US 2020 Survey Writeup

Dave Armstrong

10/23/2020

The data have been collected by Qualtrics between 2020-09-09 and 2020-10-19. We collected data in 12 cities - Atlanta, Charlotte, Chicago, Columbus (OH), Dallas, Denver, Fresno, Los Angeles, Milwaukee, New York City, Portland (OR) and Raleigh. These cities were selected based on historical conflict data. We performed a cluster analysis on historical conflict trajectories and found three different groups - those that had historically high, middling and low levels of conflict, particularly incidents of protest and riots. We chose four cities from each of these three groups based on having a minimum population of around 400,000. We hoped this would enable us to get good quality samples in each city.

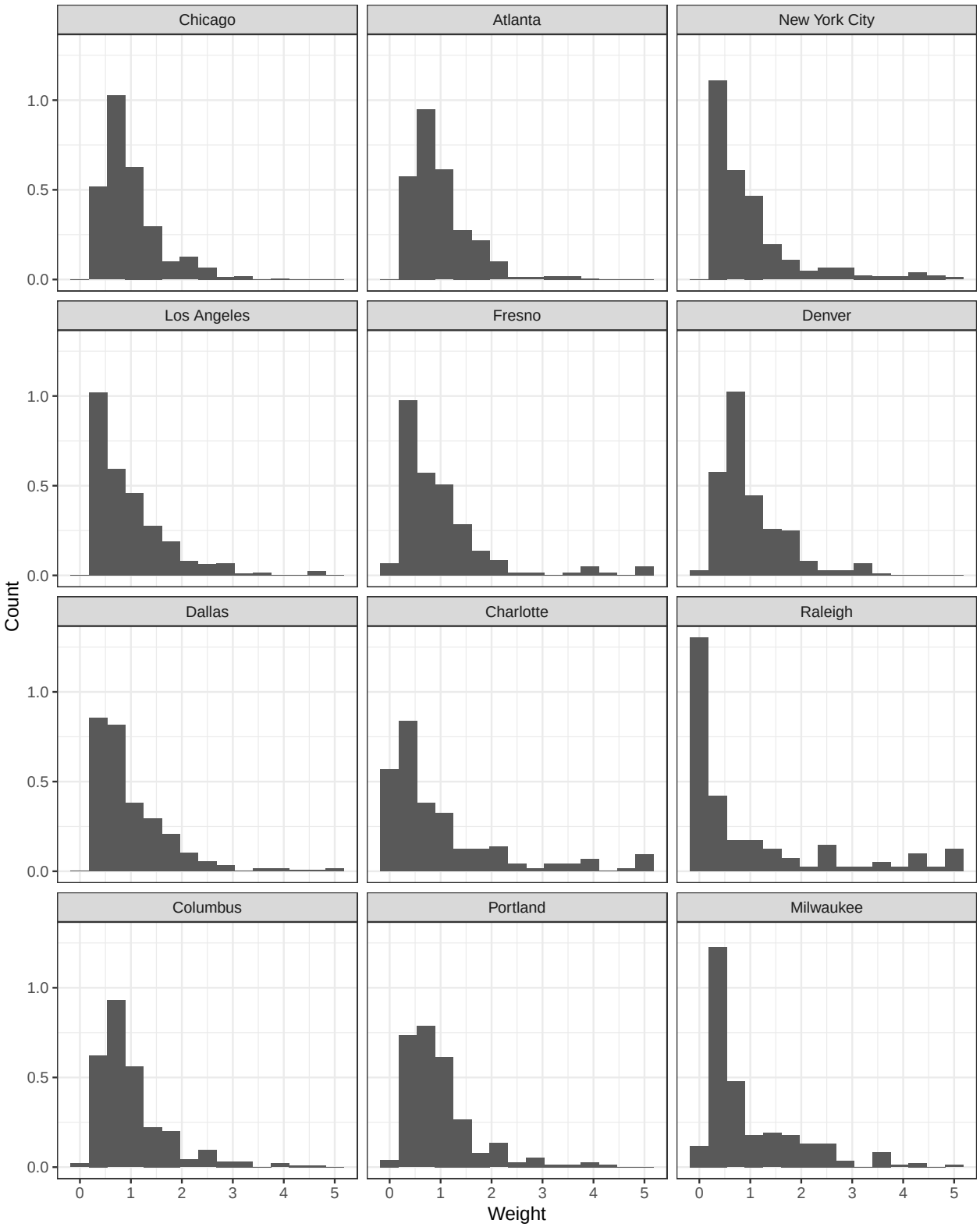
Qualtrics implemented a quota sampling strategy in each of our cities using the following quotas:

	variable	category	Chicago	Atlanta	NY	LA	Denver	Dallas	Charlotte	Columbus	Portland	Raleigh	Milwaukee	Fresno
1	gender	Male	0.49	0.49	0.48	0.5	0.5	0.5	0.48	0.49	0.49	0.48	0.48	0.49
2	gender	Female	0.51	0.51	0.52	0.5	0.5	0.5	0.52	0.51	0.51	0.52	0.52	0.51
3	race	White (NH)	0.33	0.38	0.32	0.29	0.54	0.29	0.42	0.56	0.71	0.53	0.35	0.27
4	race	Black (NH)	0.3	0.52	0.21	0.09	0.09	0.24	0.35	0.28	0.06	0.28	0.38	0.07
5	race	Hispanic	0.29	0.04	0.29	0.49	0.3	0.42	0.14	0.06	0.1	0.11	0.19	0.49
6	race	Asian	0.06	0.04	0.14	0.12	0.04	0.03	0.07	0.06	0.08	0.05	0.04	0.14
7	race	Other	0.02	0.02	0.04	0.01	0.03	0.02	0.02	0.04	0.05	0.03	0.04	0
8	income	<$25000	0.28	0.28	0.27	0.32	0.23	0.29	0.27	0.3	0.26	0.2	0.28	0.36
9	income	$25-50000	0.25	0.31	0.25	0.29	0.28	0.3	0.33	0.29	0.29	0.3	0.32	0.3
10	income	$50-75000	0.21	0.16	0.17	0.16	0.19	0.19	0.16	0.19	0.18	0.2	0.2	0.14
11	income	$75-100000	0.09	0.09	0.11	0.09	0.13	0.07	0.08	0.1	0.1	0.1	0.11	0.1
12	income	$100-150000	0.09	0.08	0.09	0.07	0.08	0.07	0.08	0.07	0.11	0.1	0.07	0.05
13	income	$150000 +	0.08	0.08	0.11	0.07	0.09	0.08	0.08	0.05	0.06	0.1	0.02	0.05
14	educ	<HS	0.16	0.11	0.2	0.26	0.14	0.24	0.12	0.12	0.08	0.09	0.11	0.25
15	educ	HS Grad	0.23	0.19	0.25	0.19	0.18	0.21	0.2	0.26	0.16	0.16	0.28	0.23
16	educ	Some College	0.18	0.17	0.15	0.18	0.18	0.18	0.21	0.22	0.22	0.18	0.21	0.24
17	educ	2yr Degree	0.06	0.05	0.06	0.06	0.05	0.05	0.07	0.07	0.07	0.08	0.08	0.08
18	educ	4yr Degree	0.22	0.27	0.2	0.21	0.28	0.2	0.27	0.22	0.28	0.32	0.21	0.14
19	educ	Grad Degree	0.15	0.21	0.14	0.1	0.17	0.12	0.13	0.11	0.19	0.17	0.11	0.06
20	age	18-24	0.13	0.12	0.13	0.17	0.1	0.13	0.12	0.17	0.1	0.17	0.18	0.17
21	age	25-34	0.25	0.19	0.19	0.23	0.28	0.25	0.17	0.25	0.24	0.24	0.23	0.23
22	age	35-44	0.18	0.2	0.19	0.17	0.2	0.19	0.2	0.17	0.21	0.19	0.16	0.17
23	age	45-54	0.16	0.2	0.18	0.17	0.15	0.16	0.19	0.16	0.16	0.16	0.16	0.16
24	age	55-64	0.14	0.15	0.15	0.13	0.13	0.14	0.15	0.13	0.15	0.12	0.14	0.13
25	age	65+	0.14	0.14	0.16	0.13	0.14	0.13	0.17	0.12	0.14	0.12	0.12	0.13

At the time of our initial analysis, many of the quotas had not been met. To produce results that are roughly representative of the cities, we used developed a set of rake weights using the `anesrake` (Pasek 2018) package in R. We used the proportions above as the targets for the weighting, except for race, where Other, Asian and

Hispanic were grouped together to avoid having groups that were too small. We also grouped together those who had not graduated high school and those who had only graduated high school because there were too few of the former in our survey. We also grouped together those making $100,000-150,000 and those making more than $150,000 per year for the same reason.

In all cases, except one, the weighting procedure was able to produce weights that mimicked the targets within rounding. The case where that was not true was the gender breakdown in Los Angeles where the target was 50/50 and the weighted percentages were 51.3% Female and 48.7% Male. The general design effects ranged from 1.31 in Chicago to 3.17 in Raleigh. The figure below plots a histogram of the weights in each city (weights were capped on the high end at 5).



We intended to collect 500 observations from each city for a total of 6000 respondents, but by the time of our preliminary analysis, we had collected just over 4000 observations, rather unevenly, from the 12 cities. The table below gives the number of observations collected city.

City	N
Chicago	423
Atlanta	448
New York City	902
Los Angeles	487
Fresno	166
Denver	282
Dallas	354
Charlotte	207
Raleigh	114
Columbus	265
Portland	210
Milwaukee	235

Pasek, Josh. 2018. *Anesrake: ANES Raking Implementation.* https://CRAN.R-project.org/package=anesrake.