Forecasting Population and Economic Growth

UNC CHARLOTTE, School of Continuing Studies

DATA ANALYTICS BOOTCAMP

Executive Summary

Overview: This project strives to forecast population growth and analyze its economic implications in various U.S. states. Using data sets related to population growth, cost of living, new businesses, and wages, we have developed a machine-learning model and interface to rank the Best States for 2024. Further, we have employed visualization techniques, such as Tableau dashboards, correlation matrices, and linear regression charts, to present our conclusions effectively.

Objective: The direct objective of this project is to deliver insights into the connection between population growth and economic factors in different states. By forecasting population growth and evaluating its impact on economic indicators like cost of living, new business opportunities, and wages, we aim to identify the Best States for 2024. The project's secondary objective is to showcase the results through visualizations, including a Tableau dashboard, correlation matrix, and linear regression charts, to enhance understanding and foster data-driven decision-making.

The Process

Step 1 → *Data Collection (2020-2022)*

Step 2 → *Data Cleaning & Organization*

Step 3 → *Data Visualization*

+

Step 4 → *Machine Learning Interface*

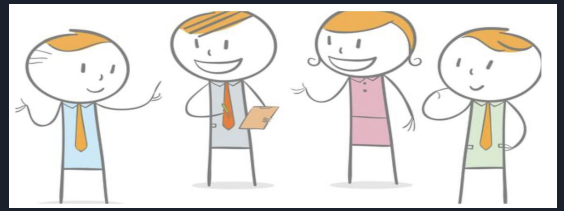
MEMBER	ROLE	RESPONSIBILIITIES
Rahiem Brooks	Machine Learning Specialist	Utilized Machine Learning techniques to develop a user-friendly interface.
Ross Branch	Data Visualization Specialist	Created Scikit-Learn charts to visualize and analyze data.
Karen Leswing	Data Analysis Specialist	Generated linear regression charts to examine relatioships between variables.
Nicholas Best	Data Visualization Specialist	Developed a Tableau dashboard for visualizing and presenting data.
All Members	Data Collection and Cleaning Team	Collaboratively collected and cleaned data to ensure accuracy and consistency.

Project Goals!

01 Our objective was to establish a correlation between the Total Population in each State and the Total Household Income.

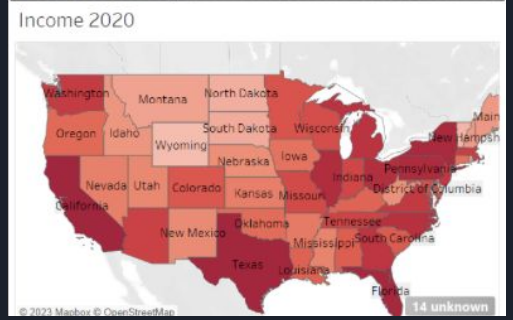
02 Our aim was to present a comprehensive visualization of the correlations between various variables, including Births and Deaths, for each State.

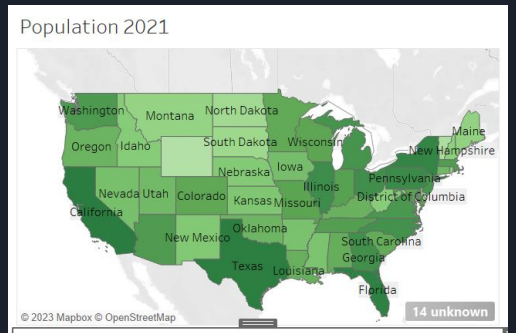
03 Our goal was to create a user-friendly interface that accurately predicts the Best States using data collected from 2020 to 2022.

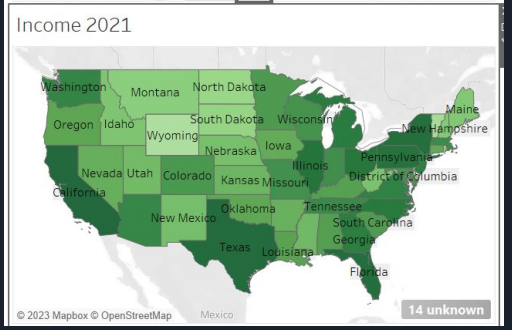


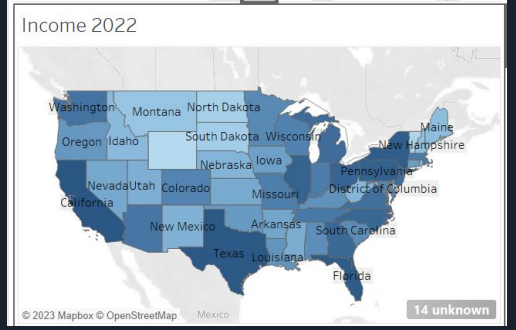
VS

Visualized Comparisons









[Population Dashboard]

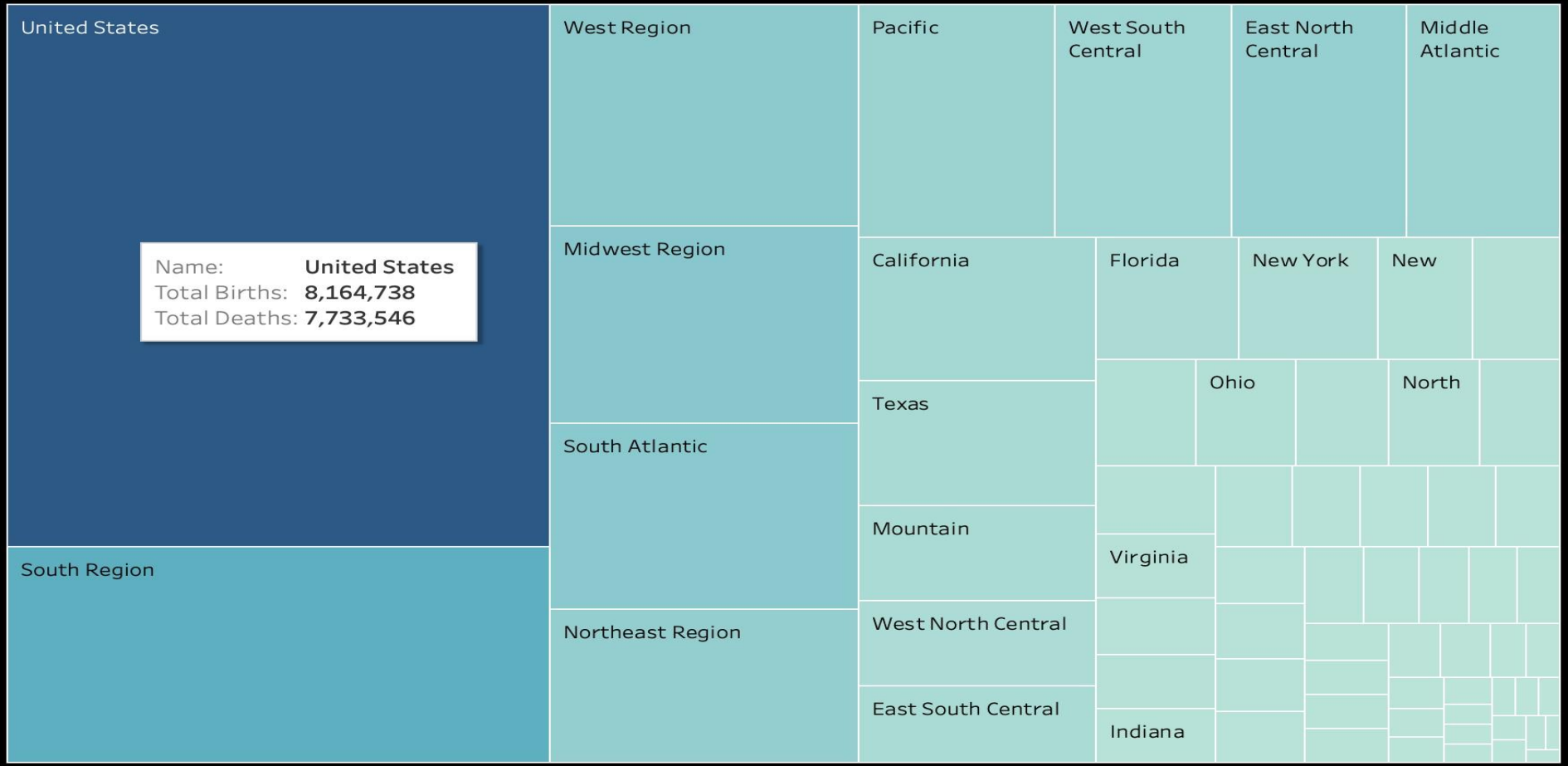
Births and Deaths from 2020-2022

In the course of our analysis, a meticulous examination was conducted to assess the variables in relation to the entire population. The following steps were undertaken:

- Importing the csv file into Tableau.
- Sorting the information to obtain the counts of births and deaths per year, categorized by "NAME".
- Constructing a bar chart to facilitate visual comparisons across different years.
- Creating a treemap visualization to depict the magnitude of the aggregated totals, categorized by "NAME".

By following this systematic approach, we were able to gain valuable insights into the distribution and trends related to births and deaths within the dataset.

Total Births and Deaths per "NAME"



Births and Deaths Total Tableau Link

Births and Deaths Dashboard

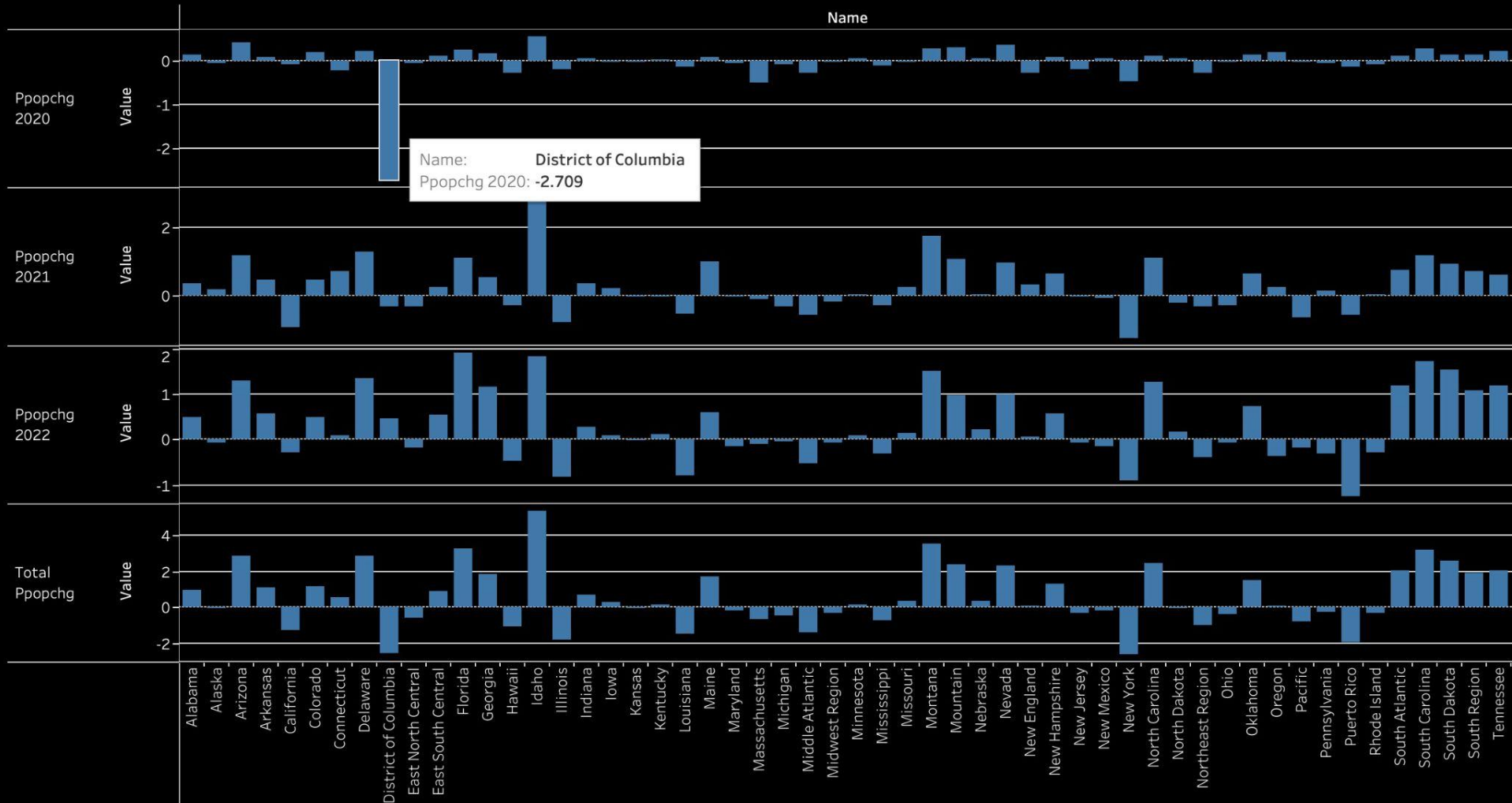
Estimated Population Totals from 2020-2022

During the course of our analysis, we undertook a systematic categorization of the population, focusing on key indicators such as the estimated total resident population, the change in resident total population, and the numerical change in resident total population. Our methodology involved the following steps:

- Importing the relevant dataset in CSV format into the Tableau software platform.
- Extracting the "Popestimate" field, which represents the population estimate for each year, commencing on July 1st.
- Organizing the data by sorting the "Ppopchg" field, which captures the population change on a yearly basis.
- Determining the "Npopchg" values, which represent the numeric population change for each year. This calculation involved subtracting the total population (based on the "NAME" field) from the estimated population total for the year 2020.
- Constructing visually informative bar charts that facilitate comparisons across years and locations (as denoted by the "NAME" field).
- The bar charts effectively highlight both positive and negative values in the "Ppopchg" and "Npopchg" variables, thereby enabling a comprehensive understanding of the dataset.

By adhering to this rigorous and methodical approach, we ensured a robust and professional analysis of the population data.

Population Change in Resident Total Population



Estimated Population Totals Tableau Link

Estimated Population Dashboard

Linear Regression Model

Household income average from 2020-2022

Population change from 2020-2022

```
X = Train_Test_df.loc[:,['AVG_HH_Total_Income', 'AVG_PERCENT_UNDER10K', 'AVG_PERCENT_OVER200K']]
y = Train_Test_df['AVG_PPOP_CHG'].values
```

