

University of Denver

Final Project: Kronos

COMP 4449: Data Science Capstone

Andy Tyler and Duncan Ferguson
3-7-2023

Contents

Background	2
Data Assessment	2
Preliminary Review of the Data	3
Article Reduction	4
Timeline Development	4
Question 1: Provide a clear analysis of the structure of the POK network, with supporting evidence.	5
Who are the leaders? (Fig 6)	5
Who is part of the extended network? Where are the potential connections between the POK and GASTech? (Fig 6)	5
How has the group structure and organization changed over time? (Fig 6)	7
Question 2: Describe the events of January 20-21 st , 2014. What is the timeline of events?	7
January 20th (Fig 1 & 2)	7
January 21 st (Fig 3)	8
Question 3: Provide at least two possible explanations why the GASTech employees may be missing. What evidence do you have to support each of these explanations?	9
Hypothesis 1: The “Kidnapped” GASTech employees are part of the POK and staged the kidnapping to expose activity of the leadership.	9
Hypothesis 2: The POK did successfully perform the kidnapping, but the ransom was met by the executives.	10
Conclusion	10
Appendix	12
Fig 1). January 20 th AM Timeline	12
Fig 2). January 20 th PM Timeline	13
Fig 3). January 21 st Timeline	13
Fig 4). PoK Cluster	14
Fig 5). Graph Database	14
Fig 6). POK Structure Timeline	15
Fig 7). Tableau Dashboard	15
Fig 8) Article Word Frequencies	16
Fig 9) Timeseries word frequencies example	16

Background

An unknown number of people associated with the organization GASTech disappeared on the 20th of January 2014 during a celebration held by the organization. The goal is to use the various data sources provided to determine the cause of the disappearance and reconstruct the timeline of events that occurred. Two hypotheses are to be constructed and explored what actually happened during the disappearance.

Data Assessment

When first approaching the project, data from both TP-1_Kronos and TP-2_Kronos was considered. The theory was that the two folders of data would be able to complement each other, which they did.

The TP-1 Data folder contained a folder of 845 .txt files with dates ranging from 6/15/2001 to 3/26/2014. These files all contained a short news clip with the new paper media organization, title, date, and a short article about current events. These articles could be difficult to understand at times but provided insights to the disappearance and laid out the timeline of events from 1/19/2014 to 1/21/2014. The two Historical Documents “5 year report clean.docx” and “10 year historical document clean.docx” provided insights into POK and GASTech background information. There were 35 resumes in .docx format that provided information on GASTech employees. A Map of Kronos was provided in JPG format that helped conceptualize the locational distances. An organizational chart of GASTech was also provided in PDF format. In addition to this data two CSV files; “email headers.csv” and “EmployeeRecords.csv” helped display which GASTech employees were communicating with each other as well as their backgrounds. The email headers file contained email senders and receivers as well as subject lines and timeline data which spanned from 1/6/2014 to 1/17/2014.

The TP-2 Data folder contained four csv files; “car_assignments.csv”, “cc_data.csv”, “gps.csv”, and “loyalty_data.csv”. In addition to these CSV files was a map of Kronos and Geospatial data given in KML format that included a background shape for Kronos and street locations. The car assignments were paired with the GPS data to identify cars and locations. The credit card and loyalty data also helped supplement locational data.

Preliminary Review of the Data

The best plan of action was set to divide the two different data folder and conduct preliminary research. TP-1 data was first used to create a nodal network of email data and cluster groups of employees. This set of data contained mostly communication channels and articles. A graph database was also constructed to make exploration of the email data more efficient and to generally identify the clusters of individuals and their communication patterns within the GasTech email server.

TP-2 was used to create locational data of the different GASTech employees. To do this multiple different data sources were joined together. The files included bits and pieces of the different employee information and timeseries data paired from GPS, credit card and loyalty data (_8_Tableau_GIS.ipynb). The original goal was to take the clusters from the TP-1 data and pair them together with the locational data of TP-2. A list of victims and suspects were discovered from in ‘_4f_final_9_suspects.ipynb’ and incorporated into a tableau workbook (“Kronos_Tableau_Workbook.twbx”) for exploration. It was during this exploration that 7 suspects (Hennie Osvaldo, Hideki Corinaro, Isia Vann, Kanon Herrero, Loreto Bodrogi, Minke Mies, Varja Lago) can all be seen gathering within the proximity of Katerina’s Café and Guy’s Gyros all on the 18th hour of January 19th (Fig 7). This help confirm suspicions that they were all involved together in something but did not prove that they were the kidnappers.

