VAST Challenge 2018 Reviewer Guide: Mini-Challenge 3

This document provides information to support peer review of submissions to VAST Challenge 2018, Mini-Challenge 3. It covers background about the submission structure, the challenge problem, tasks and questions presented to participants, potential answers, and evidence found in the Challenge data supporting these answers. For a full description of the challenge problems and to access the data provided to the participants, please visit http://vacommunity.org/VAST+Challenge+2018.

Submissions

Participants are required to submit their entries on a standard answer form, along with a video explaining how visual analytics were used to help solve the challenges. Please consider both parts of the submission in your review. If you have difficulty reading the answer form or playing the video, please contact us at vast-challenge@ieee.org for assistance.

Scenario

We have oriented the VAST Challenge 2018 around an invented environmental scenario. This year's scenario extends the one we created in 2017, and if you would like additional information about that work, please see the 2017 Challenge material at http://www.vacommunity.org/VAST+Challenge+2017 . The 2018 participants did not have to participate in the 2017 Challenge, nor would they have needed to extensively review that material to submit this year. The storyline may make a little more sense if some of the 2017 narrative was reviewed, however.

Overview

Mistford is a mid-size city located to the southwest of the Boonsong Lekagul Wildlife Preserve. The city has a small industrial area with four light-manufacturing endeavors. Mistford and the wildlife preserve are struggling with the possible endangerment of the Rose-Crested Blue Pipit, a locally loved bird. The bird's nesting pairs seem to have decreased alarmingly, prompting an investigation last year implicating Kasios Office Furniture, a Mistford manufacturing firm. Kasios presents itself as an extremely eco-friendly organization. But they and their international parent company may be hiding something. It's time to apply your visual analytics expertise to help illuminate the path to good science.

Mini-Challenge 3, "Big graph challenge"

The Kasios Furniture Company is affecting the environment near Mistford. But this is only a small part of the global Kasios footprint. Grad student Mitch Vogel has travelled to northern Europe and discovered telltale signs of Methylosmoline damage at a nearby aviary sanctuary. He has also learned that the parent company of Mistford's malfeasant furniture company is Kasios International, which is a huge multinational conglomerate with factories all over the world. Mitch is immediately suspicious that the local subsidiary—EuroKasios—is contaminating the aviary sanctuary and endangering the Greater Eurasian Red-Throated Pipit. Mitch reaches out to the company so he can confront them about their

chemical use, but he gets no response. Then Mitch receives an anonymous letter from someone who's willing to help. A fellow pipit lover who works at Kasios International has gathered up a variety of company data and identified a suspicious group within the company. Attached to the letter Mitch receives is a USB drive with phone, email, meeting, and procurement records for Kasios International over the past 2 1/2 years. Mitch wonders if the fate of the Eurasian Pipit lies somewhere in that data. Contestants were asked to put this data together and use visual analytics to see if the problems with Kasios are much larger than initially suspected.

Ground Truth

Data

Challenge participants were provided with the data that the Kasios Insider passed to Mitch Vogel. There are call records, emails, purchases, and meetings. The data only includes the source of each transaction, the recipient (destination), and the time of the transaction. Contents of emails or phone calls are not available. There is also a company index that shows the name of everyone in the company and their associated ID#. There are 642,631 individuals in the index. There are four data files (one for each type of transaction) that cover the whole company. And four more data files that have the transactions that the insider has indicated are suspicious. For the stretch assignment (question 4) there is a file that contains a list of 4 individuals who made 7 suspicious purchases.

Graph metadata:

Graph 1: Calls – 534,000 nodes and 10.6 million edges

Graph 2: Emails – 526,000 nodes and 14.6 million edges

Graph 3: Purchases – 227,000 nodes and 762,000 edges

Graph 4: Meetings – 101,000 nodes and 127,000 edges

Narrative

A leader inside Kasios International has learned from some angry customers that AGOC-3A is not as effective as Methylosmoline. He's been trying to do the right thing—development and marketing of AGOC-3A was one of his major professional accomplishments—but now he's feeling pressure to restart production of Methylosmoline. He calls a meeting with a few other key people in the company and asks them to secretly buy Methylosmoline. A larger group forms around this task. They communicate, meet and ultimately purchase some Methyl precursors after the leader gives his final approval.

There are several key events that participants may identify:

The signal starts on 11/3/2017 at 9:34:53.

The sale of the banned chemical occurs on 9/21/2019 at 3:38:53

The signal ends on 11/23/2019 at 4:38:53

The group of suspects has an overall leader who is the instigator and three core groups (finance/logistics, builders, and operations).

The overall leader of the suspect group (Richard Fox, ID:857138) calls three key people in the builders and finance groups to motivate them. Meryl Pastuch (ID:1690582) is the leader of the builders group. Lindsy Henion (ID: 1108217) is the most connected person in the suspect group and is a trusted confidante of the leader/instigator. Ferne Hards (ID: 944354) is also close to the leader.

Richard Fox meets with Sherrell Biebel, Julie Tierno, and Madeline Nindorf (of the finance and operations groups).

The leaders of the groups meet with key team members: Richard Fox met with Madeline Nindorf. Meryl Pastuch calls Sheilah Stachniw, Rosalia Larroque, and Lindsy Henion and they have several follow up calls. And Tobi Gatlin emails Ferne Hards and Sherrell Biebel. Richard, Meryl, and Tobi lead the operations, builders, and finance groups, respectively.

Richard Fox meets with a builder and finance (Sherrell Biebel and Meryl Pastuch) to approve the chemical purchase on 8/30/2019.

