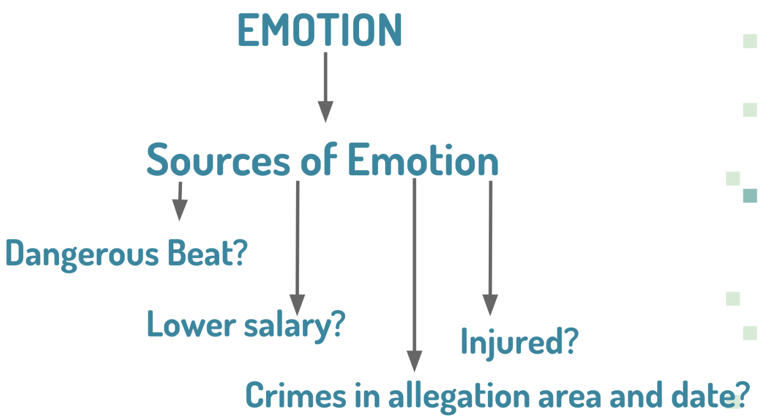
Project Report **GROUP 5**
MSAI 339 **Jack R**
Prof. Jennie Rogers **Quincia H**
December 11, 2018 **Ikhlas A**

Purpose

This report will identify what we've learned from building connections between the results of each checkpoint in this project. This report will also illustrate what we found interesting, surprising, and that which was regrettably, left unanswered.

Focus and Motivation

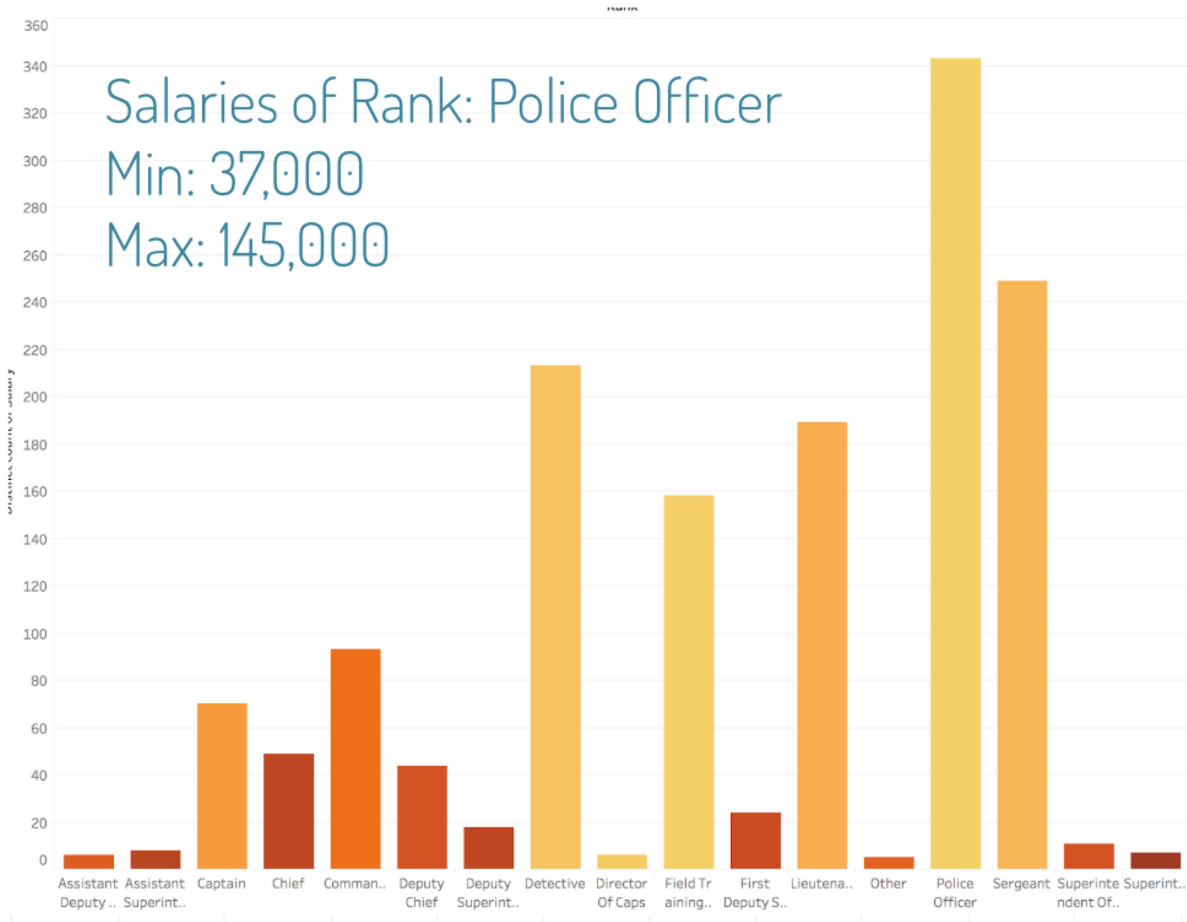
We wanted to stray away from the other common project theses in order to approach a connection that would provide more unique results. To accomplish this, we focused on the relationships between allegations, crimes, and *signs of compelling emotion*. Being a police officer requires maintaining discipline and composure in exceptionally stressful and potentially life-threatening situations. Police officers consistently deal with disobedience, frustration, and an ever-present worry for their safety. On the other hand, the motivation that pushes someone to file a complaint against an officer has the potential to be emotionally driven. While this doesn't mean all complaints are, there should be a unique threshold, or point until action, that a person must meet before they willingly carry out the process of filing a complaint. As emotion cannot yet be accurately quantified, we instead focused on separating and predicting how certain, more emotion-inducing, stimuli can have affected officers and complainants.



