Exploring the relationship between misconduct and career development

The Silent Fox

-RESULT AND ANALYSIS

Theme

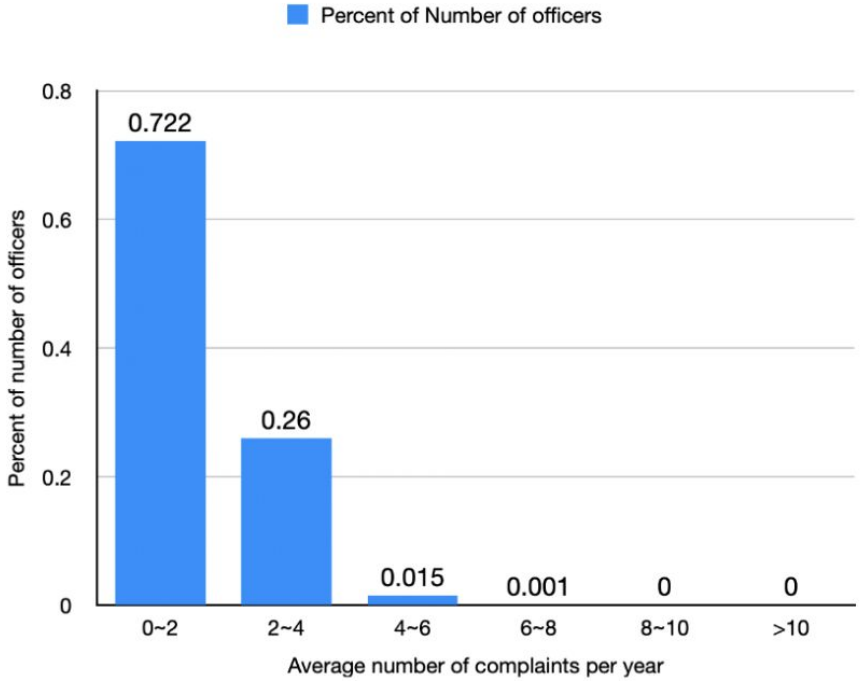
- Our theme is to explore the relationship between misconduct in police officers and their career development.
- To be more specific, we are going to find if the misconduct in enforcement of a police officer can be reflected properly by his/her career.
  - Salary
  - Award
  - Rank

About our research

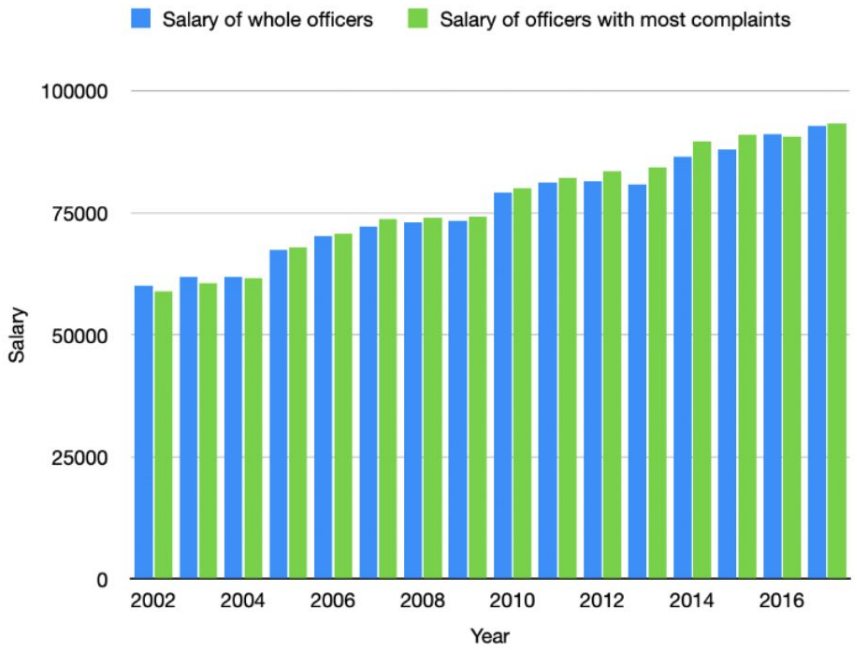
1. Relational Analytics
2. Data exploration
3. Interactive visualization
4. Graph Analytics
5. Natural Language Processing

