Report 7: How does media coverage change through time?
A casual report

Ricky Heinrich for Mélanie Méthot

2023-09-02

Contents

Overview, cleaning

It looks like there are 747 cases that don't have the number of articles recorded. The cases with empty cells for number of articles occur across five of the states, shown in the table below. A quick review in the original spreadsheet of Queensland, Western Australia, and NSW shows that at least some of these cases do have links. The lack of value in this cell thus comes from a deficiency in the transcriber. The two cases in South Australia don't have any links, and the first few in Victoria also don't, but I didn't look at all of the cases.

Var1	Freq
NSW	535
Queensland	164
South_Australia	2
Victoria	23
Western_Australia	23

I don't have a quick an easy way right now to check the number of links and input a value, although I am sure it is doable. I may try to think about it and find a way later.

There are 647 cases where no entry was recorded for "# of articles" AND there was nothing recorded in "link1". I am adding the entry '0' in those cases.

There are 61 cases where there's no entry in the number of articles column, but there is one link or what seems like an article title. 39 cases had something along the lines of "yorick's info" but didn't include any links, so for now have remained as having no entry for # of articles. Since there there isn't an easy straightforward way to count how many links there are per cases, and a quick overview showed that some included more than 5 links, I will not change these values to '1' article as it would skew the results. I will instead just not include them in the analysis, and leave them as 'no entry', which is valid since it is a small proportion of the data. Maybe later I'll go back and take the time to manually add the number of links since I guess it would be manageable.

