

## Indian Administrative Services Dataset Codebook v1.2

<https://github.com/tcpd/ias-release>

Last Updated: 11 January, 2022

### Contents

| Title  | Page |
|--|------|
| 1. Introduction  | 2    |
| 2. Terms and Conditions  | 3    |
| 3. Table 1: Variable names, type, and description for profile of IAS officers            | 4    |
| 4. Table 2: Variable names, type, and description for education details of IAS officers  | 6    |
| 5. Table 3: Variable names, type, and description for experience details of IAS officers | 7    |
| 6. Table 4: Data Sources   | 9    |
| 7. Documentation Process   | 10   |
| 8. Exceptions  | 13   |
| 9. Version History   | 15   |

## 1. Introduction

This dataset contains information on the officers of the Indian Administrative Service (IAS). The data has been carefully extracted and verified from official government websites which documents their names and tenure dates. The data for the TCPD-IAS dataset was first retrieved from government sources in May-June 2021 and further updated in December 2021. See Table 4 for the data sources and section 9 for a complete version history.

We table it in an orderly manner, to systematically place the information under variables such as the unique IAS officer ID, name of the IAS officer, Service, Cadre, Allotment Year, Date of Birth, Date of Joining, Source of Recruitment, Gender, Place of Domicile, Mother Tongue, Languages Known, as well as variables around their experience history and educational qualifications. The data also mentions the sources of information for each instance captured, so that users can verify the same.

We have provided the documentation process in the following sections which lists the various sources and how they were used to populate the data. Moreover, we have also listed detailed exceptions that we came across while populating and checking the dataset, clarifying the assumptions we have made.

Lastly, we are aware that the information from these sources is updated periodically. We will do our best to keep the dataset updated and post announcements accordingly. We encourage our users to get in touch if they observe any discrepancies.

## 2. Terms and conditions

The TCPD-IAS dataset consists of information on IAS officers from 1951 to 2020. 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 Indian Administrative Service Officers Dataset (TCPD-IAS), 1951-2020". Trivedi Centre for Political Data, Ashoka University.

**Codebook Citation:** Neelesh Agrawal, Srishti Gupta, Vedant Jumle, Mohit Kumar, SV Sai Vikas. 2022. "TCPD Indian Administrative Service Officers Dataset (TCPD-IAS), 1951-2020 Codebook 1.2", Trivedi Centre for Political Data, Ashoka University.

**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.

**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](mailto: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.

### 3. Table 1: Profile of IAS officers

| No. | Variable Name                    | Variable Type | Variable Description  |
|-----|----------------------------------|---------------|---|
| 1   | ID                               | string        | Unique IDs given to the IAS officers.   |
| 2   | Name                             | string        | Name of the the IAS officer   |
| 3   | Cadre                            | string        | State/Union territory to which the IAS officer was appointed                                  |
| 4   | Allotment_Year                   | numeric       | The year in which the IAS officer was allotted the state cadre.                               |
| 5   | Date_of_Birth                    | date          | Date of Birth of the IAS officer. The format is: yyyy-mm-dd.                                  |
| 6   | Date_of_Joining                  | date          | Date of Joining of the IAS officer. The format is: yyyy-mm-dd.                                |
| 7   | Source_of_Recruitmment           | string        | Source of Recruitment of the IAS officer.   |
| 8   | Gender                           | string        | Gender of the IAS officer   |
| 9   | Place_of_Domicile                | string        | State in which the IAS officer holds a permanent residence                                    |
| 10  | Mother_Tongue                    | string        | Mother Tongue of the IAS officer  |
| 11  | Languages_Known                  | string        | Languages known by the IAS officer  |
| 12  | Retired                          | logical       | TRUE if retired, FALSE if active  |
| 13  | Retirement_Reason                | string        | Reason for retirement of the IAS officer. This is NA if the officer is active.                |
| 14  | Last_Education_Qualifica<br>tion | string        | Highest education qualification of the IAS officer.   |
| 15  | Last_Education_Subject           | string        | Academic discipline in which the IAS officers acquired their highest education qualification. |
| 16  | Last_Education_Division          | string        | Division in which the IAS   |

