```
Start coding or generate with AI.
```

**AGABA LUCKY**
**2024/HD05/21913U **

**RIGOROUS EXPLORATORY DATA ANALYSIS OF DATA CONCERNING HEALTHCARE, AND HEALTH POLICY ISSUES AFFECTING AMERICANS AGED 50 AND OLDER UTILIZING PYTHON LIBRARIES INCLUDING PANDAS, MATPLOTLIB, AND SEABORN**

```
Start coding or generate with AI.
```

Double-click (or enter) to edit

```
#import the necessary python libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Read the csv file
df = pd.read_csv('/content/NPHA-doctor-visits.csv')
```

```
#check for the first five rows of the dataset
df.head()
```

| | Number of Doctors Visited | Age | Phyiscal Health | Mental Health | Dental Health | Employment | Stress Keeps Patient from Sleeping | Medication Keeps Patient from Sleeping | Pain Keeps Patient from Sleeping | Bathroom Needs Keeps Patient from Sleeping | Uknown Keeps Patient from Sleeping | Trouble Sleeping | Prescription Sleep Medication | Ra |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 3 | 2 | 4 | 3 | 3 | 3 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | |
| **1** | 2 | 2 | 4 | 2 | 3 | 3 | 1 | 0 | 0 | 1 | 0 | 3 | 3 | |
| **2** | 3 | 2 | 3 | 2 | 3 | 3 | 0 | 0 | 0 | 0 | 1 | 3 | 3 | |

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
#check for the last five rows of the dataset
df.tail()
```

| | Number of Doctors Visited | Age | Phyiscal Health | Mental Health | Dental Health | Employment | Stress Keeps Patient from Sleeping | Medication Keeps Patient from Sleeping | Pain Keeps Patient from Sleeping | Bathroom Needs Keeps Patient from Sleeping | Uknown Keeps Patient from Sleeping | Trouble Sleeping | Prescription Sleep Medication |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **709** | 2 | 2 | 2 | 2 | 2 | 3 | 0 | 0 | 0 | 1 | 0 | 3 | 3 |
| **710** | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 2 | 3 |
| **711** | 3 | 2 | 4 | 2 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 3 | 3 |
| **712** | 3 | 2 | 3 | 1 | 3 | 3 | 1 | 0 | 1 | 1 | 1 | 3 | 3 |

```
#check for the number of rows and columns in the dataset
df.shape
```

```
(714, 15)
```

```
#check for the columns together with their datatypes
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 714 entries, 0 to 713
Data columns (total 15 columns):
 #   Column                                Non-Null Count  Dtype
---  ------                                --------------  -----
 0   Number of Doctors Visited             714 non-null    int64
 1   Age                                   714 non-null    int64
 2   Phyiscal Health                       714 non-null    int64
 3   Mental Health                         714 non-null    int64
```

```
 4   Dental Health                           714 non-null    int64
 5   Employment                              714 non-null    int64
 6   Stress Keeps Patient from Sleeping      714 non-null    int64
 7   Medication Keeps Patient from Sleeping  714 non-null    int64
 8   Pain Keeps Patient from Sleeping        714 non-null    int64
 9   Bathroom Needs Keeps Patient from Sleeping  714 non-null  int64
 10  Uknown Keeps Patient from Sleeping      714 non-null    int64
 11  Trouble Sleeping                        714 non-null    int64
 12  Prescription Sleep Medication           714 non-null    int64
 13  Race                                    714 non-null    int64
 14  Gender                                  714 non-null    int64
dtypes: int64(15)
memory usage: 83.8 KB
```

```
#check for the number of unique elements in the dataset
df.nunique()
```

| | 0 |
|---|---|
| **Number of Doctors Visited** | 3 |
| **Age** | 1 |
| **Phyiscal Health** | 6 |
| **Mental Health** | 6 |
| **Dental Health** | 7 |
| **Employment** | 4 |
| **Stress Keeps Patient from Sleeping** | 2 |
| **Medication Keeps Patient from Sleeping** | 2 |
| **Pain Keeps Patient from Sleeping** | 2 |
| **Bathroom Needs Keeps Patient from Sleeping** | 2 |
| **Uknown Keeps Patient from Sleeping** | 2 |
| **Trouble Sleeping** | 4 |
| **Prescription Sleep Medication** | 4 |
| **Race** | 5 |
| **Gender** | 2 |

```
#obtain a summary of the dataset/ descriptive statistics
df.describe().T
```

| | count | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|---|
| **Number of Doctors Visited** | 714.0 | 2.112045 | 0.683441 | 1.0 | 2.0 | 2.0 | 3.0 | 3.0 |
| **Age** | 714.0 | 2.000000 | 0.000000 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| **Phyiscal Health** | 714.0 | 2.794118 | 0.900939 | -1.0 | 2.0 | 3.0 | 3.0 | 5.0 |
| **Mental Health** | 714.0 | 1.988796 | 0.939928 | -1.0 | 1.0 | 2.0 | 3.0 | 5.0 |
| **Dental Health** | 714.0 | 3.009804 | 1.361117 | -1.0 | 2.0 | 3.0 | 4.0 | 6.0 |
| **Employment** | 714.0 | 2.806723 | 0.586582 | 1.0 | 3.0 | 3.0 | 3.0 | 4.0 |
| **Stress Keeps Patient from Sleeping** | 714.0 | 0.247899 | 0.432096 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| **Medication Keeps Patient from Sleeping** | 714.0 | 0.056022 | 0.230126 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| **Pain Keeps Patient from Sleeping** | 714.0 | 0.218487 | 0.413510 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| **Bathroom Needs Keeps Patient from Sleeping** | 714.0 | 0.504202 | 0.500333 | 0.0 | 0.0 | 1.0 | 1.0 | 1.0 |
| **Uknown Keeps Patient from Sleeping** | 714.0 | 0.417367 | 0.493470 | 0.0 | 0.0 | 0.0 | 1.0 | 1.0 |
| **Trouble Sleeping** | 714.0 | 2.407563 | 0.670349 | -1.0 | 2.0 | 3.0 | 3.0 | 3.0 |
| **Prescription Sleep Medication** | 714.0 | 2.829132 | 0.546767 | -1.0 | 3.0 | 3.0 | 3.0 | 3.0 |
| **Race** | 714.0 | 1.425770 | 1.003896 | 1.0 | 1.0 | 1.0 | 1.0 | 5.0 |
| **Gender** | 714.0 | 1.550420 | 0.497800 | 1.0 | 1.0 | 2.0 | 2.0 | 2.0 |

```
df.describe()
```

| | Number of Doctors Visited | Age | Phyiscal Health | Mental Health | Dental Health | Employment | Stress Keeps Patient from Sleeping | Medication Keeps Patient from Sleeping | Pain Keeps Patient from Sleeping | Bathroom Needs Keeps Patient from Sleeping | Uknown Keeps Patient from Sleeping | S |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **count** | 714.000000 | 714.0 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 714.000000 | 71 |
| **mean** | 2.112045 | 2.0 | 2.794118 | 1.988796 | 3.009804 | 2.806723 | 0.247899 | 0.056022 | 0.218487 | 0.504202 | 0.417367 | |
| **std** | 0.683441 | 0.0 | 0.900939 | 0.939928 | 1.361117 | 0.586582 | 0.432096 | 0.230126 | 0.413510 | 0.500333 | 0.493470 | |
| **min** | 1.000000 | 2.0 | -1.000000 | -1.000000 | -1.000000 | 1.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | - |
| **25%** | 2.000000 | 2.0 | 2.000000 | 1.000000 | 2.000000 | 3.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | |
| **50%** | 2.000000 | 2.0 | 3.000000 | 2.000000 | 3.000000 | 3.000000 | 0.000000 | 0.000000 | 0.000000 | 1.000000 | 0.000000 | |
| **75%** | 3.000000 | 2.0 | 3.000000 | 3.000000 | 4.000000 | 3.000000 | 0.000000 | 0.000000 | 0.000000 | 1.000000 | 1.000000 | |
| **max** | 3.000000 | 2.0 | 5.000000 | 5.000000 | 6.000000 | 4.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | 1.000000 | |

```
#checking for missing values
df.isnull().sum()
```

| | 0 |
|---|---|
| **Number of Doctors Visited** | 0 |
| **Age** | 0 |
| **Phyiscal Health** | 0 |
| **Mental Health** | 0 |
| **Dental Health** | 0 |
| **Employment** | 0 |
| **Stress Keeps Patient from Sleeping** | 0 |
| **Medication Keeps Patient from Sleeping** | 0 |
| **Pain Keeps Patient from Sleeping** | 0 |
| **Bathroom Needs Keeps Patient from Sleeping** | 0 |
| **Uknown Keeps Patient from Sleeping** | 0 |
| **Trouble Sleeping** | 0 |
| **Prescription Sleep Medication** | 0 |
| **Race** | 0 |
| **Gender** | 0 |

```
#checking for missing values
df.isnull().sum()/df.shape[0]*100
```

| | 0 |
|---|---|
| **Number of Doctors Visited** | 0.0 |
| **Age** | 0.0 |
| **Phyiscal Health** | 0.0 |
| **Mental Health** | 0.0 |
| **Dental Health** | 0.0 |
| **Employment** | 0.0 |
| **Stress Keeps Patient from Sleeping** | 0.0 |
| **Medication Keeps Patient from Sleeping** | 0.0 |
| **Pain Keeps Patient from Sleeping** | 0.0 |
| **Bathroom Needs Keeps Patient from Sleeping** | 0.0 |
| **Uknown Keeps Patient from Sleeping** | 0.0 |
| **Trouble Sleeping** | 0.0 |
| **Prescription Sleep Medication** | 0.0 |
| **Race** | 0.0 |
| **Gender** | 0.0 |

