The Kronos Incident

Simple Geospatial, Temporal, and Transaction data analysis using R-Shiny and Rstudio.
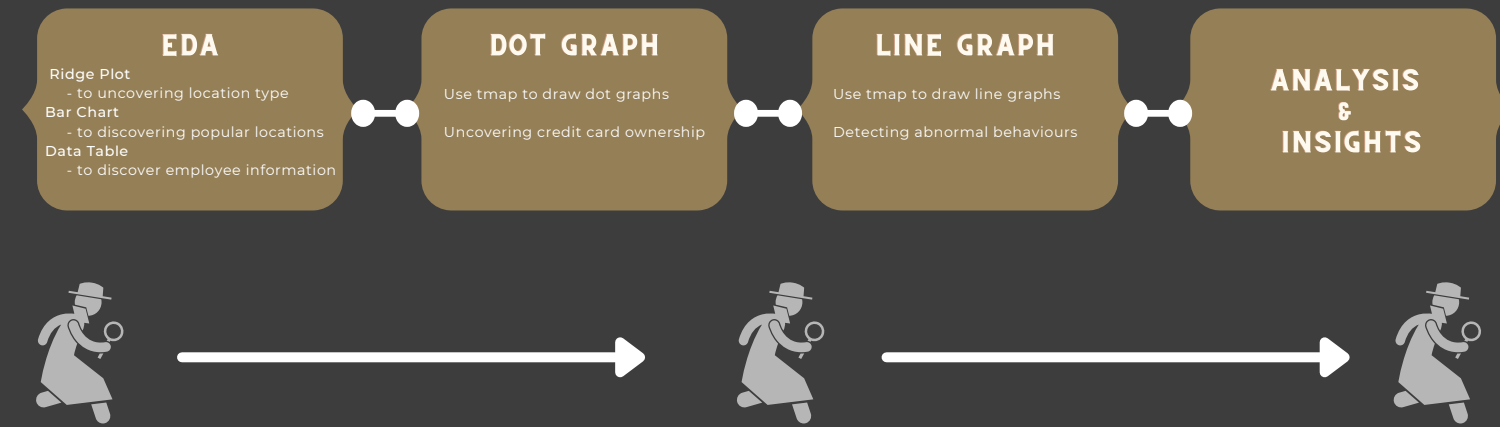
Group 8:

Bai Yining | Lin Peirong | Liu Manye

04 APPLICATION FLOWCHART

EDA	DOT GRAPH	LINE GRAPH	ANALYSIS & INSIGHTS
Ridge Plot - to uncovering location type; Bar Chart - to discovering popular locations; Data Table - to discover employee information	Use tmap to draw dot graphs. Uncovering credit card ownership	Use tmap to draw line graphs. Detecting abnormal behaviours	

INTRODUCTION

» CONTENT

- Problem Statement
- Project Objective
- Motivation
- Application Flowchart
- Result
- Future Work

» TOOLS WE USED

- Rstudio: To develop and share our work
- tidyverse: Used in data wrangling techniques
- dplyr: Provide operators/functions for data wrangling techniques
- ggplot2: Basic package in plotting various types of graphs
- plotly: To create interactive graphical visuals
- tmap: To visualize geospatial data
- lubridate: To convert and deal with unstructured datetime
- DT: To provide interactive data tables
- sf: Loading of Geospatial shapefiles
- Shiny: To build interactive web app

01 ISSUES AND PROBLEMS

GAStech has been operating a natural gas production site in the island country of Kronos for the past 20 years. It has been remarkably profitable but GAStech has not been as successful in demonstrating environmental stewardship.

In January 2014, several employees of GAStech go missing in the midst of the leaders' celebration of their new-found fortune gained from the initial public offering of the company. An organization known as the Protectors of Kronos (POK) is suspected of the disappearance, and further investigation is still required.