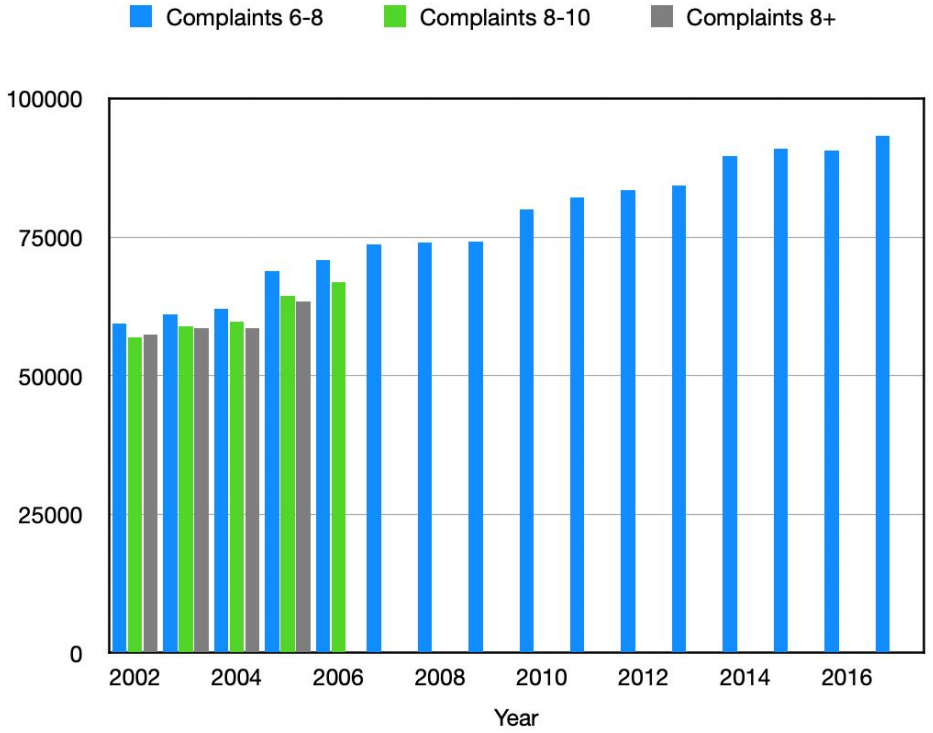
Relational Analytics

Some statistics of salary for the officers with the most complaints



Relational Analytics





Relational Analytics

What percentage of total complaints are these officers who have received an honor mention or award responsible for?

- Total number of complaints from citizens:
  - 235263
- Total number of complaints received by officers who get awards:
  - 225182
- Total number of officers who received complaints:
  - 23098
- Total number of officers who received both complaints and awards:
  - 20818

- **Percentage of complaints:**
  - 0.9572
- **Percentage of officers:**
  - 0.9013

Relational Analytics

Can the total number or frequency of complaints received by the officer be reflected by his/her career advancement(whether an officer holding an award in a year receives less or no complaint)?

- Since the award record in the database starts at 2005, we only analyze the complaints and awards received by officers from 2005 to 2008
- We divide the data into 3 sections:
  - 2005-2006: 75.36% decrease from 1315 to 324 (number of officers)
  - 2006-2007: 77.28% decrease from 911 to 207
  - 2007-2008: 64.03% decrease from 934 to 336

Relational Analytics

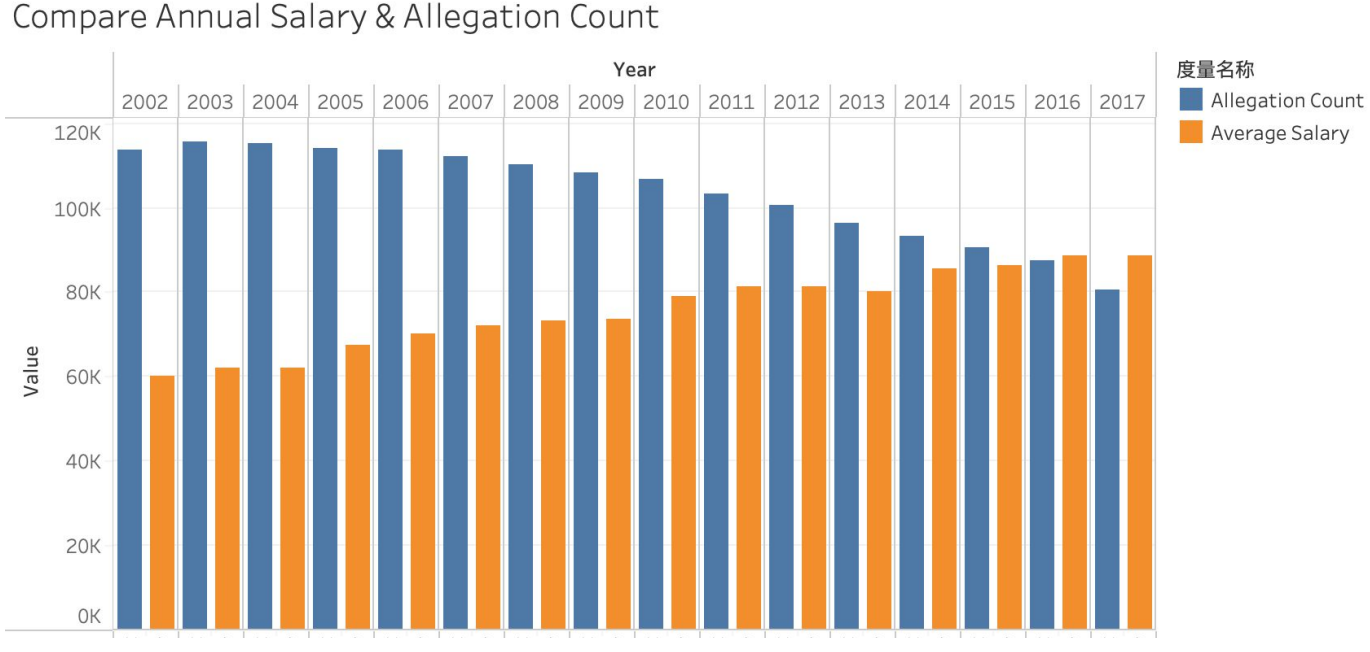
Is there a relationship between the change of frequency of an officer who gets allegations and the change of his/her salary(if any)?