```
#checking for duplicates in the dataset
df.duplicated().sum()
```

```
42
```

```
#dropping or eliminating the duplicated values
df.drop_duplicates()
```

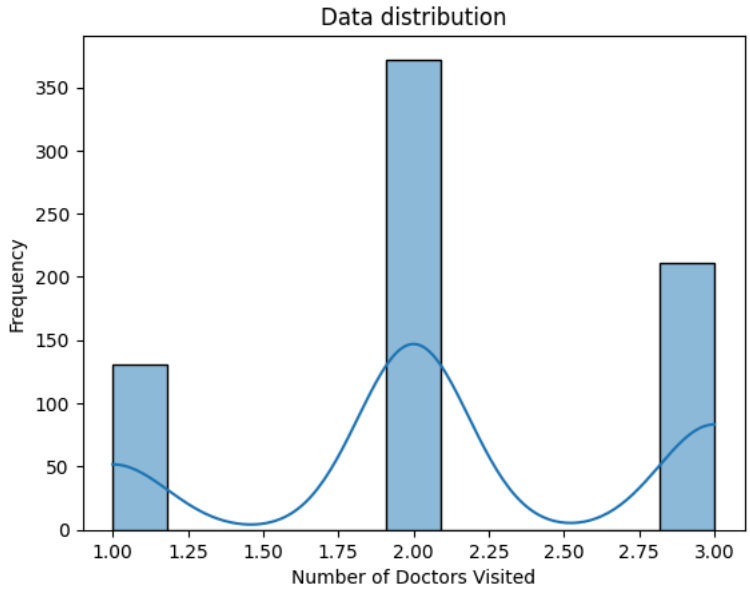
| | Number of Doctors Visited | Age | Phyiscal Health | Mental Health | Dental Health | Employment | Stress Keeps Patient from Sleeping | Medication Keeps Patient from Sleeping | Pain Keeps Patient from Sleeping | Bathroom Needs Keeps Patient from Sleeping | Uknown Keeps Patient from Sleeping | Trouble Sleeping | Prescription Sleep Medication |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 3 | 2 | 4 | 3 | 3 | 3 | 0 | 0 | 0 | 0 | 1 | 2 | 3 |
| **1** | 2 | 2 | 4 | 2 | 3 | 3 | 1 | 0 | 0 | 1 | 0 | 3 | 3 |
| **2** | 3 | 2 | 3 | 2 | 3 | 3 | 0 | 0 | 0 | 0 | 1 | 3 | 3 |
| **3** | 1 | 2 | 3 | 2 | 3 | 3 | 0 | 0 | 0 | 1 | 0 | 3 | 3 |
| **4** | 3 | 2 | 3 | 3 | 3 | 3 | 1 | 0 | 0 | 0 | 0 | 2 | 3 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **706** | 3 | 2 | 4 | 2 | 2 | 3 | 0 | 0 | 1 | 1 | 0 | -1 | 3 |
| **710** | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | 2 | 3 |
| **711** | 3 | 2 | 4 | 2 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 3 | 3 |
| **712** | 3 | 2 | 3 | 1 | 3 | 3 | 1 | 0 | 1 | 1 | 1 | 3 | 3 |
| **713** | 3 | 2 | 3 | 2 | 3 | 1 | 1 | 0 | 1 | 1 | 0 | 2 | 3 |

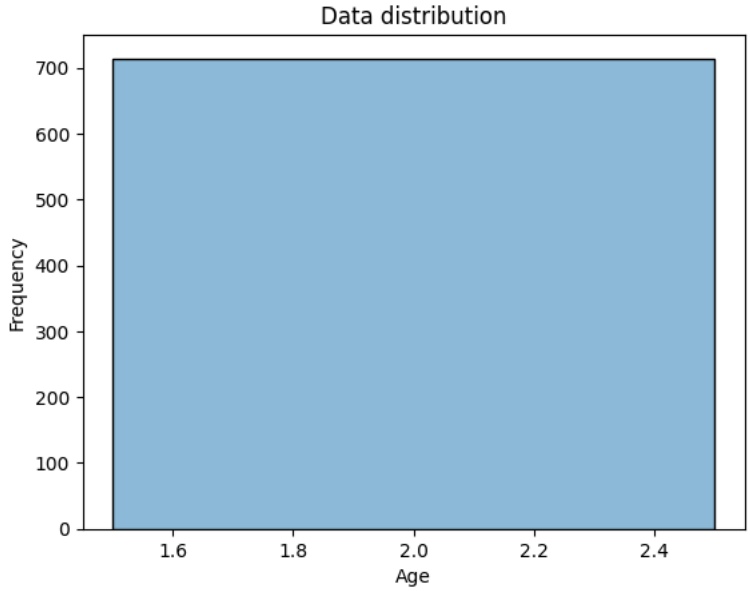
```
#checking for garbage values
for i in df.select_dtypes(include='object').columns:
  print(df[i].value_counts())
  print("***"*10)
```

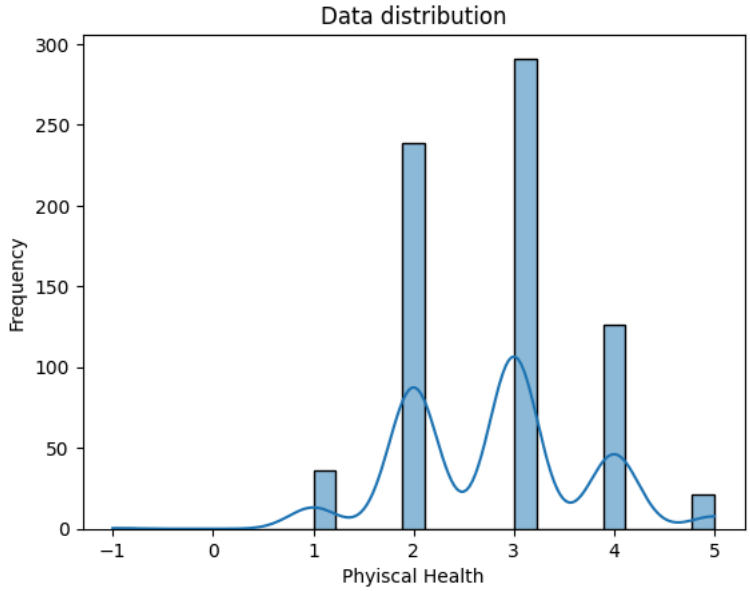
**Exploratory Data Analysis (EDA)**

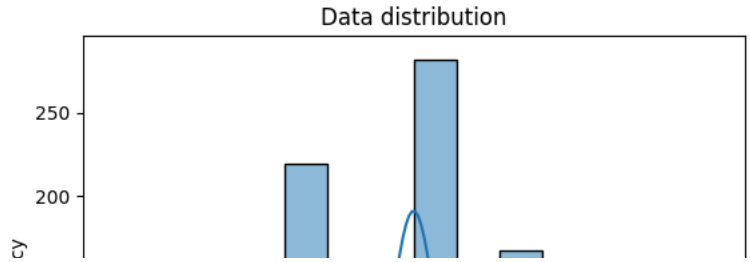
```
#understanding the distribution of the data for each numerical column using a histogram
for i in df.select_dtypes(include='number').columns:
  sns.histplot(data=df, x=i, kde=True)
  plt.ylabel('Frequency')
  plt.title('Data distribution')
  plt.show()
```

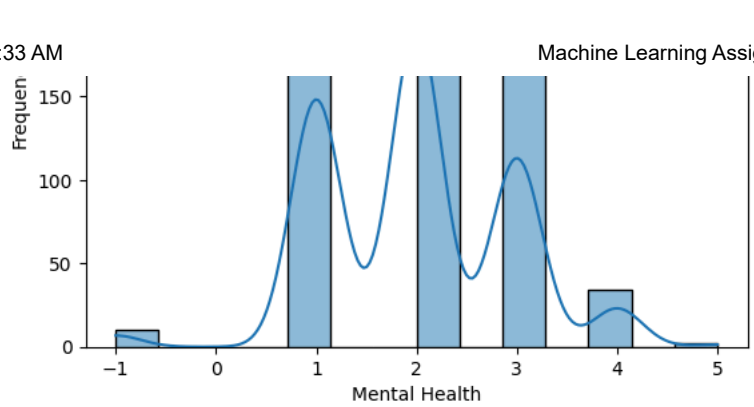
[+]

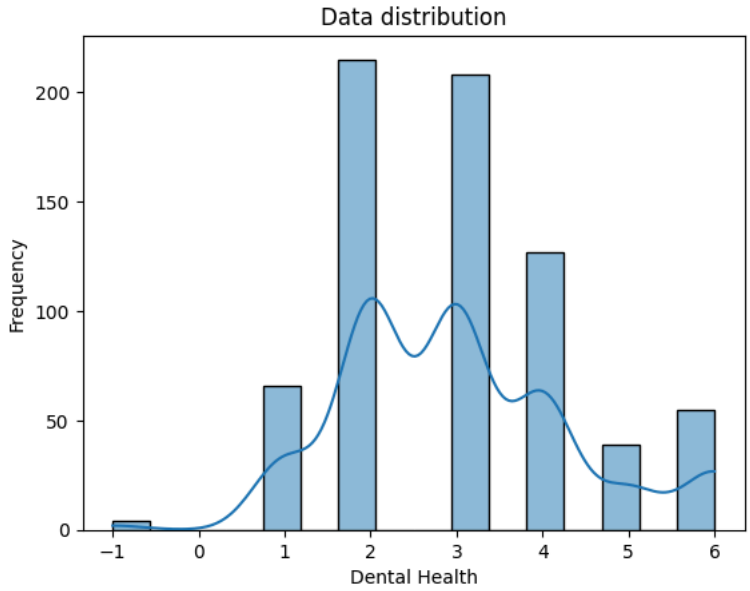

Data distribution
Frequency
Number of Doctors Visited

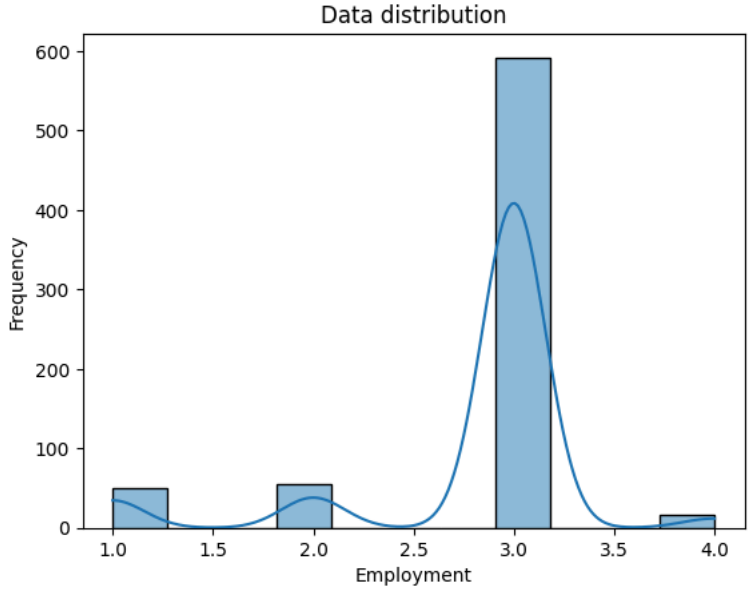

Data distribution
Frequency
Age

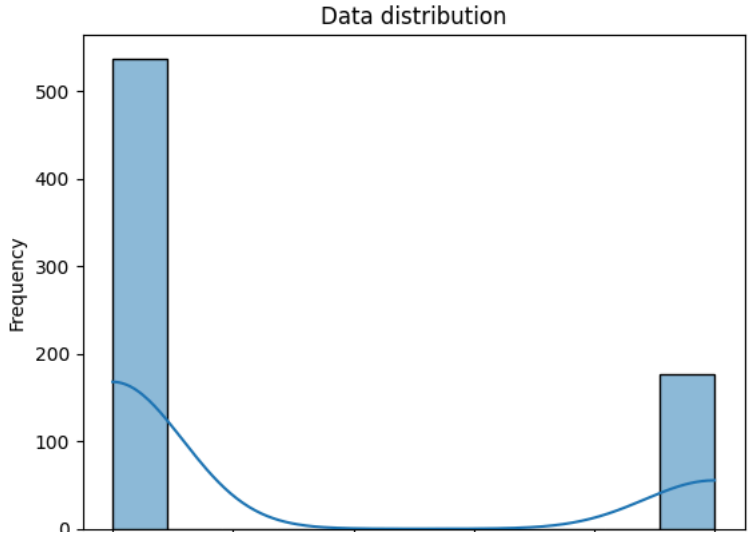

Data distribution
Frequency
Phyiscal Health

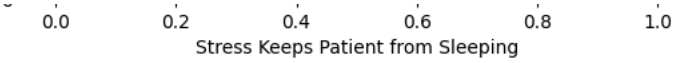

Data distribution

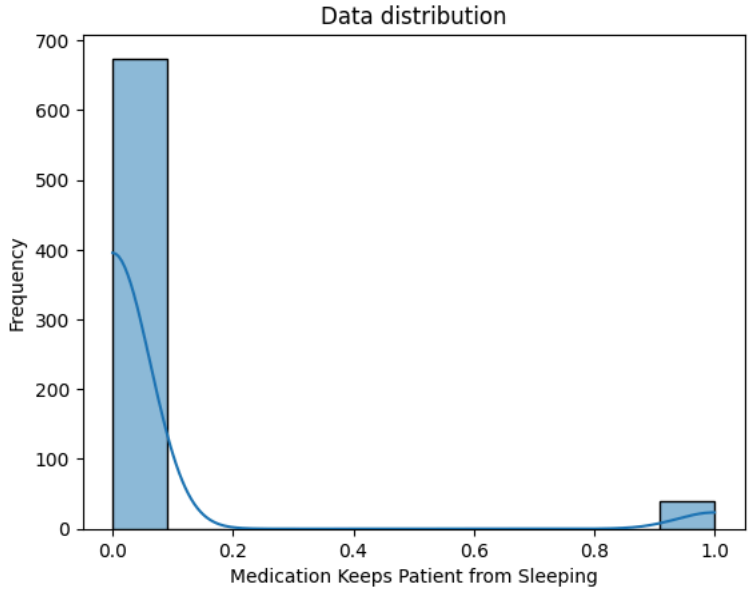

Frequen
150
100
50
0
−1
0
1
2
3
4
5
Mental Health

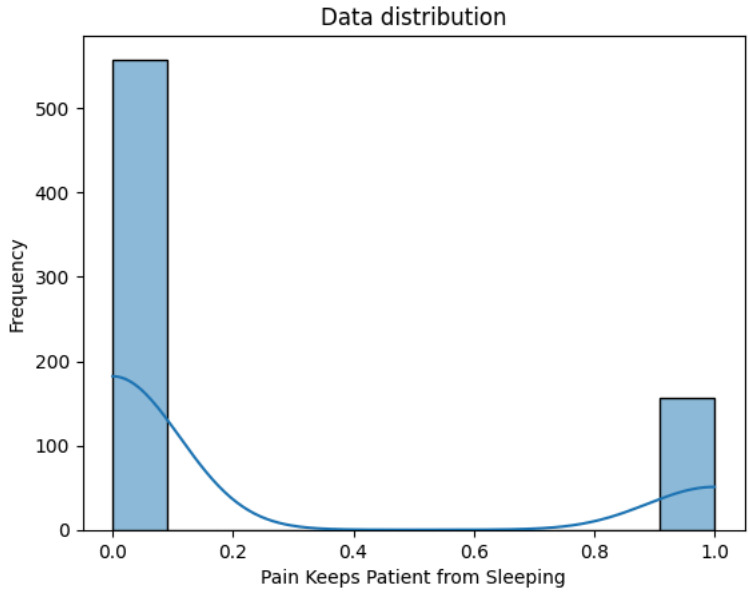

Data distribution
Frequency
200
150
100
50
0
−1
0
1
2
3
4
5
6
Dental Health

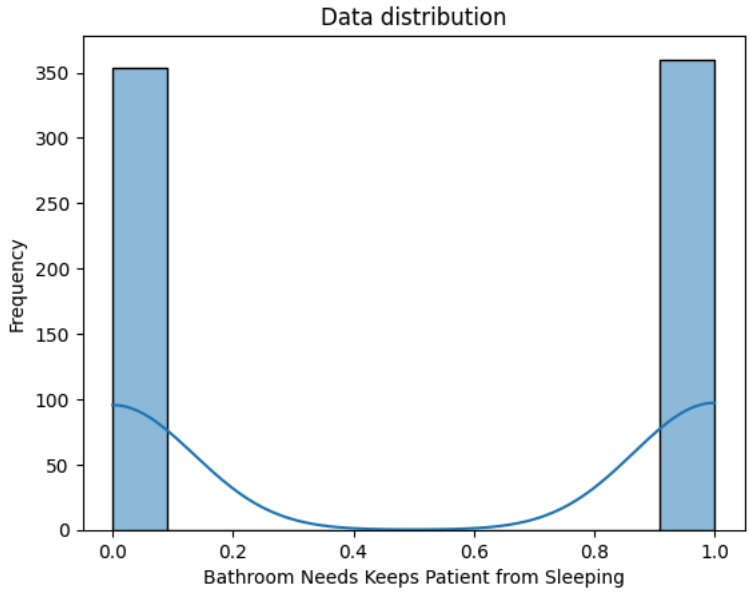

Data distribution
Frequency
600
500
400
300
200
100
0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Employment

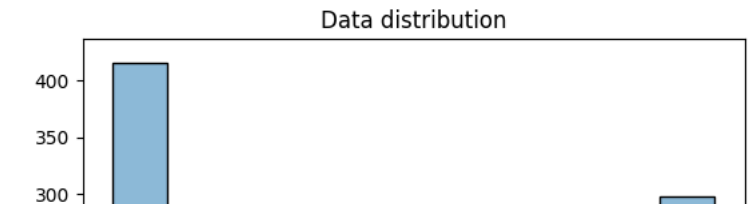

Data distribution
Frequency
500
400
300
200
100
0

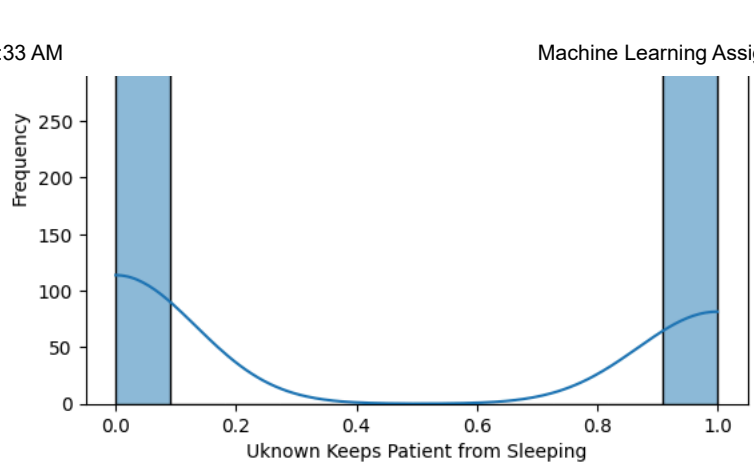

0.0
0.2
0.4
0.6
0.8
1.0
Stress Keeps Patient from Sleeping

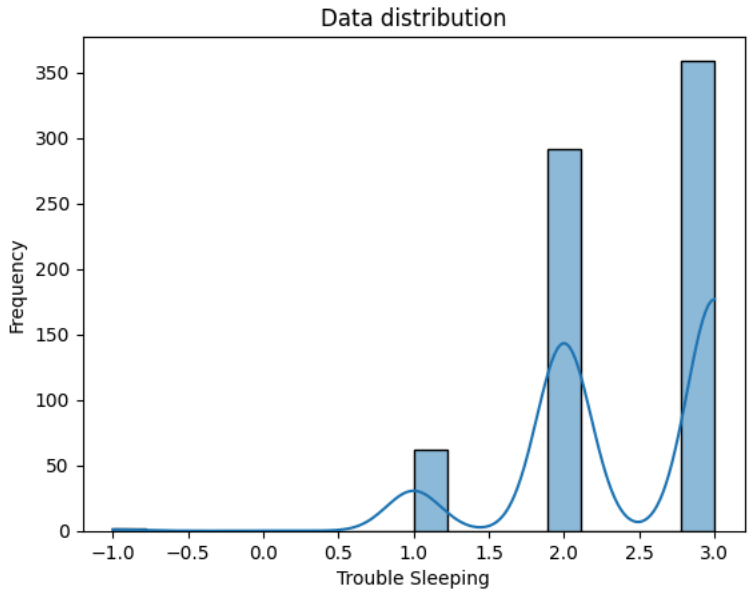

Data distribution
Frequency
Medication Keeps Patient from Sleeping

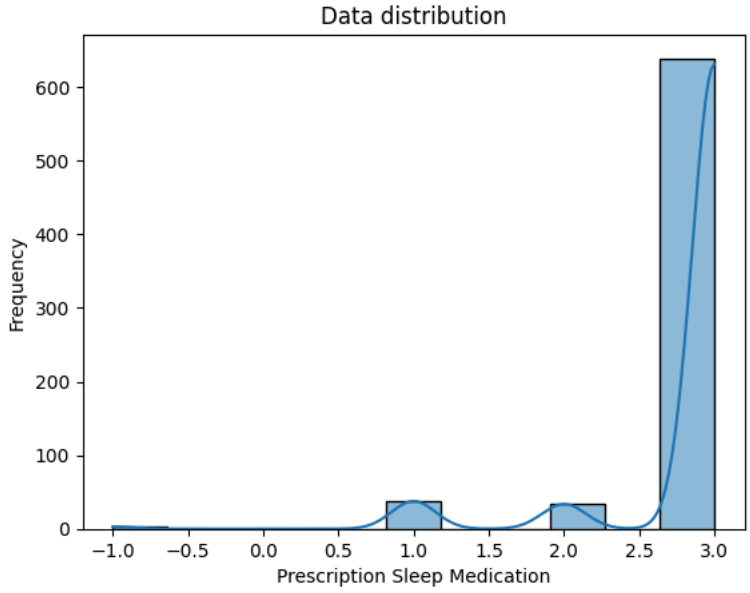

Data distribution
Frequency
Pain Keeps Patient from Sleeping

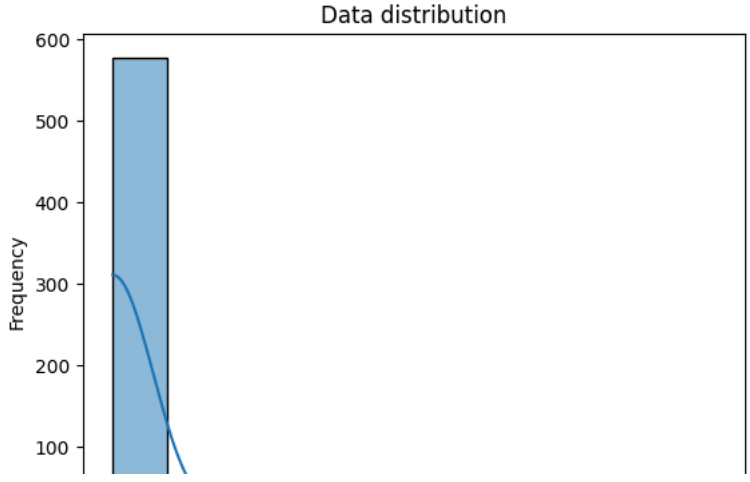

Data distribution
Frequency
Bathroom Needs Keeps Patient from Sleeping

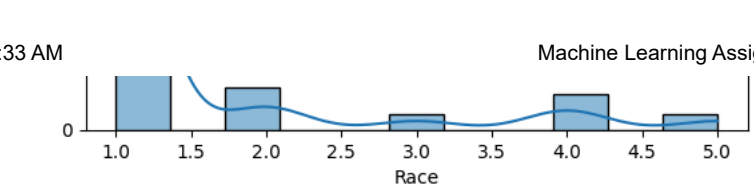

Data distribution

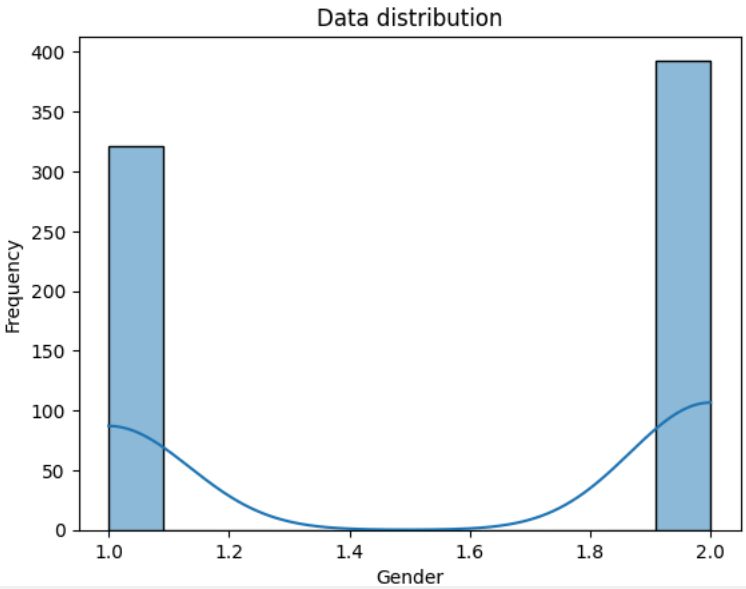

Frequency
Uknown Keeps Patient from Sleeping


Data distribution
Frequency
Trouble Sleeping

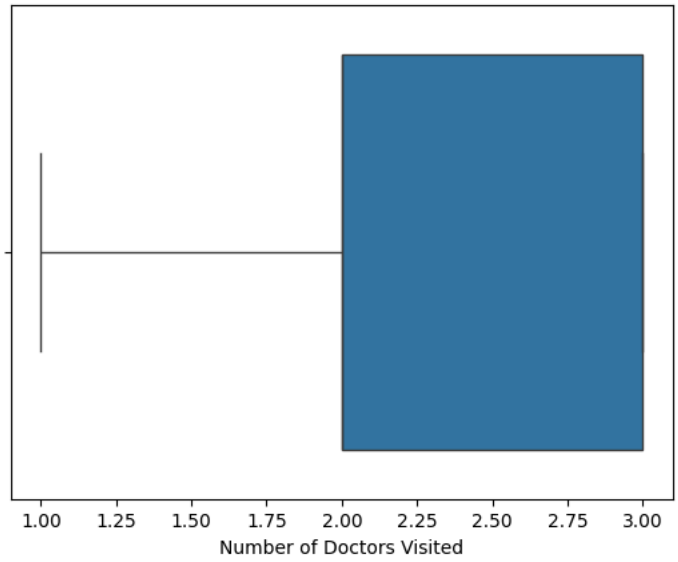

Data distribution
Frequency
Prescription Sleep Medication

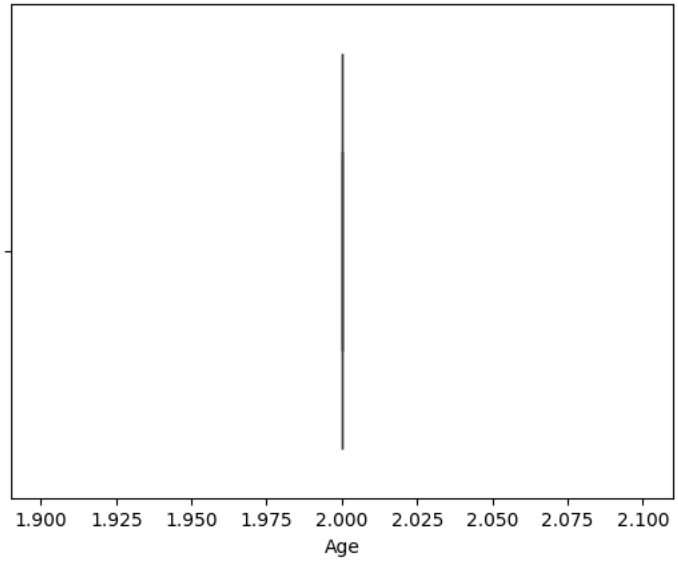

Data distribution
Frequency

