Does It Pay To Be Good?

Dhruv Narayan, Keshav Bhandari and Akshay Kumar

December 7, 2021

Abstract

We hope to answer the question with the data - Does Chicago PD reward their good police officers? We will categorize a "good" police officer with a custom defined complaint severity score (CSS), which is as follows: The average annual complaints of an officer multiplied by custom defined weights based on the complaint category. The weights are defined based on research by our previous cohort [1]. We look at different percentile blocks of police officers categorized by CSS and identify 'good' police officers. We also look at different components that comprise "rewards" for an officer. These include benefits, awards and promotions that are available to us through the CPDP dataset.

Past research [2, 3] has shown us that there were many officers who had filed tactical reports and also had a number of complaints filed against them, but were rewarded by the police department. We try to extend this research by asking the inverse: "Whether the Chicago police department rewards its 'good' officers to encourage good behavior?"

We check whether officers in a low CSS percentile bracket receive higher awards compared to officers in a high one. Under different reward categories, we observe the distribution of CSS of award winners. We visualize the relationship between CSS of an officer and their salary to understand if CSS is an important metric. Also, the first degree connections of officers provides insight into group behavior. With graph analytics, we aim to observe the CSS and rewards for first degree connections of police officers. We find out whether there is any relationship between the CSS metric and the sentiment polarity of allegation summaries. Finally, we include a text auto-completion system developed using Deep Learning. Chicago PD officers can benefit heavily from a system like this, which can help cut down the time to file an official report.

Contents

1 Relational Analytics:

For this checkpoint, our most interesting finding comes from the following plots which show how rewards change for the police officers under different brackets of CSS. Brackets were determined based on the percentile of the CSS (0-10, 10-20, etc.). The plots show us that the majority of awards have an increasing trend where officers who have a higher CSS score tend to receive more awards than those who score lower. The outside governmental agency award and police blue shield award are the only two awards which decrease as CSS scores increase. This tells us that the outside governmental agency tends to look at different factors, weighing complaints more severely compared to the internal award factors.

