Descriptive analysis of the Mortality Rate of different states in the US for the year 2010.

Rajkumar Gosai, Nidhi Paliwal, Aakash Dogra

ILLINOIS
School of Information Sciences

Introduction

Despite all healthcare facilities present today, we can see that several counties have a high mortality rate. These numbers are surprisingly high, despite the fact that the United States has the best healthcare facility in the world.

For profoundly understanding the reasons and depicting why different states in the US have different mortality rates, we try to analyze the Community Health Status Indicators (CHSI) dataset. This dataset provides detailed data of mortality rate on basis of Demographic factors, Leading causes of deaths, infant mortality rates due to unavoidable factors, a population vulnerable to fatal diseases and how the quality of air has been a contributing factor.
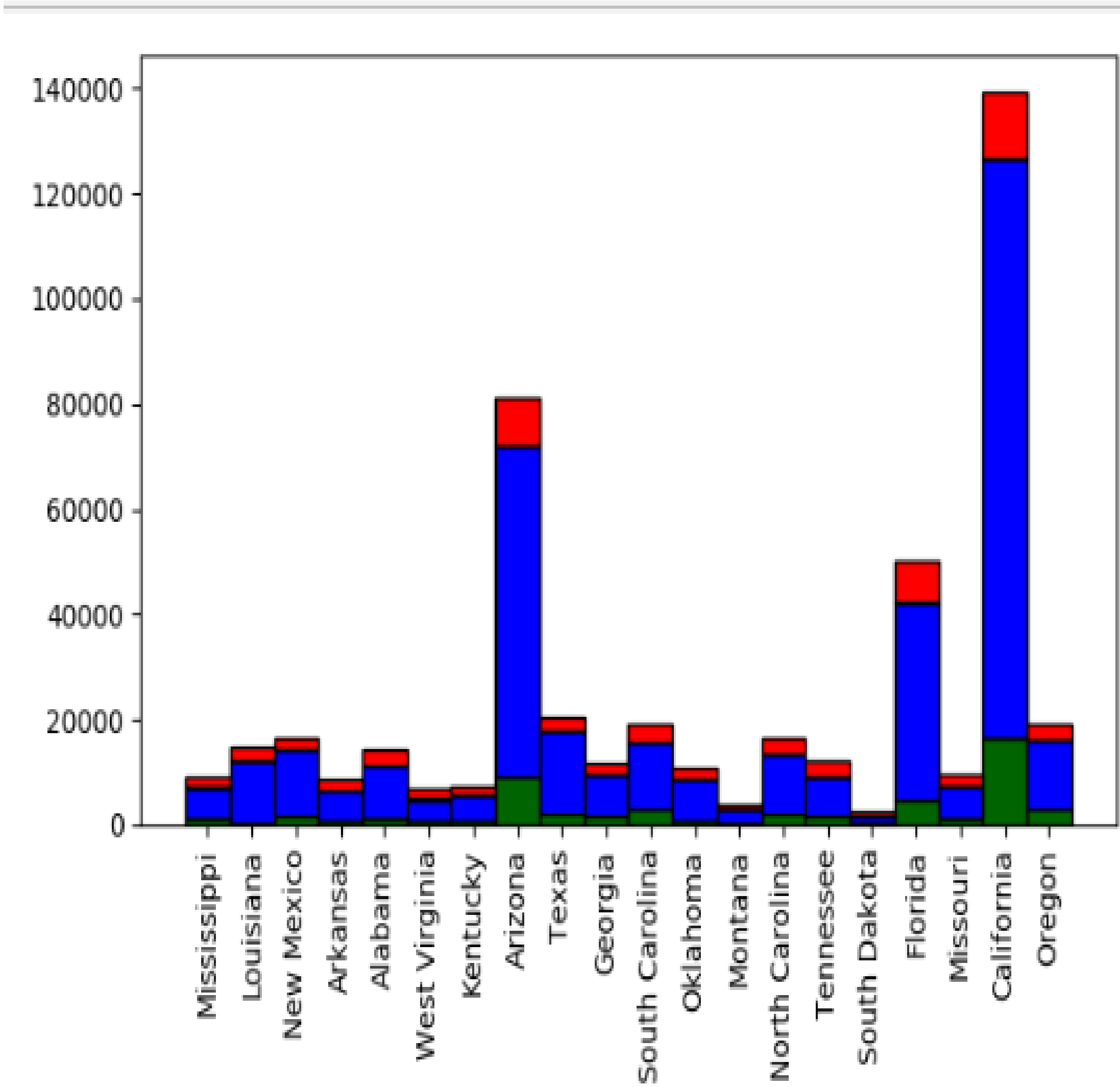
Hypothesis

