Checkpoint 5

Theme:

Our theme is analyzing the influence that supervisors have on the officers they manage.

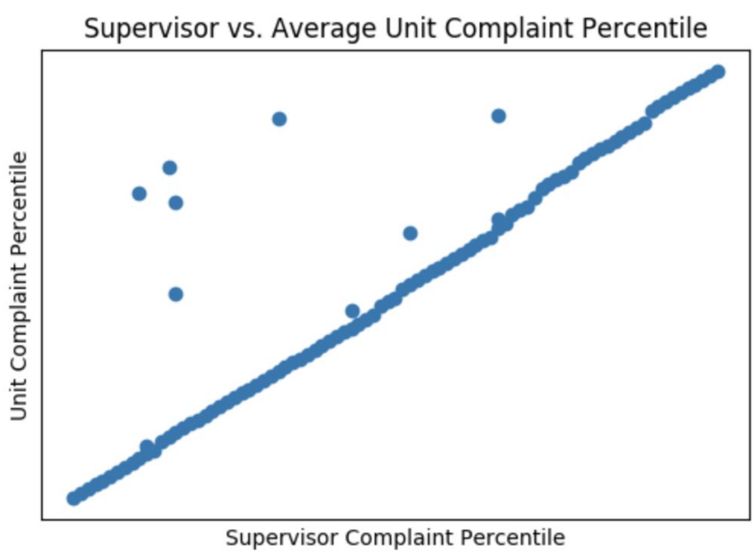
Machine Learning questions

Can we predict a unit's complaint percentile based on a supervisor's complaint percentile?

Throughout this project, supervisors with high complaint percentiles (complaint percentiles over 75) have been associated with managing units that also have high complaint percentiles. However, we have found that when a unit has a high complaint percentile, it does not guarantee that the supervisor has a high complaint percentile.

Our indicator of misconduct in our project has been a high complaint percentile. Keeping with this indicator, if we can predict a unit's complaint percentile based on their supervisor's, we may be able to make recommendations on whom to promote to supervisors to keep misconduct in the unit, and overall in the department, low.

Our hypothesis for this question was that supervisors with high complaint percentiles would be predicted to manage units that have high complaint percentiles, and supervisors with lower complaint percentiles would be predicted to manage units with lower complaint percentiles, although we weren't completely sure about this second part.



When we plotted the data (as shown above) with unit's complaint percentile on the y axis and supervisor's complaint percentile on the x axis, we observed a very strong linear relationship that supported our hypothesis, despite a few outliers.

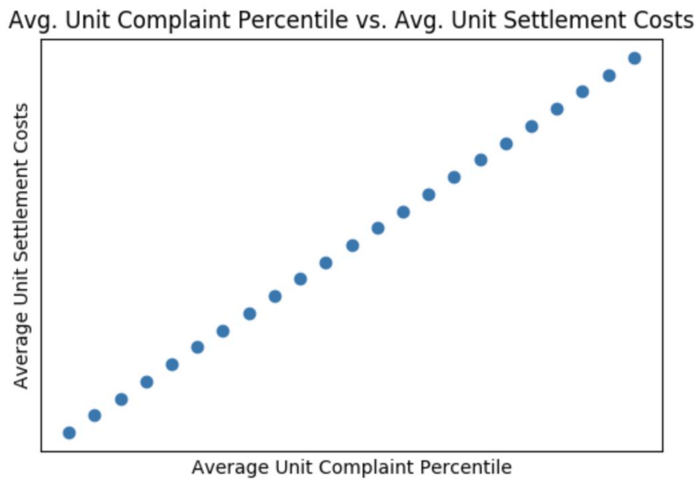
Unfortunately, the results from MSE and R^2 metrics showed that our model did not score well, with high error and a low coefficient of determination (MSE=298.65; R^2=0.132). We attribute this to not having enough data to train and test on. There are 98 units total, so once we split that up into training and testing, it's not very much for our model to learn from; the previous metrics were reported using a 70/30 train/test split. Other train/test splits did not fare much better. However, based on the appearance of the plotted data, we decided to run a Pearson's correlation test. Our data has a highly statistically significant positive correlation ($r=0.323$; $p < 0.001$) of supervisor complaint percentile to their average unit complaint percentile.

Thus, despite not being able to run predictive statistics due to a lack of training data, our correlational results still support our hypothesis (as does the clear plot). Hopefully, these results could still be useful to a police department when making promotion decisions, although correlation should not be mistaken for causation. While more research is required, it would be recommended that police departments thoroughly vet an officer with a high complaint percentile before making them supervisor. In future work, more data would be required in order to draw any meaningful predictive conclusions.