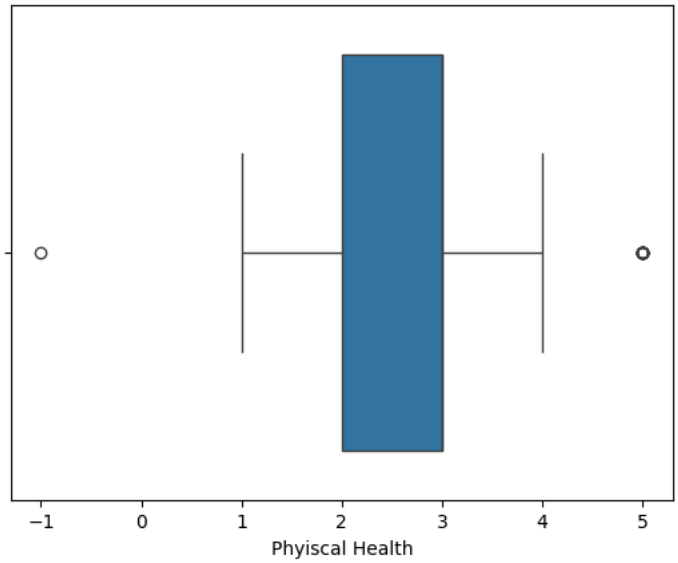

0
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Race

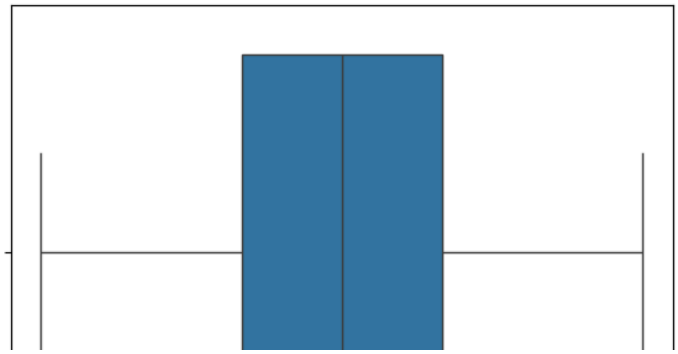

Data distribution
400
350
300
250
200
150
100
50
0
Frequency
1.0
1.2
1.4
1.6
1.8
2.0
Gender

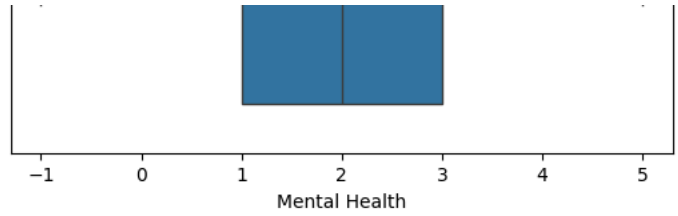
```
#identifying outliers in the dataset
for i in df.select_dtypes(include='number').columns:
  sns.boxplot(data=df, x=i)
  plt.show()
```

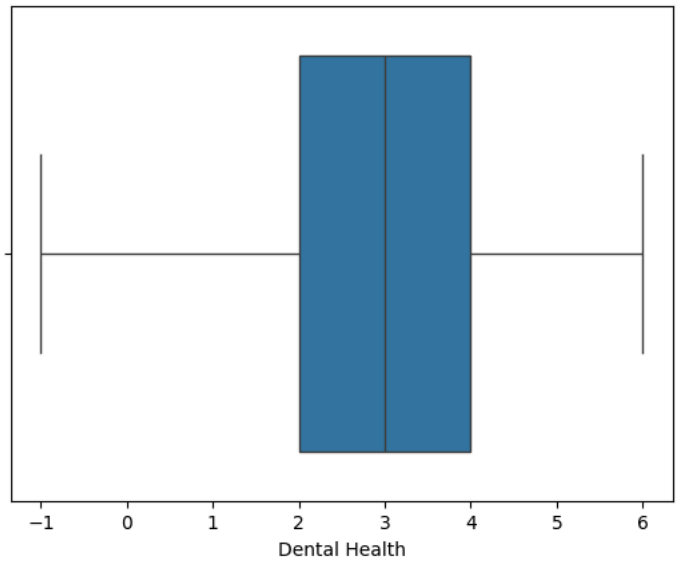

1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
Number of Doctors Visited

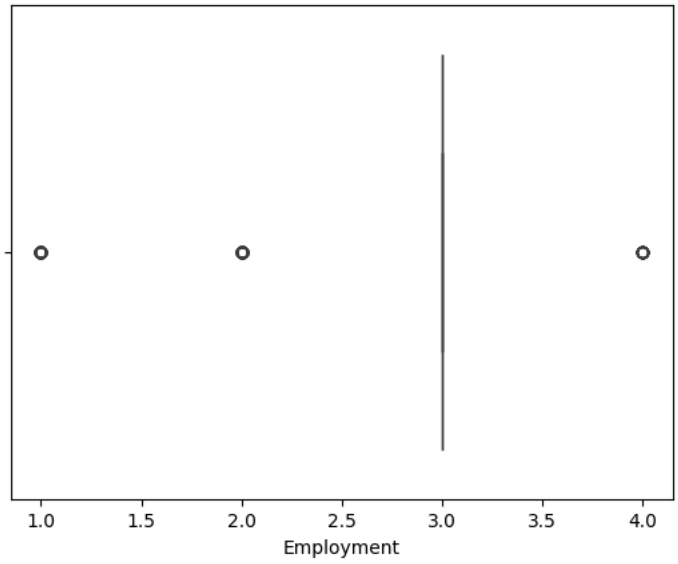

1.900
1.925
1.950
1.975
2.000
2.025
2.050
2.075
2.100
Age

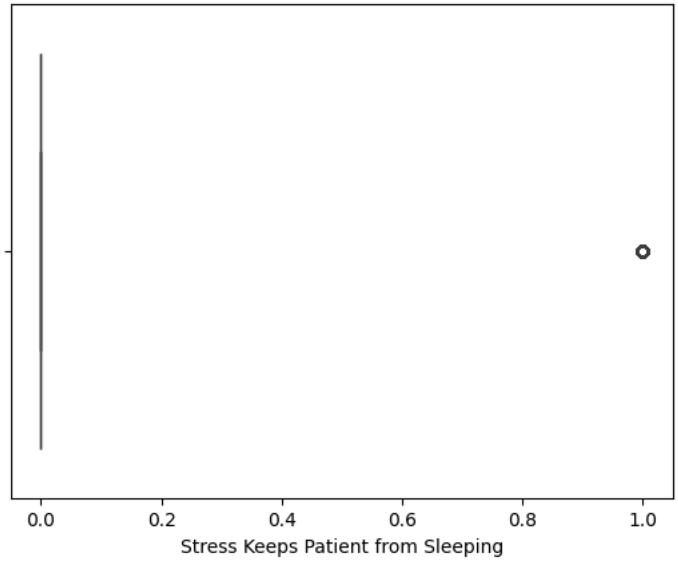

−1
0
1
2
3
4
5
Phyiscal Health

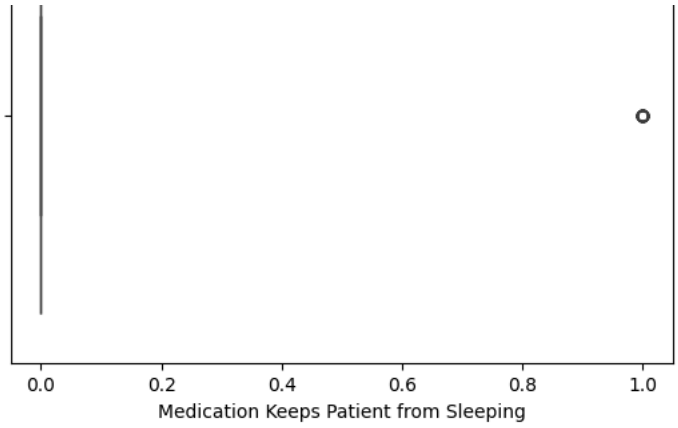


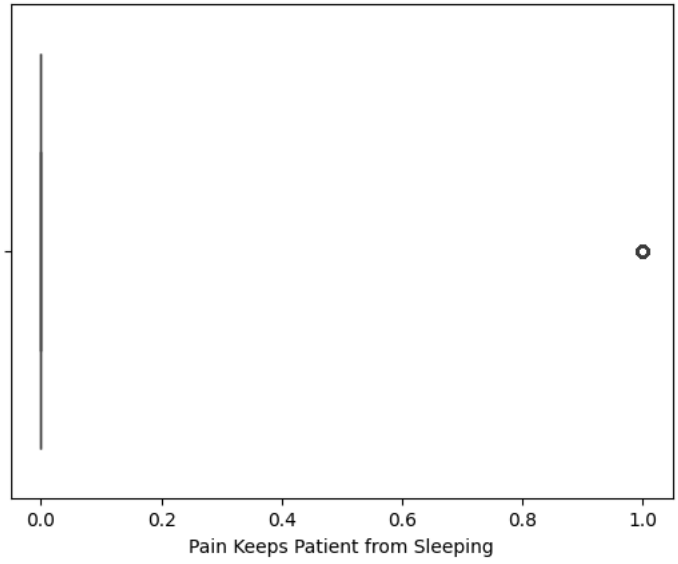

−1
0
1
2
3
4
5
Mental Health

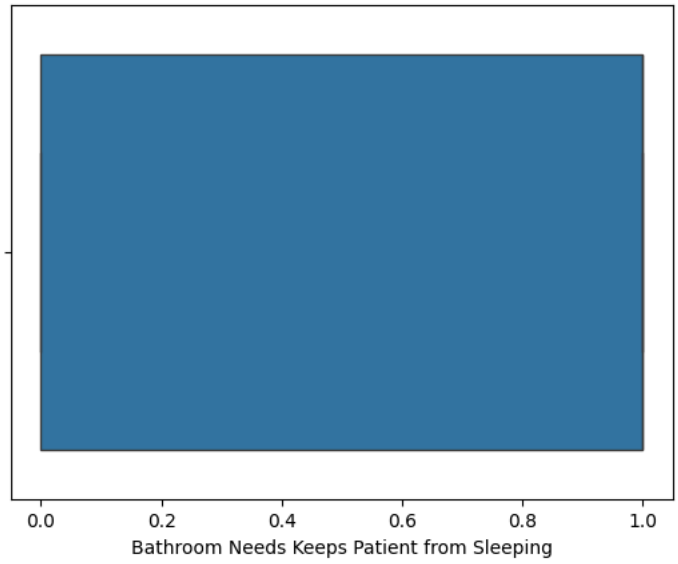

−1
0
1
2
3
4
5
6
Dental Health

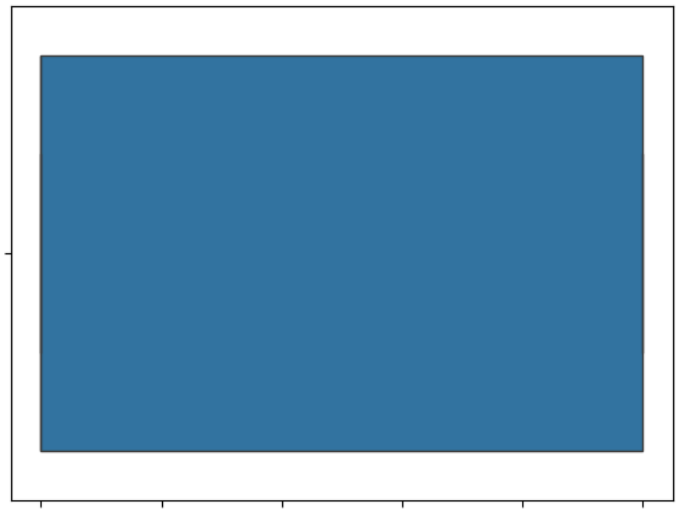

1.0
1.5
2.0
2.5
3.0
3.5
4.0
Employment

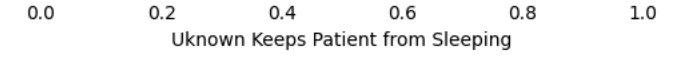

0.0
0.2
0.4
0.6
0.8
1.0
Stress Keeps Patient from Sleeping

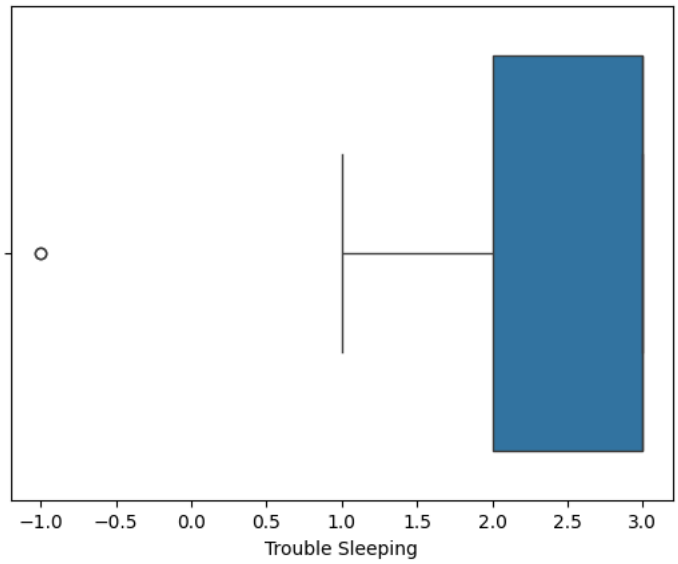

0.0
0.2
0.4
0.6
0.8
1.0
Medication Keeps Patient from Sleeping

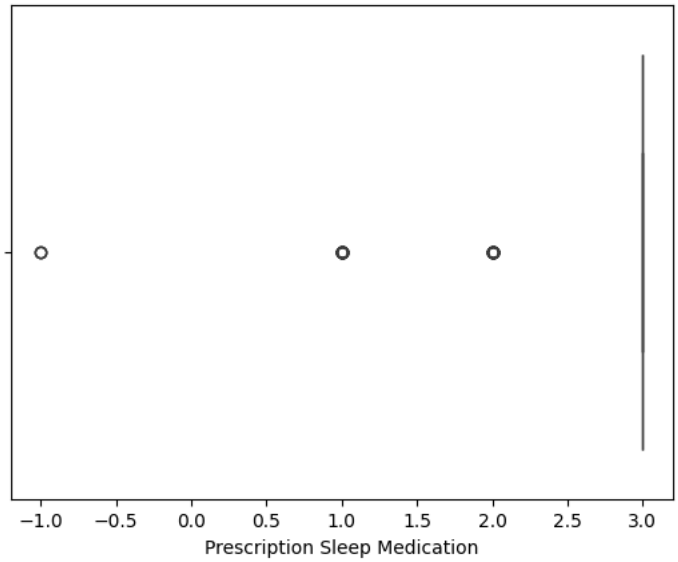

0.0
0.2
0.4
0.6
0.8
1.0
Pain Keeps Patient from Sleeping

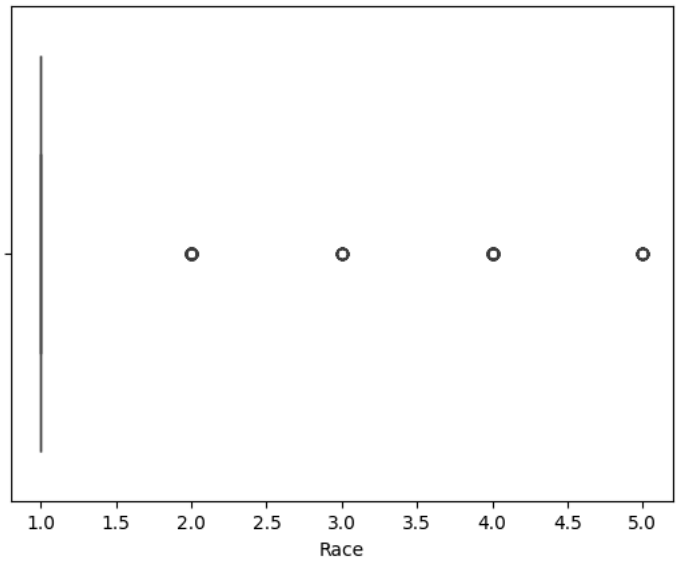

0.0
0.2
0.4
0.6
0.8
1.0
Bathroom Needs Keeps Patient from Sleeping

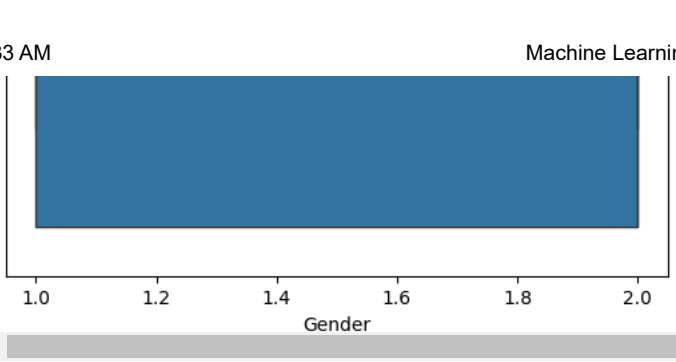



0.0
0.2
0.4
0.6
0.8
1.0
Uknown Keeps Patient from Sleeping


−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Trouble Sleeping


−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Prescription Sleep Medication


1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Race


1.0
1.2
1.4
1.6
1.8
2.0
Gender