1. It has been observed that the mortality rate among individuals of diverse ethnic backgrounds and race is disturbingly varying. This means that individuals from different ethnic backgrounds and races do not have equal access to medical facilities and are prone to fatal diseases.

2. Pollution and genetic defects have been the leading causes of cancer and major respiratory diseases. In addition to this drug abuse, smoking and unhealthy lifestyle choices have been major contributing factors in increasing mortality rate.

3. The mortality rate in states with higher poverty index and higher risk factors are greater as compared to the ones with lower risk factors.

Future Analysis

• The dataset does not possess any quantifiable column for understanding the duration of the census. Since the data was updated recently in February 2019, we can assume the data to be latest.

• The genetic defects considers only the birth defects in the general population(up to the age 1). Genetic defects for the age group 14 and above can be analyzed with more data.
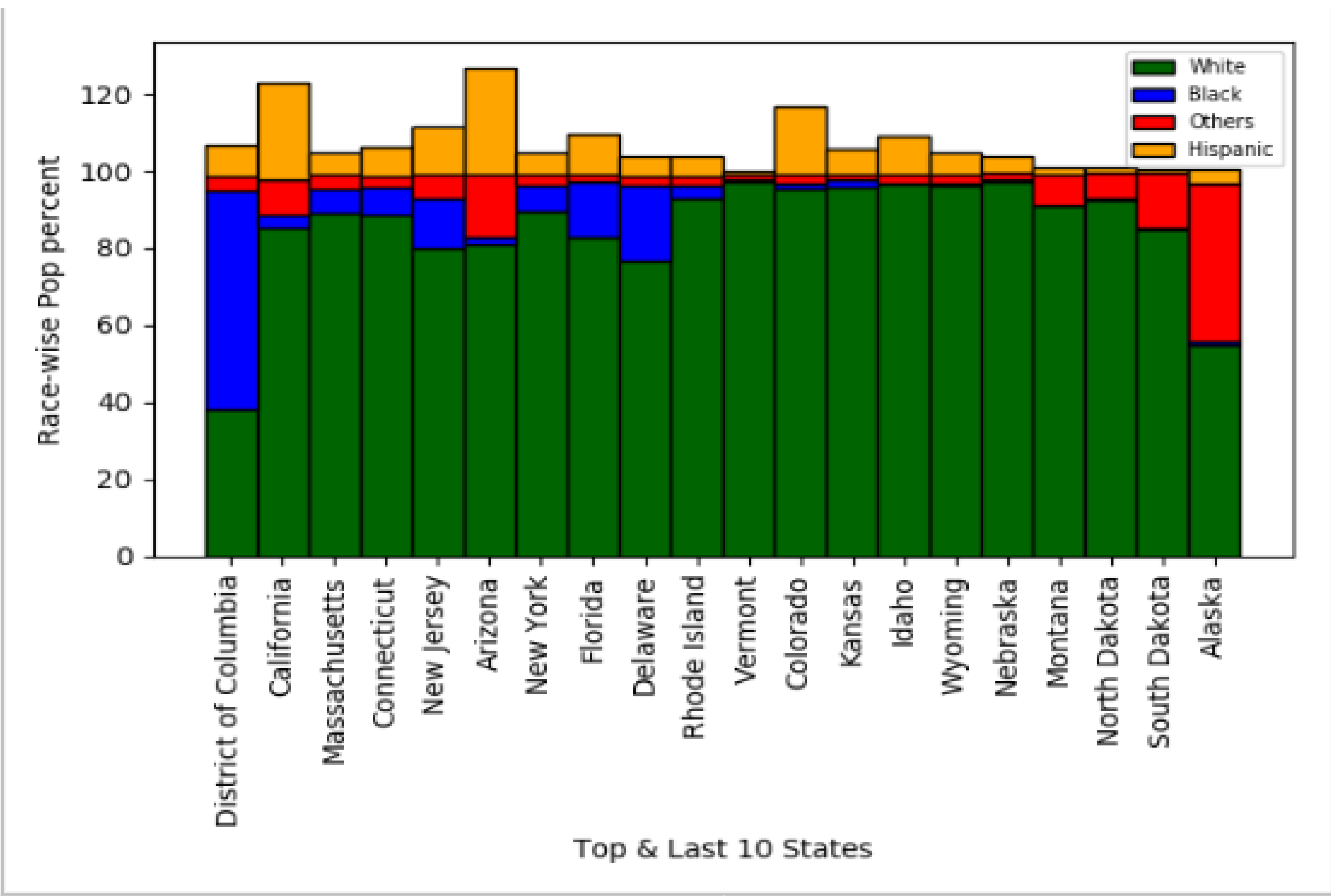
Analysis and Results

