

Officer Codebook

Officer Roster

- **birth__year**: Birth year of the officer.
- **appointed__month**: The month and year the officer was made an officer in YYYY-MM-DD format. The day is always the first day of the month.
- **officer__id**: Unique identifier for each officer.
- **officer__race**: Race of the officer.
- **officer__gender**: Sex of the officer.
- **spanish**: Does the officer speak Spanish or not?
- Uniquely identified by **officer__id**. The unit of observation is an officer.
- **Number of officers**: 33645

Creating the officer roster

“The administrative data from the CPD used in this study span multiple datasets collected in collaboration with the Invisible Institute, Sam Stecklow, and Emma Herman over the course of three years (2016-2019). We obtained these records from the Chicago Police Department or Chicago Department of Human Resources via Freedom of Information Act (FOIA) or through court ordered releases stemming from requests made by Invisible Institute and Jaime Kalven. CPD provided the following data: rosters of all available current and past officers up to 2018, unit history data for individual officers from the 1930s to 2016, Tactical Response Reports from 2004 to 2018 (i.e. use of force reports), and arrest data with arresting officers and arrestee demographic information from 2001 to 2017. The Chicago Department of Human Resources provided data on officers’ language skills up to 2019. We supplement our core data with data on ‘Stop, Question and Frisk’ (SQF) activity between 2012-2015, which was shared by the Lucy Parson’s Lab. Finally, the Automated Daily Attendance and Assignment sheet data for each police district between 2012 and 2015 was obtained via a FOIA request to the CPD and shared by Rachel Ryley.” pages 5-6 of Appendix Section S1.2 in Ba et al. 2021.

“These data and others have been used to construct rich profiles of Chicago Police Officers. While no file contains a unique identifier (star numbers change over time, names are common, etc.), we constructed unique officer profiles through a successive merge process described here. Each file contains some identifying information such as of demographic data (birth year, race, gender) or other characteristics (name, start/badge number, appointed date, resignation date, current unit). We used these identifying characteristics to first de-duplicate officers within a file and to then merge to pre-existing officer data with inter-file unique identifiers. The merging process itself is an iterative-pairwise matching method, where the officers in each dataset are repeatedly merged on identifying characteristics and any successful 1-to-1 match in a round removes the matched officers from the next round of merging.” page 6 of Appendix Section S1.2 in Ba et al. 2021.

Officer Race

“We determine race/ethnicity of CPD officers based on demographic data obtained from the CPD through FOIA. The CPD usually classifies race/ethnicity in at most 7 mutually exclusive groups: White/Caucasian, White Hispanic, Black/African American, Black Hispanic, Asian/Pacific Islander, Native American/Native Alaskan, and unknown/missing. However, there are inconsistencies in how races and ethnicities are coded across files. For example, some files do not include ‘Black Hispanic’ as a racial category (very few officers are ever classified as Black Hispanic), and some files contain outdated racial categories which we update to the best of our ability. For consistency, we classify ‘Hispanic’ and ‘White Hispanic’ as ‘Hispanic’; ‘Black’ and ‘Black Hispanic’ (rare cases) as ‘Black.’ ‘White’ in our analysis refers to non-Hispanic White. If an officer has multiple races associated with them across different datasets, we aggregate by most common non-missing races.” page 5 of Appendix Section S1.1 in Ba et al. 2021.

variable	value	nr	prop	nr_unique
officer_gender	MALE	27,869	0.828	3
officer_gender	FEMALE	5,771	0.172	3
officer_gender		5	0.000	3
officer_race	officer_white	20,675	0.615	6
officer_race	officer_black	7,634	0.227	6
officer_race	officer_hisp	4,578	0.136	6
officer_race	officer_aapi	538	0.016	6
officer_race		153	0.005	6
officer_race	officer_native	67	0.002	6
spanish	FALSE	30,801	0.915	2
spanish	TRUE	2,844	0.085	2

variable	mean	sd	min	p25	median	p75	max	iqr	mad	nr_missing	prop_missing
birth_year	1,956.503	18.135	1,916	1,944	1,956	1,971	1,996	27	20.756	926	0.028

variable	nr_rows	nr_unique	nr_missing	prop_missing
officer_id	33,645	33,645	0	0

variable	min	p25	median	p75	max	nr_missing	prop_missing
appointed_month	1936-03-01	1969-05-01	1988-01-01	2000-08-01	2018-02-01	70	0.002