There are a few cases where the entry in "# of articles" is not just a numerical digit, shown in the table below. For those that have a variation of 'more than x', I'll just put in the x. For 'cannot find articles' and 'none' I will put '0'.
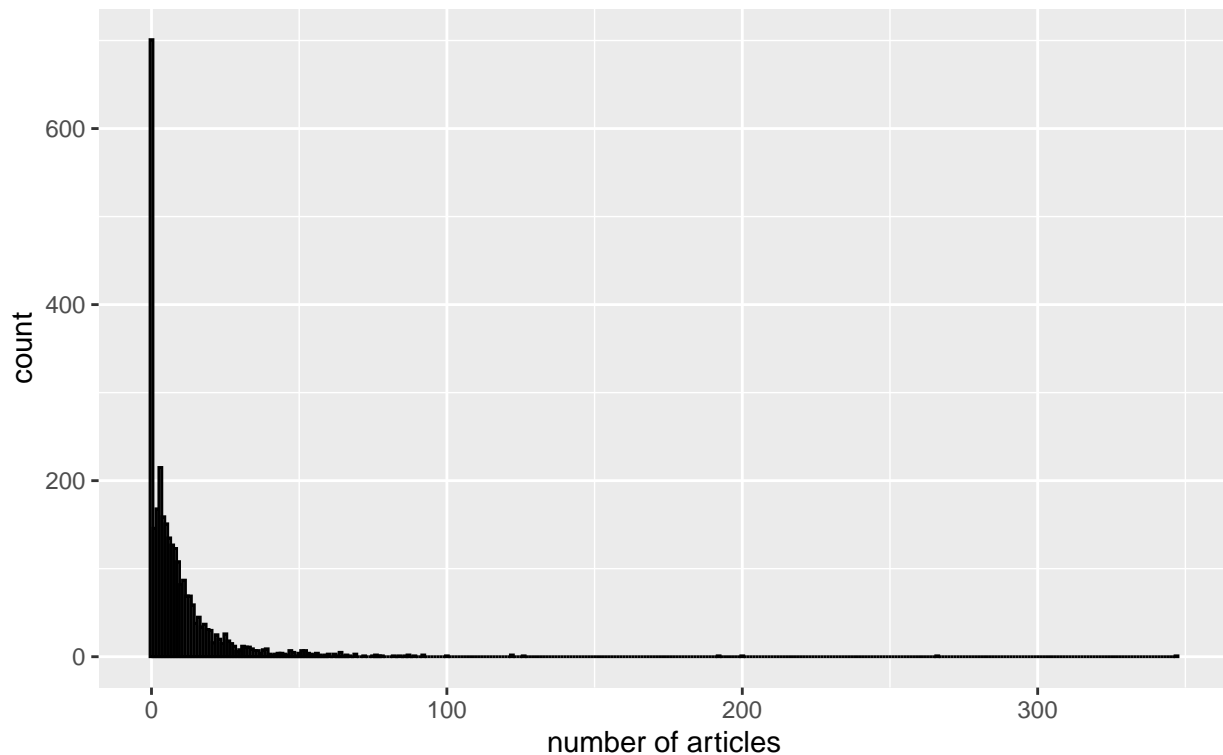
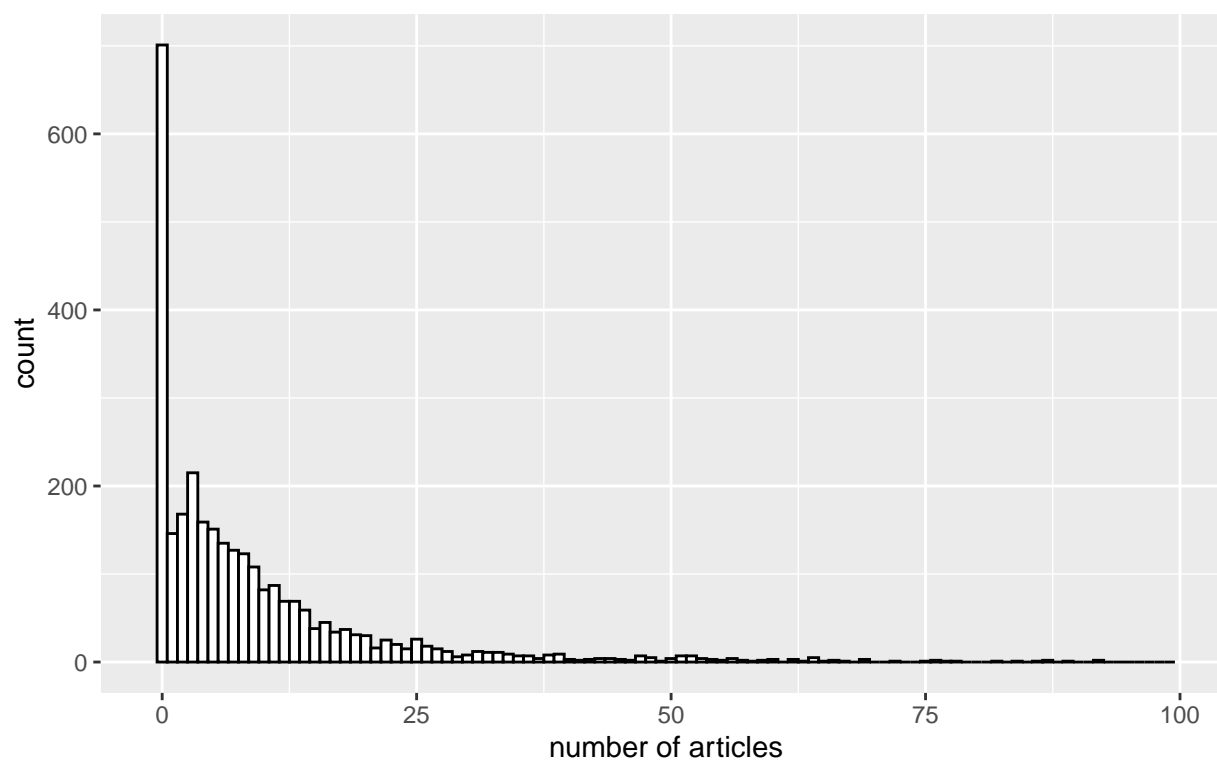| Var1	Freq
50+	1
6 (but 20 for his youth offences!) will also ask for a divorce later on	1
9 (2 in 1862 and 7 in 1867)	1
at least 12	1
cannot find articles on Trove currently	1
MORE THAN 122	1
more than 92	1
none	9

Count Bar Plot

The maximum number of articles is 347. The plot below shows the count of article numbers, and the two after show it zoomed in. We see that the most occurring is '0', which occurs 701 times. These includes all the times there was no entry in "# of articles" and no entry in "link1". There's somewhat of a peak at 3 articles, then a somewhat gradual decline, with small peaks at 11, 16, 17, 22, 25, ... these peaks are relatively small differences: the number of cases with 10 articles is 82, and that with 11 is 87.

