In [1]:
```
# Import the necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
# Load the dataset
df = pd.read_csv('Key_indicator_statewise.csv')
```

In [3]:
```
# Basic data exploration
print(df.head())
```

```
       State_Name  AA_Sample_Units_Total  AA_Sample_Units_Rural  \
0           Assam                  1784                   1412
1           Bihar                  2356                   1981
2    Chhattisgarh                  1255                    926
3       Jharkhand                  2108                   1513
4  Madhya Pradesh                  2557                   1660

   AA_Sample_Units_Urban  AA_Households_Total  AA_Households_Rural  \
0                    372               388853               319766
1                    375               612684               568030
2                    329               287085               226554
3                    595               392734               318142
4                    897               519811               357179

   AA_Households_Urban  AA_Population_Total  AA_Population_Rural  \
0                69087              1809610              1518639
1                44654              3227867              2993906
2                60531              1264309               994416
3                74592              2019298              1644036
4               162632              2389787              1629355

   AA_Population_Urban  ...  \
0               290971  ...
1               233961  ...
2               269893  ...
3               375262  ...
4               760432  ...

   ZZ_Under_Five_Mortality_Rate_U5MR_Rural_Lower_Limit  \
0                                                 76
1                                                 71
2                                                 63
3                                                 56
4                                                 91

   ZZ_Under_Five_Mortality_Rate_U5MR_Rural_Upper_Limit  \
0                                                 79
1                                                 73
2                                                 67
3                                                 58
4                                                 94

   ZZ_Under_Five_Mortality_Rate_U5MR_Urban_Lower_Limit  \
0                                                 33
1                                                 48
2                                                 37
3                                                 27
4                                                 55

   ZZ_Under_Five_Mortality_Rate_U5MR_Urban_Upper_Limit  \
0                                                 40
1                                                 55
2                                                 43
3                                                 31
4                                                 59

   ZZ_Sex_Ratio_At_Birth_Total_Lower_Limit  \
```

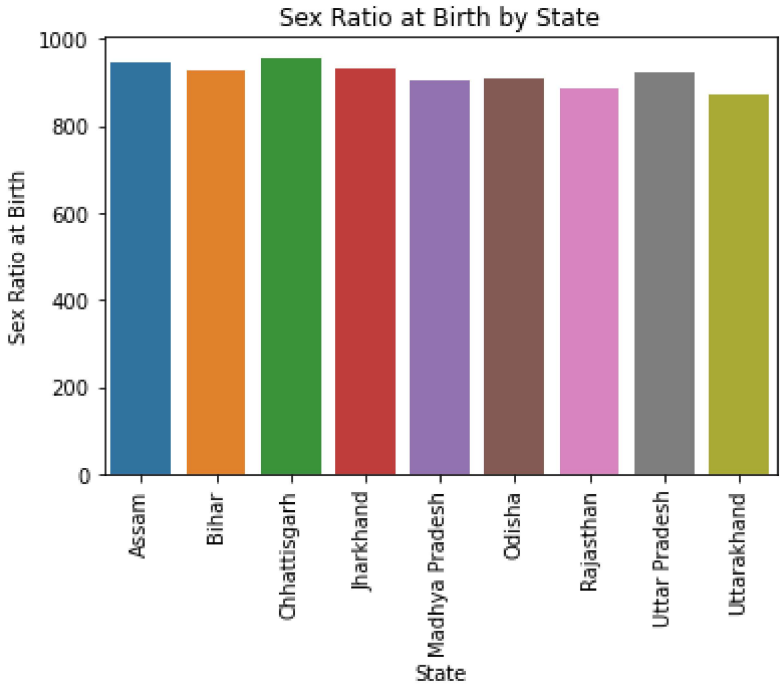
```
0                                        936
1                                        918
2                                        943
3                                        920
4                                        897

   ZZ_Sex_Ratio_At_Birth_Total_Upper_Limit  \
0                                        958
1                                        932
2                                        969
3                                        940
4                                        914

   ZZ_Sex_Ratio_At_Birth_Rural_Lower_Limit  \
0                                        930
1                                        918
2                                        944
3                                        932
4                                        906

   ZZ_Sex_Ratio_At_Birth_Rural_Upper_Limit  \
0                                        954
1                                        933
2                                        973
3                                        954
4                                        926

   ZZ_Sex_Ratio_At_Birth_Urban_Lower_Limit  \
0                                        945
1                                        893
2                                        916
3                                        860
4                                        860

   ZZ_Sex_Ratio_At_Birth_Urban_Upper_Limit
0                                     1009
1                                      953
2                                      975
3                                      909
4                                      893

[5 rows x 643 columns]
```

