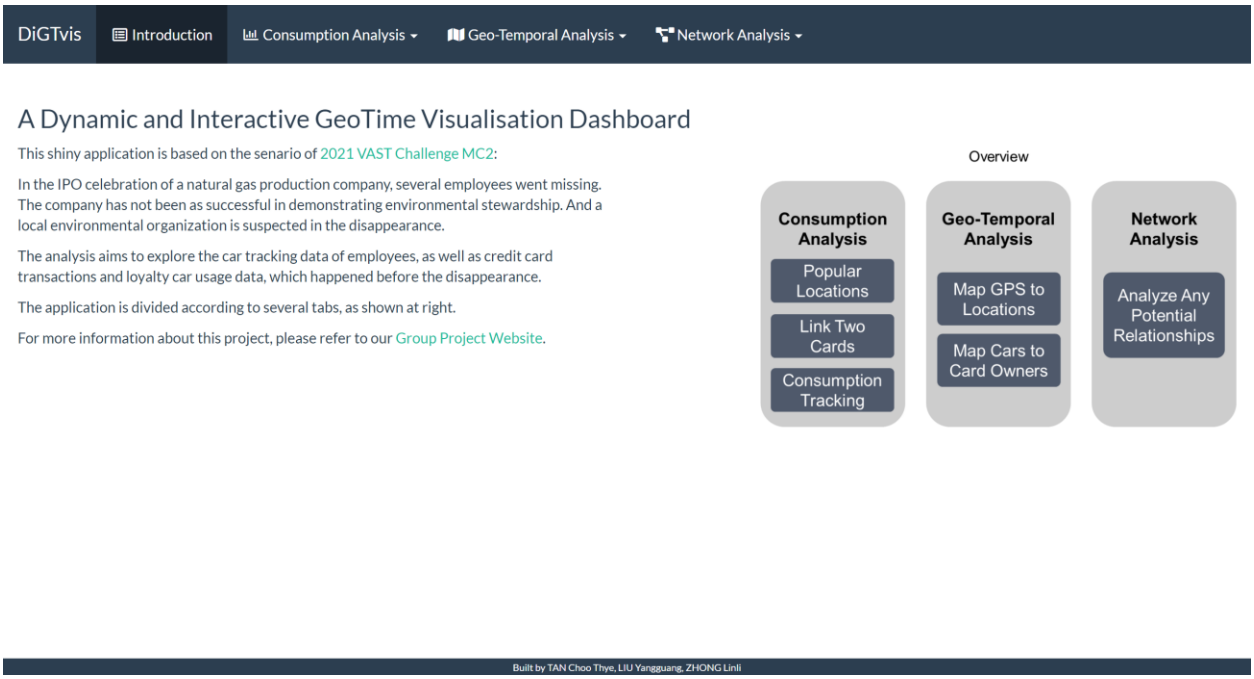
User Guide – 2021 Vast Challenge MC2

1. Introduction Page

This is the home page of the shiny application. Users can directly read the overview along with the story behind the application. This page provides the link of Vast Challenge and Group Project Website for users at the same time.