|    |                             |        |  |
|----|-----------------------------|--------|--|
|    |                             |        | officers secured their highest academic qualification.   |
| 17 | Last_Designation            | string | The last administrative position held by the IAS officer.  |
| 18 | Last_Level                  | string | The last level of the administrative position held by the IAS officer                                      |
| 19 | Last_Office                 | string | The last ministry or department or district where the IAS officer was working.                             |
| 20 | Last_Organisation           | string | The last organization where the officer was working.   |
| 21 | Last_Field_of_Experience    | string | This category shows the administrative department in which the IAS officer held the last designation.      |
| 22 | Last_Category_of_Experience | string | This category shows the associated departments within the broad category for the Last_Field_of_Experience. |
| 23 | Last_Start_Date             | date   | Date of beginning of latest appointment. The format is: yyyy-mm-dd.  |
| 24 | Last_End_Date               | date   | Date of the end of latest appointment. The format is: yyyy-mm-dd   |
| 25 | Source                      | string | Website from which the data was extracted.   |
| 26 | Gender_Source               | string | Indicates whether gender value was assigned by TCPD or derived directly from the source.                   |

**4. Table 2: Education Details of IAS officers**

| No. | Variable Name       | Variable Type | Variable Description  |
|-----|---------------------|---------------|---|
| 1   | ID                  | string        | Unique IDs given to the IAS officers.   |
| 2   | Name                | string        | Name of the IAS officer   |
| 3   | Cadre               | string        | State/Union territory to which the IAS was allotted   |
| 4   | Reference_Value     | numeric       | These are numeric values assigned to the IAS officers qualifications. Value of "0" is assigned to the first level of qualification. For every increase in the level of qualification, values are assigned incrementally starting from 0.  |
| 5   | Qualification       | string        | Education qualification of the IAS officer.   |
| 6   | Subject             | string        | Academic discipline in which the IAS officers acquired their education qualifications.  |
| 7   | Category_of_Subject | string        | Category (assigned by TCPD) that the subject falls under. There are eight categories:<br><ol style="list-style-type: none"> <li>1. Social Sciences and Liberal Arts</li> <li>2. Governance, Law and Management</li> <li>3. Engineering, Technology and Mathematics</li> <li>4. Natural Sciences</li> <li>5. Languages and Fine Arts</li> <li>6. Finance</li> <li>7. Medical</li> <li>8. Others</li> </ol> |
| 8   | Division            | string        | Division in which the IAS officers secured their academic qualifications.   |
| 9   | Source              | string        | Website from which the data was extracted.  |

**5. Table 3: Experience Details of IAS officers**

| No. | Variable Name          | Variable Type | Variable Description  |
|-----|------------------------|---------------|---|
| 1   | ID                     | string        | Unique IDs given to the IAS officers.   |
| 2   | Name                   | string        | Name of the IAS officer   |
| 3   | Cadre                  | string        | State/Union territory to which the IAS was allotted   |
| 4   | Reference_Value        | numeric       | These are numeric values assigned to the IAS officers designation. Value of "0" is assigned to the latest designation held by the IAS officer. The values are assigned incrementally from 0 and assigned to each designation held by the IAS officer, ending with their first designation when they started their tenure. |
| 5   | Designation            | string        | Administrative position held by the IAS officer.  |
| 6   | Level                  | string        | Level of the administrative position held by the IAS officer  |
| 7   | Office                 | string        | The ministry or the department or the district where the IAS officer was working.   |
| 8   | Organisation           | string        | The organisation where the officer was working.   |
| 9   | Field_of_Experience    | string        | This category shows the administrative department in which the IAS officer held the designation.  |
| 10  | Category_of_Experience | string        | This category shows the associated departments within the broad category for the Field_of_Experience.   |
| 11  | Start_Date             | date          | Date of beginning of appointment. The format is: yyyy-mm-dd   |

|    |          |        |   |
|----|----------|--------|---|
| 12 | End_Date | date   | Date of the end of appointment. The format is: yyyy-mm-dd |
| 13 | Source   | string | Website from which the data was extracted.                |



**6. Table 4: Data Sources**

| No. | Name of the source                          | Link to the Source  |
|-----|---|---|
| 1.  | Department of Personnel and Training (DOPT) | <a href="https://easy.nic.in/civilListIAS/YrCurr/FinalCL.asp">https://easy.nic.in/civilListIAS/YrCurr/FinalCL.asp</a> |
| 2.  | SUPREMO                                     | <a href="https://supremo.nic.in/KnowYourOfficerIAS.aspx">https://supremo.nic.in/KnowYourOfficerIAS.aspx</a>           |

**Disclaimer from SUPREMO:**

Every effort has been made to ensure accuracy of data on the ER Sheet. However these need to be verified by and supplemented with documents issued finally by the concerned authorities.