I would subjectively suggest the following categories: 0, 1-3, 4-9, 10-15, 16-27, 28-69, and 70+.
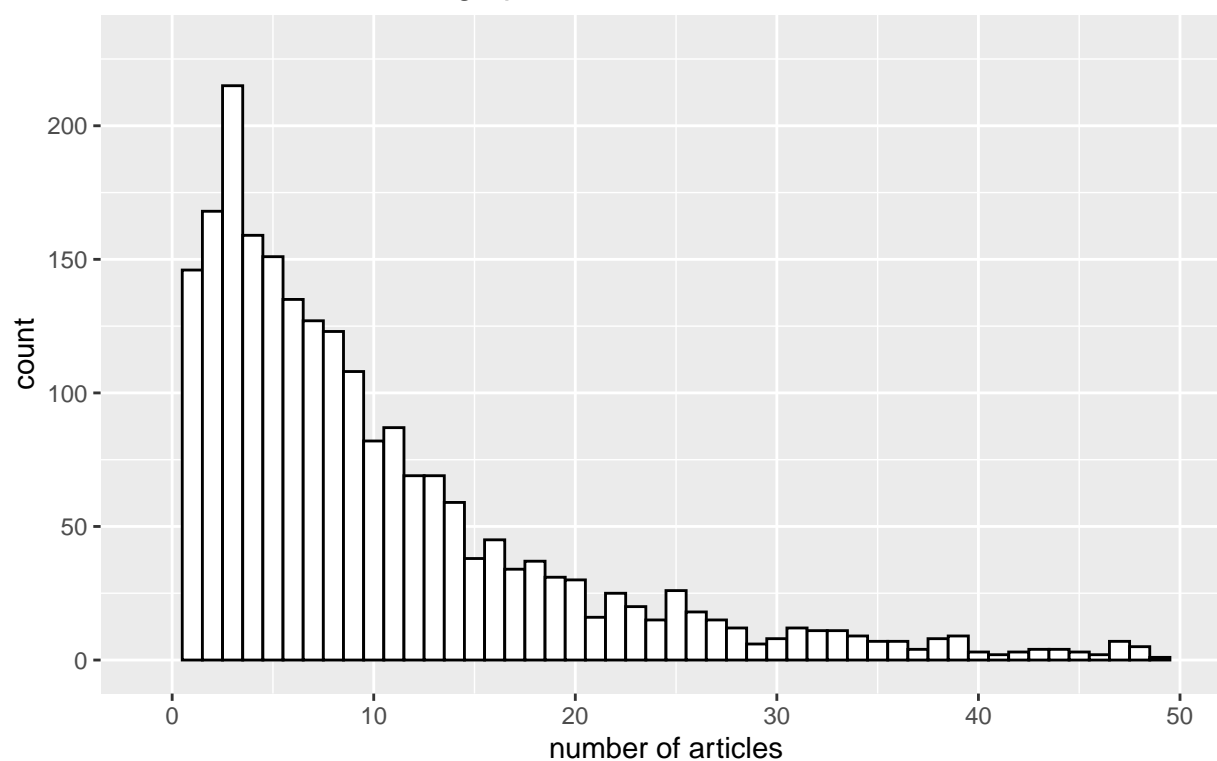

