

Chen and Ash

Final Version of Figures

Nov 22nd 2018

This markdown is for Chen and Ash figures.

Chen and Ash 1

K-means clustering with outlier detection (ignore code output)

Beginning k-means-- clustering

.....

Clustering complete in 24 iterations.

k = 14 l = 200 Cluster sizes: 80 44 55 79 56 37 43 120 106 51 20
(between cluster sum sqr / total sum sqr = 96.60315 %)

Centroids:

	x	y
[1,]	19.501945	-3.180028
[2,]	-2.807261	26.315234
[3,]	6.411435	19.824719
[4,]	7.048351	31.347035
[5,]	-10.126216	13.388597
[6,]	-17.661710	23.608754
[7,]	-10.713017	-8.491994
[8,]	14.717682	-19.232978
[9,]	-23.502363	-1.287189
[10,]	-2.263348	-17.384235
[11,]	-29.859900	38.031529
[12,]	3.114528	-1.186931
[13,]	-7.921285	-24.861331
[14,]	13.869833	8.101443

Outliers:

A tibble: 200 x 2

	x	y
	<dbl>	<dbl>
1	-27.5	-18.9
2	7.09	-35.0
3	32.3	-14.6
4	31.3	-15.0
5	32.0	-14.1
6	-37.0	23.0
7	17.5	-35.6
8	-33.9	21.2
9	-23.3	-18.5

```
10 -36.2    23.4
# ... with 190 more rows
```

Available components:

[1]	"k"	"l"	"C"
[4]	"C_sizes"	"C_ss"	"L"
[7]	"L_dist_sqr"	"L_index"	"XC_dist_sqr_assign"
[10]	"within_ss"	"between_ss"	"tot_ss"
[13]	"iterations"		

figure 1 (sampled half the data set ($n = 1089$) and using different point shapes)

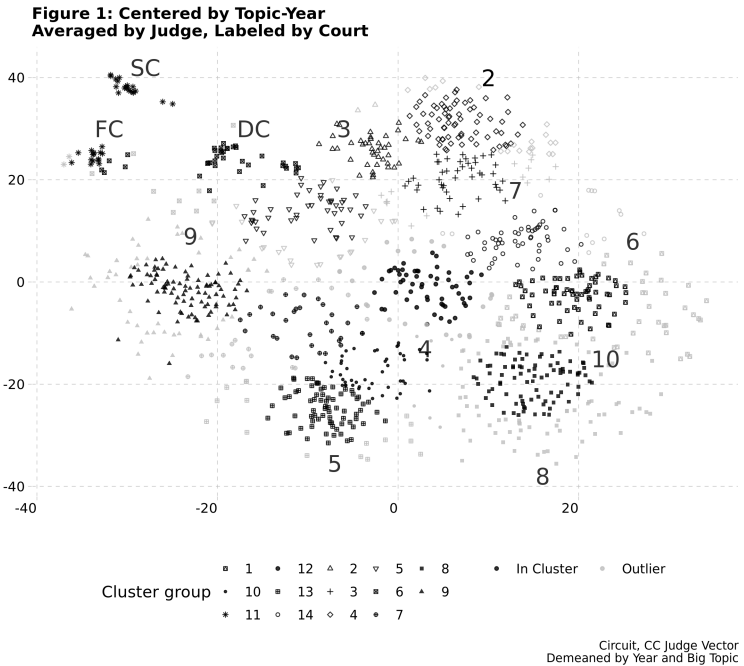


figure 1 (sampled half the data set ($n = 1089$) and using different point shapes)

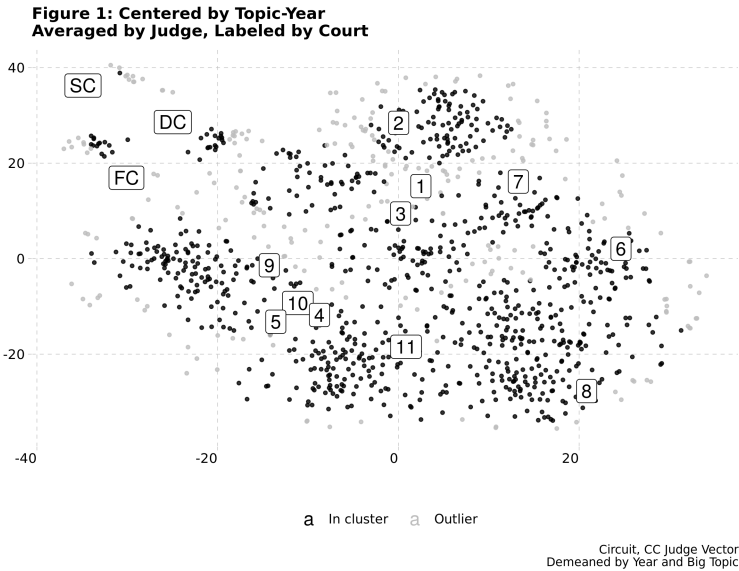
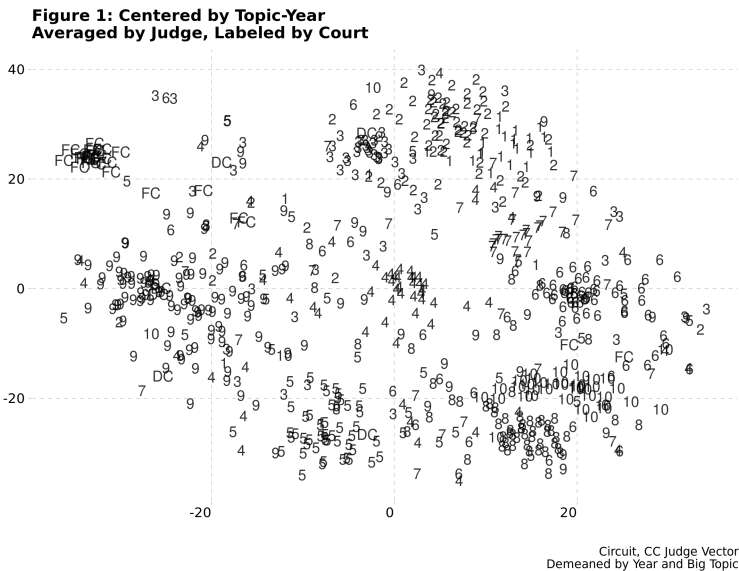


