US COVID-19 Hospitalizations Analysis

Lindsey Roeder, Steve Riehl, Dennis Stoliaryk
May 21, 2021

Problem:

Can we predict which COVID-19 probable cases will be admitted to the hospital?

Model Construction Timeline

- Data
- Feature selection
- Fitting a model
- Evaluating the model

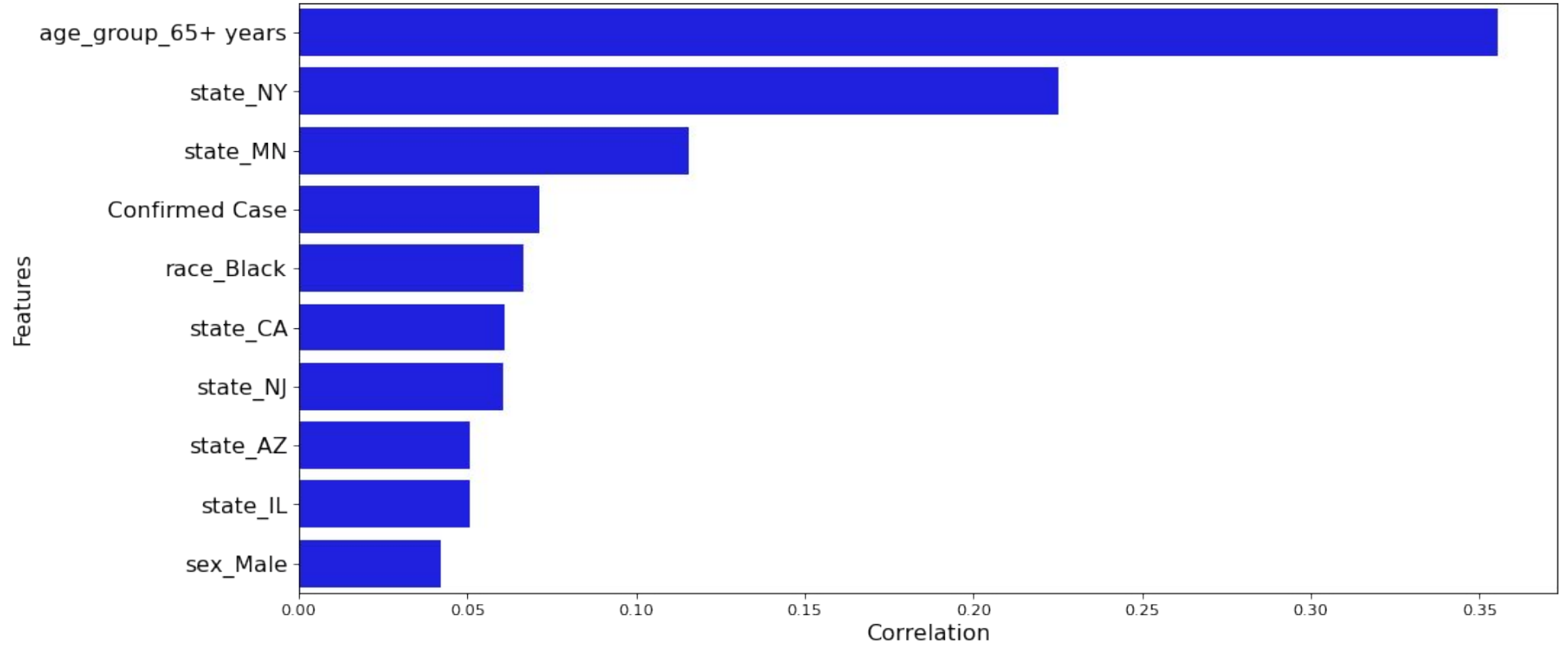
Data

Data Collection & Preparation