In [4]:
```
print(df.info())
```

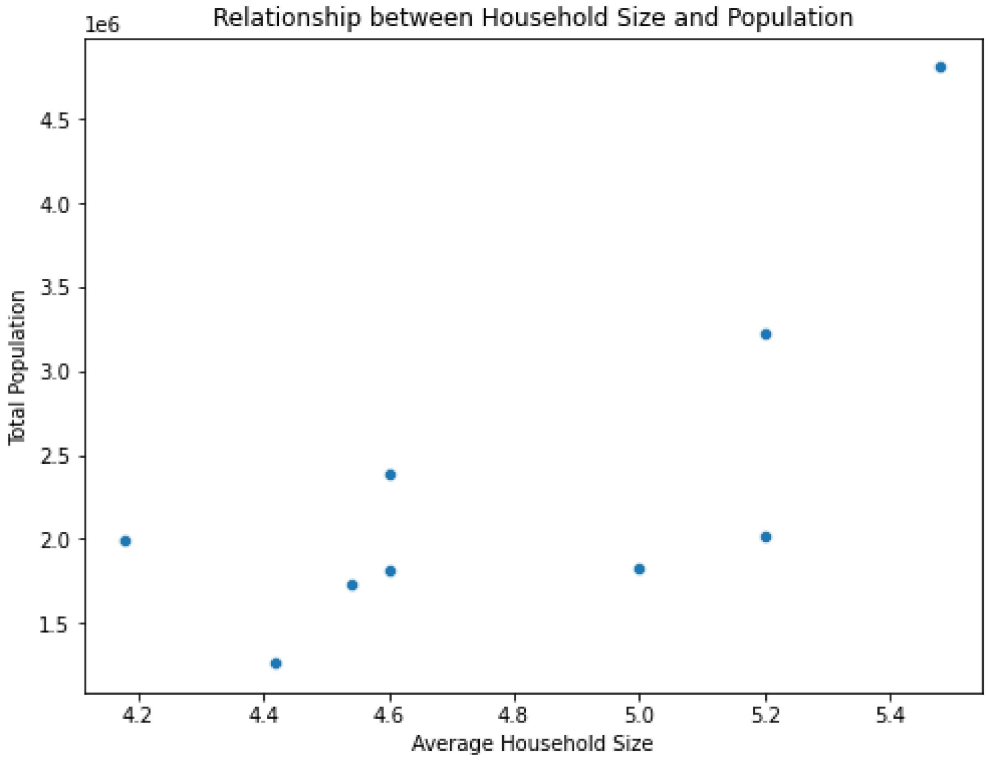
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9 entries, 0 to 8
Columns: 643 entries, State_Name to ZZ_Sex_Ratio_At_Birth_Urban_Upper_Limit
dtypes: float64(609), int64(33), object(1)
memory usage: 45.3+ KB
None
```

In [5]:
```
# Data cleaning (if necessary)
df.dropna(inplace=True)   #  Drop rows with missing values
```

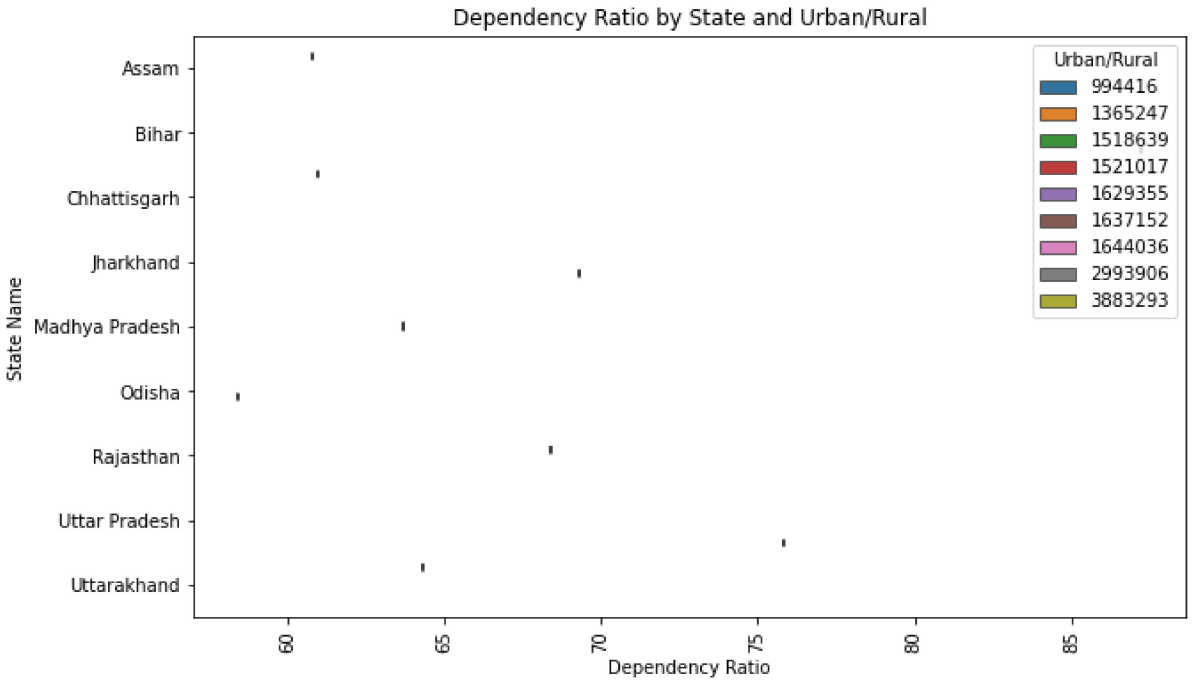
