Risi - Merging Stops, Arrests, and Uses of Force to Officer Assignments

Changes to the data - Risi

Officer Shift Assignments

1. **Rank**: Chief, Commander, Deputy Chief, and Lieutenant are combined to create a new **leadership** category. **Sergeant** and **Police Officer** remain the same. All other ranks are combined into an **other** category. Percentage-wise there are few shift assignments for those in leadership, but there are still tens of thousands of shifts assigned to those in leadership. Depending on one's analysis, this category may still be useful or not.
2. I drop the **months_from_start_sq** column.
3. I format the police officer race column to be more concise.

n	prcnt	name	value
3050853	0.8668383	rank	police officer
281965	0.0801147	rank	sergeant
111393	0.0316501	rank	other
53765	0.0152762	rank	NA
21542	0.0061207	rank	leadership

Stops

1. I decided to use the **contact_type** variable for categorizing stop types rather than **stop_type** in line with Ba et al. 2021. To see how contact type is created from stop type, see data-schemas.Rmd or the Ba et al. 2021 paper.
2. **civ.race**: I combine Asian American/Pacific Islanders and Native Americans/Alaskan Natives into an **other** category.
3. I drop the **civilian_race_short** column.
4. **civ.gender**: Any civilians whose gender is categorized as **X** gets recoded as missing.

n	prcnt	name	value
1124118	0.6600198	civ.race	black
346558	0.2034797	civ.race	hispanic
210545	0.1236204	civ.race	white
21937	0.0128802	civ.race	other

n	prcnt	name	value
1322801	0.7766754	civ.gender	M
379401	0.2227632	civ.gender	F
956	0.0005613	civ.gender	NA

Arrests

1. I drop the **statute_description** column because there are too many unique values for it to be of use.
2. **civ.race**: I combine Asian American/Pacific Islanders and Native Americans/Alaskan Natives into an **other** category.
3. I drop the **civilian_race_short** column.

n	prcnt	name	value
214048	0.6650097	civ.race	black
70132	0.2178879	civ.race	hispanic
34528	0.1072725	civ.race	white
2494	0.0077484	civ.race	other
670	0.0020816	civ.race	NA

Uses of Force

1. **civ.race**: I combine Asian American/Pacific Islanders and Native Americans/Alaskan Natives into an **other** category.
2. I drop the **civilian_race_short** column.

n	prcnt	name	value
6774	0.7289358	civ.race	black
1387	0.1492521	civ.race	hispanic
949	0.1021199	civ.race	white
106	0.0114064	civ.race	NA
77	0.0082858	civ.race	other

Merge process on Risi data using Ba et al. 2021 merging method

For information on how the Ba et al. 2021 merging method works, see **ba-merge_outcomes_officer-assignments**.

Merge stops - Risi using Ba et al. 2021

- **Number of resulting rows**: 4318675
- **Number of stops which occurred during a shift**: 1465791
  - **Percentage of stops** (1703158): 86.0631251%
- **Number of shift assignments retained:** 3515226
  - **Percentage of shift assignments**: 99.8780515%
- 8511 shift assignments matched with at least one stop **without having** a start time and an end time.
  - This represents 0.2418229% of all shift assignments.
  - 20134 stops are affected or 1.1821569% of all stops.
- There are 345 stops which match to more than one shift assignment.
  - These **duplicate** stops represent 0.0202565% of all stops (1703158).
  - 364 shift assignments are affected or 0.0103423% of all shift assignments.
  - In total, there are 690 **duplicate stop-shift observations** representing 0.0159771% of all stop-shift observations.

Merge arrests - Risi using Ba et al. 2021

- **Number of resulting rows**: 3551941
- **Number of arrests which occurred during a shift**: 268554
  - **Percentage of arrests** (321872): 83.4350301%
- **Number of shift assignments retained:** 3518508
  - **Percentage of shift assignments**: 99.9713029%
- 3635 shift assignments matched with at least one arrest **without having** a start time and an end time.
  - This represents 0.1032812% of all shift assignments.
  - 4153 arrests are affected or 1.2902645% of all arrests.
- There are 114 arrests which match to more than one shift assignment.
  - These **duplicate** arrests represent 0.0354178% of all arrests (321872).
  - 208 shift assignments are affected or 0.0059099% of all shift assignments.
  - In total, there are 228 **duplicate stop-shift observations** representing 0.006419% of all arrest-shift observations.