02 OBJECTIVE

As visual analytics students, we want to identify which GASTech employees made which purchases and identify suspicious patterns of behavior. We will cope with uncertainties that result from missing, conflicting, and imperfect data to make recommendations for further investigation. Our ultimate goal is to shed some light on the people and the locations that the law enforcement should investigate by analyzing the movement and tracking data provided, thereby contributing to the investigation towards finding the missing staff members.

03 MOTIVATION

- The interesting background story of the case.
- Availability of GPS tracking dataset which does not often found in visual analytics cases.
- The attractiveness of performing geospatial analysis apart from traditional analytics. Geospatial analysis has developed a lot in recent years. Data visualization using geographical mapping methods can bring extra relationships and insights than traditional visualization methods.
- Opportunities to provide interactive analytics tools that allow users to investigate the case without understanding programming.

05 RESULTS

1. EXPLORATORY DATA ANALYSIS

1.1 **Interactive Ridge Plot** build in RShiny allows selection of location category. Ridge plot on the right side displays the visiting time distribution feature of different location categories. One location with an "unknown" location category can be chosen for plot on the right side, so users can compare two plots and uncovering its category.
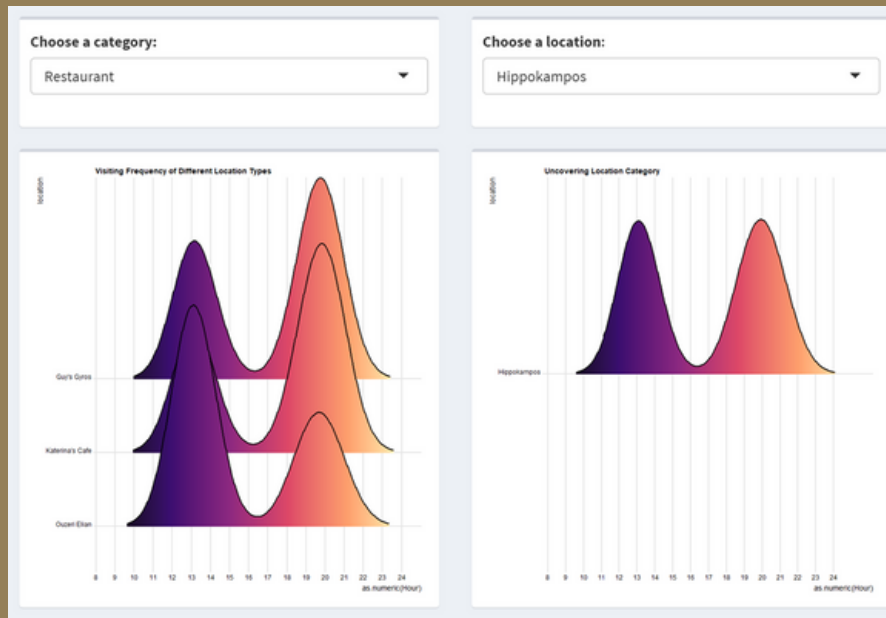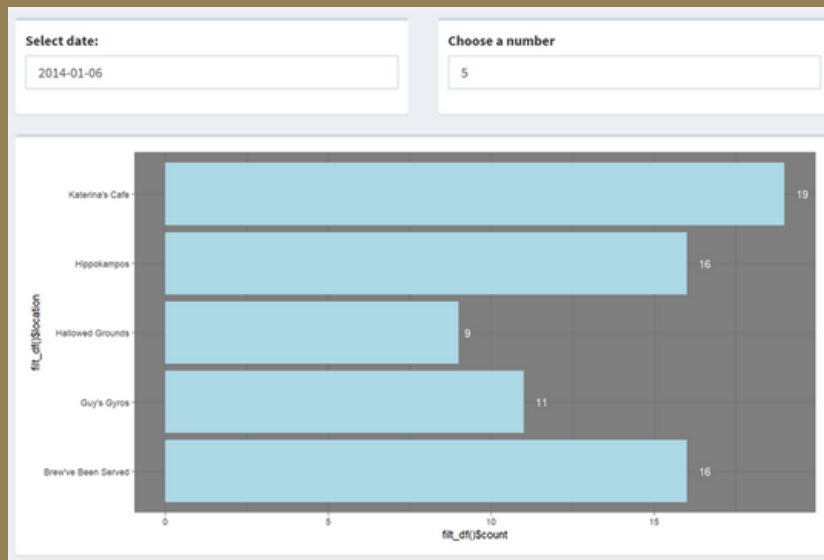

