Report 5: Judges & Sentences

A casual report

Ricky Heinrich for Mélanie Méthot

2023-08-03

Contents

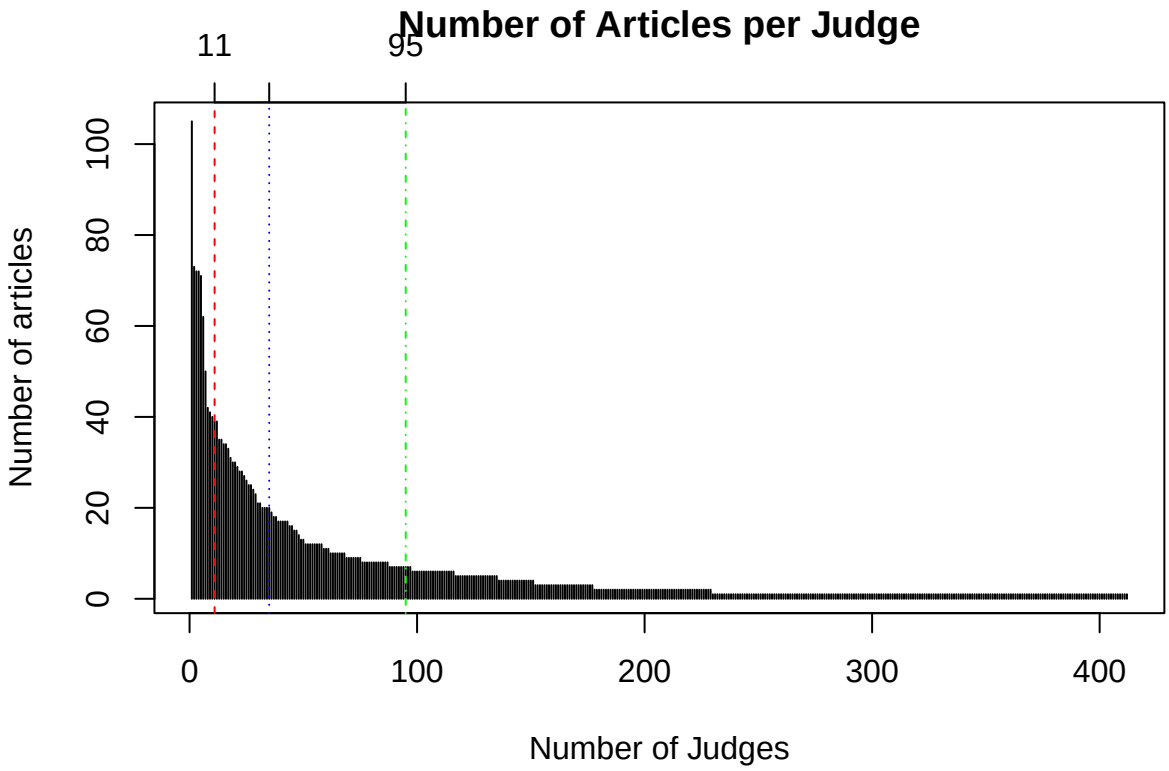
Judges

2637 out of the 3070 (85.9%) cases across Australia have a Judge recorded. There are 1026 unique inputs; way too many to go through manually. A lot of cleaning was done to end up with just one name, the details of which are in the code. In cases where multiple names were present, the last name mentionned was kept. (An exeption being where I caught one time the first judge had (supreme court) after it whereas the second judge had a different court, and I assumed that supreme court would be the one we should keep). Once 'clean', we end up with 413 unique judges. Some care was taken to fix obvious typos, but there may remain cases where a judge is meant to be the same but has different spelling. **Let me know if you would like a list of these to review**.

There are 20 judges that show up over 30 times. They are shown in the following table. These judges account for for nearly one third of all cases (0.32), despite only being 4.85% of the judges. I did the rest of the report before coming back here and adding the year ranges for each judgse. I had previously made the unconscious assumption that each judge name corresponded to the same judge, which may not be the case. For example I don't think it's likely that Douglas practiced from 1864 to 1948, a range of 84 years ... This means that I will have to go back and find a way while cleaning to distinguish between different judges with the same name. Maybe I'll separate by state first under the assumption that they never move and that there is only one of that name per state... I'll think about it and try things out, but it'll take a long time so I will send this over as is for now.

