

Group 1 ShinyVA - A Shiny Application for Crime Detection

Li Nan
Singapore Management University
nan.li.2020@mitb.smu.edu.sg

Li Yueting
Singapore Management University
ytli.2020@mitb.smu.edu.sg

ABSTRACT

This is the abstract.

It consists of two paragraphs.

1. INTRODUCTION

This paper is based on analytics for the Mini-Challenge 1 & Mini-Challenge 2 of the VAST Challenge 2021.

In a fiction scenario, a gas-production company named Tethys-based GASTech has been operating a natural gas production site in the island country of Kronos and 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. And in January, 2014, the leaders of GASTech are celebrating their new-found fortune, but 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 disappearance, but things may not be what they seem. For this fiction scenario, we focus on mini-challenge1 and mini-challenge2 of VAST Challenge 2021.

- Mini-challenge1:

Mini-challenge1 provides (adding more...)

- Mini-challenge2

Mini-challenge2 provides 4 datasets to data analysts for exploration. These datasets contain 3 aspects. (Transaction/GPS/Care Assignment) Through data manipulation and visualization, it can be detected anomalies that appear in unmatched transaction records between loyalty card and debit/credit card and gps data also shows the unfamiliar movement for some employees. Through the background stated, this paper develops interactive visualization approaches to provide evidence and suspicious behaviors of GASTech employees. The

introduction of this paper is followed by an explanation of our motivation and objectives in Section 2, then followed by Section 3 which details the data used and methodology selected. And Section 4 provides a visual overview of the final application and finally section 5 provides the conclusion and insights for this paper.

2. MOTIVATION AND OBJECTIVES

This project was motivated by a desire to i) identify the complex relationships among all of these people and organizations. ii) track data for the two weeks leading up to the disappearance, as well as credit card transactions and loyalty card usage data. To achieve these aims, the interactive tools are developed to addresses the following requirements: (adding graphs and utility)

- What is the suspicious pattern in company cars for employees' personal and professional use? - i) utilize geographical visualization techniques to create car/truck routes of GASTech employees based on gps dataset and also label parking points of car (parking more than 5 minutes) to dig out suspicious movement patterns.
- how to utilize gps data to match card owners and the debit/credit card owners - ii) Create interactive transaction data table to track unusual credit/debit card transactions records of GASTech employees and compare them with car parking points shown in visualization map.

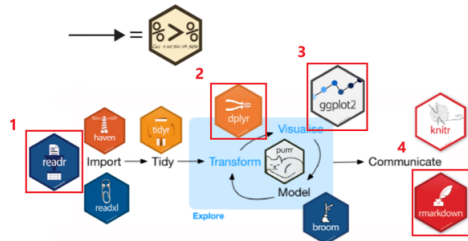
3. METHODOLOGY

The paper utilize data in three aspects: i) Find relationships between different data tables and try to use different data manipulation methods (e.g. group and filter data) to deepen understandings of data story. ii) Based on the understanding of data, find explicit and effective data visualization methods to deal with varied data types (e.g. data table – currency data; spatial data – geospatial visualization). iii) Point out and try to give reasonable explanations and insights of data anomalies based on data visualization tools. To enable readers to have deeper understanding about data manipulation in this project, an explanation of the different methods used follows.

3.1 Data Manipulation Flow

The raw data manipulation is followed in the following 4 steps by using tidyverse package in R (shown in Figure 1.1): i) Step 1 is to read varied dataset (e.g. csv file, spatial data

Data Science Workflow with Tidyverse



Shiny is an R package that makes it easy to build interactive web apps straight from R. Users can host standalone apps on a webpage or embed them in R Markdown documents or build dashboards. The reason why choosing Shiny as interactive tool is as follows: i) Input values can be changed by the user at any time, through interaction with customizable widgets. ii) Output values and graphs react to changes in inputs, with the resulting outputs being reflected immediately. iii) Table and graph can be interactive by setting logics in server code. The output values react to changes in inputs, with the resulting outputs being reflected immediately. The design of the Shiny web app flowed from the key business questions of the data: i) First tab – Network Analysis (EDA) to form relationships between). ii) Second tab – Geospatial visualization and interactive data table help the user understand and match the relationship between card owners and car drivers and also provide insights of unusual movements of car drivers to detect suspicious persons. To give better understandings and guidance from users, we included a user guide to accompany each analysis technique featured in the app.