Example 1.1: Hippokampos should be a Restaurant
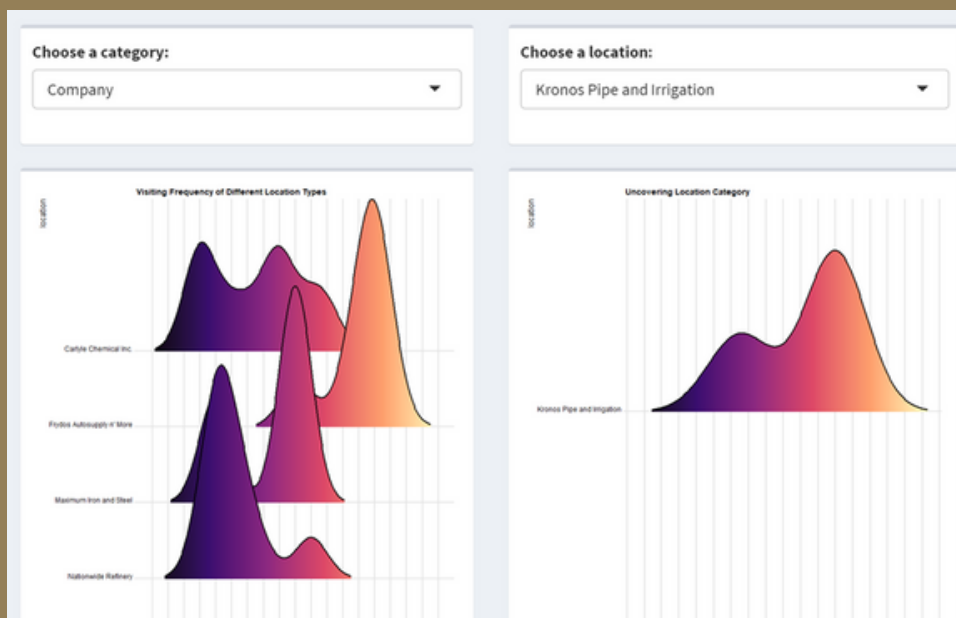
Example 1.2: Kronos Pipe and Irrigation should be a Company

1.2 **Interactive Bar Chart** build in RShiny allows selection of date and top n number. Chart plot below will display popular locations in a specific date. User can discover popular locations in different weekdays, hence identify abnormal phenomenons.


1.3 **Interactive Data Table** build in RShiny allows filter of Name, CarID, Employee Type, and Employee Title. Users can explore the basic information of GAStech's employees by filtering any conditions we need.

