

The Shortage of Rural Hospitals in Tennessee?

Data acquisition and cleaning

- The list of Tennessee hospitals came from downloading a csv file provided by the Tennessee Department of Safety and Homeland Security.
- Tennessee county demographic information was scraped from en.wikipedia.org/wiki/List_of_Tennessee_locations_by_per_capita_income and then merged with the hospital list based upon with a left join to the hospital list. Demographic data that represented dollar amounts was imported as text, so the \$ was removed, and the column type set to float.
- The determination of whether a county is rural or not is based upon the US Department of Agriculture Rural-Urban Continuum Code.
- There are 95 counties in Tennessee.
- 25 Counties in Tennessee do not have a hospital.

County data

Without a hospital

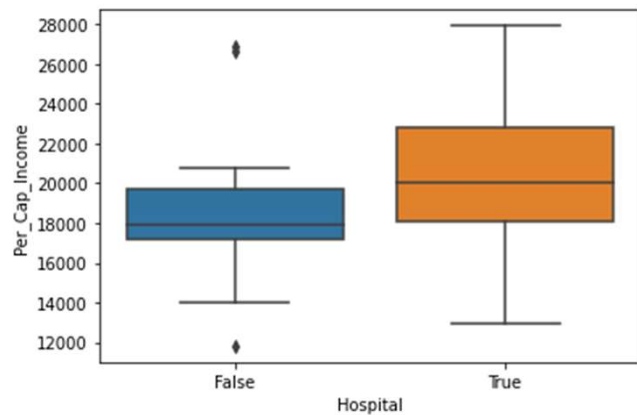
	Per_Cap_Income	Median_House_Income	Median_Family_Income	Population	RUCC
count	25.000000	25.000000	25.000000	25.000000	25.000000
mean	18409.840000	34402.200000	43347.360000	17679.960000	6.240000
std	3217.424911	6515.983809	6780.212576	10580.052286	2.454248
min	11813.000000	24700.000000	32394.000000	5077.000000	3.000000
25%	17160.000000	30193.000000	39767.000000	11753.000000	3.000000
50%	17883.000000	33506.000000	41948.000000	16825.000000	6.000000
75%	19742.000000	36772.000000	46670.000000	21987.000000	9.000000
max	26898.000000	56729.000000	63186.000000	52266.000000	9.000000

With a hospital

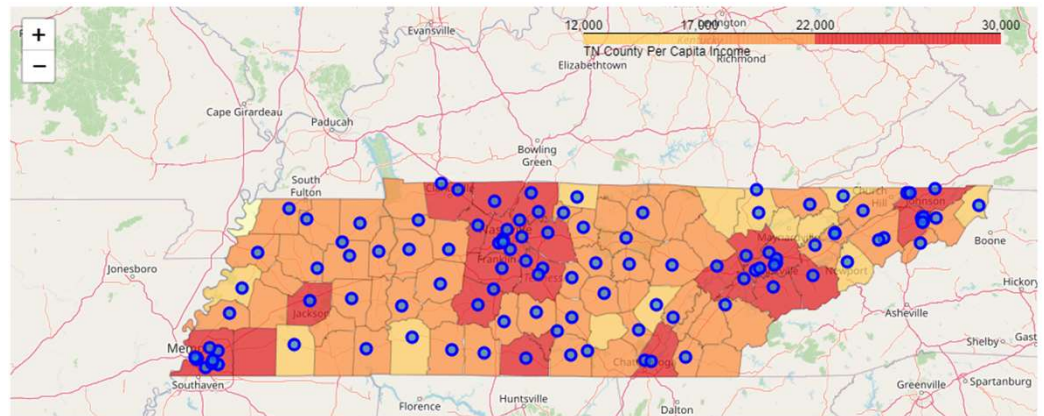
	Per_Cap_Income	Median_House_Income	Median_Family_Income	Population	RUCC
count	70.000000	70.000000	70.000000	70.000000	70.000000
mean	20505.614286	39950.514286	48779.357143	84344.371429	4.685714
std	3401.777597	9043.666955	9988.577433	142453.976415	1.757327
min	12907.000000	23125.000000	31495.000000	6819.000000	3.000000
25%	18088.000000	34357.000000	42732.500000	26332.750000	3.000000
50%	20030.500000	38641.500000	47840.000000	40884.000000	4.000000
75%	22772.750000	43953.750000	52913.750000	71448.500000	6.000000
max	27925.000000	87832.000000	100407.000000	927644.000000	8.000000