avg_2005	avg_2006	avg_2007	avg_2008
67448.293124863009	70243.065004324013	72148.680703277621	73078.770459081836

sum_2005	sum_2006	sum_2007	sum_2008
6720	7045	8687	8819

Data Exploration

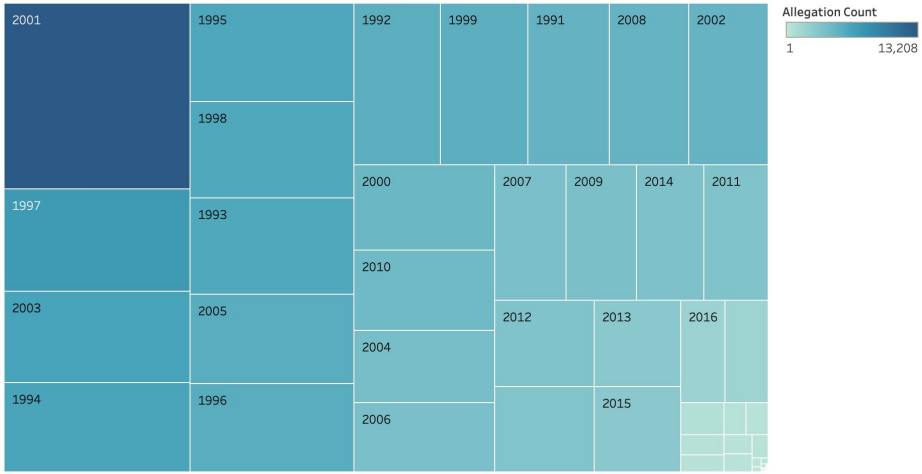
Is there a correlation between the salary of the officer and the number of complaints?



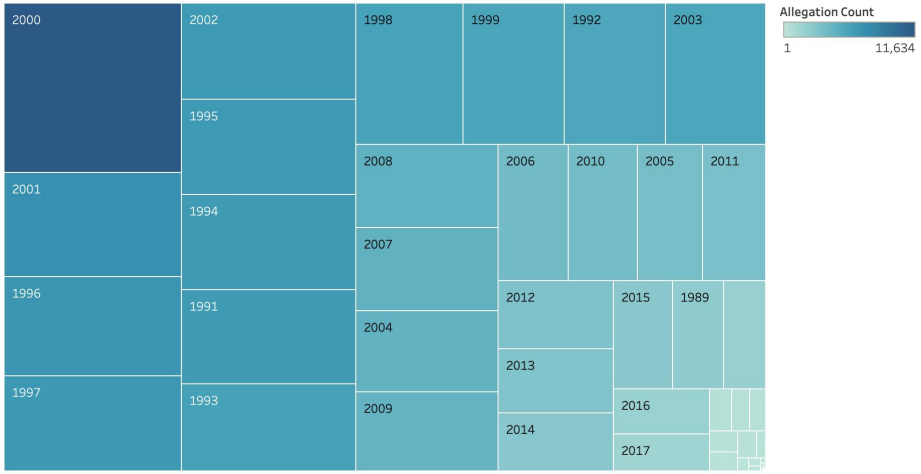
Data Exploration

Is there a correlation between the number of allegations received before an officer's first settlement case and number of allegations received after?