2. DOT GRAPH

The GPS information during movement is actually not useful for location verification, so we can use the stop points data only to draw the dot graph.
The interactive geospatial graph on Shiny allows users see where the selected person is at the specific time selected.
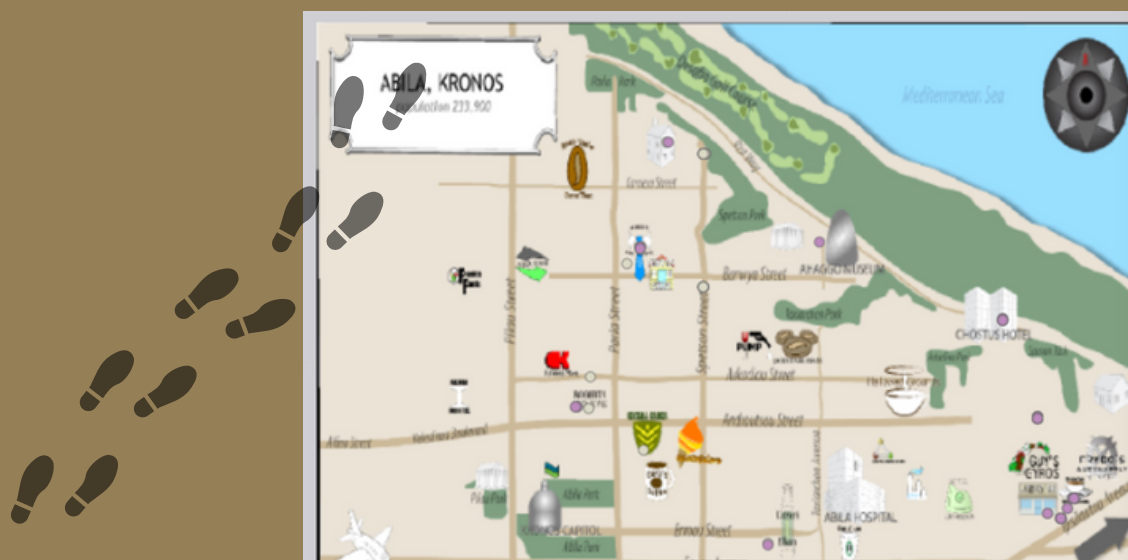

3. LINE GRAPH

The Line graph on Shiny dashboard allows user to investigate the whereabouts of a specific person on the specific day selected.
This graph is actually very useful to assist law enforcement from Kronos and Tethys to identify abnormal trace and suspects who may participate in this Kidnapping case.

4. ANALYSIS & INSIGHTS

The first suspicious pair is Card ID 33 and 7. The two visited the Chostus Hotel several time at 13pm and visited restaurants at evening. And they even have some tracks in the same apartment. The owners of the cars are Elsa Orilla and Brand Tempestad, who have the same employment type and employment title.
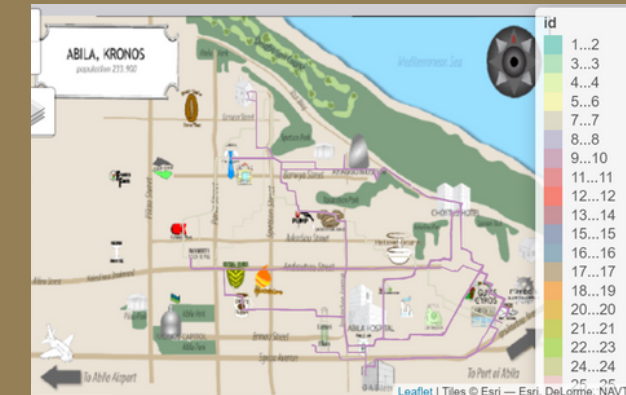
The second suspect group is Car ID 22, 30 and 15. There is a lot of overlap in the tracks of these cars. The owners of the three cars are Nubarron Adra, Loreto Bodrogi and Felix Resumir. These three employees are all security guards. Therefore, we can assume that they have comparably close relationship.

Dot graph below shows tracing spots at midnight when generally very few people are outside. 15, 21, 24, and 16 gathering in the Frydo's Autosupply N More in late night. Someone even drive to parks far from the downtown center. Suspicious activity locations can be identified: Spetson Park, Taxiarchon Park and Frydo's Autosupply N More.

05 FUTURE WORK

There are some visualizations that might be suitable for static but not too good to show in the shiny as well as interactive way, such as heatmap. Therefore, some static visualizations can be added to help us analyze in the future.

More data can be collected, such as email or text messages to better defining the relationship between employees. Since the linkages of credit cards and loyalty cards may be not very accurate to define the relationship between employees. Some text data will be helpful to find the relationship and the social networks.