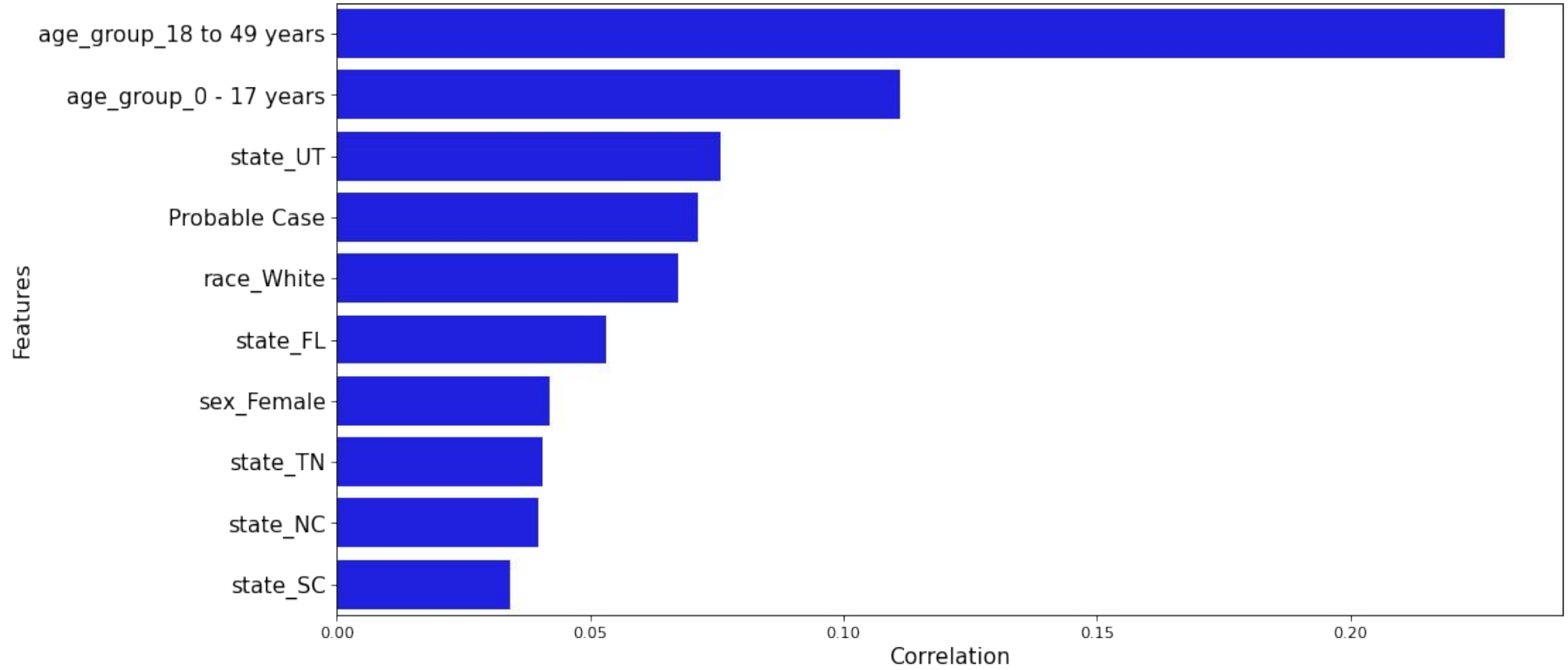
Source	Centers for Disease Control and Prevention https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4/data	10,000,000 Rows
Date Range	January 2020 – April 2021	
Data Cleansing	Features with over **80%** null values removed.	3,000,000 Rows ---------- 70 Columns
	Null values in remaining rows dropped	
	Additional features engineered	