Crime and Allegation Relationship

Crime is often emotionally driven and it can quickly get out of hand when an officer approaches a suspect. A condition for this project to be successful was to accurately identify some correlation between crimes and allegations when integrating them together. Ultimately, we connected allegations to an array of crimes that happened on the same exact day and street as the allegation. While this does not reach the granular accuracy of specific times in the day, it allows us to form a relationship to work with.

Salary



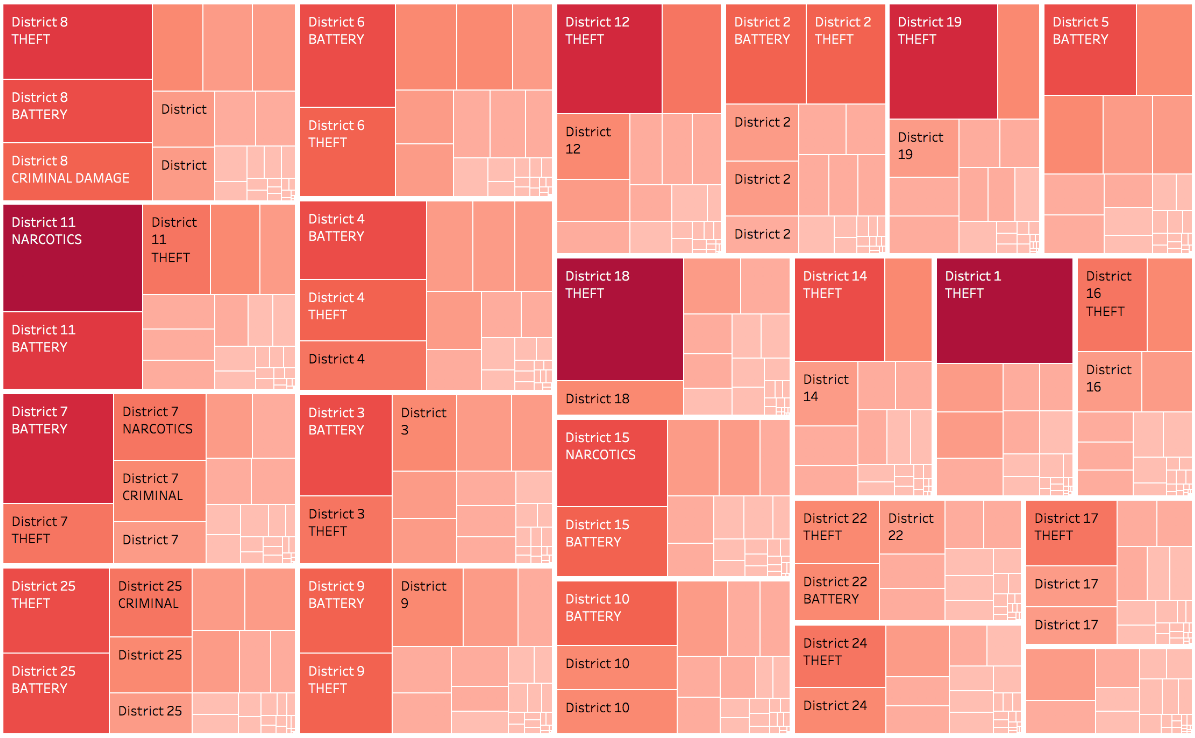
There have been numerous reports written by the *National Bureau of Economic Research* that catalog the relationship between an officer's performance and their compensation. We wanted to see if there was a relationship between police officer salary and the number of allegations against them. This would ultimately transfer to predicting the salary of an officer given the number of allegations against them.

