Qualitative to Quantitative Sub Theme Model

Aaron Quinton, Ayla Pearson, Fan Nie
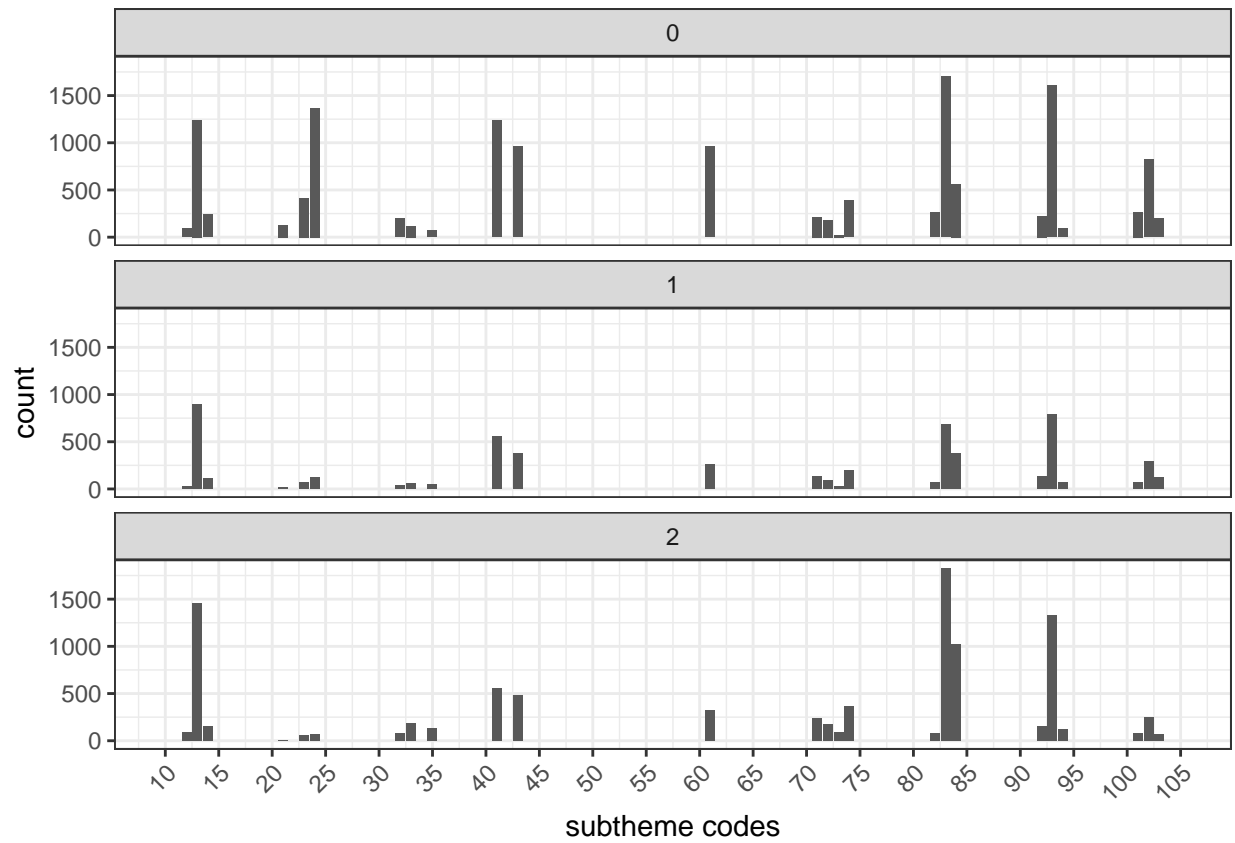
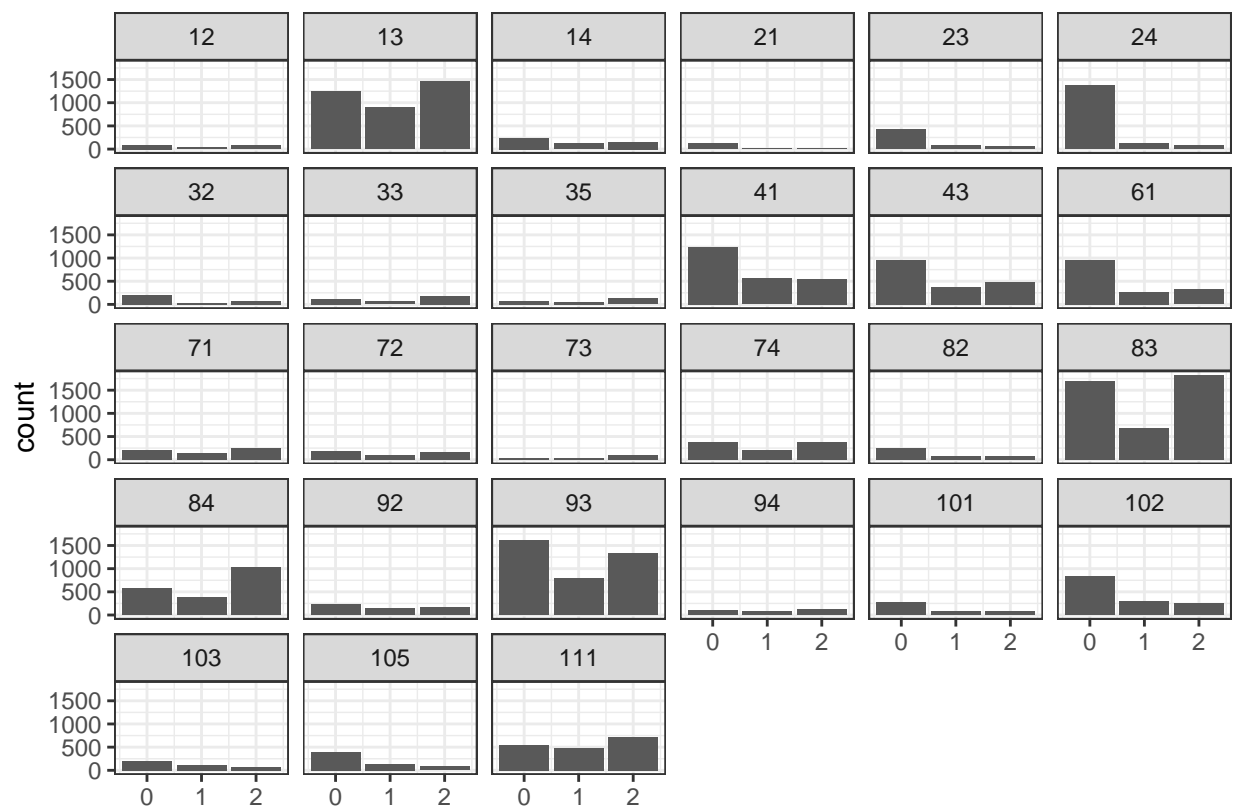*June 2019*

Table 1: Overall Summary
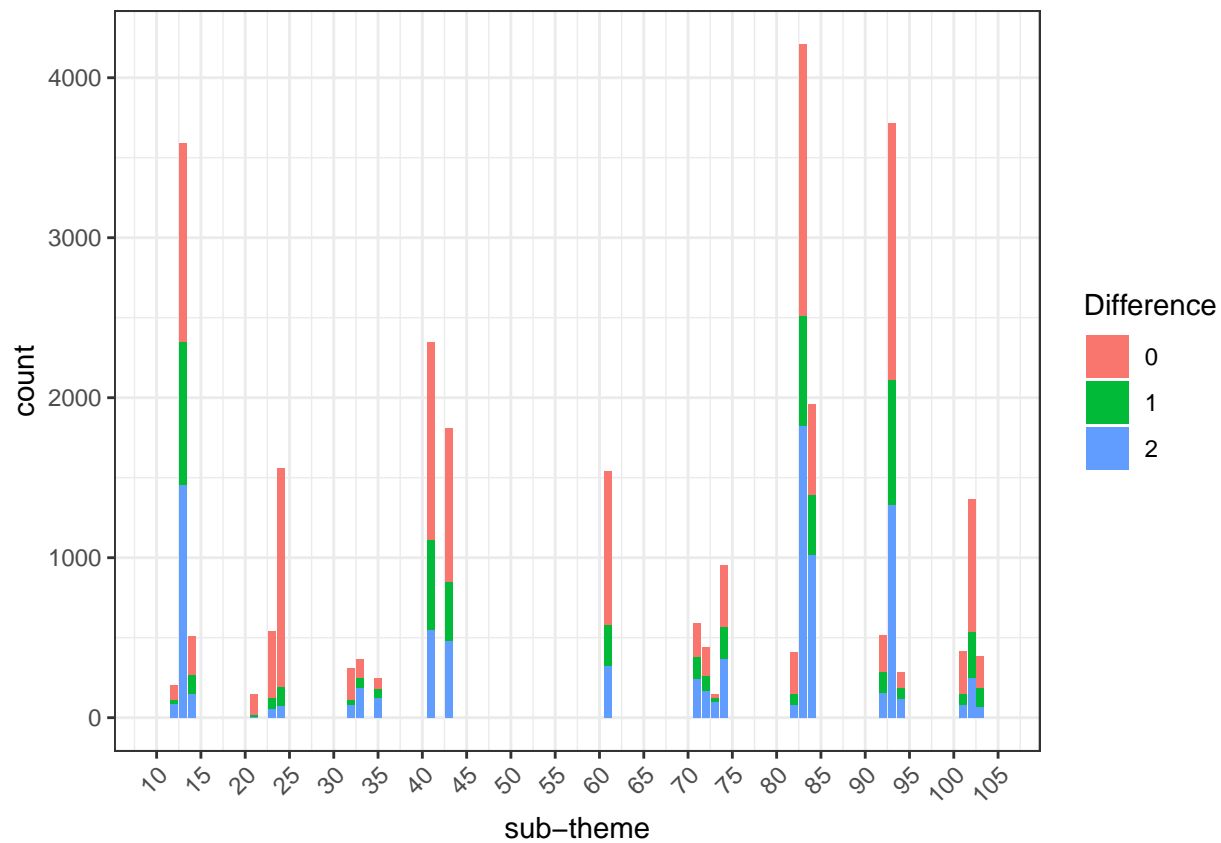
Qualitative Sentiment	Quantitative Sentiment	Difference	Count	Total Percentage
negative	negative	0	14503	46.92
negative	neutral	1	6236	20.17
negative	positive	2	10173	32.91

Table 2: Differences by theme

Theme	Total Number	Difference of 0 (%)	Difference of 1 (%)	Difference of 2 (%)
12	202	44.55	12.87	42.57
13	3587	34.57	24.90	40.54
14	507	47.53	22.29	30.18
21	147	87.07	9.52	3.40
23	541	76.71	13.12	10.17
24	1563	87.52	7.93	4.54
32	307	63.52	10.42	26.06
33	364	32.42	16.76	50.82
35	247	27.53	20.65	51.82
41	2346	52.60	23.91	23.49
43	1809	52.90	20.51	26.59
61	1544	62.24	16.90	20.85
71	590	35.93	22.88	41.19
72	443	40.86	20.54	38.60
73	150	16.00	20.67	63.33
74	955	40.52	20.94	38.53
82	409	63.33	17.60	19.07
83	4209	40.37	16.27	43.36
84	1959	28.79	19.14	52.07
92	514	43.97	25.49	30.54
93	3719	43.16	21.08	35.76
94	281	33.81	23.49	42.70
101	414	64.73	16.43	18.84
102	1362	60.65	21.07	18.28
103	385	51.69	30.65	17.66
105	616	63.64	22.56	13.80
111	1742	31.63	27.32	41.04

Comments and To do's

- not sure why "Warning message: Unknown or uninitialised column: 'theme'." keeps appearing
- write some more unit tests/double check filter conditions