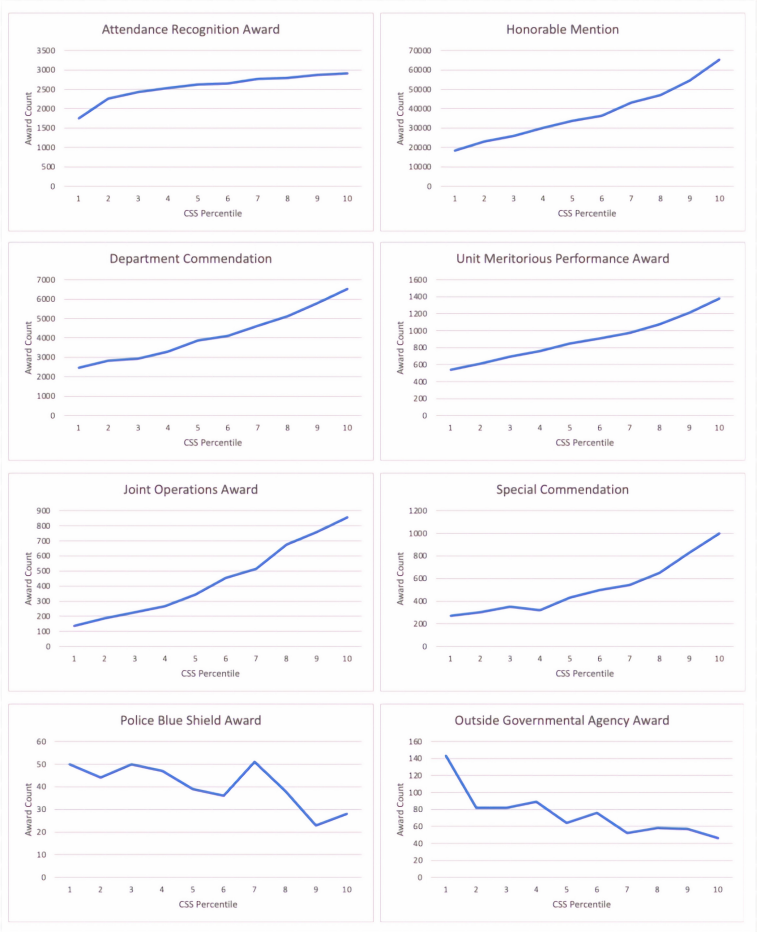


Figure 1: Relationship of CSS and Awards

2 Visualization

The most common rewards are listed below in decreasing order of reward counts -

1. Honorable Mention
2. Complimentary Letter
3. Department Commendation
4. Attendance Recognition Award
5. Emblem of Recognition- Physical Fitness

These reward categories account for 79.98% of total rewards.