3.3.1 Data Table

Transaction Data						
last4ccnum	loyaltynum	location	price	Day	hour	
All	All	All	All	All	All	
1286	L3572/L3288	Brew've Been Served	14.97	06	08	
1286	L3572/L3288	Abila Zacharo	50.14	06	13	
1286	L3572/L3288	Brew've Been Served	11.92	07	07	
1286	L3572/L3288	Kalamí Kafénion	45.05	07	13	
1286	L3572/L3288	Brew've Been Served	7.26	08	08	

Thematic maps are geographical maps in which spatial data distributions are visualized. This package `tmap` offers a flexible, layer-based, and easy to use approach to create thematic maps, such as route map. Shiny package takes a fresh, interactive approach to telling the data story. Another package used in this map creation, `sf`, supports for simple features, a standardized way to encode spatial vector data.



awaiting to add...

The final interactive graphs in the published web application, and an illustration of the potentially wide range of insights that can be gleaned from the application are briefly described below. What info can gain from Interactive route map - Residence: Here we can identify the living place by changing the input of “Tracking Day,” and we can see every day Car ID2(Azada_Las) will appear in this location, so we can identify its living place is here

Car ID

2

Tracking Day

11

☒ Show data table

Figure 1 ID2_Day11_Starting point



Figure 4 ID2_Day13

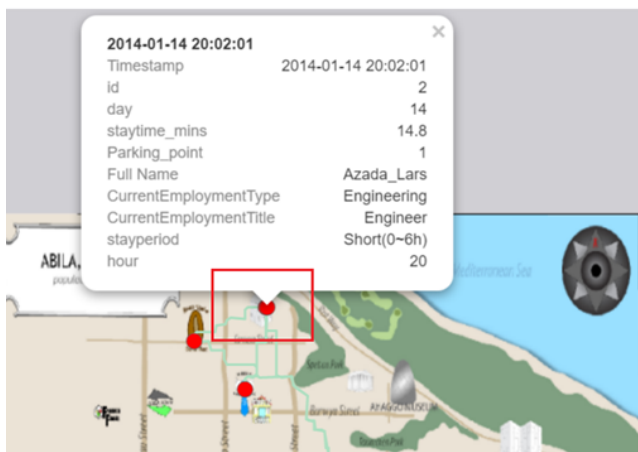
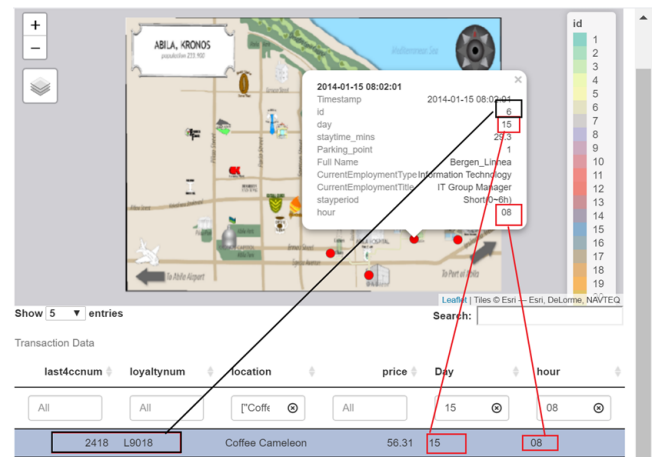


Figure 4ID2_Day14

- Match credit card owner and car owner

Through the stopping points which show info about day and hour and the corresponding transaction record, users can easily see the relationships between car owners and credit card owner, which provides strong support for future detecting work for unusual transaction data.



- Can detect people who gather together in one day

If we finally notice some suspicious people we can see their behaviors by clicking multiple Car IDs, we can detect in one day people who gather together in one place.



5. CONCLUSION AND FUTURE IMPROVEMENT

This paper set out the development of a web application targeted at providing an interactive visualization tool to help minimize time to have a thorough understanding of person's background (e.g. car owner/credit card owner/daily movement/daily spend) and detect suspicious patterns.

Adding on:...

However, shiny app also have some demerits that needs to be improved in the future. For the route map app: 1. Consider add "select all" action button. This will help to detect each day all the driver's movement. 2. Instead of using Car ID as input, consider using Employee title and Employee department as input, which can help to find relationship between different departments. 3. Consider Add starting time and end time in the tooltip, which can be more helpful for detecting suspicious gathering groups.

6. ACKNOWLEDGMENTS

The authors thank Associate Professor, KAM Tin Seong, Singapore Management University for his support and guidance, Classmate Liu Yang Guang for giving advice and suggestions for data visualization and Classmate Syed Ahmad Zaki whose assignment expand ideas and logics.

References

- [1] Fenner, M. 2012. One-click science marketing. *Nature Materials*. 11, 4 (Mar. 2012), 261–263.

- [2] Meier, R. 2012. *Professional Android 4 Application Development*. John Wiley & Sons, Inc.