```
#dealing with the outliers
#def out_liers(col):
#  q1,q3=np.percentile(col,[25,75])
#  iqr=q3-q1
 # upper_bound=q3+(1.5*iqr)
  #lower_bound=q1-(1.5*iqr)
  #return upper_bound,lower_bound
```

```
#for i in ['Phyiscal Health', 'Employment', 'Stress Keeps Patient from Sleeping', 'Medication Keeps Patient from Sleeping',
 #         'Pain Keeps Patient from Sleeping', 'Trouble Sleeping', 'Prescription Sleep Medication', 'Race']:
  #upper_bound,lower_bound=out_liers(df[i])
  #df[i]=np.where(df[i]>upper_bound,upper_bound,df[i])
  #df[i]=np.where(df[i]<lower_bound,lower_bound,df[i])
```

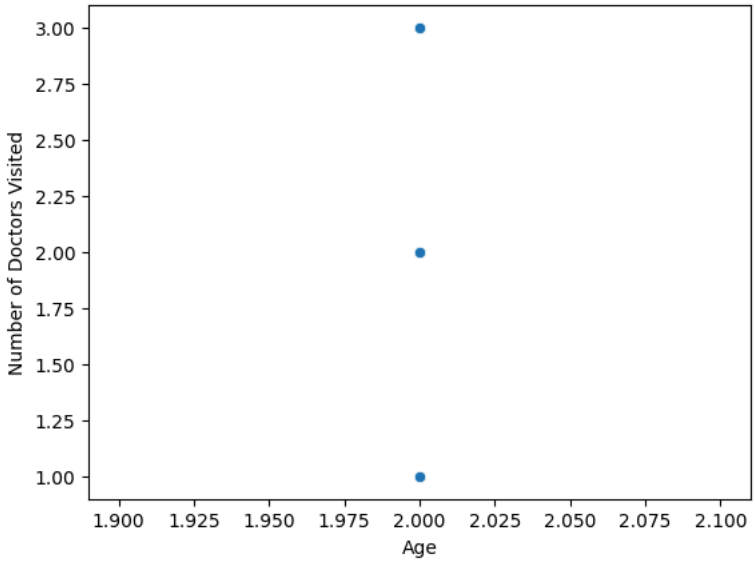
```
#for i in ['Phyiscal Health', 'Employment', 'Stress Keeps Patient from Sleeping', 'Medication Keeps Patient from Sleeping',
    #      'Pain Keeps Patient from Sleeping', 'Trouble Sleeping', 'Prescription Sleep Medication', 'Race']:
    #      sns.boxplot(df[i])
     #     plt.show()
```

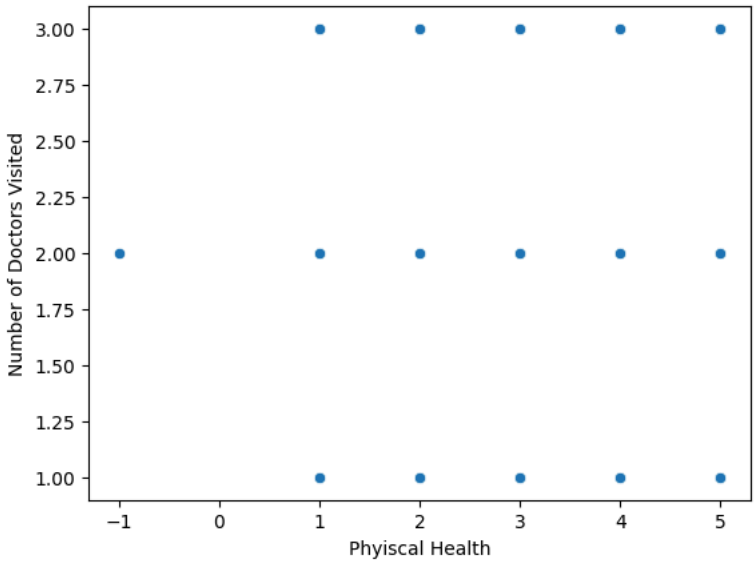
```
df.select_dtypes(include='number').columns
```

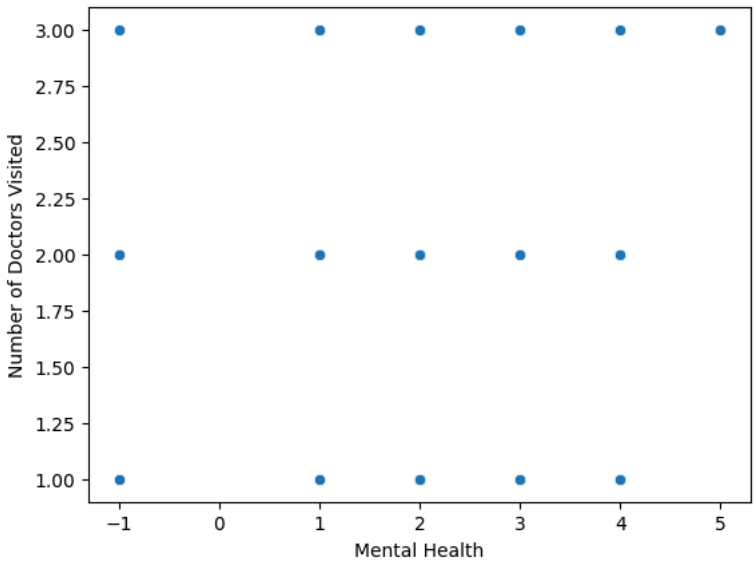
```
Index(['Number of Doctors Visited', 'Age', 'Phyiscal Health', 'Mental Health',
       'Dental Health', 'Employment', 'Stress Keeps Patient from Sleeping',
       'Medication Keeps Patient from Sleeping',
       'Pain Keeps Patient from Sleeping',
       'Bathroom Needs Keeps Patient from Sleeping',
       'Uknown Keeps Patient from Sleeping', 'Trouble Sleeping',
       'Prescription Sleep Medication', 'Race', 'Gender'],
      dtype='object')
```

