

Data from the United States Census Bureau discussed below focuses on biographical information, including but not limited to, overall population count, different age groups, sex, and race of different subsets of the population dataset, per state.

Name	Definition	Data type	Possible values	Required?
StateFIPS	Identification state code	Integer	1, 2, 3, 4, ..., 55	Yes
State	Name of the state	String	AL, CA, NC	Yes
Age_Groups	Subsets of Population that falls under a specific age group as a percentage	Float	0.1, 5.3, 16.8, ..., 100.0	Yes
Sex	Subset of Population that falls under a specific sex, 'T' stands for total	String	'M', 'F', 'T'	Yes

This dataset can be merged with the primary COVID-19 dataset using the StateFIPS or by a simple data wrangling process, where a derived variable can be generated from the State Name and converted to the unique state identifier.

An approach that can be addressed while diving into the comprehensive dataset would be using the age groups in different states and analyse the actual numbers in those different age groups, moreover, a great conclusion would be comparing death rates among the different age groups, with the hypothesis of senior citizens are the most vulnerable.

Another aspect worth exploring is if there are any correlations between sex and infection rates, and death rates.