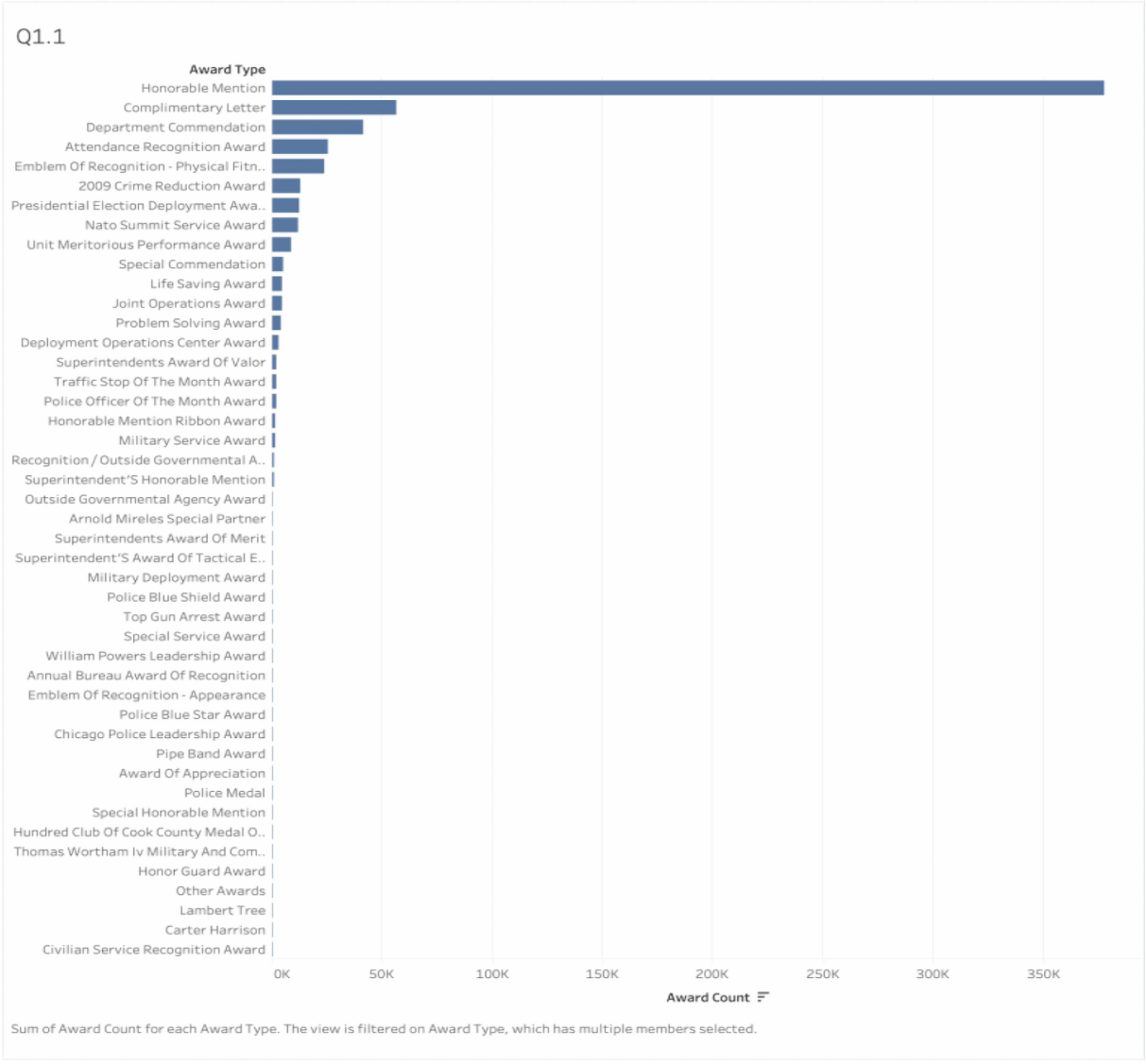


Figure 2: Distribution of awards

Next, the relationship between reward categories and average CSS of officers through a box and whisker plot, which allows us to visualize the variance between each category while also showing us the outliers.

From the figure shown below, we see that reward categories and CSS of officers have a very non-uniform relationship. Several categories have outliers for the CSS metric. This indicates that despite having a high CSS, several cops have received awards. On the other hand, categories such as the Chicago police leadership award and Outside Governmental Agency Award have relatively lowered CSS values.

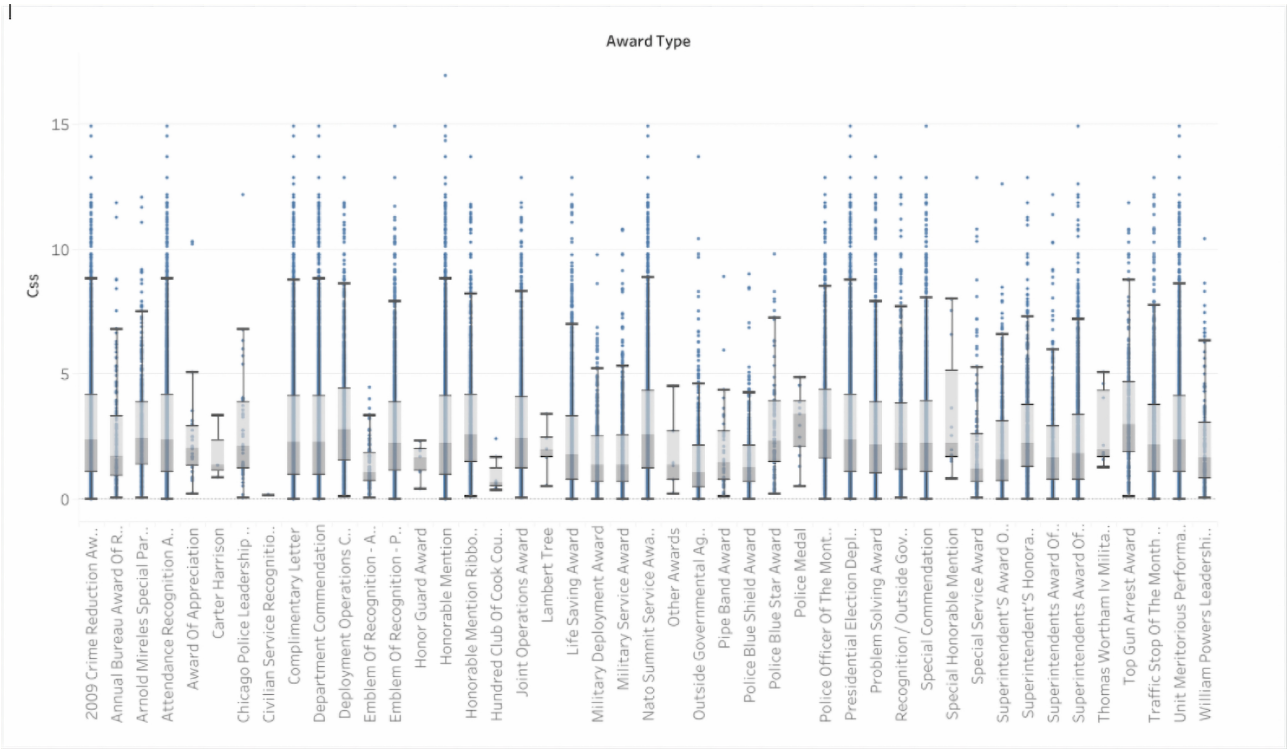


Figure 3: Award Type

3 Interactive visualization

We try to find whether there is any inherent pattern in the relationship of officer's CSS and salary.