Then Meryl meets with Rosalia Larroque. Subsequently, there are several emails and calls between Rosalia Larroque and Kerstin Belveal and before and after Rosalia purchases the chemicals from Jenice Savaria 9/21/19.

The CEO is connected to the seller of the chemical through this chain:

Seller (Jenice Savaria) → buyer (Rosalia Larroque) → trusted confidante of the leader (Meryl Pastuch) → leader (Richard Fox)

This table contains all of the people within Kasios who are involved in the plot to purchase banned chemicals:

Name	ID	Role/group	Leadership and other roles
Alex Hall	505444	finance/logistics	Drops out of group
Lizbeth Jindra	623027	builders	Drops out of group
Madeline Nindorf	175354	builders	
Patrick Lane	1023940	operational	
Richard Fox	857138	operational	Overall leader, operations leader
Sara Ballard	1663285	operational	
Martha Harris	2037156	operational	
May Burton	387676	operational	
Glen Grant	1642962	operational	Drops out of group
Dylan Ballard	1039322	builders	Drops out of group
Meryl Pastuch	1690582	builders	Builders subgroup leader, trusted confidante of leader, Final coordination of purchase
Melita Scarpaci	51780	builders	Drops out of group
Augusta Sharp	387200	builders	
Kerstin Belveal	728286	builders	Instructs buyer
Sheilah Stachniw	786361	builders	Drops out of group later
Rosalia Larroque	1847246	builders	Buyer
Lindsy Henion	1108217	builders	High degree (most connected), trusted confidante of leader
Julie Tierno	1886447	finance/logistics	

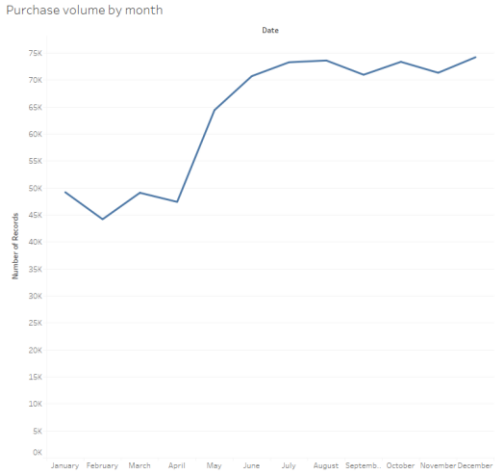
Ferne Hards	944354	finance/logistics	Trusted confidante of leader
Sherrell Biebel	981554	finance/logistics	Final coordination of purchase
Jose Ringwald	696047	finance/logistics	
Ramiro Gualt	234164	finance/logistics	Drops out of group later
Tobi Gatlin	969089	finance/logistics	Finance subgroup leader
Refugio Orrantia	28520	finance/logistics	Drops out of group
Jenice Savaria	2038003		Seller

People listed in black were given to the participants as suspects. People who are highlighted in Red were not provided to participants as part of the signal. They only appear in the big graph and can be found through their connections to the larger group.

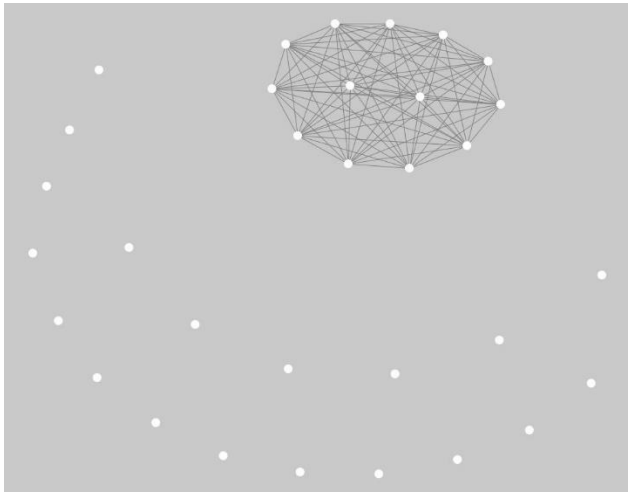
Contestant Questions

1. *Using the four large Kasios International data sets, combine the different sources to create a single picture of the company. Characterize changes in the company over time. According to the company's communications and purchase habits, is the company growing? Limit your responses to 5 images and 500 words.*