```
#scatter plot to understand the relationship between my target variable and other variables
for i in ['Age', 'Phyiscal Health', 'Mental Health',
       'Dental Health', 'Employment', 'Stress Keeps Patient from Sleeping',
       'Medication Keeps Patient from Sleeping',
       'Pain Keeps Patient from Sleeping',
       'Bathroom Needs Keeps Patient from Sleeping',
       'Uknown Keeps Patient from Sleeping', 'Trouble Sleeping',
       'Prescription Sleep Medication', 'Race', 'Gender']:
       sns.scatterplot(data=df, x=i, y='Number of Doctors Visited')
       plt.show()
```

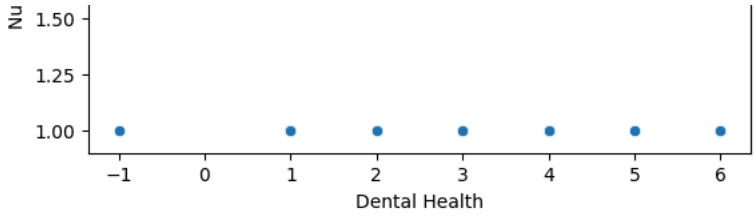
[4]

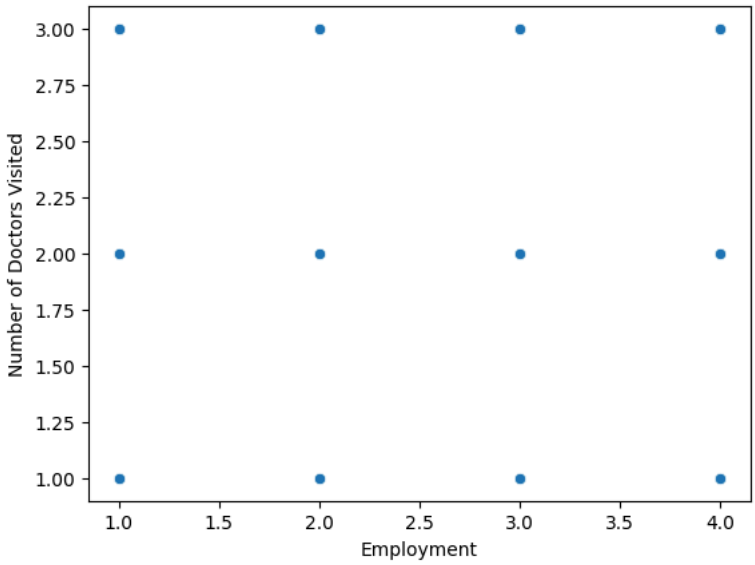

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
1.900
1.925
1.950
1.975
2.000
2.025
2.050
2.075
2.100
Age

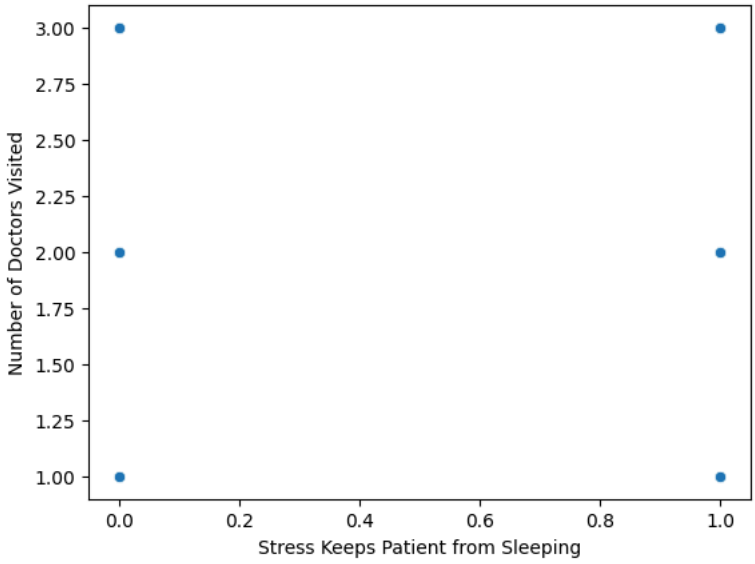

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
−1
0
1
2
3
4
5
Phyiscal Health

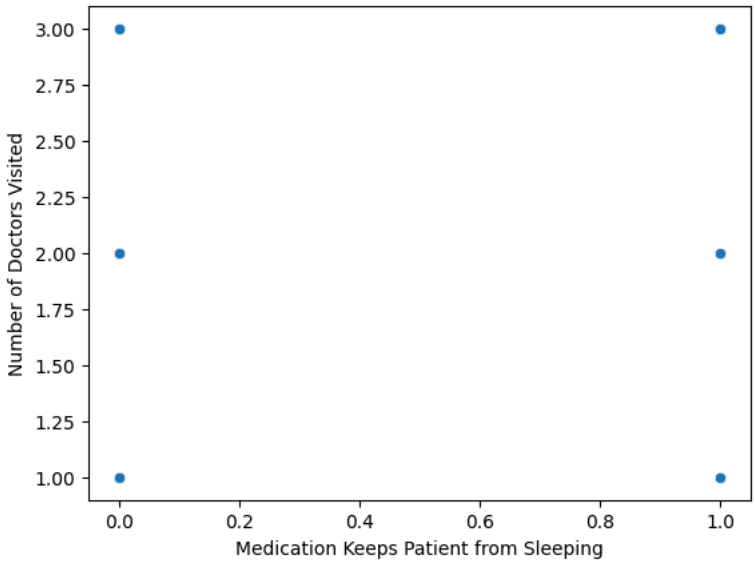

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
−1
0
1
2
3
4
5
Mental Health


mber of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75

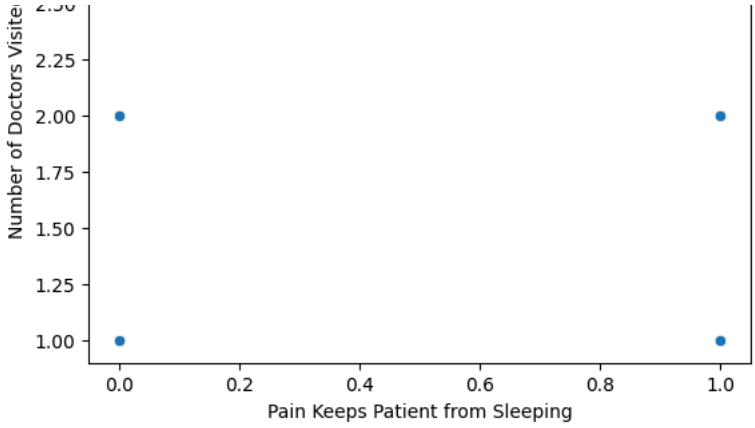

Nu
1.50
1.25
1.00
−1
0
1
2
3
4
5
6
Dental Health

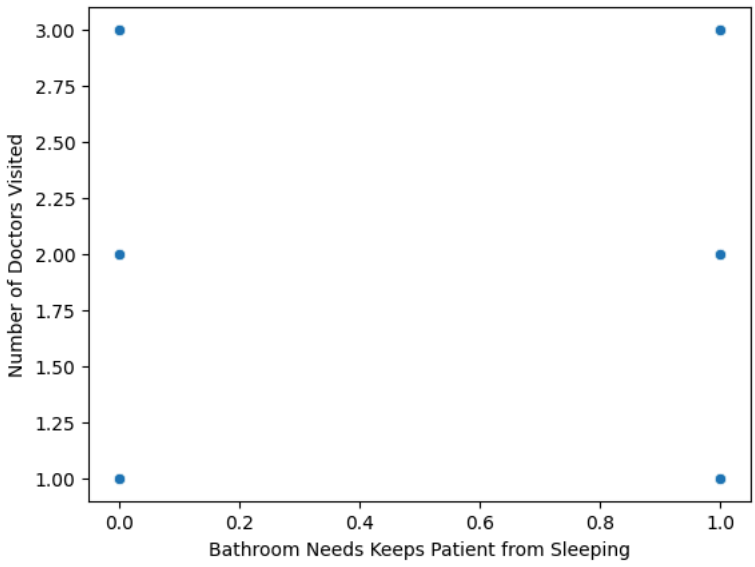

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Employment

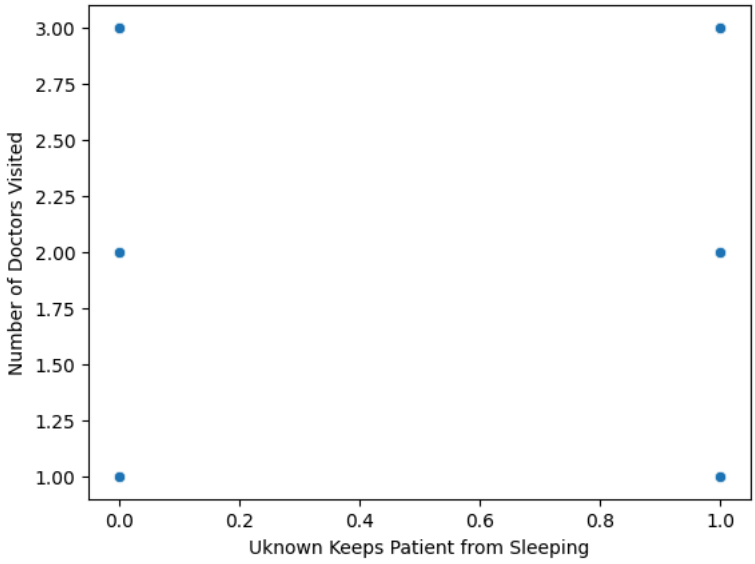

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Stress Keeps Patient from Sleeping

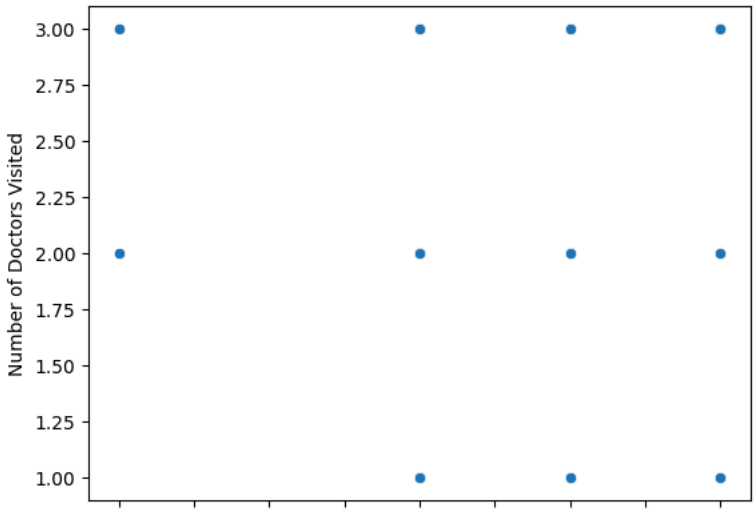

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Medication Keeps Patient from Sleeping

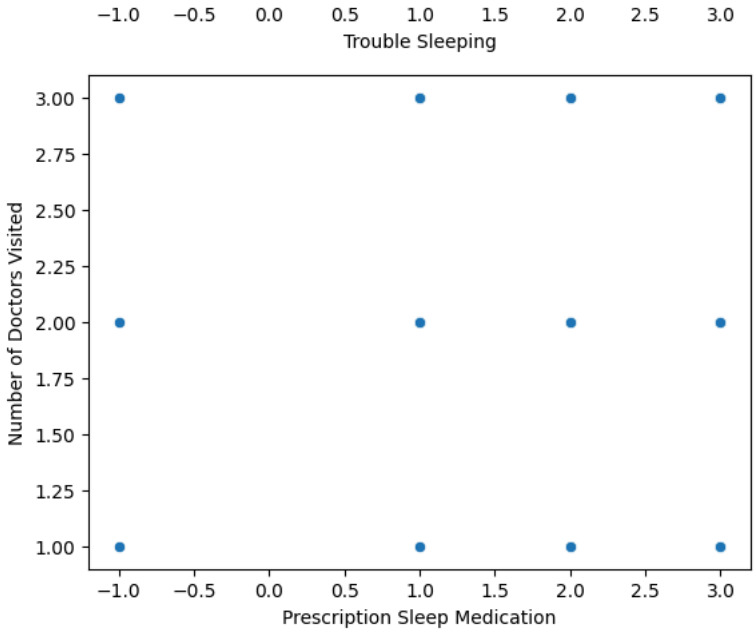

3.00
2.75

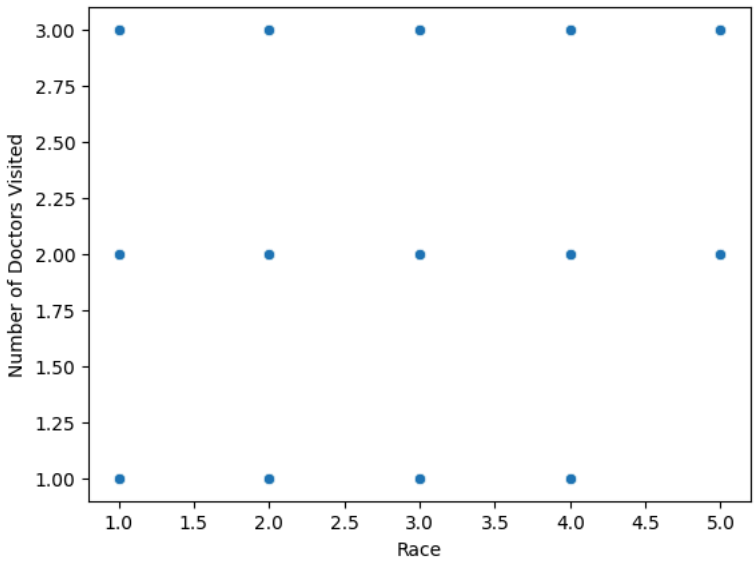

Number of Doctors Visited
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Pain Keeps Patient from Sleeping

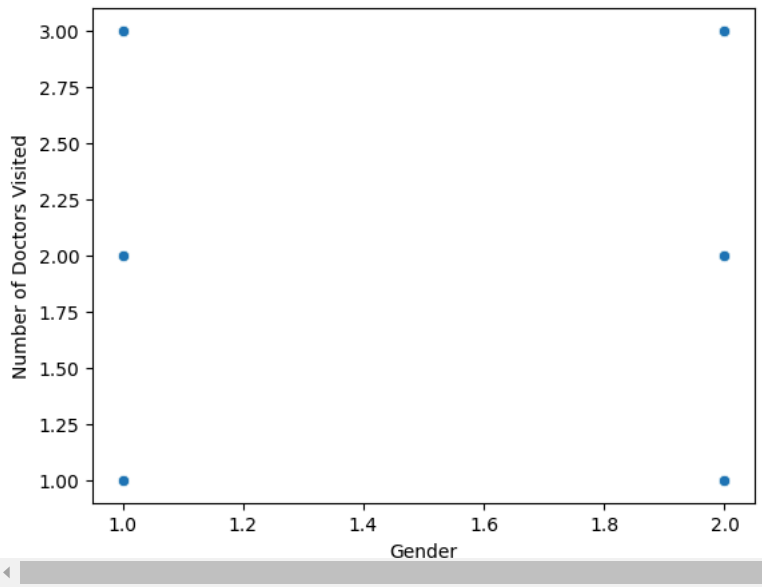

Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Bathroom Needs Keeps Patient from Sleeping


Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0.0
0.2
0.4
0.6
0.8
1.0
Uknown Keeps Patient from Sleeping


Number of Doctors Visited
3.00
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00


Trouble Sleeping
Number of Doctors Visited
Prescription Sleep Medication


Number of Doctors Visited
Race


Number of Doctors Visited
Gender