Allegation Annual Settlement Count

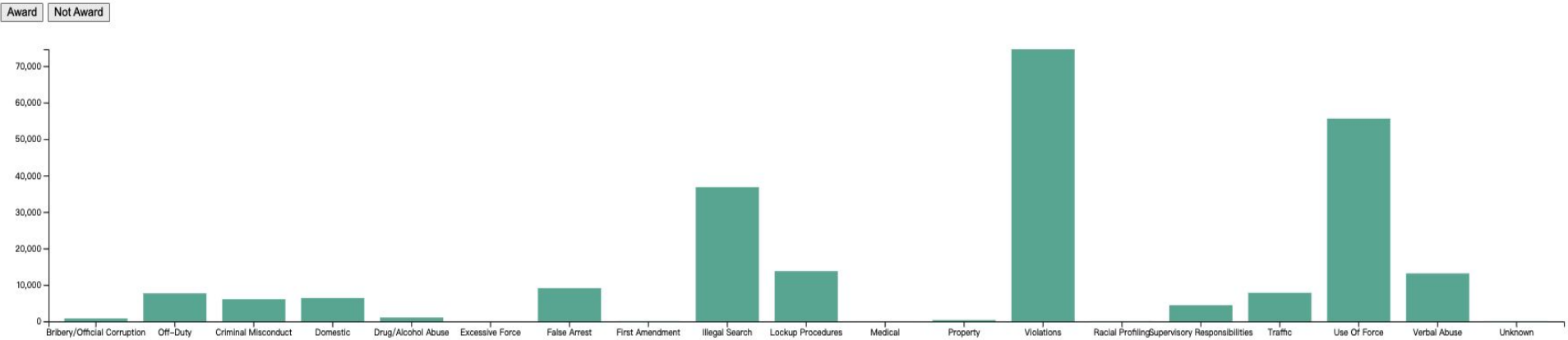


Allegation Annual Received Count



Interactive Visualization

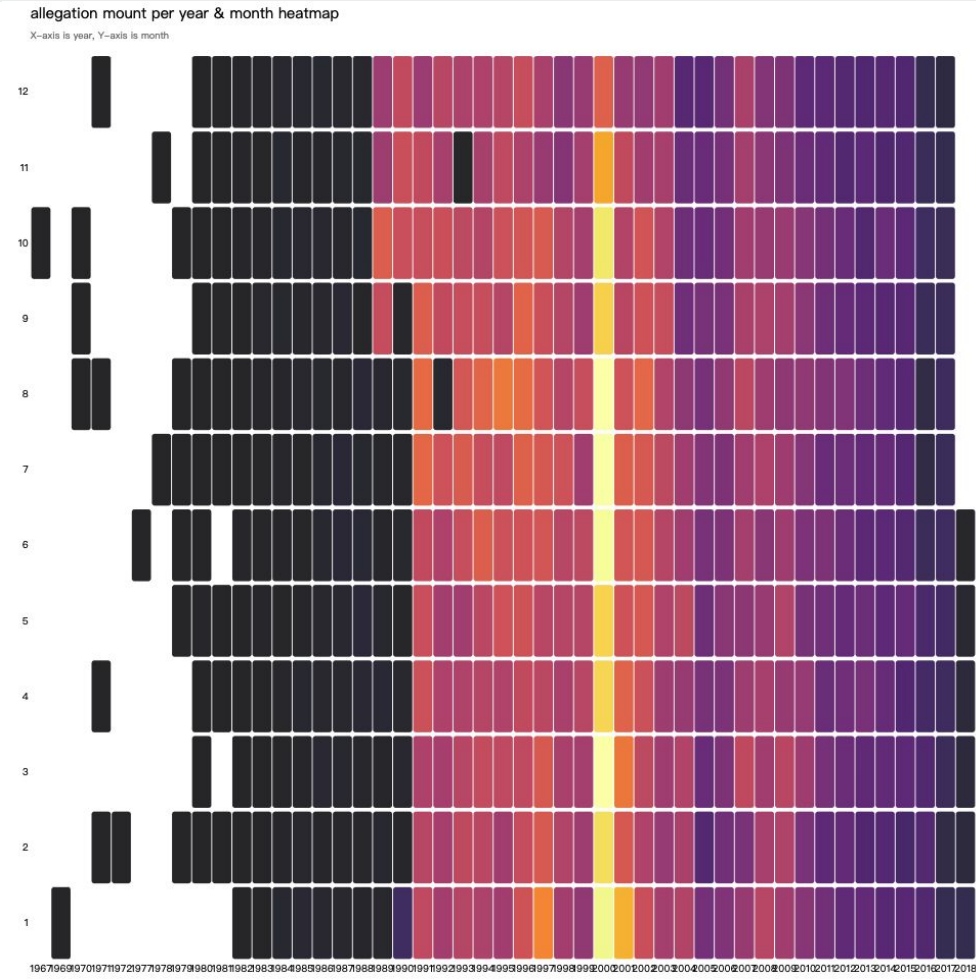
What is the distribution of the categories of the misconducts (illegal search, use of force, etc.) for officers with or without awards?



Interactive Visualization

Can we compare misconduct maps according to different time lines(monthly, yearly)?

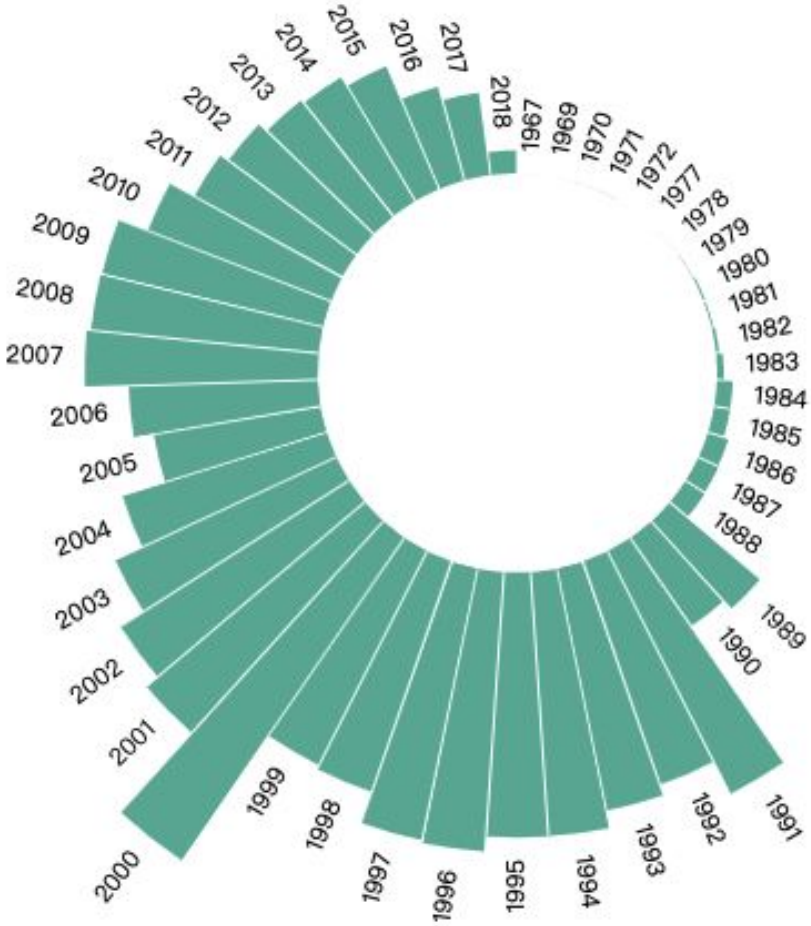
- Heatmap of allegation amount over years and months



