SAIPAVAN TADIKONDA

TASK-2

Census Demographic ACS:

The Census Demographic ACS dataset contains demographic information of all states of the United States by county level from the United States Census Bureau site for the year 2019. The demographic information includes population estimations based on sex, age (different age groups), races (different races), eligible voters based on sex.

Variable dictionary:

Column Variables	Data types	Variable description
countyFIPS	int64	Unique five-digit code for every county
County Name	object	Name of the county
State	object	Name of the State
Total population	object	Estimation of total population
Males	object	Estimation of total male population
Females	object	Estimation of total female population
Sex ratio (males per 100 females)	object	Estimated ratio of males per 100 females
Under 5years	object	Total population estimate of ages under 5years
5-9years	object	Total population estimate of ages from 5-9 years
10-14years	object	Total population estimate of ages from 10-14years
15-19years	object	Total population estimate of ages from 15-19years
20-24years	object	Total population estimate of ages from 20-24years
25-34years	object	Total population estimate of ages from 25-34years
35-44years	object	Total population estimate of ages from 35-44years
45-54years	object	Total population estimate of ages from 45-54years
55-59years	object	Total population estimate of ages from 55-59years
60-64years	object	Total population estimate of ages from 60-64years

65-74years	object	Total population estimate of ages from 65-74years
75-84years	object	Total population estimate of ages from 75-84years
85years and over	object	Total population estimate of ages 85years and over
Median age(years)	object	Median age (years)
Under 18years	object	Total population estimate of age under 18years
18years and over	object	Total population estimate of age 18years and over
18years and over Male	object	Total population estimate of age 18years and over (Male)
18years and over Female	object	Total population estimate of age 18years and over (Female)
Sex ratio (18years and over)	object	Estimated Sex ratio of 18years and over
White population	object	Total population estimate of White population
Black population	object	Total population estimate of Black population
Asian population	object	Total population estimate of Asian population
Voting population	object	Estimate of Voting population
Voting population (Male)	object	Estimate of Voting population (Male)
Voting population (Female)	object	Estimate of Voting population (Female)

Merge with Covid Dataset:

The census demographic dataset has around 358 columns. I considered the columns of population estimations based on sex, age groups, different races to merge with covid dataset which would be further helpful in performing analysis related to covid spread.

I observed GEO_ID and Name column variables in the enrichment dataset. The GEO_ID has the unique five-digit countyFIPS code and the Name column has both the County Name and the State name. These two columns in enrichment dataset are in common with the covid dataset. So, I extracted the countyFIPS code from GEO_ID using slice and I obtained the County Name and State by using split. Finally, using the countyFIPS, County Name, State I merged both enrichment dataset and covid dataset.

Importance of Census Demographic enrichment data in Covid analysis:

This enrichment data helps us to know which age groups are mostly affected across different counties. Also, which gender got mostly affected with the covid spread. It also helps us to know which races were affected with this covid spread. Some awareness can be shared among people through government, private organizations and NGOs with this analysis.

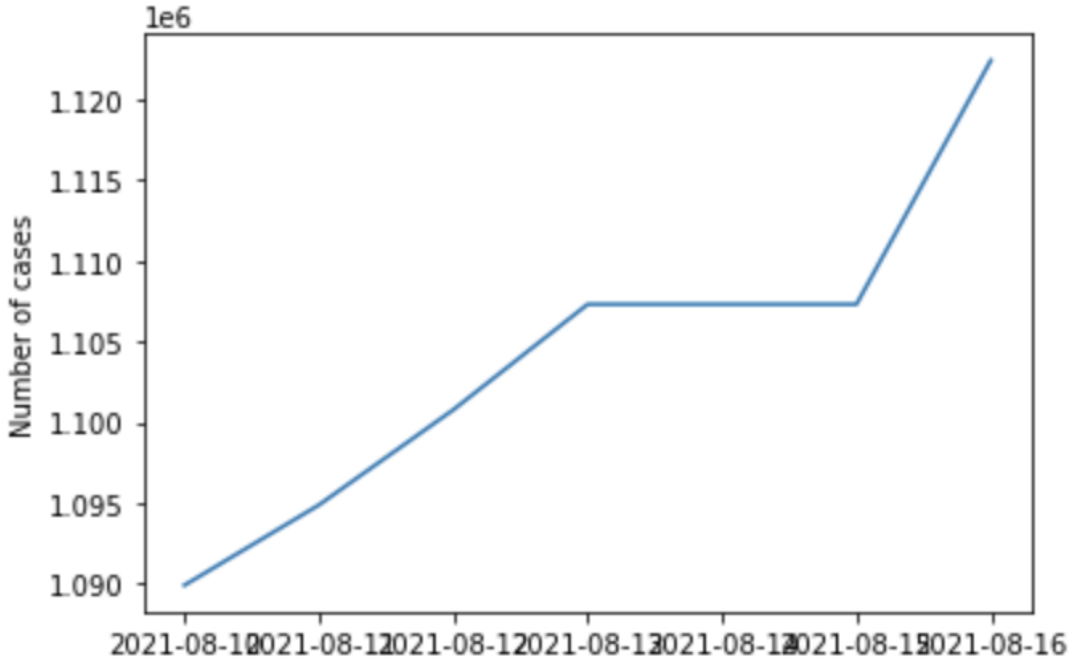
INITIAL HYPOTHESIS QUESTIONS:

1) Which age group was badly affected by this virus?
2) Which race was mostly affected with this virus?
3) Which gender was more prone to the virus?
4) Is there more death rate among the higher age groups?
5) Were there high cases among the children under 5 years?
6) Which county was mostly affected with this virus?
7) Which state has a smaller number of cases compared to other states?
8) Does the chance of getting infected increases with the increase in age?

TASK-3

NC state Last week covid cases trends:

I chose the North Carolina State for observing the last week trends of the covid cases i.e., from 10th August 2021 to 16th August 2021.



From the above plot we can say that the cases increased initially from 10th August 2021 to 12th August 2021 and cases remained stable for the next three days i.e., from 13th August 2021 to 15th August 2021. But on the last day i.e., on 16th August 2021 cases increased.