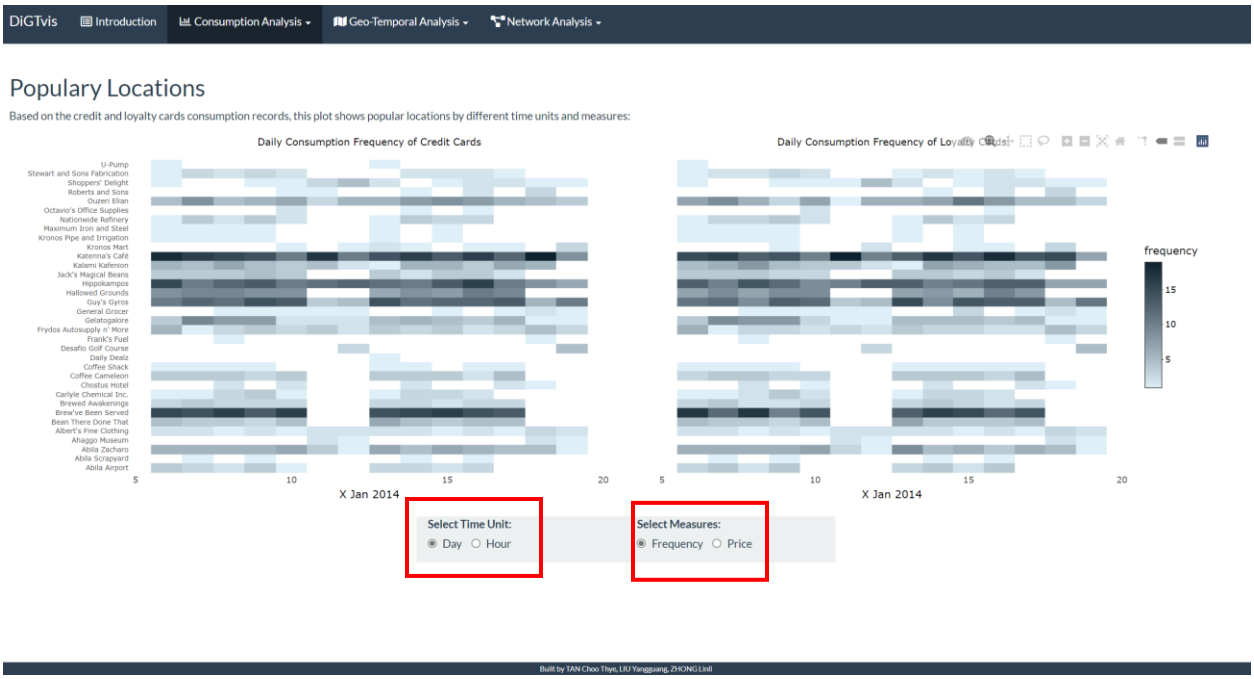
2. Consumption Analysis

2.1 Heatmap

Heatmap is a data visualization technique that shows magnitude of a phenomenon as color in two dimensions. The DiGTvis provides two heatmap, one for credit card, another for loyalty card under the same tab with fixed locations variable in the Y axis. Another dimension can be defined by users.

[1] Time Unit: Users can flexibly choose the time range as day or hour for Credit Card Heatmap, the graph will automatically change the x axis accordingly. If users choose hour in the “Time Unit”, the Loyalty Cards Heatmap will disappear because loyalty card csv didn’t provide information about time.

[2] Measures: Users can define popularity by two scales, transaction frequency and price purchased. Here we provide these two measurements for users to choose. After selecting one of the two measurements, the tooltip of the graph will also be updated accordingly.



2.2 Bar Chart

Bar chart is a graph that represents categorical data with rectangular bars with heights proportional to the values that they represent.

In this panel, users can view the frequency of transactions with the kind of transaction ways, namely employees used credit card, loyalty card or both cards. This graph is in descending order. Hence, users can compare the popularity of locations easily.

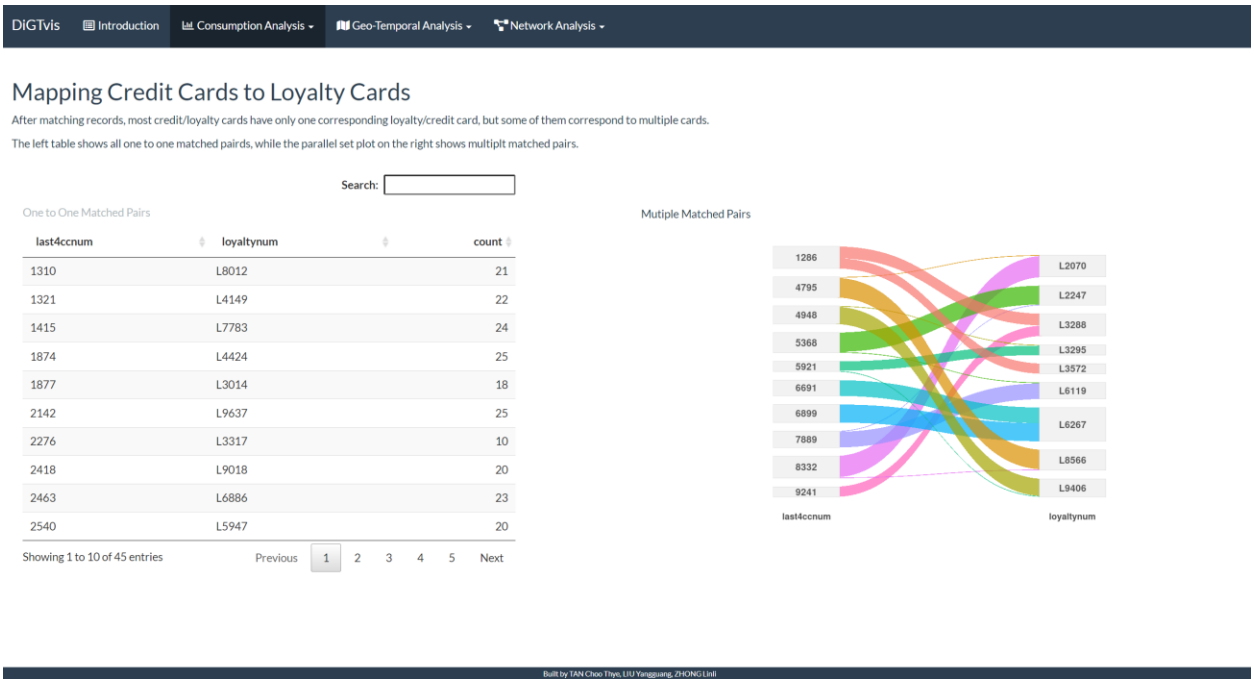
2.3 Parallel Set

Parallel Set Chart is to show data frequencies instead of the individual data points. Each flow path aims to show the relationships between different categories.

