

Spatial Internet Shutdowns Data for India

A Spatio-Temporal Data Set of Internet
Shutdowns in India, 2016 to 2022

Jens Koning ^a

^a Peace Research Institute Oslo (PRIO)

https://github.com/jens-koning/internet_shutdowns_india

1 Introduction

India holds the record for the highest number of government-induced internet shutdowns globally, surpassing all other countries (AccessNow, 2023a; Ruijgrok, 2022). From 2016 to the end of 2022, there were a staggering 1,978 district-level shutdowns in the country. To gain deeper insights into these shutdowns and their impact on political and economic aspects in India, but also other countries experiencing similar restrictions to the internet, it is crucial to have a comprehensive spatio-temporal data set that provides information on the locations and duration of these events (Keremoğlu and Weidmann, 2020). To bridge this informational gap, we have developed an data set that aims to make internet shutdown data easily accessible and usable for analysis by aligning available online information with the spatial coding of the Database of Global Administrative Areas (GADM) administrative level 2 naming conventions, enhancing its usability.

2 Extended Access Now Data

The process of creating this data set involved expanding the existing Shutdown Tracker Optimization Project (STOP) data set for the years 2016-2022 from the Access Now webpage (Access Now, 2023b). Access Now is a New York-based non-governmental organization defending and extends the digital rights of users at risk around the world. The following paragraphs briefly outlines the methodology and limitations of the STOP data set.¹.

¹ A complete breakdown of Access Now's methodology can be read in Access Now (2023c)

2.1 STOP Methodology and Limitations

In the STOP data, an internet shutdown refers to a deliberate disruption of internet or electronic communications, rendering them inaccessible or unusable for a specific population or within a particular location. STOP's definition, established in 2016 with input from technologists, policymakers, and activists, encompasses complete network shutdowns, bandwidth throttling, and service-based blocking of two-way communication platforms. Instances where the cause of disruption is uncertain are included in a separate category until confirmed or disproven as intentional. Their tracker encompasses global internet shutdowns caused by both government and non-state actors, with the country column typically reflecting the perpetrator. However, in cases where non-government actors impose a shutdown in a country, the country column specifies the affected nation, while the «ordered_by» and «decision_maker» columns provide details about the actual perpetrator.

Each shutdown instance refers to a disruption event lasting over one hour or a series of related events attributed to the same circumstances, justifications, methods, and perpetrators. The instance may persist even if internet services are restored and subsequently shut down again, or if the scope of the shutdown expands or contracts. For example, a "digital curfew" involving nightly shutdowns for several consecutive days would be considered a single instance. This grouping enables us to achieve policy goals, attract media attention, and pursue mitigation and remedy measures.

The recorded instances of shutdowns were derived from various sources, including local and international news reports, reports from local actors through Access Now's Digital Security Helpline or the STOP Coalition e-mail list, and direct input from telecommunication and internet companies. If a shutdown occurs without a specific triggering event or in response to a broader political struggle, each shutdown is recorded as a separate instance once the service has been restored for at least 24 hours before subsequent disruptions.

It is important to note that there may be slight variations in the number of shutdowns between trackers due to methodological differences and ongoing updates. Unconfirmed shutdowns are not included in the STOP tracker to maintain accuracy. Additionally, the tracker does not cover network or service disruptions caused by factors beyond its scope, such as natural disasters or technical issues. For shutdowns that follow a "curfew" style pattern with multiple instances occurring over a period of time, STOP treat it as separate instances unless technical measurement data confirms a clear and continuous pattern attributed to the same cause.

3 Changes Made to the STOP Data

We meticulously reviewed all 1,978 observations (shutdowns) and extracted the "state" and "district" information based on the "area_name_string." We automated district name matching using the *fuzzyjoin*-package for the R programming language. To ensure accuracy, we manually

verified matched-names using internet searches (Wikipedia), cross-referencing the names with GADM level 2 administrative naming conventions, utilizing the data available for India from GADM's official website (GADM, 2023). In cases of state-wide shutdowns, we triangulated information based on the URLs provided in the original STOP original data set to identify the corresponding districts. These URLs contained important additional information regarding the locality of shutdowns beyond what was denoted in the original STOP data set. The additional districts identified were manually added to the "districts" column.

Subsequently, we merged the different years into a comprehensive time series spanning from 2016 to 2022 and addressed any inconsistencies present in each column. Lastly, we calculated the duration of each shutdown using the *lubridate*-package for R whenever data was available, providing additional insights into the duration of these events.