Feature Selection Process

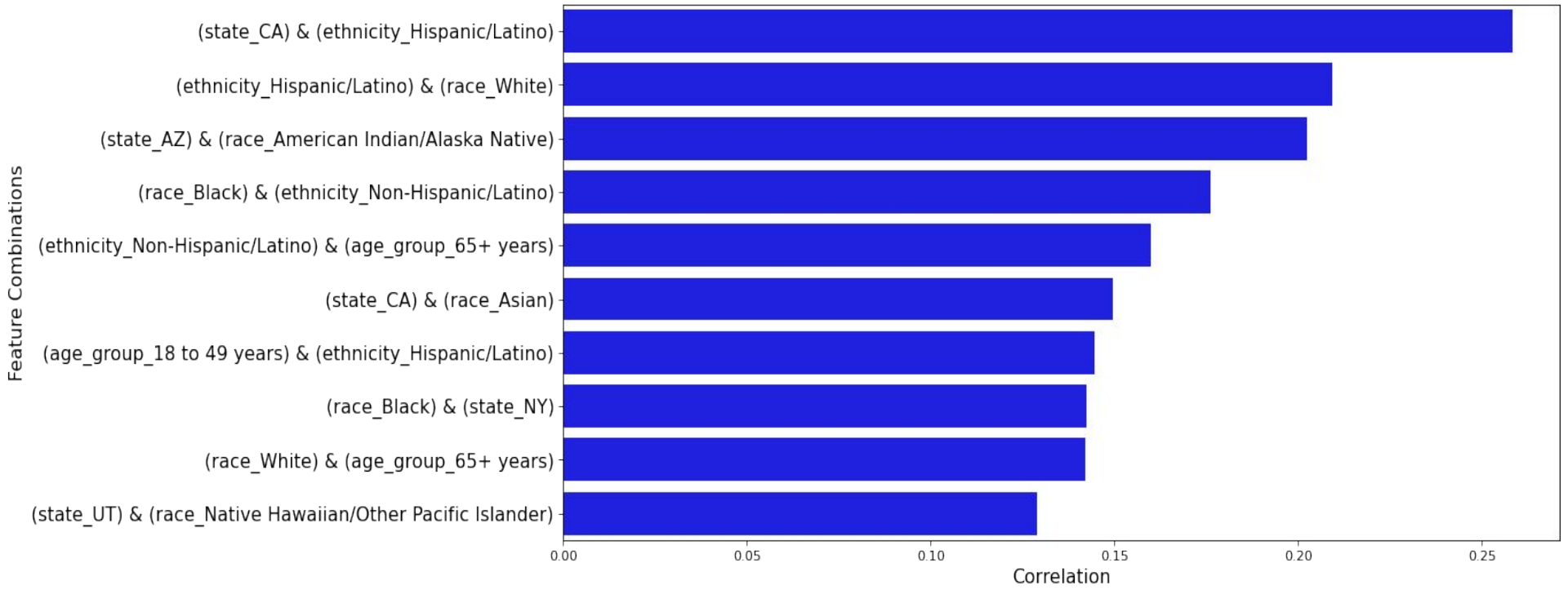
Features Highly Correlated with Hospitalizations



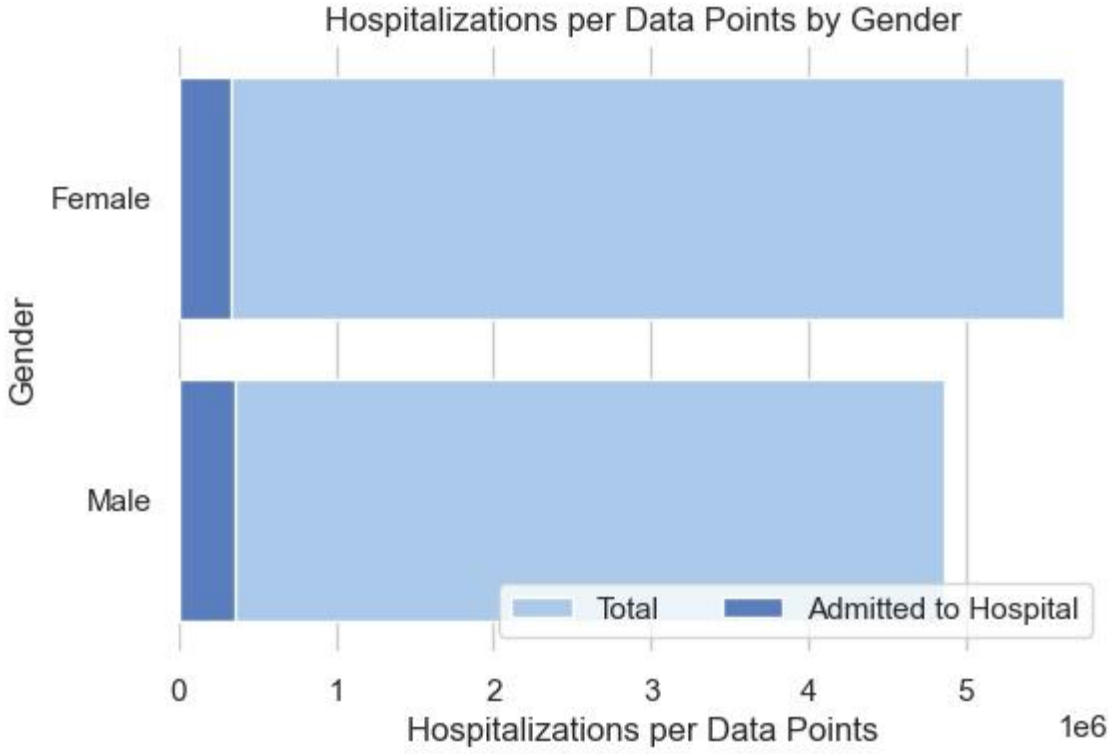
Features Least Correlated with Hospitalizations



