

## INTRODUCTION

India, the world's largest democracy, unsurprisingly has a huge and complex judicial system. Judgements of the higher judiciary in India (Supreme Court of India and the Provincial High Courts) touch upon almost every aspect of an individual's life. Several judgements from these courts have a significant bearing on socio-cultural identities of communities. Judicial decisions also affect the economic policies of the country. While judges in the Supreme Court are subject to greater public attention, despite playing a critical role in the life of the nation, there is limited to no data available on the background of Judges in the High Courts. Greater public awareness about our Judges is important for continued public trust in the judicial process. It is also critical as it informs the lawyers as well as the citizenry of potential areas of bias and conflicts of interest. The conceptualization of a dataset on High Court Judges was a response to this clear gap in India's information and judicial architecture.

The logistics of the data collection exercise was supported by Agami, CivicDataLab and the Centre for Public Policy, Law and Good Governance (National Law University Odisha).

We launched the "Summer of Data" campaign to start building this public dataset about 43 indicators related to Judges of the High Courts. In this initiative, more than 30 students collaborated on crowd sourcing data from multiple High Courts, Department of Justice websites, and other publicly available documents on official websites. We then conducted several rounds of verification to test the accuracy of the gathered data points.

## **OVERVIEW OF KHOJ DATASET**

While some information about the Judges is available on the official High Court websites and other official websites, the information is not standardized. Moreover, nuanced details with respect to their education, area of legal expertise and how they entered the service, among other information, is not available.

The KHOJ dataset focuses on four categories of information which have been sourced from openly accessible official sources (High Court Websites, Website of Department of Justice, Ministry of Law & Justice, Government of India, Website of the Indian Supreme Court and Websites of official bodies like tribunals, Lok Sabha etc.)

Background Information - There are 7 variables centered around their names, place of birth, date of birth, gender etc.

Education related information - There are 10 variables through which we have captured their education background, ranging from schooling, university and higher education.

Though we have attempted to capture all the publicly available information, this section is not exhaustive and there may be data points that are not available in the public domain.

Experience before Judicial Appointments - There are 10 variables capturing the professional experiences of judges, including litigation and administrative experience, if they were empanelled by PSUs, banks, statutory bodies or private companies.

Information related to Judicial Appointment(s) - There are 13 variables centred around their appointments, transfers from high courts, appointments in the Supreme Court etc.

The dataset contains 26 files, with 25 files dedicated to each High Court and one integrated master file with details of 2111 judges. The master file features each judge only once, even if they feature many times across different High Courts (which is the case with many judges due to their transfers).

In this brief report, we present some key findings from the KHOJ dataset across 43 parameters, along with some statistics and trends. A state name represents the High Court of that state. Eg - If we write Allahabad, we mean "Allahabad High Court" in the data visualisations. To maintain consistency, we have limited the scope of our research to only official sources of information. The data curation exercise is a volunteer-driven effort and you might come across a data point that is incorrect or not available. We will be working to periodically update this dataset as per the inputs shared by the community.

## **ACKNOWLEDGEMENTS**

We would like to acknowledge the contribution of several people who have contributed to this dataset over the course of one year.

Firstly, we would like to thank the student volunteers who were the original data collectors and analysers in June 2021 - Abhishek Anand, Anaika Martin, Ananya Malaviya, Ankita Kamath, Antara Balaji, Anugra Anna Shaju, Asna D, Biswajeet Mishra, Daaman kaushal, Gokul Suresh Nair, Gunjan Rathore, Harshita Tyagi, Indrasish Majumder, Ishita Khandelwal, Kathi Thriveni, Krishna Nair, Mahi Jaiswal, Manya Thapliyal, Mohnish Solanki, Nikita Bansal, Nisha Munshi, Parv Pancholi, Piyush Sinha, Prakhar Gupta, Rishi Saraf, Shaswata Kapat, Shivam Kumar Mishra, Shivam Patel, Shreyas Rao, Smarak Samal, Smita Gupta, Sneha Rath, Swati Banerjee, Varun Ahuja, Vidit Parmar, Viraj Aditya and Yash Prakash Yadav.

Secondly, the volunteers from TresVista have played a big role in cleaning and standardizing the datasets. Thank you *Danish Ahmed, Sneha Balpande, Sayali Mate, Amit Mane, Pooja Kumari, Shubham Soner, Yash Bhatt, Akash Bodhale, Kunal Bhegade and Ayushi Ninave.* 



bit.ly/khoj-india