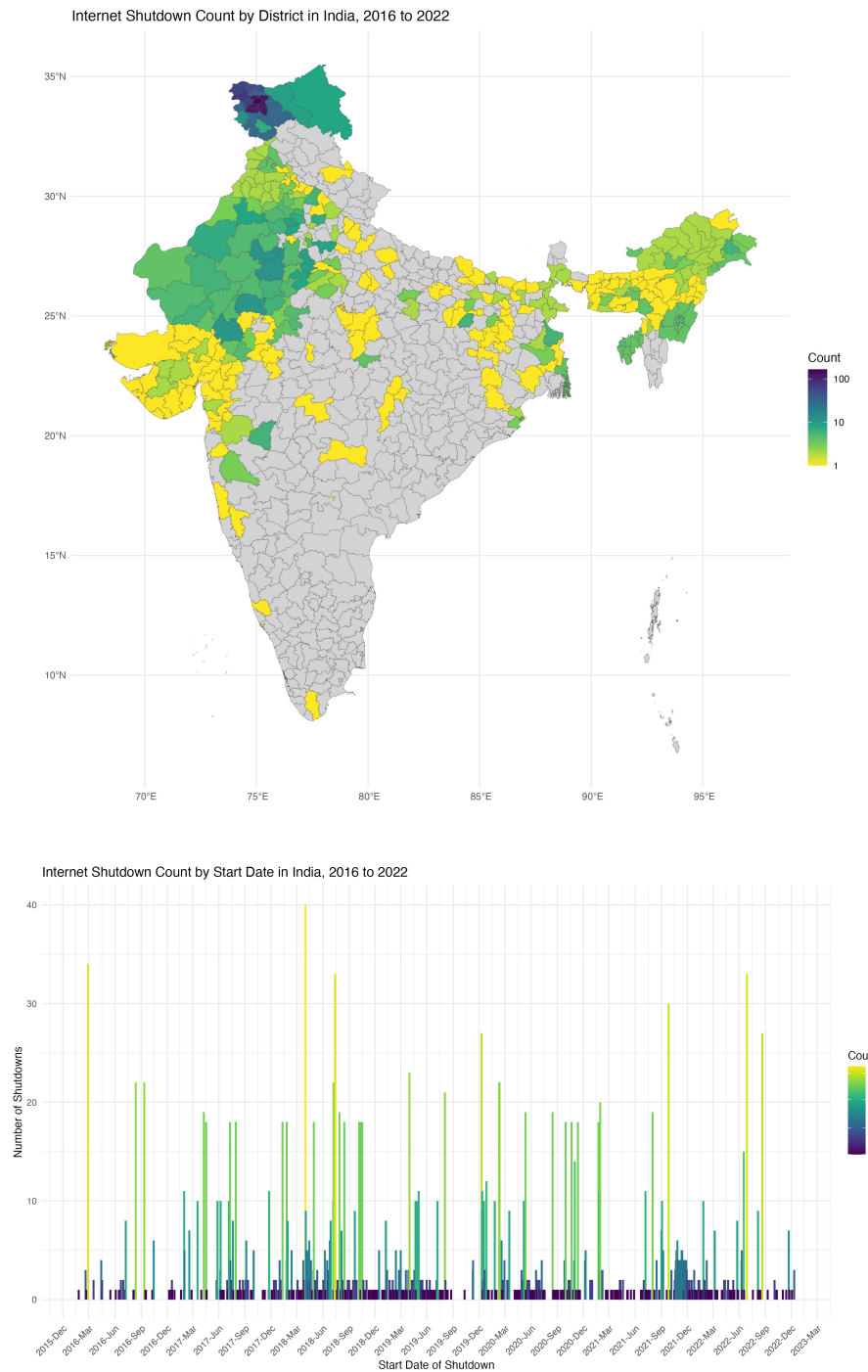
4 Description of Output Data Set

The output file for end users is "shutdowns_india_2016_22.rds" and can be accessed through the github repository. The following list provides an overview of the variables contained in the data set. The main unit of observation is *district-event*, meaning the shutdown event that took place in a district for a specified period of time. Each event is given one row in the data set, and where possible, end dates have been denoted.

- **start_date**: start date of shutdown in "%Y/%m/%d" format.
- **end_date**: end date of shutdown in "%Y/%m/%d" format.
- **duration_days**: the duration of the shutdown in days (1 day is 24 hours).
- **duration_hours**: the duration of the shutdown in hours.
- **country**: India.
- **state**: state in India. GADM level 1 naming.
- **districts**: district in India. GADM level 2 naming.
- **event**: what happened where the shutdown took place.
- **area_name_string**: original string denoting the area where it took place in STOP.
- **ordered_by**: the government authority (Local, State etc.) who issued the shutdown.
- **gov_justification**: the government justification for the shutdown (if any).
- **affected_network**: the network affected by the shutdown.
- **actual_cause**: actual cause of the shutdown as estimated by STOP.
- **source_link**: source of information (URL).

- **gov_ack_source**: government acknowledgement or official document describing the shutdown (URL).

5 Spatio-temporal Distribution



References

- AccessNow (2023a). Five years in a row: India is 2022's biggest internet shutdowns offender. <https://www.accessnow.org/press-release/keepiton-internet-shutdowns-2022-india/>.
- AccessNow (2023b). KeepItOn: fighting internet shutdowns around the world. <https://www.accessnow.org/campaign/keepiton/>.
- AccessNow (2023c). #KeepitOn STOP data set. https://www.accessnow.org/wp-content/uploads/2023/03/Read-Me_STOP_data_methodology.pdf.
- GADM (2023). GADM data set. <https://gadm.org/data.html>.
- Keremoğlu, E. and Weidmann, N. B. (2020). How Dictators Control the Internet: A Review Essay. *Comparative Political Studies*, 53(10-11):1690–1703. <https://doi.org/10.1177/0010414020912278>.
- Ruijgrok, K. (2022). The authoritarian practice of issuing internet shutdowns in India: the Bharatiya Janata Party's direct and indirect responsibility. *Democratization*, 29(4):611–633. <https://doi.org/10.1080/13510347.2021.1993826>.