Interactive Visualization

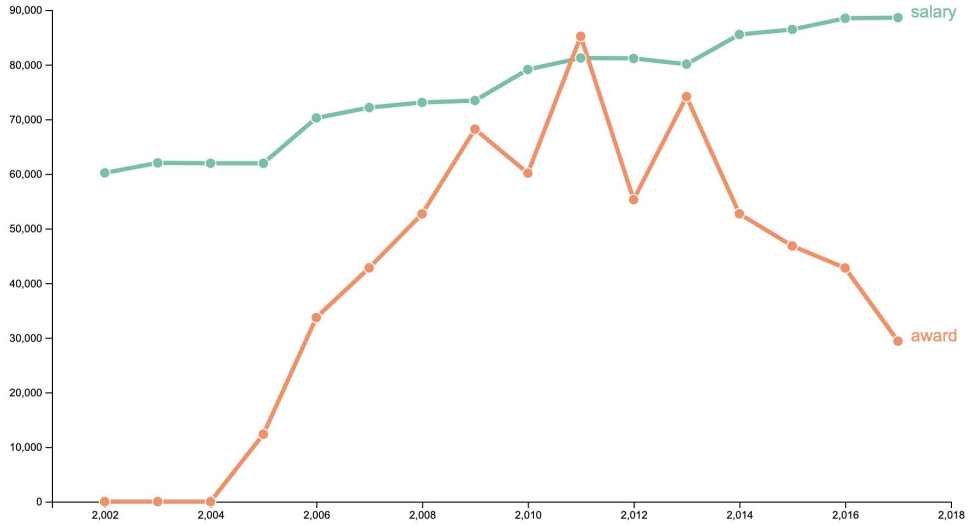
Can we compare misconduct maps according to different time lines by showing the continuous change?

- Barplot of allegation amount over years



Interactive Visualization

More information about cold zone(misconducts happened lower annually/on average)

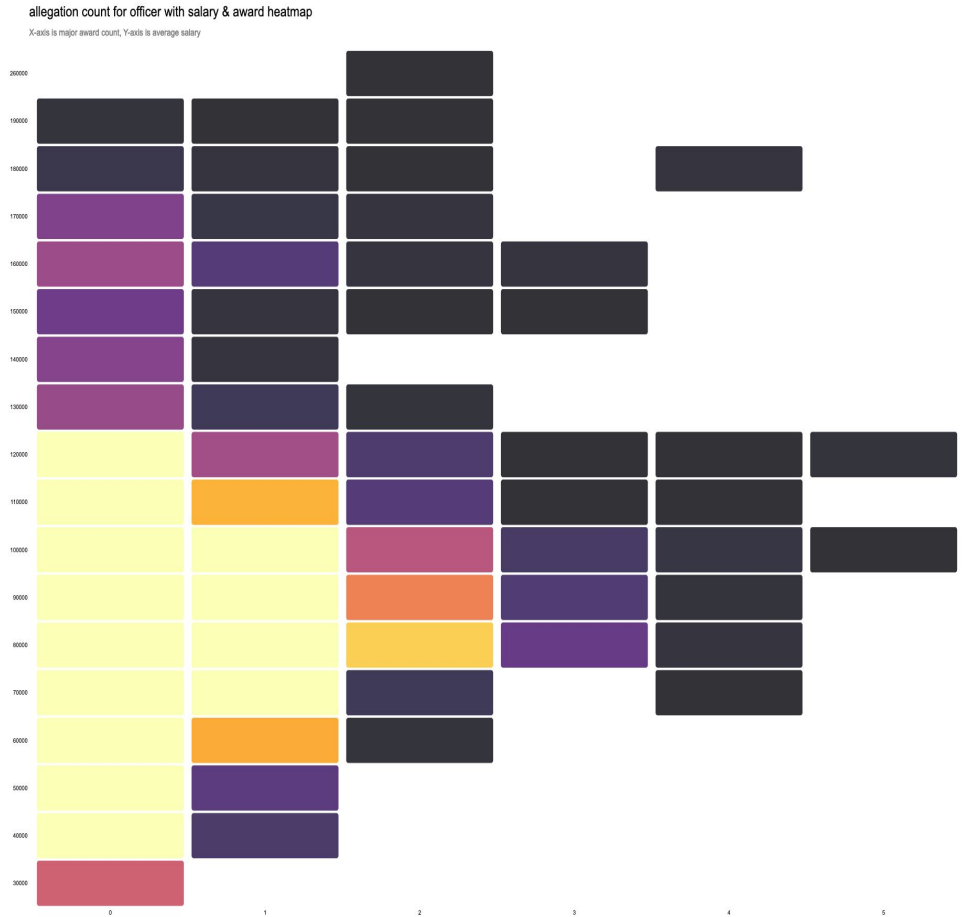


Exact value: 60134

Interactive Visualization

What are the common/different patterns(focus on salary and awards) for those cold zone shares?

