

Merge Officers To Their Work Assignments

Officers

- **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.

Work Assignments

- **officer_id**: Unique identifier for each officer.
- **month**: Month of the work assignment in YYYY-MM-DD format. The day is always the first day of the month.
- **rank**: Rank of the officer assigned to the work assignment.
- **unit**: Unit of the officer assigned to the work assignment.
- **date**: Date of the work assignment in YYYY-MM-DD format.
- **shift**: The timing of the work assignment (i.e., morning, afternoon, evening).
- **start_time**: Hour start time of the work assignment in military time.
- **end_time**: Hour end time of the work assignment in military time.
- **weekday**: Day of the week of the work assignment.

- **beat_assigned**: The beat the officer is assigned to.
- **months_from_start**: The number of months between the officer’s appointment date and their work assignment date.
- **months_from_start_sq**: The number of months between the officer’s appointment date and their work assignment date, squared.
- **duration**: Length of the work assignment in hours.
- Uniquely identified by **officer_id** and **date**. The unit of observation is a specific work assignment for a specific officer.
- **Number of officer-work assignments**: 3519518

Join the officer roster to work assignments - Ba et al. 2021

Pre-filtering totals

- **Number of officers**: 8502
- **Number of officer-work assignments**: 3519518
- **Number of work assignments (unit + beat + shift)**: 9813
- **Number of work assignments (beat + shift)**: 9335
- **Number of MDSBs (month/year + day of week + shift + beat)**: 800311

Filtering by work assignment

“Within this subset, we further drop non-standard assignments (notably including “protest detail,” “station supervisor,” and “station security” assignments, as well as special assignments for training, compensatory time, and excused sick leave). Patrol assignments in which officers are indicated as non-present are also dropped... we drop double shifts (patrol assignment slots in which the assigned officer served for more than one shift on the same day) to address the possibility that officers behave differently due to fatigue in these circumstances. We also eliminate officers assigned to non-standard watches (i.e., other than first through third watches).” page 7 of Appendix Section S1.4 in Ba et al. 2021.

I did not find any evidence in their code that they did this.

Filtering by race/ethnicity

“We restrict analysis to patrol assignments in which Black, Hispanic, or White officers serve. Asian/Pacific Islander and Native American/Alaskan Native officers are not examined due to small sample sizes.” page 7 of Appendix Section S1.4 in Ba et al. 2021.

- Number of officer-work assignments: 3390915
 - Officer-work assignments dropped: 128603
 - Percentage reduction: 3.654%
- Number of unique officers: 8203
 - Number Of Officers dropped: 299
 - Percentage reduction: 3.517%

Filtering by rank

“Finally, we drop officers at ranks other than ‘police officer.’ This step eliminates police sergeants, who serve in 8% of beat assignments but make very few stops and arrests, as well as legal officers, helicopter pilots, explosives technicians, and canine handlers.” page 7 of Appendix Section S1.4 in Ba et al. 2021.

- Number of officer-work assignments: 3050853
 - Officer-work assignments dropped: 468665
 - Percentage reduction: 13.316%
- Number of unique officers: 7216
 - Officers dropped: 1286
 - Percentage reduction: 15.126%

Final results

- **Number of officers:** 6947
 - **Percentage reduction:** 18.29%
- **Number of officer-work assignments:** 2932680
 - **Percentage reduction:** 16.674%
- **Number of work assignments (unit + beat + shift):** 8275
 - **Percentage reduction:** 15.673%
- **Number of work assignments (beat + shift):** 7976
 - **Percentage reduction:** 14.558%
- **Number of MDSBs:** 653087
 - **Percentage reduction:** 18.396%

Missings

- 40064 officer-work assignments are missing their start and end time.
- 17 officer-work assignments are missing the work experience of the officer.

Join the officer roster to work assignments - Risi & Zhao

1. For officer race, combine AAPI and Native American officers into one category.
2. Create a start date time and end date time column which combines date with start time and end time respectively. This creates two columns in the YYYY-MM-DD HH:MM:SS format.
3. For officer rank, I combine Chief, Commander, Deputy Chief, and Lieutenant into a **leadership** category. **Sergeant** and **Police Officer** remain the same. All other ranks are combined into an **other** category. Percentage-wise there are only a few work assignments for those in leadership, but there is still a large absolute number of patrols shifts conducted by those in leadership (in the tens of thousands).

Results of join

- Number of officer-work assignments: 3519518
- Number of officers: 8502
- Number of work assignments (unit + beat + shift): 9813
- Number of work assignments (beat + shift): 9335
- Number of MDSBs (month/year + day of week + shift + beat):800311
- Number of non-matching officers: 25143
 - Percentage of officers who don't match: 74.73%

Missings

- 2 officers are missing their race (7 officer-work assignments).
- 0 officers are missing their sex (0 officer-work assignments).
- 0 officers are missing their age (0 officer-work assignments).
- 0 officers are missing their Spanish-speaking ability (0 officer-work assignments).
- 0 officers are missing the month they started (0 officer-work assignments).
- 48210 work assignments are missing their start time/end time.
- 53765 work assignments are missing their rank.
- 778 work assignments are missing the amount of experience of the officer. This is due to a data error where the officer's appointment date is **AFTER** their work assignment date. I.e., they began working before they were officially a police officer.