

DEFATING

COVID - 19

With
Data!

Featuring: John Loucks, Ryan Muir, and Zach Adams

Covid and the Sleepy Upper Midwest

Local concerns and outcomes
during the global pandemic:

- Where are the hotspots of Covid cases?

We expect cities to have higher incidence rates

- Are incidence rates affected by lower income counties?

We expect higher income counties to have less cases and deaths than average

- How does mask usage affect covid rates?

We expect areas with high mask usage will have lower contraction rates than average

We tested our hypothesis over 5 Midwest states:

Missouri, Illinois, Indiana, Ohio, and Michigan



SOURCES:

- CENSUS - ACS5
 - [HTTPS://API.CENSUS.GOV/DATA/2018/ACS/ACS5.HTML](https://api.census.gov/data/2018/acs/acs5.html)
- JOHNS HOPKINS COVID-19 CSV
 - [HTTPS://GITHUB.COM/CSSEGISANDDATA/COVID-19](https://github.com/CSSEGISANDDATA/COVID-19)
- NYT MASK SURVEY CSV
 - [HTTPS://GITHUB.COM/NYTIMES/COVID-19-DATA](https://github.com/nytimes/covid-19-data)

Data Cleanup & Exploration

insights, difficulties, & resolutions

#1

Perform API call and loaded CSV files

```
# Census API Key
from config import census_key
c = Census(census_key, year=2018)
```

#2

Decided to merge data from all sources to single dataframe on FIPS county code and removed all null values

```
mo_census_data = c.acs5.get(("NAME", "B01003_001E", "B19301_001E"),{'for': 'county:', 'in': 'state:29'})
mo_census_pd = pd.DataFrame(mo_census_data)

# Column Reordering
mo_census_pd = mo_census_pd.rename(columns={"B01003_001E": "Total Population",
                                             "B19301_001E": "Per Capita Income",
                                             "NAME": "County, State"})

mo_census_pd['FIPS']=mo_census_pd['state']+mo_census_pd['county']
mo_census_pd=mo_census_pd.drop(columns=['state', 'county'])
mo_census_pd=mo_census_pd.merge(covid_91120_data,how='left',on='FIPS').drop(
    columns={'County, State','Combined_Key'}).rename(
    columns={'Admin2':'County','Province_State':'State'})
```

#3

Had to change FIPS codes to string for merging in CSV data

```
#Import covid data
covid_91120_data=pd.read_csv('CSV_Data/09-11-2020.csv')
covid_91120_data=covid_91120_data.loc[covid_91120_data['FIPS'].isnull()==False]
covid_91120_data['FIPS']=covid_91120_data['FIPS'].astype(int).astype(str)
covid_91120_data.head()
```

#4

Because our data is geographical, we decided to use plotly express to construct our figures, including choropleths!

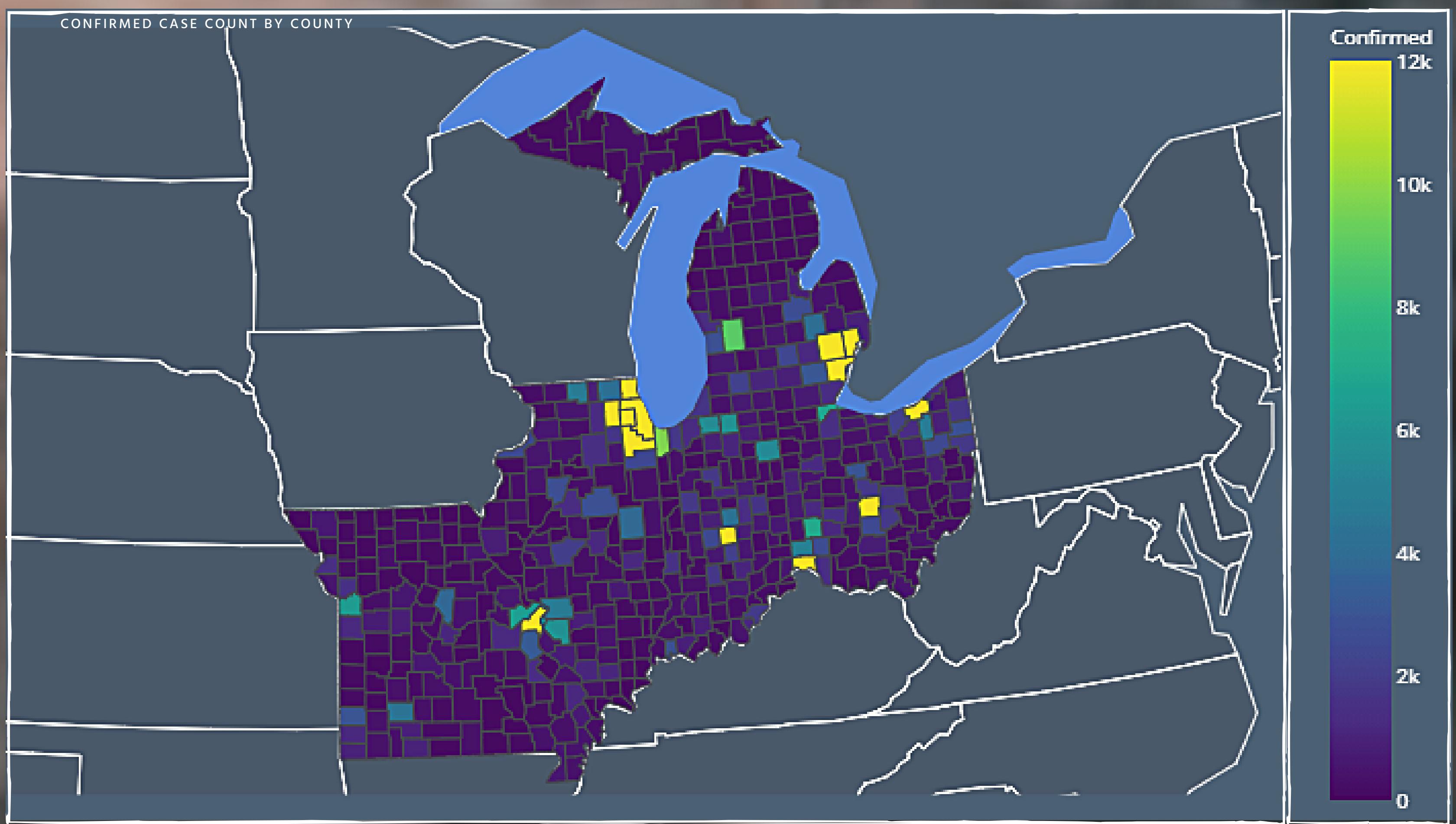
#5

Removed Un-needed columns from our Dataframe

Cleaned DataFrame

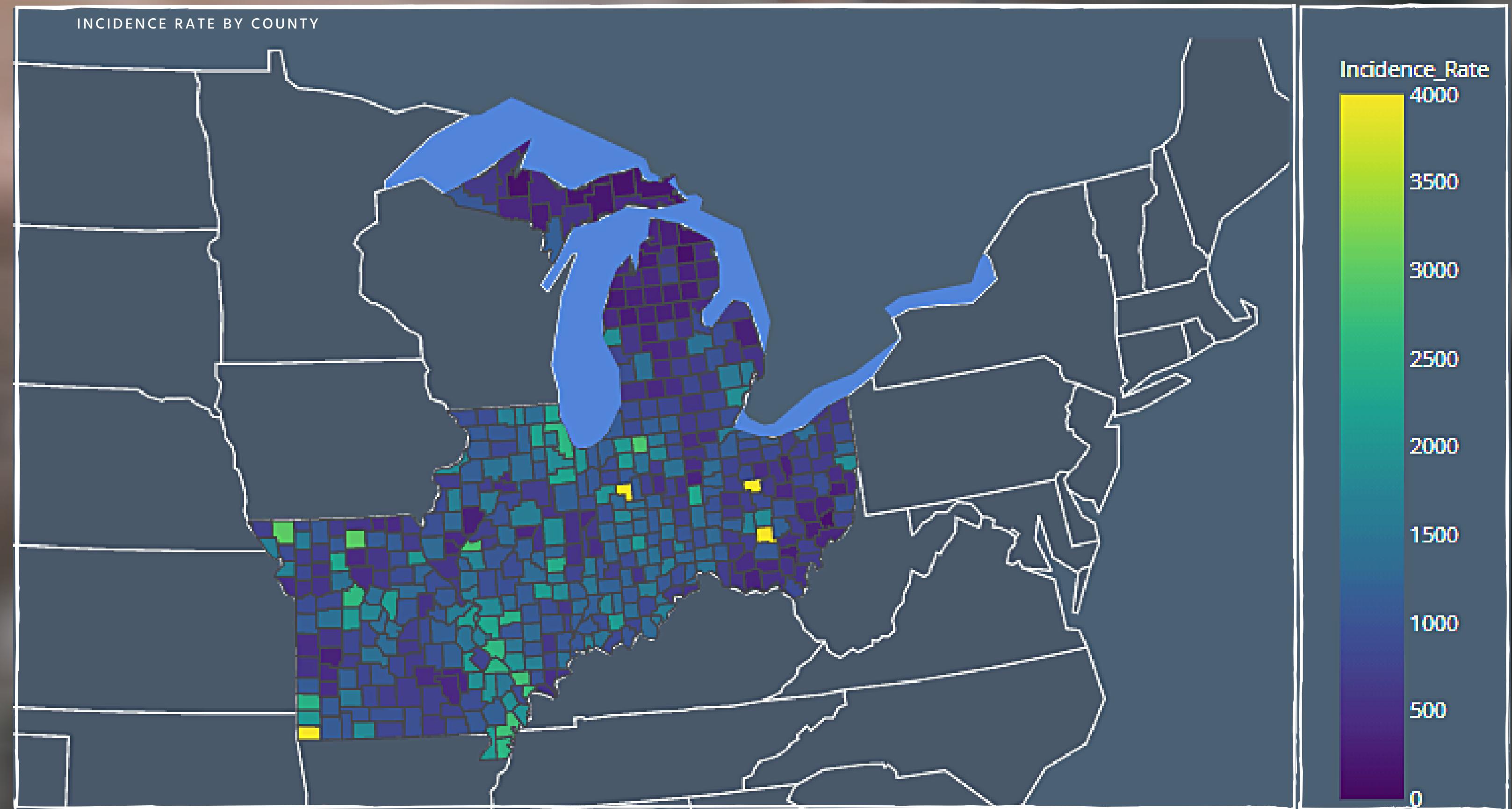
	Total Population	Per Capita Income	FIPS	County	State	Confirmed	Incidence_Rate	Case-Fatality_Ratio	NEVER	RARELY	SOMETIMES	FREQUENTLY	ALWAYS
0	23102.0	22802.0	29195	Saline	Missouri	599	2631.694565	1.502504	0.083	0.230	0.112	0.186	0.389
1	12205.0	22119.0	29123	Madison	Missouri	229	1894.440768	0.000000	0.082	0.124	0.092	0.296	0.406
2	18293.0	19909.0	29229	Wright	Missouri	127	694.406474	0.000000	0.419	0.139	0.187	0.100	0.155
3	20691.0	24388.0	29217	Vernon	Missouri	121	588.435540	0.000000	0.164	0.105	0.110	0.125	0.496
4	22547.0	22151.0	29147	Nodaway	Missouri	663	3001.086366	1.055807	0.093	0.193	0.160	0.283	0.272
...
475	19994.0	28430.0	18015	Carroll	Indiana	262	1293.380066	4.961832	0.059	0.104	0.170	0.220	0.447
476	36378.0	25386.0	18069	Huntington	Indiana	194	531.215772	1.546392	0.150	0.182	0.053	0.199	0.416
477	24217.0	27621.0	18181	White	Indiana	465	1929.300473	2.795699	0.056	0.177	0.194	0.189	0.384
478	20993.0	22470.0	18075	Jay	Indiana	149	729.105500	0.000000	0.141	0.139	0.157	0.235	0.328
479	115702.0	28525.0	18019	Clark	Indiana	1947	1645.787899	2.722137	0.068	0.060	0.141	0.228	0.501

Population Density And Covid-19



CONFIRMED CASE COUNT BY COUNTY

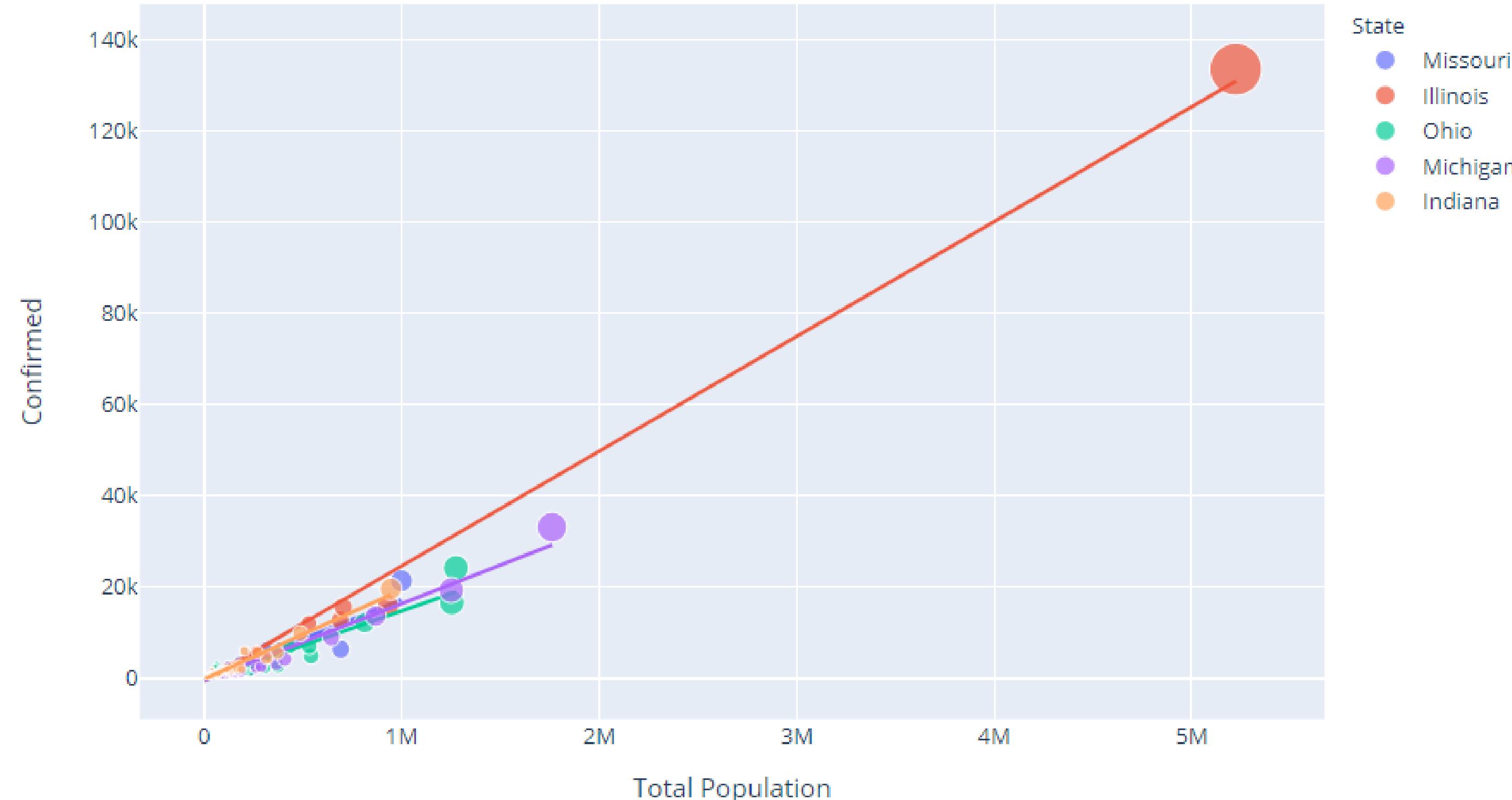
MIDWEST COVID-19



INCIDENCE RATE BY COUNTY

MIDWEST COVID-19

Total Population and Confirmed Cases by State

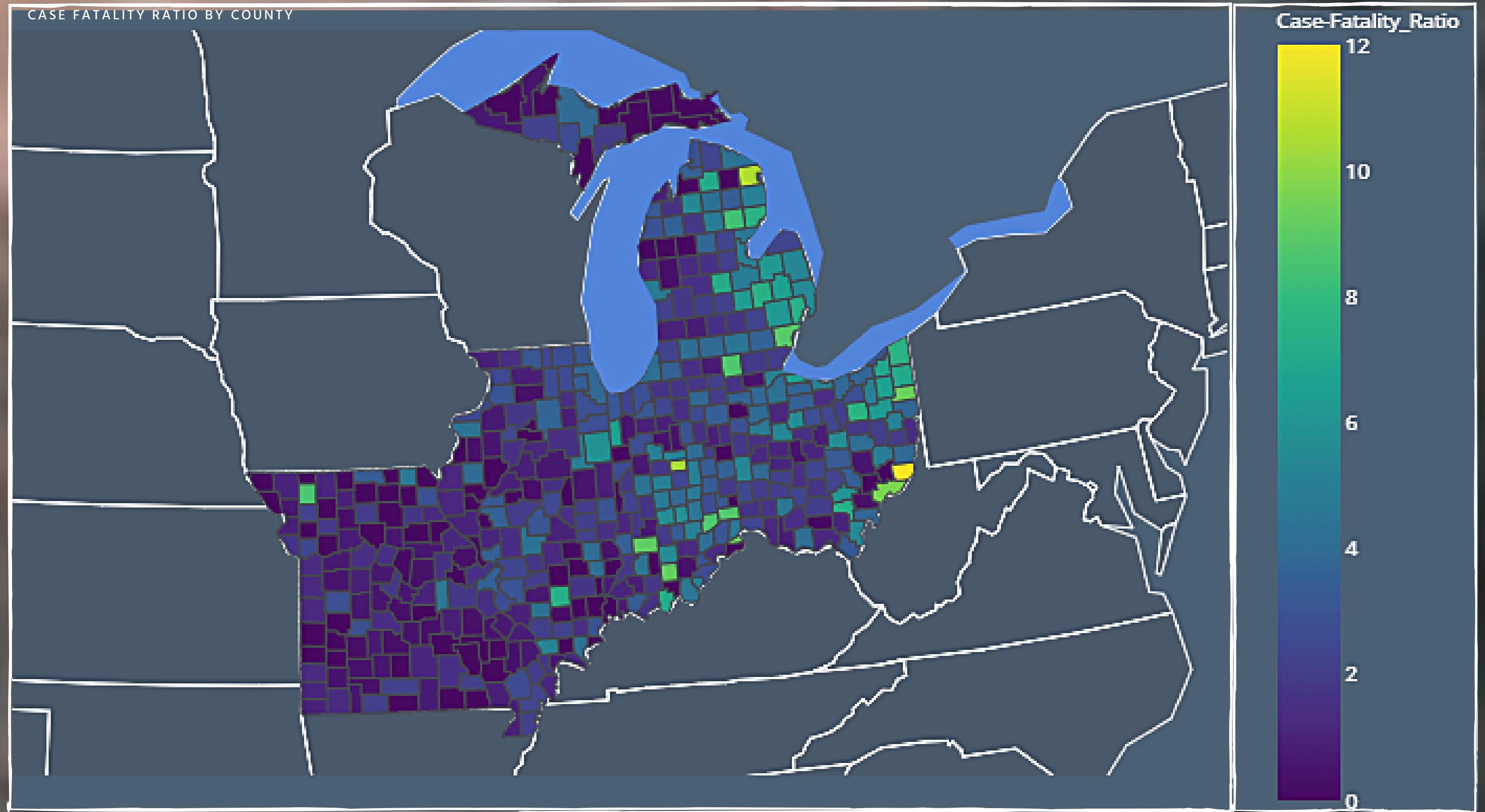


CONFIRMED CASES / TOTAL POPULATION BY STATE

MIDWEST COVID-19

Income, Deaths, And Covid-19

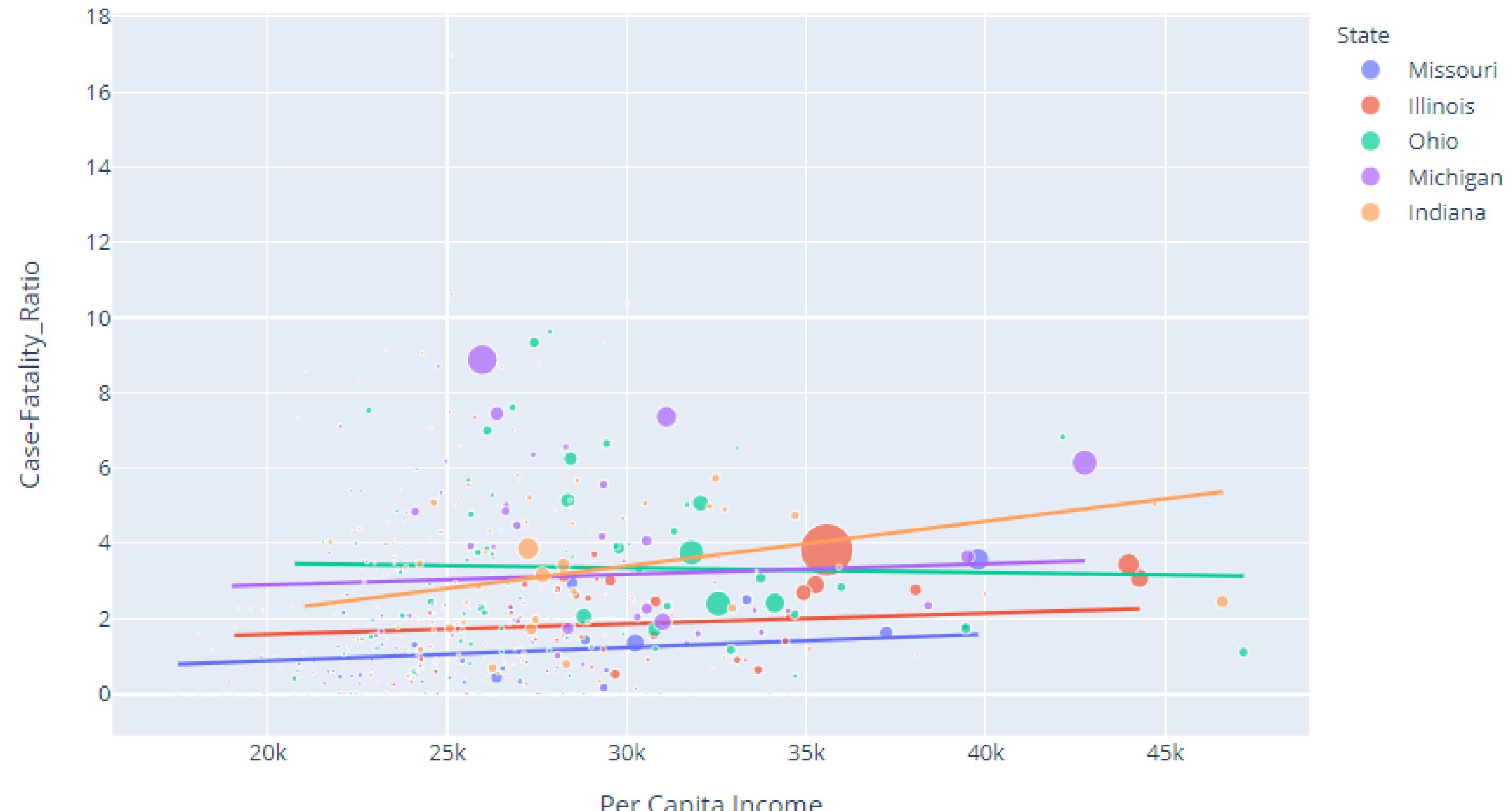
CASE FATALITY RATIO BY COUNTY



CASE FATALITY RATIO BY COUNTY

MIDWEST COVID-19

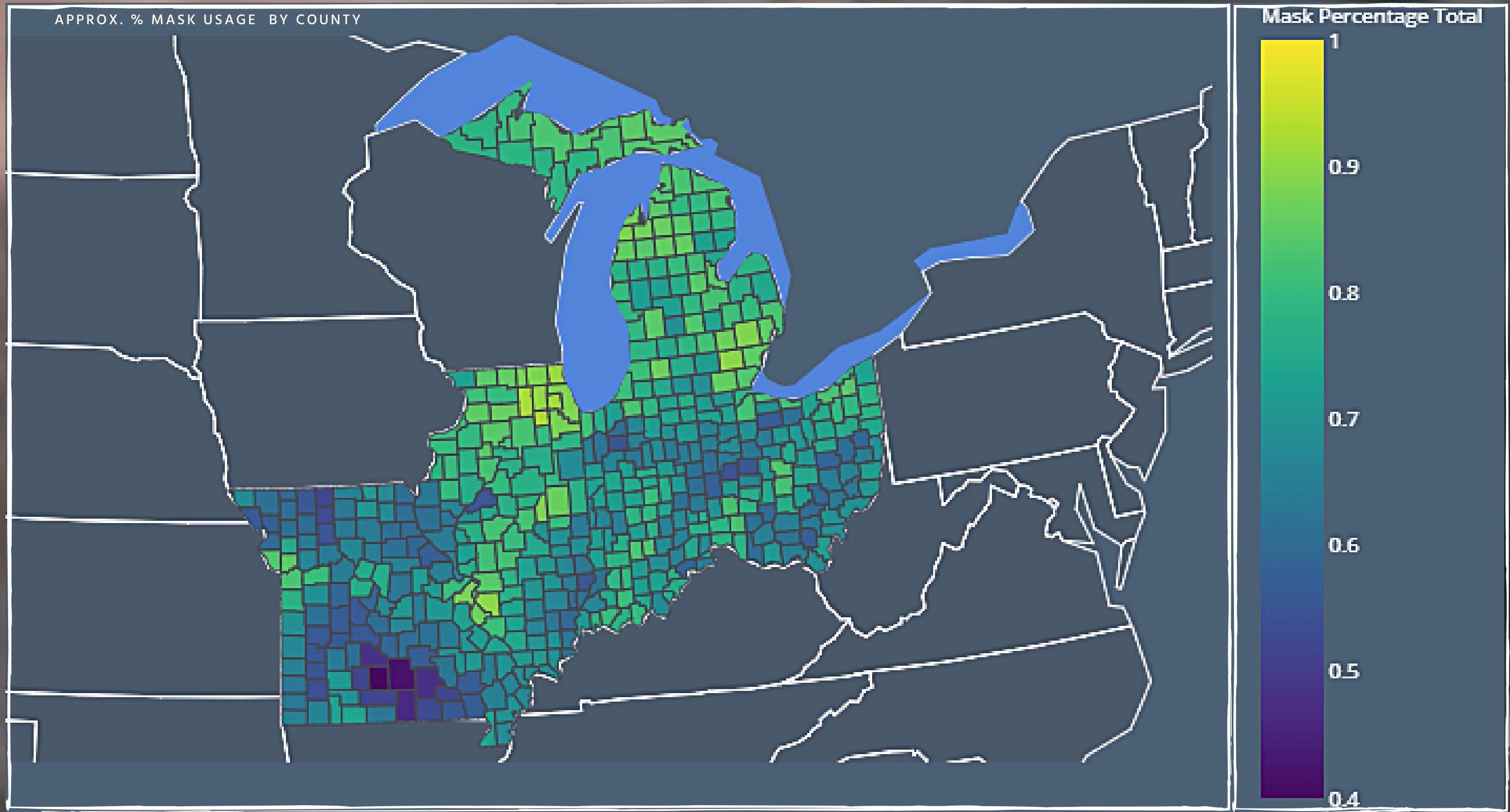
Per Capita Income and Case-Fatality Ratio by State



Self-Reported Mask Usage and Covid-19

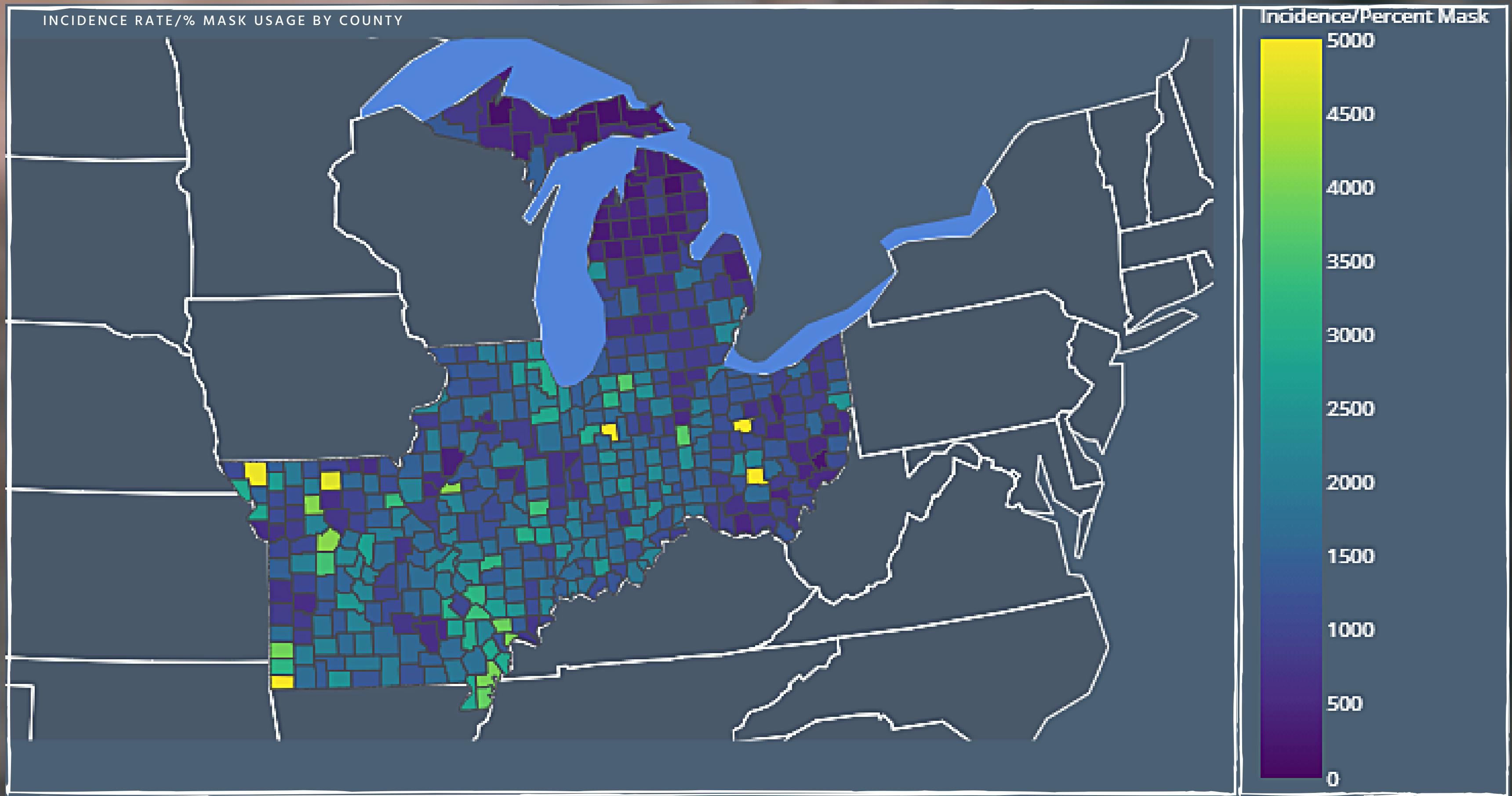
Cleaned DataFrame

Total Population	Per Capita Income	FIPS	County	State	Confirmed	Incidence_Rate	Case-Fatality_Ratio	NEVER	RARELY	SOMETIMES	FREQUENTLY	ALWAYS	Mask Percentage Total
23102.0	22802.0	29195	Saline	Missouri	599	2631.694565	1.502504	0.083	0.230	0.112	0.186	0.389	0.64200
12205.0	22119.0	29123	Madison	Missouri	229	1894.440768	0.000000	0.082	0.124	0.092	0.296	0.406	0.70500
18293.0	19909.0	29229	Wright	Missouri	127	694.406474	0.000000	0.419	0.139	0.187	0.100	0.155	0.35825
20691.0	24388.0	29217	Vernon	Missouri	121	588.435540	0.000000	0.164	0.105	0.110	0.125	0.496	0.67100
22547.0	22151.0	29147	Nodaway	Missouri	663	3001.086366	1.055807	0.093	0.193	0.160	0.283	0.272	0.61250
...
19994.0	28430.0	18015	Carroll	Indiana	262	1293.380066	4.961832	0.059	0.104	0.170	0.220	0.447	0.72300
36378.0	25386.0	18069	Huntington	Indiana	194	531.215772	1.546392	0.150	0.182	0.053	0.199	0.416	0.63725
24217.0	27621.0	18181	White	Indiana	465	1929.300473	2.795699	0.056	0.177	0.194	0.189	0.384	0.66700
20993.0	22470.0	18075	Jay	Indiana	149	729.105500	0.000000	0.141	0.139	0.157	0.235	0.328	0.61750
115702.0	28525.0	18019	Clark	Indiana	1947	1645.787899	2.722137	0.068	0.060	0.141	0.228	0.501	0.75750



APPROXIMATE PERCENTAGE OF MASK USE BY COUNTY

MIDWEST COVID-19



INCIDENCE RATE TO PERCENTAGE OF MASK USAGE BY COUNTY

MIDWEST COVID-19

Mask Usage Percentage and Incidence Rate



MASK USAGE / INCIDENCE RATE

MIDWEST COVID-19

Covid-19 in the Midwest

- While midwest cities had higher numbers, incidence rate was relatively low or close to average
