User Guide – A Visual Exploration of the employees' behaviour in GAStech

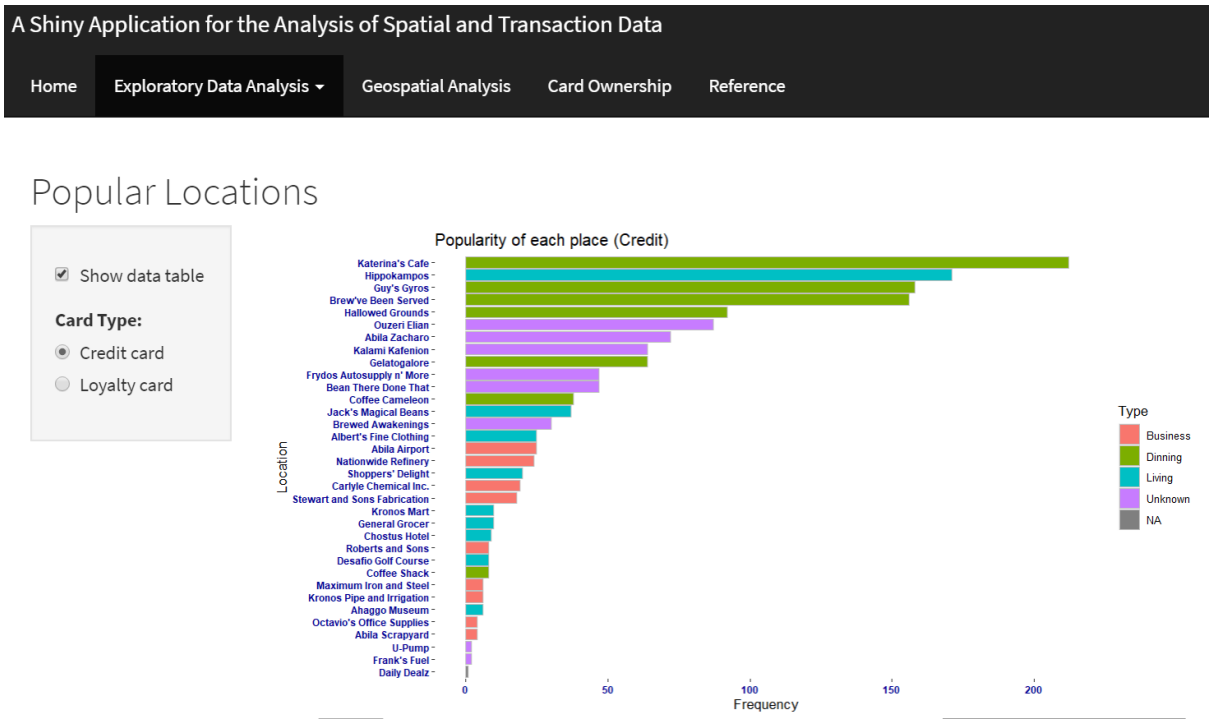
1. Introduction Page

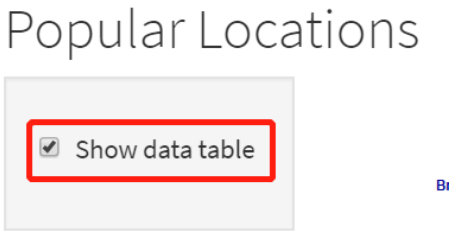
On this page, there is a short description of the application and this whole case.

