Spectacular Cranes CP5 Findings

Note: Much of our analysis and “findings” are written in the submission notebook.

Question 1: Cluster allegations and try to find cluster and group allegations based on similar inputs (i.e. maybe there are certain allegations that are commonly being changed to certain new categories with certain officers).

Findings were limited with this approach, although there may be some merit to the idea if one has enough time to inspect the clusters more. The model seemed to have group allegations based mostly on a particular column. In some cases, the column was **old_category**, in other cases, it was **new_allegation_name.** Though difficult to verify, for allegations that had categories/names that differed from the predominant category/name in the cluster, they may have been grouped there for a more interesting reason, such as officers being co-accused with the rest of the officers showing up on allegations in that cluster.

	officer_id	officer_fname	officer_lname	complaint_percentile	crid	old_category	new_allegation_name	allegation_mapping_boolean
1277	22695	Mary	Platt	54.6237	1020838	CLOSED HAND STRIKE PUNCH	Use Of Profanity	False
4154	10931	Bolivar	Guaman	65.3815	1035341	INADEQUATE FAILURE TO PROVIDE SERVICE	Reports	False
4455	8072	Rodrigo	Espinoza	72.1142	1037593	TASER PROBE DISCHARGE	Excessive Force On Duty No Injury	False
4466	32067	Eric	Jackson	95.9650	1037717	TASER PROBE DISCHARGE	Excessive Force On Duty No Injury	False
4467	32269	Victor	Portis	91.3683	1037717	TASER PROBE DISCHARGE	Excessive Force On Duty No Injury	False
4672	28358	Tewelde	Tesfai	65.3815	1038974	TASER PROBE DISCHARGE	Miscellaneous	False
4673	18275	Shane	Mc Hugh	72.9854	1038974	TASER PROBE DISCHARGE	Use Of Profanity	False
4674	19801	George	Moussa	95.8695	1038974	TASER PROBE DISCHARGE	Use Of Profanity	False
4675	28358	Tewelde	Tesfai	65.3815	1038974	TASER PROBE DISCHARGE	Use Of Profanity	False
4695	6713	Michael	Demcak	96.6451	1039073	TASER PROBE DISCHARGE	Use Of Profanity	False

One can see above that most allegations in this cluster (6) have to do with TASER. But the first two allegations may have been assigned to this cluster for a different reason.

Question 2: Predict Allegation count monthly by allegation category by looking at monthly allegation data.

We collected the Allegation Data from CPDB database for count of allegation under different category type by year from 2001 until the latest data available year. We cleaned and transformed the data to replace Null/NaN values and removed the bad Category type like Unknown to make it easier for the model to train and test.

Data Before Cleaning

▸ (1) Spark Jobs

Out[229]:

category year	Bribery / Official Corruption	Conduct Unbecoming (Off-Duty)	Criminal Misconduct	Domestic	Drug / Alcohol Abuse	Excessive Force	False Arrest	First Amendment	Illegal Search	Lockup Procedures	Operation/Personnel Violations	Racial Profiling	Supervisory Responsibilities	Traffic	Unknown	Use Of Force	Verbal Abuse
2001.0	21.0	324.0	308.0	735.0	39.0	NaN	237.0	NaN	1978.0	857.0	3708.0	NaN	294.0	657.0	NaN	2246.0	895.0
2002.0	13.0	207.0	61.0	736.0	28.0	NaN	511.0	7.0	1910.0	508.0	3684.0	NaN	224.0	431.0	NaN	2278.0	905.0
2003.0	13.0	168.0	59.0	444.0	27.0	NaN	470.0	NaN	1834.0	346.0	3418.0	2.0	195.0	216.0	NaN	1800.0	804.0
2004.0	6.0	127.0	80.0	365.0	30.0	NaN	451.0	2.0	1988.0	428.0	2752.0	5.0	235.0	81.0	NaN	1448.0	347.0
2005.0	14.0	137.0	80.0	277.0	33.0	NaN	463.0	6.0	1102.0	332.0	1880.0	8.0	221.0	52.0	NaN	1603.0	60.0
2006.0	20.0	141.0	54.0	247.0	28.0	NaN	686.0	6.0	1506.0	479.0	2264.0	16.0	129.0	193.0	NaN	1574.0	110.0
2007.0	36.0	167.0	71.0	190.0	30.0	NaN	794.0	10.0	2072.0	463.0	3052.0	8.0	49.0	205.0	NaN	1963.0	216.0
2008.0	36.0	175.0	198.0	237.0	45.0	NaN	569.0	NaN	1623.0	466.0	3233.0	2.0	50.0	174.0	NaN	1627.0	184.0
2009.0	67.0	115.0	136.0	291.0	27.0	NaN	631.0	2.0	2014.0	432.0	2410.0	NaN	38.0	129.0	NaN	1828.0	150.0
2010.0	26.0	116.0	110.0	253.0	28.0	NaN	512.0	NaN	1486.0	419.0	2179.0	NaN	64.0	97.0	NaN	1645.0	186.0
2011.0	26.0	96.0	25.0	232.0	33.0	1.0	452.0	1.0	1162.0	324.0	1975.0	NaN	41.0	117.0	NaN	1465.0	160.0
2012.0	24.0	91.0	21.0	201.0	41.0	1.0	553.0	3.0	974.0	320.0	1877.0	NaN	35.0	107.0	NaN	1088.0	137.0
2013.0	20.0	189.0	31.0	182.0	24.0	5.0	622.0	NaN	1028.0	437.0	1508.0	NaN	22.0	85.0	NaN	870.0	127.0
2014.0	28.0	89.0	28.0	110.0	16.0	3.0	640.0	2.0	1017.0	376.0	1491.0	NaN	30.0	112.0	NaN	705.0	76.0
2015.0	19.0	88.0	38.0	73.0	27.0	6.0	324.0	NaN	664.0	326.0	1426.0	NaN	18.0	155.0	NaN	415.0	65.0
2016.0	3.0	17.0	61.0	30.0	10.0	5.0	51.0	NaN	145.0	106.0	495.0	2.0	7.0	46.0	1.0	144.0	30.0
2017.0	NaN	1.0	NaN	17.0	NaN	2.0	56.0	NaN	28.0	NaN	10.0	NaN	NaN	NaN	NaN	194.0	11.0
2018.0	NaN	NaN	NaN	1.0	NaN	NaN	3.0	NaN	NaN	NaN	1.0	1.0	NaN	NaN	NaN	9.0	NaN

Command took 0.93 seconds -- by sundar@northwestern.edu at 12/11/2019, 7:45:17 PM on My Cluster