1. Looking at the majority of complaint categories, there is a pattern of high salary for low CSS scores with some outliers, i.e. officers with high CSS scores have low salaries and vice versa (which is good).

   Some examples include - Illegal Search, Operation/Personnel Violations, and Use of Force. From the example shown above from the illegal search, we divided the scatter plot into 4 quadrants and we can see in the first quadrant (top right), there are very few officers who have a high CSS score with a high income.

2. Several complaint categories like Bribery/Official Corruption, Verbal Abuse, and

Drug/Alcohol Abuse do not show any visible relationship between the CSS and salary of officers.

3. It is interesting to note that in the majority of cases, the highest earning bracket of police officers have no association with CSS, i.e. the highest earning members (possibly the highest ranking officers) of Chicago PD have a wide range of CSS scores.

   This tells us that they receive a high salary irrespective of them having a 'good'/'bad' CSS metric.

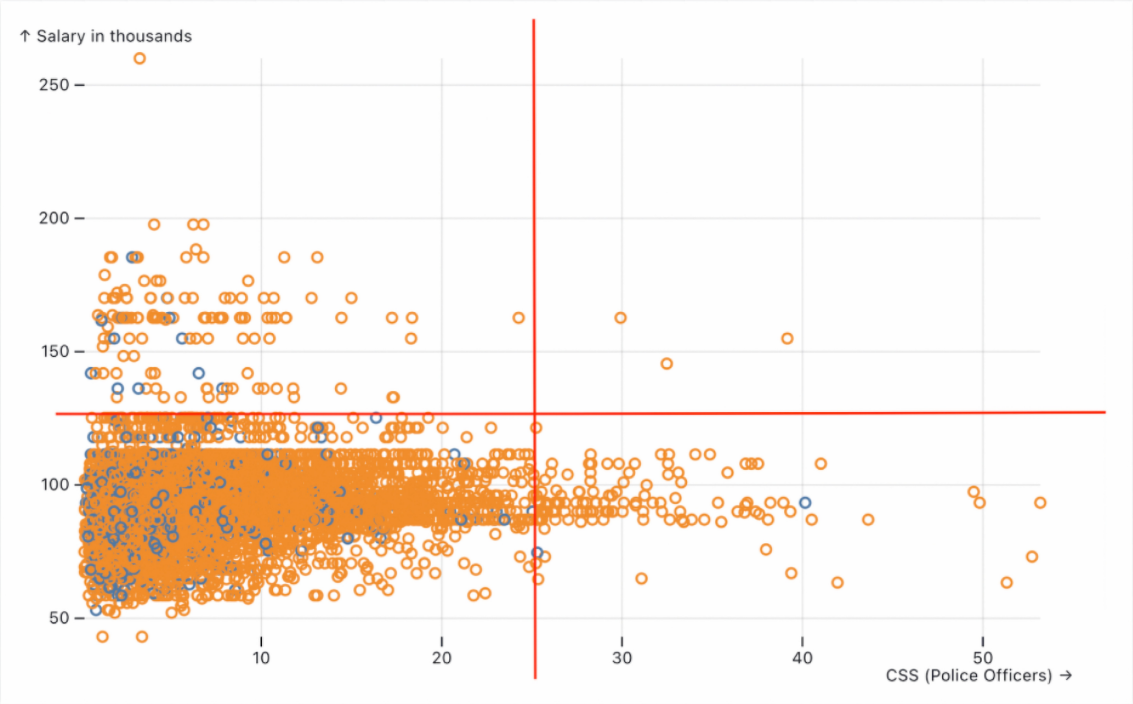


Figure 4: Relationship between CSS and Salary of Police Officers

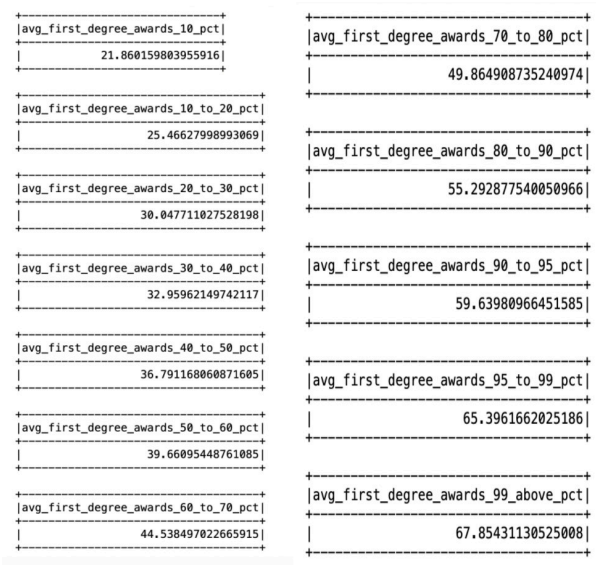
4 Graph Analytics

We observed that 'good' cops i.e. those belonging to buckets of low CSS percentile have a lower reward count.

We divide 'src' police officers under different buckets based on CSS percentile. TABLE (in Figure 5) represents the same for 10th, 25th, 50th, 75th, 95th and 99th percentile- The reward count increases as the CSS of cops increases, i.e. 'bad' cops are more rewarded which is a troubling sign and doesn't look good.

Furthermore (from Figure 6), the most connected officers i.e. those who have multiple counts of being co-accused with other officers have higher rewards as compared to the least connected officers.

