

KKT COVID-19 Dataset*

Varun Vasudevan¹, Jithin K. Sreedharan², Varsha Sankar³, and Siddarth A. Vasudevan⁴

¹Stanford University, Stanford, California, USA, (Email: devan@stanford.edu)

²Purdue University, West Lafayette, Indiana, USA, (Email: jithinks@purdue.edu)

³SAP, Palo Alto, California, USA (Email: s.varsha100@gmail.com)

⁴Philip Morris International Research and Development, Philip Morris Products S.A., Neuchâtel, Switzerland (Email: devan.siddarth@gmail.com)

April 8, 2020

DISCLAIMER: This report is provisional and will be updated in the coming weeks. The updated report will contain sections on Karnataka and Tamil Nadu.

Abstract

India reported its first case of COVID-19 on Jan 31, 2020 in Kerala. In KKT-COVID19, KKT stands for the three south Indian states Karnataka, Kerala and Tamil Nadu.

1 Introduction

2 Kerala

- The format of the bulletin has changed/evolved over time. Therefore it seemed practically infeasible to write a single script to automatically get data from the bulletin ([Vasudevan et al., 2020](#)).
- Number of positive cases reported per district is obtained from the "Test Results" section of the bulletin and then cross-checked with the numbers mentioned for that date in the "Chronology of Positive cases" table in annexure 2.
- Number of recovered cases reported per district is obtained from the "Test Results" section of the bulletin.

Primary data source ([DHS, 2020](#))

2.1 Timeline of events

- First case
- Second wave
- Incoming planes were stopped

*Working title

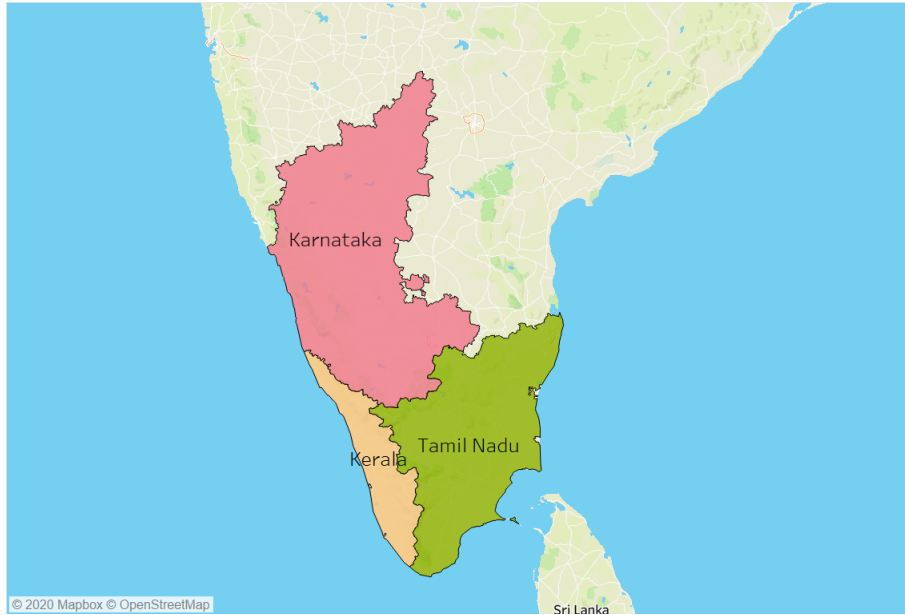


Figure 1: Map of south India showing the three neighboring states Karnataka, Kerala and Tamil Nadu. (Generated using Tableau Desktop 2020.)

- Lockdown in Kerala
- Lockdown in India

3 Acknowledgment

The authors would like to thank

References

- [1] Varun Vasudevan, Jithin K. Sreedharan, Varsha Sankar, and Siddarth A. Vasudevan. KKT COVID-19 dataset, 2020.
- [2] Kerala DHS. Daily bulletin. <http://dhs.kerala.gov.in/>, 2020. [Online; accessed 01-April-2020].

A Kerala

A.1 Format changes and inconsistencies in data reporting

In this section we describe the major format changes and inconsistencies in data reporting in the Directorate of Health Services (DHS) daily bulletin.

- Jan 31, 2020: The first DHS bulletin was published on this day.
- Mar 08, 2020: This bulletin reported five positive cases from Pathanamthitta. This information was also reported in the evening news¹ on that day. There is a mistake in the later bulletins where these five cases are marked against Mar 09, 2020 in the chronology of positive cases table of Annexure-2.
- Mar 10, 2020: For gaining clarity on the number of people reported positive in the English version of the bulletin, we also referred to the Malayalam version of the bulletin.
- Mar 13, 2020: First bulletin to have Annexure-2 – chronology of positive cases.
- Mar 22, 2020: Until Mar 22, the last table in the bulletin didn't have a title. It was assumed that the numbers in that table represented the active cases from a district. When the authors contacted the government staff for a clarification they added the title "District wise distribution based on hospital admission" to the table starting Mar 23.
- Mar 23, 2020: As per the description under the "Test Results" section, a person from Kannur was declared negative and discharged. But this information was not recorded in the "Present Status" column of the chronology of positive cases table until Mar 24.
- Mar 24, 2020: A date column was added to the chronology of positive cases table.
- Mar 26, 2020:
 - (a) The "Test Results" section of the bulletin mentions that a patient from Pathanamthitta tested negative. This information wasn't updated in the "Present Status" column of the chronology of positive cases table until Mar 27.
 - (b) The "Test Results" section of the bulletin mentions that two foreigners under treatment in Ernakulam tested negative. This information wasn't updated in the "Present Status" column of the chronology of positive cases table until Mar 28.
 - (c) A third column titled "Other Districts" was added to the "District wise distribution based on hospital admission" table. This made it seem like one could calculate the number of active cases from a district using the entries in this table. But that isn't true (as of Apr 06, 2020) due to the inconsistencies in the reporting of entries in the newly added column. The authors validated the presence of inconsistencies by contacting the government staff².
- Mar 28, 2020: The "Test Results" section mentions that a foreigner from Ernakulam tested negative, whereas in the chronology of positive cases table a foreigner from Idukki is marked negative. As per the evening news¹ this foreigner had travelled from Munnar in Idukki district to Ernakulam before testing positive on Mar 15. We believe that this is the reason for the

¹ "Malayala Manorama" news on TV at 6 pm.

² Email us for more info on this.

confusion. To be consistent with the bulletin from Mar 15, we treat this person as a confirmed case from Idukki.