	CHSI_State_Name	No_Exercise	Obesity	Poverty	High_Blood_Pres		Smoker	Diabetes	Total_Deaths	Total_Births
0	Mississippi	27.175610	23.491463	19.991463	12.398780		21.408537	9.720732	2070.853659	2981.573171
1	Louisiana	30.493750	24.660937	19.129687	18.023438		22.328125	8.700000	2922.328125	4446.406250
2	New Mexico	23.430303	20.278788	18.884848	16.324242		19.957576	6.993939	1961.484848	3604.242424
3	Arkansas	26.678667	22.648000	17.793333	12.289333		23.740000	8.068000	2290.093333	2790.693333
4	District of Columbia	22.200000	21.300000	17.500000	26.700000		20.400000	7.800000	17375.000000	22742.000000

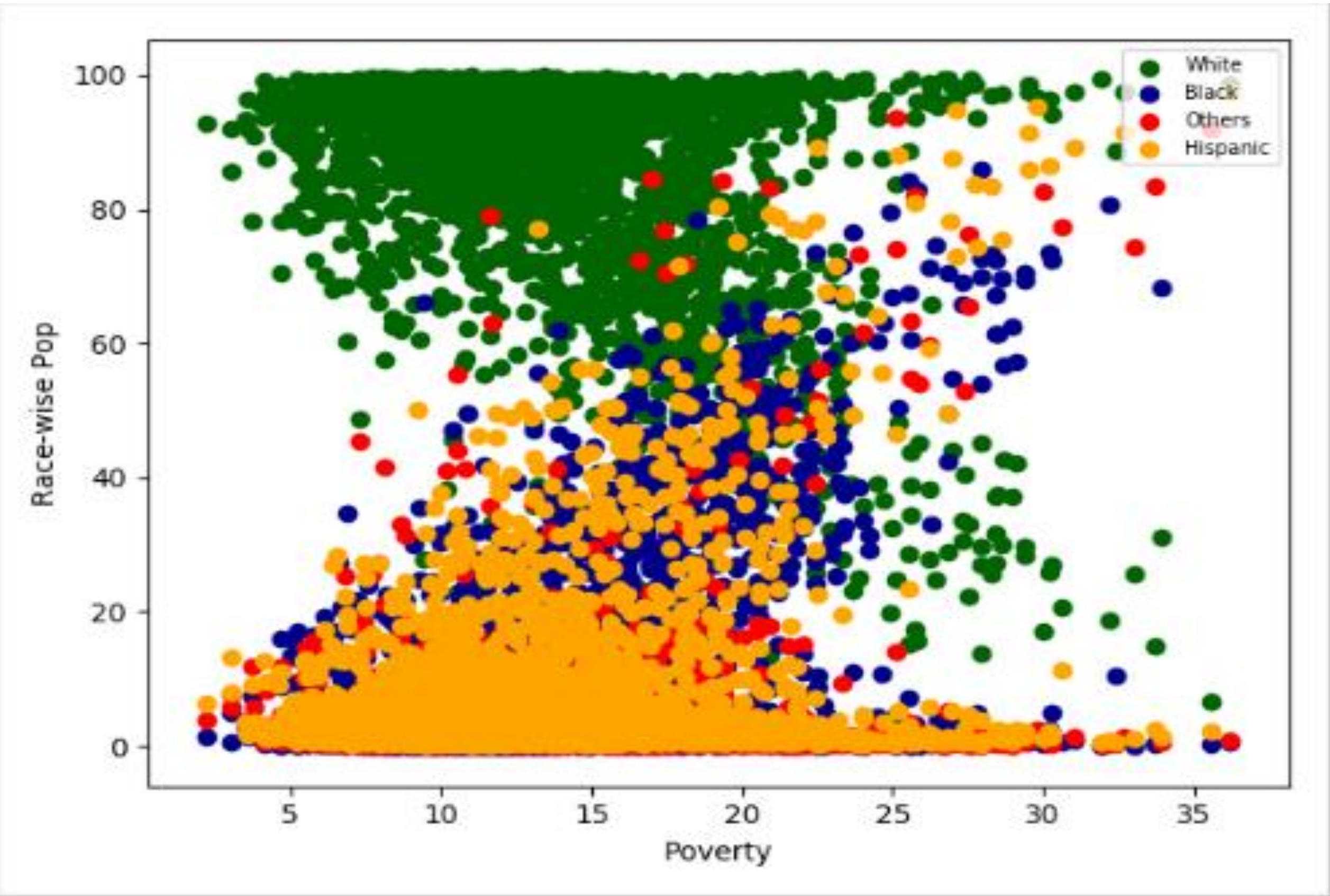|	CHSI_State_Name	No_Exercise	Obesity	Poverty	High_Blood_Pres		Smoker	Diabetes	Total_Deaths	Total_Births
45	Minnesota	12.333333	13.622989	8.527586	9.594253	45	13.713793	5.156322	2153.298851	3384.436782
46	Wisconsin	17.966667	19.098611	8.487500	12.379167	46	18.536111	5.934722	2835.611111	3885.458333
47	Rhode Island	22.080000	17.940000	8.440000	27.820000	47	20.760000	6.240000	6556.800000	8335.600000
48	New Hampshire	20.870000	20.500000	7.190000	24.600000	48	22.680000	6.280000	3565.200000	5003.700000
49	Connecticut	20.937500	19.312500	7.062500	24.850000	49	19.437500	5.987500	11197.000000	15940.250000

• The mortality rate is higher in places where the poverty index is higher and lower in the states where the poverty index is lower.
• Also in addition to this we assumed that the mortality rate depends directly on various risk factors like obesity , diabetes, smoking, etc.
• After the descriptive analysis we found out that this does not hold true in all the cases.
• The mortality rate will not always be higher for the regions with higher poverty index or risk factor. Though it did turn out to be true that mortality rate was lower where these factors were low.

• This graph represents death percentage the top 10 and the last 10 states based on race.

• It is clear that the states with higher death percentage have a higher black and Hispanic population whereas the states with lower death percentage have higher white and other race population.