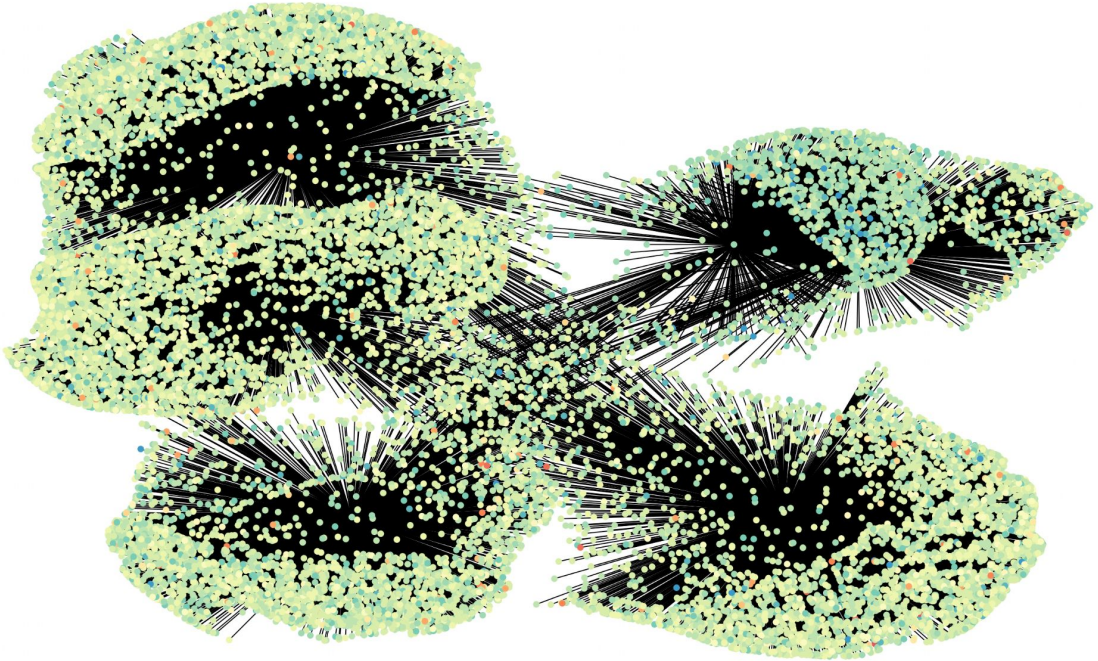
- Heat map of allegation amount with officers' average salary and total award received



Graphic Analytics

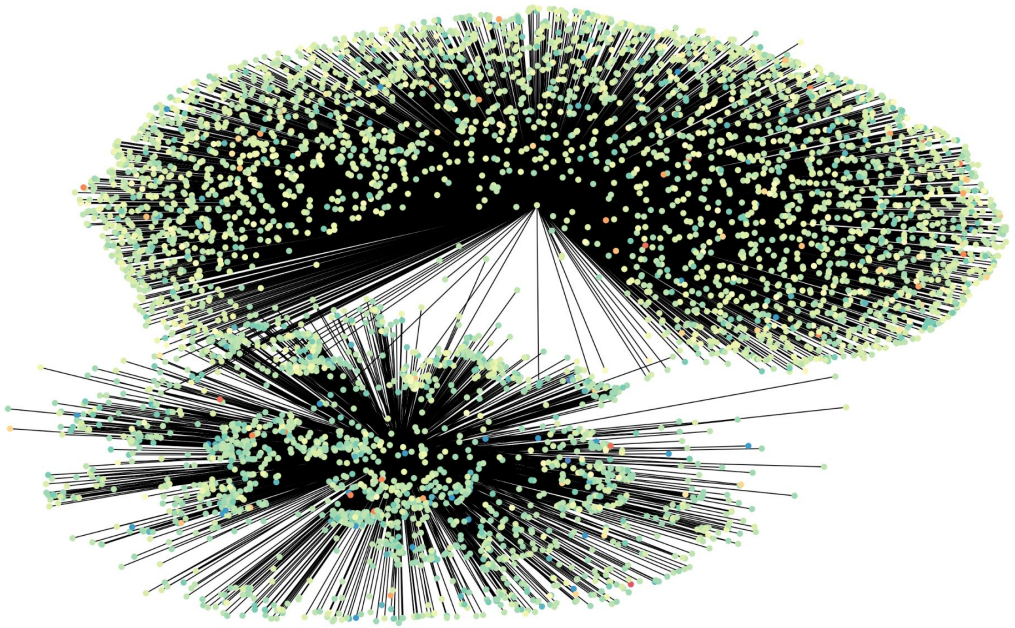
Is there a relationship between the clusters of officers with most complaints and their salary?

