IS428 Visual Analytics for Business Intelligence

Assignment 2 (20 Marks)

1. Overview

The assignments require you to put the concepts, methods and techniques you had learned in this course to solve real world problem using visual analytics techniques. Students should also use the assignments to gain hands-on experience on using the data visualization software or packages we have learned to complete the assignment. The assignment topic is based on Mini-Challenge 2 (https://vast-challenge.github.io/2021/MC2.html) of VAST Challenge 2021 (https://vast- challenge.github.io/2021/description.html). Below is an introduction on the background of this assignment adapted from the official website:

> *In the roughly twenty years that Tethys-based GAStech has been operating a natural gas production site in the island country of Kronos, it has produced remarkable profits and developed strong relationships with the government of Kronos. However, GAStech has not been as successful in demonstrating environmental stewardship.*
>
> *In January, 2014, the leaders of GAStech are celebrating their new-found fortune as a result of the initial public offering of their very successful company. In the midst of this celebration, several employees of GAStech go missing. An organization known as the Protectors of Kronos (POK) is suspected in the disappearances.*
>
> *Many of the Abila, Kronos-based employees of GAStech have company cars which are approved for both personal and business use. Those who do not have company cars have the ability to check out company trucks for business use, but these trucks cannot be used for personal business.*
>
> *Employees with company cars are happy to have these vehicles, because the company cars are generally much higher quality than the cars they would be able to afford otherwise. However, GAStech does not trust their employees. Without the employees' knowledge, GAStech has installed geospatial tracking software in the company vehicles. The vehicles are tracked periodically as long as they are moving.*
>
> *This vehicle tracking data has been made available to law enforcement to support their investigation. Unfortunately, data is not available for the day the GAStech employees went missing. Data is only available for the two weeks prior to the disappearance.*
>
> *To promote local businesses, Kronos based companies provide a Kronos Kares benefit card to GASTech employees giving them discounts and rewards in exchange for collecting information about their credit card purchases and preferences as recorded on loyalty cards. This data has been made available to investigators in the hopes that it can help resolve the situation. However, Kronos Kares does not collect personal information beyond purchases.*

Assignment 2 asks you to analyse movement and tracking data. GAStech provides many of their employees with company cars for their personal and professional use, but unbeknownst to the employees, the cars are equipped with GPS tracking devices. You are given tracking data for the two weeks leading up to the disappearance, as well as credit card transactions and loyalty card usage data. From this data, can you identify anomalies and suspicious behaviours? Can you identify which people use which credit and loyalty cards?

The dataset can be found in "**MC2.zip**" on elearn. You are required to develop visual analytics techniques to finish different analysis tasks.

2. Tasks and Questions

1) Using just the credit and loyalty card data, identify the most popular locations, and when they are popular. What anomalies do you see? What corrections would you recommend to correct these anomalies? **Please limit your answer to <=8 images and <=500 words**. **(4 marks)**

2) Add the vehicle data to your analysis of the credit and loyalty card data. How does your assessment of the anomalies in question 1 change based on this new data? What discrepancies between vehicle, credit, and loyalty card data do you find? **Please limit your answer to <= 8 images and <= 500 words**. **(4 marks)**

3) Can you infer the owners of each credit card and loyalty card? What is your evidence? Where are the uncertainties in your method? Where are the uncertainties in the data? **Please limit your answer to <= 8 images and <= 500 words**. **(4 marks)**

4) Given the data sources provided, identify potential informal or unofficial relationships among GASTech personnel. Provide evidence for these relationships. **Please limit your response to <= 8 images and <= 500 words**. **(4 marks)**

5) Do you see evidence of suspicious activity? Identify **1 - 8 locations** where you believe the suspicious activity is occurring, and why. **Please limit your response to <= 10 images and <= 500 words**. **(4 marks)**

Apart from writing your responses to the above five tasks, you are asked to provide the link to your demo or your code. **(1 mark)**

Also, you are asked to write a brief description on **the challenges of visualization designs** and introduce **the procedures** on data processing and how you have implemented the necessary visualizations. **Please limit your description to <= 10 images and <= 800 words**. **(4 marks)**

Notes:

i) You can use either Tableau or D3 to develop your visual analytics system.
ii) For each task, please make sure your conclusions are supported upon the attached visualization images. No marks will be provided for a conclusion/observation without valid visualization image(s) to support it.
iii) For each image, please add a brief caption for it to illustrate the overall content in the image/figure.
iv) If you are not able to finish all the five tasks above, you may try to finish as more tasks as possible.

3. Deliverables

A report in pdf or word format and you are asked to use the provided word template ("**Report-template.docx**").

If you are using Tableau, please also provide the link on Tableau public or directly runnable twbx file of your Tableau dashboard(s) with your word/pdf report; if you are using D3, please also submit your code (and the necessary datasets) to elearn and provide a brief explanation on how to run your code in your word/pdf report.

4. Submission

This assignment must be your independent effort. You should NOT refer to or copy other students' work. Plagiarism is strictly prohibited.

Deadline: Mar. 24th, 2023 (Friday), 11:59pm (mid-night)