In [6]:
```
# Data visualization
# Chart 1
# Bar chart to visualize the Sex Ratio at Birth
sns.barplot(x='State_Name', y='CC_Sex_Ratio_At_Birth_Total', data=df)
plt.xticks(rotation=90)
plt.title('Sex Ratio at Birth by State')
plt.xlabel('State')
plt.ylabel('Sex Ratio at Birth')
plt.show()
```



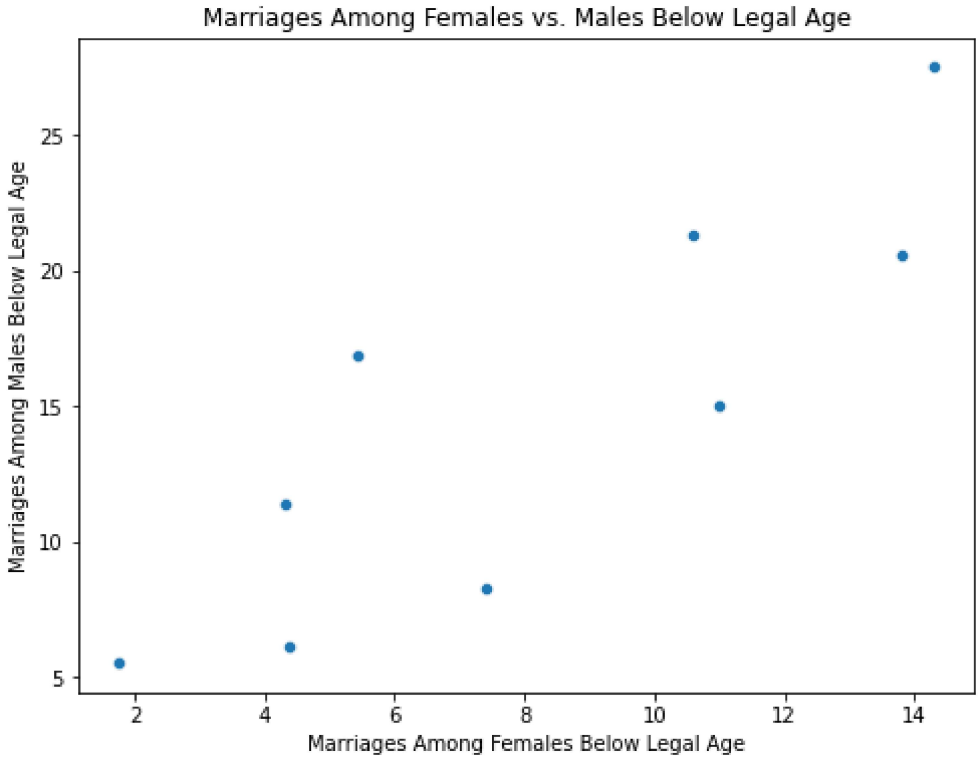
In [7]:
```
# Chart 2: Scatter plot to assess the relationship between average household si

