

Rajya Sabha Codebook 1.0

<https://github.com/tcpd/rs-data-release>

Last Updated: 29 June, 2022

Contents

Title	Page
1. Introduction	2
2. Terms and Conditions	3
3. Table 1: Variable names, type, and description	4-6
4. Documentation Process	7
5. Exceptions	8
6. Acknowledgements	8



1. Introduction

The Rajya Sabha Dataset is a corpus of data concerning all members of the Rajya Sabha of India from **March 1952** to **September 2021**. The data comes from the official website of Rajya Sabha (<https://rajyasabha.nic.in/>). This official website contains information about members' political history, positions held, sociological information, etc. We first extracted this information, and then cleaned and structured it to fit into a tabular format.

Each individual member of the Rajya Sabha has been assigned a unique identification number (PID) to help track their political history. We have made the greatest of efforts to keep this PID common across the Rajya Sabha and Lok Sabha datasets. This means that if an individual contested any seat in the national-level General Elections (GE), state-level Assembly Elections (AE), and/or Rajya Sabha Elections, they will ideally have the same PID across all TCPD datasets. However, please let us know if we might have missed out on a few individuals.

This codebook provides information about the variables included in the Rajya Sabha dataset. The codebook will be updated as and when new information is added to the dataset.

We hope that you will find this repository useful and request that you cite the data as indicated in the codebook, whenever you use the data in a publication or public presentation.

With thanks,

TCPD Team

2. Terms and conditions

The Rajya Sabha dataset is a list of all members of the Rajya Sabha from March 1952 to September 2021. The users are free to download, display or include the data in other products for non-commercial purposes at no cost subject to the following limitations:

Data Citation: "TCPD Rajya Sabha dataset (TCPD-RSD), 1952 - 2021". Trivedi Centre for Political Data, Ashoka University.

Codebook Citation: Ananay Agarwal, Neelesh Agrawal, Saloni Bhogale, Vedant Jumle, Shruti Kraleti, Mohit Kumar, Gilles Verniers. 2022. "TCPD Rajya Sabha Dataset (TCPD-RSD), 1952 - 2022 Codebook", Trivedi Centre for Political Data, Ashoka University.

No Endorsement: The user must not claim or imply that the Trivedi Centre for Political Data endorses the user's use of the data or use of the Centre's logo(s) or trademarks(s) in conjunction with the same.

No Warranty: The Centre makes no warranties with respect to the accuracy of the data. While we do our best to ensure that the numbers are accurate, there is room for error. And as such the user must agree that the Centre shall not be held responsible or liable to the user for any errors, omissions, misstatements and/or misrepresentations of the data though the user is encouraged to report the same to us ([send us a message](#) or write us at tcpd-contact@ashoka.edu.in).

Personal data: The Centre may record visits to the TCPD website without collecting the personal information of the users. The records shall be used for statistical reports only.

Acknowledgement: The user must agree that the use of the Data presented within the application can be seen as the acknowledgement of unconditionally accepting the Terms of Use presented by the Centre.

3. Table 1: Variable names, type, and description

No.	Variable Name	Variable Type	Variable Description
1.	pid	string	A unique identification number assigned to each Individual member to help track their political career across national-level General Elections, state-level Assembly Elections, and Rajya Sabha Elections
2.	Member_Name	string	The name of the individual member [As of March 1952 - September 2021, there are 2314 unique Individuals in the Rajya Sabha.]
3.	Image	string	The link to the individual's image stored on the official Rajya Sabha website
4.	Father_Name	string	The name of the individual's father if available
5.	Mother_Name	string	The name of the individual's mother if available
6.	Date_of_Birth	date(dd-mm-yyyy)	The day, month, and year the individual was born.
7.	Marital_Status	string	Describes if the individual is/ was single, unmarried, married, or divorced, as per the Rajya Sabha website If available, also mentions the date and/or month and/or year when the Individual was married
8.	Spouse_Name	string	The name of the individual's spouse if available
9.	Children	string	The name and gender of all the Individual's offspring
10.	Educational_Qualification	string	Describes in detail what degrees the individual has attained and from which educational

			institutions
11.	Profession	string	Describes the occupational experience of the individual
12.	Positions_Held	string	Describes all the positions the individual has held within the Indian Legislature, the Executive, or the Judiciary, along with the years in which the said positions were held
13.	State	string	The state which the individual represents in the Rajya Sabha for the particular tenure If the individual was Nominated, this variable will instead say <i>Nominated</i> .
14.	Nominated	logical	Describes if the individual was Nominated (TRUE) or represents a State (FALSE) in the Rajya Sabha for the particular tenure
15.	Party	string	The party which the individual represents in the Rajya Sabha for the particular tenure
16.	Term_From	date(dd-mm-yyyy)	The day, month, and year from which the individual started the particular tenure in the Rajya Sabha
17.	Start_Year_Appointment	date(yyyy)	The Year when the individual was appointed to the Rajya Sabha for a particular tenure
18.	Term_To	date(dd-mm-yyyy)	The day, month, and year when the individual is/ was officially supposed to end the particular tenure in the Rajya Sabha
19.	End_Year_Appointment	date(yyyy)	The Year when the

			<p>individual will have officially retired, if they complete/ed their full term</p> <p>The actual date when they retire may or may be this.</p>
20.	Term_Number	numeric	Describes the <i>n</i> th term for this individual for the particular tenure.
21.	Reason_of_Vacation	string	<p>If the individual vacated a particular tenure before the official date, this variable gives the reason as to why. The following reasons have been declared on the official website :</p> <ol style="list-style-type: none"> 1. Retirement 2. Resignation 3. Death 4. Disqualification 5. Elected to Lok Sabha 6. Reorganisation of the State 7. Seat Declared Vacant
22.	Date_of_Vacation	date(dd-mm-yyyy)	If the individual vacated a particular tenure before the official date, this variable gives that date
23.	Type	string	Describes if the individual is currently a member of the Rajya Sabha (Current) or not (Former)

4. Documentation Process

All the information for this dataset was taken from the official website of the Rajya Sabha (<https://rajyasabha.nic.in/>). The data was first extracted by scraping using the Python libraries BeautifulSoup4 and Selenium. It was then cleaned and formatted to fit into a tabular format.

Afterwards, the PID was assigned by the Centre's open-source software called SURF to track an Individual's political career through national-level General Elections, state-Level Assembly Elections, and Rajya Sabha Elections. For more details on SURF, please [see here](#).

5. Exceptions

The following are the manual changes made by the Centre due to missing data on the official website :

1. Member : Shri Manohar Kant Dhyani
 - a. Added state "Uttarakhand" for the tenure 09/11/2000 - 25/11/2002
 - b. Source : https://www.wikiwand.com/en/Manohar_Kant_Dhyani

6. Acknowledgements

We thank our colleagues - Shivam Gangwani, Dipanita Malik, Supriya Sanjay, and Priyamvada Trivedi who provided invaluable assistance to the project.