   Purchasing data shows that the company is growing

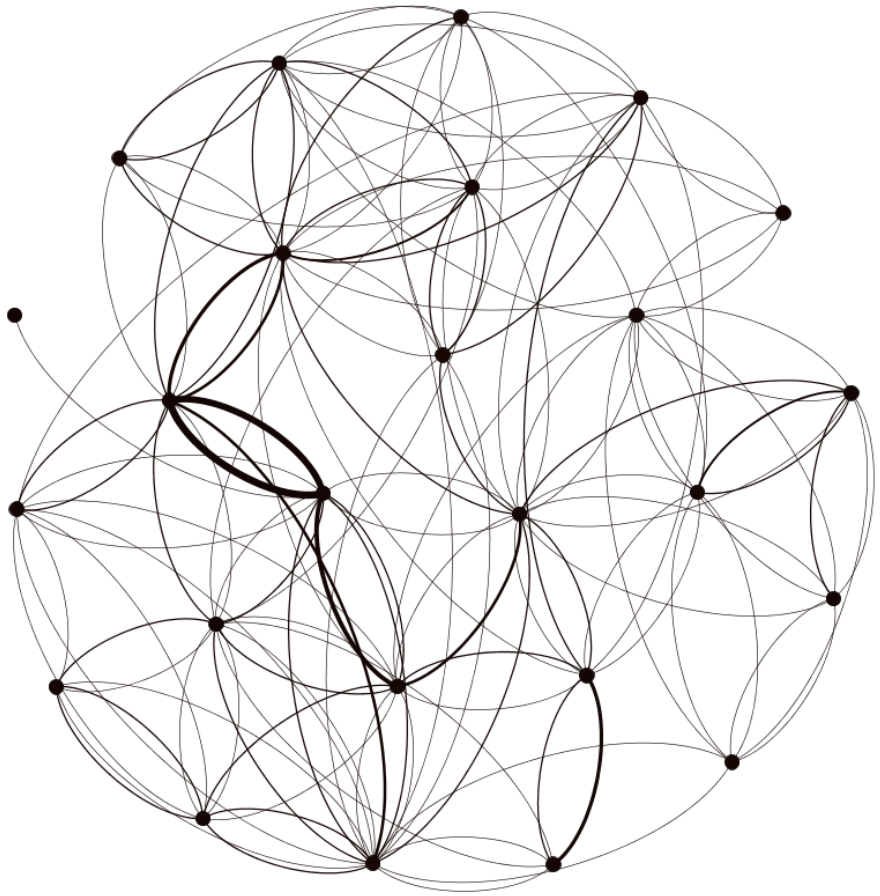


   Looking only at one month of the data one can cluster the company into a smaller set of groups. These could be the subsidiaries in different locations. There are also a number of independent and unconnected groups. These could be external people or people who have left the company.

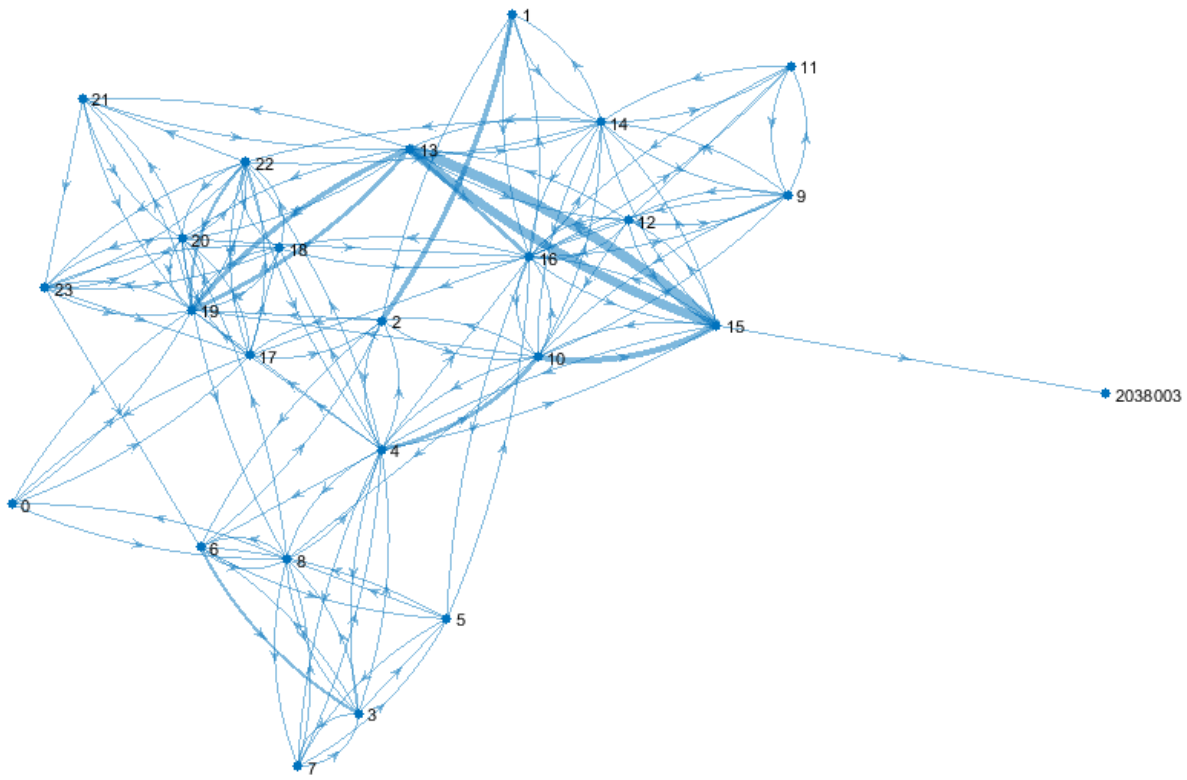


2. *Combine the four data sources for group that the insider has identified as being suspicious and locate the group in the larger dataset. Determine if anyone else appears to be closely associated with this group. Highlight which employees are making suspicious purchases, according to the insider's data. Limit your responses to 8 images and 500 words.*

Here are several ways to represent the group of suspects with graphs that show connections between the group:



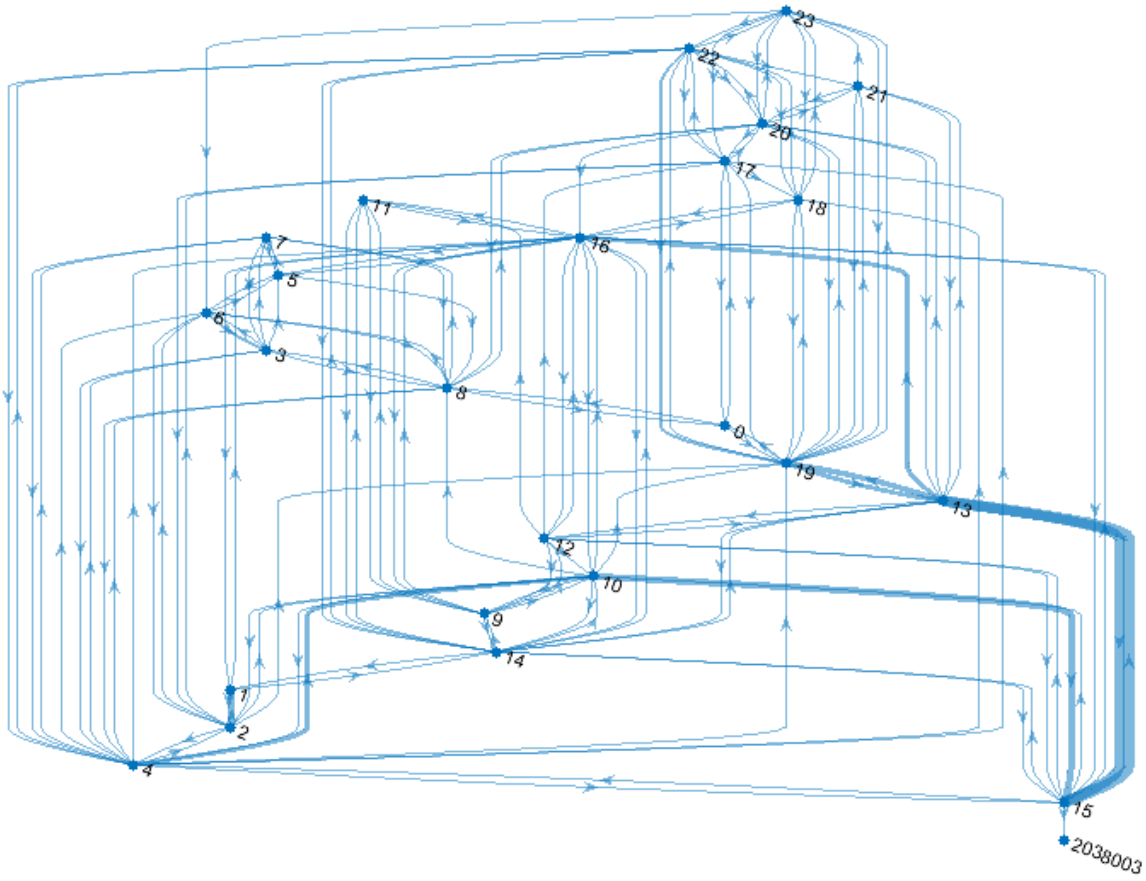
The seller is the single point at the left. It's hard to tell the groups apart with this image.



Now the seller is off to the right. The structure of the group is a little clearer. It's easy to see who communicates the most because the lines connecting them are thicker. Note that the numbering for this graph lists the participants in the plot in increasing numerical order from 0 (with the exception of the seller). The table below maps the ID provided to the participants with the short numbering

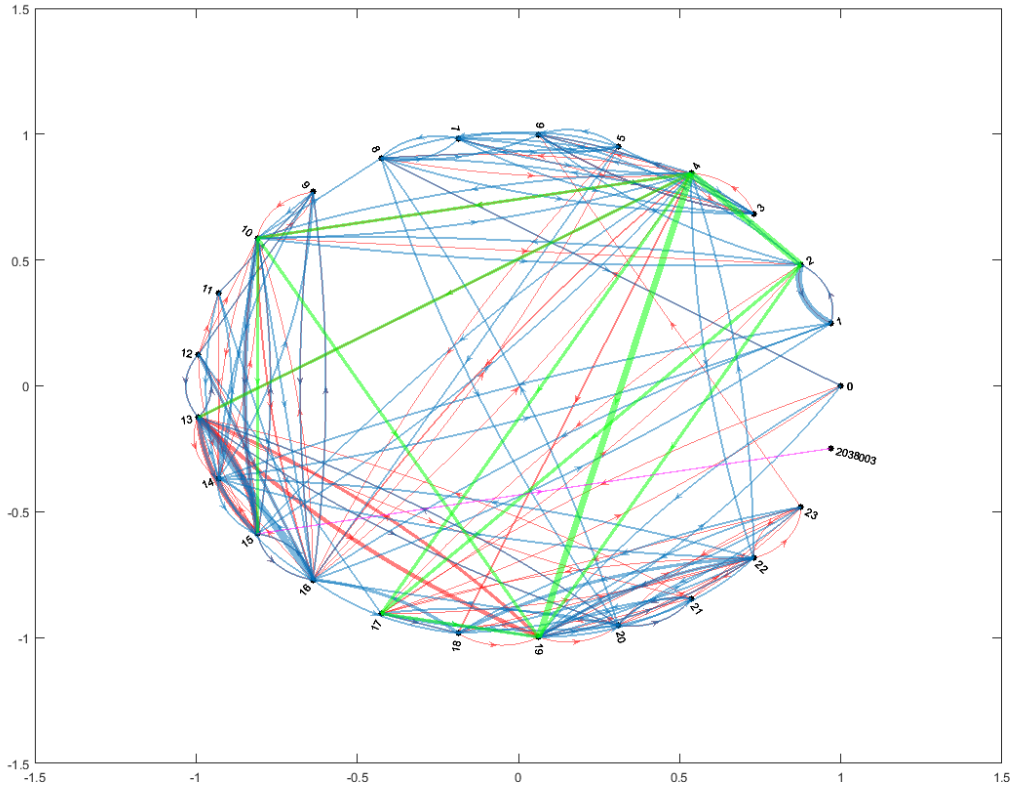
Name	ID	Short ID	Name	ID	Short ID
Alex Hall	505444	0	Kerstin Belveal	728286	13
Lizbeth Jindra	623027	1	Sheilah Stachniw	786361	14
Madeline Nindorf	175354	2	Rosalia Larroque	1847246	15
Patrick Lane	1023940	3	Lindsy Henion	1108217	16
Richard Fox	857138	4	Julie Tierno	1886447	17
Sara Ballard	1663285	5	Ferne Hards	944354	18
Martha Harris	2037156	6	Sherrell Biebel	981554	19
May Burton	387676	7	Jose Ringwald	696047	20
Glen Grant	1642962	8	Ramiro Gualt	234164	21
Dylan Ballard	1039322	9	Tobi Gatlin	969089	22
Meryl Pastuch	1690582	10	Refugio Orrantia	28520	23
Melita Scarpaci	51780	11	Jenice Savaria	2038003	2038003
Augusta Sharp	387200	12			