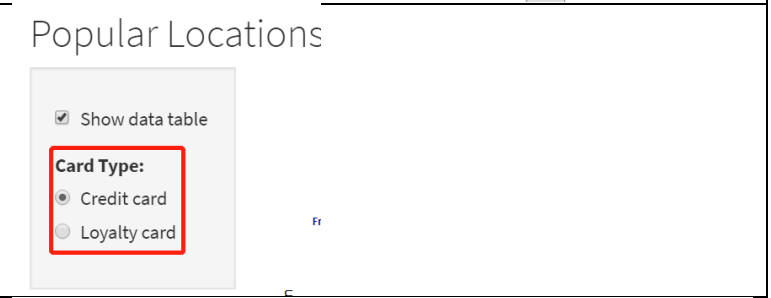
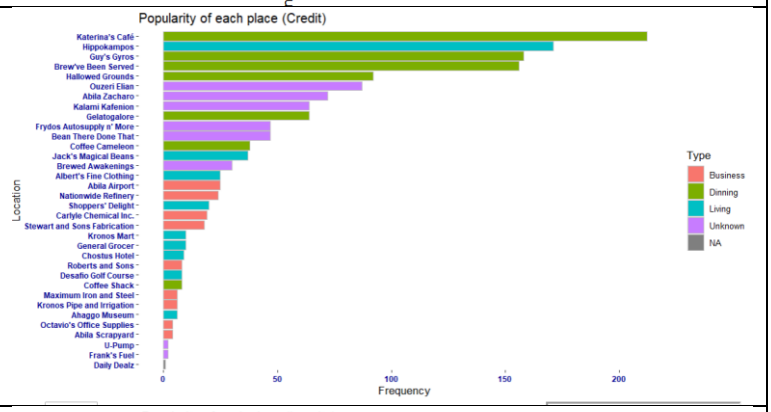
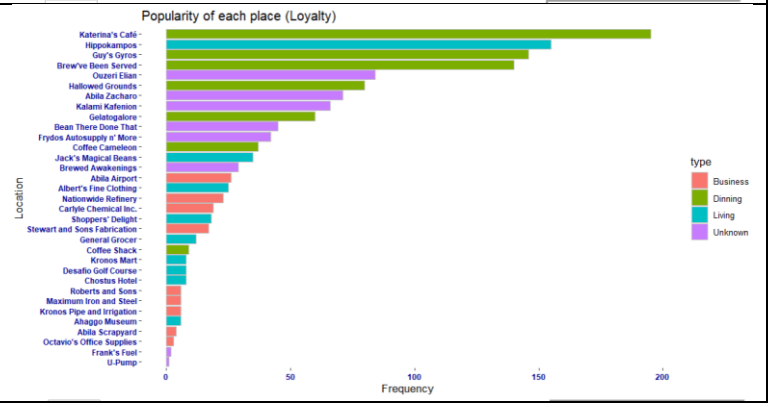
2. Exploratory Data Analysis (EDA)

2.1 Bar Chart

The bar chart showed the popularity of each location by the number of times people visit it.