	judge_simple	Freq	first_year	last_year	range
1	markell	105	1929	1949	20
2	lowe	73	1927	1947	20
3	duffy	72	1933	1947	14
4	macfarlan	72	1922	1947	25
5	martin	71	1935	1947	12
6	barton	62	1933	1949	16
7	o'bryan	50	1939	1947	8
8	cohen	42	1895	1929	34
9	shortland	41	1942	1948	6
10	curlewis	40	1925	1948	23
11	backhouse	39	1889	1921	32
12	mann	39	1919	1943	24
13	holt	35	1902	1949	47
14	scholes	35	1913	1928	15
15	douglas	34	1864	1948	84
16	white	34	1921	1938	17
17	hodges	33	1890	1919	29
18	dwyer	31	1916	1957	41
19	armstrong	30	1908	1932	24
20	edwards	30	1916	1936	20

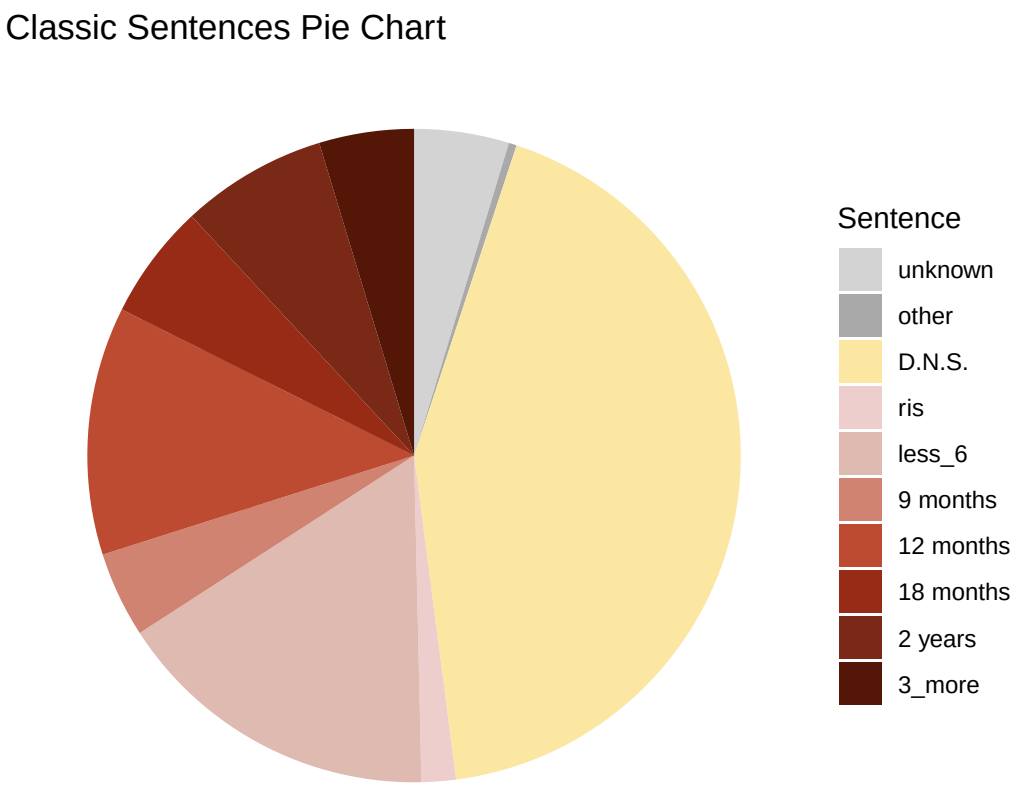
The following graph shows the distribution of number of articles per judge, sorted in a decreasing manner. The lines indicate select percentiles: 25% of articles are attributed to top 11 judges, 50% to top 35 (I'm not sure why the number isn't showing up above the blue line), and 75% to top 95. We see that more than half of the judges have less than 2 articles associated with them, and nearly half have only 1.



Sentences

3066 out of the 3070 (99.87%) cases across Australia have an entry in the 'Sentence' column ('?' was reported when unknown). I've cleaned the sentences down to the same categories as we've dealt with before. It's worth noting that I put sentences that were 1 day or less under 'rising of the court'. Months were rounded as I pleased - details are in the code and available upon request. The counts for each category are shown in the following table. A classic pie chart for the whole dataset follows it, where I still need to figure out how to add value labels.