This graph is a hierarchical layout:



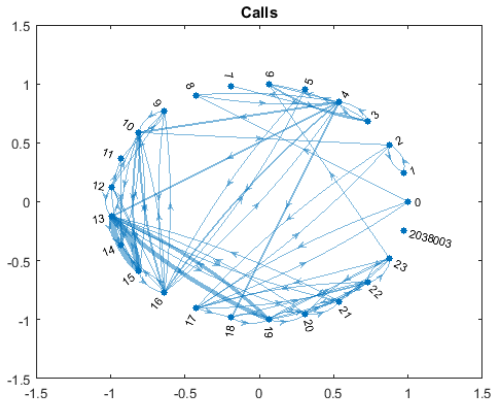
The leader (listed as #4 in this graph) is clearly shown at the left while the buyer and seller are shown at the right.

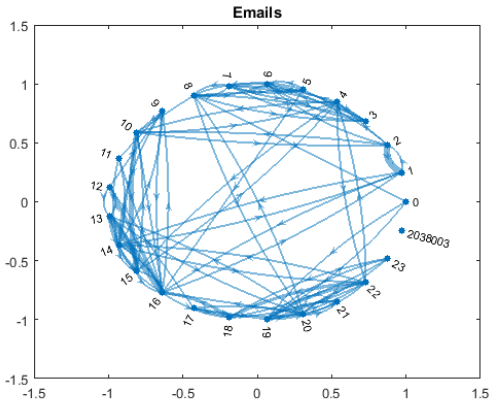
This graph shows the suspects in an oval layout:

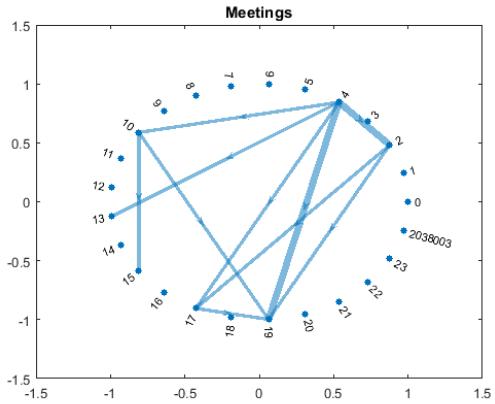


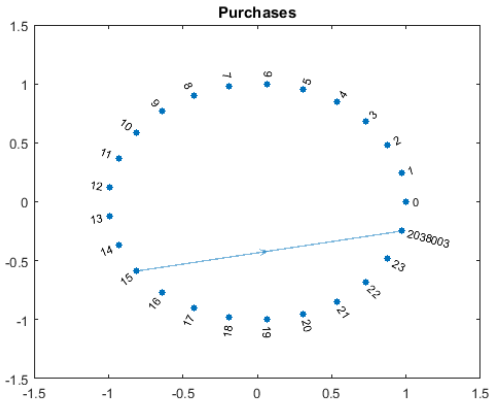
Colors: Calls = red, Emails = blue, Meetings = green, Purchases = magenta. It's easy to see that some people prefer meetings and others prefer email or calls. The groups are evident through more dense interactions.

Comparing calls, emails, meetings and purchases reveals that the suspects prefer to coordinate their activities over email. There are relatively few meetings.

