Understanding the Structure of the Justice Gap Survey Data

What do the **ROWS** represent?

For **respondent-level** data, each row represents a unique case. However, the structure for the household and individual-level data files is quite different.

For the **household-level** data, each row represents a legal problem (identified in the “PROBNUM” variable), meaning each respondent has 88 rows – one for each of 88 potential legal problems. For example:

- Row 1: Respondent A Problem 1
- Row 2: Respondent A Problem 2
- ... through ...
- Row 87: Respondent A Problem 87
- Row 88: Respondent A Problem 88
- Row 89: Respondent B Problem 1
- Row 90: Respondent B Problem 2
- ... and so on ...

For the **individual-level** data, each row again represents a legal problem, but in this case only problems personally experienced by the respondent are included.

What do the **COLUMNS** represent?

For **respondent-level** data, each column contains a variable describing characteristics of the respondent and her or his household.

For the **household-level** and **individual-level** files, columns contain information about the row’s specific legal problem (identified in “PROBNUM”) and about the respondent’s experience with that specific legal problem. For example, this may include:

- the categorization of the legal problem (“block”),
- who in the household experienced the legal problem (“who_exp”),
- how severe the issue was (“severity”),
- the person from whom legal help was sought, and so forth.

The responses can be said to “cascade” across the columns in that having a response in one column may depend on a response in one or more other columns.