Economic Rural Jobs Employment Data

Team:

Paris Lee
Joy Weishan
Laura Zapata
Sam Hoemann

Paris

Paris is a newly minted Medical Laboratory Scientist from Cleaveland, Ohio She is studying data analystics with the hopes of incorporating it into her future career as a physician, with the larger overarching goal of being a point of understandable scientific and medical infomration. She aims to support her community and loved ones, an ultimately help people to be informed and lead healthy and fulfilling lives.

Joy

Joy is an experienced IT professional from rural Wisconsin. She is studying data well as full stack development to accomodate data visualization presentation. H incorporate better understanding of Big Data and full stack development to add numerous other IT skills. She aims to support data analytics using an ethical ap fully implementing applications with regulatory compliance, data quality, effici retrieval and to avoid data bias when possible.

Laura

Laura is a Legal Assistant from Houston, Texas, who is passionate about studying data analytics. She hopes to integrate this knowledge into her legal career, with the larger goal of becoming a source for clear and accessible information.

Sam

Sam is a an aspiring science fiction and slice of life author. Wanting to learn a new trade in the form of Data Analytics to diversify their life and skill set. They are from St.louis, Missouri. Go Tigers.

Introduction:

This project is centered on developing a full-stack web application aimed at visualizing significant trends in job and income dynamics across rural America. The core objective was to create an interactive dashboard that offers a view of these economic trends through a variety of data visualizations, including bar charts, sunburst charts, bubble charts, and a multi-layered interactive map.

The backend of the project involved thorough data cleaning, data normalization and database creation with data imports with attempts at using flask. The frontend included analysis conducted in Jupyter Notebook to prepare the dataset for meaningful insights to include the creation of the visualizations that would be displayed on the dashboard.

The final application, built using HTML, JavaScript an interface that allows participants to explore the visualizations, interact with various controls, and gain insights into the economic factors affecting rural America. By integrating these components, our project not only highlights the importance of data in understanding rural economic employment trends while demonstrating the effective use of modern web technologies in data presentation.

Analysis

From our initial analysis it's easy to see that rural population centers within the continental United States are beginning to show a population decline tied to a decrease in economic opportunity. While this development is not unique to specifically the United States nor its rural communities, it is the focus of our project. As a side tangent, we can see this same shift in demographics within New York city and its boroughs.

Our primary portion of the analysis focused on the macro scale, namely demographic changes of states rather than counties. Counties being something that could be further explored to show a more detailed picture of every shifting economics of the Rural American experience and life.

Color Palette

We used a pastel color pallet for consistency across the graphs.

Dataset Cleaning

Our initial dataset included jobs, employment and unemployment in a flattened CSV. In order to do thorough analysis this dataset was split into three new datasets. One with the jobs data, the second with the employment data and the third with the unemployment data. From there, we initialized null data and renamed columns with appropriate headers to match the database created.

We chose a normalized database with three tables; jobs, employment and unemployment that were able to be joined on the common data item fips. We created a primary key for employment and unemployment as there was now a one to many relationship for the employment and unemployment table to the jobs table.

The data was by County. Therefore, SQL queries with subselects were needed to first sum up the employment and unemployment data by county prior to joining the data to the jobs table.

The data was quite large and due to time constraints the queries were limited to just a few columns in the data post import to the database. Queries were built upon all tables and as well by joining all tables together.

We also incorporated a state table in the database with state name, latitude and longitude for use in maps.

