

PROJECT

STAGE – 1

By

Nikhil Bolisetty

ENRICHMENT DATASET

- ACS SELECTED SOCIAL CHARACTERISTICS

DATA DESCRIPTION:

The data is taken from the United Census Bureau. This data is about the social characteristics of the American households living in the country. The data contains information about Population, Average household size, Total number of high school Enrollments, Households with internet Subscriptions etc.,

S.NO	COLUMN NAME	DATA TYPE	Data DESCRIPTION
1	GEO_ID	Object	Geography id and the last fivedigits contains the FIPS code
2	DP02_0001E	Int64	Total Households living in a family
3	DP02_0002E	Int64	Total number of Married-couple family
4	DP02_0017E	Float64	Average family Size
5	DP02_0153E	Int64	Households with Internet subscription

6	DP02_0089E	Int64	People born in US
7	DP02_0066E	Int64	Population with Higher Education
8	DP02_0057E	Int64	Number of Enrollments in High School
9	DP02_0037E	Int64	People with Age of 15-50 Given Birth Last Year
10	DP02_0015E	Int64	Households with Senior living with them

How to Merge:

To merge the social characteristics Enrichment data with the covid 19 we need extract the county code which is present in the GEO_ID in social characteristics data and perform the join operation between both data sets on the same county code column.

Importance of Enrichment Data in analysis of covid 19:

The social characteristics data contains the information about the Population count, Internet subscriptions, high school enrollments. We can analyze how these got affected by covid.

Hypothesis Questions:

- What is the impact on Internet Subscriptions?
- Effect on the Population count due to covid 19?
- Increase or decrease in number of high school enrollments?