- Network
  - Iterations = 10
  - K = 0.1
  - Total of 17991 nodes
  - Total of 17985 edges
  - Darker the color, higher salary
  - Group officers with complaint percentile 0,20,..., 100



Graphic Analytics

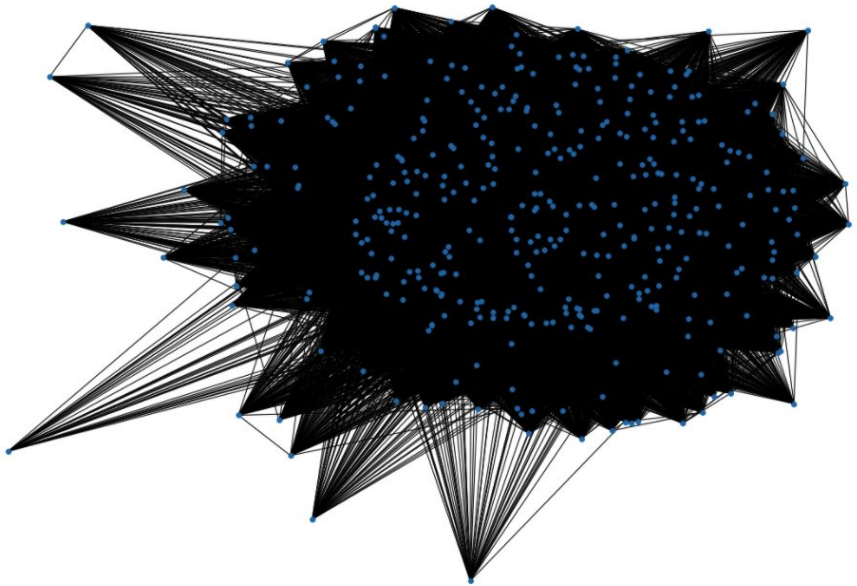
- **Adding more edges**
  - Work environment
  - 17991 nodes
  - 472157 edges
    - Increased from 17985 edges



Graphic Analytics

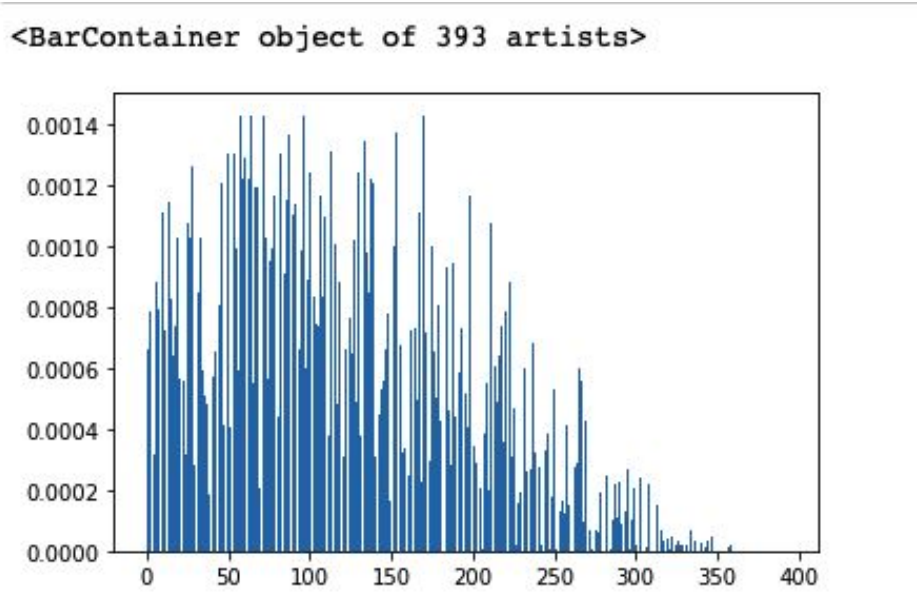
Can we identify clusters of officers who are likely to be co-accused by fact related to their career?

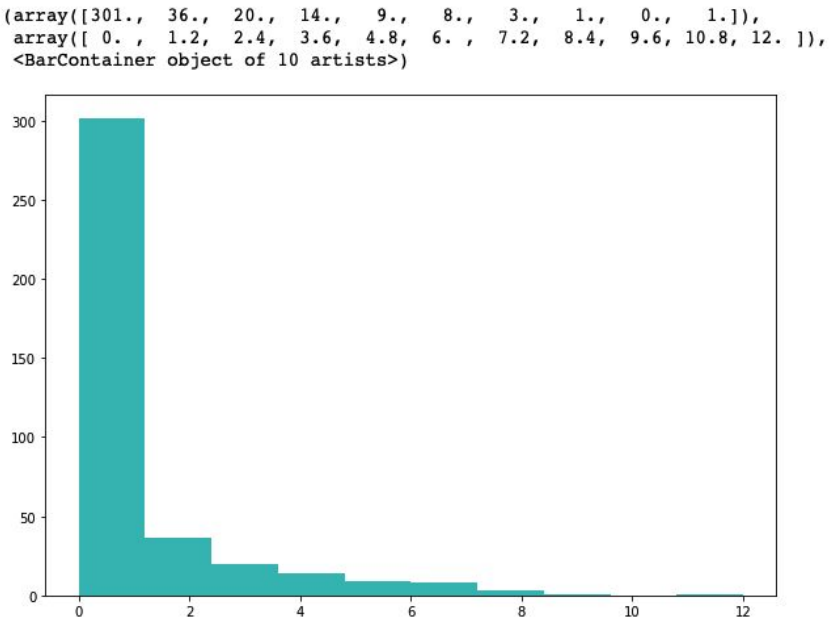
- Salary
  - network of salary data
  - officers who work in the same un
  - 393 nodes, 62343 edges.
  - average degree is 317.2672.