Economic Rural Atlas

Portion of uncleaned data

	A	B	C	D	E	F	G	H	I	J
1	FIPS	State	County	UnempRate2020	PctEmpChange1920	UnempRate2019	PctEmpChange1819	UnempRate2018	UnempRate2017	UnempRate2016
2	0	US	United Sta	8.1	-6	3.7	1.3	3.9	4.4	4.9
3	1000	AL	Alabama	5.9	-3.2	3	2.3	3.9	4.6	5.9
4	1001	AL	Autauga	4.9	-3.9	2.7	1.7	3.6	4	5.1
5	1003	AL	Baldwin	5.6	-3.6	2.8	3.7	3.6	4.2	5.4
6	1005	AL	Barbour	7	-2.4	3.8	2.9	5.1	6	8.4
7	1007	AL	Bibb	6.6	-3.6	3.1	1.7	4	4.5	6.5
8	1009	AL	Blount	4.1	-3.5	2.7	1.8	3.5	4.2	5.4

Portions of cleaned data

Jobs:

	A	B	C	D	E	F	G	H
1	fips	econ_state	county	pctemp_agriculture	pctemp_mining	pctemp_construction	pctemp_manufacturing	pctemp_trade
2	0	US	United States	1.259202071	0.51272268	6.592261663	10.10800771	13.74533368
3	1000	AL	Alabama	0.993189611	0.398210342	6.604989835	14.33256857	14.08373479
4	1001	AL	Autauga	0.517902292	0.35478346	6.072098524	12.95163527	12.44596689
5	1003	AL	Baldwin	0.952771556	0.257647937	8.585460243	9.249035135	16.47790012
6	1005	AL	Barbour	5.717342209	0	6.810887912	23.04766433	12.81350291
7	1007	AL	Bibb	1.967330392	1.895791105	9.848575176	16.89519494	12.90091809
8	1009	AL	Blount	1.446365835	0.634210887	9.718483369	17.15563261	14.82410914
9	1011	AL	Bullock	4.908543644	0	5.140078722	36.25839315	5.556841862
10	1013	AL	Butler	2.496847415	0.075662043	3.94703657	25.01891551	12.66078184

Unemployment:

	A	B	C	D	E
1	id	fips	econ_year	unemp_rate	num_unemployed
2	1	0	2020	8.1	12933704
3	2	1000	2020	5.9	131056
4	3	1001	2020	4.9	1262
5	4	1003	2020	5.6	5425
6	5	1005	2020	7	605
7	6	1007	2020	6.6	573
8	7	1009	2020	4.1	1008
9	8	1011	2020	5.5	265
10	9	1013	2020	8.8	801
11	10	1015	2020	7.1	3260
12	11	1017	2020	6.8	1078

Employment:

	A	B	C	D	E
1	id	fips	econ_year	num_civ_labor_force	num_employed
2	16391	0	2020	160611064	147677360
3	16392	1000	2020	2230118	2099062
4	16393	1001	2020	25838	24576
5	16394	1003	2020	96763	91338
6	16395	1005	2020	8587	7982
7	16396	1007	2020	8640	8067
8	16397	1009	2020	24661	23653
9	16398	1011	2020	4818	4553
10	16399	1013	2020	9056	8255
11	16400	1015	2020	46240	42980
12	16401	1017	2020	15865	14787