plt.figure(figsize=(8, 6))
sns.scatterplot(x='BB_Average_Household_Size_All_Total', y='AA_Population_Total
plt.title('Relationship between Household Size and Population')
plt.xlabel('Average Household Size')
plt.ylabel('Total Population')
plt.show()
```



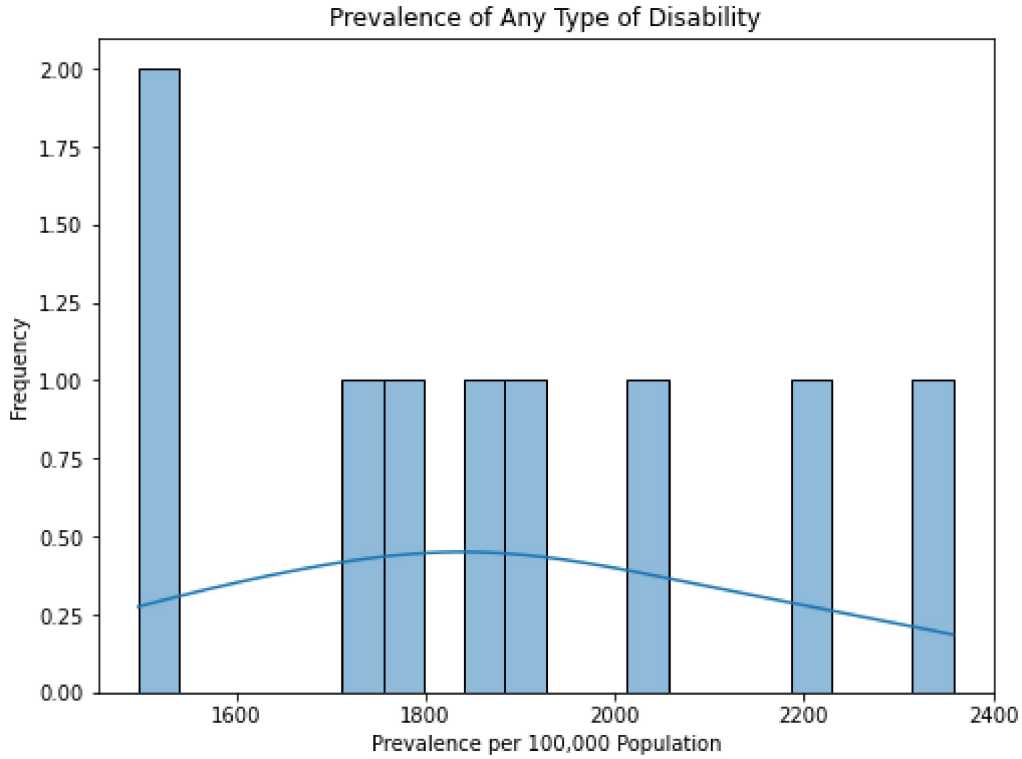
In [8]:
```
# Chart 3 Box Plot for Dependency Ratio by Rural and Urban Areas
plt.figure(figsize=(10, 6))
sns.boxplot(x='BB_Dependency_Ratio_Total', y='State_Name', data=df, hue='AA_Pop
plt.title('Dependency Ratio by State and Urban/Rural')
plt.xlabel('Dependency Ratio')
plt.ylabel('State Name')
plt.legend(title='Urban/Rural', loc='upper right')
plt.xticks(rotation=90)
plt.show()
```



In [9]:
```
# Chart 4 Scatter Plot for the Relationship between Marriages Among Females Bel
plt.figure(figsize=(8, 6))
sns.scatterplot(x='EE_Marriages_Among_Females_Below_Legal_Age_18_Years_Total',
                y='EE_Marriages_Among_Males_Below_Legal_Age_21_Years_Total',
                data=df)
plt.title('Marriages Among Females vs. Males Below Legal Age')
plt.xlabel('Marriages Among Females Below Legal Age')
plt.ylabel('Marriages Among Males Below Legal Age')
plt.show()
```



In [10]:
```
# Chart 5 Histogram for Prevalence of Any Type of Disability per 100,000 Popula
plt.figure(figsize=(8, 6))
sns.histplot(df['HH_Prevalence_Of_Any_Type_Of_Disability_Per_100000_Population_
plt.title('Prevalence of Any Type of Disability')
plt.xlabel('Prevalence per 100,000 Population')
plt.ylabel('Frequency')
plt.show()
```



In [11]:
```
# Statistical analysis