Count of Media Coverage per Case

Count of Media Coverage per Case

Count of Media Coverage per Case

Here is the exact numbers in case you wanted to refer to them directly; every second row is the count.

```
##
##    0    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16   17   18   19
##  701  146  168  215  159  151  135  127  123  108   82   87   69   69   59   38   45   34   37   31
##   20   21   22   23   24   25   26   27   28   29   30   31   32   33   34   35   36   37   38   39
##   30   16   25   20   15   26   18   15   12    6    8   12   11   11    9    7    7    4    8    9
##   40   41   42   43   44   45   46   47   48   49   50   51   52   53   54   55   56   57   58   59
##    3    2    3    4    4    3    2    7    5    1    4    7    7    4    3    2    4    2    1    2
##   60   62   63   64   65   66   67   69   72   75   76   77   78   82   84   86   87   89   92  100
##    3    3    1    5    1    2    1    3    1    1    2    1    1    1    1    1    2    1    2    1
##  122  126  192  200  266  347
##    2    1    1    1    1    1
```
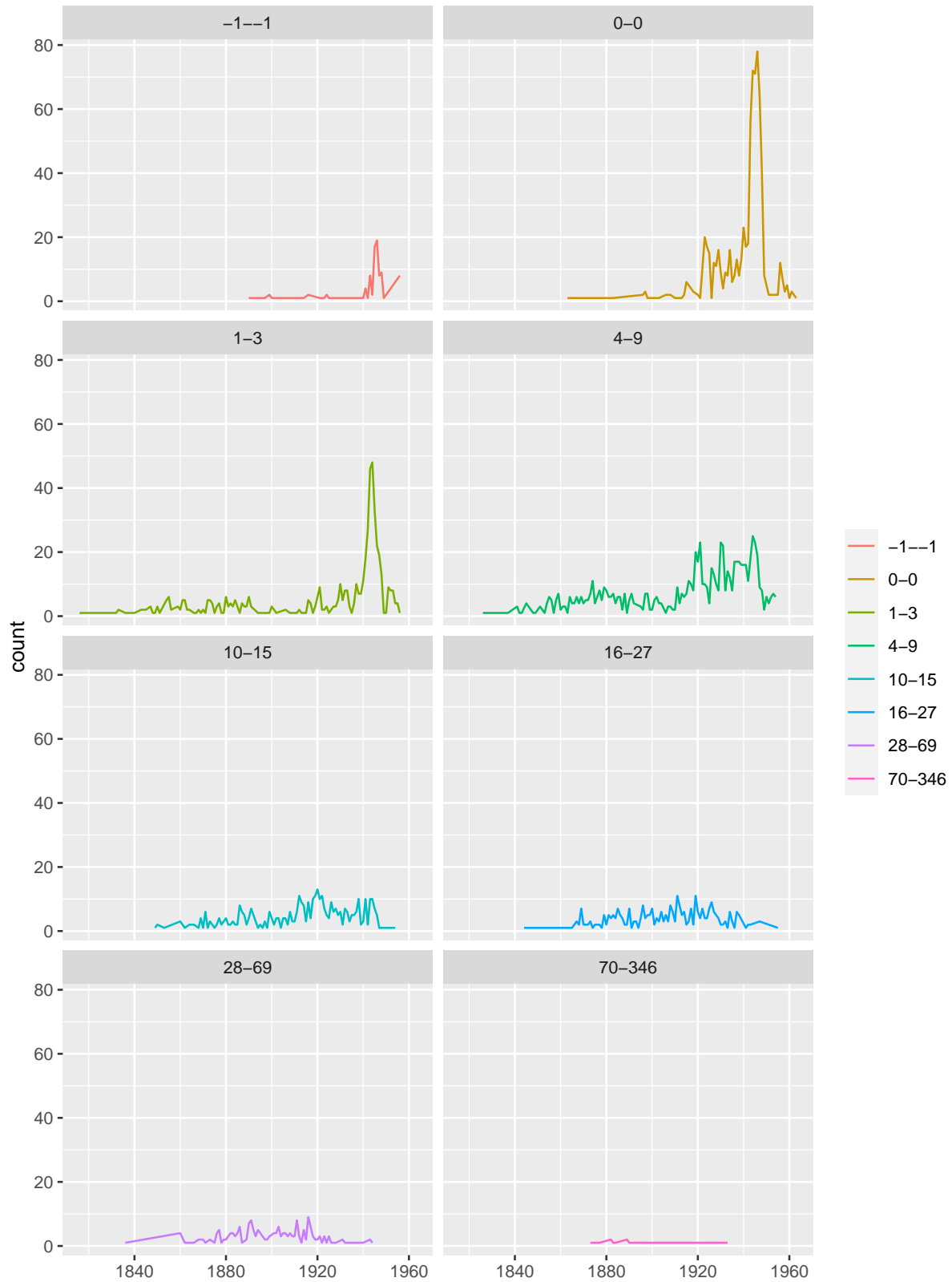
Coverage Through Time