Department of Personnel & Training or the National Informatics Centre will not be responsible for any decision or claim that is based on the basis of displayed data. This disclaimer is valid for all pages in the ER Sheet.

## 7. Documentation Process

For the data, the primary data sources were the official government websites with domain names *gov* and *nic* (as referred in table 4). The Department of Personnel and Training (DOPT) website provides data on (current) IAS officers who were appointed between the years 1983 to 2020 whereas the data collected from the SUPREMO website has the data of IAS officers who were appointed between 1951 to 2020.

Data from the [DOPT](#) & [SUPREMO](#) websites was scraped into tables. The data was sourced first during May-June 2021 and updated in December 2021. The DOPT and SUPREMO datasets were combined, and duplicate rows removed, to generate the final dataset.

The DOPT website has data only on the latest experience of the IAS officer. Therefore, when combining the datasets, entries in the SUPREMO dataset were considered to compare similar or duplicate entries.

### Formatting fixes:

1. An important part to the combining of the two dataset was the 'Identity No.' Identity No. is a unique alpha-numeric string that was assigned to each IAS officer. It is an 8-character long string where the first 2 characters are short-forms of the officer's cadre and the last 6 digits are numbers (Example: MH16100). In the combined dataset, the column called 'Officer IDs' contains all the relevant Identity Nos.
2. Some rows in the SUPREMO data refer to the same person. These rows are in the form of pairs where one row contains the information on education, while the other contains the information on experience. Both these rows were combined into one single row and the first 2-digits in the ID value were removed.

### Gender assignment:

For rows where the gender was not available, it was assigned by the TCPD team using the name titles (Shri, Smt etc), and secondary research.

### Category of Subject assignment:

All the different education subjects for the officers were assigned a category of subject from eight categories:

1. Social Sciences and Liberal Arts
2. Governance, Law and Management
3. Engineering, Technology and Mathematics
4. Natural Sciences
5. Languages and Fine Arts
6. Finance
7. Medical
8. Others

The exact mappings are available on:

<https://docs.google.com/spreadsheets/d/1u6eJ2sEXz06zxhThK0oNfdvtmfNKe0vuFwk0TRZS5qA/edit#gid=0>

The final dataset contains 13570 entries. It has been split into three tables for better comprehension.

1. ias-profile.csv:

This is a csv file that contains the data that was scraped and cleaned. This does not include the education and experience details of the officers. There are 13570 rows in this file.

2. ias-education.csv:

This is a csv file that has data on the educational background of the IAS officers. Every row contains information on one specific educational detail for each IAS officer. There are 22149 rows in this file.

3. ias-experience.csv:

This is a csv file that has data on the past and the current experience of the IAS officers. Each row presents the information on the IAS officer's experience. There are 151568 rows in this file. There are 211 officers for whom there is no information on experience.

This data is also represented in a hierarchical json format, combining the three files. This allows more technically advanced users to only use one file to analyse the data. The format for the json file is as follows.

```
{
  "<ID>":{
    "Name": "",
    "Cadre": "",
    "Allotment_Year": "",
    "Date_of_Birth": "",
    "Date_of_Joining": "",
    "Source_of_Recruitment": "",
    "Gender": "",
    "Place_of_Domicile": "",
    "Mother_Tongue": "",
    "Languages_Known": "",
    "Retirement_Reason": "",
    "Source": "",
    "Gender_Source": "",
    "Education":[
      {
        "Qualification": "",
        "Subject": "",
        "Category_of_Subject": "",
        "Division": ""
      }, {<repeat>}
    ],
    "Experience": [
```

```

    {
      "Designation": "",
      "Level": "",
      "Office": "",
      "Organisation": "",
      "Field_of_Experience": "",
      "Category_of_Experience": "",
      "Start_Date": "",
      "End_Date": ""
    }, {<repeat>}
  ]
},
"<ID>": {<repeat>}
}

```

Example queries for the json file:

- `officer[MH0123][Education][0][Qualification]`  
This will return the first educational qualification for the officer with ID "MH0123"
- `officer[RJ1209][Experience][1][Organisation]`  
This will return the second organisation that the officer with ID "RJ1209" has worked with. (May give error if officer has held only one position)

## 8. Exceptions

### Data discrepancies:

For the following officers, the Date of Joining seems to be erroneous in the source (SUPREMO), as it is logically improbable.