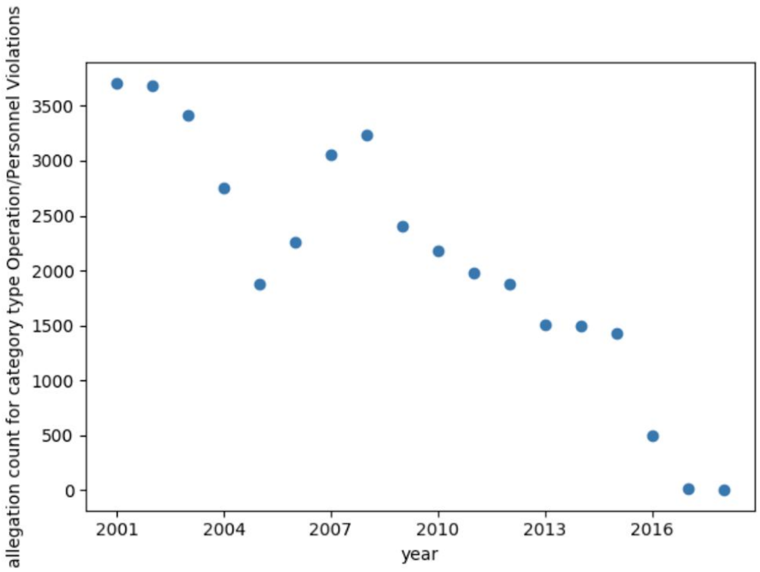
Data after cleaning

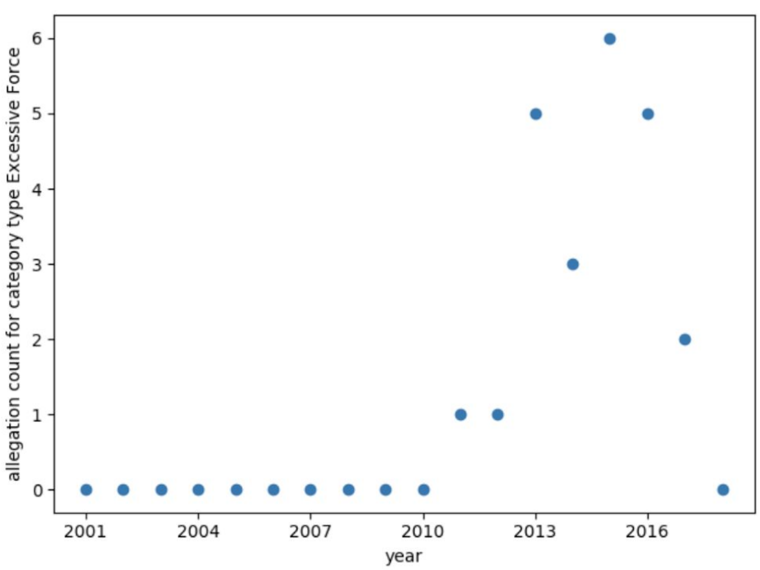
Out[343]:

category year	Bribery / Official Corruption	Conduct Unbecoming (Off-Duty)	Criminal Misconduct	Domestic	Drug / Alcohol Abuse	Excessive Force	False Arrest	First Amendment	Illegal Search	Lockup Procedures	Operation/Personnel Violations	Racial Profiling	Supervisory Responsibilities	Traffic	Use Of Force	Verbal Abuse
2001.0	21.0	324.0	308.0	735.0	39.0	0.0	237.0	0.0	1978.0	857.0	3708.0	0.0	294.0	657.0	2246.0	895.0
2002.0	13.0	207.0	61.0	736.0	28.0	0.0	511.0	7.0	1910.0	508.0	3684.0	0.0	224.0	431.0	2278.0	905.0
2003.0	13.0	168.0	59.0	444.0	27.0	0.0	470.0	0.0	1834.0	346.0	3418.0	2.0	195.0	216.0	1800.0	804.0
2004.0	6.0	127.0	80.0	365.0	30.0	0.0	451.0	2.0	1988.0	428.0	2752.0	5.0	235.0	81.0	1448.0	347.0
2005.0	14.0	137.0	80.0	277.0	33.0	0.0	463.0	6.0	1102.0	332.0	1880.0	8.0	221.0	52.0	1603.0	60.0

Command took 0.07 seconds -- by sundar@northwestern.edu at 12/11/2019, 9:07:09 PM on My Cluster

Plots for different Category Types





Trained the model using sklearn Linear Regression model

Command took 0.02 seconds -- by sundar@northwestern.edu at 12/11/2019, 8:59:38 PM on My Cluster

Cmd 28

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
regressor = LinearRegression()
regressor.fit(x_train, y_train)
y_pred = regressor.predict(x_test)
print(x_test,y_pred)
```

```
[[2007]
 [2004]
 [2014]
 [2003]
 [2015]
 [2008]] [[ 0.6352381 ]
 [ 0.10380952]
 [ 1.8752381 ]
 [-0.07333333]
 [ 2.05238095]
 [ 0.81238095]]
```

Command took 0.03 seconds -- by sundar@northwestern.edu at 12/11/2019, 8:52:18 PM on My Cluster

Cmd 29

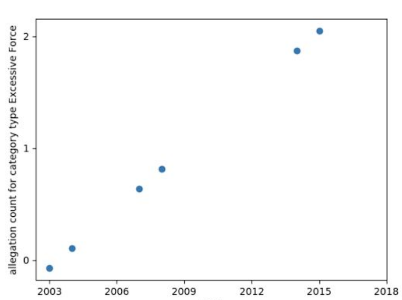
Analysis : Our model worked but we couldn;t find great correlation on the predicted data and the existing data which we attribute to insufficient and randomness in the data of the allegation count. We also conclude that further collaboration with Invisible Institute and investigation is needed for fine tuning the model and pruning the data.

Thoughts

After running regression on multiple category types, we learned that the category type allegation data isn't really in order and linear regression models can't really predict the number of allegations by category per year as shown below

Cmd 30