Var1	Freq
unknown	144
other	12
D.N.S.	1316
ris	52
less_6	498
9 months	130
12 months	377
18 months	175
2 years	222
3_more	144

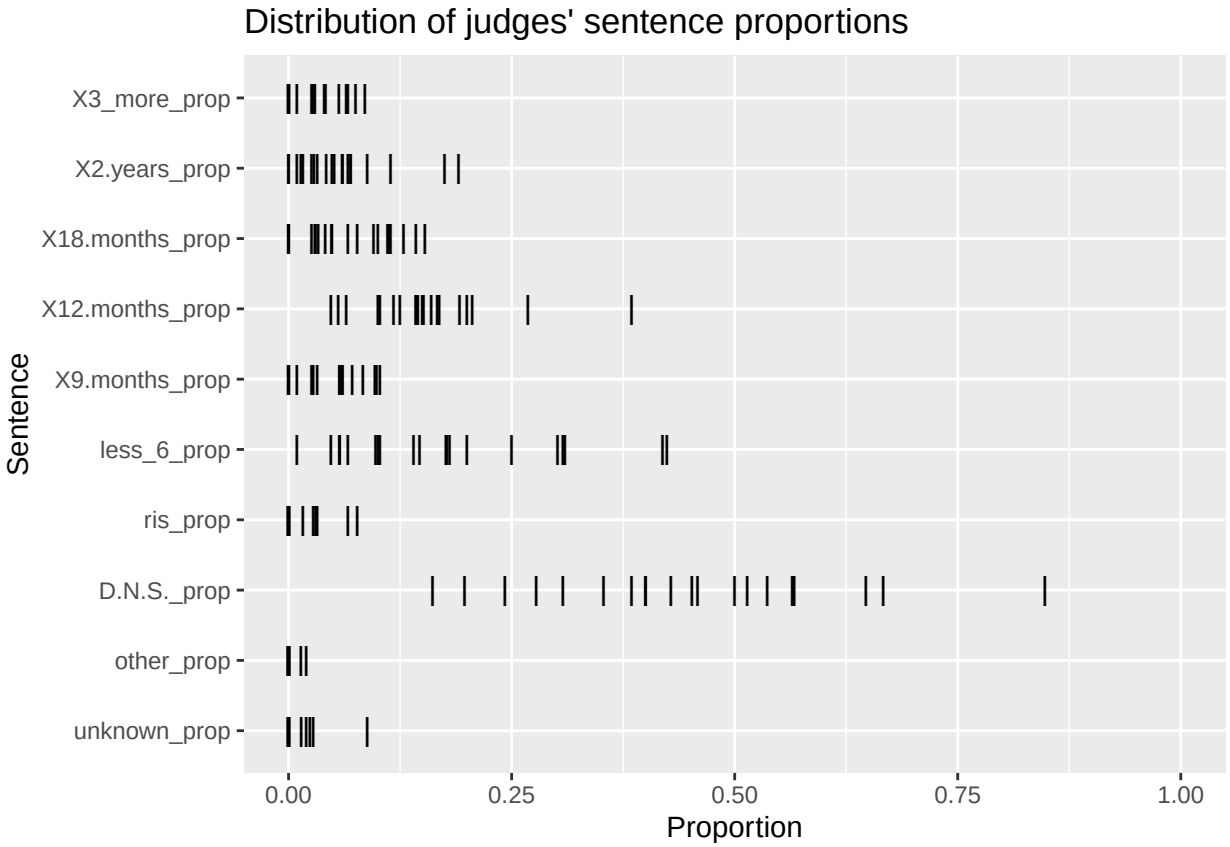


Intersection

For my intersection analysis, I am only including the judges that show up 30 times and over, as 30 is the magic number that makes a sample size 'large' and statistically valid.