Feature Combinations Highly Correlated with Hospitalizations



Hospitalizations by Gender

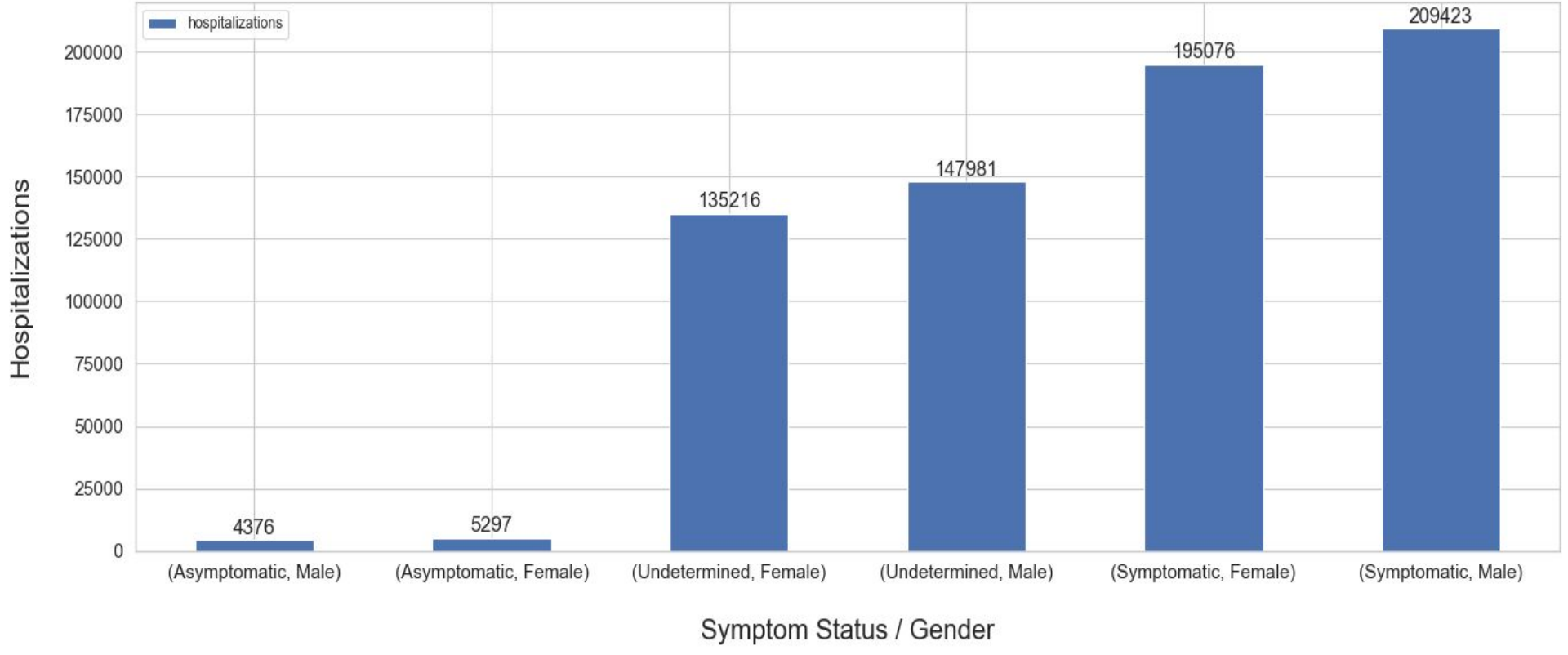


% Hospitalized

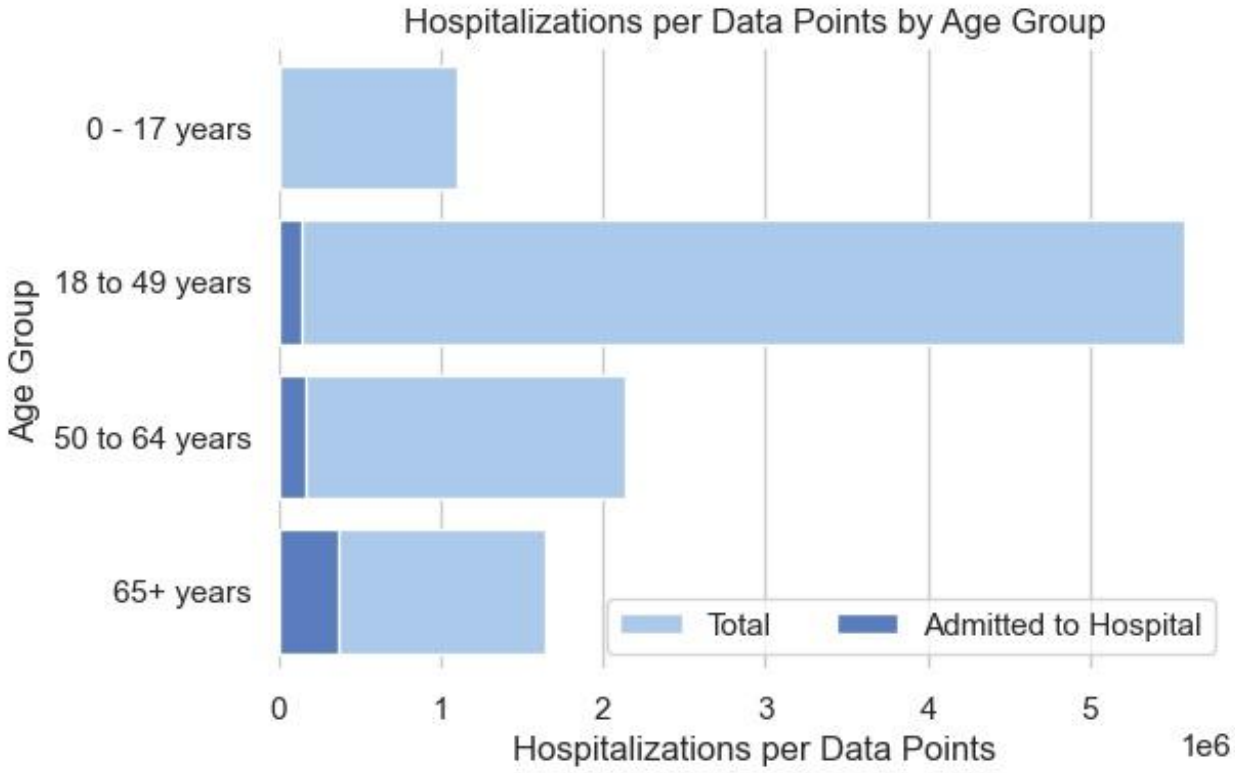
8%

23%

Hospitalizations by Symptom Status



Hospitalizations by Age Group



% Hospitalized

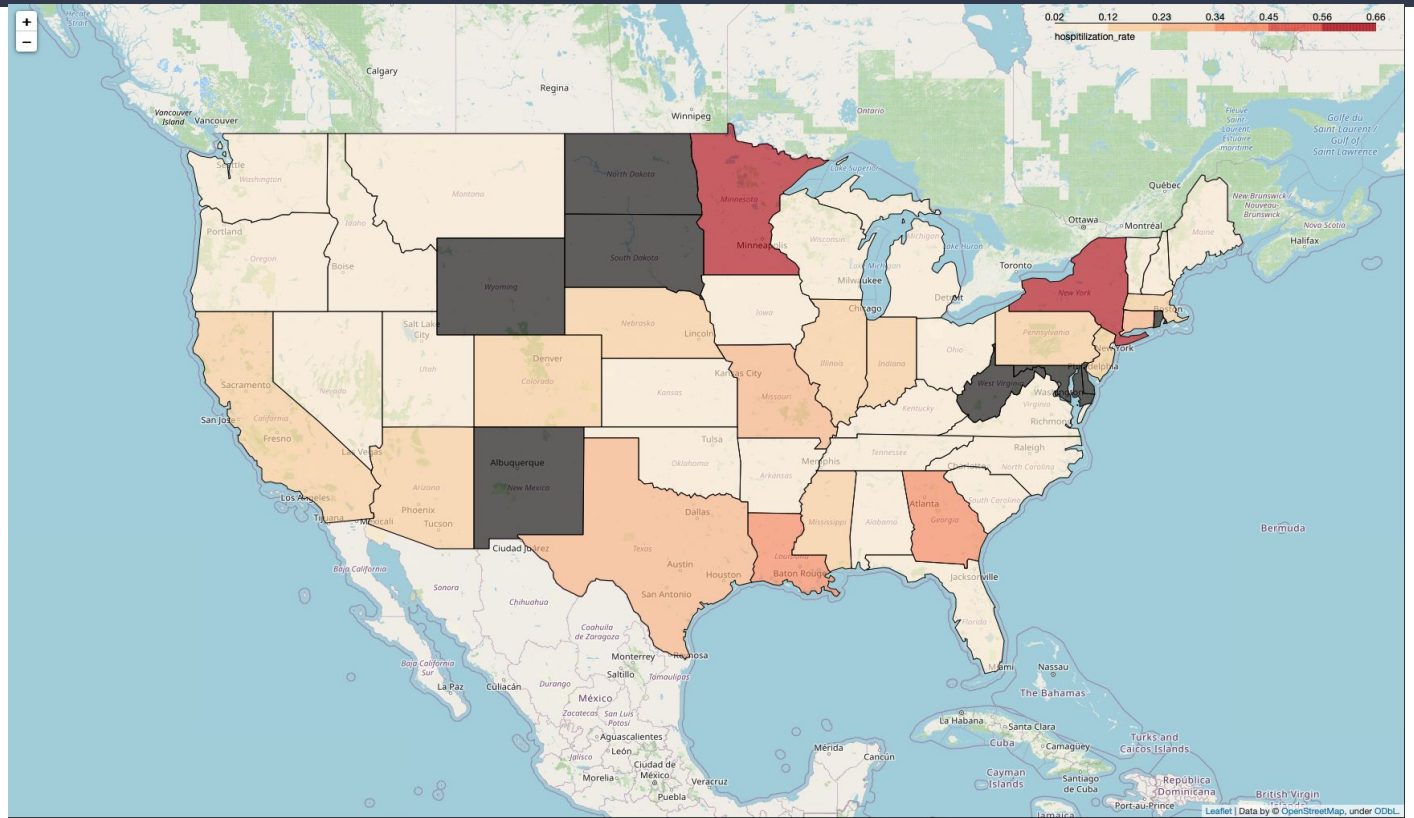