```
fig, ax = plt.subplots()
ax.locator_params(integer=True)
plt.ylabel("allegation count for category type "+category_type)
plt.xlabel("year")
ax.scatter(x_test, y_pred)
#plt.hist(list(allegation_df["Bribery / Official Corruption"].index), allegation_df["Bribery / Official Corruption"])
plt.xticks(np.arange(min(x_test), max(x_test)+5, 3.0))
display()
```



Command took 0.36 seconds -- by glgasser@gmail.com at 12/11/2019, 9:11:26 PM on My Cluster

Cmd 31

Question 3: Compare allegation category strings where the mapping was violated (suspicious changes). Of these allegations, which ones have similar category strings (using cosine similarity), meaning that although the category changed, the names of the categories still mean roughly the same thing.

Here, for each allegation, we converted the allegation's **old_category** and **new_allegation_name** to tfidf vectors in order to try to compare their similarity. The entire premise of this approach was to try to identify category changes where the categories are completely different (don't have similar meanings in vector space.) The main thing we found, which is consistent with a pattern we had noticed in previous analysis, is that the category: MISCELLANEOUS is commonly used to mask the content of allegations. Of course, the word miscellaneous is boring and uninformative, so it doesn't look bad at all.

	officer_id	officer_fname	officer_lname	complaint_percentile	crid	old_category	new_allegation_name	allegation_mapping_boolean
2	10322	Joel	Gonzalez	89.8819	1012611	MISCELLANEOUS	Inadequate Failure To Provide Service	False
3	31291	Willard	Wright	52.8214	1012622	EEO INVESTIGATIONS	Miscellaneous	False
4	16729	Jerry	Luke	37.9466	1012622	EEO INVESTIGATIONS	Miscellaneous	False
5	20033	Kevin	Murphy	80.4570	1012622	EEO INVESTIGATIONS	Miscellaneous	False
6	31291	Willard	Wright	52.8214	1012622	EEO INVESTIGATIONS	Miscellaneous	False
7	16729	Jerry	Luke	37.9466	1012622	EEO INVESTIGATIONS	Miscellaneous	False
8	20033	Kevin	Murphy	80.4570	1012622	EEO INVESTIGATIONS	Miscellaneous	False
9	13604	Brian	Johnson	97.5202	1012629	CHOKED	Excessive Force On Duty No Injury	False
10	7233	John	Dougherty	92.7369	1012629	CHOKED	Excessive Force On Duty No Injury	False
11	13604	Brian	Johnson	97.5202	1012629	CHOKED	Excessive Force On Duty No Injury	False
12	7233	John	Dougherty	92.7369	1012629	CHOKED	Excessive Force On Duty No Injury	False
13	30336	Kenneth	Weller	82.4004	1012648	MISCELLANEOUS	Excessive Force On Duty No Injury	False
14	15657	Timothy	Lange	64.2047	1012648	MISCELLANEOUS	Excessive Force On Duty No Injury	False
16	27997	Brian	Swiatkowski	98.5896	1012654	PUSHPULLGRAB	Use Of Profanity	False
17	5328	Markee	Cooper	95.2734	1012654	PUSHPULLGRAB	Use Of Profanity	False
18	12177	John	Higgins	88.5216	1012654	PUSHPULLGRAB	Use Of Profanity	False
19	27997	Brian	Swiatkowski	98.5896	1012654	PUSHPULLGRAB	Unnecessary Physical Contact On Duty No Injury	False
20	5328	Markee	Cooper	95.2734	1012654	PUSHPULLGRAB	Unnecessary Physical Contact On Duty No Injury	False
21	12177	John	Higgins	88.5216	1012654	PUSHPULLGRAB	Unnecessary Physical Contact On Duty No Injury	False
22	27997	Brian	Swiatkowski	98.5896	1012654	PUSHPULLGRAB	Fail To Obtain A Complaint Register Number	False
23	5328	Markee	Cooper	95.2734	1012654	PUSHPULLGRAB	Fail To Obtain A Complaint Register Number	False
24	12177	John	Higgins	88.5216	1012654	PUSHPULLGRAB	Fail To Obtain A Complaint Register Number	False
25	23190	Hakeem	Qazi	80.8987	1012689	PUSHPULLGRAB	Excessive Force On Duty No Injury	False
26	909	Lindbergh	Askew	78.5694	1012690	MISCELLANEOUS	SecondarySpecial Employment	False
27	1360	Donna	Barnes Simmons	49.6573	1012691	MISCELLANEOUS	SecondarySpecial Employment	False
28	1586	Ronald	Beach	56.9371	1012692	MISCELLANEOUS	Seat Belts	False
