Project Intro

Anuhya Nallapati	2021101076
Harshit Aggarwal	2021111015
Shreyas Reddy Palley	2021101016

Overview

Our project involves analysing a judicial dataset, which is a collection of data related to the legal system. This dataset includes information about court cases, judges, and legal proceedings, and it can be used for a variety of purposes such as analysing trends and patterns in legal disputes, developing predictive models to forecast case outcomes, and identifying areas for improvement in the legal system. Some common types of data included in the judicial dataset include case details such as charges, judges involved, date of filing, etc. By leveraging the power of data analysis, judicial datasets can provide valuable insights into the legal system and inform efforts to improve legal processes and outcomes. Our project aims to leverage this data to gain insights into the legal system and inform decision-making processes.

Another interesting area of analysis is comparing the sentencing patterns of male and female judges. By examining the length and severity of punishments given by judges of different genders, we can observe any potential biases in the legal system. This analysis can also be conducted over time to identify any trends in sentencing patterns. Such insights can be valuable for promoting gender equality in the legal system and ensuring fair treatment for all individuals involved in legal disputes.

Columns

Judges Column Descriptions

File	Variable name	Label	Type	Format	Definition	Encoding
1	judges_clean.dta	ddl_judge_id	Unique judge id - created by DDL	float	Judge identifier	Identifier created by DDL. Merge with case dataset using this id, and the judge_cases_merge_key.dta to merge judge-case data.
2	judges_clean.dta	state_code	State code - name in state key	str	Geographic identifier	E-courts identifier
3	judges_clean.dta	dist_code	District code - name in district key	str	Geographic identifier	E-courts identifier
4	judges_clean.dta	court_no	Court number - name in court key	str	Geographic identifier	E-courts identifier
5	judges_clean.dta	judge_position	Formatted judge designation	str	Judge characteristic	Cleaned judge designation (e.g. "Chief Judicial Magistrate")
6	judges_clean.dta	female_judge	Judge gender classification	int	Judge characteristic	1 if judge is female, 0 otherwise, -9998 = unclear
7	judges_clean.dta	start_date	Tenure start date	int	Judge characteristic	Judge tenure start date in court-position

8	judges_clean.dta	end_date	Formatted tenure end date	int	Judge characteristic	Judge tenure end date in court-position

Cases Column Descriptions

File	Variable name	Label	Type	Format	Definition	Encoding
1	cases_[year].dta	ddl_case_id	Case ID - generated by DDL	string	Unique ID	This ID can be used to link with acts_sections.dta & judges_clean.dta
2	cases_[year].dta	year	Case dataset year	int	Dataset characteristic	This is the filing year of almost all cases in this dataset
3	cases_[year].dta	state_code	State code - E-courts, name in state key	str	Geographic identifier	E-courts identifier - link with cases_state_key on state_code, year to get state_name
4	cases_[year].dta	dist_code	District code - E-courts, name in district key	str	Geographic identifier	E-courts identifier - link with cases_district_key on state_code, district_code, year to get district_name
5	cases_[year].dta	court_no	Court number - E-courts, name in court key	str	Geographic identifier	E-courts identifier - link with cases_court_key on state_code, district_code, court_no, year to get court_name
6	cases_[year].dta	cino	Unique CRN number - use this to link to E-courts	str	Case identifier	E-courts unique 16 digit case number: this ID can be used to view case details on E-courts
7	cases_[year].dta	judge_position	Judge position	string	Case type	Cleaned judge designation
8	cases_[year].dta	female_defendant	Defendant gender	int	Case characteristic	Defendant gender classified using a neural network
9	cases_[year].dta	female_adv_def	Defendant advocate gender	int	Case characteristic	Defendant advocate gender classified using a neural network
10	cases_[year].dta	female_adv_pet	Petitioner advocate gender	int	Case characteristic	Petitioner advocate gender classified using a neural network
11	cases_[year].dta	type_name	Case type ID	int	Case characteristic	Case type - link with type_name_key using type_name, year to get string value
12	cases_[year].dta	purpose_name	Case purpose ID	int	Case characteristic	Case purpose - link with purpose_name_key using purpose_name, year to get string value
13	cases_[year].dta	disp_name	Disposition ID	byte	Case characteristic	Case disposition - link with disp_name_key

						using disp_name, year to get string value
14	cases_[year].dta	date_of_filing	Filing date	string	Date	Date when the case was filed
15	cases_[year].dta	date_of_decision	Decision date	string	Date	Date when the case reached a decision
16	cases_[year].dta	date_first_list	First hearing date	string	Date	Date when the case was first heard in court
17	cases_[year].dta	date_last_list	Most recent hearing date	string	Date	Date of most recent hearing
18	cases_[year].dta	date_next_list	Next hearing date	string	Date	Date of next scheduled hearing

Act Sections Column Description

File	Variable name	Label	Type	Format	Definition	Encoding
1	acts_sections.dta	ddl_case_id	Unique case id - generated by DDL	str	Case identifier	Identifier created by DDL
2	acts_sections.dta	act	Act - ID	int	Case characteristic	Can be linked to act_key.dta to extract string values
3	acts_sections.dta	section	Section - ID	float	Case characteristic	Can be linked to section_key.dta to extract string values
4	acts_sections.dta	bailable_ipc	Whether offense is bailable (only available for IPC cases)	byte	Case characteristic	Whether the section the crime was filed under is bailable in the IPC
5	acts_sections.dta	number_sections_ipc	No. of sections filed under (only available for IPC cases)	byte	Case characteristic	Number of sections filed under, if the act was the IPC
6	acts_sections.dta	criminal	Indicator whether criminal case	byte	Case characteristic	Indicator for whether the case is a criminal case

Tasks & Relations

1. Total time duration of the cases
    - The total time for the cases i.e. the difference between the time the case was filled and the time when the last hearing took place.
    - We will analyse this with respect to time i.e. year wise and with respect to the different places in the country.
2. Difference between the previous with next visit dates of courts
    - Court Visits may have next visit dates, a date and time for additional court hearings/visits as necessary for further judgements. This distribution can be caused by a variety of factors.
    - The time after which the judge announces the next hearing date i.e. the difference between the current hearing date and the next hearing date.
    - We will analyse this with respect to time i.e. year wise and with respect to the different places in the country.

3. Distribution of courts per area
    - Making a heat map according to the number of courts per area.
    - This task could be useful to understand the accessibility of the legal system in different areas
4. How many cases per area
    - Counting the number of legal cases within specific geographical areas and comparing them over time.
    - This task could be useful for a variety of purposes, such as identifying areas with higher rates of legal disputes.
5. Distribution of courts per area (point 4) and number of cases per area (point 5)
    - Can be analysed together to determine where additional legal resources might be needed.
6. Time Duration per Judge Position
    - Each Judge position deals with different types of cases. Ex: Chief Justice, High Court judges, Additional District & Session Judge, Chief Judicial Magistrate, etc.. We can compare the length of cases for these positions of Judges.
    - This can be useful to show the difficulty and carefulness of every position of Judges.
7. Males vs Females sentences harshness
    - The length and severity of punishments given by Male and Female judges for all cases
    - This can be shown over time to observe any general trends of sentencing periods.
    - Males and Females can be compared to determine if there is any significant difference between the 2 genders, which could help signify bias in some areas and fields.
8. Number of Female Judges versus time
    - We will analyse the trend of the number of female judges over time.
    - This task could be useful for understanding the progress of gender diversity in the legal system and identifying potential areas for improvement.

We can also analyse a lot of different things and find many more relations. We can add those if we have time later on.