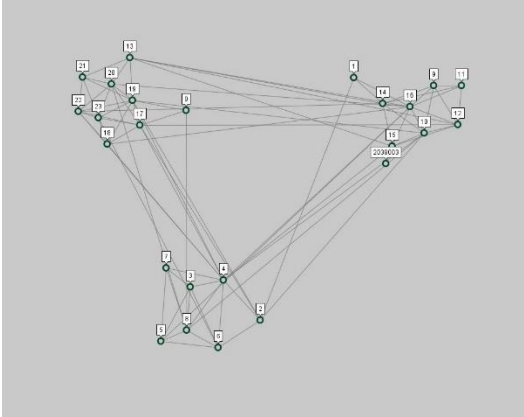






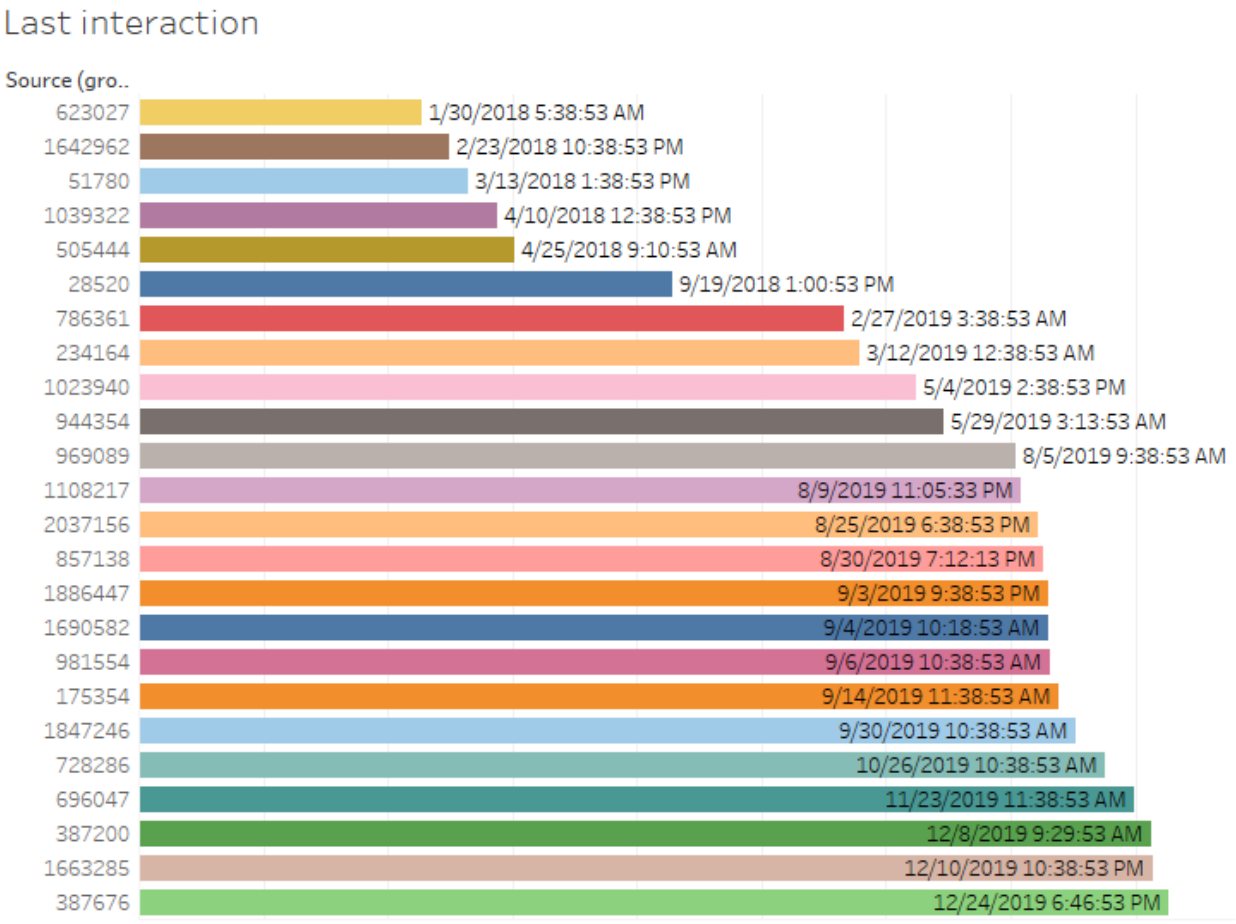


3. *Using the combined group of suspected bad actors you created in question 2, show the interactions within the group over time. Limit your responses to 10 images and 1000 words*
    a. *Characterize the group's organizational structure and show a full picture of communications within the group.*
    The group of bad actors can be separated into three groups based on the individuals they interact most with.



    b. *Does the group composition change during the course of their activities?*

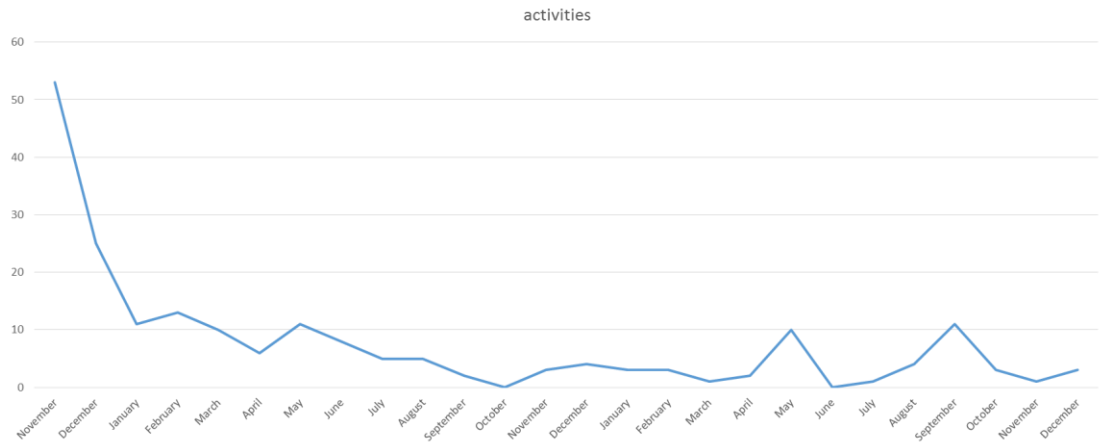
There are several group dropouts that can be seen by their last interaction with any of the other suspects. At least 8 of the suspects don't have any contact with the group after March of 2019.



This table is a list of the suspects who drop out of the group

Name	ID	Role
Alex Hall	505444	finance/logistics
Lizbeth Jindra	623027	builders
Glen Grant	1642962	operational
Dylan Ballard	1039322	builders
Melita Scarpaci	51780	builders
Sheilah Stachniw	786361	builders
Ramiro Gault	234164	finance/logistics
Refugio Orrantia	28520	finance/logistics

c. *How do the group's interactions change over time?*



After the leader's initial meeting there is a burst of activity from the whole group. That eventually calms down over the next few months. More than a year later the leader has to motivate the group again, which spurs more activity. Then there is a burst of activity right around the time that the purchase is made.

This video shows the group's interactions over time:



Suspect graph over time.webm (Click this link to open the video – Windows media player or VLC media player may be used to play it)

4. *The insider has provided a list of purchases that might indicate illicit activity elsewhere in the company. Using the structure of the first group noted by the insider as a model can you find any other instances of suspicious activities in the company? Are there other groups that have structure and activity similar to this one? Who are they? Each of the suspicious purchases could be a starting point for your search. Provide examples of up to two other groups you find that appear suspicious and compare their structure with the structure of the first group. The structures should be presented as temporal not just structural (i.e., the sequence of events—A is followed by B one or two days later—will be important). Limit your responses to 10 images and 1200 words*

   Participants were provided this list of suspicious transactions:

Source	Etype	Dest	Time
580766	2	2038138	53206509
580766	2	2038138	69572943
437025	2	2037766	65109062
437025	2	2037766	81692792
695013	2	2037860	66679529
720233	2	2038031	78658500
720233	2	2038031	81994361

   Participants would find that these 7 transactions are separated into four groups:

Source	Etype	Dest	Time	Group
580766	2	2038138	53206509	1
580766	2	2038138	69572943	1
437025	2	2037766	65109062	2
437025	2	2037766	81692792	2
695013	2	2037860	66679529	3
720233	2	2038031	78658500	4
720233	2	2038031	81994361	4

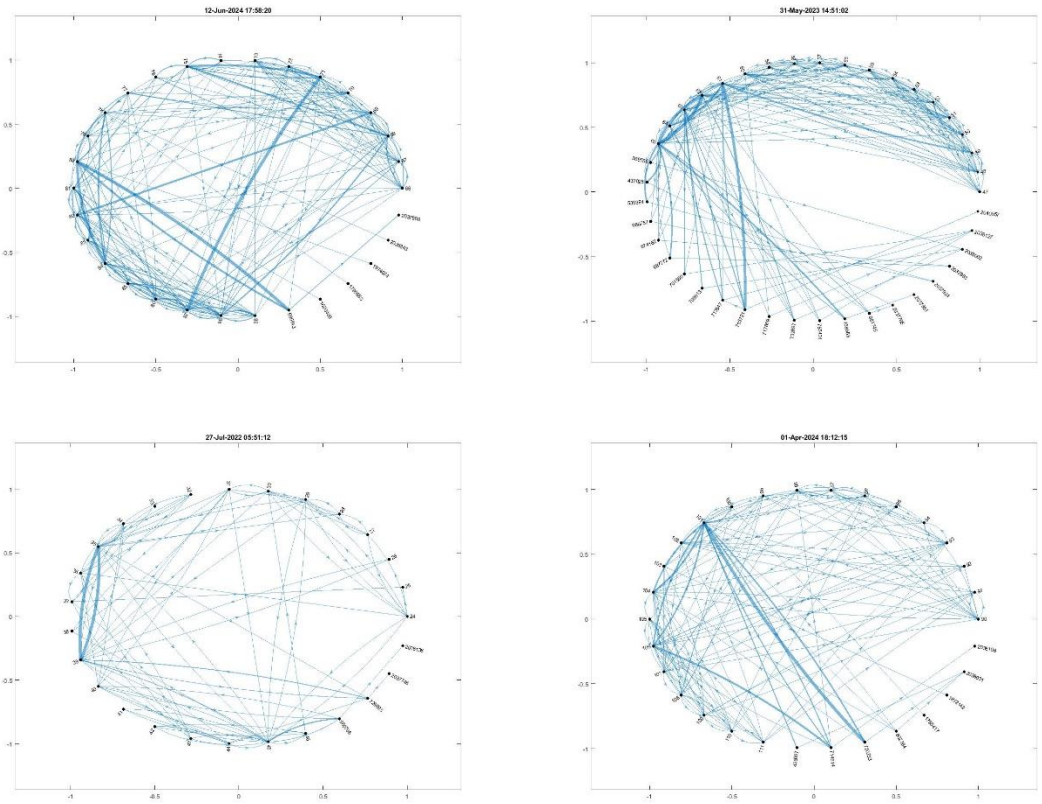
   Participants could find a number of Kasios employees associated with each group (the groups have a structure similar to the group described earlier in the challenge) and even some additional illicit purchases. In total there are 115 individuals and 13 illicit purchases in these “hidden signals.”

Group	# of Participants	Total # of Illicit Purchases	Hidden Illicit Purchases
1	27	3	1
2	34	6	4
3	26	1	0
4	28	3	1
Total	115	13	6

This table contains a full list of illicit purchases that are included in the big data set

Source	Dest	Date	Group
580766	2038138	1/17/2019 9:35:09	1
726693	2037746	6/3/2019 19:49:03	1
580766	2038138	7/25/2019 19:49:03	1
701550	2038127	6/1/2018 20:09:31	2
717869	2038127	10/29/2018 19:18:33	2
437025	2037766	6/4/2019 3:51:02	2
758683	2037861	8/11/2019 3:51:02	2
732897	2038002	11/14/2019 3:51:02	2
437025	2037766	12/13/2019 2:26:32	2
695013	2037860	6/22/2019 8:05:29	3
714514	2038108	9/10/2019 23:35:00	4
720233	2038031	11/7/2019 23:35:00	4
720233	2038031	12/16/2019 14:12:41	4
Highlighted rows were not provided to the participants			

The structure of each group can also be visualized:



Group 1 at the top left had a large number of intergroup transactions whereas Group 3 (lower left) had far fewer interactions. The larger number of purchases by Group 2 (upper right) has a strong influence on the appearance of this graph. In each case it appears that sellers only have contact with a select set of intermediaries within the company.

