

Provenance: VAST Challenge 2018

Title: **Mini-Challenge 3**

### **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**

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.

### **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

Download data using your Rowan account at

<https://drive.google.com/file/d/1ogqlr3n0fCiFwLBnRswZJy6qg-glSYpz/view?usp=sharing>

### **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.

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.
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
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