**Examining Statute References in Chevron Case Law**

**Goal:**

The goal of this project was to attempt to quickly point to the relevant statute being interpreted, specifically for level 2 and level 3 analysis. This could have potentially saved hundreds of man-hours of reading cases and manually filling in the relevant statutes. The project tries to examine what sort of relationships exists between the most frequently referenced statute in a judgement and the statute under investigation, particularly for level 2 and level 3 analysis.

**The Data:**

There was a total of 1558 statutes under investigation, of which Chevron deference applied to 1166. For the 1166 statutes, there are a total of 1002 cases under investigation, which means some cases have multiple statutes being interpreted. The cases date from 2003 and 2013. If cases were overturned during that timeframe, both the original case and the overturned cases are included in the dataset. This brings the total number of cases under textual analysis to 1012, the vast majority of which have only 1 statute under examination. However, for the final prediction results, only 402 cases had level 2 or level 3 attributes analyzed.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Min** | **Max** | **Median** | **Mean** | **% of cases that examine more than 1** |
| 1 | 7 | 1 | 1.137 | 8.78% |

Number of statutes under examination in the dataset

Some cases were missing identification and had multiple cases under the same name. In those scenarios, a case was chosen based on the timeframe of the study (2003 – 2013) and the number of references a given case made to “Chevron deference.”

**The Analysis:**

The dataset includes the case, the statute under investigation, and analysis of the investigation, broken down into different levels of attributes (from level 1 to level 5). For example, level 2 indicates the specific subsection of the statute under investigation. The specific definition of a level of analysis can be found in the code book. Not all statutes under investigation have attributes on all levels, although some statutes can have multiple attributes on the same level.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Name** | **Min** | **Max** | **Median** | **Mean** | **% of cases that examine more than 1** |
| Level 1 | 0 | 4 | 1 | .86768 | 4.58% |
| Level 2 | 0 | 7 | 1 | 1.216 | 14.5% |
| Level 3 | 0 | 7 | 1 | 1.0763358 | 6.6% |
| Level 4 | 0 | 4 | 1 | .363867 | 5.08% |
| Level 5 | 1 | 4 | 1 | 1.055 | 3.56% |

Number of attributes cases have on each level of analysis

The main focus of this project was to attempt to quickly identify the level 2 and level 3 attributes of the statute being debated. Level 2 attributes are defined as the specific subsection of the statute being interpreted by the agency. For example, 18 U.S.C. § 923(g)(5)(A) in *10 Ring Precision, Inc. v. Jones.* Level 3 attributes are broader, they are the entire relevant section of the statute being interpreted by the agency. So for the previous example, the level 3 attribute would be 18 U.S.C. § 923.

**The Results:**

The approach was simplistic and the results were lackluster. For each case, all referenced statutes were collected, standardized, and counted. The method of standardization varied between level 2 and level 3 attributes. Then, the most frequently mentioned statute for each level in the case was selected as the statute under interpretation. The table demonstrates the rate of success for this technique in selecting the correct statute under interpretation for specific level, using the analysis of the 402 completed cases.

|  |  |
| --- | --- |
| **Type** | **Accuracy (Percent Correct)** |
| Level 2 | 49.2% |
| Level 3 | 54.65% |

Results from simplistic approach

**The Conclusion:**

The results were not great, but roughly half of the statutes under discussion in a judgement are also the most frequently mentioned statutes in the judgement. Level 3 is slightly more accurate for one pretty obvious reason; it requires less specificity than level 2 attributes.

One major issue faced was the inconsistency in data collection and the creation of a standard format for analysis. Judges have their own styles, so it is difficult to collect and count all statutes referenced in a judgement, specifically when attempting to differentiate between specific subsections and more general sections. Another issue was being unable to accurately account for cases with multiple statutes under interpretation.

Improving on these issues would yield more accurate results, but it is clear that frequency alone is an insufficient predictor.

Files and Code Descriptions

**Articles Directory:**

Contains three files for each Chevron case

1. <Case Id>.txt – contains the text of the judgement
2. <Case Id>\_meta.txt – contains a list of citations that the case makes
3. <Case Id>\_struct.txt – contains the structure of the judgement

**raw\_text Directory:**

Contains raw text (<Case Id>.txt) of each judgement

**meta\_text Directory:**

Contains meta text (<Case Id>\_meta.txt) of each judgement

**statutes Directory:**

Contains the statute frequency count for level 2 (l2) and level 3 (l3) analysis for each case

<Case Id>.txt-l2.txt indicates it is the level 2 statute counts

<Case Id>.txt-l3.txt indicates it is the level 3 statute counts

<Case Id>--1.txt-l3.txt indicates it is the level 3 statute counts of an overturned case

**sorted\_analyzed\_cases.csv:**

Contains the current list of cases that have been analyzed by the team, sorted and with

their respective analysis

**sorted\_final\_total\_list.txt:**

Contains the current list of all cases that have been analyzed by the team, sorted

**grab\_google\_scholar.py:**

For every case in the list of cases, extracts the case id, searches for it on Google scholar,

and extract the text, meta text, and struct text

**sort\_data\_sets.py:**

Sorts the cases alphabetically and counts which cases have been analyzed

on the l2 or l3 level

**law\_analysis.py:**

Contains a number of functions that analyze the law texts:

flatten() – flattens the Article directory into a new directory

average\_length() – gets the length of the shortest judgement as well as the average

length of the judgements and the median length of them

footnotes() – operates on meta text files and examines the number of footnotes,

references, and outside citations in a text, as well as some

meaningful statisitics

find\_statutes() – attempts to find all the l2 and l3 statutes referenced in the raw

texts of the judgements. Stores them in the relevant statute file

and returns some meaningful statistics.

**case\_analysis\_parser.py:**

Parses the case analysis csv, and returns meaningful statistics on the number of

cases that have multiple statutes under analysis as well as the statistics for each particular

level of analysis.

**relevant\_statute\_analyzer:**

Analyzes whether the most frequently referenced statute is the one under inspection.

Done for level 2 and level 3 analysis.