List of participants in each of the signal groups:

ID	First name	Last Name	Group
1057015	Stacia	Wanat	1
145671	Shayne	Germershausen	1
132540	Artie	Rotkovecz	1
1695111	Katharine	Busenbark	1
1258010	Johanna	Becker	1
1678729	Anne	Parsons	1
1817340	Audrie	Balangatan	1
1909365	Jacklyn	Mashek	1
391229	Lorrie	Crismon	1
1841190	Shalon	Pella	1
1216766	Estelle	Schwartz	1
1256646	Brent	Roberson	1
1680332	Jeremiah	Pearson	1
1090708	Tillie	Hazlett	1
414381	Lelah	Spell	1
1981017	Terrilyn	Overkamp	1
1271503	Jerome	Jordan	1
1425987	Anabel	Coston	1
1476791	Alesha	Aschenbrenner	1
1060156	Ignacio	Thieklin	1
1141575	Lucy	Herrera	1
320914	Cora	Cross	1
1152569	Donnetta	Lapoint	1
580766	Trevor	Webb	1
726693	Archie	Griffies	1
2037746	Nicholas	Newton	1
2038138	Tyree	Barreneche	1
1424784	Virgie	Pratt	2
1714881	Concha	Goodall	2
1499804	Katharine	Santos	2
1590376	Olivia	Brown	2
1505718	Valentine	Klette	2
1422630	Amada	Faul	2
1938283	Lupe	Gullatt	2
1608171	Ilona	Barros	2
1446867	Fonda	Bursch	2
224204	Indira	Fugua	2
103229	Sherri	Brensnan	2
468954	Abbey	Rhead	2

1710613	Amelia	Colon	2
32081	Edgar	McCormick	2
1771151	Karyl	Snobeck	2
160091	Cora	Gonzalez	2
1371959	Zachary	Hampton	2
447025	Renae	Hilbrand	2
1172172	Angelic	Graetz	2
369655	Layla	Mostad	2
437025	Beth	Wilensky	2
701550	Florencio	Schwaller	2
709913	Birdie	Ploch	2
713547	Augustus	Francescone	2
715721	Steven	Palmer	2
717869	Nikia	Wilebski	2
732897	Cecilia	Pichette	2
755431	Ollie	Andrews	2
758683	Anjelica	Hoger	2
981745	Prudence	Rosol	2
2037766	Gregory	Russell	2
2037861	Jewell	Dronet	2
2038002	Janette	Laffoon	2
2038127	Harris	Dearmas	2
250565	Jesica	Pokoj	3
1376314	Marc	Bowen	3
1229528	Regina	Bordoy	3
114786	Elva	Ingram	3
115539	Dina	Fairy	3
312722	Courtney	Wiedemann	3
40236	Lulu	Larson	3
1706578	Sha	Pardoe	3
1266039	Dorothea	Kulback	3
1066593	Marjorie	Halbach	3
200342	Sharmaine	Lofredo	3
1862079	Cathleen	Kucinski	3
221502	Arthur	Fox	3
293560	Kathie	Matheu	3
1147062	Yong	Wilbert	3
9391	Roger	Beck	3
1572181	Alan	Sedotal	3
1740389	Sherlyn	Delcine	3
1880061	Merlene	Tessier	3

206602	Rena	Jerabek	3
1200868	Jayden	Walters	3
1564399	Daria	Housten	3
170126	Virginia	Buchanan	3
350756	Omar	Tako	3
695013	Laure	Pelkey	3
2037860	Carlos	Morris	3
1924635	Carrie	Knight	4
1931527	Dolly	Shimura	4
1721909	Shavonne	Roper	4
1092214	Vernon	Waters	4
1135817	Carrol	Jilek	4
1394563	Sharda	Colonel	4
490696	Royal	Jarratt	4
1187158	Alfonso	Vanslander	4
230932	Nathaniel	Wolfe	4
1708920	Shandi	Kupstas	4
287612	Louise	Ramsey	4
178026	Lillia	Steffensmeier	4
527592	Jen	Masten	4
244450	Bertha	Jackson	4
1448580	Reagan	Engard	4
499479	Lakesha	Gilfillan	4
1609888	Nell	Dixon	4
151930	Elnora	Wallace	4
452527	Ethelyn	Siska	4
2667	Hoa	Mech	4
387938	Jefferson	Kestle	4
292182	Lydia	Russell	4
426067	Huong	Lipari	4
714514	Robbie	Tanzi	4
720233	Alpha	Chessor	4
802384	Aaron	Harper	4
2038031	Kia	Halonen	4
2038108	Conrad	Mira	4

Additional Reviewer Considerations

MC3 Questions and Approach:

- Was the submission able to sort out the connections in the suspicious group within Kasios (relevant to the scenario)? Were multiple variations in patterns identified?
- If there were hypotheses that remained unresolved, did the submission specify actions to be taken to resolve them?

MC3 Application of visual analytics:

- Did the team develop an innovative visual analytic tool? Alternatively, did they use an existing tool in an innovative way?
- Did visualizations enable the analysis process? Or did they merely illustrate conclusions? Did the submission rely more heavily on non-visual analytic approaches?
- Did their tool allow useful interactions?
- Did they use all the available data?
- Was the submission clear?