1%

3%

8%

23%

Hospitalization Rates of Mainland United States



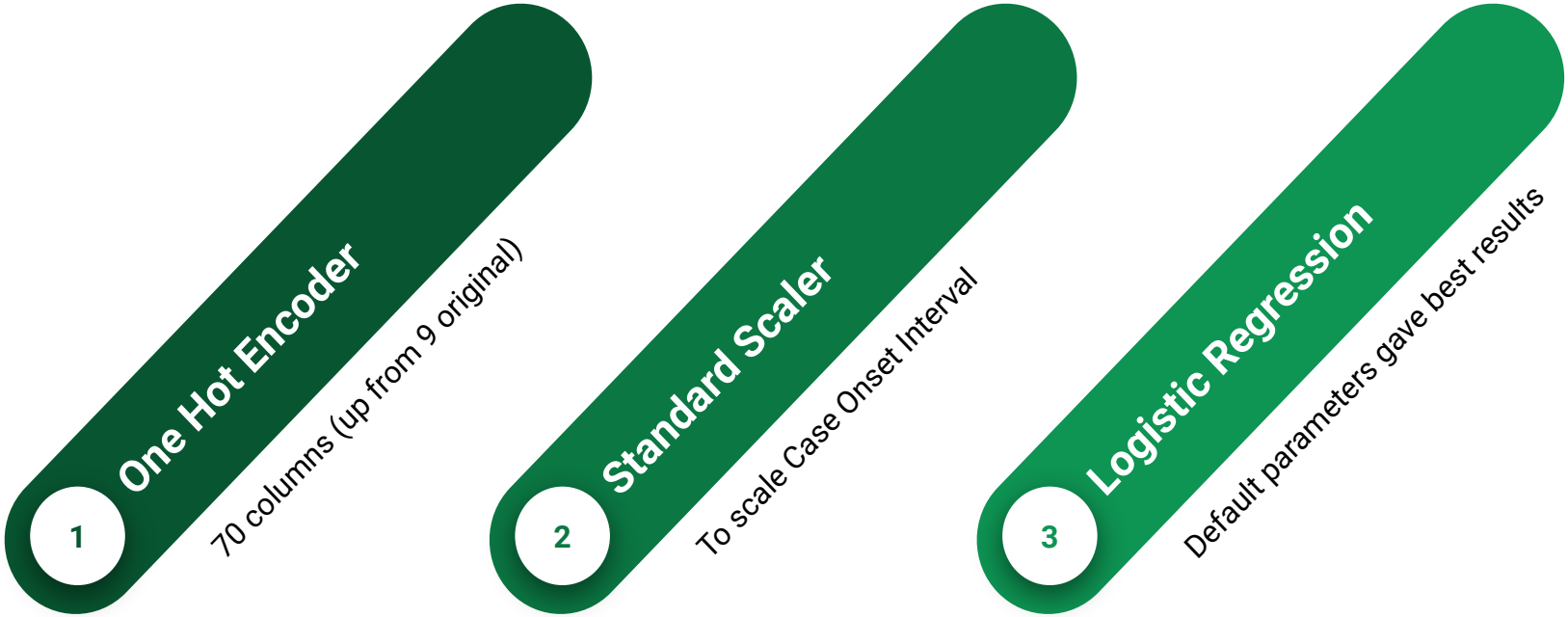
Fitting the Model

Model Selection Criteria:

Accuracy
Low False Negatives
Processing Time

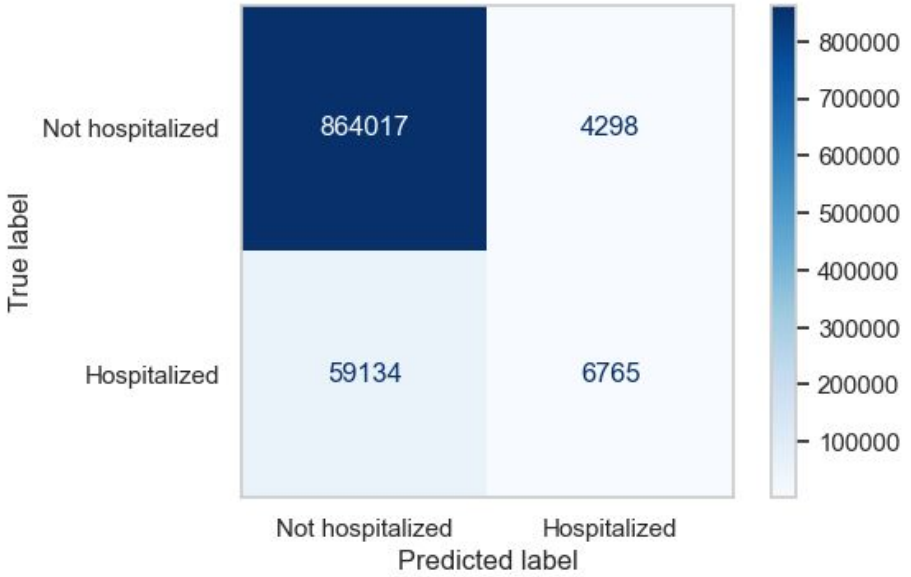
Estimator	Accuracy	FN
Sample Baseline	**92.6%**	**185**
Logistic Regression	93.2%	157
AdaBooster	93.2%	157
Simple NN	93.0%	167
Complex NN	92.6%	185
Random Forest	92.4%	154
KNN	92.2%	170
Bagging	91.4%	148
Decision Tree	91.2%	156

Final Model Selection



Evaluating the Model

Evaluating the Model



	precision	recall	f1-score	support
0	0.94	1.00	0.96	868315.00
1	0.61	0.10	0.18	65899.00
accuracy	0.93	0.93	0.93	0.93
macro avg	0.77	0.55	0.57	934214.00
weighted avg	0.91	0.93	0.91	934214.00

Baseline: 92.6% Accuracy & 65,899 False Negatives

Conclusion & Further Research

Further Research

- Bring in more features
- Conduct Analysis again a year from now
- Compare US hospitalization rates to that of other countries/regions