Merge uses of force - Risi using Ba et al. 2021

- **Number of resulting rows**: 3519703
- **Number of uses of force which occurred during a shift**: 9293
  - **Percentage of uses of force** (9293): 100%
- **Number of shift assignments retained:** 3519516
  - **Percentage of shift assignments**: 99.9999432%
- 146 shift assignments matched with at least one use of force **without having** a start time and an end time.
  - This represents 0.0041483% of all shift assignments.
  - 149 uses of force are affected or 1.6033574% of all uses of force.
- There are 6 uses of force which match to more than one shift assignment.
  - These **duplicate** uses of force represent 0.0645647% of all uses of force (9293).
  - 12 shift assignments are affected or 3.4095578×10^{-4}% of all shift assignments.
  - In total, there are 12 **duplicate stop-shift observations** representing 3.4093786×10^{-4}% of all force-shift observations.

Merge process - Risi

For notes on the differences between the Risi merging process and the Ba et al. 2021 merging process, see **ba-merge_outcomes_officer-assignments**.

Merge stops - Risi

- **Number of resulting rows**: 4306505
- **Number of stops which occurred during a shift**: 1439082
  - **Percentage of stops** (1703158): 84.4949206%
  - **Number of fewer stops matched than Ba et al. 2021**: 26709
  - **Percent reduction from Ba et al. 2021**: 1.8221561%
- **Number of shift assignments retained:** 3519518
  - **Percentage of shift assignments**: 100% **NOTE** this should be 100%.
- 0 shift assignments matched with at least one stop **without having** a start time and an end time. **NOTE** that this should be 0.
- There are 60 stops which match to more than one shift assignment.
  - These **duplicate** stops represent 0.0035229% of all stops (1703158).
  - 66 shift assignments are affected or 0.0018753% of all shift assignments.
  - In total, there are 120 **duplicate stop-shift observations** representing 0.0027865% of all stop-shift observations.
  - As noted above, **all these entries are dropped**.

Merge arrests - Risi

- **Number of resulting arrests**: 3552174
- **Number of arrests which occurred during a shift**: 263026
  - **Percentage of arrests** (321872): 81.7175772%
  - **Number of fewer arrests matched than Ba et al. 2021**: 5528
  - **Percent reduction from Ba et al. 2021**: 2.0584315%
- **Number of shift assignments retained:** 3519518
  - **Percentage of shift assignments**: 100% **NOTE** this should be 100%.
- 0 shift assignments matched with at least one arrest **without having** a start time and an end time. **NOTE** that this should be 0.
- There are 10 arrests which match to more than one shift assignment.
  - These **duplicate** arrests represent 0.0031068% of all arrests (321872).
  - 20 shift assignments are affected or 5.6825963×10^{-4}% of all shift assignments.

  - In total, there are 20 **duplicate arrest-shift observations** representing 5.6303548×10^{-4}% of all arrest-shift observations.
  - As noted above, **all these entries are dropped**.

Merge force - Risi

- **Number of resulting rows**: 3519699
- **Number of uses of force which occurred during a shift**: 9096
  - **Percentage of uses of force** (9293): 97.8801248%
  - **Number of fewer uses of force matched than Ba et al. 2021**: 197
  - **Percent reduction from Ba et al. 2021**: 2.1198752%
- **Number of shift assignments retained:** 3519518
  - **Percentage of shift assignments**: 100% **NOTE** this should be 100%.
- 0 shift assignments matched with at least one use of force **without having** a start time and an end time. **NOTE** that this should be 0.
- There are 0 uses of force which match to more than one shift assignment.
  - These **duplicate** uses of force represent 0% of all uses of force (9293).
  - 0 shift assignments are affected or 0% of all shift assignments.
  - In total, there are 0 **duplicate force-shift observations** representing 0% of all force-shift observations.
  - As noted above, **all these entries are dropped**.