

Table 1: Descriptive Statistics for Continuous Variables (At the County Level)

Variable (Per County)	NAs	Mean	Std. Dev.	1st Quartile	Median	3rd Quartile
Average Life Expectancy	0	76.32	2	75	76.5	77.7
Population Density	1	250.07	1703.27	17	44	109.75
Population Size	0	94427.06	306520.4	11220	25270	64111
Percent with Age <19	0	24.81	3.28	22.7	24.6	26.4
Percent with Age 19-64	0	60.28	3.35	58.3	60.3	62.3
Percent with Age 65-84	0	12.79	3.33	10.7	12.5	14.7
Percent with Age 85+	0	2.12	0.95	1.5	1.9	2.6
Percent White	0	87.04	16.14	82.8	94.1	97.6
Percent Black	0	8.99	14.55	0.5	2.1	10.3
Percent Native American	0	1.95	7.62	0.2	0.4	0.9
Percent Asian	0	1.12	2.76	0.3	0.5	1
Percent Hispanic	0	7.02	12.47	1.1	2.3	6.3
Death Rate	3	905.64	131.21	814.25	898.6	989.8
Percent with Health Status	662	17.32	6.09	12.9	16.4	20.9
Number of Monthly Unhealthy Days	0	0.019	0.36	0.021	0.025	0.03
Percent not Exercising	933	26.51	6.7	21.9	26	30.8
Percent with Few Fruits or Vegetables	1235	78.92	5.16	75.5	79	82.4
Percent Obese	915	24.15	4.9	21.1	24.3	27.2
Percent with High Blood Pressure	1617	26.48	5.44	22.8	26.2	29.9
Percent Smoking	872	23.11	5.73	19.4	23	26.7
Percent with Diabetes	420	7.81	2.76	5.9	7.5	9.45
Physician Rate	0	57.6	44.78	30.55	50.6	74.7
Dentist Rate	1	32.19	21.5	18.7	30	43.3
Percent Poverty	1	13.35	4.88	9.8	12.6	16.2
Percent Unemployed	543	6.11	1.34	5.2	6	6.8
Percent Uninsured	0	0.14	0.36	0.11	0.13	0.17
Percent Medicare Based on Disability	0	0.017	0.79	0.017	0.023	0.031
Percent Medicare Based on Elderly Status	0	0.098	0.8	0.11	0.14	0.17
Percent with Major Depression	0	0.061	0.0066	0.056	0.059	0.065
Percent with No High School Diploma	0	0.15	0.058	0.11	0.14	0.19
Percent with Recent Drug Use	0	0.052	0.011	0.045	0.051	0.057
Percent Severely Disabled	0	0.025	0.062	0.02	0.027	0.034

Table 2: Model Comparison by Cross-validation RMSE

Model	Mean RMSE	Variance RMSE	1st Quartile	Median RMSE	3rd Quartile
Least Squares	1.3060	1.7205	0.8451	0.8839	0.9840
Ridge	1.2115	0.9278	0.8545	0.9064	0.9938
GAM	1.1308	0.1676	0.9819	1.0259	1.0550
Regression Tree	1.0179	0.0166	0.9609	1.0091	1.0925
PLS	1.0167	0.1299	0.8595	0.8855	1.0013
KNN	0.9863	0.0088	0.9094	1.0191	1.0444
LASSO	0.9033	0.0099	0.8668	0.8993	0.9601
PCR	0.9018	0.0103	0.8485	0.8789	0.9788
MARS	0.8783	0.0115	0.8053	0.8676	0.9717
Bagging	0.8594	0.0177	0.7742	0.8874	0.9274
Boosting	0.8329	0.0153	0.7300	0.8510	0.9340
Random Forest	0.8309	0.0142	0.7490	0.8502	0.8946