Injury

One potential motivation for bystanders or victims of an unlawful officer to complain is whether someone was injured during the interaction. Regardless of the history of the suspect an officer was interacting with, if a bystander witnessed excessive force, they may be more inclined to file an allegation.

Transitional Anxiety

Crimes By District and Type



The term *Transitional Anxiety* stems from the time when an officer changes units to a more dangerous beat/district/unit. This is similar to being transferred to a new sector in a company or having to move to an entirely new location of a country, except this transition could be putting an officer in a much more dangerous environment. The labels in the figure above correspond to the district number plus one.

Checkpoint Takeaways

Relational Analysis

The first checkpoint developed the building blocks necessary for understanding the Chicago Police Database. Of the four questions we set out to answer, we paid the most attention to an officer's salary; we felt there was a strong relationship between motivation in work and the salary you are compensated with. The other three questions were indirectly replicated in every following checkpoint and beneficial for future results.

Data Cleaning and Integration

column2	column3	column4	column5	column6
01/24/2008	2008-01-24	W 63RD ST	W 63RD ST	6038616
01/24/2008	2008-01-24	W GRAND AVE	W GRAND AVE	6038641
01/23/2008	2008-01-23	S HALSTED ST	S HALSTED ST	6038642
01/22/2008	2008-01-22	S MICHIGAN AVE	S MICHIGAN AVE	6038659
01/01/2008	2008-01-01	S THROOP ST	S THROOP ST	6038668
01/23/2008	2008-01-23	W 77TH ST	W 77TH ST	6038672
01/20/2008	2008-01-20	S HALSTED ST	S HALSTED ST	6038689
01/23/2008	2008-01-23	S PULASKI RD	S PULASKI RD	6038695
06/15/2007	2007-06-15	S MICHIGAN AVE	S MICHIGAN AVE	6038709
01/24/2008	2008-01-24	S STATE ST	S STATE ST	6038750
01/24/2008	2008-01-24	S STATE ST	S STATE ST	6038824
01/24/2008	2008-01-24	W GRAND AVE	W GRAND AVE	6038886
01/24/2008	2008-01-24	S STATE ST	S STATE ST	6039037
01/24/2008	2008-01-24	S STATE ST	S STATE ST	6039044
01/22/2008	2008-01-22	S HALSTED ST	S HALSTED ST	6039045

The second checkpoint was paramount to exploring anything pertaining to crimes having an impact on an allegation. During this checkpoint, we were able to connect 82% of the allegations to corresponding crimes. What was interesting, however, was that the relationship of allegation to crime was one to many. We found that there were multiple crimes committed in the exact same location on the same day of each allegation. This connection was fundamental in the success of the following checkpoints, especially when machine learning was used to predict the allegation type.