Next we have these groups in time. We see that some groups appear to have smaller ranges in variations throughout the years, like cases with 10-15 and 16-27 articles. As expected, cases with 69+ articles do not occur very often. They don't appear to be clustered either. (keep in mind the dates of the articles themselves may be wide ranging, the date plotted here is the 'indictment' main date associated with the case). There appears to be only two years where there were multiple (ie 2) cases that had more than 70 articles; all the others occured in separate years. The biggest trend is number of cases with 0 articles (same warning as above - may be due to no actual articles, or failure to find). There is a large peak during WWII. Overall more cases with less media coverage (0-9 articles) occur during WWII. The point plot below the line plot is to show exactly where the cases occur, as the line plot doesn't go down to 0 when there are none in that year.
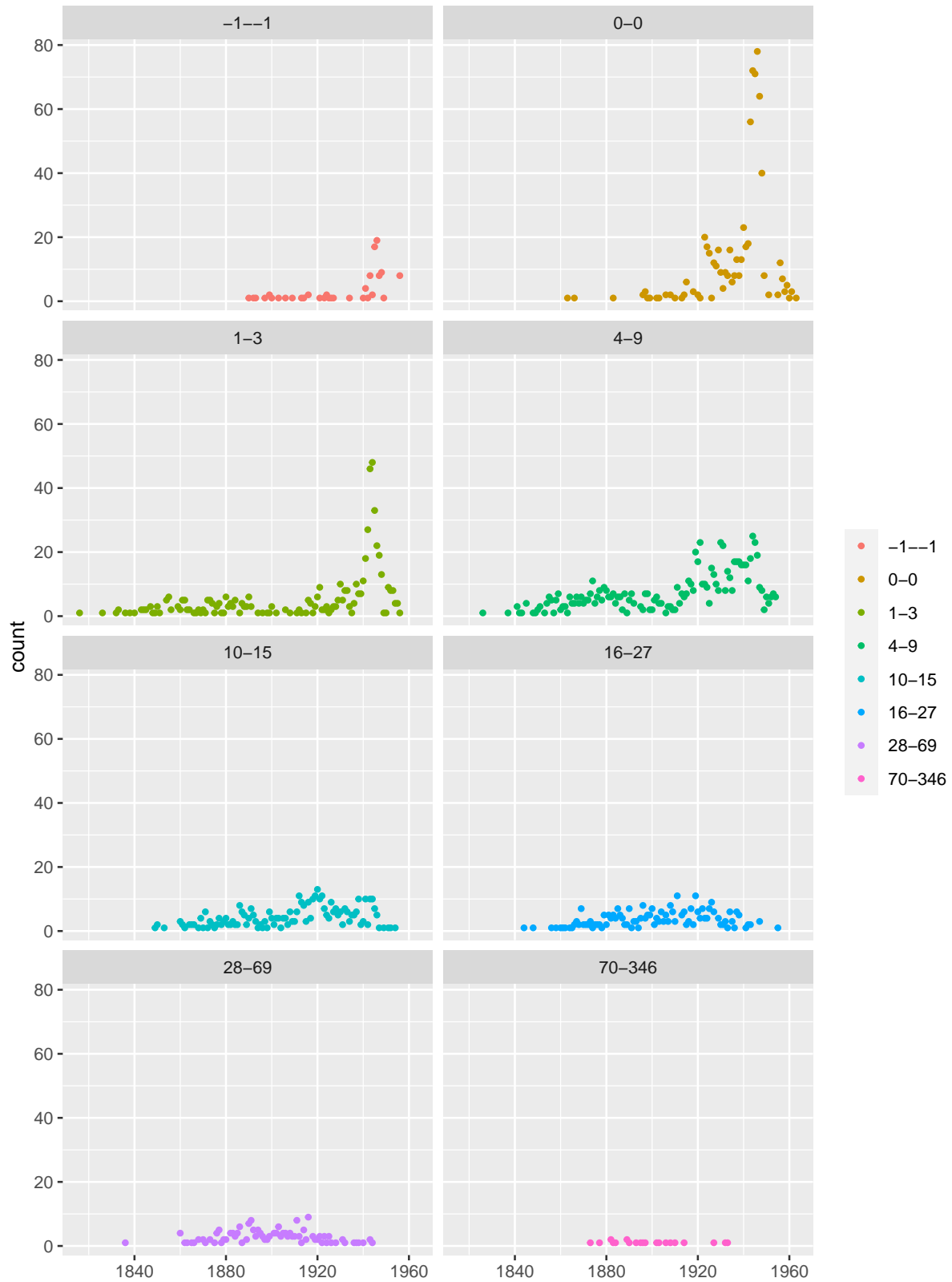
Count by 'number of articles' groups through time