Article Reduction

After completing the preliminary research, the choice was to either dig further into patterns of locations or solve a mystery. The geospatial data only went up to January 19th whereas the TP-1 data covered all points in time. It was then decided to approach TP-1: “The Kronos Incident the Kidnapping.” The data that was compiled in this project played more to the groups strengths and offered a better chance at understanding what happened during the kidnapping. The email data was used to create a graph network that showed which GASTech employees were communicating with each other (Fig 4 & 5). This information confirmed findings in articles and using geospatial data that the actors within Protectors of Kronos were sharing emails related to their involvement with the organization. This helped zero in on the key players and their connections for additional text analytics.

The next approach was to construct a Jupyter notebook that was aimed at looking at the sentiments of the articles and scrapping the articles time and dates to identify which articles were pertinent to constructing a timeline of the events (“_5.2f_sentiment_rip.ipynb”, “_7_Article_Dates_and_more.ipynb”). The articles were boiled down from 845 to 105 for the timeline events of January 19th – 21st and then a word frequency was constructed to help illustrate important events that happened leading up to the kidnapping (Fig 9). While these were interesting, they did not provide as much insight as hoped. It was decided to just read boiled down list of articles to help develop a moving timeline (Fig 1,2,3).

Timeline Development

By bringing down the number of articles it became much easier to sift through and line up the different event timelines. These timelines help provide useful into what was occurring during the kidnapping but left many questions as to what was happening. This led to a stall in progress that

eventually led to the shift of researching the questions below. To start this process, reading the historical documents helped create founding POK Structure timeline. The names used from the POK founders was then used to sift through the other articles to identify other articles that help display the change of the POK over time. This provided immensely useful in connecting GASTech employees to POK members (“_1.f_kidnapping_rip.ipynb”). The answer was in the last names of GASTech employees leading to the Last names of POK Members. The more articles that we sifted through the more evidence was mounted that there was a strong connection between the two.

Question 1: Provide a clear analysis of the structure of the POK network, with supporting evidence.

Who are the leaders? (Fig 6)

There were seven founding members of the POK: Henk Bodrogi, Joreto Katel, Carmine Osvaldo, Ale L. hanne, Jeroen Karel, Valentine Mies, and Yanick Cato. Henk Bodrogi started as the president from 1997-2002. The leadership was then passed to Elian Karel in 2002. Karel then passed away in jail (while being held for tax fraud) shortly before being released. After this death Silvia Marek then became the leader of the POK. Silvia remained the leader of the POK all the way through the kidnapping.

Who is part of the extended network? Where are the potential connections between the POK and GASTech? (Fig 6)

There are quite a few notable characters that were part of the extended network. It should first be noted that Henk Bodrogi wife passed away due to illness in 1997 likely associated with the byproducts of GasTech’s environmental negligence. Juliana Vann also passed away from illness and was used a poster child for POK’s movement into trying to fight GASTech pollution. Following this death, the POK began to change structure and tactics. Lucio Jakab was an SOW member that was recruited to the

POK by Elian Karel. Michale Kraft was the POK Attorney, Mestache Vardshvna was a POK protestor and Lorenz Di Stefano wrote pro POK articles in a blog post.

The main break in the investigation came when comparing the founding POK members chart to the graph database of emails. When looking at the two lists together it became evident that a group GASTech employees that were clustered together shared matching last names with the POK founders (Fig 4 & 6). Furthermore, these employees seemed to be sharing email traffic with subject matter consistent with involvement in the PoK organization. Matching the email traffic to the last names of the PoK founders offered abundant evidence that the dual GasTech/PoK members were likely related to the founders and part of the overall scheme taking place on January 20th.

Carmine Bodrogi, the daughter of Henk and Mira Bodrogi was arrested for drugs. The POK's involvement with drugs was commonly sighted and could help point to the fact that she was a POK member. She definitely had a motive with her mother passing away due to GASTech pollution. Loreto Bodrogi, a GASTech employee, is also believed to be a relative of Henk Bodrogi. Further developing the theory of founders and children employed by GASTech is Hennie Osvaldo who is related to the founder Carmine Osvaldo and can be assumed to be a POK member. Rucella Mies Haber was a GASTech employee and shares a maiden name with Valentine Miles a founder of POK. A weaker linkage, but one that plays to suspicions. Minke Mies can be assumed to be a relative of Valentine and Rucella and was also a GAS Tech employee. Lastly Isa Vann (GASTech) was the brother of Isa Vann (who died from supposed GASTech pollution). Mandor Vann was also related to Isa Vann and had military ties of which overlapped with many of the GASTech employees. The pieces fit themselves together that this group had ties between both GASTech and the POK.