Workflow Analytics

The workflow analytics checkpoint marked the first time during our project that we began to look for predictable attributes to complaints such as salary or high crime/allegation producing districts. There were two important takeaways from this checkpoint. The first was that there was a positive correlation between allegations and crimes in districts. This ultimately confirmed that even though there were over 6 million crimes recorded in the crime database, there are still certain districts that consistently see higher crime rates than others. If this was not the case, we would not be able to pursue transitional anxiety as a potential weight for the emotion heuristic. The other takeaway was that we did not find any overwhelming evidence that suggested that salary was predictable based off of the number of allegations against a given officer. Before turning any attention away from salary, it is important to note that the data was poorly distributed, and ultimately there were too many unique salary values for police officers. This was seen in the following checkpoints as well.

Machine Learning

We attempted to predict the following two features: *Allegation Category* and *Officer Salary.* An analysis of each can be seen below:

Allegation Category

```
#Use Random Forest Classifier for Allegation Category
assembler = VectorAssembler(inputCols=["IsOfficerComplaint","ARSON","ASSAULT","BATTERY","BURGLARY","CRIM SEXUAL ASSAULT", "CRIMINAL DAMAGE", "CRIMINAL TRESPASS", "DECEPTIVE PRACTICE", "GAMBLING", "INTERFERENCE WITH PUBLIC OFFICER", "KIDNAPPING", "LIQUOR LAW VIOLATION", "MOTOR VEHICLE THEFT", "NARCOTICS", "OFFENSE INVOLVING CHILDREN", "OTHER OFFENSE", "PROSTITUTION", "PUBLIC PEACE VIOLATION", "ROBBERY", "SEX OFFENSE", "STALKING", "THEFT", "WEAPONS VIOLATION"],outputCol="features")
output = assembler.transform(dt)
featureIndexer = VectorIndexer(inputCol="features", outputCol="indexedFeatures", maxCategories=4).fit(output)
(trainingData, testData) = output.randomSplit([0.7, 0.3])
labelIndexer = StringIndexer(inputCol="category", outputCol="indexedCategory").fit(output)
rf = RandomForestClassifier(labelCol="indexedCategory", featuresCol="indexedFeatures", numTrees=10)
labelConverter = IndexToString(inputCol="prediction", outputCol="predictedLabel",labels=labelIndexer.labels)
pipeline = Pipeline(stages=[labelIndexer, featureIndexer, rf, labelConverter])
model = pipeline.fit(trainingData)
predictions = model.transform(testData)
evaluator = MulticlassClassificationEvaluator(labelCol="indexedCategory", predictionCol="prediction", metricName="accuracy")
accuracy = evaluator.evaluate(predictions)
print("Test Error = %g" % (1.0 - accuracy))
```

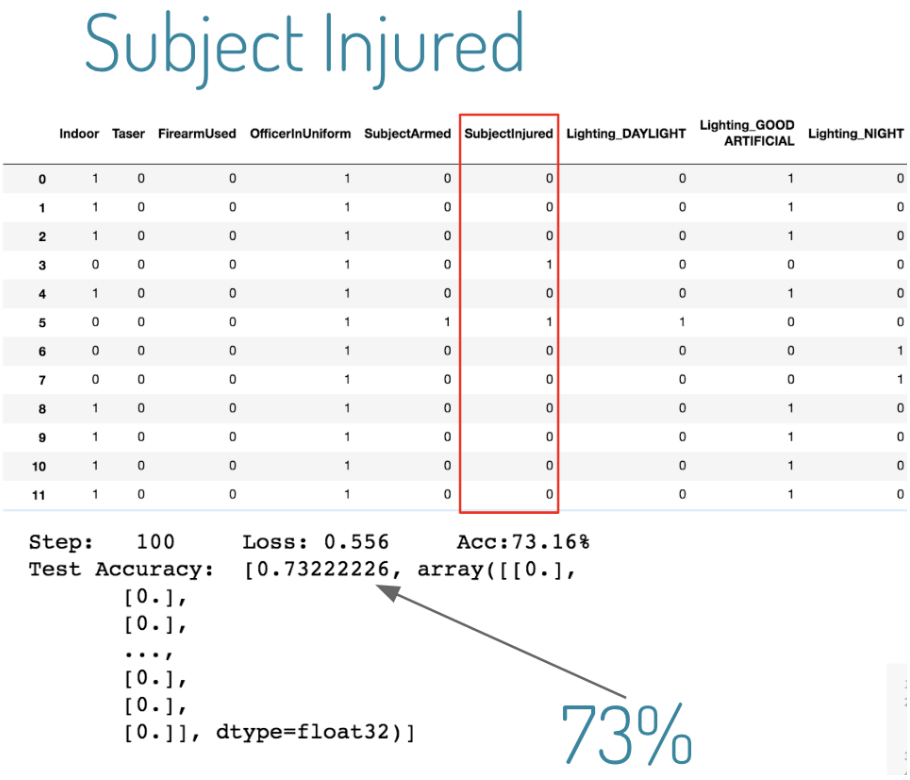
```
(14) Spark Jobs
output: pyspark.sql.dataframe.DataFrame = [Address: string, Date: timestamp ... 26 more fields]
trainingData: pyspark.sql.dataframe.DataFrame = [Address: string, Date: timestamp ... 26 more fields]
testData: pyspark.sql.dataframe.DataFrame = [Address: string, Date: timestamp ... 26 more fields]
predictions: pyspark.sql.dataframe.DataFrame = [Address: string, Date: timestamp ... 32 more fields]
Test Error = 0.646915
```