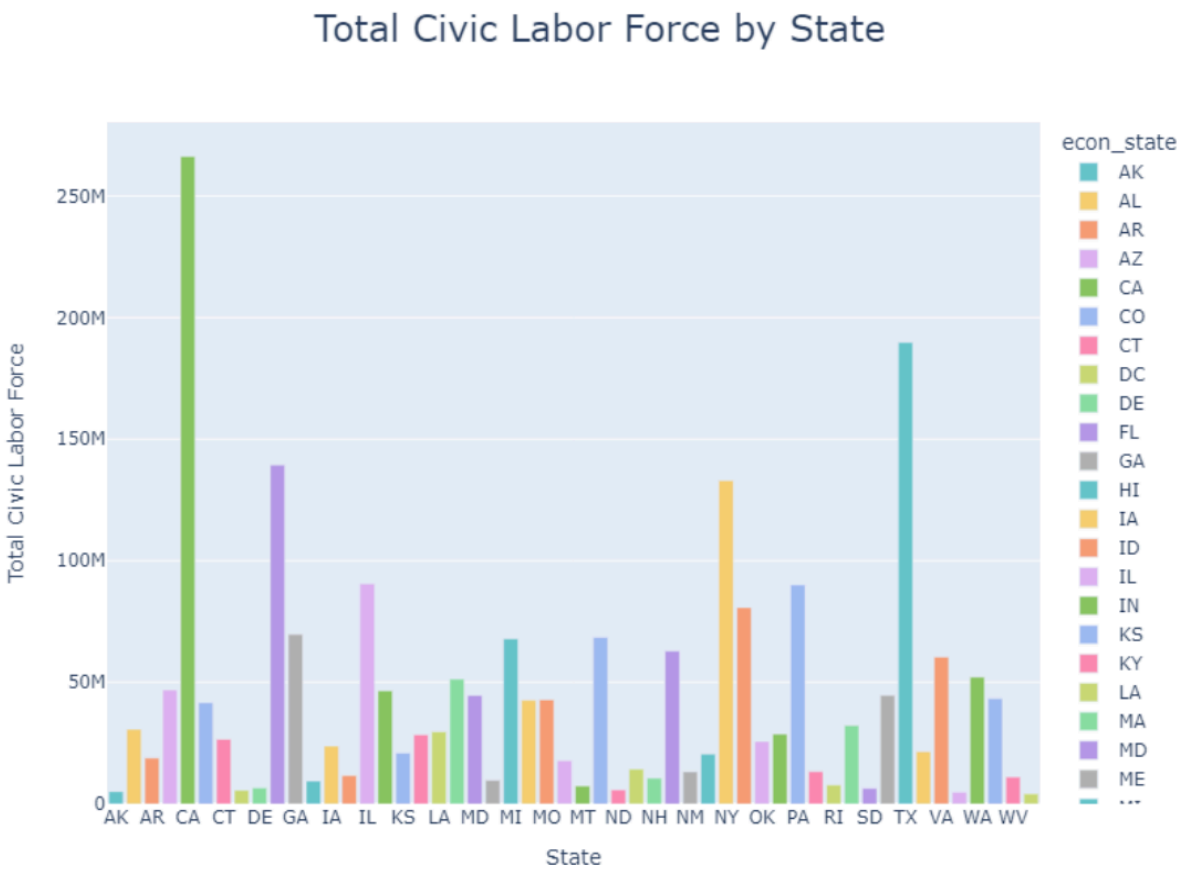
State:

	A	B	C	D
1	econ_stat	latitude	longitude	name
2	AK	63.588753	-154.493062	Alaska
3	AL	32.318231	-86.902298	Alabama
4	AR	35.20105	-91.831833	Arkansas
5	AZ	34.048928	-111.093731	Arizona
6	CA	36.778261	-119.417932	California
7	CO	39.550051	-105.782067	Colorado
8	CT	41.603221	-73.087749	Connecticut
9	DC	38.905985	-77.033418	District of Columbia
10	DE	38.910832	-75.52767	Delaware
11	FL	27.664827	-81.515754	Florida
12	GA	32.157435	-82.907123	Georgia

Database Diagram

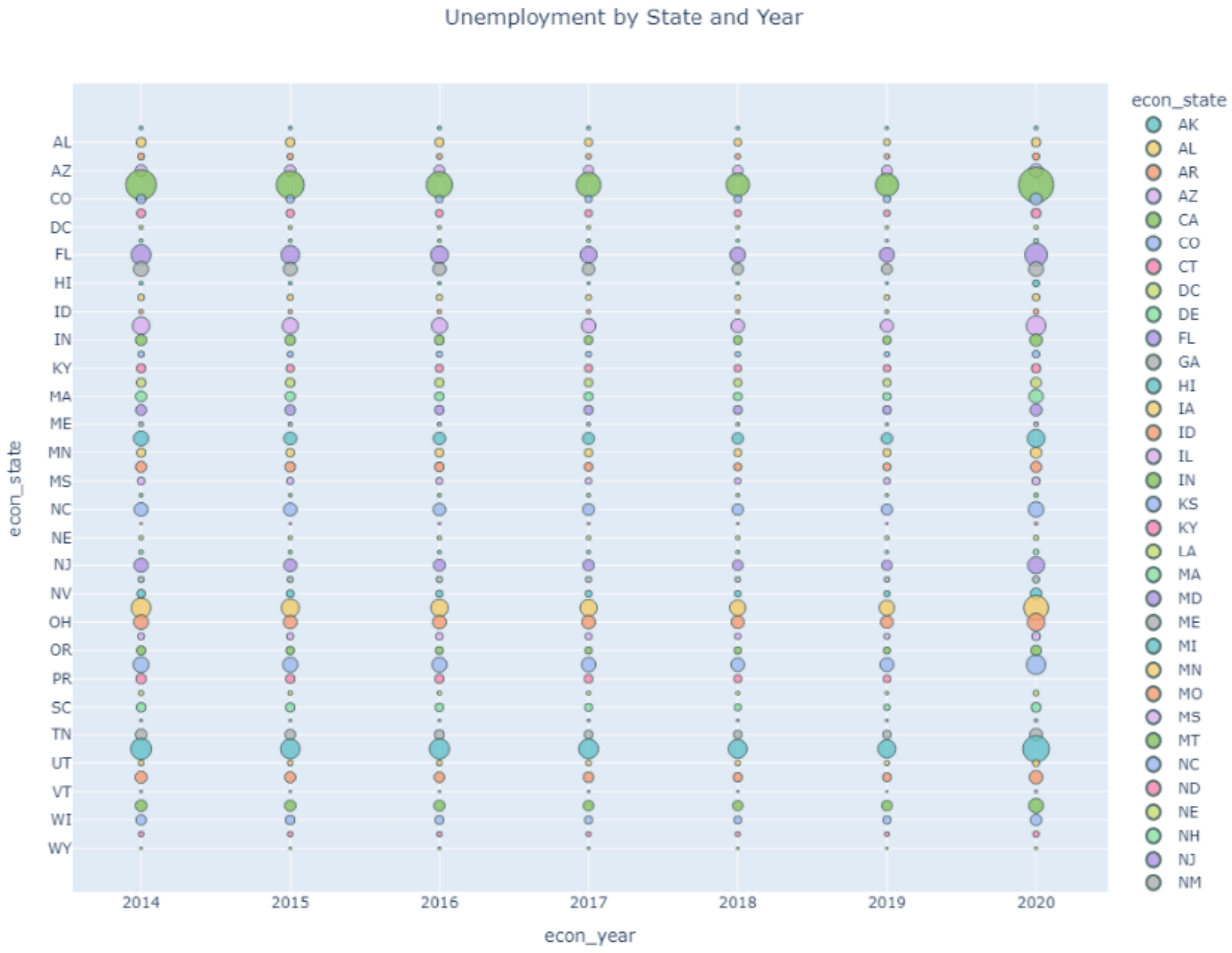
Bar Chart

In this chart we can see the total sum of the civic labor force per state. With little surprise we see the dominant economic power houses of the United States having the largest share of the total national labor force. That being California, Texas, New York, and Florida. To note about these four states, namely what they have in common in terms of their employment sectors, is tourism. Texas, California, and Florida owing as well as to Military bases that provide civilian jobs alongside Federal jobs. In our data we do see some fluctuation between states like Texas and California during the covid years, namely due to things like remote jobs and cost of living differences.