Linear Regression Model

Predicted Population Change

VS

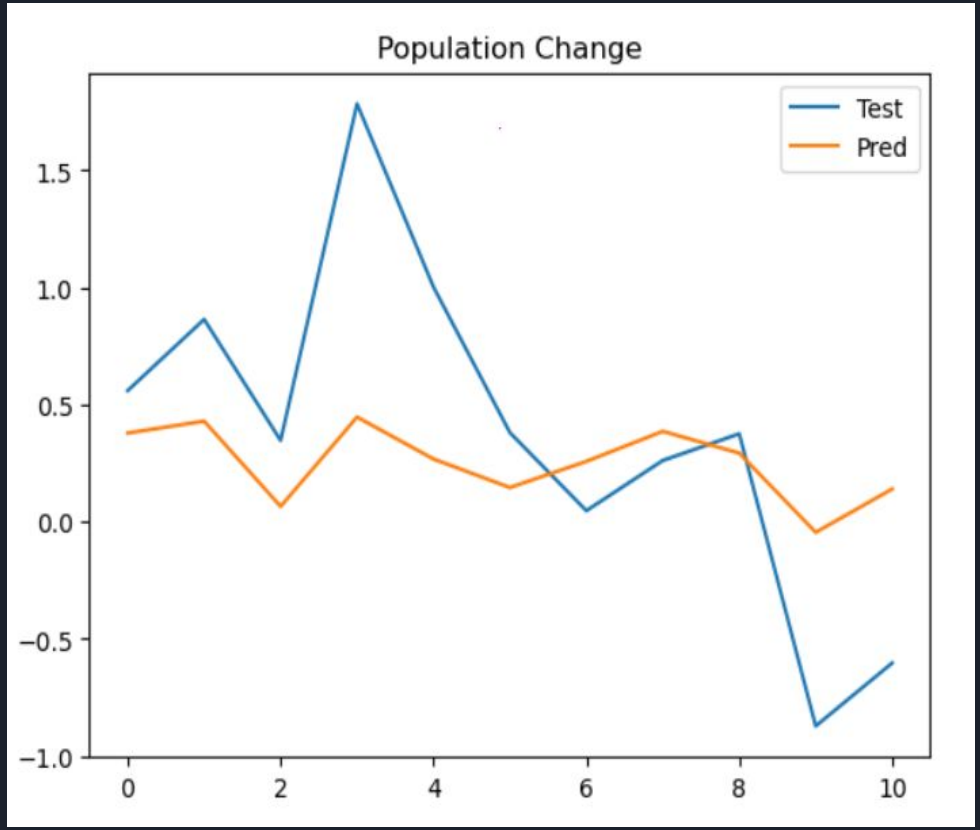
Actual Population Change

```
# show predicted values
pred
```

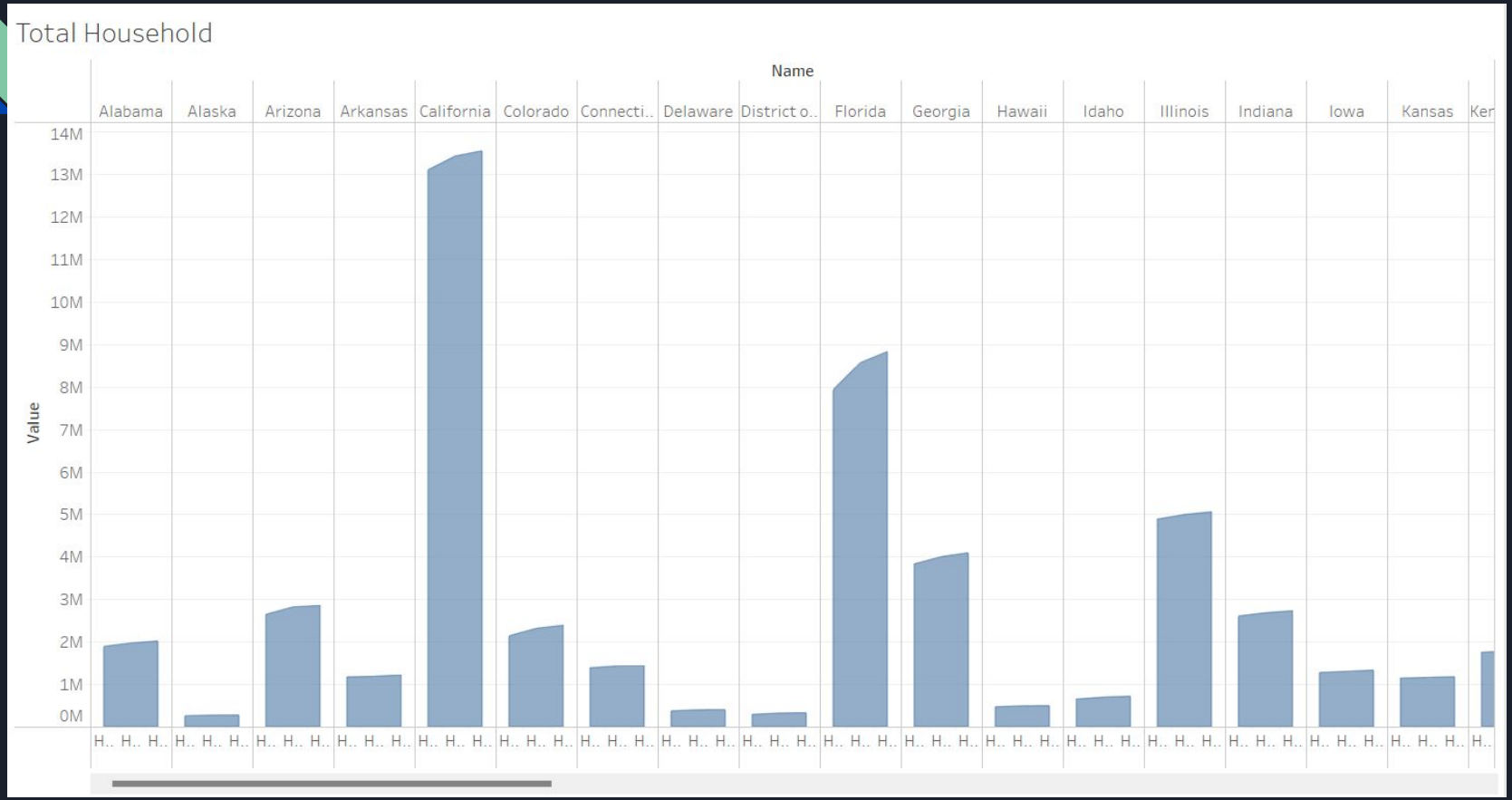
```
array([ 0.37893826,  0.42987143,  0.06576367,  0.44735089,  0.26807433,
        0.14638223,  0.25806814,  0.38648526,  0.29314355, -0.04480252,
        0.13990485])
```

```
# show true values
y_test
```