Per Capita Income

- State mean \$19,954
- No hospital mean: \$18,410
- With hospital: \$20,506

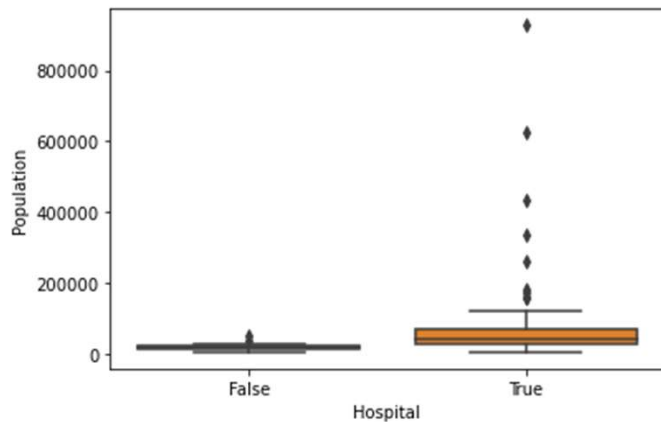


- County map with blue dots denoting hospitals

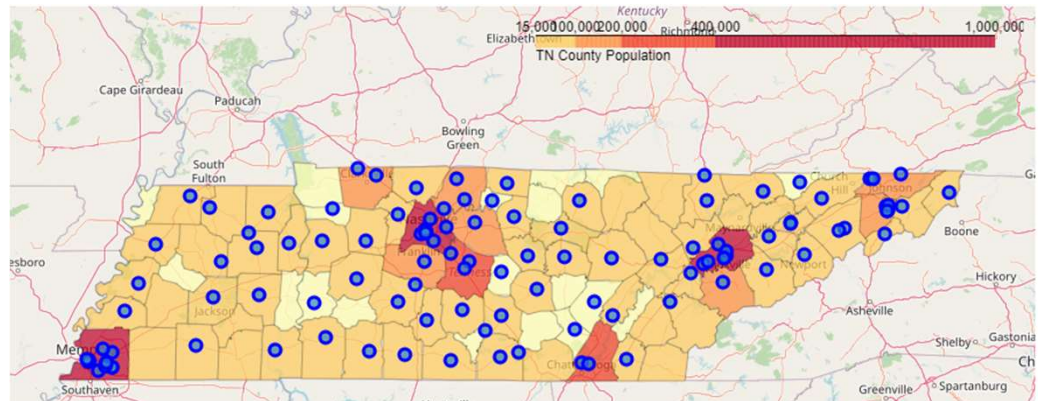


Population

- State mean 66,801
- No hospital mean: 17,680
- With hospital: 84,344

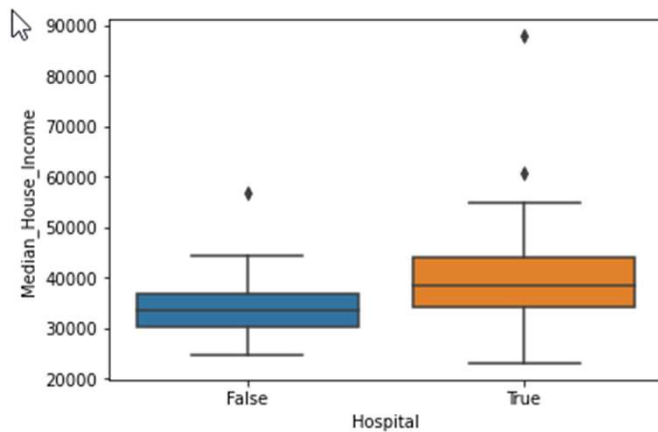


- County map with blue dots denoting hospitals

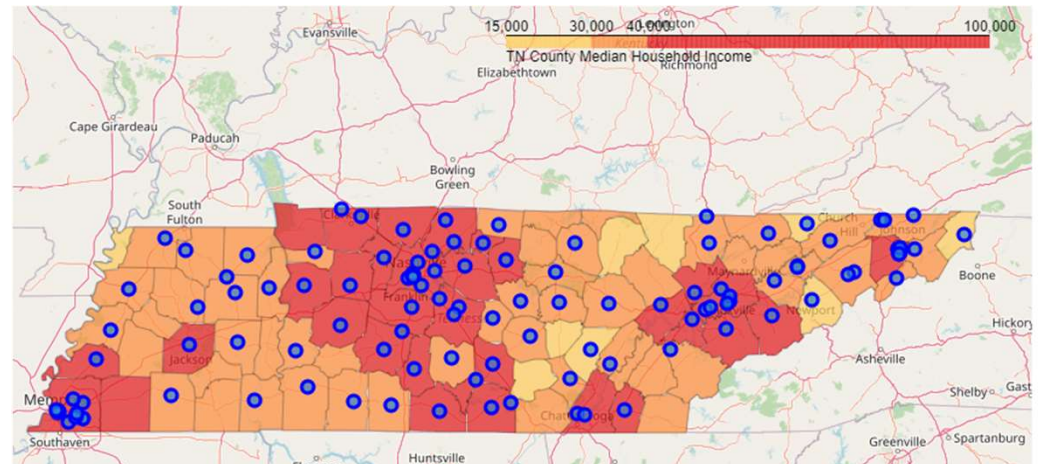


Median Household Income

- State mean \$38,490
- No hospital mean: \$34,402
- With hospital: \$39,950

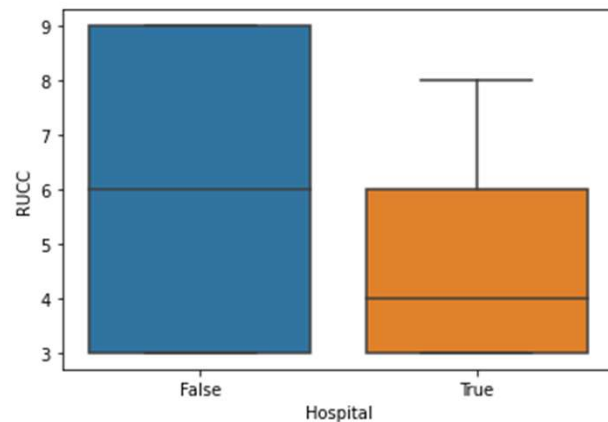


- County map with blue dots denoting hospitals

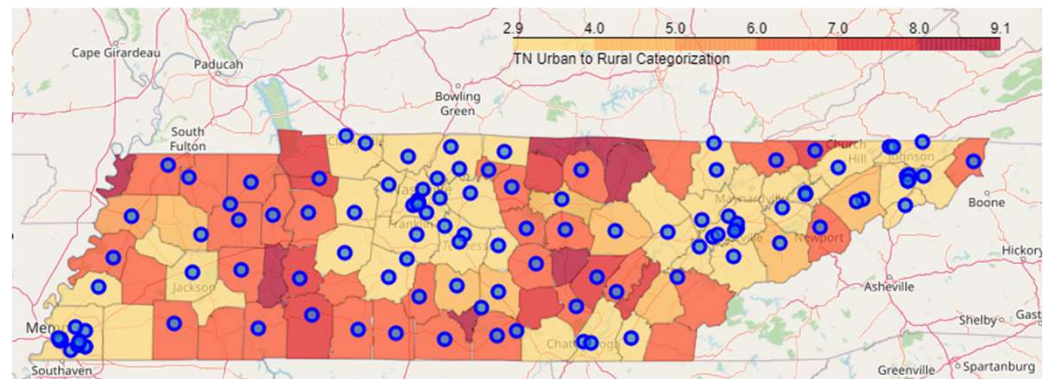


Median Household Income

- State mean: 5
- No hospital mean: 6.24
- With hospital: 4.69

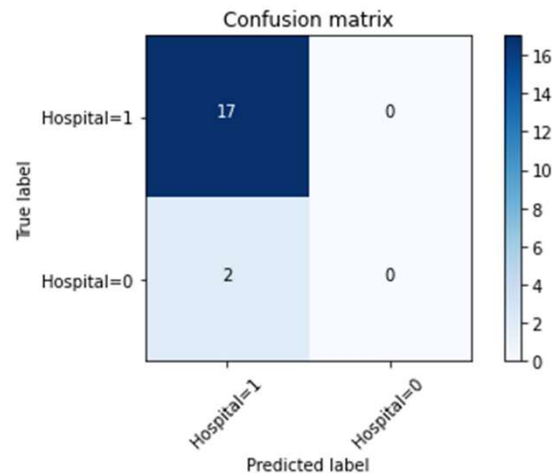


- County map with blue dots denoting hospitals



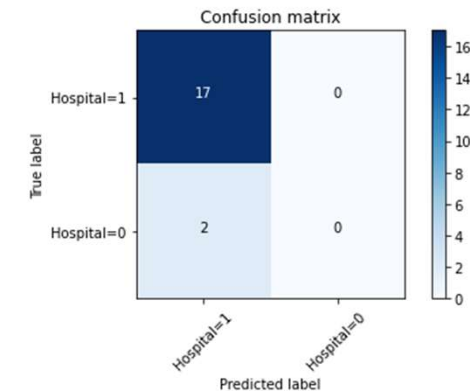
Logistic Regression vs SVC using RUCC

	precision	recall	f1-score	support
0	0.00	0.00	0.00	2
1	0.89	1.00	0.94	17
accuracy			0.89	19
macro avg	0.45	0.50	0.47	19
weighted avg	0.80	0.89	0.85	19



	precision	recall	f1-score	support
0	0.00	0.00	0.00	2
1	0.89	1.00	0.94	17
accuracy			0.89	19
