cdc

September 20, 2023

[16]:
```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
```

[20]:
```
# Fetch the data
url = "./cdc_places_ny.csv"
df = pd.read_csv(url)
df.columns
```

[20]:
```
Index(['Year', 'StateAbbr', 'StateDesc', 'LocationName', 'DataSource',
       'Category', 'Measure', 'Data_Value_Unit', 'Data_Value_Type',
       'Data_Value', 'Data_Value_Footnote_Symbol', 'Data_Value_Footnote',
       'Low_Confidence_Limit', 'High_Confidence_Limit', 'TotalPopulation',
       'LocationID', 'CategoryID', 'MeasureId', 'DataValueTypeID',
       'Short_Question_Text', 'Geolocation'],
      dtype='object')
```

[21]:
```
# Filter for 'BINGE' as measureid and 'Age-adjusted prevalence' as␣
↪data_value_type
df = df[(df['MeasureId'] == 'BINGE') & (df['Data_Value_Type'] == 'Age-adjusted␣
↪prevalence')]
df
```

[21]:

	Year	StateAbbr	StateDesc	LocationName	DataSource	\
335	2021	NY	New York	Orange	BRFSS	
437	2021	NY	New York	Steuben	BRFSS	
490	2021	NY	New York	Oswego	BRFSS	
499	2021	NY	New York	St. Lawrence	BRFSS	
541	2021	NY	New York	Orleans	BRFSS	
…	…	…	…	…	…	
4286	2021	NY	New York	Cortland	BRFSS	
4301	2021	NY	New York	Cattaraugus	BRFSS	
4356	2021	NY	New York	Lewis	BRFSS	
4431	2021	NY	New York	Greene	BRFSS	
4502	2021	NY	New York	Hamilton	BRFSS	

	Category	Measure	\

```
335   Health Risk Behaviors  Binge drinking among adults aged >=18 years
437   Health Risk Behaviors  Binge drinking among adults aged >=18 years
490   Health Risk Behaviors  Binge drinking among adults aged >=18 years
499   Health Risk Behaviors  Binge drinking among adults aged >=18 years
541   Health Risk Behaviors  Binge drinking among adults aged >=18 years
…                     …                                          …
4286  Health Risk Behaviors  Binge drinking among adults aged >=18 years
4301  Health Risk Behaviors  Binge drinking among adults aged >=18 years
4356  Health Risk Behaviors  Binge drinking among adults aged >=18 years
4431  Health Risk Behaviors  Binge drinking among adults aged >=18 years
4502  Health Risk Behaviors  Binge drinking among adults aged >=18 years

     Data_Value_Unit          Data_Value_Type  Data_Value  …  \
335                %  Age-adjusted prevalence        17.1  …
437                %  Age-adjusted prevalence        20.0  …
490                %  Age-adjusted prevalence        19.7  …
499                %  Age-adjusted prevalence        20.3  …
541                %  Age-adjusted prevalence        18.1  …
…              …                        …         …  …
4286               %  Age-adjusted prevalence        18.5  …
4301               %  Age-adjusted prevalence        19.4  …
4356               %  Age-adjusted prevalence        19.1  …
4431               %  Age-adjusted prevalence        21.0  …
4502               %  Age-adjusted prevalence        21.5  …

      Data_Value_Footnote  Low_Confidence_Limit  High_Confidence_Limit  \
335                   NaN                  14.6                   19.9
437                   NaN                  17.2                   23.0
490                   NaN                  17.1                   22.7
499                   NaN                  17.5                   23.3
541                   NaN                  15.4                   21.1
…                     …                    …                      …
4286                  NaN                  15.6                   21.5
4301                  NaN                  16.5                   22.4
4356                  NaN                  16.3                   22.2
4431                  NaN                  17.9                   24.3
4502                  NaN                  18.2                   25.1

      TotalPopulation  LocationID  CategoryID MeasureId DataValueTypeID  \
335            404525       36071     RISKBEH     BINGE       AgeAdjPrv
437             92948       36101     RISKBEH     BINGE       AgeAdjPrv
490            117387       36075     RISKBEH     BINGE       AgeAdjPrv
499            108051       36089     RISKBEH     BINGE       AgeAdjPrv
541             40191       36073     RISKBEH     BINGE       AgeAdjPrv
…                 …          …         …          …              …
4286            46311       36023     RISKBEH     BINGE       AgeAdjPrv
4301            76426       36009     RISKBEH     BINGE       AgeAdjPrv
```