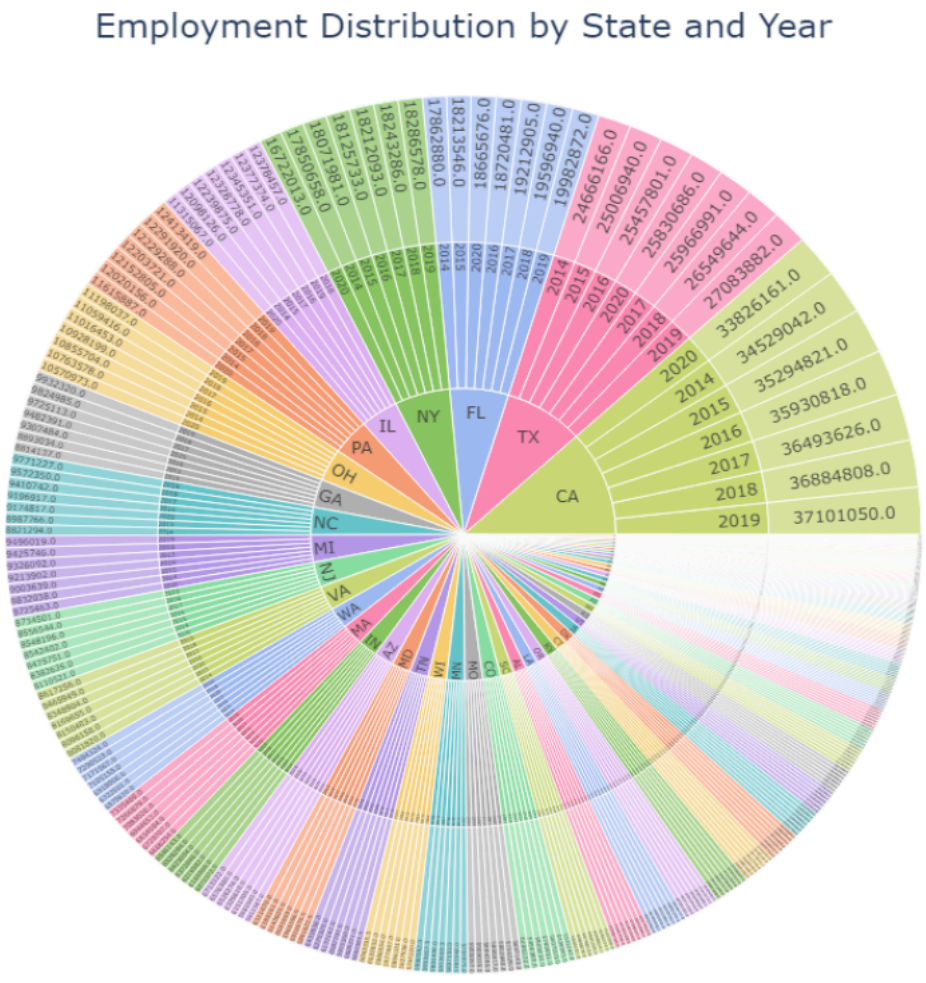
Bubble Chart:

Unemployment trends show that larger, tech-driven states like Texas and California are experiencing the most significant impacts, possibly due to policy-related factors. The most notable differences in unemployment rates occurred between 2014 to 2015 and 2019 to 2020, with the latter likely driven by the effects of COVID-19.