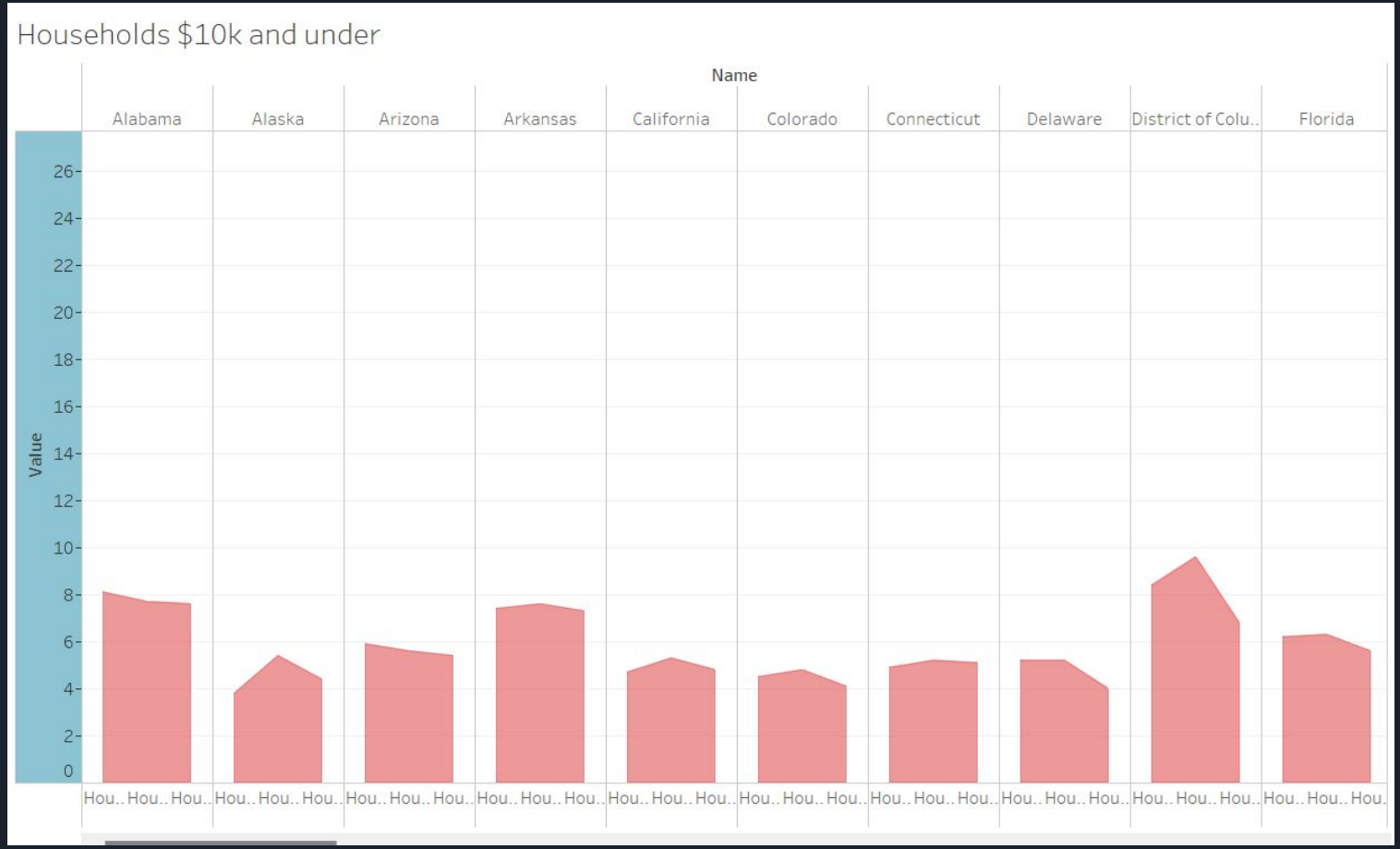
```
array([ 0.56028957,  0.86434408,  0.34734686,  1.78440554,  1.00254532,
        0.38078819,  0.04746681,  0.2619363 ,  0.37604025, -0.87182791,
       -0.60293555])
```



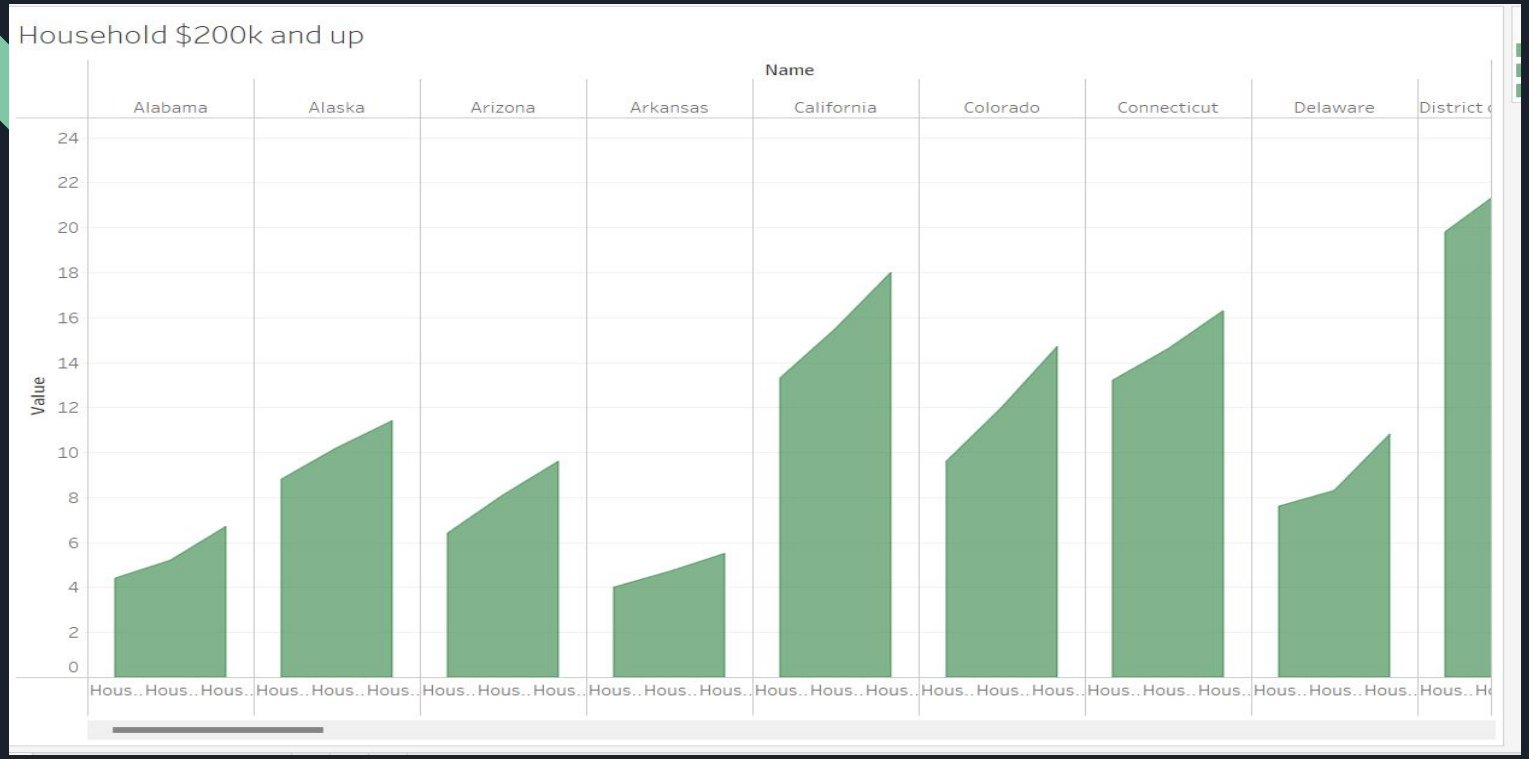
Population with Total Household Income for Years 2020, 2021, 2022



Population with Household income $10k and under per State 2020, 2021, 2022



Population with Household income $200k and up per State 2020, 2021, 2022



[Population Income Dashboard]

Group 4

Best States

2024 Forecast

Machine Learning Model

- Evaluating state performance across multiple factors
- Utilizing a comprehensive methodology with 3,000 data points
- Assessing economic strength, population distribution, business environment, crime and corrections, and opportunities for residents
- Importance of healthcare and education in the rankings
- Consideration of data from 2020-2022, accounting for the COVID-19 pandemic
- Prioritizing state GDP, population growth, employment, and opportunities in the rankings
- Significance of one-year home appreciation in the analysis
- Incorporating measures of crime and corrections, as well as birth and death rates in the assessment

METHODOLOGY

Alabama State GDP Example

2020	235,118.3
2021	257,986.5
2022	281,569

Minimum 235,118.3

Maximum 281,569

NORMALIZATION FORMULA

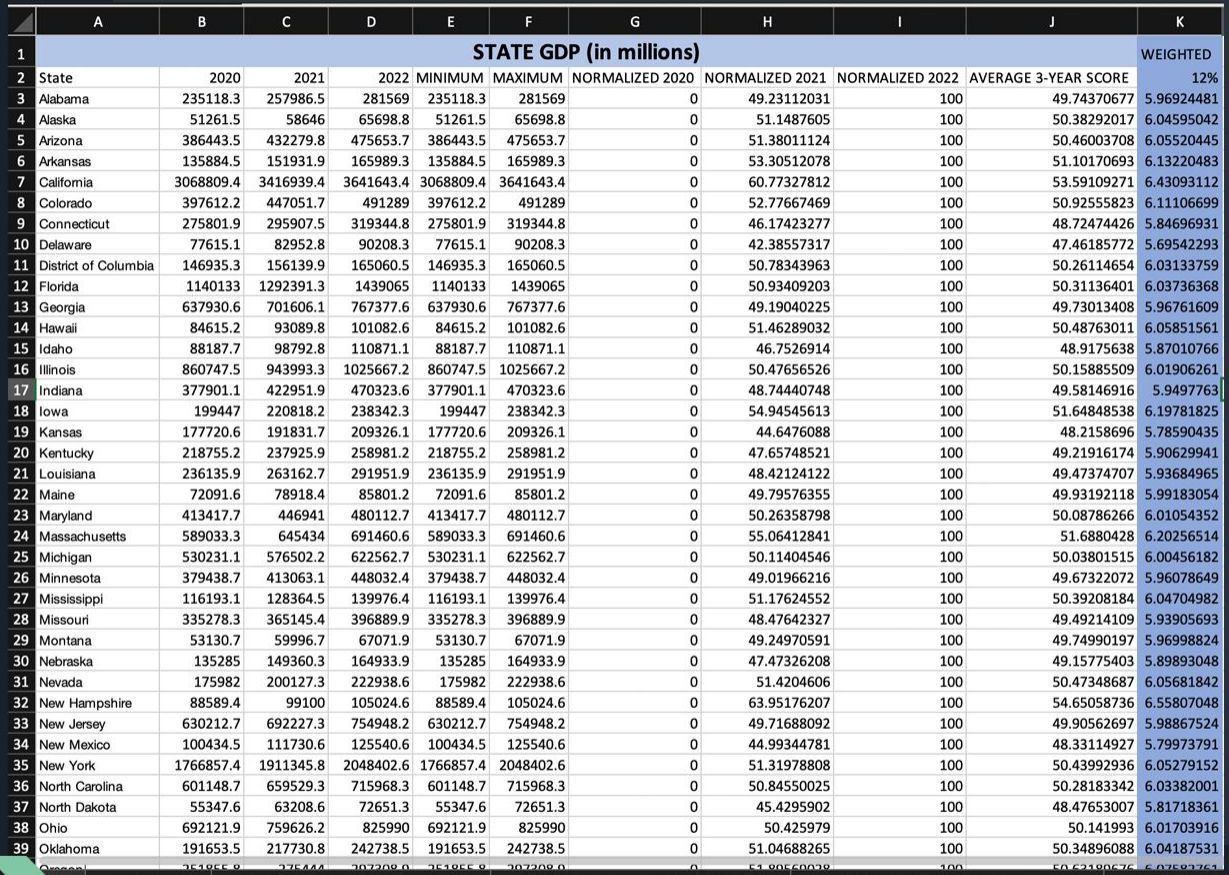
(YEAR-MIN) / (MAX-MIN)*100

THREE YEAR AVERAGE NORMALIZATION

(2020 + 2021 + 2022) /3

WEIGHT FORMULA

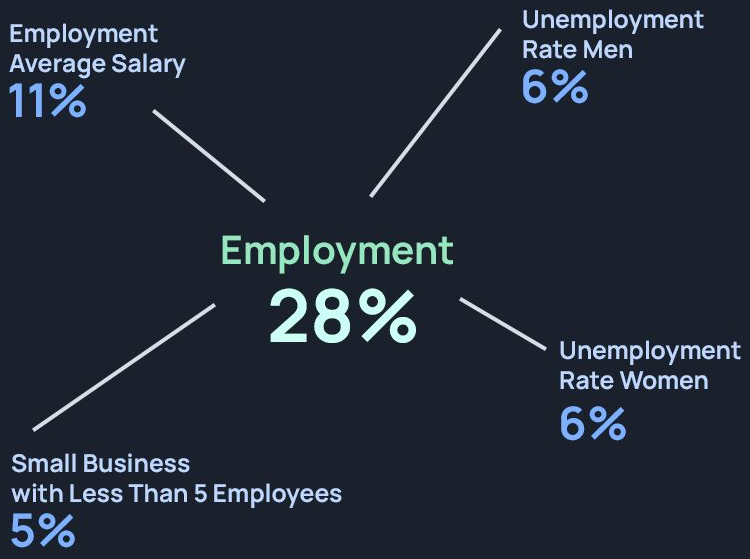
THREE YEARS AVERAGE NORMALIZATION * WEIGHT = SCORE

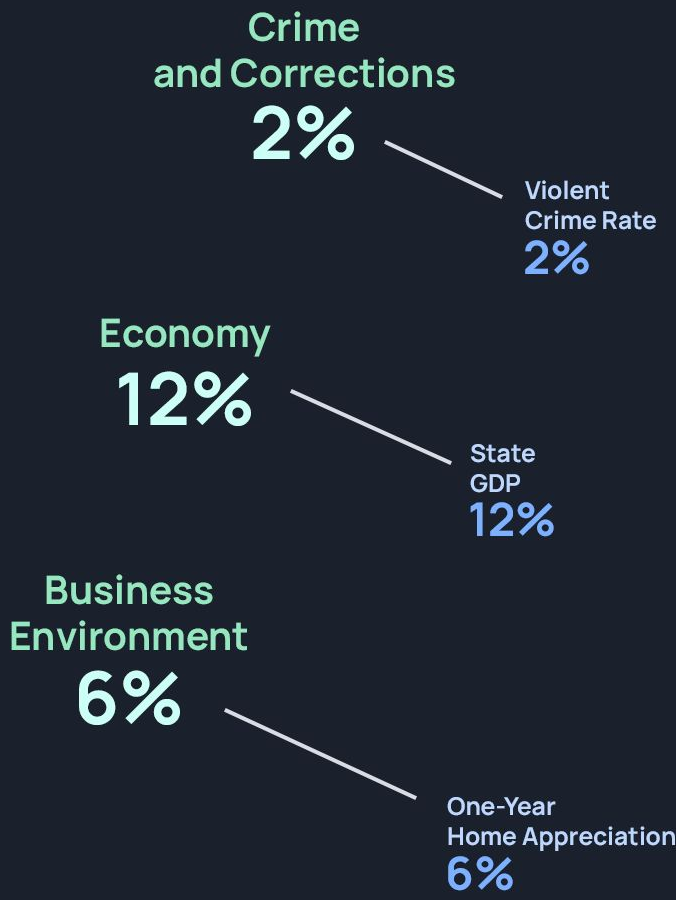


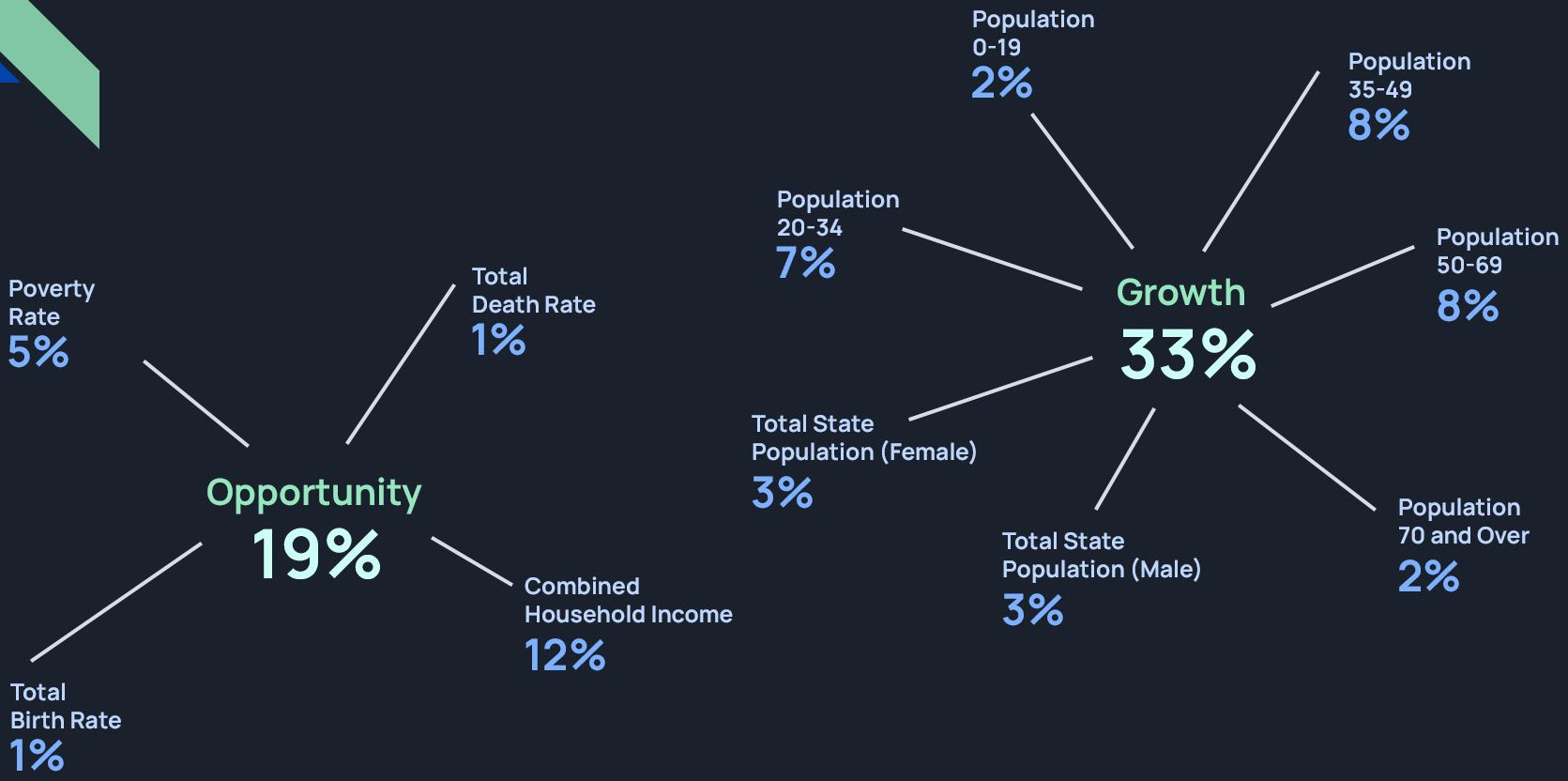
	A	B	C	D	E	F	G	H	I	J	K
1	STATE GDP (in millions)										WEIGHTED
2	State	2020	2021	2022	MINIMUM	MAXIMUM	NORMALIZED 2020	NORMALIZED 2021	NORMALIZED 2022	AVERAGE 3-YEAR SCORE	12%
3	Alabama	235118.3	257986.5	281569	235118.3	281569	0	49.23112031	100	49.74370677	5.96924481
4	Alaska	51261.5	58646	65698.8	51261.5	65698.8	0	51.1487605	100	50.38292017	6.04595042
5	Arizona	386443.5	432279.8	475653.7	386443.5	475653.7	0	51.38011124	100	50.46003708	6.05520445
6	Arkansas	135884.5	151931.9	165989.3	135884.5	165989.3	0	53.30512078	100	51.10170693	6.13220483
7	California	3068809.4	3416939.4	3641643.4	3068809.4	3641643.4	0	60.77327812	100	53.59109271	6.43093112
8	Colorado	397612.2	447051.7	491289	397612.2	491289	0	52.77667469	100	50.92555823	6.11106699
9	Connecticut	275801.9	295907.5	319344.8	275801.9	319344.8	0	46.17423277	100	48.72474426	5.84696931
10	Delaware	77615.1	82952.8	90208.3	77615.1	90208.3	0	42.38557317	100	47.46185772	5.69542293
11	District of Columbia	146935.3	156139.9	165060.5	146935.3	165060.5	0	50.78343963	100	50.26114654	6.03133759
12	Florida	1140133	1292391.3	1439065	1140133	1439065	0	50.93409203	100	50.31136401	6.03736368
13	Georgia	637930.6	701606.1	767377.6	637930.6	767377.6	0	49.19040225	100	49.73013408	5.96761609
14	Hawaii	84615.2	93089.8	101082.6	84615.2	101082.6	0	51.46289032	100	50.48763011	6.05851561
15	Idaho	88187.7	98792.8	110871.1	88187.7	110871.1	0	46.7526914	100	48.9175638	5.87010766
16	Illinois	860747.5	943993.3	1025667.2	860747.5	1025667.2	0	50.47656526	100	50.15885509	6.01906261
17	Indiana	377901.1	422951.9	470323.6	377901.1	470323.6	0	48.74440748	100	49.58146916	5.9497763
18	Iowa	199447	220818.2	238342.3	199447	238342.3	0	54.94545613	100	51.64848538	6.19781825
