

TEAM MEMBERS:

1. Ankitha Chavan
2. Ashrita Dasari
3. Deepika Dasaroju
4. Lakshmi Manogna Bobbili
5. Venkatesh Pulibandla

COVID DATA SETS

1. COVID County Population DataSet

Name	Definition	Data type	Possible Values	Required?
County FIPS	Specific code for the county	Integer	1001,1003,1005 etc.	Yes(Primary Column)
County Name	Name of the County	String	Barbour County, Baldwin County etc.	Yes
State	Name of the State	String	AL,AK etc.	Yes
Population	Number of people.	Integer	6893,1183 etc.	Yes

2. COVID County Confirmed Cases DataSet:

Name	Definition	Data Type	Possible values	Required?
countyFIPS	Code for that particular county	Int	starts from 1000 for all odd value	Yes(Primary Key)
County Name	Name of the county	String	Can be of any name depending on the county (Multiple names possible)	Yes
State	Name of the state	String	Can be of any name depending on the state	Yes
StateFIPS	Code for that particular state	Int	starts from 1 alphabetically for all the states	Yes
Date	No. of confirmed cases on that date	Int	from 0 to a max of the number of people in that county	Yes

3. COVID County Confirmed Death cases DataSet:

Name	Definition	Data type	Possible values	Required?
------	------------	-----------	-----------------	-----------

CountyFIPS	Code for that particular county	Integer	starts from 1000 for all odd value	Yes (Primary Key)
CountyName	Name of the county	String	Can be of any name depending on the county (Multiple names possible)	Yes
State	Name of the state	String	Can be of any name depending on the state	Yes
StateFIPS	Code for that particular state	Integer	starts from 1 alphabetically for all the states	Yes
Date	No. of confirmed death on the particular data	Integer	from 0 to a max of the number of people in that county	yes

From these data sets we can find the three main factors that are a result of COVID namely population, the number of COVID cases that are confirmed each day for the given time frame, and the number of confirmed deaths that have occurred because of the COVID-19 virus for those each given dates in the time frame.

With the help of this data, we can see how COVID-19 has affected many components surrounding, for example, because of covid patients started flooding into hospitals and hence there was a shortage in hospital facilities. Another example, since the rise of COVID-19, there have been multiple lockdowns, and the employment rate started to fluctuate drastically.

Like the above examples, we can analyze different factors that have been affected by COVID-19.

REFERENCES:

1. <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/>
2. <https://data.census.gov/cedsci/table?q=dp&tid=ACSDP1Y2018.DP05>
3. <https://data.census.gov/cedsci/table?q=dp&tid=ACSDP1Y2018.DP05>
4. <https://www.bls.gov/cew/downloadable-data-files.htm>
5. <https://www.kaggle.com/unanimad/us-election-2020>
6. <https://protect-public.hhs.gov/pages/hospital-utilization>
7. <https://analystanswers.com/what-is-a-data-dictionary-a-simple-thorough-overview/>
8. <https://bulldogjob.com/news/449-how-to-write-a-good-readme-for-your-github-project>