Can we predict how much a unit will cost the department in settlements based on a unit's complaint percentile?

This question is important because if we can find an indicator that a unit will cost the department a lot of money in settlements, there can possibly be intervention to prevent that from happening.

Our hypothesis was that units with high complaint percentiles will cost the department more money than units with lower complaint percentiles. We defined cost as the sum of payment and fees_cost. We decided to look at cost in terms of average cost for each unit.



We used linear regression for our model because when we plotted the data, it had a surprisingly strong linear trend, as seen in the graph above. We want to note that we were very suspicious about how linear this data was. We checked to make sure our data was mapped correctly throughout our process, and regardless of data conversions we confirmed that we preserved column relationships. We found no indication that any new mapping had occurred.

Yet again, the results from MSE and R^2 metrics showed that our model did not score well, with extremely high error and a low coefficient of determination (MSE > 5×10^9; $R^2 = -0.081$). We had to do a 50/50 train/test split, or else the error was much worse. We attribute this to not having enough data to train and test on. There are 98 units total, but we don't have settlements data for every unit, coming out to a total of only 23 units with settlement data. Once we split the data into training and testing, it's very little for our model to learn from. We found this interesting because there is such a clear correlation in this data. Thus, we tried running a Pearson's r test again, however this time the correlation was not significant ($r=0.139$, $p=0.527$). Some resources suggest calculating Pearson's correlation coefficient is not advised for data under 25 samples, so this could explain the inconclusive p value.

Thus, no meaningful relationship could be discerned between average unit complaint percentiles and average unit settlement costs. Future research with more data would be advised so that departments could determine exactly how financially damaging high-misconduct units are.

Conclusion for Machine Learning

Our original thought was that if a supervisor with a high complaint percentile is more likely to lead a unit with a high complaint percentile, and a unit with a high complaint percentile is predicted to cost the department a lot of money, promoting an officer with a history of misconduct (high complaint percentile) to a supervisor role will end up costing the department a lot of money. However, when we don't have enough data to create these models, we can't provide enough insight to inform any sort of recommendations in these instances. That said, there is still strong correlational evidence that a high supervisor complaint percentile is related to a high average unit complaint percentile, and requires more investigation by police departments and data scientists.

Text Analytics questions

What is the top tag associated with allegations for supervisors above the 75th complaint percentile?

We wanted to get an idea of the type of allegations supervisors that we've been studying (those above the 75th complaint percentile) were engaged in. The document tags were able to provide this sort of context. We decided to find the top tag that was associated with the allegations the supervisor has been involved in.

Unfortunately, while our code and pipeline works, there was not enough tagged allegation data; there was no data associated with this subset of supervisors. When expanded to include all supervisors, it was evident that none of the supervisors we've identified had tagged allegation documents. However, we feel that our work on this is still important to include. As more tagged allegation data becomes available, our pipeline will already be set up and ready to use to find the top tag associated with supervisors, and importantly, those above the 75th percentile. This lays the groundwork for future groups hoping to find this information.

If they had this information, police departments could potentially use the data to predict which supervisors are most likely to commit high levels of misconduct, which is important in restructuring a department.

What is the top tag associated with allegations for the units the supervisors oversee from the question above?

To understand the influence a supervisor has on their unit, we wanted to see if the unit's top tag was the same as the supervisor's. This allows us to see if the incidents officers are involved in are related to their supervisor's. Without the supervisors' tags, the full comparison was not possible. However, we were able to find each unit's top tags. You can see the full results in src/top_allegation_label_per_unit.csv. The most frequent top tag is tasers, which is unsurprising given what we learned earlier in this course; police officers were given tasers with little to no training and had a lot of issues with them. We were not able to glean a lot of information from this data, mostly because it is pretty scarce right now. A subset of our results showing the prevalence of the "tasers" label per unit is shown below:

unit_id	max_label
2	tasers
3	tasers
4	tasers
5	tasers
6	trespass
7	trespass
8	tasers
9	tasers
10	tasers
11	tasers
12	tasers

Conclusion for Text Analysis

Despite our inconclusive findings, we still think this work is valuable for future data analysis by the Invisible Institute or future groups in this course. If the top tag for a unit were the same for its supervisor, it's possible there could be a relationship such that the supervisors are influencing the behaviors of the officers in their unit. As the tagged data becomes more populated, more insight will be gained from running our code.