19	Kansas	177720.6	191831.7	209326.1	177720.6	209326.1	0	44.6476088	100	48.2158696	5.78590435
20	Kentucky	218755.2	237925.9	258981.2	218755.2	258981.2	0	47.65748521	100	49.21916174	5.90629941
21	Louisiana	236135.9	263162.7	291951.9	236135.9	291951.9	0	48.42124122	100	49.47374707	5.93684965
22	Maine	72091.6	78918.4	85801.2	72091.6	85801.2	0	49.79576355	100	49.93192118	5.99183054
23	Maryland	413417.7	446941	480112.7	413417.7	480112.7	0	50.26358798	100	50.08786266	6.01054352
24	Massachusetts	589033.3	645434	691460.6	589033.3	691460.6	0	55.06412841	100	51.6880428	6.20256514
25	Michigan	530231.1	576502.2	622562.7	530231.1	622562.7	0	50.11404546	100	50.03801515	6.00456182
26	Minnesota	379438.7	413063.1	448032.4	379438.7	448032.4	0	49.01966216	100	49.67322072	5.96078649
27	Mississippi	116193.1	128364.5	139976.4	116193.1	139976.4	0	51.17624552	100	50.39208184	6.04704982
28	Missouri	335278.3	365145.4	396889.9	335278.3	396889.9	0	48.47642327	100	49.49214109	5.93905693
29	Montana	53130.7	59996.7	67071.9	53130.7	67071.9	0	49.24970591	100	49.74990197	5.96998824
30	Nebraska	135285	149360.3	164933.9	135285	164933.9	0	47.47326208	100	49.15775403	5.89893048
31	Nevada	175982	200127.3	222938.6	175982	222938.6	0	51.4204606	100	50.47348687	6.05681842
32	New Hampshire	88589.4	99100	105024.6	88589.4	105024.6	0	63.95176207	100	54.65058736	6.55807048
33	New Jersey	630212.7	692227.3	754948.2	630212.7	754948.2	0	49.71688092	100	49.90562697	5.98867524
34	New Mexico	100434.5	111730.6	125540.6	100434.5	125540.6	0	44.99344781	100	48.33114927	5.79973791
35	New York	1766857.4	1911345.8	2048402.6	1766857.4	2048402.6	0	51.31978808	100	50.43992936	6.05279152
36	North Carolina	601148.7	659529.3	715968.3	601148.7	715968.3	0	50.84550025	100	50.28183342	6.03382001
37	North Dakota	55347.6	63208.6	72651.3	55347.6	72651.3	0	45.4295902	100	48.47653007	5.81718361
38	Ohio	692121.9	759626.2	825990	692121.9	825990	0	50.425979	100	50.141993	6.01703916
39	Oklahoma	191653.5	217730.8	242738.5	191653.5	242738.5	0	51.04688265	100	50.34896088	6.04187531
	Oregon	251855.8	275444	297308.9	251855.8	297308.9	0	51.88560928	100	50.62189676	6.07587763