AM806102, AM806103, AP032721, AP032730, AP909110, BH036402, BH037612, BH040719, CG039311, CG046201, GJ023711, GJ026218, JH106C04, MH032111, MH034015, MH036911, MP039312, MP040915, OR905700, OR905800, UD051516, UD055211, UD056512, UD057613, UP051518, UP051520, UP051521, UP051522, UP055013, UP055015, UP055212, UP056513, UP056514, UP057612, UP057622, UP057623, UP917800, UT903000, WB027825, WB029012, WB029013, WB029016, WB029021, WB029022, WB030011, WB030012, WB030015, WB030019, WB030028, WB031111

For the following officers, there are duplicate entries with conflicting values as they appear in the source (SUPREMO).

BH015000, MT011411, RJ029800, MT007500, MT011514, UP020000, TN010700, PB008203, PB004300, MT004500, PB005200

The following rows seemed to have erroneous "Start\_Date" value (as derived from SUPREMO) in the ias-experience.csv file:

| ID       | Name                      | Cadre      | Reference_Value | Start_Date |
|----------|---------------------------|------------|-----------------|------------|
| TN014601 | Shri Vinay Jha            | Tamil Nadu | 18              | 1915-10-01 |
| KN015600 | Dr. Pj Nayak              | Karnataka  | 13              | 1900-01-01 |
| HY017700 | Shri Varinder Singh Kundu | Haryana    | 29              | 1906-05-01 |
| OR114Q07 | Shri Amrit Ruturaj        | Odisha     | 2               | 1900-01-01 |
| TN810005 | Shri S Jayandhi           | Tamil Nadu | 0               | 2029-02-22 |

The following rows seemed to have erroneous "End\_Date" values (as derived from SUPREMO) in the ias-experience.csv file:

| ID | Name | Cadre | Reference_Value | End_Date |
|----|------|-------|-----------------|----------|
|----|------|-------|-----------------|----------|

|          |                    |             |   |            |
|----------|--------------------|-------------|---|------------|
| MH013300 | Shri Ck Modi       | Maharashtra | 0 | 1900-08-31 |
| RJ020511 | Shri K S Galundia  | Rajasthan   | 9 | 1902-02-01 |
| RJ025516 | Shri Rc Roongta    | Rajasthan   | 0 | 1915-10-31 |
| RJ027117 | Shri P D Paliwal   | Rajasthan   | 0 | 1938-07-31 |
| KN034900 | Shri Anurag Tewari | Karnataka   | 0 | 2041-05-31 |
| KN034900 | Shri Anurag Tewari | Karnataka   | 1 | 2041-05-31 |

## 9. Version History

| S.No. | Description  | Release Date | Version |
|-------|--|--------------|---------|
| 1     | Initial Release  | 29 Nov 2021  | v1.0    |
| 2     | <ol style="list-style-type: none"> <li>1. Date of data scrape added</li> <li>2. Date_of_Joining updated for source discrepancy</li> <li>3. Clarification added that DOPT releases exist for active officers' list</li> <li>4. Retirement column updated</li> <li>5. New column added for education - category of subject</li> <li>6. Date format updated to yyyy-mm-dd in ias-experience.csv as well as ias-profile.csv</li> </ol> | 10 Dec 2021  | v1.1    |
| 3     | <ol style="list-style-type: none"> <li>1. Data scraped again from SUPREMO and merged with the existing dataset.</li> <li>2. New discrepancies detected as attached in exceptions.</li> <li>3. Erroneous rows were removed from the dataset.</li> <li>4. Removed the "Service" column.</li> <li>5. Missing gender values were updated by TCPD.</li> </ol>   | 11 Jan 2022  | v1.2    |