

TASK - 2

Aman Tej Vidapu

ACS Economic Survey Data:

This enrichment dataset contains the economic information of every state by county level from the American Community Survey for the year 2019. The Information includes Modes of transportation used by people, Occupation details, Health insurance information, Households and their income and Per capita income.

Variable Dictionary:

~Mentioned datatypes are Initial datatypes.

Column Variable	Datatype	Variable description
GEO_ID/countyFIPS	Object/Int 64	Geography id and the last five digits contains the fips code
County Name	Object	Name of county
State	Object	Name of State
Total people commuting to work	Object	Total Number of people commuting to work
Driving alone	Object	Total Number of people driving to work alone
Carpooled	Object	Total Number of people sharing rides while going to work
Only public transports	Object	Total Number of people commuting to work by public transport (other than taxi)
Walked	Object	Total Number of people commuting to work by walking
Other means	Object	Total Number of people commuting to work by other means
Work from home	Object	Total Number of people working from home
Mean travel time to work	Object	Mean time taken to travel
Total people has occupation	Object	Total Number of people having an occupation

Management&business&science&arts occupations	Object	Total Number of people having an occupation in Management, business, science and arts fields
Service occupations	Object	Total Number of people having an occupation in Service field
Sales and office occupations	Object	Total Number of people having an occupation in Sales and office fields
Natural resources, construction, and maintenance occupations	Object	Total Number of people having an occupation in Natural resources, construction, and maintenance fields
Production, transportation, and material moving occupations	Object	Total Number of people having an occupation in Production, transportation, and material moving fields
Private health insurance	Object	Number of people having private health insurance
Public health insurance	Object	Number of employed people having public health insurance
No health insurance object	Object	Number of employed people having no health insurance
Total households	Object	Total number of households in the county
Mean household income	Object	Mean household income of the county
Per capita income	Object	Per capita income for each county

How to Merge:

This particular enrichment data have more than hundreds of columns. So, we have to read and understand the data before merging with the super covid dataset. I considered estimates of commuting to work, health insurance, occupations, household and per capita income-related columns to merge with the super covid dataset to perform further analysis. Also, thought about what hypothesis can be made by keeping these particular columns.

To merge, I used (County Name, State, countyFIPS) as common columns in both datasets. In enrichment data GEO_ID has the last five digits as county fips, has to extract there and NAME has both county name and state, has to split one column to two required columns. So that we can perform the merge.

Importance of Economic enrichment data in Covid analysis:

By using the household income, modes of transport, insurance status, occupation details combined with the super covid dataset, we can understand which economic state people are more prone to covid and who are the reason for the spread of the virus. So that, moderate rules can be implemented as per the observations made to stop the spread.

Hypothesis Questions:

1. Commuting time plays a role in the spread of covid and by which means of transport we are more prone to covid?
2. Which occupational fields are more prone to covid. Closed workspaces are more secured than onfield workspaces?
3. Counties with a fewer number of health insurances has more covid spread?
4. Does private/public insurance makes the difference in the spread and which insurance has more difference.
5. Counties with low mean household incomes are more prone to covid spread than high mean household incomes?
6. Does Per capita Income play a role in the spread of covid?
7. People who walk to work are more prone to covid spread?
8. Which occupation field has the potential to increase the covid spread.