```
#checking for the correlation with heatmap to interpret the relation and multicollinearity
df.select_dtypes(include='number').corr()
```

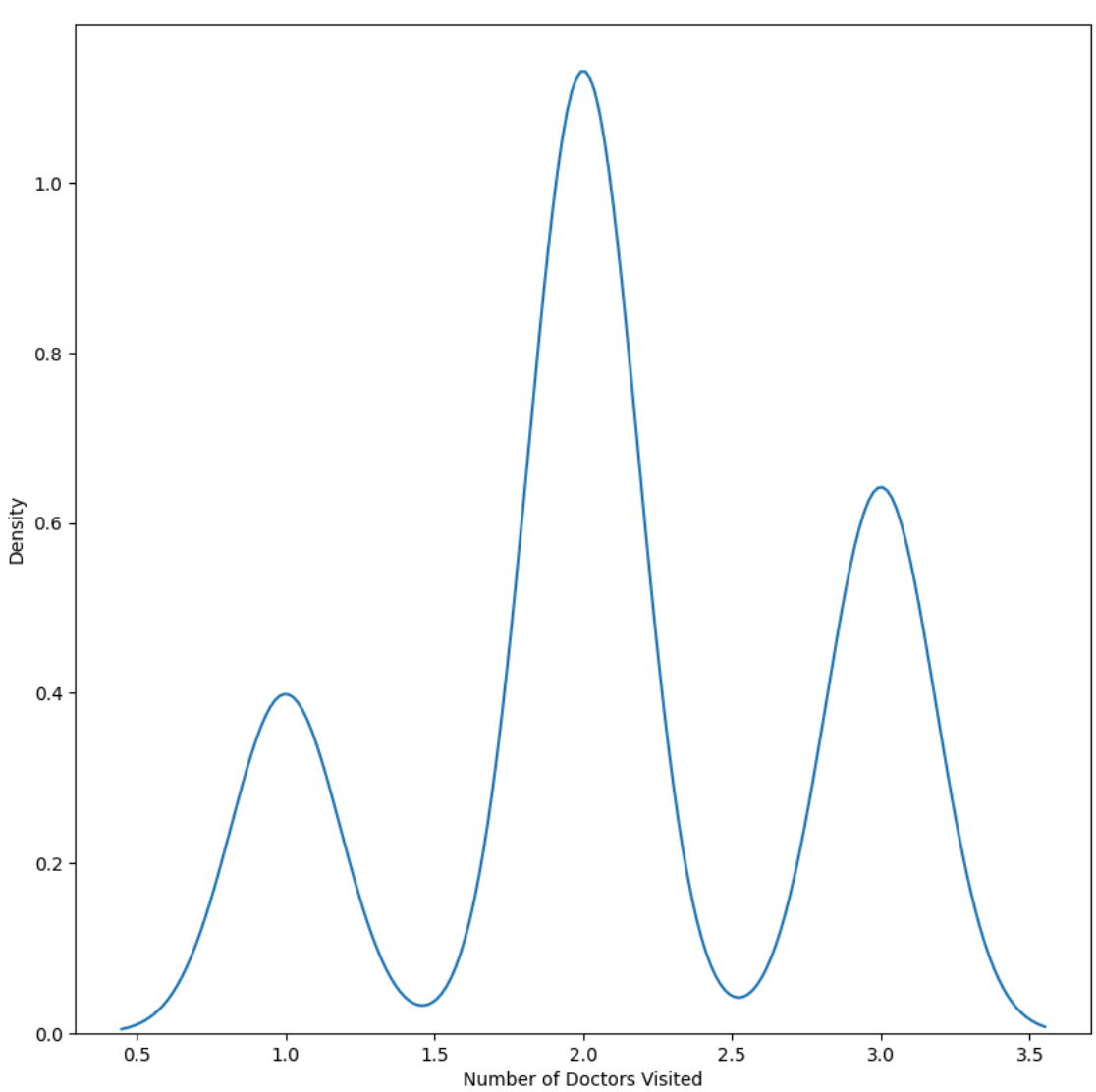
| | Number of Doctors Visited | Age | Phyiscal Health | Mental Health | Dental Health | Employment | Stress Keeps Patient from Sleeping | Medication Keeps Patient from Sleeping | Pain Keeps Patient from Sleeping | Bathroom Needs Keeps Patient from Sleeping | Uknown Keeps Patient from Sleeping | Trouble Sleeping |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Number of Doctors Visited** | 1.000000 | NaN | 0.169629 | 0.049990 | 0.009371 | 0.092578 | 0.053040 | 0.120549 | 0.081990 | 0.056043 | -0.014095 | -0.063079 |
| **Age** | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| **Phyiscal Health** | 0.169629 | NaN | 1.000000 | 0.404705 | 0.404238 | 0.147526 | 0.034014 | 0.109827 | 0.275266 | 0.003477 | -0.011505 | -0.213855 |
| **Mental Health** | 0.049990 | NaN | 0.404705 | 1.000000 | 0.269770 | 0.077469 | 0.138074 | 0.139072 | 0.121780 | 0.044835 | -0.038285 | -0.110718 |
| **Dental Health** | 0.009371 | NaN | 0.404238 | 0.269770 | 1.000000 | 0.076156 | -0.018446 | 0.029588 | 0.080913 | -0.007269 | -0.014453 | -0.044351 |
| **Employment** | 0.092578 | NaN | 0.147526 | 0.077469 | 0.076156 | 1.000000 | -0.043106 | 0.059546 | -0.004908 | 0.012329 | 0.017427 | -0.116836 |
| **Stress Keeps Patient from Sleeping** | 0.053040 | NaN | 0.034014 | 0.138074 | -0.018446 | -0.043106 | 1.000000 | 0.029395 | 0.136015 | -0.001581 | -0.314897 | -0.150775 |
| **Medication Keeps Patient from Sleeping** | 0.120549 | NaN | 0.109827 | 0.139072 | 0.029588 | 0.059546 | 0.029395 | 1.000000 | 0.165965 | 0.071039 | -0.119734 | -0.148217 |
| **Pain Keeps Patient from Sleeping** | 0.081990 | NaN | 0.275266 | 0.121780 | 0.080913 | -0.004908 | 0.136015 | 0.165965 | 1.000000 | 0.144695 | -0.213823 | -0.235680 |
| **Bathroom Needs Keeps Patient from Sleeping** | 0.056043 | NaN | 0.003477 | 0.044835 | -0.007269 | 0.012329 | -0.001581 | 0.071039 | 0.144695 | 1.000000 | -0.382029 | 0.038795 |
| **Uknown Keeps Patient from Sleeping** | -0.014095 | NaN | -0.011505 | -0.038285 | -0.014453 | 0.017427 | -0.314897 | -0.119734 | -0.213823 | -0.382029 | 1.000000 | -0.023123 |
| **Trouble Sleeping** | -0.063079 | NaN | -0.213855 | -0.110718 | -0.044351 | -0.116836 | -0.150775 | -0.148217 | -0.235680 | 0.038795 | -0.023123 | 1.000000 |

```
plt.figure(figsize=(10,10))
sns.heatmap(df.select_dtypes(include='number').corr(), annot=True)
```

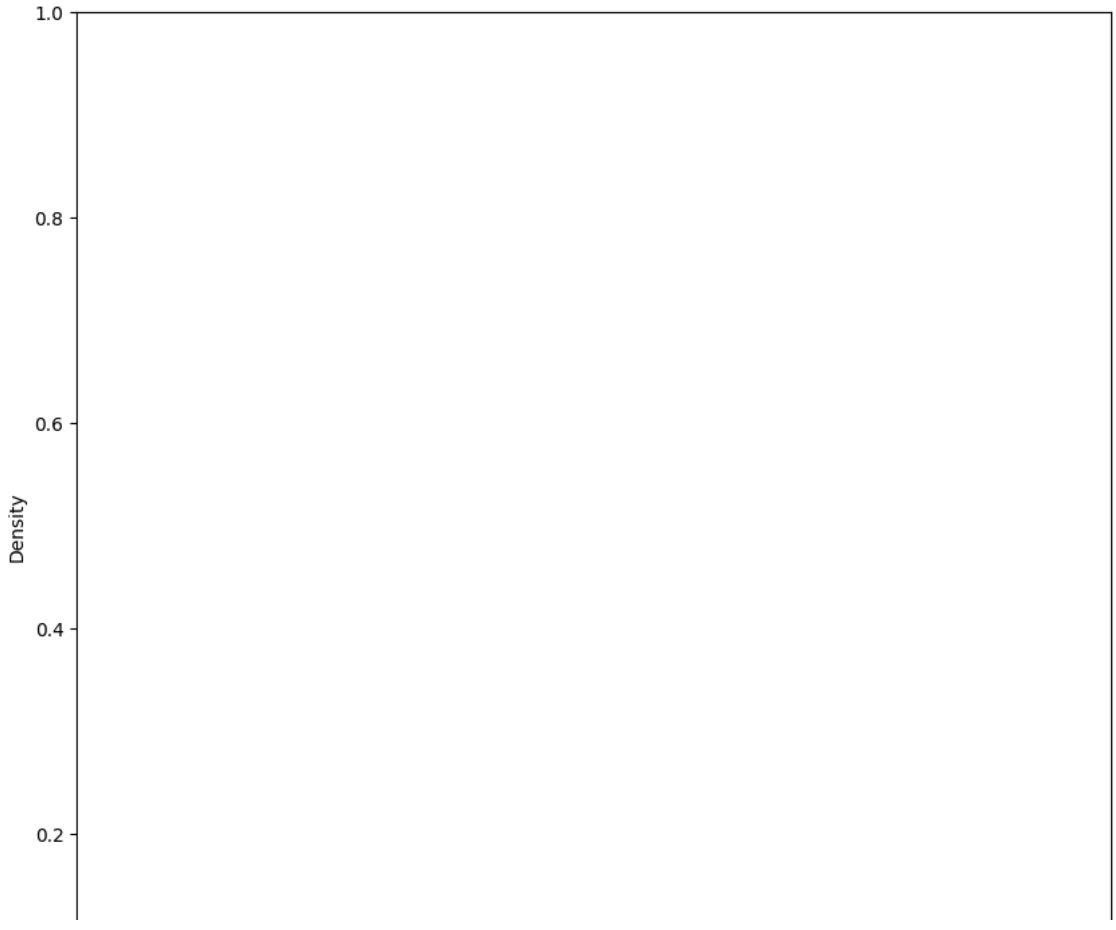
```
<Axes: >
```

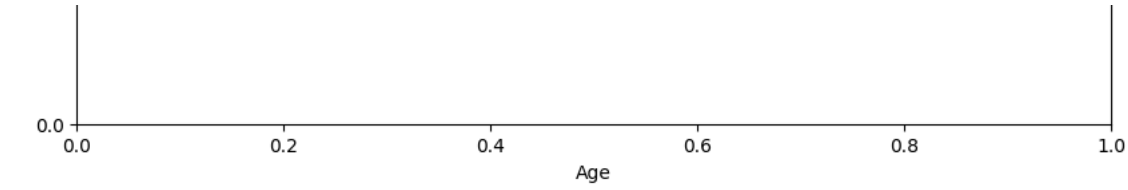