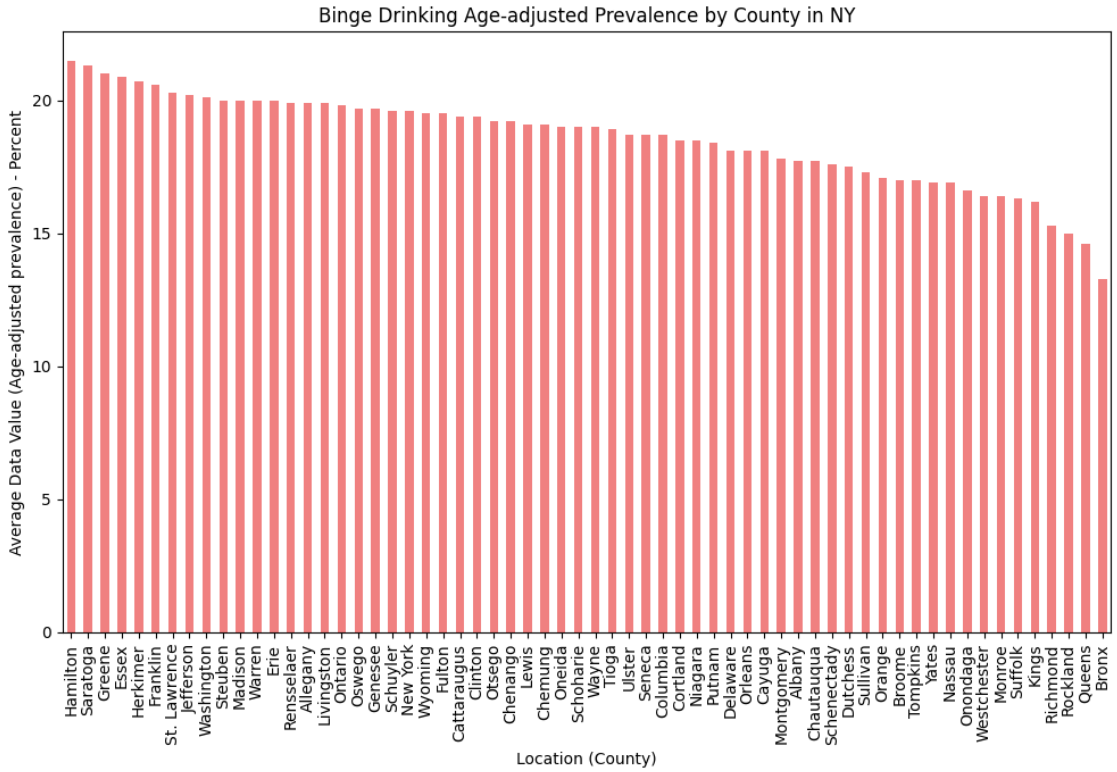
```
4356            26573       36049     RISKBEH     BINGE       AgeAdjPrv
4431            48499       36039     RISKBEH     BINGE       AgeAdjPrv
4502             5119       36041     RISKBEH     BINGE       AgeAdjPrv

     Short_Question_Text                      Geolocation
335       Binge Drinking  POINT (-74.3062522 41.4024096)
437       Binge Drinking  POINT (-77.3855253 42.2667252)
490       Binge Drinking  POINT (-76.2092618 43.4614431)
499       Binge Drinking   POINT (-75.074311 44.4881125)
541       Binge Drinking  POINT (-78.2070281 43.3399059)
…                    …                              …
4286      Binge Drinking   POINT (-76.0762398 42.594039)
4301      Binge Drinking   POINT (-78.662332 42.2390986)
4356      Binge Drinking  POINT (-75.4441402 43.7826811)
4431      Binge Drinking  POINT (-74.1420253 42.2798211)
4502      Binge Drinking  POINT (-74.5024556 43.6578786)

[62 rows x 21 columns]
```

[24]:
```
# Group by 'LocationName' and get the average (or sum) 'Data_Value'
grouped = df.groupby('LocationName').Data_Value.mean().
 ↪sort_values(ascending=False)
grouped
```

[24]:
```
LocationName
Hamilton    21.5
Saratoga    21.3
Greene      21.0
Essex       20.9
Herkimer    20.7
            …
Kings       16.2
Richmond    15.3
Rockland    15.0
Queens      14.6
Bronx       13.3
Name: Data_Value, Length: 62, dtype: float64
```

[27]:
```
# Plotting
plt.figure(figsize=(10, 7))
grouped.plot(kind='bar', color='lightcoral')
plt.ylabel('Average Data Value (Age-adjusted prevalence) - Percent')
plt.xlabel('Location (County)')
plt.title('Binge Drinking Age-adjusted Prevalence by County in NY')
plt.xticks(rotation=90)
plt.tight_layout()
plt.savefig("binge_drinking_per_location.png")  # Saving the plot as an image
```

```
plt.show()
```


Binge Drinking Age-adjusted Prevalence by County in NY
Average Data Value (Age-adjusted prevalence) - Percent
0
5
10
15
20
Location (County)
Hamilton
Saratoga
Greene
Essex
Herkimer
Franklin
St. Lawrence
Jefferson
Washington
Steuben
Madison
Warren
Erie
Rensselaer
Allegany
Livingston
Ontario
Oswego
Genesee
Schuyler
New York
Wyoming
Fulton
Cattaraugus
Clinton
Otsego
Chenango
Lewis
Chemung
Oneida
Schoharie
Wayne
Tioga
Ulster
Seneca
Columbia
Cortland
Niagara
Putnam
Delaware
Orleans
Cayuga
Montgomery
Albany
Chautauqua
Schenectady
Dutchess
Sullivan
Orange
Broome
Tompkins
Yates
Nassau
Onondaga
Westchester
Monroe
Suffolk
Kings
Richmond
Rockland
Queens
Bronx