variable	unit_of_time	value	n	prop
appointed_month	month	1	2,150	0.064
appointed_month	month	2	3,185	0.095
appointed_month	month	3	2,888	0.086
appointed_month	month	4	2,048	0.061
appointed_month	month	5	2,561	0.076

variable	unit_of_time	value	n	prop
appointed_month	month	6	3,025	0.090
appointed_month	month	7	2,571	0.076
appointed_month	month	8	2,518	0.075
appointed_month	month	9	2,905	0.086
appointed_month	month	10	3,632	0.108
appointed_month	month	11	2,684	0.080
appointed_month	month	12	3,408	0.101
appointed_month	month		70	0.002
appointed_month	year	1,936	1	0.000
appointed_month	year	1,945	11	0.000
appointed_month	year	1,946	206	0.006
appointed_month	year	1,947	135	0.004
appointed_month	year	1,948	90	0.003
appointed_month	year	1,949	30	0.001
appointed_month	year	1,950	80	0.002
appointed_month	year	1,951	101	0.003
appointed_month	year	1,952	164	0.005
appointed_month	year	1,953	319	0.009
appointed_month	year	1,954	358	0.011
appointed_month	year	1,955	599	0.018
appointed_month	year	1,956	677	0.020
appointed_month	year	1,957	736	0.022
appointed_month	year	1,958	299	0.009
appointed_month	year	1,959	226	0.007
appointed_month	year	1,960	96	0.003
appointed_month	year	1,961	578	0.017
appointed_month	year	1,962	471	0.014

variable	unit_of_time	value	n	prop
appointed_month	year	1,963	81	0.002
appointed_month	year	1,964	204	0.006
appointed_month	year	1,965	346	0.010
appointed_month	year	1,966	966	0.029
appointed_month	year	1,967	519	0.015
appointed_month	year	1,968	856	0.025
appointed_month	year	1,969	404	0.012
appointed_month	year	1,970	927	0.028
appointed_month	year	1,971	486	0.014
appointed_month	year	1,972	475	0.014
appointed_month	year	1,973	859	0.026
appointed_month	year	1,974	64	0.002
appointed_month	year	1,975	181	0.005
appointed_month	year	1,976	515	0.015
appointed_month	year	1,977	529	0.016
appointed_month	year	1,978	254	0.008
appointed_month	year	1,979	4	0.000
appointed_month	year	1,980	445	0.013
appointed_month	year	1,981	393	0.012
appointed_month	year	1,982	680	0.020
appointed_month	year	1,983	104	0.003
appointed_month	year	1,985	458	0.014
appointed_month	year	1,986	1,309	0.039
appointed_month	year	1,987	485	0.014
appointed_month	year	1,988	329	0.010
appointed_month	year	1,989	247	0.007
appointed_month	year	1,990	623	0.019

variable	unit_of_time	value	n	prop
appointed_month	year	1,991	1,005	0.030
appointed_month	year	1,992	474	0.014
appointed_month	year	1,993	451	0.013
appointed_month	year	1,994	1,004	0.030
appointed_month	year	1,995	959	0.029
appointed_month	year	1,996	630	0.019
appointed_month	year	1,997	589	0.018
appointed_month	year	1,998	852	0.025
appointed_month	year	1,999	913	0.027
appointed_month	year	2,000	770	0.023
appointed_month	year	2,001	641	0.019
appointed_month	year	2,002	556	0.017
appointed_month	year	2,003	460	0.014
appointed_month	year	2,004	505	0.015
appointed_month	year	2,005	526	0.016
appointed_month	year	2,006	722	0.021
appointed_month	year	2,007	408	0.012
appointed_month	year	2,008	133	0.004
appointed_month	year	2,009	139	0.004
appointed_month	year	2,010	185	0.005
appointed_month	year	2,011	52	0.002
appointed_month	year	2,012	448	0.013
appointed_month	year	2,013	589	0.018
appointed_month	year	2,014	314	0.009
appointed_month	year	2,015	442	0.013
appointed_month	year	2,016	620	0.018
appointed_month	year	2,017	1,058	0.031

variable	unit_of_time	value	n	prop
appointed_month	year	2,018	210	0.006
appointed_month	year		70	0.002