```
for i in df.select_dtypes(include='number').columns:
  plt.figure(figsize=(10,10))
  sns.kdeplot(data=df, x=i)
  plt.show()
```

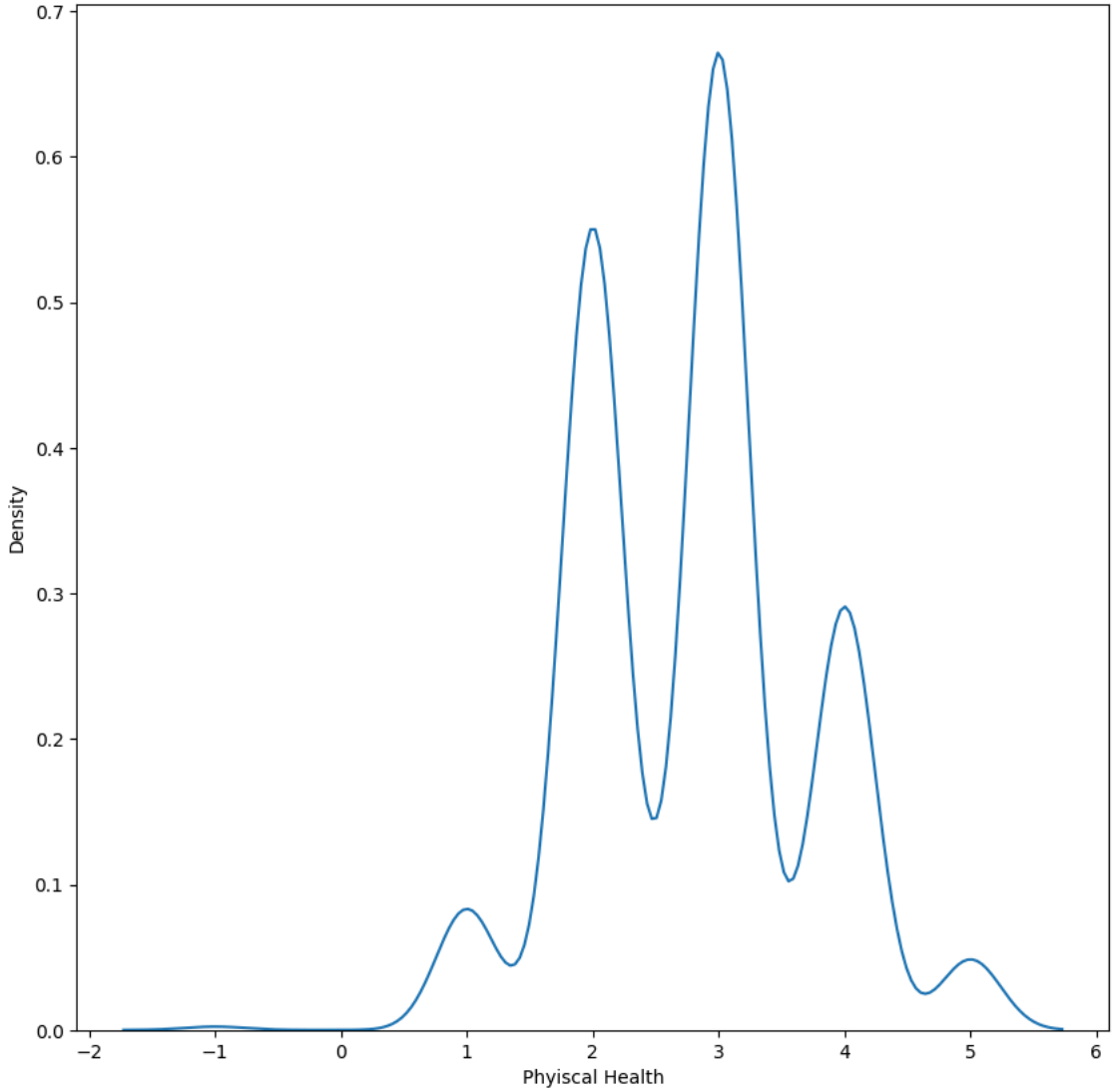
[+]

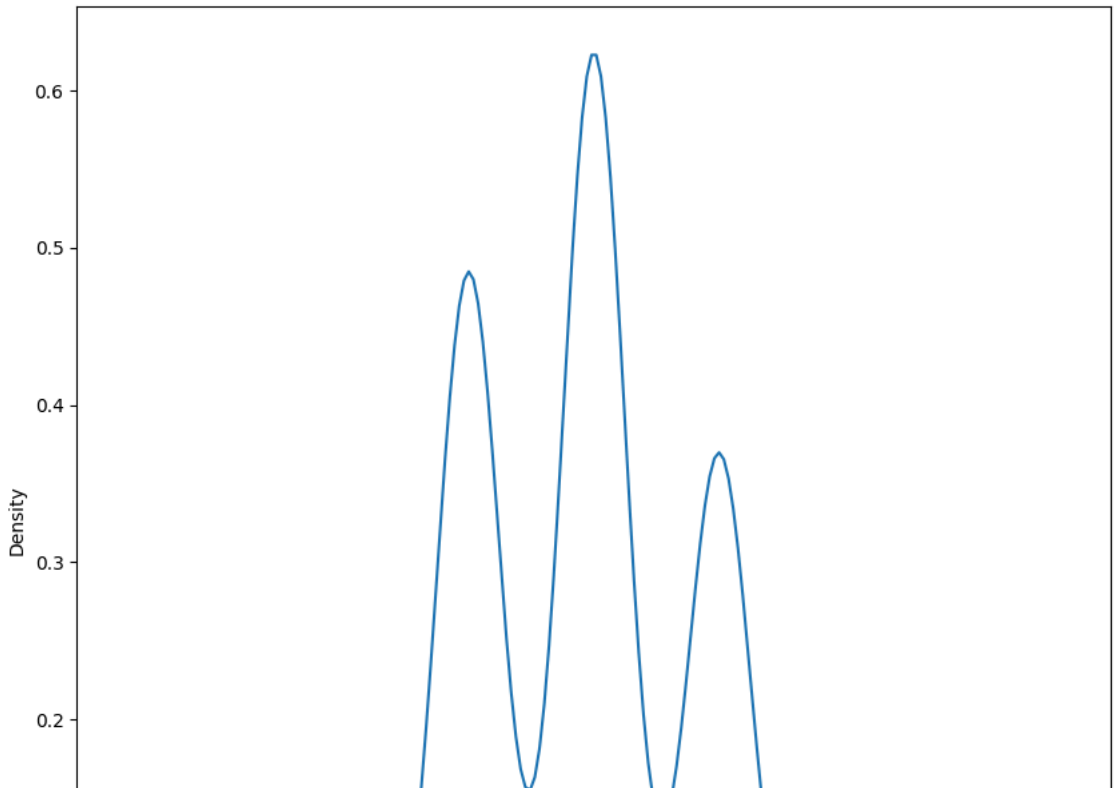


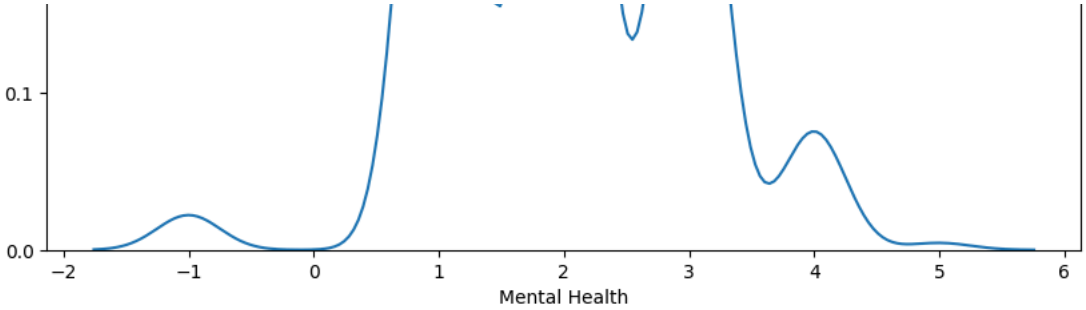

```
<ipython-input-15-8369a3293a54>:3: UserWarning: Dataset has 0 variance; skipping density estimate. Pass `warn_singular=False` to
  sns.kdeplot(data=df, x=i)
```

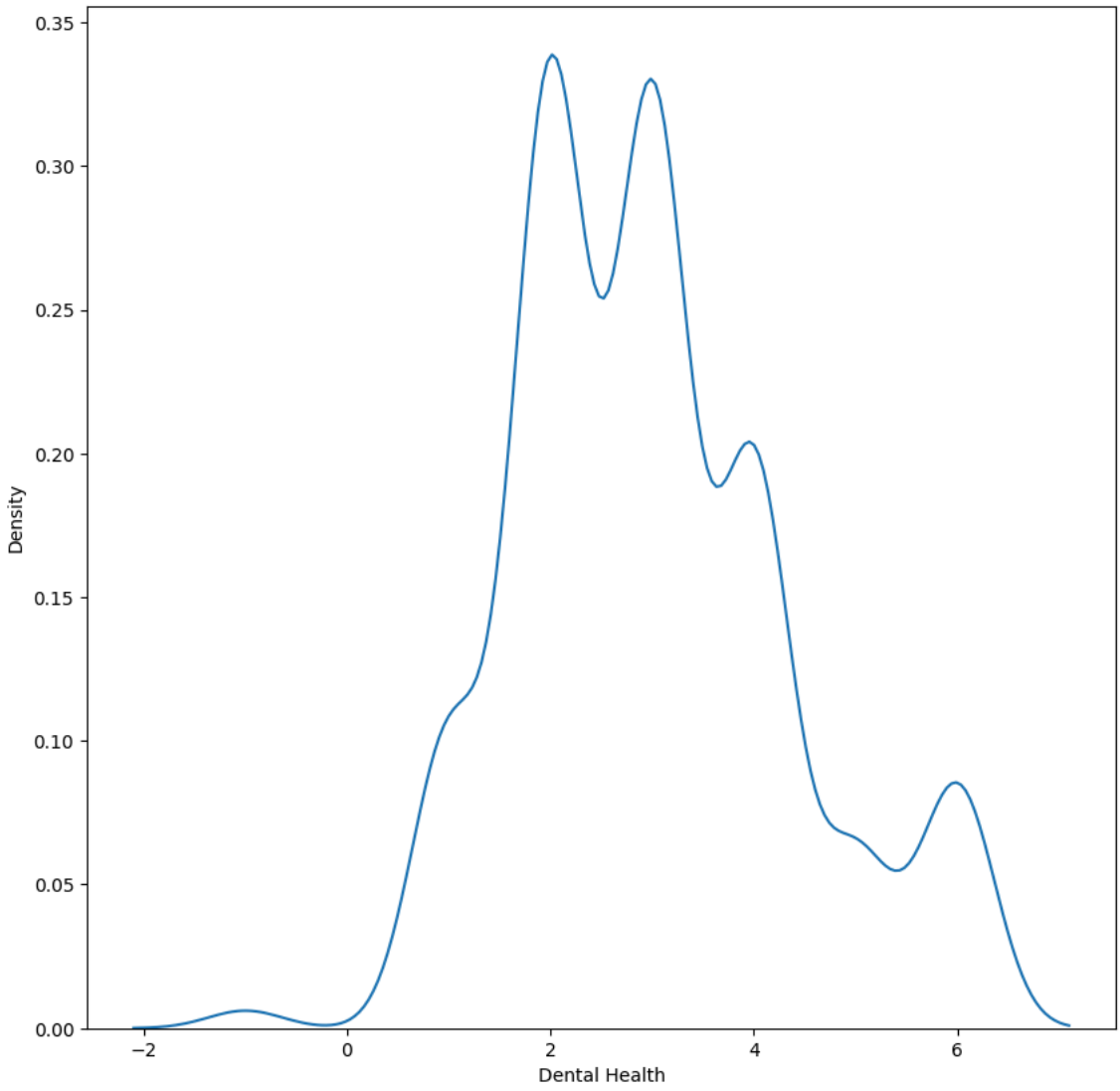


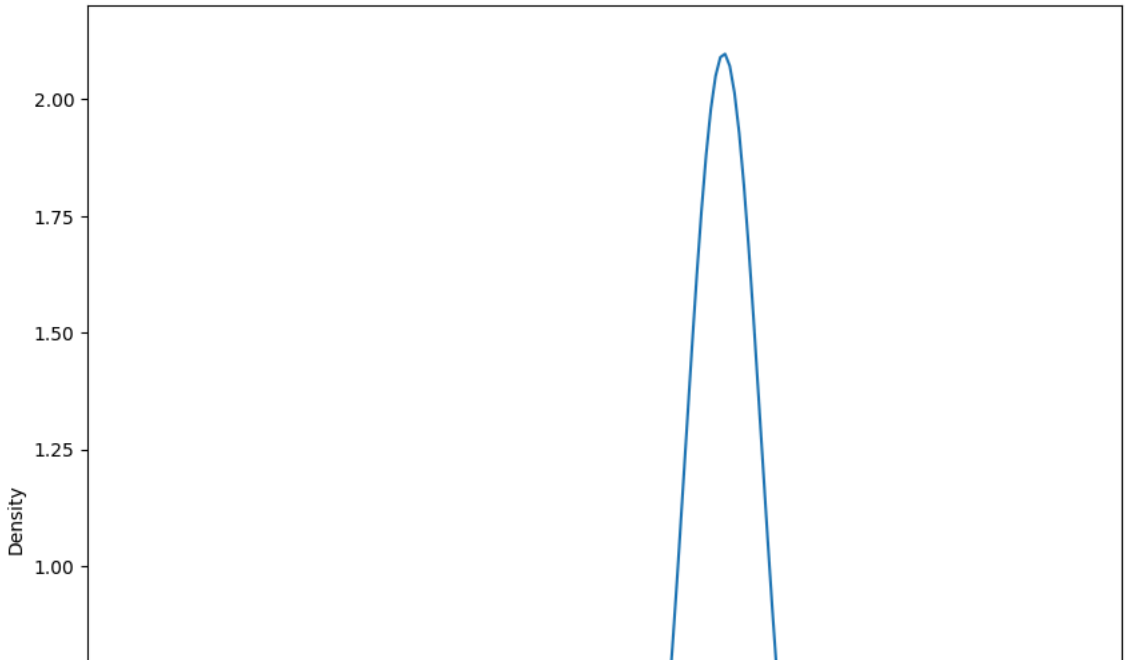

0.0
0.0
0.2
0.4
0.6
0.8
1.0
Age


0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Density
−2
−1
0
1
2
3
4
5
6
Phyiscal Health


0.6
0.5
0.4
0.3
0.2
Density


0.1
0.0
−2
−1
0
1
2
3
4
5
6
Mental Health


0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Density
−2
0
2
4
6
Dental Health


2.00
1.75
1.50
1.25
1.00
Density