−1−−1 represents either yorick cases or cases with links but no entry
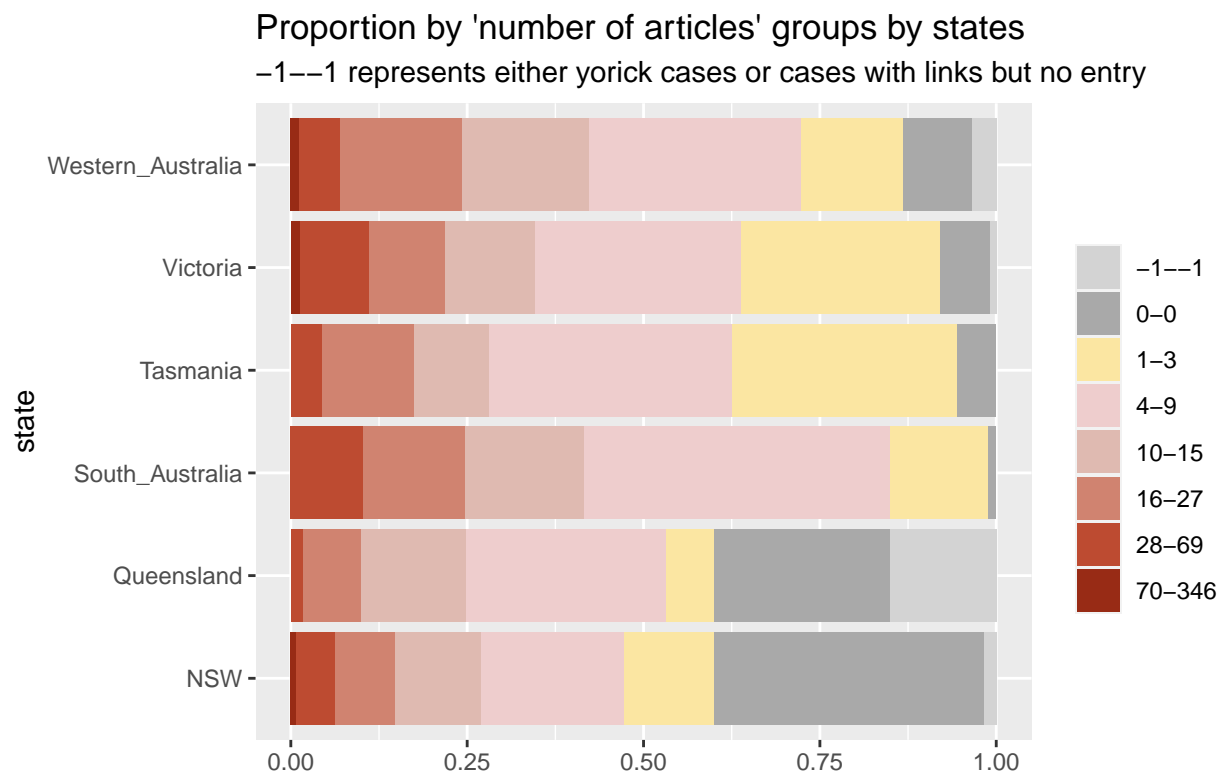
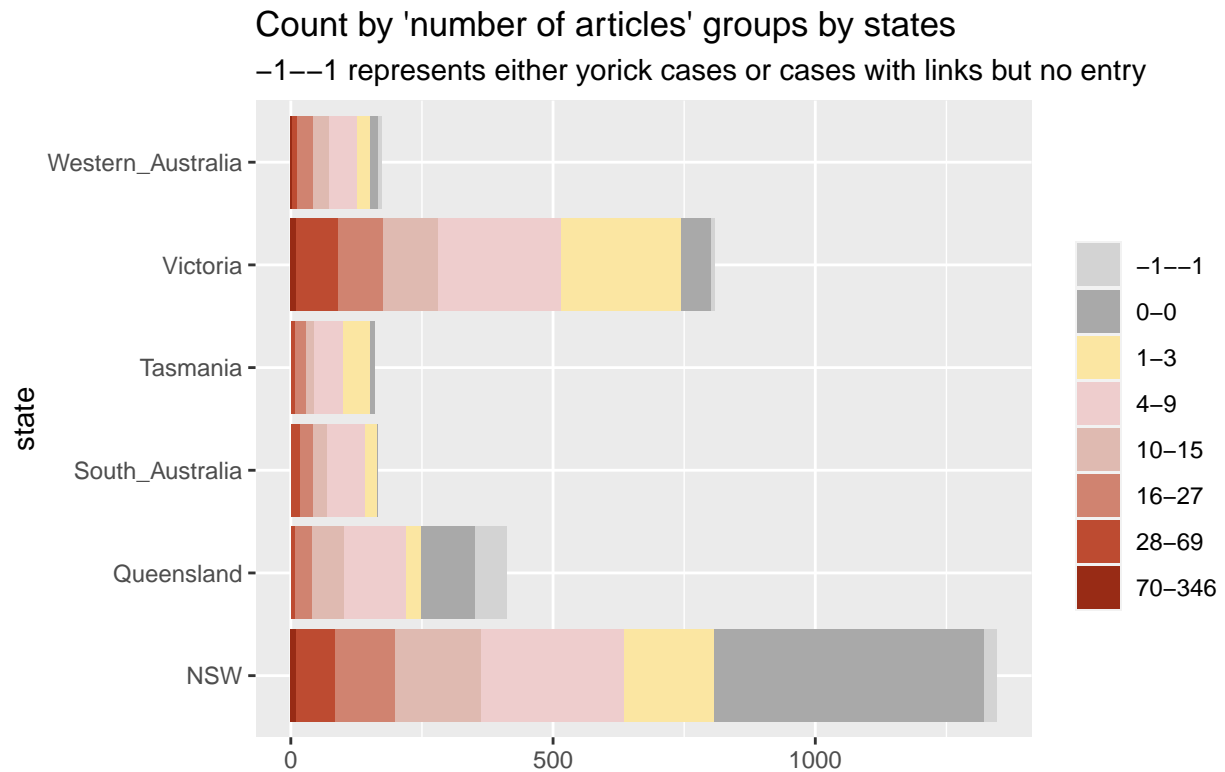# Count by 'number of articles' groups through time