How has the group structure and organization changed over time? (Fig 6)

The POK started off with seven peaceful founders whose intent was to hold GASTech accountable for pollution. They started by trying to lobby the government, and even convinced Cesare Nespola the Minister of Health to introduce legislation that would hold GASTech responsible. Unfortunately, Nespola mysteriously died on 6/15/2001 and the President of Krono's (Luis Araullo) nephew Vincent Kapelou was appointed into the position. It is also suspected that at this point, the GasTech leadership began strengthening ties to the Abilan local government.

This marked the start of a structural change in the POK. Henk turned over the reins to Elain Karel. Many other founders turned over their leadership positions. Karel then proceeded to switch the POKs focus into government corruption and accountability in addition to fighting GASTech. By 2009 Karel had grown the POK to 200-300 people. POK also became more associated with drugs and violence within the article word sentiment as the timeline approached the kidnapping.

Question 2: Describe the events of January 20-21st, 2014. What is the timeline of events?

After reviewing the relevant articles and reducing them down from 845 to 105, the decision was made to reapply NLTK to the articles and construct the timeline of the events that happened on 20th & 21st. A word frequency by date and time helped illustrate which times were most important to look into (Fig 9, '_7_Article_Dates_and_more.ipynb') such as 10am on Jan 20th.

January 20th (Fig 1 & 2)

The first articles that provided an insight into the timeline were articles 348 and 834. These articles stated that there was a GASTech IPO celebration that was scheduled for 10am with CEO Sten Sanjorge, Ingrid Bassaranco, Ada Campo-Corrente, Orhan Strum, Willem Vasco-Paris to be in attendance at the reception held after words. Right at 10am a fire alarm went off at Gastech put people were not

noted to be exiting the building other than a helicopter that left the roof (ART 763). Shortly after the fire alarm went off the fire department arrived and then bomb squad arrived for inspection. At 10:52 am the CEO Sanjorge arrived at the capitol building but did not reappear. It was also reported that there was no sign of the other GASTech employees. Captain Federigo Accardi gave all is clear sometime around 11am. After all clear was given the Police arrived at GAS Tech around noon. The police blocked off GASTech and started to question the employees.

At 2pm the news articles 522 and 828 reported that there were GASTech employees that were missing. Shortly after the report of missing employee's Government officials arrived at gas tech (ART 811, 521, 719). At 2:30pm a Privat jet with 8 people left for Rome Italy (ART 427, 56, 772). The police and GASTech employees started to leave the scene around 3pm (ART 21, 320, 426, 484, 486). The remaining police started to ask questions about the breakfast crew that was at GASTech earlier that morning. Not many employees seemed to know about other suspects other than these individuals (ART 434, 43, 417).

Later at around 5pm the media started to report that the employees had been kidnapped. The media then speculates about the breakfast crew being the kidnappers at around 6pm and reports that one of the two jets that left during the day came back. Evard Vann was set free at 7pm after 6 hours of interrogation (ART 412). He was held due to his last name 'Vann' and its association with the POK. Ironically, they did not question Isa Vann who had been arrested multiple times following his sister's death. During this time the police reported that there were approximately 14 people that were unaccounted for (ART 567, 485, 262). At 8pm a Rogue GASTech employee went against company policy and stated that the employees had been kidnapped (ART 172, 386, 384). At 10pm a jet from Tethys arrived where the occupants went straight to the police (ART 118, 744).

January 21st (Fig 3)

At 7am news articles started to report that 14 missing GASTech employees were missing and that there was a GASTech plane that arrived late in the night (142, 738). This report was then modified to 10 people that were missing around 8am and 9am (ART 418, 583). At 10 am CEO Sten Sanjorge was reported not missing and an occupant on the flight that returned from Tethys (ART 566, 624). At 11am news correspondent Petrus was contacted by the POK to deliver a message (ART 110, 793). At noon the POK Claimed Responsibility for the kidnapping and demanded a ransom of \$20 million (ART 407, 708).

Question 3: Provide at least two possible explanations why the GASTech employees may be missing. What evidence do you have to support each of these explanations?

Hypothesis 1: The “Kidnapped” GASTech employees are part of the POK and staged the kidnapping to expose activity of the leadership.

The POK wants to embarrass and expose some activities that the GASTech executives were engaged in. There are many articles that reference GASTech pollution early on before 1997. Multiple deaths occurred due to these deaths such as Mire Bodrogi and Juliana Vann. POK leader Elian Karel even suffered a strange death before being released from jail during a legislation push to hold GASTech accountable. As the POK grew in age there were more and more articles associated with violence and outbreaks, especially during the anniversary of Elian Karels death on June 20th (`'_7_Article_Dates_and_more.ipynb'`).

In addition to the POK motive there are multiple articles that state that the executives were missing during the celebration and subsequent bomb threats but were later accounted for when they turned up at the police station the night of January 20th. The POK claimed responsibility for the kidnapping, but the executives were never actually kidnapped. If the POK had insider knowledge that the executives would not be attending the celebration and were engaged in activity that could be

damaging to their reputation, GASTech, or associated parties, then claiming responsibility for the kidnapping would then expose their absence and limit their pool of alibis related to their whereabouts.

While the discovery of an exact list of missing persons (other than suspected executives) was never found, the tie between POK founders and GASTech employees was discovered. These GASTech employees were all spotted together the night before the kidnapping (Fig 9) and had a high frequency of communication amongst themselves. Furthermore, their utilization of GasTech's internal email system and the specific language of the communications – FW: ARISE- Protectors of Kronos – suggests that they had some action planned for the 20th that elicited an inspiring email to prepare the PoK actors for their actions on the day of the celebration.

Hypothesis 2: The POK did successfully perform the kidnapping, but the ransom was met by the executives.

In multiple public letters the POK claims responsibility for the disappearance of the executives and demands a \$20m ransom for their release. It is possible that using the IPO money recently secured by GASTech, the executives paid the ransom themselves and were released by the kidnappers. If the ransom was paid, then the executives may avoid some scrutiny for their activities and the goal of inflicting financial damage to GASTech, as well as securing significant funding for future endeavors, has been met by the POK. This might explain why the executives were unaccounted for during the original news reports but were then found.

Conclusion

In conclusion, hypothesis 1 is most likely. There was a cluster of GASTech Employees that were all contacting each other via email before the "Kidnapping" happened. These individuals all bear last names that are related to the founding members of the POK. To further this there are no actual reports of "Kidnappers" there are only speculations that it was the breakfast crew, but none of the other

GASTech employees report that anything happened other than confusion about why the fire alarm went off. It was only when a rogue GASTech employee reported to the news that there was a kidnapping that kidnapping became plausible. Until then the police were only reporting that there were people that were unaccounted. The number of people missing happened to fit the one node of GASTech employees that had POK connections. It is this strong correlation with lack of other significant evidence that suggests that hypothesis one is correct.

Lots of time was spent looking into both Kronos Project 1 & 2 during the beginning of the project. This led to a dispersion of resources into analysis. While it was convenient to find the suspects in the same location the day before the kidnapping, the amount of work that was conducted to find this did not outweigh the efforts. If the project was to be continued, the patterns and frequency of the group meeting together would be further explored. It was a bit disappointment that the GPS data did not go through the kidnapping timeline. When full resources were directed toward Project 1 the answer popped out after solving the prescribed questions. A deeper dive into name associations in the articles would be nice to help further prove hypothesis 1, but it was felt that enough evidence was gathered to construct a plausible story as to what happened on Jan 20th. The relatives/children of the POK founders infiltrated GASTech as employees and staged the disappearance to highlight that the GasTech executives were not present during the event and to raise concerns regarding their whereabouts during that timeline.

Fig 2). January 20th PM Timeline

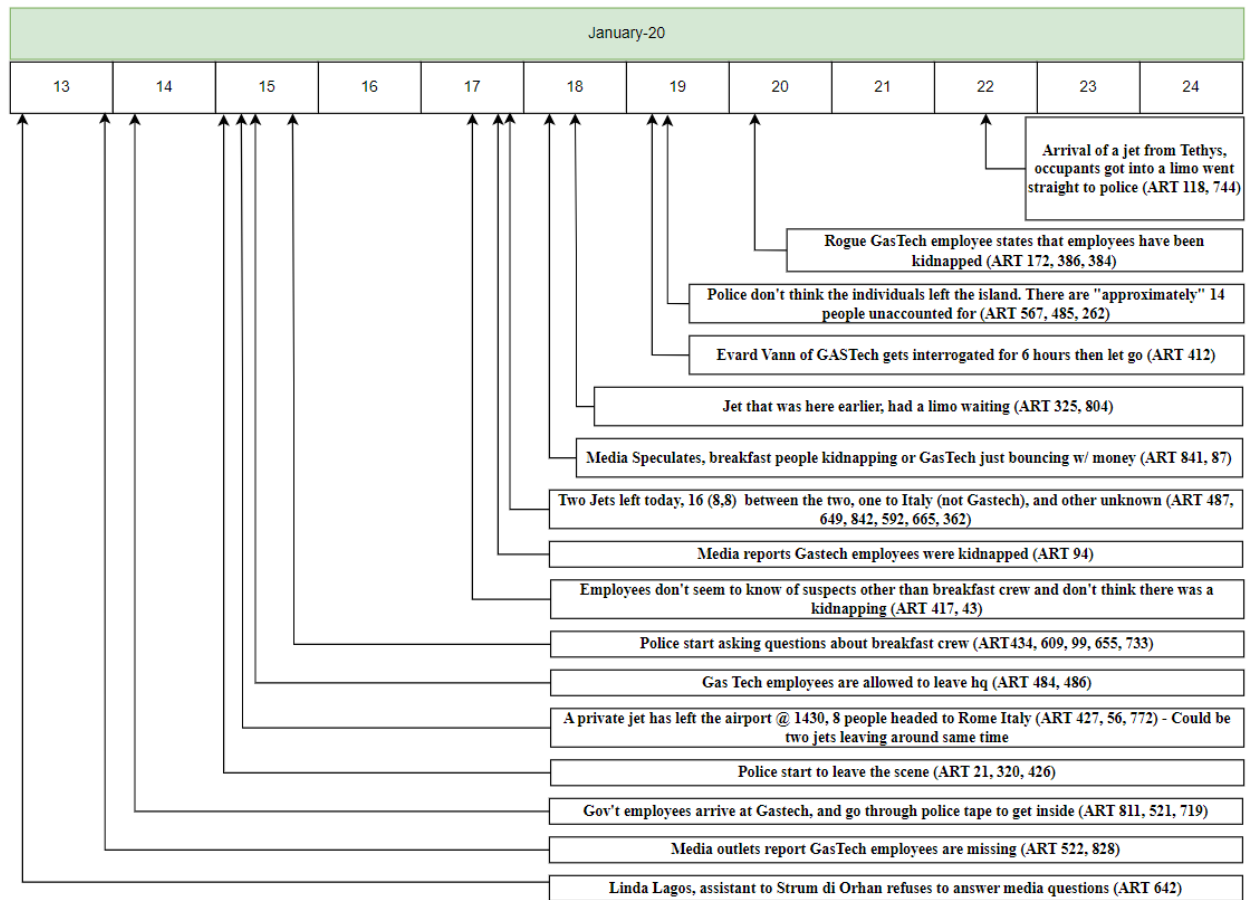


Fig 3). January 21st Timeline

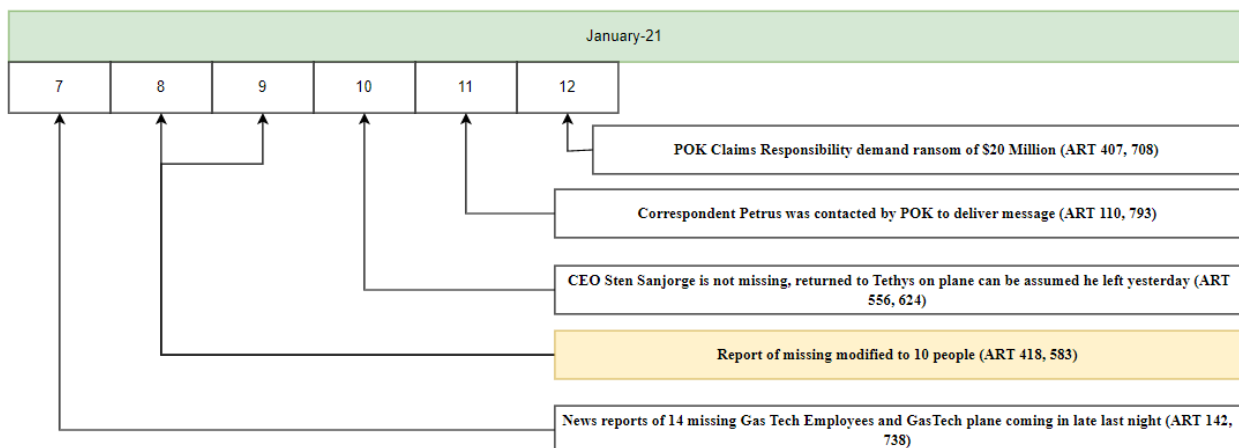


Fig 4). PoK Cluster

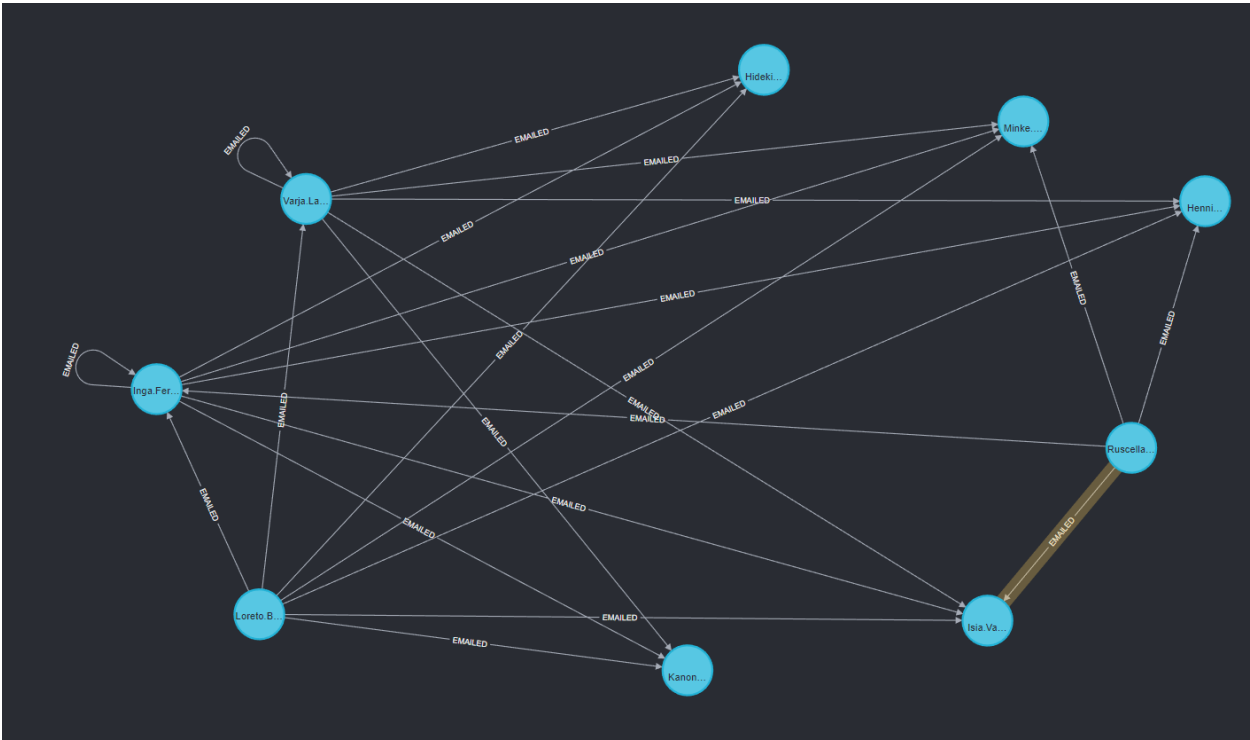


Fig 5). Graph Database

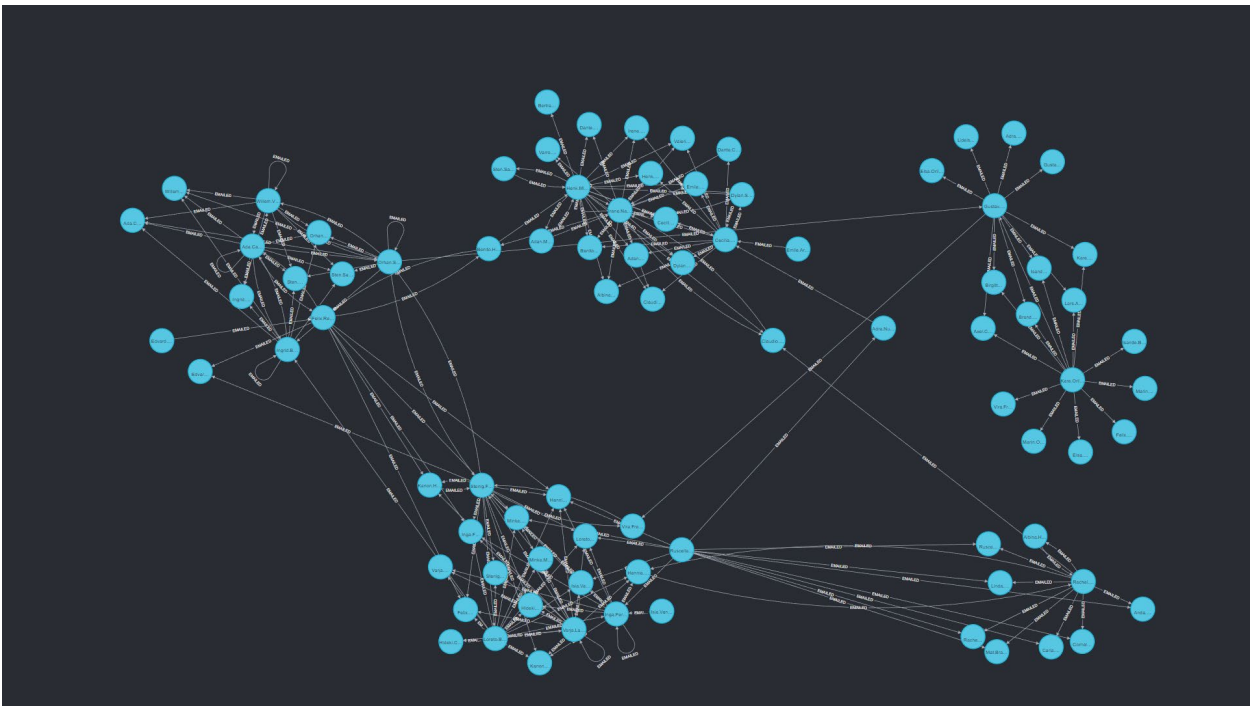


Fig 6). POK Structure Timeline

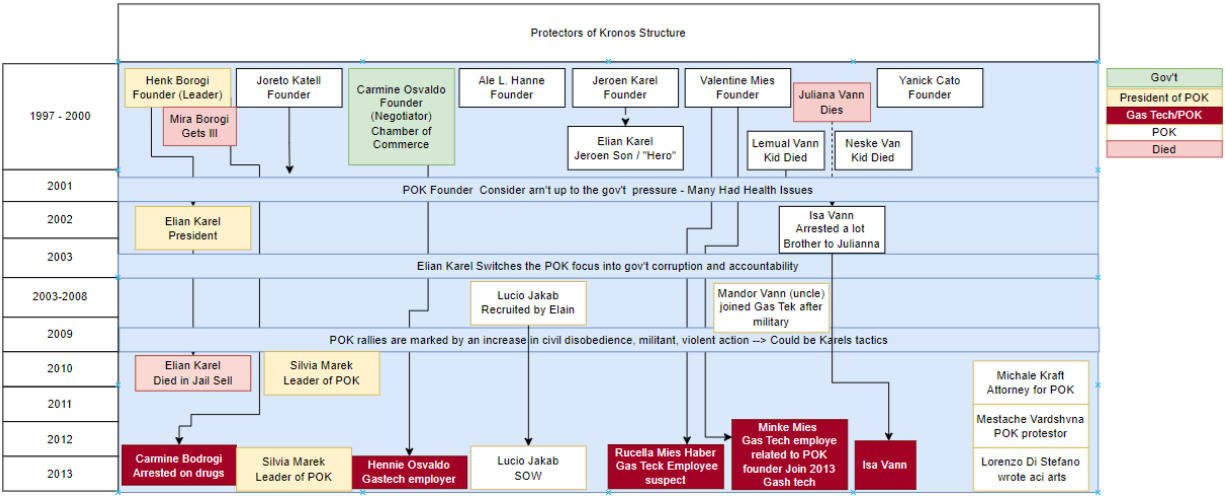


Fig 7). Tableau Dashboard

https://public.tableau.com/views/Kronos_Tableau_Workbook/MapDash?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

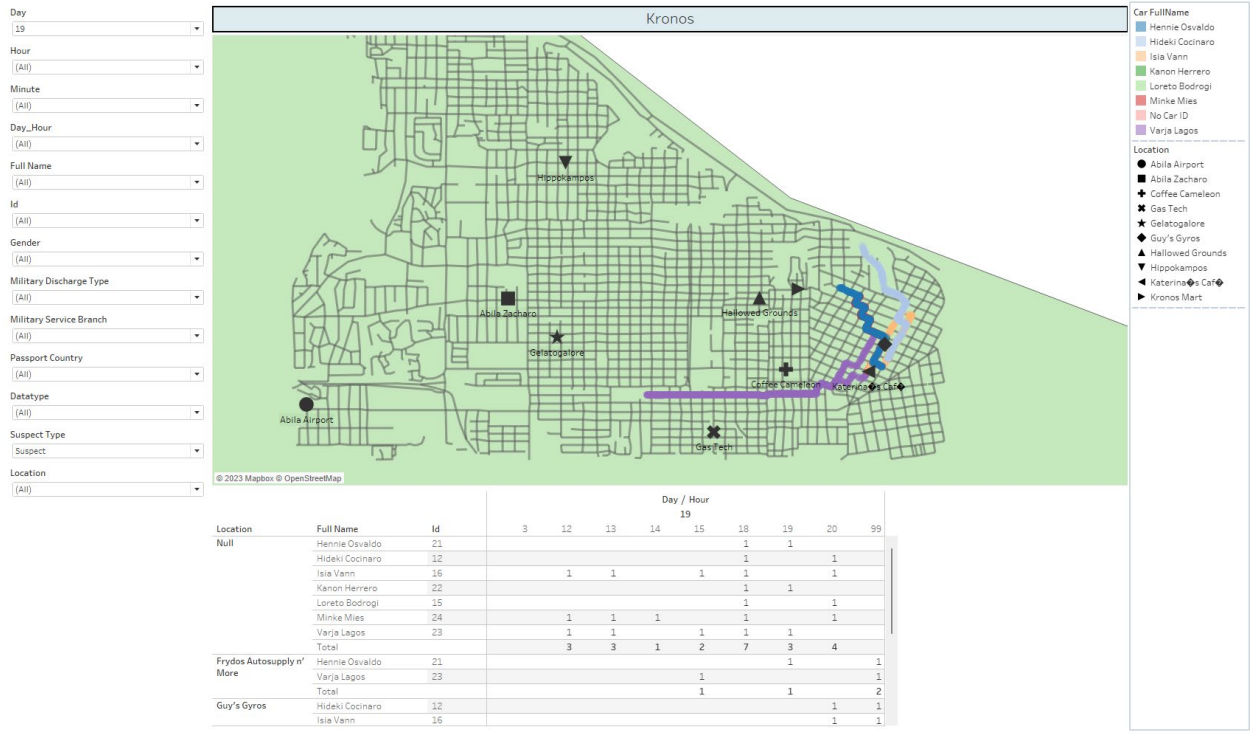


Fig 8) Article Word Frequencies

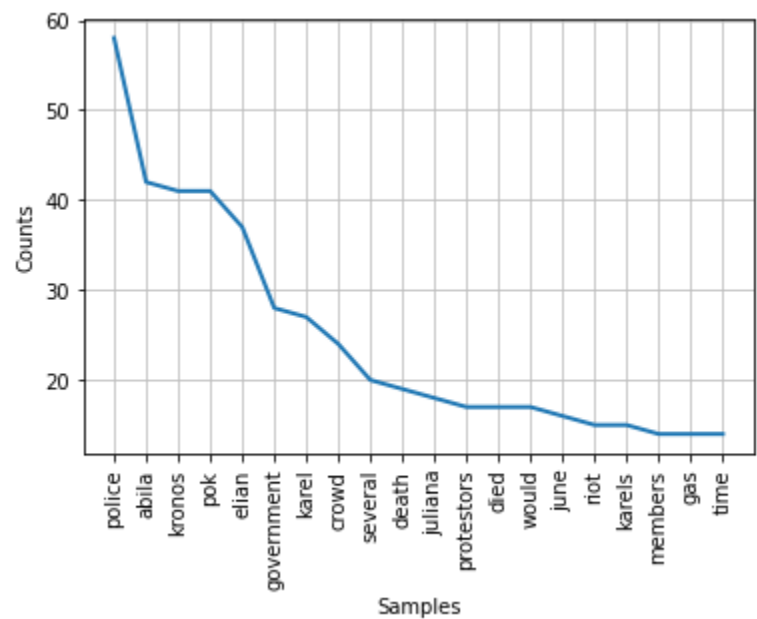


Fig 9) Timeseries word frequencies example