RANKINGS | MALE UNEMPLOYMENT RATE | FEMALE UNEMPLOYMENT RATE | COMBINED HOUSEHOLD INCOME | EMPLOYME

CATEGORIZATION



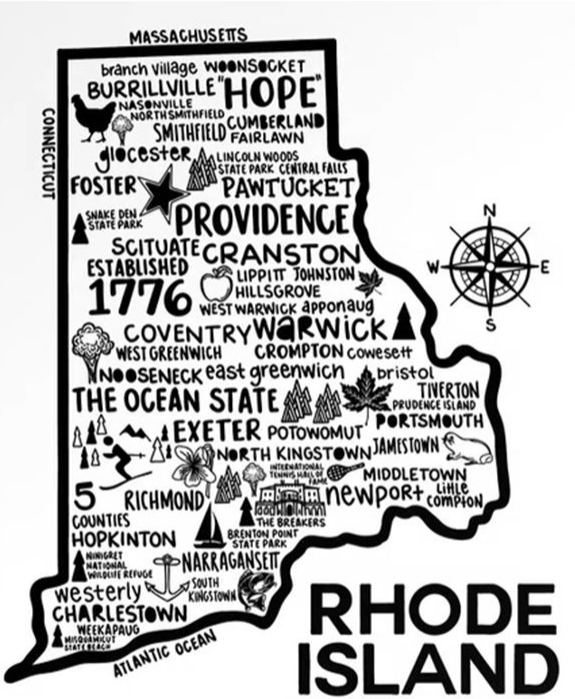



Poverty Rate
5%
Total Death Rate
1%
Opportunity
19%
Total Birth Rate
1%
Combined Household Income
12%
Population 0-19
2%
Population 35-49
8%
Population 20-34
7%
Growth
33%
Population 50-69
8%
Total State Population (Female)
3%
Total State Population (Male)
3%
Population 70 and Over
2%

49 Utah
49 Indiana
49 Florida
48 Alabama
47 Missouri
46 Colorado
45 South Carolina
44 Mississippi
43 Nebraska
42 Tennessee
40 North Carolina
40 Michigan
39 Kentucky
38 Ohio
37 Arizona
36 District of Columbia
35 Vermont
34 Alaska
33 New Mexico
32 Iowa
31 Idaho
30 Georgia

29 Wisconsin
28 Delaware
27 Texas
26 Virginia
25 Connecticut
24 Pennsylvania
23 Wyoming
22 Maryland
21 South Dakota
20 Illinois
19 New Hampshire
18 New Jersey
17 Louisiana
16 Oklahoma
15 Washington
14 Maine
13 California
12 Oregon
11 Massachusetts
10 Minnesota
09 North Dakota
08 Arkansas
07 Montana
06 Hawaii

5.


MASSACHUSETTS
CONNECTICUT
branch village
WOONSOCKET
BURRILLVILLE
"HOPE"
NASONVILLE
NORTH SMITHFIELD
SMITHFIELD
CUMBERLAND
FAIRLAWN
glocester
LINCOLN WOODS STATE PARK
CENTRAL FALLS
FOSTER
PAWTUCKET
SNAKE DEN STATE PARK
PROVIDENCE
SCITUATE
ESTABLISHED
1776
CRANSTON
LIPPITT
JOHNSTON
HILLSGROVE
WEST WARWICK
apponaug
COVENTRY
WARWICK
WEST GREENWICH
CROMPTON
cowesett
NOOSENECK
east greenwich
bristol
TIVERTON
PRUDENCE ISLAND
THE OCEAN STATE
EXETER
POTOWOMUT
PORTSMOUTH
NORTH KINGSTOWN
JAMESTOWN
INTERNATIONAL TENNIS HALL OF FAME
MIDDLETOWN
5
RICHMOND
newport
little compton
COUNTIES
THE BREAKERS
HOPKINTON
BRENTON POINT STATE PARK
NARRAGANSET
MINIGRET NATIONAL WILDLIFE REFUGE
SOUTH KINGSTOWN
westerly
CHARLESTOWN
WEEKAPAUG
MISQUAMICUT STATE BEACH
ATLANTIC OCEAN
N
W
E
S
RHODE ISLAND

4.


N
W
E
S
WHEELING
PENNSYLVANIA
WESTOVER
MORGANTOWN
WEST VIRGINIA UNIVERSITY
fairmont
VIENNA
PLEASANT VALLEY
GRAFTON
KINGWOOD
PARKERSBURG
clarksburg
BRIDGEPORT
PHILIPPI
SHINNSTON
STONEWOOD
LITTLE KANAWHA RIVER
MONONGAHELA RIVER
WESTON
TYGART VALLEY RIVER
BUCKHANNON
ELKINS
ravenswood
FONZO
SPENCER
GLENVILLE
RIPLEY
ORMA
"MOUNTAINEERS ARE ALWAYS FREE"
POINT PLEASANT
AMMA
ELKVIEW
BULLTOWN
WHEELER
NITRO
ST ALBANS
HURRICANE
CAMDEN PARK
CHARLESTON
KUMBRABOW STATE FOREST
RICHWOOD
MONONGAHELA NATIONAL FOREST
SNOWSHOE MOUNTAIN
GAULEY RIVER
huntington
UNIVERSITY OF CHARLESTON
DUNBAR
SOUTH CHARLESTON
KANAWHA RIVER
SUMMERSVILLE
MADISON
FOUNDED 1863
OAK HILL
NEW RIVER
GORGE NATIONAL RIVER
NEW RIVER
HAWKS NEST STATE PARK
BECKLEY
HINTON
HARTS
CHAPMANVILLE
SOPHIA
BEAVER
LEWISBURG
WILLIAMSON
"THE MOUNTAIN STATE"
LOGAN
OCEANA
GILBERT
PINEVILLE
ATHENS
CONCORD UNIVERSITY
WELCH
princeton
BLUEFIELD
MARTINSBURG
HARPERS FERRY
RANSON
CHARLES TOWN
KEYSER
"TAKE ME HOME COUNTRY ROADS"
55 COUNTIES
APPALACHIAN MOUNTAINS
ROMNEY
WEST VIRGINIA

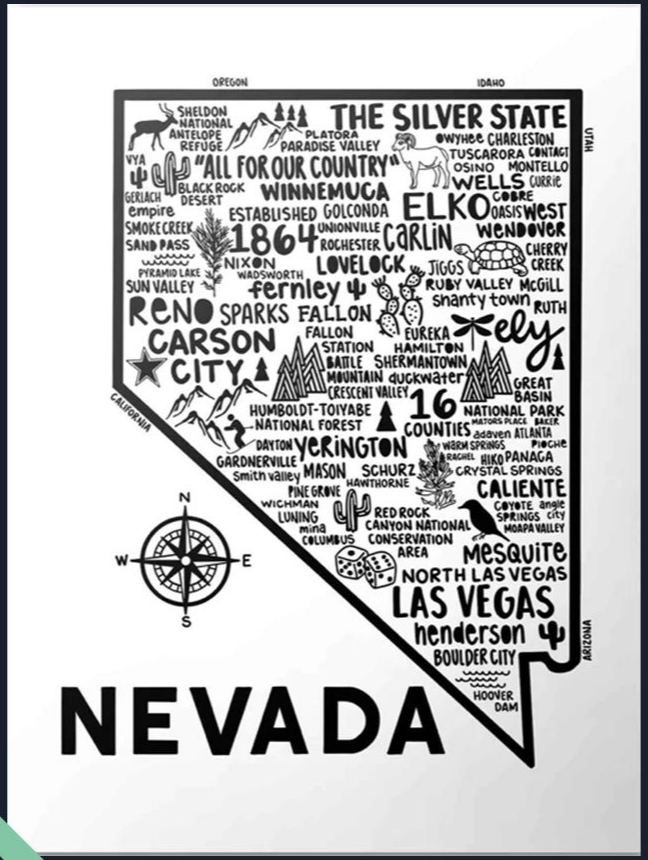
3.