#  Correlation between Average Household Size and Population
correlation = df['BB_Average_Household_Size_All_Total'].corr(df['AA_Population_
print(f'Correlation between Household Size and Population: {correlation}')
```

```
Correlation between Household Size and Population: 0.7230712033713329
```

In [13]:
```
#  T-Test for Sex Ratio at Birth in Rural and Urban Areas

from scipy.stats import ttest_ind

rural_sex_ratio = df['CC_Sex_Ratio_At_Birth_Rural']
urban_sex_ratio = df['CC_Sex_Ratio_At_Birth_Urban']

t_stat, p_value = ttest_ind(rural_sex_ratio, urban_sex_ratio)
print(f'T-Test: Sex Ratio at Birth (Rural vs. Urban)')
print(f'T-statistic: {t_stat}')
print(f'P-value: {p_value}')

if p_value < 0.05:
    print('There is a significant difference in sex ratio at birth between rura
else:
    print('There is no significant difference in sex ratio at birth between rur
```

```
T-Test: Sex Ratio at Birth (Rural vs. Urban)
T-statistic: 0.9385782134434763
P-value: 0.3619010363816274
There is no significant difference in sex ratio at birth between rural and ur
ban areas.
```

In [14]:
```
#  Chi-Square Test for Marriages Among Females Below Legal Age

from scipy.stats import chi2_contingency

contingency_table = pd.crosstab(df['EE_Marriages_Among_Females_Below_Legal_Age_

chi2, p, dof, expected = chi2_contingency(contingency_table)
print(f'Chi-Square Test: Marriages Among Females Below Legal Age (Rural vs. Urb
print(f'Chi-Square Statistic: {chi2}')
print(f'P-value: {p}')

if p < 0.05:
    print('There is a significant association between rural and urban marriages
else:
    print('There is no significant association between rural and urban marriage
```

```
Chi-Square Test: Marriages Among Females Below Legal Age (Rural vs. Urban)
Chi-Square Statistic: 71.99999999999999
P-value: 0.23025670024966835
There is no significant association between rural and urban marriages among f
emales below legal age.
```

In [ ]: