
Anxiety                 0.322352
Initial_days           34.428010
dtype: float64
```

In [117…

```python
# Replace the NaN values with the Mean
new_Med_DF = new_Med_DF.fillna(new_Med_DF[["Population", "Children", "Age", "Household_Income", "Soft_drink", '
print(round(new_Med_DF[35:45]))
```

```
    State                County    Zip   Lat   Lng  Population  Area  \
36     23                Orange   5043  44.0 -72.0      1009.0     1
37     20             Middlesex   2468  42.0 -71.0      5688.0     2
38     20             Middlesex   2138  42.0 -71.0     38835.0     1
39     15               Douglas  68164  41.0 -96.0     29143.0     1
40     16            Clearfield  16878  41.0 -78.0      1816.0     2
41     21                  Bath  40360  38.0 -84.0      6875.0     2
42     15               Douglas  68132  41.0 -96.0     14569.0     3
43     24  District of Columbia  20228  39.0 -77.0     10072.0     1
44     16            Lackawanna  18504  41.0 -76.0     20531.0     1
45      1                  Bibb  35074  33.0 -87.0     10072.0     2

    Timezone                                        Job  Children  \
36        -5                             Radio producer       0.0
37        -5               Designer, television/film set       1.0
38        -5                  Conference centre manager       1.0
39        -6               Advertising account executive       3.0
40        -5                  Civil Service fast streamer       3.0
41        -5              Training and development officer       3.0
42        -6                              Buyer, retail       2.0
43        -5                             Event organiser       0.0
44        -5                IT technical support officer       2.0
45        -6  Historic buildings inspector/conservation officer       0.0

    ...  Daily_Average_Charges  Additional_charges  Timely_Admission  \
36  ...                 2806.0             10730.0                 4
37  ...                 2040.0              4205.0                 4
38  ...                 2776.0              9647.0                 5
39  ...                 3247.0              9712.0                 2
40  ...                 2286.0             11658.0                 5
41  ...                 3319.0             13348.0                 3
42  ...                 3360.0             14749.0                 3
43  ...                 3523.0              9442.0                 4
44  ...                 3575.0              4621.0                 3
45  ...                 4784.0              7361.0                 3

    Timely_Treatment  Timely_Visits  Reliability  Options  Hours_Treatment  \
36                 4              4            3        4                5
37                 4              5            5        1                3
38                 5              4            4        3                5
39                 1              2            6        2                4
40                 3              2            4        4                4
41                 4              2            3        4                3
42                 3              4            4        4                3
43                 3              4            4        2                4
44                 2              4            4        3                3
45                 5              3            1        6                3

    Courteous_Staff  Active_Listening
36                3                 4
37                5                 5
38                5                 6
39                4                 3
40                3                 4
41                4                 4
42                5                 3
43                4                 6
44                2                 5
45                3                 2

[10 rows x 47 columns]
```

In [118…

```python
nan_values = new_Med_DF.isnull()
nan_columns = nan_values.any()
columns_with_nan = new_Med_DF.columns[nan_columns].tolist()
print(columns_with_nan)
```

```
[]
```

In [119…

```python
print(new_Med_DF[["Population", "Children", "Age", "Household_Income", "Soft_drink", "Overweight", "Anxiety", '
```

```
Population          10071.892214
Children                2.097423
Age                    53.295676
Household_Income    40481.492257
Soft_drink              0.258099
Overweight              0.709105
Anxiety                 0.322352
Initial_days           34.428010
dtype: float64
```

In [120…

```python
new_Med_DF = new_Med_DF.fillna(new_Med_DF[["Population", "Children", "Age", "Household_Income", "Soft_drink", '
print(round(new_Med_DF[35:45]))
```

```
    State                County    Zip   Lat   Lng  Population  Area  \
36     23                Orange   5043  44.0 -72.0      1009.0     1
37     20             Middlesex   2468  42.0 -71.0      5688.0     2
38     20             Middlesex   2138  42.0 -71.0     38835.0     1
39     15               Douglas  68164  41.0 -96.0     29143.0     1
40     16            Clearfield  16878  41.0 -78.0      1816.0     2
41     21                  Bath  40360  38.0 -84.0      6875.0     2
42     15               Douglas  68132  41.0 -96.0     14569.0     3
43     24  District of Columbia  20228  39.0 -77.0     10072.0     1
44     16            Lackawanna  18504  41.0 -76.0     20531.0     1
45      1                  Bibb  35074  33.0 -87.0     10072.0     2

    Timezone                                        Job  Children  \
36        -5                             Radio producer       0.0
37        -5               Designer, television/film set       1.0
38        -5                  Conference centre manager       1.0
39        -6               Advertising account executive       3.0
40        -5                  Civil Service fast streamer       3.0
41        -5              Training and development officer       3.0
42        -6                              Buyer, retail       2.0
43        -5                             Event organiser       0.0
44        -5                IT technical support officer       2.0
45        -6  Historic buildings inspector/conservation officer       0.0

    ...  Daily_Average_Charges  Additional_charges  Timely_Admission  \
36  ...                 2806.0             10730.0                 4
37  ...                 2040.0              4205.0                 4
38  ...                 2776.0              9647.0                 5
39  ...                 3247.0              9712.0                 2
40  ...                 2286.0             11658.0                 5
41  ...                 3319.0             13348.0                 3
42  ...                 3360.0             14749.0                 3
43  ...                 3523.0              9442.0                 4
44  ...                 3575.0              4621.0                 3
45  ...                 4784.0              7361.0                 3

    Timely_Treatment  Timely_Visits  Reliability  Options  Hours_Treatment  \
36                 4              4            3        4                5
37                 4              5            5        1                3
38                 5              4            4        3                5
39                 1              2            6        2                4
40                 3              2            4        4                4
41                 4              2            3        4                3
42                 3              4            4        4                3
43                 3              4            4        2                4
44                 2              4            4        3                3
45                 5              3            1        6                3

    Courteous_Staff  Active_Listening
36                3                 4
37                5                 5
38                5                 6
39                4                 3
40                3                 4
41                4                 4
42                5                 3
43                4                 6
44                2                 5
45                3                 2

[10 rows x 47 columns]
```

In [121…

```python
print(new_Med_DF.head(1).transpose())
new_Med_DF = new_Med_DF.drop(columns = ['State', 'County','Zip', 'Lat', 'Lng','Job'])
new_Med_DF.describe()
```

```
                                                     1
State                                                1
County                                          Morgan
Zip                                              35621
Lat                                            34.3496
Lng                                          -86.72508
Population                                      2951.0
Area                                                 2
Timezone                                            -6
Job                   Psychologist, sport and exercise
Children                                           1.0
Age                                               53.0
Education                                           13
Employment                                           1
Household_Income                              86575.93
Marital                                              1
Gender                                               1
ReAdmis                                              0
VitD_levels                                   17.80233
Doc_visits                                           6
Full_meals_eaten                                     0
VitD_supp                                            0
Soft_drink                                    0.258099
Initial_admin                                        1
HighBlood                                            1
Stroke                                               0
Complication_risk                                    2
Overweight                                         0.0
Arthritis                                            1
Diabetes                                             1
Hyperlipidemia                                       0
BackPain                                             1
Anxiety                                            1.0
Allergic_rhinitis                                    1
Reflux_esophagitis                                   0
Asthma                                               1
Services                                             1
Initial_days                                  10.58577
Daily_Average_Charges                      3191.048774
Additional_charges                         17939.40342
Timely_Admission                                     3
Timely_Treatment                                     3
Timely_Visits                                        2
Reliability                                          2
Options                                              4
Hours_Treatment                                      3
Courteous_Staff                                      3
Active_Listening                                     4
```

Out[121…

	Population	Area	Timezone	Children	Age	Education	Employment	Household_Income	Marital	Ger
count	9999.000000	9999.000000	9999.000000	9999.000000	9999.000000	9999.000000	9999.000000	9999.000000	9999.000000	9999.00
mean	10071.892214	1.993299	-5.860786	2.097423	53.295676	13.526153	1.996200	40481.492257	3.001100	1.480
std	14785.192945	0.816816	1.001610	1.854792	17.994275	3.145866	1.414279	24883.852453	1.407087	0.540
min	1.000000	1.000000	-10.000000	0.000000	18.000000	0.000000	1.000000	154.080000	1.000000	0.000
25%	743.000000	1.000000	-6.000000	1.000000	41.000000	12.000000	1.000000	23948.445000	2.000000	1.000
50%	2950.000000	2.000000	-6.000000	2.097423	53.295676	14.000000	1.000000	40481.492257	3.000000	2.000
75%	13938.000000	3.000000	-5.000000	3.000000	65.000000	16.000000	3.000000	46463.540000	4.000000	2.000
max	122814.000000	3.000000	-4.000000	10.000000	89.000000	24.000000	5.000000	207249.130000	5.000000	2.000

8 rows × 41 columns

In [122…

```python
#Move ReAdmis to end of Dataset and make target
new_Med_DF = new_Med_DF[['Population', 'Area', 'Timezone', 'Children', 'Age', 'Education',
         'Employment', 'Household_Income', 'Marital', 'Gender',
         'VitD_levels', 'Doc_visits', 'Full_meals_eaten', 'VitD_supp',
         'Soft_drink', 'Initial_admin', 'HighBlood', 'Stroke', 'Complication_risk',
         'Overweight', 'Arthritis', 'Diabetes', 'Hyperlipidemia', 'BackPain',
         'Anxiety', 'Allergic_rhinitis', 'Reflux_esophagitis', 'Asthma',
         'Services', 'Initial_days', 'Daily_Average_Charges', 'Additional_charges',
         'Timely_Admission', 'Timely_Treatment', 'Timely_Visits', 'Reliability',
         'Options', 'Hours_Treatment', 'Courteous_Staff', 'Active_Listening', 'ReAdmis']]

df = new_Med_DF.columns
print(df)
new_Med_DF[['Population', 'Area', 'Timezone', 'Children', 'Age', 'Education', 'Employment', 'Household_Income',
plt.savefig('medical_pyplot.jpg')
plt.tight_layout()
plt.show()
```

```
Index(['Population', 'Area', 'Timezone', 'Children', 'Age', 'Education',
       'Employment', 'Household_Income', 'Marital', 'Gender', 'VitD_levels',
       'Doc_visits', 'Full_meals_eaten', 'VitD_supp', 'Soft_drink',
       'Initial_admin', 'HighBlood', 'Stroke', 'Complication_risk',
       'Overweight', 'Arthritis', 'Diabetes', 'Hyperlipidemia', 'BackPain',
       'Anxiety', 'Allergic_rhinitis', 'Reflux_esophagitis', 'Asthma',
       'Services', 'Initial_days', 'Daily_Average_Charges',
       'Additional_charges', 'Timely_Admission', 'Timely_Treatment',
       'Timely_Visits', 'Reliability', 'Options', 'Hours_Treatment',
       'Courteous_Staff', 'Active_Listening', 'ReAdmis'],
      dtype='object')
```