- Income had no effect on incidence rate or Covid-related deaths
- Average mask usage seemed to have no specific effect on data
It is possible that high mask usage is a response to high incidence rates, particularly in cities

No Particular state seemed to differ significantly from other observed states

Future Concepts

Future Questions could address time series relationships between our tested hypotheses, or how our measures compared with Covid infections since the first case.

We could additionally compare this data to other regions or states in the US, or to more populous regions.

Questions?

Thank you for your time!

Additional Code Examples

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import requests
from census import Census
import plotly.express as px
```

```
fig = px.choropleth(states_tot, geojson=counties_geo, locations='FIPS', color='Confirmed',
                     color_continuous_scale="Viridis",
                     range_color=(0, 12000),
                     scope = 'usa',
                     )
fig.update_geos(fitbounds="locations", visible=True)
fig.update_layout(width=800, height=400, margin={"r":0,"t":0,"l":0,"b":0})
fig.update_layout(geo=dict(bgcolor= 'rgba(0,0,0,0)',lakecolor = '#5583D7',landcolor='rgba(51,17,0,0.1)'),
                  font = {"size": 10, "color":"White"},paper_bgcolor="#4E5D6C",plot_bgcolor="#4E5D6C")
fig.show()
```

```
val=[0,.25,.5,.75,1]

mask_df['Mask Percentage Total']=mask_df['NEVER']*val[0]+mask_df['RARELY']*val[1]+mask_df['SOMETIMES']*val[2]+
                                mask_df['FREQUENTLY']*val[3]+mask_df['ALWAYS']*val[4]
mask_df
```