In this panel, credit card and loyalty card table are joint together.

[1] Data table. In the left-hand side of this panel, we provide the result of matching two tables. In each row, users can get the credit card and loyalty card used by one employee and how many times of this credit card and loyalty card had been used at the same time.

[2] Parelle set. From the matched data table, users can view 10 credit cards and 9 loyalty cards are matched with more than one cards with each other.

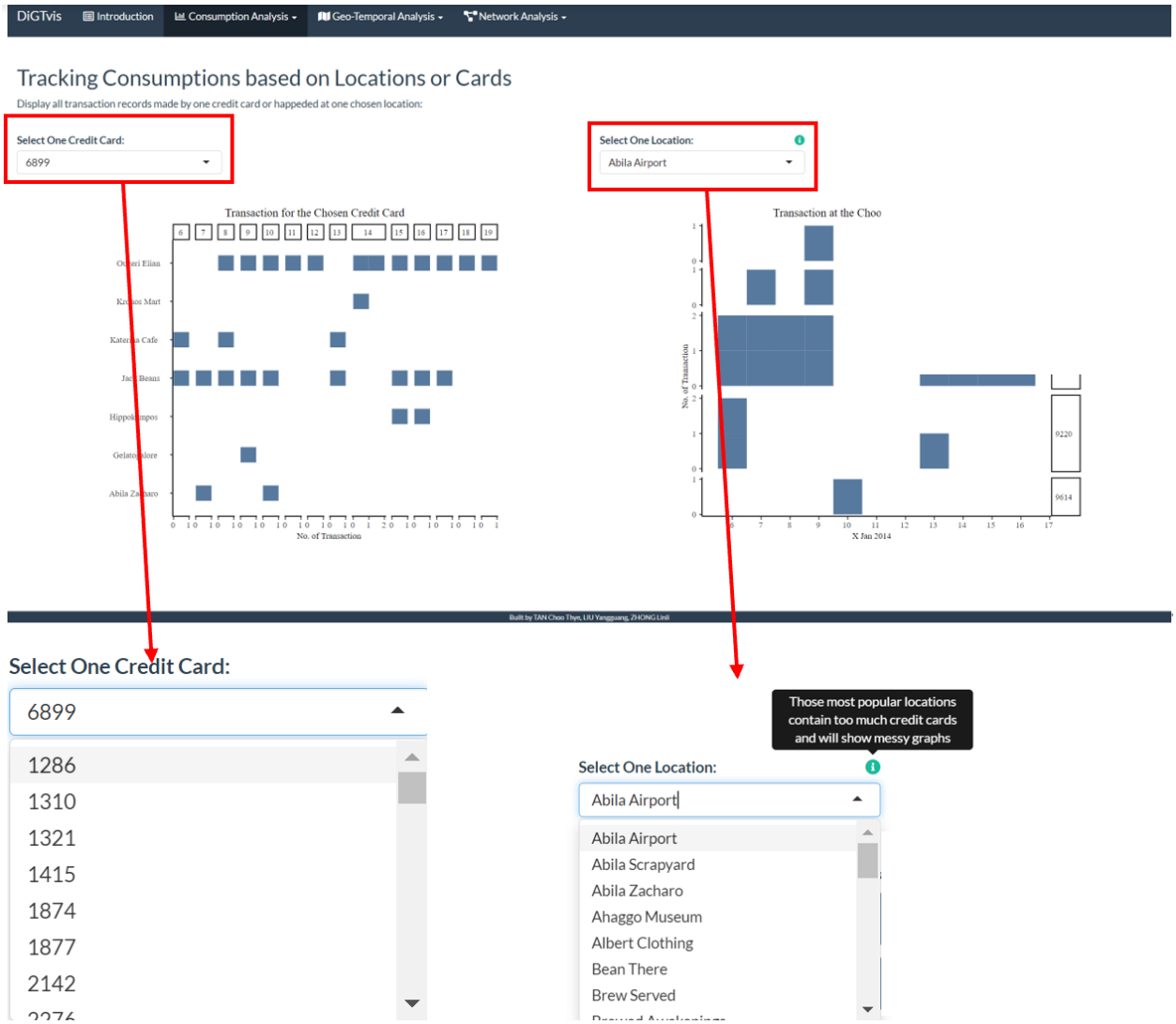


2.4 Facet Plot

Under this tab, DiGTvis Shiny app aims to show the faceted bar plot for users to detect the exacted transaction time by choosing transaction credit card or location.

