

Spectacular Cranes CP2 Findings

Questions

1. Is there a significant difference in the distribution of allegation categories in the new data (CPDB) compared to the old data? Visualization of allegation data distribution by category and compare with different versions of allegation tables - i.e. find allegation category distribution from previous version and compare with the current version of cpdb database (density charts)
2. How are the allegation categories of the new data (CPDB) distributed geographically (can we create points on a map to show the distribution?)
3. Similar to question 2, how are allegation categories distributed among beat id's? Can we compare the both the distribution of the old data and the new data? (maybe bubble density chart)
4. How have the allegation categories changed from the old data to the new (CPDB)? Are there more or less categories? What is the distribution for the categories, old and new? (pie chart or donut chart)

Answers

1. See `Question1`, `Distribution (Old DB)`, `Distribution (CPDB)` dashboard and sheets in the workbook. You may also look at the attached pictures: `% Distribution (Old data).png` and `% Distribution (CPDB).png` The first thing to note is the change in the allegation category names from the old data to the new data (CPDB). In the old data, there were many more categories and commonly used ones were "MISCELLANEOUS", "ARESTEE-DURING ARREST", and "UNECESSARY PHYSICAL CONTACT ON-DUTY." Not only is there a significant difference because of the category names, but one can also see there is also a difference where category names seem to be the same. For example, "Excessive Force" in the old data is only 5%

whereas "Use of Force" in the new data is 35%. This could be due to name conventions changing, which is the main challenge of our theme.

2. See `Question2` sheet in the workbook or see `Q2_heatmap.png`. Here we can see where allegations occurred and normalized categories for those allegations.

3. See `Question3` and `Question3A` sheets in the workbook or `Q3.png` and `Q3A.png`. For each category, marked by color, you can see the distribution of allegation categories for CPDB where the size of the bubble indicates the % of allegations in a category in that particular beat. For example, beat 132 has a larger 'Operations/Personnel Violations' red bubble than others at 8.5% indicating that a larger percentage of those violations are committed in beat 132.

4. See `/src/d3/index.html` or the two "Screen Shots" of the donut charts to see what the interactive vizualization might look like. Here, with "Old Category Dist" selected, you can see many more categories and a much more skewed distribution. When "New Category Dist" is selected, you can see less categories and a somewhat more even distribution where for example, the purple part is "Use of Force", which confirms what we have seen in previous visualizations such as the one for Question 1.