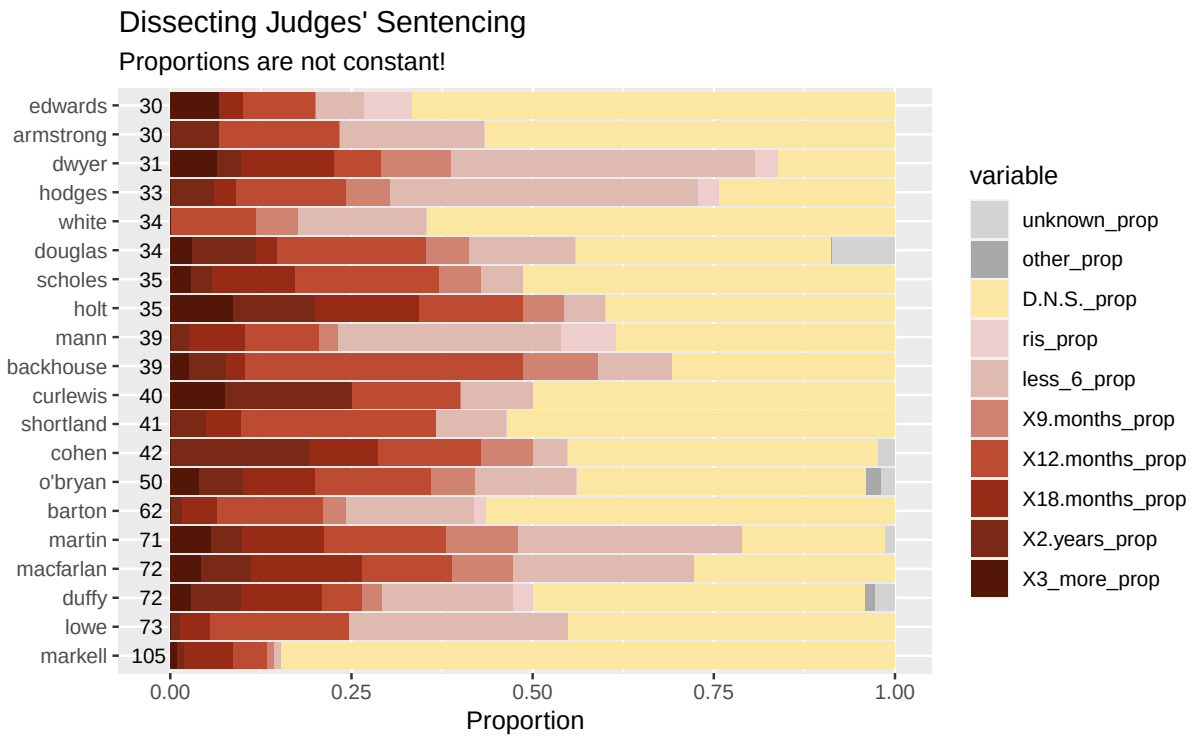
Tick plot

The following plot shows the distribution of judges' sentence proportions. For example, you see that a sentence of 3 years or more (top row) never exceeds 10% of a judge's sentences, whereas the range for other sentences varies. Most of the 'unknown' sentence proportions lie near 0%, except for one case that lies nearer 10%. The range of D.N.S. sentences proportions varies much more: the first tick (belonging to Dwyer in this case) mark lies at 16%, whereas the last one (Markell's) at 85%. This means that 85% of Markell's sentencing are D.N.S, but only 16% of Dwyer's. You could be inclined to think that this may arguably be in part due to the number of cases each judge is mentioned in: Markell has 104, whereas Dwyer only had 31. But as the plot upcoming shows, that is not the case.



Color barplot

The next plot dissects each Judge's sentencing 'habits'. We see that proportions are not constant, and don't seem to follow a trend regarding the number of articles a judge is mentionned in, as the bars are sorted and the color lenghts seem to flip flop. You can see that although Edwards and Dwyer have similar number of articles mentioning them (30 and 31, respectively), their proportions are pretty dissimilar: Edwards has a much larger proportion of D.N.S. (67%), whereas Dwyer's is much smaller. Another difference is that Dwyer's proportion of sentences that are 6 months or less is much larger. It seems like their proportions of sentences that are greater than 3 years are about the same: around 6.5%. Although Lowe and Barton have 73 and 62 articles each, they both never gives the harshest sentence, which is also the case with Cohen, Shorthand, Mann, White, and Hodge. White's colorbar appears particularly light, never sentencing over 1 year. Let me know if you'd like more analysis regarding this graph or need more help to interpret it. **Let me know if you'd like a table with the exact values**.



Sentences in Time

Count

Next we'll look into the time of these sentences: are there trends through the years? In the following plot, we can see where the cases for a certain sentence are dense, and where they are sparse. We can see that the 'rising of the court' are somewhat sparse throughout, and that the first instance is near 1862. A disadvantage of this plot is that the areas may become too dense, like in the case of D.N.S. where you can't see how many cases are present as the ticks all merge together.