'Bad' cops stick together and repeat offenses, yet they are heavily rewarded. Whereas, we see that 'good' cops are not rewarded a lot.



avg_first_degree_awards_10_pct
21.860159803955916

avg_first_degree_awards_10_to_20_pct
25.46627998993069

avg_first_degree_awards_20_to_30_pct
30.047711027528198

avg_first_degree_awards_30_to_40_pct
32.95962149742117

avg_first_degree_awards_40_to_50_pct
36.791168060871605

avg_first_degree_awards_50_to_60_pct
39.66095448761085

avg_first_degree_awards_60_to_70_pct
44.538497022665915

avg_first_degree_awards_70_to_80_pct
49.864908735240974

avg_first_degree_awards_80_to_90_pct
55.292877540050966

avg_first_degree_awards_90_to_95_pct
59.63980966451585

avg_first_degree_awards_95_to_99_pct
65.3961662025186

avg_first_degree_awards_99_above_pct
67.85431130525008

Figure 5: Distribution of awards by percentile

id	officer_name	allegation_count	CSS	award_count	pagerank
32442	John Zinchuk	23.0	12.8494	97	122.85961890322304
32440	Mark Zawila	34.0	13.4627	25	82.596692708314
32425	Perry Williams	27.0	11.5465	118	72.0279045279787
32410	Joseph Watson	29.0	10.8425	62	68.47949785853883
32430	Michael Wrobel	22.0	16.7592	165	67.9187831349547
32350	Robert Spiegel	20.0	13.0244	183	65.39503775647411
32351	Boonserm Srisuth	25.0	11.0281	75	63.01146602153837
32433	Kenneth Yakes	29.0	9.8227	53	62.051713039664
32284	Mark Reno	76.0	32.1284	142	61.834556455615996
32419	Eric Wier	18.0	7.1819	90	58.028946264273195
32384	Edwin Utreras	47.0	20.8848	43	57.37383713588734
32435	Mohammad Yusuf	22.0	15.2373	148	56.708114377778166
32074	Ronald Jenkins	46.0	13.5129	64	55.47174776302173
32431	Albert Wyroba	15.0	9.888	200	55.44272739297056
32337	Louis Silva	21.0	14.8829	131	55.12138144372036
32413	Carl Weatherspoon	69.0	23.1244	55	54.75312241275796
32289	John Rivera	44.0	12.922	66	53.69168834565139
32401	Joshua Wallace	45.0	22.8745	50	53.6111678756193
32375	James Triantafillo	31.0	18.6742	108	48.42330492775258
32436	Edmund Zablocki	28.0	16.2565	51	46.89728707125255