[1] Credit Card Selection panel: the drop-down list shows the last 4-digits of credit card number. Users can select one credit card number. System will return the chosen credit card transaction information in the left-hand side of the panel. In the x axis, it will return the number of transactions within a day, and in the y axis, it will return exact locations. Users can check the transaction time in the tooltip by highlight action.

[2] Location Selection panel: Users can select one location from the drop-down list. The x axis is day of the time, and the y axis is the number of transactions.



3. Spatial Analysis

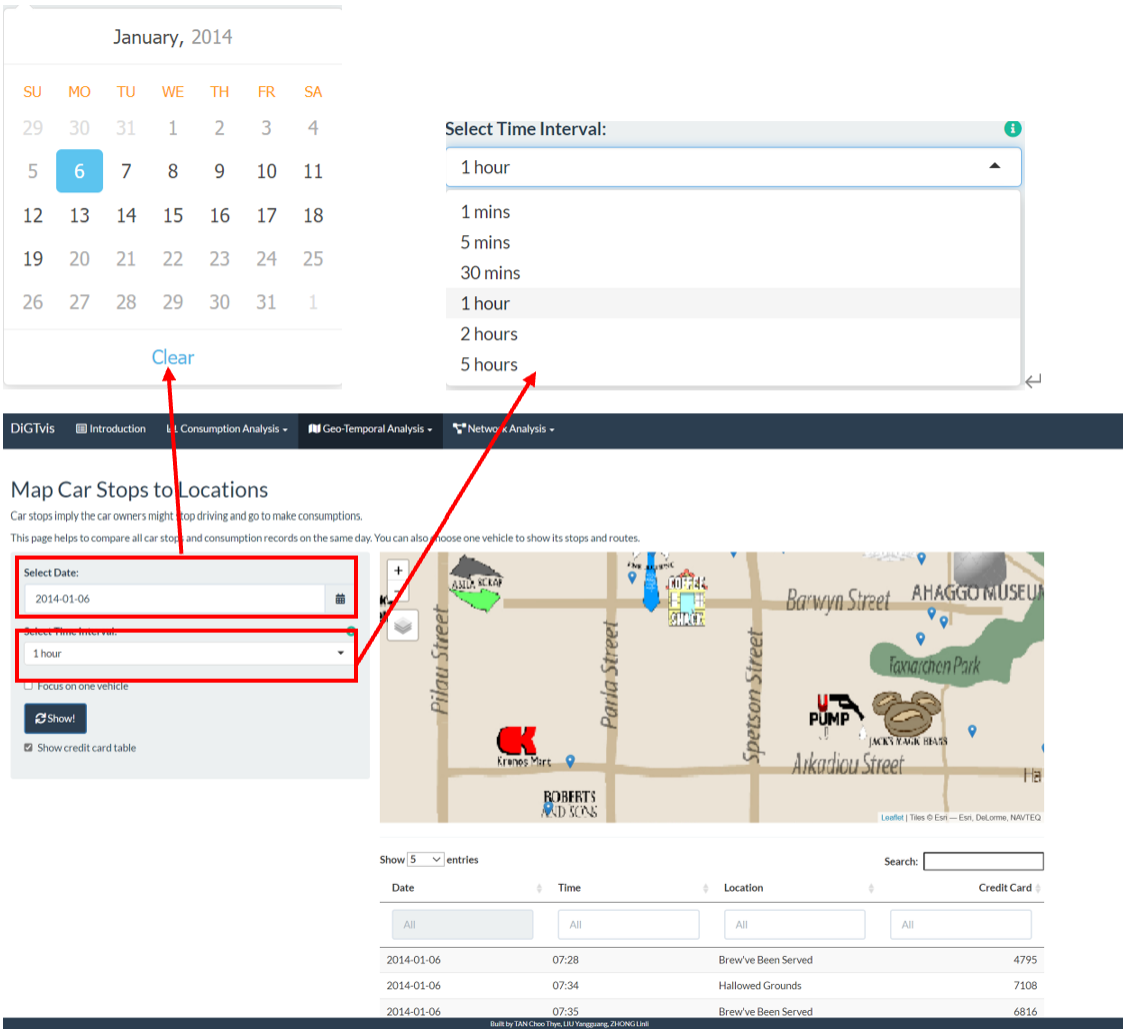
3.1 GPS Tracking

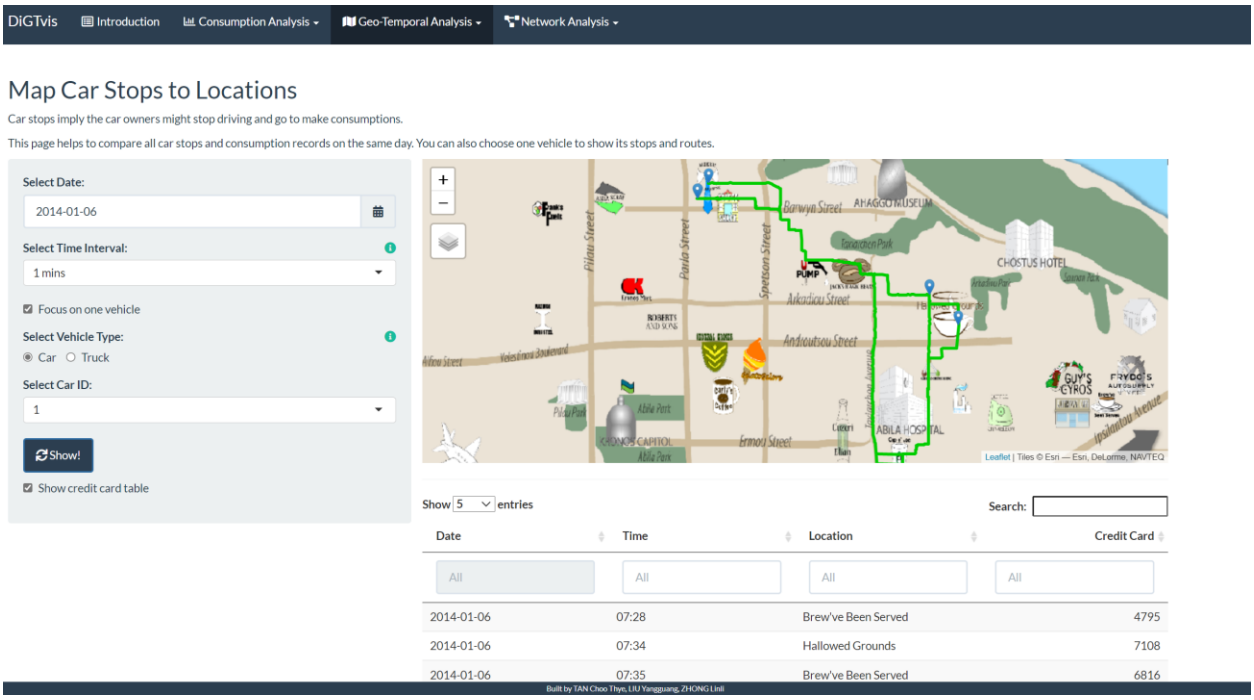
Maps provide two ways to visual suspicious activities. Parking spots are to show the staytime in one position compared with other cars' spots. Route tracking can track the travel route of specific cars by date to find out the discrepancies from daily routes.

[1] Parking spots: Users can customize the date and time interval (time range between arrival and departure). After setting these two variables, users need to click show button. R shiny will automatically return all the vehicles (company cars and trucks) stopping spots on the map. Tooltip will provide the car owners information if it is a company car. On the contrary, company trucks will not get the information of the truck owner because 5 trucks are assigned to 7 truck drivers. At the same time, users can get the information about date and time from the tooltip.

[2] Route tracking: above show tab, there is a checkbox named “focus on one vehicle”. There will be two additional variables appear as long as users tick the “focus on one vehicle” checkbox. When users choose car, the “select Car ID” dropdown list’s values will be 1 to 35. When users choose truck, the dropdown list will be automatically updated to the truck id (101, 104, 105, 106, 107). Users are free to choose the car id or vehicle id as you wish. After defining these two variables, “show” tab need to be clicked to show the travel route of the defined car.

[3] Credit Card Data Table: credit card table has 4 columns to show the date/ time, location and credit card 4-digits number. Users can compare the table value with the map to spot the location name in the map. This table can be filtered by select date tab or use the search bar under each column.