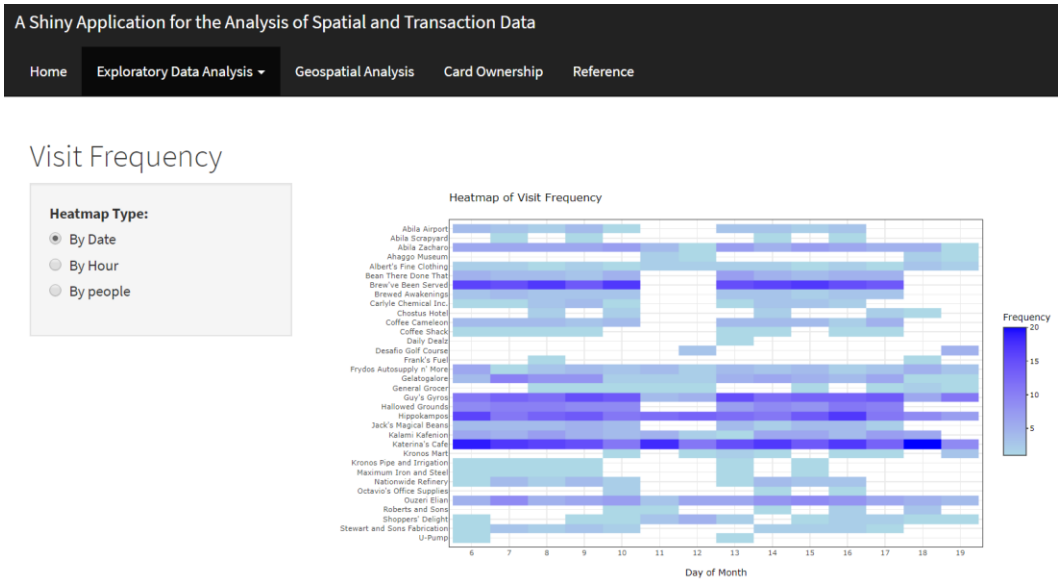


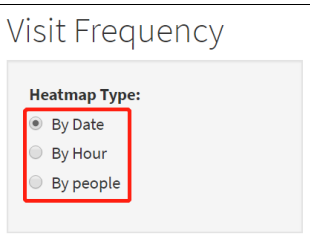
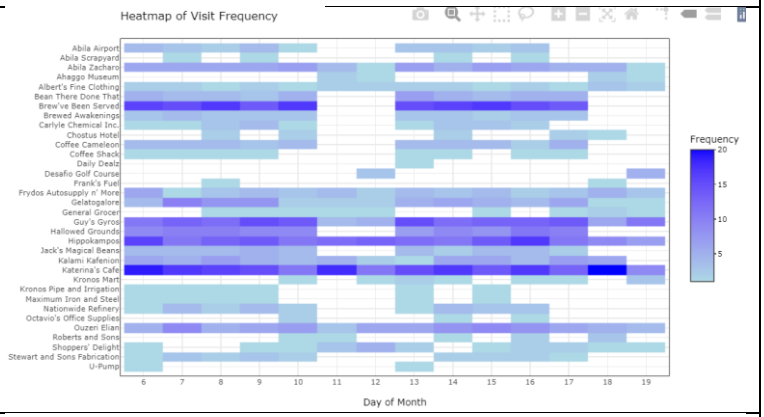
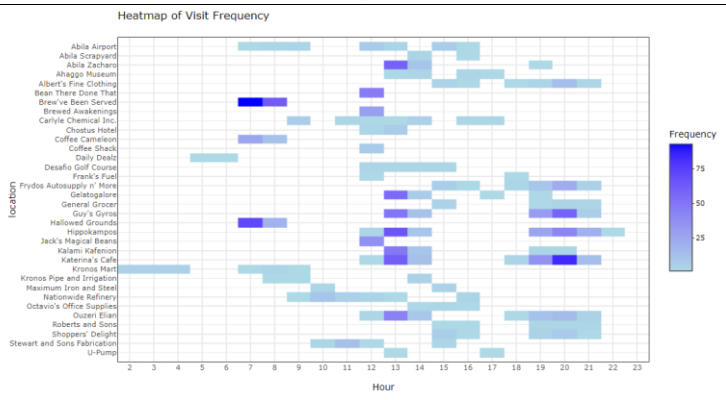
No.	Step	notes
1	Click "Show data table" to generate a table shown the detailed information, for example, type of location and number of times that people visit, of each location.	