id	officer_name	allegation_count	CSS	award_count	pagerank
9427	Robert Gallegos	1.0	0.2248	0	0.3235123432519 66
30702	Bobby Williams	1.0	0.1727	0	0.3235123432519 66
812	Patrick Arens	10.0	1.0	6	0.3235123432519 66
4184	Edward Castellano	1.0	0.3044	1	0.3235123432519 66
9534	Gerardo Garcia	10.0	1.5739	8	0.3235123432519 66
16767	Gregory Luszowiak	2.0	0.6833	26	0.3235123432519 66
32434	Steven Yee	2.0	0.3359	25	0.3235123432519 66
32263	Jeffrey Pineda	1.0	0.8333	21	0.3235123432519 66
16144	Don Lewis	2.0	0.4546	2	0.3235123432519 66
32448	James Martin	5.0	2.0261	0	0.3235123432519 66
25570	John Scatchell	1.0	0.1827	4	0.3235123432519 66
3279	Robert Bullock	12.0	2.8877	38	0.3235123432519 66
178	Richard Aguilar	38.0	8.8356	4	0.3235123432519 66
545	Alvin Amos	15.0	3.2273	6	0.3235123432519 66
1640	Sanford Becker	1.0	0.1291	0	0.3235123432519 66
28534	Jeffery Thompson	4.0	0.5677	3	0.3235123432519 66
11388	Leon Hardeman	1.0	0.1815	3	0.3235123432519 66
17954	Garry Mc Carthy	2.0	2.831	2	0.3235123432519 66
15381	Donald Kumiga	2.0	0.2963	9	0.3235123432519 66
6021	William Czahor	3.0	0.4473	0	0.3235123432519 66

Figure 6: First Degree Connection Of Police Officers

5 Natural Language Processing

For the natural language processing checkpoint, we wanted to find out whether there is any relationship between the CSS metric and the sentiment polarity of allegation summaries. If there is indeed a relationship between these two variables, then it tells us that CSS may be a strong and reliable metric by itself.

However, a cloudy formation on the scatter plot and / or a high p-value and R-square on the regression model summary would mean that these two variables are not related and potentially closes this avenue of exploring any association between allegation summaries and CSS by themselves. We plot the scatter plot using a log scaled CSS metric with the compound score and find that there is a cloudy pattern as seen below:

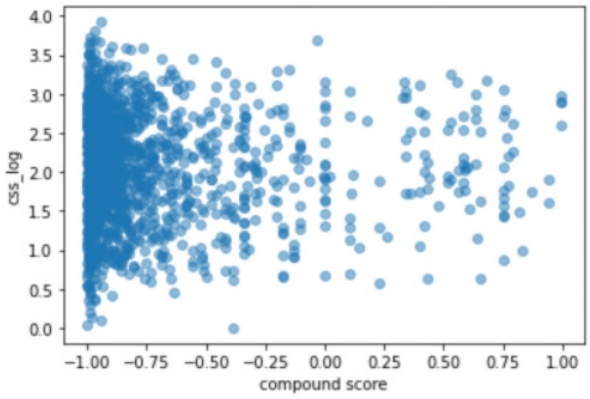


Figure 7: Sentiment Analysis

This is further supported by the regression analysis that we did taking the response as the log scaled CSS metric and predictor variables such as positive, negative, neutral probability scores and the compound polarity score. As seen below, the independent variables do not explain the variance in the dependent variable and their coefficients are not significant either:

OLS Regression Results

Dep. Variable:	css	R-squared:	0.002
Model:	OLS	Adj. R-squared:	-0.000
Method:	Least Squares	F-statistic:	0.7910
Date:	Wed, 01 Dec 2021	Prob (F-statistic):	0.531
Time:	21:37:53	Log-Likelihood:	-5575.2
No. Observations:	1681	AIC:	1.116e+04
Df Residuals:	1676	BIC:	1.119e+04
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	328.5081	398.948	0.823	0.410	-453.980	1110.997
compound	0.4046	0.561	0.721	0.471	-0.696	1.506
pos	-324.5849	398.748	-0.814	0.416	-1106.681	457.511
neg	-315.7638	398.819	-0.792	0.429	-1097.999	466.472
neu	-320.0278	399.016	-0.802	0.423	-1102.649	462.594

Omnibus:	507.668	Durbin-Watson:	1.917
Prob(Omnibus):	0.000	Jarque-Bera (JB):	1458.060
Skew:	1.555	Prob(JB):	0.00
Kurtosis:	6.339	Cond. No.	7.43e+03

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.43e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Figure 8: Regression Analysis

Hence, we can conclude that there is no significant association between CSS and sentiment of allegation summaries and therefore this analysis closes up future work along these lines.

6 Exploring Other Avenues Using Natural Language Processing

Officers may potentially spend a lot of time writing a report and filling up an allegation summary form when in fact their presence would be more valuable on the field fighting crime. To minimize their time writing such reports and filling up allegation summaries, we built an autocorrect algorithm using deep learning that will automatically predict and recommend the next 3 words for their report. This would allow them to finish up the writing faster, thereby minimizing the time needed on such administrative tasks.

We used a LSTM model to generate text character-by-character. After running the model for 25 epochs, the model started generating locally coherent sounding texts. We ran the entire code on GPU, as the recurrent networks are quite computationally intensive. The data which we used is very sparse(1681 rows) and we would have achieved a much better performance if our corpus had at least 100k characters. In the figure below we tried generating the next three words based on the inputs of the user.

```
...Diversity: 0.2
...Input Sentence: " battery, and subsequently pleaded guilt"
...Autocomplete:  y of the complainant

...Diversity: 0.5
...Input Sentence: " battery, and subsequently pleaded guilt"
...Autocomplete:  y of the department
```

Figure 9: Report AutoCompleter output

Diversity acts as a context regularization to basically show that how contextually relevant we want the auto generated texts to be based on the diversity score it generates the words. In our experiment, we found a diversity score of 0.2 to be a sweet spot where we are getting very coherent sounding auto generated texts. We think Chicago police can benefit heavily by using a system like this, which can help cut down the time it takes to file a report, which in turn will make them more productive by allowing them to focus on field work more.

7 Conclusion

'Good' police officers are not rewarded well. The majority of awards have an increasing trend where officers who have a higher CSS score tend to receive more awards than those who score lower. Reward categories and CSS of officers have a very non-uniform relationship. Despite having a high CSS, several cops have received awards. However, this could be due to officers being involved in high crime zones.

From our graph analysis, we found out that the most influential officers (based on pagerank) tend to have higher CSS values. In addition, the first degree connections of officers with a high CSS value also tend to have higher CSS scores than officers with lower CSS scores. This conclusion could have been strengthened if we had more dense values for sustained allegations, which we could have leveraged to find out whether an officer is indeed "bad".

8 References

1. Allegation Severity Weights - Prof. Jennie Rogers

2. Project by The Bold Earthquakes - https://github.com/Northwestern-Data-Sci-Seminar/Invisible-Institute-Chicago-Reporter-Collaboration-Public/tree/master/The%20Bold%20Earthquakes

3. Project by The Enchanted Badgers -https://github.com/Northwestern-Data-Sci-Seminar/Invisible-Institute-Chicago-Reporter-Collaboration-Public/tree/master/The%20Enchanted%20Badgers