Graphic Analytics

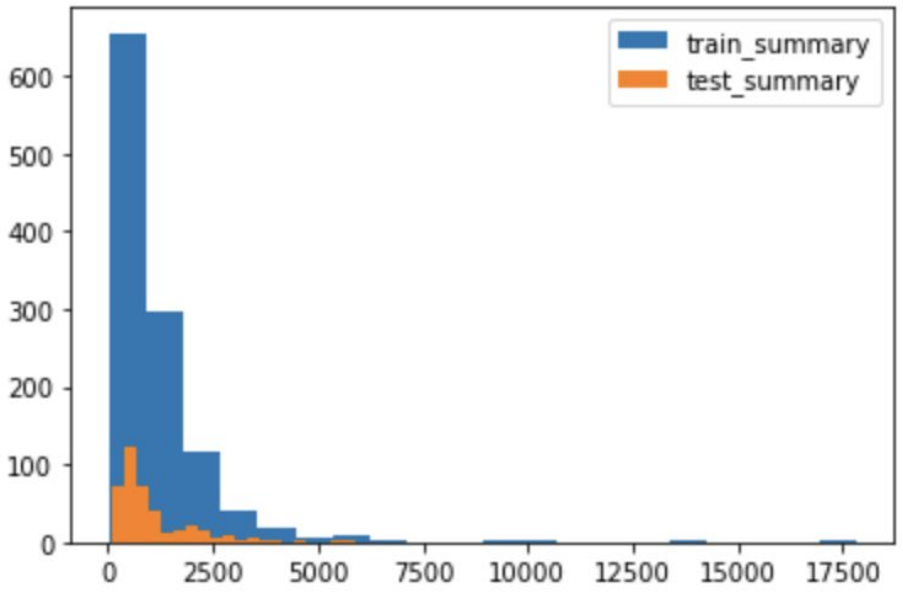
- Check the centrality of betweenness of each node and degree of each node





Natural Language Processing (NLP)

- **Train a Transformer-based Language model to identify the officer complaints from allegation narratives**
  - Text cleaning
    - Normalization
    - Stop words
    - Stemming
    - Lemmatization
  - Splitting dataset
    - 70% for training
    - 30 % for testing



Natural Language Processing (NLP)

- Genism word2vec model
  - Training
  - Compute similarity
  - Extract features from cleaning data
    - Bag-of-Words
    - TF-IDF
- Prediction & Outcome
  - F1 score: 0.695

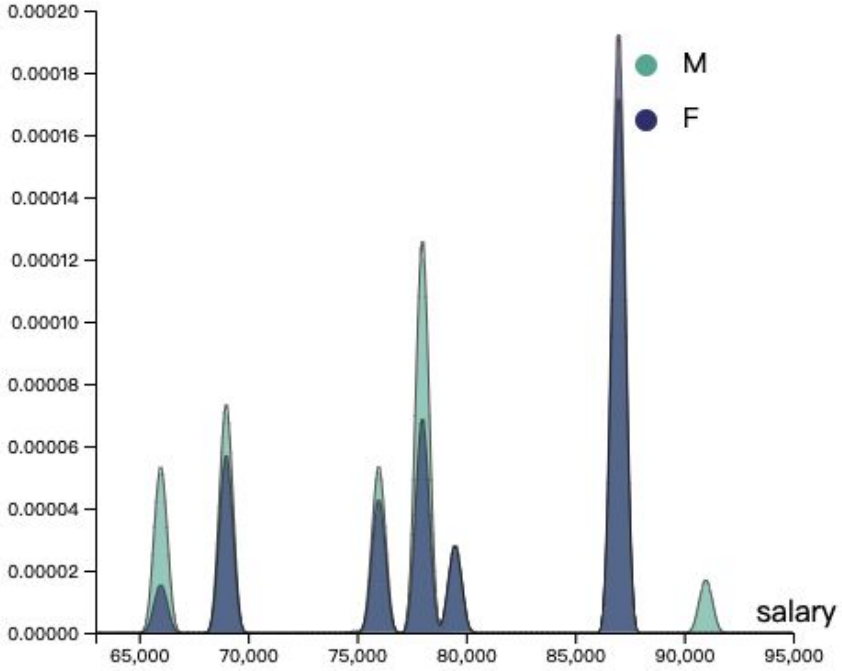
Thanks !

Appendix

Interactive Visualization

What is the distribution of officers' salary across race, gender and other discrepancies?

- **gender**



Interactive Visualization

What is the distribution of officers' salary across race, gender and other discrepancies?

- **Race**