## **SUPREME COURT DATABASE SCHEMA**

## **SUMMARY**

A schema for data from the Supreme Court Database (<u>SCDB</u>). Data was pulled from two different versions of the database:

- Case-Centered Data organized by Issue/Legal Provision
- Justice-Centered Data organized by Issue/Legal Provision

Using both databases allows us to encompass variables specific to the case-centered version and information on supreme court justices and their votes on cases since 1791.

In the schema below, we will use a total of 10 relational tables. Three of those tables will categorize the data into CASES, OUTCOMES, and JUSTICES. All three tables contain caseID, which is the primary reference key in this schema.

- CASES: data on each case, including the petitioner and respondent and the date of the decision
- · OUTCOMES: information on case outcomes useful for analysis like voting distributions and issue involved
- JUSTICES: data pertaining to the court: justices, opinion writers, vote direction (liberal, conservative)

The seven keys are to reference normalized variables. Across all three tables, **15** are normalized. Of those fifteen, eight will be manually decoded during data-preprocessing. Seven will be linked to **key tables**, which include the normalized variable and its corresponding value according to the <u>database codebook</u>. Each key table is linked to one of the three main tables that contain its normalized variable.

## **SCDB SCHEMA**



