*India state testing time series* *Steve Bachmeier* *2020-05-20*

The file ‘india\_tests.R’ uses the available APIs from <https://api.covid19india.org/> to extract testing data from Indian states.

## Notes

### APIs used

One API is used:

* “<https://api.covid19india.org/state_test_data.json>”

### Manual fixes

On occasion, manual fixes must be implemented that may or may not be required permanently. For example, there was a single bad date in the raw data; Specifically, the first data point for “Meghalaya” was dated “2020-02-16”. However, the second data point for that location was “2020-04-17”. The first date was thus manually changed to “2020-04-16” after confirming the data itself makes sense to do so.

Eventually, however, the erroneous data was fixed in the raw data pulled in and the manual fix was removed from the code.

As of 2020-06-20, the only manual fix made is to remove a duplicate observation from Uttar Pradesh on 2020-04-05:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| state | date | positive | negative | unconfirmed | totaltested |
| Uttar Pradesh | 2020-04-05 | 278 | 4796 | 181 | 5255 |
| Uttar Pradesh | 2020-04-05 | 278 | 4796 | 179 | 5255 |

I the case of such date-state duplicates, it is assumed that the last entry is the most up-to-date one and all others are dropped.

### States

The majorty of the Indian states are able to be matched with IHME’s get\_location\_metadata() output and location\_ids merged on. There are two unique locations.

#### Dadra and Nagar Haveli and Daman and Diu

The downloaded dataset contains locations “Dadra and Nagar Haveli and Daman and Diu”. IHME, meanwhile, has locations for “Dadra and Nagar Haveli” and “Daman and Diu”, but not both together. As such, the location\_id for “Dadra and Nagar Haveli and Daman and Diu” has been assigned as “NA”.

#### Combining ‘Ladakh’ and ‘Jammu and Kashmir’

To keep consistent with GBD, locations ‘Ladakh’ and ‘Jammu and Kashmir’ are combined into ‘Jammu & Kashmir and Ladakh’ (location\_id 4854). When adding the two state’s variables together, it is assumed that empty values (NA) are zero, e.g. NA + 10 = 10.

## END OF README ##