Figure 1 on the following page shows the line plot, but a failure of it is that it doesn't connect the line down to 0 when there are no values for a given year, but rather connects the points present across the years. This is very apparent in the 'other' plot; it seems like there is a constant value of '1' across all the years, but we know from the plot above that that is not the case. You can see that the D.N.S. peak aligns with the area most dense the in the plot above.

Proportions

Figure 2 upcoming shows the proportion of a sentence in a given year. You can read it as, for example, 100% of cases before 1930 were assigned a sentence of 3 years or more; but remember the sparsity of cases in those years. 1/1 is 100%. You can see how the proportion of D.N.S. sentences rises through the years, from 1890-1940.

In the plot after, Figure 3, the proportion for each decade is plotted. You can see how some proportions like 'other' and 'ris' (rising of the court) stay mostly low and constant, whereas 'D.N.S.' is pretty varied. There doesn't seem to be obvious trends. The proportion of sentences that are less than 6 months increases up to 25% in the 1890s, and appears to be around 13% in 1920s-1940s. The proportion of sentences that are more than 3 years appear to decrease from 15% in 1830s down to 1% in 1920-1940s.

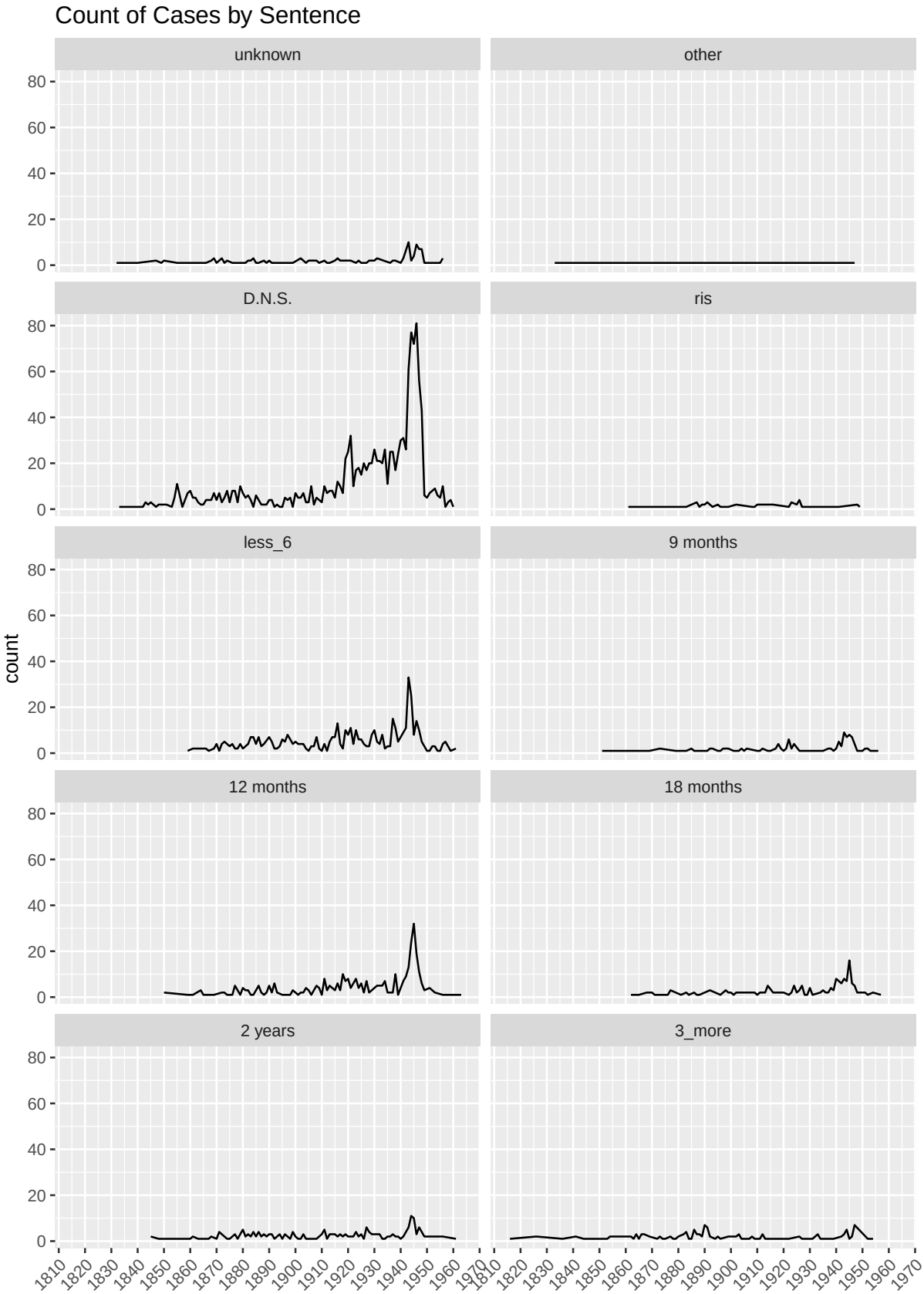


Figure 1: Sentences counts through the years

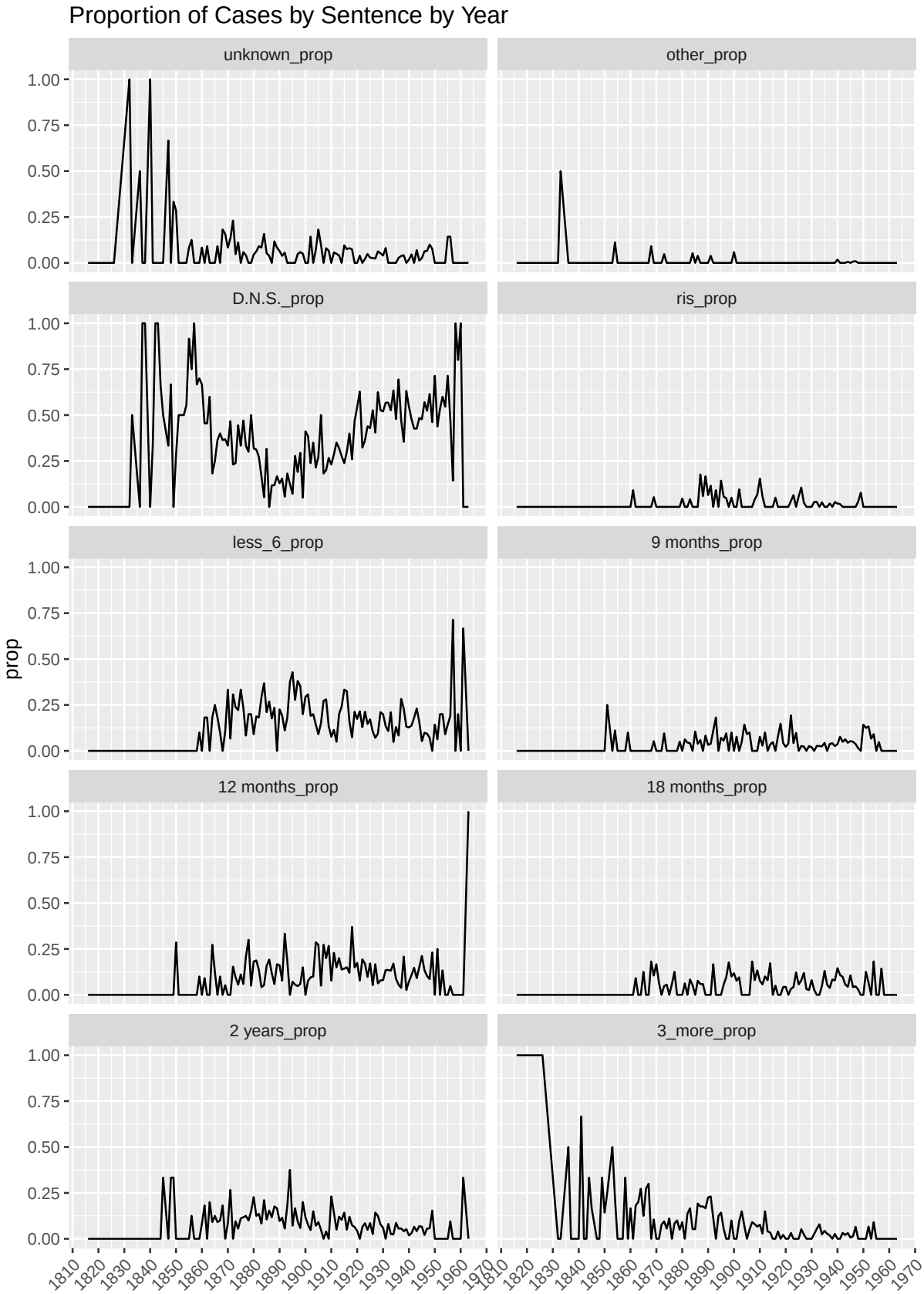


Figure 2: Sentences proportions through the years

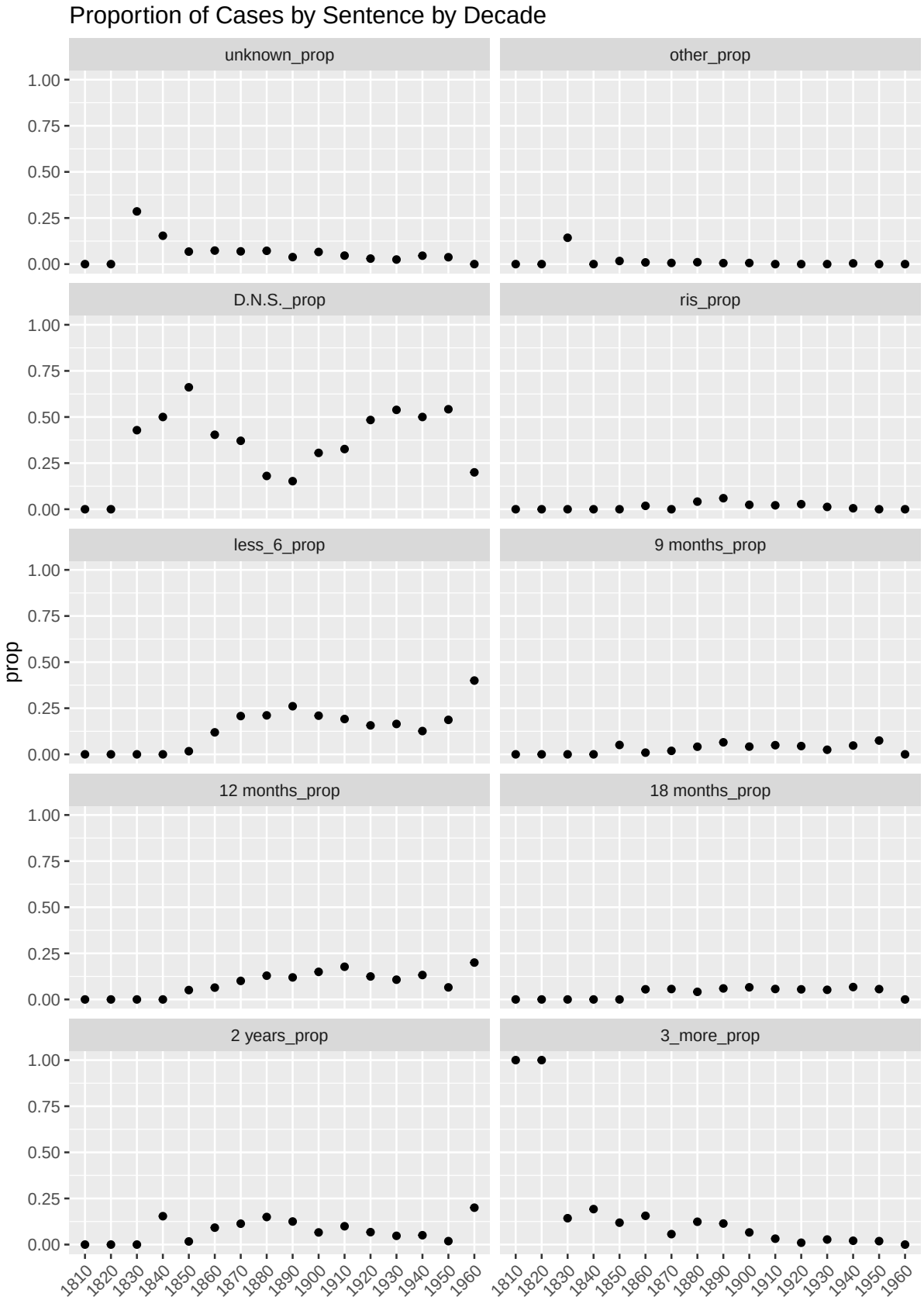


Figure 3: Sentences proportions through the decades