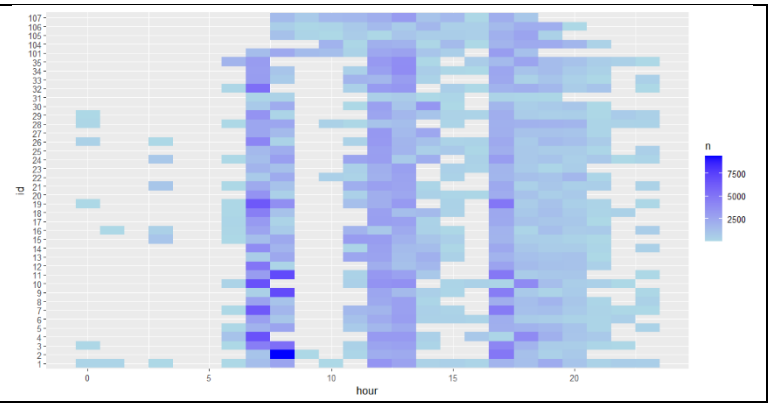
2	Results	Show 10 entries Search: Location | Number_of_Times_People_Visit | Type Abila Airport 25 Business Abila Scrapyard 4 Business Abila Zacharo 72 Unknown Ahaggo Museum 6 Living Albert's Fine Clothing 25 Living Bean There Done That 47 Unknown Brew've Been Served 156 Dinning Brewed Awakenings 30 Unknown Carlyle Chemical Inc. 19 Business Chostus Hotel 9 Living Showing 1 to 10 of 34 entries Previous 1 2 3 4 Next
3	Click different button on the side bar will changed to different bar chart	
4	There are two kinds of bar charts. The first one is credit card data	
5	The second one is loyalty card data	

2.2 Heat Map

A heat map is a data visualization technique that shows magnitude of a behavior as color in two dimensions.

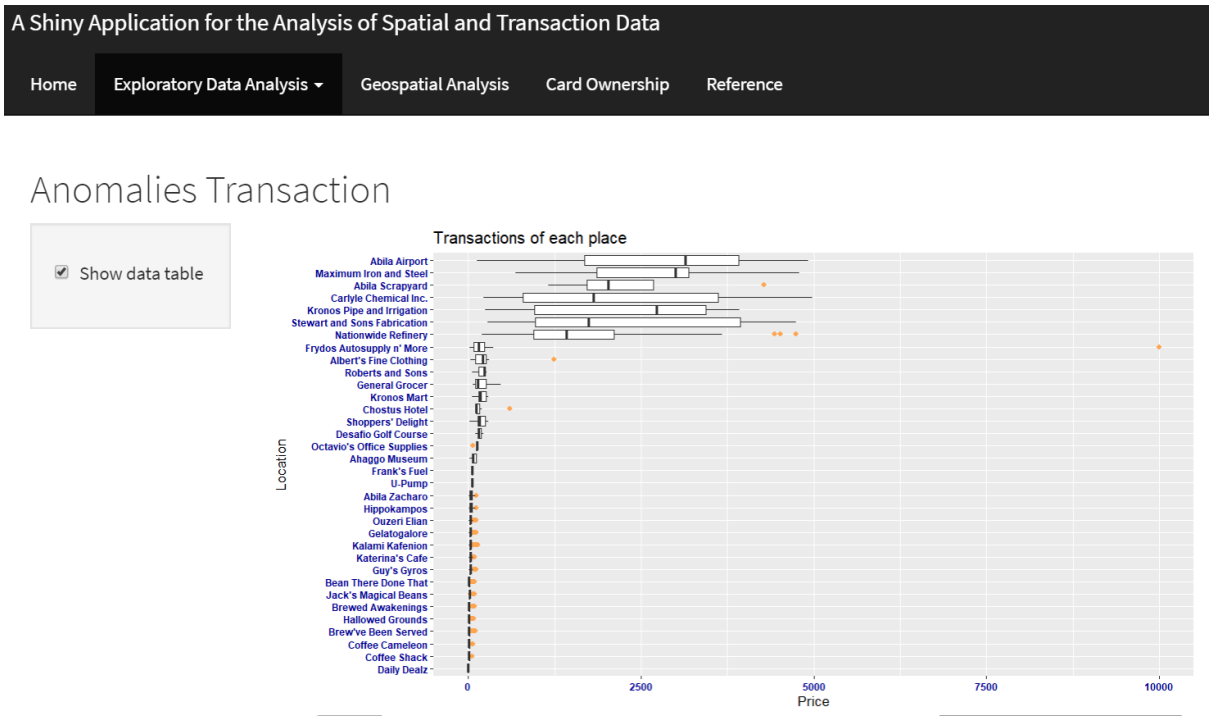


No.	Step	notes
1	Click different button on the side bar will changed to different kind of heatmap	
2	There are three kinds of heat maps. The first one shows the visit frequency of each location by date. Users can see the detailed information when put mouse on the blocks	
3	The second one shows the visit frequency of each location by hour. Users can see the detailed information when put mouse on the blocks	