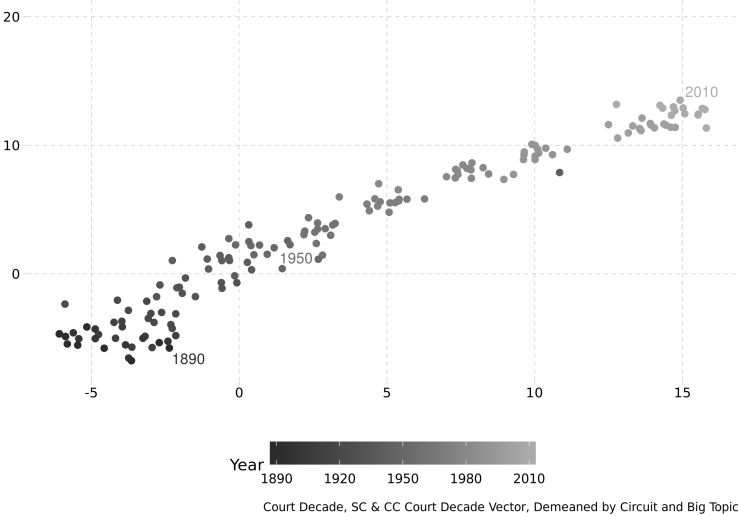
figure 1 (sampled half the data set ($n = 1089$) no clustering



Chen and Ash 2

- [1] 1980
- [1] 1990
- [1] 2000
- [1] 1970
- [1] 2010
- [1] 1960
- [1] 1920
- [1] 1910
- [1] 1930
- [1] 1900
- [1] 1950
- [1] 1940
- [1] 1890

Figure 2.
Centered by Court Topic, Averaged by Court-Year, Labeled by Decade



Chen and Ash 3

Figure 3.
Centered by Judge-Year, Averaged by Topic-Year, Labeled by Topic

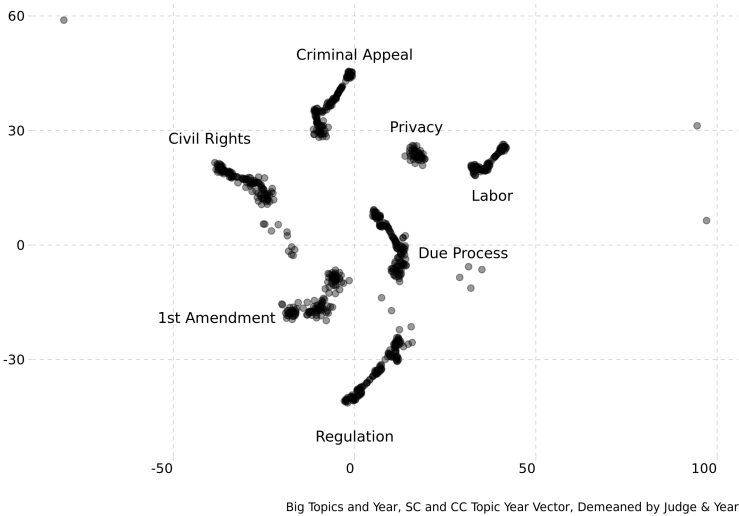
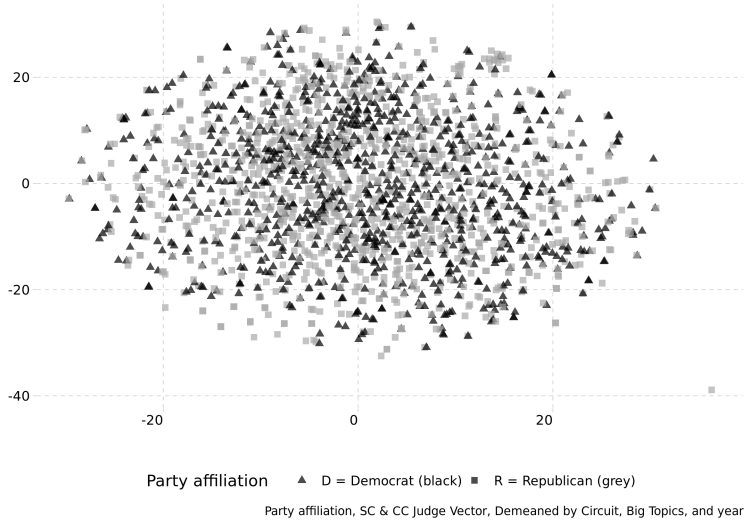


Figure 4.
Centered by Court-Topic-Year, Averaged by Judge, Labeled by Political Party



Chen and Ash 4 (black and white dots with black outline)