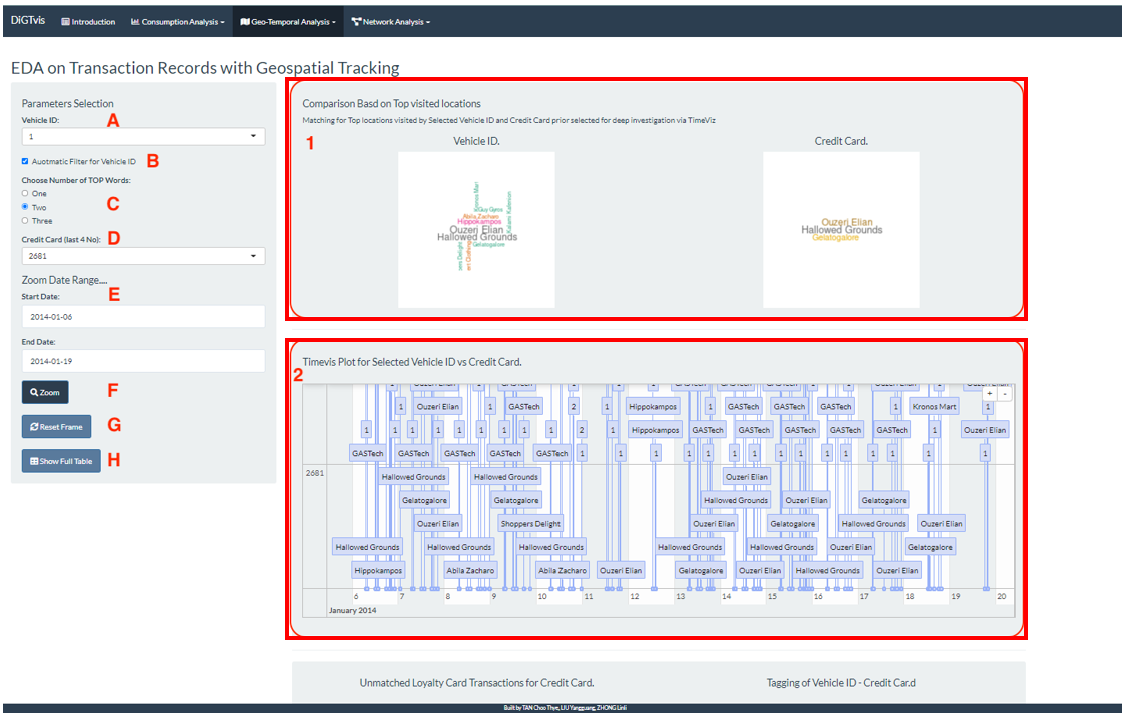




3.2 Mapping GPS

Under this tab, DiGTvis shiny aims to support an investigator in determine the Credit Card owner through mapping of transacted locations in credit card dataset with the travelled locations for vehicle id in geospatial dataset.

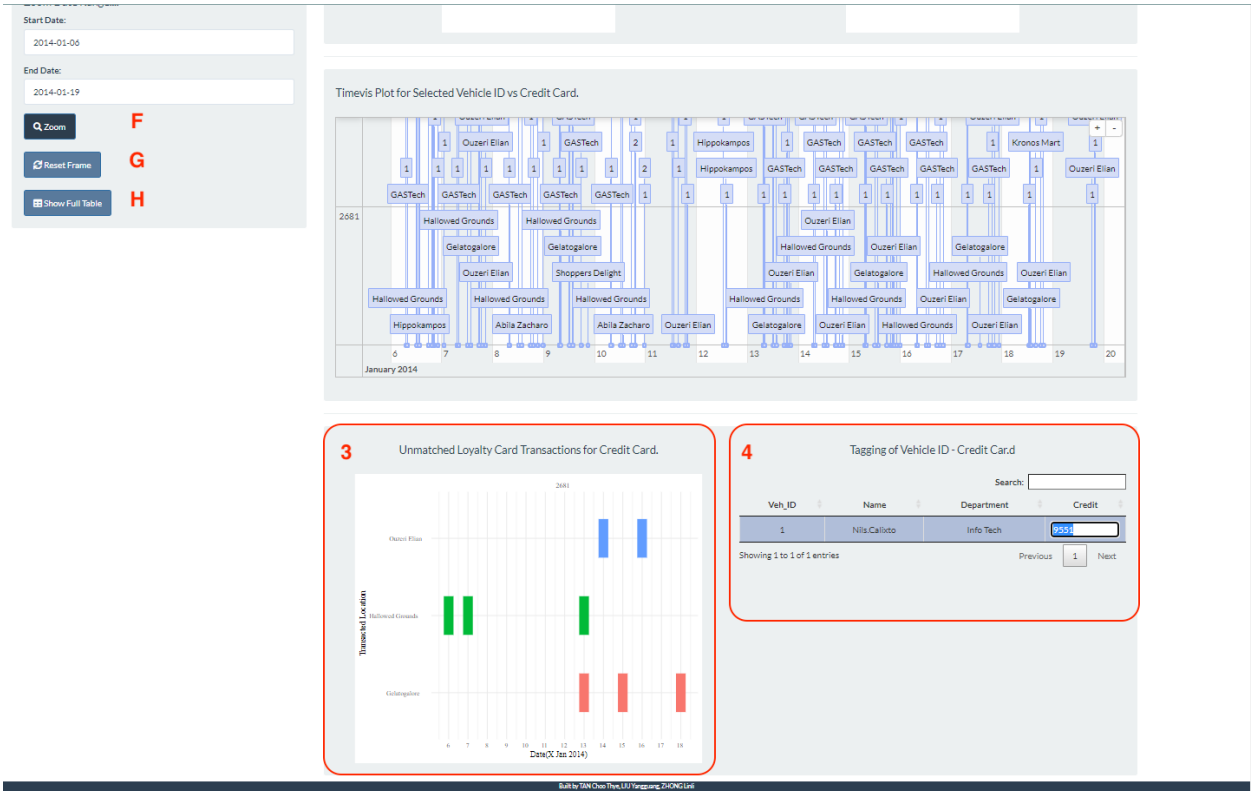
[1] wordcloud – Dropdown list of [A] comprises all vehicle IDs in the dataset. The left most word cloud shows the travelled location for the vehicle id selected via [A]. Once a vehicle ID is selected with [B] and [C], dropdown list of credit card [D] will auto filter by selecting credit card number with top [C] transaction locations contains at least the top 2 travelled locations of vehicle ID. User can then select the filtered credit card number [D] to determine the best match in both word clouds prior validation in timeviz [2]. [B] and [C] shall be changed when match could not be established. A higher value of [C] will imply a longer list of credit card for match. User can then select the filtered credit card number and observed the to determine the best match.