NEBRASKA
THE SUNFLOWER STATE
THERE'S NO PLACE LIKE HOME
WASHINGTON
BELLEVILLE
CONCORDIA
seneca
SABETHA
Hiawatha
kansas city
ATCHINSON
LANSING
LEAVENWORTH
Lawrence
OZ MUSEUM
BREWSTER
COLBY
Hill City
PHILLIPSBURG
STOCKTON
KANORADO
Wakeeney
WORLD'S LARGEST BALL OF TWINE
OSBORNE
BELOIT
LINCOLN
KANSAS RIVER
MANHATTAN
DWIGHT D EISENHOWER PRESIDENTIAL LIBRARY & MUSEUM
SMOKY HILL RIVER
JUNCTION CITY
TOPEKA
WESKAN
105 COUNTIES
HAYS
Plainville
DIGHTON
LA CROSSE
WILSON STATE PARK
SALINA
meriam
SHAWNEE
BONNER SPRINGS
PRAIRIE VILLAGE
Lenexa
OVERLAND PARK
LEAWOOD
OLATHE
garden CITY
MUSHROOM ROCK STATE PARK
TALL GRASS PRAIRIE NATIONAL PRESERVE
mission
GREAT BEND
McPherson
MONUMENT ROCKS NATIONAL NATURAL LANDMARK
COSMOSPHERE
bel aire
EMPORIA
GARDNER
OTTAWA
SPRING HILL
TRIBUNE
deerfield
pierceville
FORT LARNED NATIONAL HISTORIC SITE
HUTCHINSON
HESSTON
el dorado
COOLIDGE
SYRACUSE
JOHNSON CITY
CIMMARON
valley CENTER
NEWTON
ANDOVER
MARMATON RIVER
FORT SCOTT
FORTSCOTT NATIONAL HISTORIC SITE
COLORADO
MANTER
ULYSSES
DODGE CITY
ESTABLISHED
PARK CITY
CHENEY
WICHITA
AUGUSTA
FALL RIVER
FREDONIA
CHANUTE
ST PAUL
CIMARRON NATIONAL GRASSLAND
1861
NORWICH
HARPER
CLEARWATER
ARGONIA
HAYSVILLE
DERBY
PARSONS
PITTSBURG
Liberal
"AD ASTRA PER ASPERA"
WELLINGTON
winfield
INDEPENDENCE
SOUTH HAVEN
ARKANSAS CITY
COFFEYVILLE
OKLAHOMA
MISSOURI
KANSAS
N
W
E
S

2.


NEW YORK
CANADA
LAKE ONTARIO
LAKE ERIE
PENNSYLVANIA
VERMONT
MASSACHUSETTS
CONNECTICUT
NEW JERSEY
MASSENA
FORT COVINGTON
NORFOLK
CHATEAUGAY
CHAMPLAIN
NORWOOD
POTSDAM
Plattsburgh
OGDENSBURG
CANTON
PERU
WILMINGTON
ADIRONDACK STATE PARK
WHITEFACE MOUNTAIN
ALEXANDRIA BAY
CLAYTON
GOUVERNEUR
CAPE VINCENT
HARRISVILLE
CHAUMONT
DIANA CENTER
FORT DRUM
CARTHAGE
LAKE PLACID
WATERTOWN
LOWVILLE
CROGHAN
ADIRONDACK MOUNTAINS
ADAMS CENTER
LYONS FALLS
LAKE GEORGE
LACONA
GUEENSBURY
PULASKI
62 COUNTIES
GLENS FALLS
OSWEGO
CICERO
FULTON
ROME
AMSTERDAM
SARATOGA SPRINGS
SUNY EMPIRE STATE COLLEGE
LOCKPORT
NORTH TONAWANDA
TONAWANDA
ROCHESTER
WILLIAMSON
NEWARK
SYRACUSE
SYRACUSE UNIVERSITY
SCHENECTADY
ROTTERDAM
COHOES
GLOVERSVILLE
COLONIE
WATERVLIET
TROY
NIAGARA FALLS
BATAVIA
CANANDAIGUA
AUBURN
ONEIDA
UTICA
BUFFALO
GENEVA
MORAVIA
THE EMPIRE STATE
ALBANY
UNIVERSITY AT BUFFALO
FINGER LAKES
DRYDEN
LACKAWANNA
WATKINS GLEN STATE PARK
FINGER LAKES NATIONAL FOREST
LANSING
11TH CORTLAND TO JOIN THE UNION
COOPERSTOWN
EXCELSIOR COLLEGE
SUNY ALBANY
RENSSELAER
DUNKIRK
ASHFORD
NORWICH
ONEONTA
EST 1776
"EXCELSIOR"
CORNELL UNIVERSITY
ITHACA
OXFORD
CATSKILL
GREAT VALLEY
SALAMANCA
HORNELL
HORSEHEADS
AFTON
MOUNTAINS
FREDONIA
CHAUTAUQUA
LAKEWOOD
JAMESTOWN
ALLEGHENY STATE PARK
ALFRED
CORNING
BINGHAMTON
HUDSON RIVER
OLEAN
ELMIRA
WOODSTOCK
KINGSTON
POUGHKEEPSIE
HUDSON VALLEY STATE PARK
NEWBURGH
BEACON
MIDDLETOWN
CHESTER
PEEKSKILL
WOODBURY
RYE
WHITE PLAINS
NEW ROCHELLE
YONKERS
MONTAUK
THE HAMPTONS
LONG ISLAND
NEW YORK CITY
I ♥ NY
N
W
E
S

1.


OREGON
IDAHO
UTAH
ARIZONA
CALIFORNIA
SHELDON NATIONAL ANTELOPE REFUGE
THE SILVER STATE
PLATORA
PARADISE VALLEY
OWYHEE
CHARLESTON
TUSCARORA
CONTACT
OSINO
MONTELLO
WELLS
CURRIE
VYA
"ALL FOR OUR COUNTRY"
GERLACH
BLACK ROCK DESERT
WINNEMUCA
empire
ESTABLISHED
GOLCONDA
ELKO
COBRE
OASIS
WEST WENDOVER
SMOKE CREEK
SAND PASS
1864
UNIONVILLE
ROCHESTER
CARLIN
CHERRY CREEK
NIXON
PYRAMID LAKE
WADSWORTH
LOVELOCK
JIGGS
RUBY VALLEY
McGILL
SUN VALLEY
fernley
shanty town
RUTH
RENO
SPARKS
FALLON
EUREKA
ely
CARSON CITY
FALLON STATION
BATTLE MOUNTAIN
HAMILTON
SHERMANTOWN
duckwater
CRESCENT VALLEY
GREAT BASIN
HUMBOLDT-TOIYABE NATIONAL FOREST
16 COUNTIES
NATIONAL PARK
MATHERS PLACE
BAKER
adaven
ATLANTA
DAYTON
VERINGTON
WARM SPRINGS
RACHEL
HIKO
PANACA
PIOCHE
GARDNERVILLE
Smith valley
MASON
SCHURZ
CRYSTAL SPRINGS
HAWTHORNE
CALIENTE
PINE GROVE
WICHMAN
LUNING
mina
COLUMBUS
RED ROCK CANYON NATIONAL CONSERVATION AREA
COYOTE SPRINGS
angle city
MOAPA VALLEY
mesquite
NORTH LAS VEGAS
LAS VEGAS
henderson
BOULDER CITY
HOOVER DAM
N
W
E
S
NEVADA

- Making data-driven decisions easier and more efficient
- Transitioning to the Machine Learning Interface
- Moving from analysis to real-time predictions

Machine Learning Interface