To predict the category of an allegation, we wanted to solely use crime data, not the other features of an allegation. This meant taking the array of crimes that were carried out in the same exact location and date and using them to predict the allegation category. To do this, we initially built a binary representation of all of the crime types combined. This ultimately turned into a 20-bit hexadecimal value because there could be multiple crimes with the same crime type. With this approach, we were able to get a 65% accuracy rating for the allegation category, which is great knowing that there were over twenty classes that the model needed to choose from. The baseline used mimicked a coin-flip, which would have had a 5% accurate for a "twenty-sided coin". We believe this could be a very beneficial model to look into in the future.

Officer Salary

To predict the salary of an officer, we used the officer and officer_allegation tables in the CPDB. Specifically, we focused on the number of allegations against an officer, ie. the complaint percentile. This ultimately proved to have negative results, showing less than 50% accuracy when predicting the salary. While we tried to bucket the salaries into $2500 increments, the accuracy still did not show to be a very promising indicator for the chance of complaint. We decided to take a step back from this prediction until we built off of other aspects of our project in hopes of finding a different feature set that would produce better results. This begs the question of whether or not individual officer salary actually affected the number of allegations they received, or if salary simply had much less of an impact as described by the *National Bureau of Economic Research.*

Modeling with Neural Networks



	Indoor	Taser	FirearmUsed	OfficerInUniform	SubjectArmed	SubjectInjured	Lighting_DAYLIGHT	Lighting_GOOD ARTIFICIAL	Lighting_NIGHT
0	1	0	0	1	0	0	0	1	0
1	1	0	0	1	0	0	0	1	0
2	1	0	0	1	0	0	0	1	0
3	0	0	0	1	0	1	0	0	0
4	1	0	0	1	0	0	0	1	0
5	0	0	0	1	1	1	1	0	0
6	0	0	0	1	0	0	0	0	1
7	0	0	0	1	0	0	0	0	1
8	1	0	0	1	0	0	0	1	0
9	1	0	0	1	0	0	0	1	0
10	1	0	0	1	0	0	0	1	0
11	1	0	0	1	0	0	0	1	0

```
Step:    100      Loss: 0.556      Acc:73.16%
Test Accuracy:  [0.73222226, array([[0.],
        [0.],
        [0.],
        ...,
        [0.],
        [0.],
        [0.]], dtype=float32)]
```


The neural network checkpoint focused on a single feature in the TRR data. This feature was *SubjectInjured*. This correlates to our interest in whether or not a suspect being injured in the process of being arrested played an important part in receiving an allegation against the corresponding officer. There are a lot of potential holes in the allegation, such as: If a bystander witnesses an injury without understanding the entire scope of the situation (for example, was this suspect considered armed and dangerous), are they more likely to file an allegation? To predict if a subject was injured, we used the TRR data provided in the CPDB. We ultimately isolated our feature input to using visual and subject descriptions only, producing a 75% accuracy in predicting if a subject will be injured. This could have been built off of further with more combinations of hidden layers and feature engineering.