macro avg	0.45	0.50	0.47	19
weighted avg	0.80	0.89	0.85	19

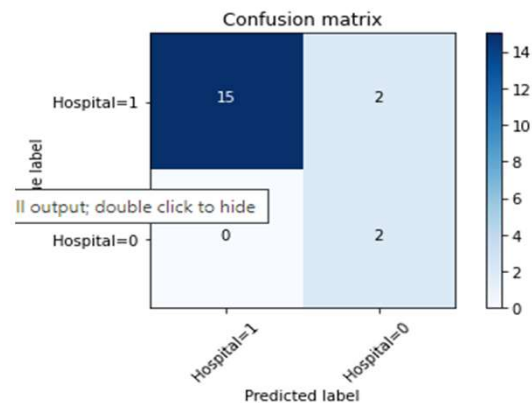
Confusion matrix, without normalization
[[17 0]
[2 0]]



Using the RUCC score only, both Logistic Regression and SVC models give strong results predicting counties with hospitals, but neither model gives good results in determining that a county does not have a hospital

Logistic Regression vs SVC using RUCC

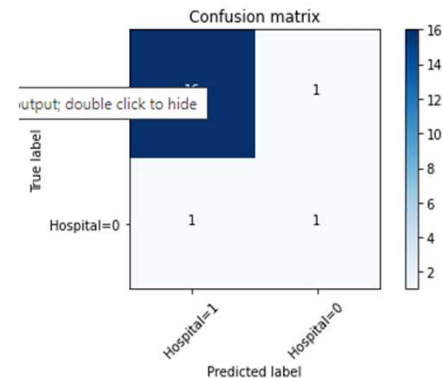
	precision	recall	f1-score	support
0	0.50	1.00	0.67	2
1	1.00	0.88	0.94	17
accuracy			0.89	19
macro avg	0.75	0.94	0.80	19
weighted avg	0.95	0.89	0.91	19



	precision	recall	f1-score	support
0	0.50	0.50	0.50	2
1	0.94	0.94	0.94	17
accuracy			0.89	19
macro avg	0.72	0.72	0.72	19
weighted avg	0.89	0.89	0.89	19

Confusion matrix, without normalization

```
[[16  1]
 [ 1  1]]
```



Result comparison of models

Weighted Avg	Precision	Recall	F1-score	Support
RUCC Logistic Reg.	0.80	0.89	0.85	19
RUCC SVM	0.80	0.89	0.85	19
All Logistic Reg.	0.95	0.89	0.91	19
All SVM	0.89	0.89	0.89	19

Using the RUCC score and demographic information, both Logistic Regression and SVC models give strong results predicting counties with hospitals, but neither model gives good results in determining that a county does not have a hospital. Logistic regression tends to have a better result with an F1-score of 0.91

The Shortage of Hospitals in Tennessee

Results

If you live in a Tennessee county that is classified as rural and not adjacent to a metropolitan area (RUCC=9), you do not have a hospital. Although a RUCC score of 9 is currently an indicator of no hospital, lower scores alone become problematic in predicting the presence of a hospital. Using the RUCC score along with population and income data allows us to better predict the absence of a hospital.

Conclusion

- Sparse populations and the lack of available capital together seem to be the best indicator of not having a hospital. Given a small hospital costs \$52,200,000 to build, the economic reality is that it takes a certain amount of people with sufficient income to make a hospital economically viable.
- There are 3006 counties in the U.S. so this study only took into account 3% of the counties available. A larger sample size might allow us to have a better understanding of what causes hospital deserts, and allow community leaders to act decisively on behalf of their citizens.