[2] timevis is used to compare the travelled against the transacted locations. A transacted location shall lie in between two similar visited locations for the vehicle id. (i.e. vehicle reach and left the location). timeviz will auto filter the its chart based on selected values in vehicle ID [A] and credit card number [B]. One can zome in and out, pull left or right through click and hold on a mouse. To zoom into a specific duration, one can select the date range [E] and click the action button “Zoom” [F]. Action button [G] will reset timevis back into the initial frame. When there are travelled location without transaction, one could verify with chart [3] to determine whether indeed there is no transaction, or it has been a loyalty card transaction. Unmatched loyalty card is not included into timevis chart due to lack of time stamp information.

[3] Unmatched Loyalty Card Transaction for Credit Card

[4] Datatable where one can double click into the table to update the credit card number based on vehicle id when matched is found. One can click action button [H] to display the full table. The table will reset to the initial form once a new vehicle id is selected

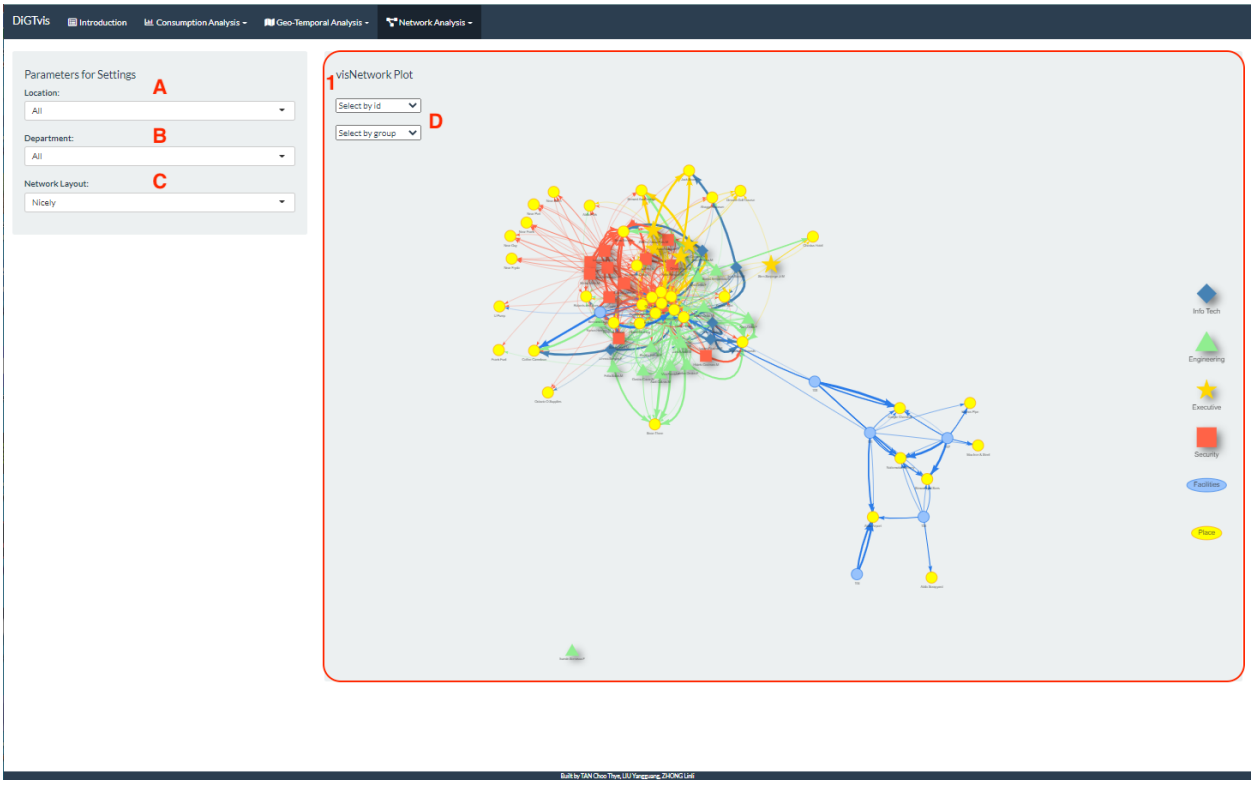


4. Network Analysis

4.1 Graph Network

A directed network structure built based on geographical tracking data. Nodes connotes vehicle id or travelled location where edge connotes the frequency a vehicle travelled to the location. It Network analysis could show the level of interactions or association among the nodes.

[1] Show the network structure based on selection of Location [A], Department [B] and Network Layout [C]. Location here refers to either the vehicle ID or location of interest where it could be transacted location, eg. Jean Beans, residential eg 3 (home of vehicle id 3), naming starts with Near i.e. locations not in transaction records. Interaction [D] is inbuilt interaction via visNetwork that allow one to highlight respective node ID or group of nodes belong to the same department. The interaction intent is for one to filter into the network to see specific network structure. Eg filter by Department [B] will show all the location and frequency of visit by personnel within the department. One can hover over the edge or node within the chart to see the node identity or visit frequency. Different layout could offer different visual effects and thus comprehension on the network structure. However, its do not change the underlying properties of the network structure.



4.2 Community Detection

To detect the present of community within a display network structure based on Location [A] selection.