Visualization

The major takeaway from the visualization checkpoint was that building a visual representation of the data can stimulate even more potential models and predictions. The three of us believe that if we had started this project with data visualization at its fore, we may have come up with more fruitful results by exploring predictions that were more promising. From this checkpoint, we were able to understand the distribution of crimes and unique officer allegations over the past 20 years, which gave insight on how well the Chicago Police are doing with reducing the crime rate.

Number of Crimes Per District Per Year

District1	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
									Date								
Null				2	1	1	2	40									1
1	14,854	17,971	17,571	17,578	17,697	16,950	15,648	15,850	13,394	12,525	11,819	12,249	12,196	11,569	11,991	13,717	15,375
2	35,401	27,046	25,096	23,758	21,959	20,185	19,366	17,774	16,746	15,951	15,323	13,606	12,825	11,514	10,726	11,401	11,559
3	24,310	23,556	23,579	24,278	24,221	23,566	23,111	21,613	20,552	19,700	18,452	17,847	16,021	13,756	13,044	12,304	12,635
4	25,115	25,831	26,143	26,142	24,986	25,984	25,457	24,973	23,047	21,200	21,415	20,005	17,917	16,692	15,827	14,969	14,694
5	20,006	21,091	19,950	20,499	20,369	19,948	20,820	19,240	18,005	16,277	15,958	15,432	13,969	12,583	11,328	11,602	11,722
6	25,478	25,314	25,460	26,104	25,613	25,150	26,997	26,037	23,346	21,392	20,658	19,433	18,309	16,517	16,012	16,163	16,509
7	26,600	28,125	27,864	28,735	27,465	27,484	27,456	27,422	23,382	22,613	21,251	20,315	18,206	15,673	15,733	14,184	13,794
8	31,342	32,092	31,706	31,399	31,349	32,263	31,098	30,840	28,626	26,505	25,363	22,691	20,253	18,247	17,285	17,486	16,505
9	24,080	25,312	25,293	23,699	22,866	23,498	21,716	20,952	19,397	18,513	18,452	16,811	14,885	13,515	12,710	12,638	11,735
10	19,641	19,460	19,086	20,364	19,269	18,310	18,331	17,717	15,910	16,113	14,960	15,205	13,965	12,552	11,739	12,536	12,462
11	27,651	30,063	30,017	29,323	29,240	28,301	26,557	25,681	23,457	22,578	21,590	22,064	21,936	20,626	19,473	18,591	17,983
12	24,878	24,678	25,000	24,226	22,290	22,042	20,913	20,501	18,136	17,264	16,723	16,022	14,094	12,597	12,319	14,050	13,459
14	22,542	22,389	20,595	19,023	16,956	17,271	16,358	16,522	15,728	15,185	13,468	12,661	10,803	9,457	8,938	10,320	10,035
15	20,118	20,443	21,331	20,259	19,134	20,418	20,347	19,124	17,927	16,706	15,515	14,564	13,805	12,927	11,715	11,378	10,157
16	16,012	16,057	15,875	14,523	14,233	14,290	14,151	14,989	13,557	11,943	11,361	10,979	10,620	9,550	9,411	9,436	8,908
17	14,484	14,430	13,975	13,058	13,136	13,057	11,926	12,560	11,648	11,280	10,456	9,783	8,547	7,430	7,716	7,768	7,888
18	22,516	22,315	20,379	20,172	18,759	18,415	18,147	17,730	16,073	15,192	14,544	14,347	12,913	11,507	11,358	13,203	15,066
19	22,755	22,901	21,485	21,038	19,528	18,883	18,590	18,510	17,048	16,173	15,477	15,771	14,064	12,199	11,559	12,238	12,041
20	9,987	9,152	8,258	8,369	8,728	8,180	7,301	7,017	6,296	6,102	5,758	5,739	4,907	4,331	4,262	4,399	4,585
21			1	3													
22	14,952	15,674	14,959	15,438	15,110	14,861	14,980	14,292	13,549	12,764	11,686	10,903	10,276	8,801	8,732	8,585	8,323
24	15,110	15,020	13,933	14,396	14,323	14,186	12,993	12,848	11,846	11,256	10,209	9,618	8,737	7,549	7,023	7,341	8,177
25	27,918	27,825	28,370	26,990	26,468	24,849	24,719	24,785	24,975	23,036	21,291	19,877	17,784	15,616	15,031	14,501	14,209
31		8	11	6	5	7	15	21	20	25	22	4	5	3	9	14	2