−1−−1 represents either yorick cases or cases with links but no entry

Coverage by State

Count by 'number of articles' groups by states
−1−−1 represents either yorick cases or cases with links but no entry

Proportion by 'number of articles' groups by states
−1−−1 represents either yorick cases or cases with links but no entry

The exact numbers are shown in the table below. You can see that the 'extreme' coverage never occurs in Queensland, SA, and Tasmania.

	-1–1	0-0	1-3	4-9	10-15	16-27	28-69	70-346
NSW	25	513	173	272	163	116	73	10
Queensland	62	103	28	117	61	34	7	0
South_Australia	0	2	23	72	28	24	17	0
Tasmania	0	9	51	55	17	21	7	0
Victoria	7	57	229	235	104	87	79	10
Western_Australia	6	17	25	52	31	30	10	2

The following shows the proportions for easier quantifying. For example, the value intersecting 0-0 and NSW 0.381 means that 38.1% of NSW's cases have article counts of 0.

	-1–1	0-0	1-3	4-9	10-15	16-27	28-69	70-346
NSW	0.019	0.381	0.129	0.202	0.121	0.086	0.054	0.007
Queensland	0.150	0.250	0.068	0.284	0.148	0.083	0.017	0.000
South_Australia	0.000	0.012	0.139	0.434	0.169	0.145	0.102	0.000
Tasmania	0.000	0.056	0.319	0.344	0.106	0.131	0.044	0.000
Victoria	0.009	0.071	0.283	0.291	0.129	0.108	0.098	0.012
Western_Australia	0.035	0.098	0.145	0.301	0.179	0.173	0.058	0.012