	CHSI_State_Name	Black	White	Others	Hispanic	Total_Deaths
8	District of Columbia	57.00	38.00	3.50	8.40	17375.0
4	California	3.63	85.09	8.92	25.20	12583.0
21	Massachusetts	6.03	89.23	3.56	6.11	12337.0
6	Connecticut	6.80	88.80	3.05	7.66	11197.0
30	New Jersey	12.85	80.08	5.89	12.67	10869.0
2	Arizona	1.89	81.09	15.95	27.88	8965.0
32	New York	6.56	89.66	2.80	5.78	8317.0
9	Florida	14.37	82.72	1.96	10.47	8121.0
7	Delaware	19.53	76.50	2.63	5.37	7014.0
39	Rhode Island	3.30	92.86	2.60	4.96	6557.0

	CHSI_State_Name	Black	White	Others	Hispanic	Total_Deaths
45	Vermont	0.51	97.31	1.16	0.96	1786.0
5	Colorado	1.46	95.18	2.49	17.45	1701.0
16	Kansas	1.89	95.71	1.56	6.72	1365.0
12	Idaho	0.41	96.49	2.22	9.75	1315.0
50	Wyoming	0.46	96.37	2.31	5.63	1150.0
27	Nebraska	0.49	97.27	1.82	4.13	1065.0
26	Montana	0.21	90.76	8.05	1.74	976.0
34	North Dakota	0.30	92.49	6.73	1.21	903.0
41	South Dakota	0.29	84.97	14.04	1.29	835.0
1	Alaska	1.17	54.49	41.01	3.67	605.0

• This graph represents the poverty of the 4 races county-wise.
• It is clear from the graph that poverty is generally random for all the races in all the counties.
• However, as the population of white people is generally greater than other races population, poverty in white people is slightly greater than the other population.
• According to the analysis based on states, it is clear that poverty is generally higher in states
• with the more white population.

	CHSI_State_Name	White	Black	Poverty	Total_Births	Total_Deaths
24	Mississippi	58.27	40.41	19.99	2981.57	2070.85
18	Louisiana	66.33	31.71	19.13	4446.41	2922.33
31	New Mexico	88.49	1.65	18.88	3604.24	1961.48
3	Arkansas	82.09	15.91	17.79	2790.69	2290.09
8	District of Columbia	38.00	57.00	17.50	22742.00	17375.00
0	Alabama	69.80	28.51	17.09	4014.48	3128.52
48	West Virginia	96.91	1.94	16.98	2114.44	2166.55
17	Kentucky	95.11	3.67	16.85	2574.76	1948.45
2	Arizona	81.09	1.89	16.84	18697.47	8965.07
43	Texas	91.00	6.85	16.79	5327.55	2551.31
10	Georgia	70.22	27.98	16.04	3892.28	2108.89
40	South Carolina	61.21	37.12	15.82	4677.02	3236.85
36	Oklahoma	82.06	3.92	15.57	2948.19	2281.32
45	Vermont	97.31	0.51	9.64	2263.71	1785.57
7	Delaware	76.50	19.53	9.60	11056.00	7014.33
20	Maryland	76.70	19.45	9.33	10010.88	6138.67
14	Indiana	95.98	2.43	9.23	3872.50	2594.72
21	Massachusetts	89.23	6.03	9.13	17513.43	12336.57
15	Iowa	97.67	0.84	9.05	2186.49	1862.42
30	New Jersey	80.08	12.85	8.54	16853.10	10868.57
23	Minnesota	95.24	0.98	8.53	3384.44	2153.30
49	Wisconsin	94.37	1.35	8.49	3885.46	2835.61
39	Rhode Island	92.86	3.30	8.44	8335.60	6556.80
29	New Hampshire	96.96	0.68	7.19	5003.70	3565.20
6	Connecticut	88.80	6.80	7.06	15940.25	11197.00

Conclusion

We have performed the initial analysis of two hypothesis using the dataset and can conclude that:
• It is not always necessary that the mortality rate of a state will be directly dependent on poverty and risk-factor indices.
• Poverty is generally higher in the county where the "White" population is much higher than other diversity groups and vice versa.
• The second hypothesis is in progress and we have not tested it yet.

Dataset Source:

https://catalog.data.gov/dataset/community-health-status-indicators-chsi-to-combat-obesity-heart-disease-and-cancer