In fact, what's especially interesting is that all but one district has shown considerably less crime rates since 2001. This checkpoint also helped us visually understand the relationship between complaint percentile and transitional anxiety. While we did not develop a model to quantify or predict this relationship, it could be seen that officers that switched to a district that was more dangerous than the previous one had a higher complaint percentile. That being said, this may be an informal fallacy drawing this connection, and further analysis is needed. Additionally the opposite should be considered as well, where officers being transferred to less dangerous districts may carry over experience and history, possibly leading to more aggressive acts for the beat and thus raising a complaint percentile from citizens that were accustomed to previous, more lax, police behavior from that area.

Surprises, Interesting Results, and Open Questions

Aggregating all of our results, there were a few things that stuck out. Firstly, we did not expect such strong results in predicting allegation category given the crimes on a given day and location. Knowing there were over 20 classes and millions of combinations of the 20-bit representation we created as a feature for the crime types, it was really surprising to see a 65% accuracy rating. This makes us wonder if we were to build representations for other features provided in the crime data, would we have been able to build an even more accurate model? Another result that was surprising was the negative results we consistently received when we delved into salary predictions. This could be due to inflation, the imbalance of data in the officer database, and laws pertaining to wages developed in Illinois. This makes us question how to replicate the results seen in the papers published by the *National Bureau of Economic Research*.

Conclusive Remarks, Insights and Future Work

To sum up the insights we gleaned from this project, we compare our results from looking at the various objective measures we can make about emotional behavior given some knowledge of an incident and correlate them to whether or not a complainant will file an allegation against an officer. We had approached our thesis by initially reading 'Negative Emotions and Their Effect on Customer Complaint Behavior' by Tronvoll, and learned that frustration was one of the strongest predictors for complaint behavior, of the most common negative emotions. We tried to speculate on the most common sources of frustration that could be present due to police or crime behavior. We learned that there is a positive correlation between crimes and allegations in a specified area and time, our first objective measure of a source of emotion. This meant that complainants who repeatedly witnessed or were exposed to crimes and/or police behavior in the same area and time were more likely to file an allegation. We also discovered that both the allegation category prediction and the dangerous beat analysis could be factored into a larger linear combination to represent emotional impact on an allegation. Some objective sources we thought would play a major factor in an officer receiving a complaint did not produce any conclusive evidence. Startlingly, an officer's salary appears to have no effect on the number of allegations they receive. Similarly for injury, while we can tell that approximately 75% of suspects will be injured in a filed TRR, we were not able find out if a chance of injury led to greater complaints (although we expect it to be a very important predictor). We ultimately saw that emotion could be quantified through heuristics represented by a linear combination of a variety of factors. This approach was similar to Ensemble Learning, a method of using multiple learning algorithms to obtain better predictors than one could achieve with a sole learner, but instead expanding to multiple predictions that would intend to sum toward a single overarching predictor. This should encapsulate similar themes of Ensemble Learning such as lower bias, rate of error, and overfitting. There were *many* questions left unanswered as we attempted to complete our project. To state a few, was there a way to tackle the issue of officer salary not having an effect on complaints that we did not come across? Frustration is not the only major negative emotion, and the article we read summarized customer behavior, so what would we have learned had we pursued the other negative emotions such as fear or anger? How did officers who transferred into a less stress-inducing district/job perform? What other visualizations did we miss that we could have used to better map our thesis of tying emotion to allegation? We came out asking more questions than we had going into this project, and we believe that while it is not the easiest task to study emotional behavior, it may also be one of the least researched. We would encourage other data scientists to consider approaching the Chicago Police Database by factoring in emotional drive, and see what discoveries it may unearth.