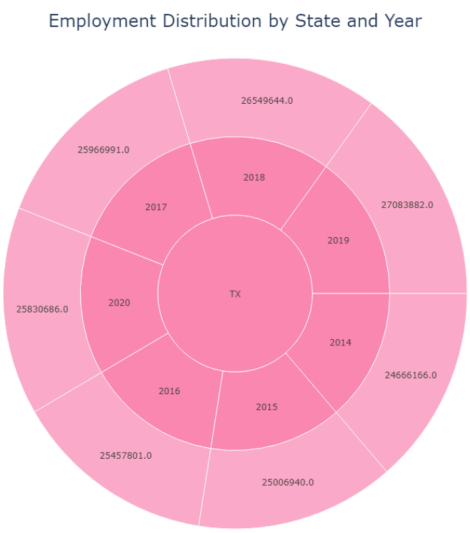


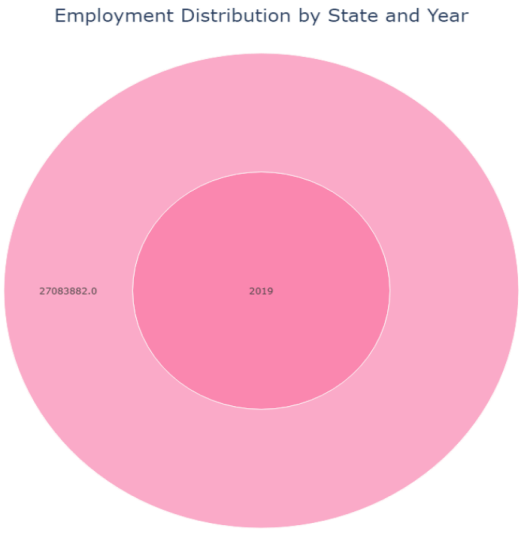
Sunburst Chart:

From 2014 to 2020, employment across states remained relatively stable. However, states like California, Texas, New York, and Florida stand out due to their significant economic activities relative to their populations. Many jobs in California are in the tech, entertainment, tourism, and national park sectors. In Texas, much of its jobs are in areas like tech, trade ports, and military. For New York, jobs are quite similar in areas like imports, exports, tech, and tourism. Finally, Florida boasts many jobs in the military and tourism industries.



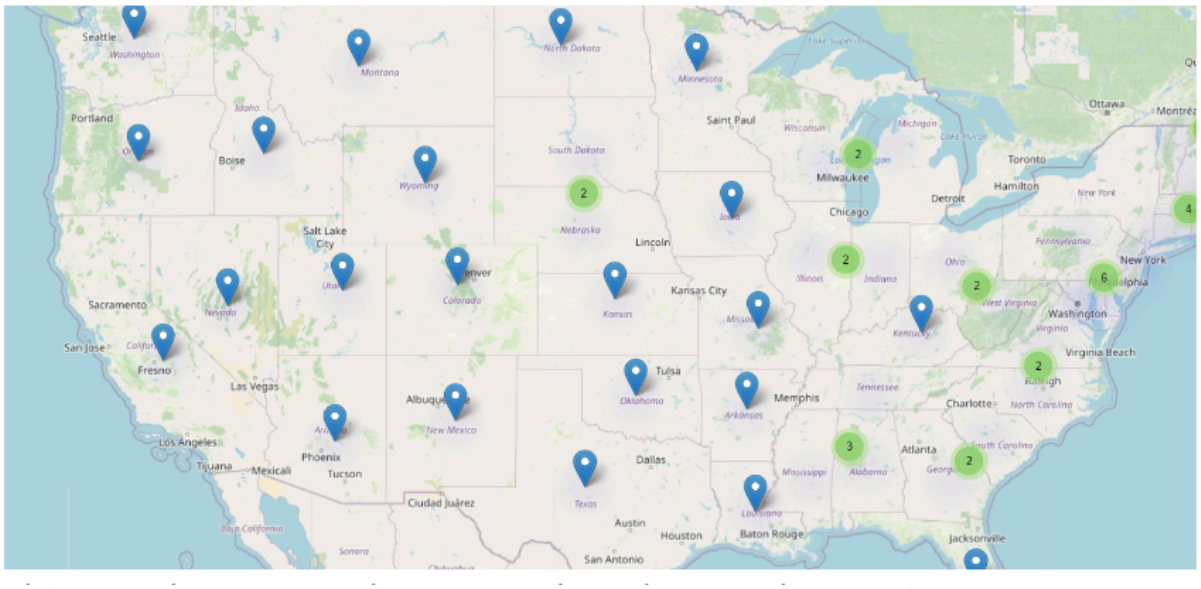
The bubble charts can be broken down to show the year and the number of employed civilians for a given state for a given year (bubble chart shown on left), or centered on state, and broken down to visualize the year and number of employed civilians for that state (bubble chart shown on right). In this example, Texas is shown to have a population of 27,083,882 employed civilians.