[1] Display the network base on selection of Location [A] and Department [B]. The Location here is a preselection of a group of nodes based on location travelled instead of individual entity like previous panel. The present Location are (1) amenities – refers to nodes that have travelled to amenities location like cafe, pump, shopping and restaurants etc. It can be further sub categorized into visitation in the morning (AM,7am to 11am), afternoon (PM, 11am to 6pm) and night (NT, after 6pm to 7am), (2) residential – refers to node that have made residential visit during the study period, (3) unknown – refers to node that have made visitation to locations not part of transaction records, (4) industrial – refers to visitation by company trucks only.

[2] The dendrogram displays the communities identified based on selection of Community Detection [C] algorithm. Only two algorithms are used here given they support direct graph. i.e. leading Egien and Walktrap. The community detection will auto compute based on the graph structure display in A.

[3] Display the modularity score and the node values of community that not singleton when action button Generate Report [D] is clicked.


DiGTvis
Introduction
Consumption Analysis
Geo-Temporal Analysis
Network Analysis
Parameters for Settings
Visitation to:
A
Amenities
Network Layout:
B
Nicely
Community Detection:
C
Leading Eglen
Generate Report
D
visNetwork Plot
1
Select by id
Select by group
E
Info Tech
Engineering
Executive
Security
Facilities
Place
Community Dendrogram Plot
2
Community Dendrogram Plot
3
MODULARITY SCORE/COMMUNITY(>1)
Modularity Score: 0.360892266483647
Community 1 :
Katerina Cafe Gelatogalore Hallowed Grounds Linnea.Bergen.F Lucas.Alcazar.M Lidelise.Dedos.F
Birgitta.Frente.F Kanon.Herrero.M Marin.Onda.M Kare.Orilla.F Bertrand.Ovan.M Coffee Cameleon
Nils.Calixto.M 105
Community 3 :
Hippokampos Albert Clothing Ingrid.Barranco.F Ada.Campo-Corrente.F Bean There Felix.Balas.M
Elsa.Orilla.F Gustav.Cazar.M Axel.Calzas.M Vira.Frente.F Lars.Azada.M Brewed Awakenings
Desafio Golf Course
Community 4 :
Ouzeri Elian Isak.Baza.M Orhan.Strum.M Jack Beans Kronos Mart Willem.Vasco-Pais.M
101
Community 2 :
Guy Gyros Abila Zacharo Kalami Kafenion Brew Served Isia.Vann.M Frydos Autosupply
Felix.Resumir.M Edvard.Vann.M Minke.Mies.M Inga.Ferro.F Loreto.Bodrogi.M Sven.Flecha.M
Adra.Nubarron.F Varja.Lagos.F Shoppers Delight Stenig.Fusil.M Brand.Tempestad.M Hideki.Cocinaro.M
Hennie.Osvaldo.M Ahaggo Museum General Grocer Sten.Sanjorge Jr.M Chostus Hotel Roberts and Sons
Octavio O Supplies U Pump Frank Fuel
Built by TAN Choo Thye, LIU Yanggang, ZHONG Lirli