Figure 1: Correlation Plot

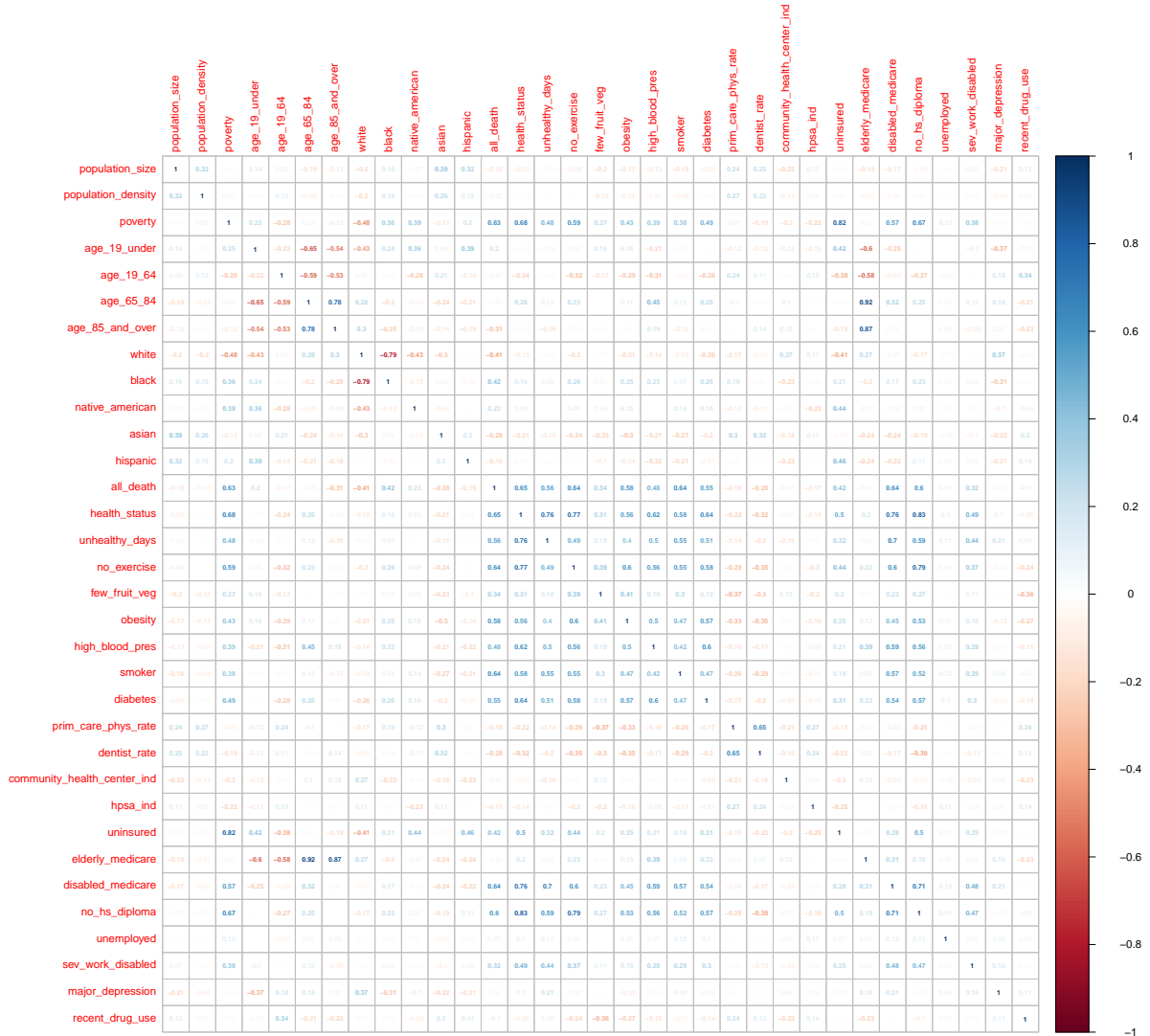


Figure 2: Hierarchical Clustering at the State Level

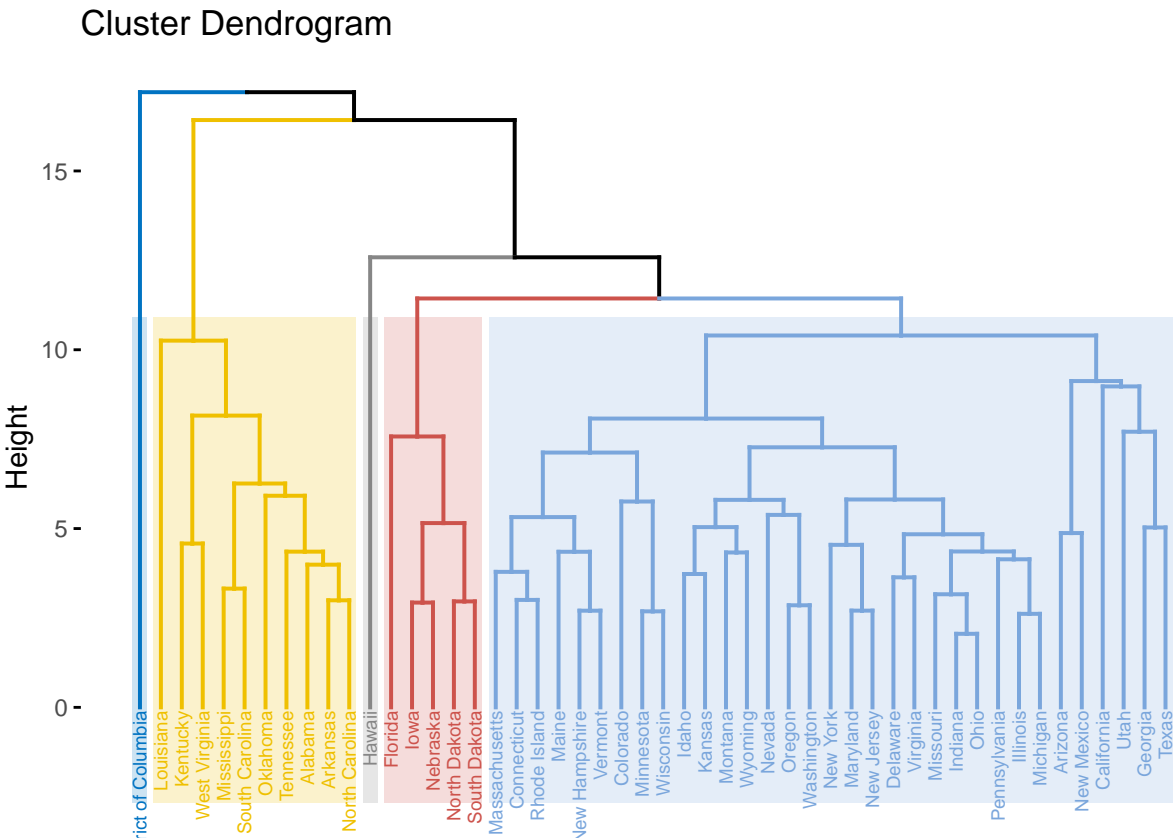


Figure 3: Heatmap at the State Level