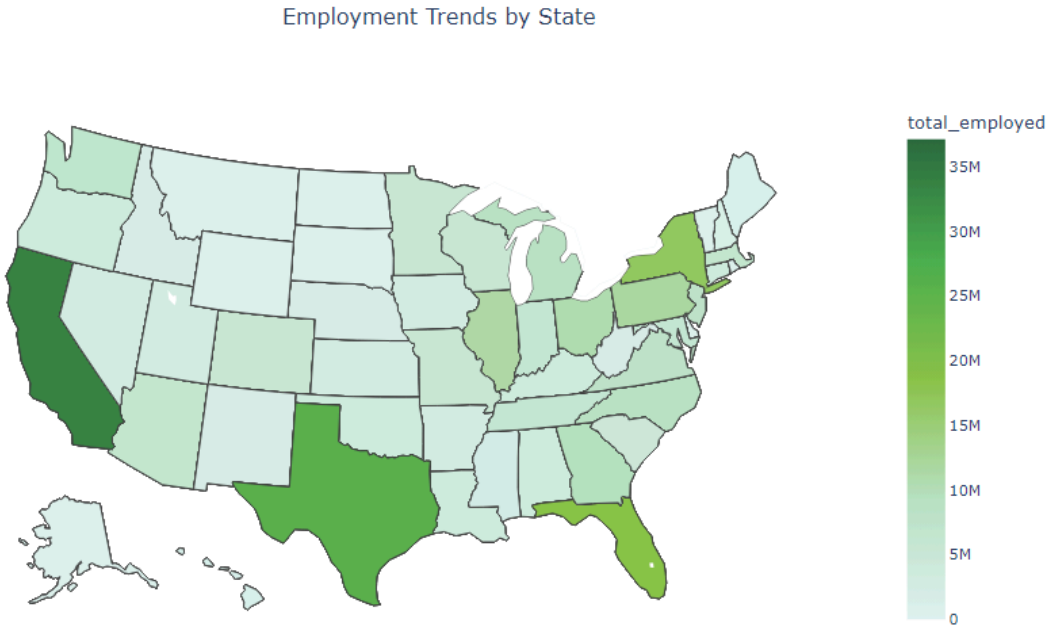


Maps

Through the use of our map we were able to see the total populations of employed and unemployed per state, which had the highest and the lowest. Looking back to other portions of this writeup, we again reflect on the big four, Texas, Florida, California and New York. While states with small overall populations displayed fewer unemployed do to the modifier of their gross population, the Big Four showed a higher unemployment rate than some of the smaller more rural areas.



Through the use of our map we were able to see the total employment by state. Showing the biggest employment states with the largest employment numbers as expected. There were not big surprises in the employment numbers. For future work, we would like to do mapping of employment data based on population to get the employment-to-population ratio per year to identify by state how the employment-to-population ratio is being impacted for each state.



Sources:

Kaggle Dataset:

- Economic Atlas of Rural and Small-Town America (Kaggle) - Source of the dataset used for the project.

Xpert Learning Assistant:

- Xpert Learning Assistant - Used for debugging certain codes.

Chat Gpt:

- Chat Gpt - Used for debugging certain codes.

GitHub Repository:

- https://github.com/cisnerosjp/project3Team 2/tree/main - Served as a reference for project organization and structure.

Texas Tribune Article:

- https://www.texastribune.org/2023/11/21/texas-immigrants-pewresearch/#:~:text=Unauthorized%20immigrants%20make%20up%208,networks%20that%20encourage%20further%20immigration - Provided valuable context and background information related to immigration trends, which informed the broader narrative of the project.

USA Today Article:

- ://www.usatoday.com/story/news/politics/2023/06/21/florida-immigration-lawbusiness-owners-fear-exodus-of-workersconstruction-landscaping/703416320 - Offered insights into the impact of immigration policies on the workforce, contributing to the understanding of economic factors in rural areas.