

DATA SOURCE DESCRIPTION

The source for data is COVID data for States in the USA . <https://drive.google.com/file/d/1ilMSAgbXIHjtnZBuRK0MpTVs0ijlezI6/view?usp=sharing>

Data source URL <https://covidactnow.org/data-api>

The Covid Act Now API provides access to comprehensive COVID data — both current and historical. The data is available for all US states (50 states plus D.C., Puerto Rico, and Northern Marina Islands), counties, and metros and is aggregated from a number of official sources, quality assured, and updated daily. It is available in JSON and CSV formats. For the current project , It is intended to use data related to the US states and aggregated data for US. The data is updated everyday (noon EST as per the website) , hence it is planned to design the data pipeline so that daily data can be pulled from API every day. The historical data will be pulled from API once and combined with daily data for analysis .

In order to access the API , we need to generate a API KEY . This key will be passed as a parameter for API calls to download the dataset.

The details of fields in the dataset are given in data definitions

Following is the API query for historical states data <https://api.covidactnow.org/v2/states.timeseries.csv?apiKey=>

The API query for current data : <https://api.covidactnow.org/v2/states.csv?apiKey=>

The API query for aggregated US data : <https://api.covidactnow.org/v2/country/US.timeseries.csv?apiKey=>