29	3914	Dena	Carli	33.6288	1012693	MISCELLANEOUS	SecondarySpecial Employment	False
30	9019	Derek	Fowler	69.2513	1012694	MISCELLANEOUS	Seat Belts	False
31	13939	Harold	Junior	51.4165	1012695	MISCELLANEOUS	Seat Belts	False
32	18213	Anthony	Mc Gowan	63.1462	1012696	MISCELLANEOUS	Seat Belts	False

One can see “MISCELLANEOUS” showing up often under **old_category** and “Miscellaneous” showing up under **new_allegation_name**. There are still quite a few interesting ones, such as "PUSHPULLGRAB" being changed to "Fail To Obtain A Complaint Register Number.” Clearly these are unrelated and the severity of these is different.

Question 4: Compare the sentiment of the old category strings vs. the new category strings (using sentiment as a proxy for severity). If the sentiment of the new category is more positive (less negative) than the previous category, that might imply that the category was changed to something less severe.

Here we thought it would be interesting to try and see which allegations were having their categories changed to something “less severe.” We decided to use *sentiment* as a proxy for severity: if the sentiment is more positive, the category is less severe (and vice versa).

Sentiment (most allegations will probably be negative 👎)

- Excessive Force: `-.25`
- Unnecessary Physical Contact On Duty No Injury: `-.2`
- Use Of Profanity: `0`
- Seat Belts: `0`
- We will compare how the sentiment (proxy for severity) changed from the old allegation categories to the new allegation categories

One can see “Excessive Force” has negative sentiment (more negative than Unnecessary Physical Contact which sort of makes sense). Use of Profanity and Seat Belts are considered neutral.

We then calculated the absolute difference between the old category and the new category, getting an idea of how the sentiment (severity) changed.

	officer_id	officer_fname	officer_lname	complaint_percentile	crid	old_category	new_allegation_name
334	5323	Ronald	Cooper	74.7927	1014535	CONDUCT UNBECOMING	Sexual Orientation
335	18907	Tomi	Methipara	73.2184	1014535	CONDUCT UNBECOMING	Sexual Orientation
336	5323	Ronald	Cooper	74.7927	1014535	CONDUCT UNBECOMING	Sexual Orientation
337	18907	Tomi	Methipara	73.2184	1014535	CONDUCT UNBECOMING	Sexual Orientation
702	28925	Joseph	Tripoli	84.5898	1017193	RACIAL ETHNIC ETC	Sexual Orientation
703	8592	Frank	Fish	52.0271	1017193	RACIAL ETHNIC ETC	Sexual Orientation
762	24506	Stan	Rogers	74.7851	1017952	SEXUAL ORIENTATION	Excessive Force On Duty No Injury
763	16660	Lawrence	Lowrey	71.4570	1017952	SEXUAL ORIENTATION	Excessive Force On Duty No Injury
764	13796	Nancy	Jones	53.9163	1017952	SEXUAL ORIENTATION	Excessive Force On Duty No Injury
765	23351	Brandon	Rambert	74.5214	1017952	SEXUAL ORIENTATION	Excessive Force On Duty No Injury
798	27725	Mark	Struke	82.2552	1018181	INADEQUATE FAILURE TO PROVIDE SERVICE	Properly Direct Subordinate
799	13909	Ted	Jozefczak	95.2046	1018181	INADEQUATE FAILURE TO PROVIDE SERVICE	Properly Direct Subordinate
800	17134	Melissa	Malm	78.4953	1018181	INADEQUATE FAILURE TO PROVIDE SERVICE	Properly Direct Subordinate
818	13637	Robert	Johnson	95.8695	1018246	RACIAL ETHNIC ETC	Sexual Orientation
819	32383	Armando	Ugarte	99.6981	1018246	RACIAL ETHNIC ETC	Sexual Orientation

It looks like many of the major differences have to do with Sexual Orientation. In this analysis, we did not find particular allegations that were grossly different. Given enough time, further analysis could be done where one uses a different sentiment analysis model, or uses the **Severity Analysis** done by another group. Using the **Severity Analysis**, one could truly compare whether allegation categories where changed from something severe to something less severe (like "Seat Belts" 😑)