4	The third one shows the visit frequency of each location by hour. Users can see the detailed information when put mouse on the blocks	

2.3 Box Plot

To explore the insights about transaction, we preferred to use boxplot, which graphically depicting groups of numerical data through their quartiles.

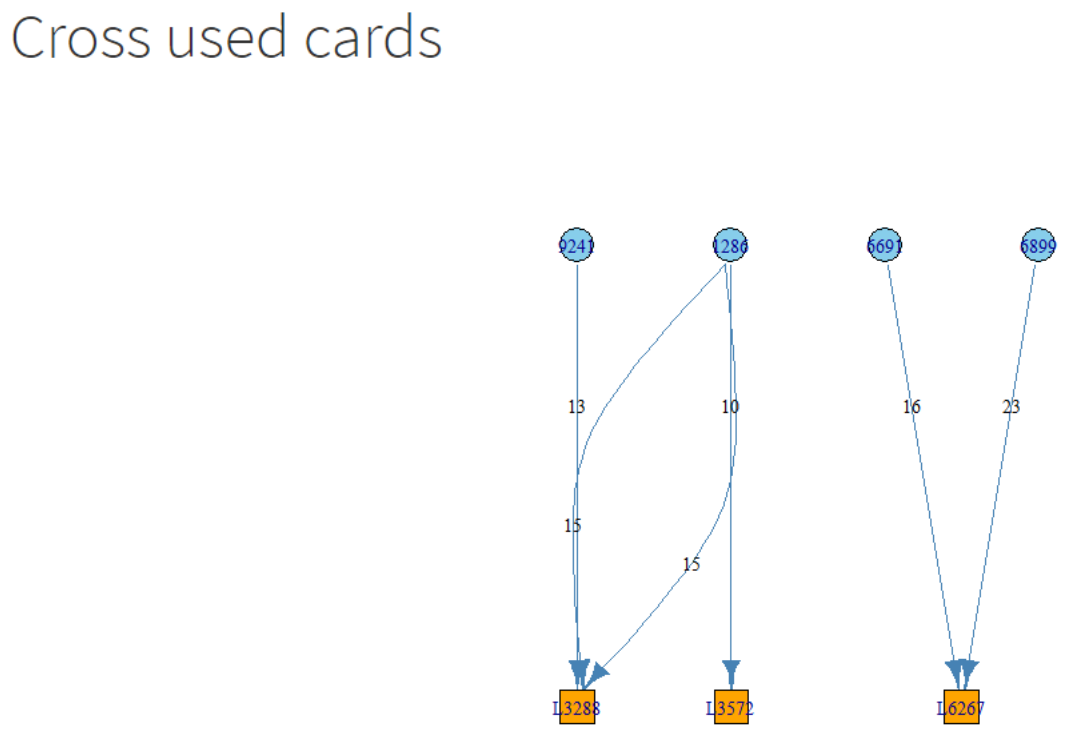


No.	Step	notes
1	Click “Show data table” to generate a table shown the suspectable transaction information, for example, date, time, price, location and card number, of each location.	

2	Results	Price Show 10 entries Search: date time price location last4ccnum 2014-01-12 03:00:00 147.3 Kronos Mart 5407 2014-01-18 03:13:00 87.66 Kronos Mart 3484 2014-01-11 03:39:00 277.26 Kronos Mart 8156 2014-01-18 03:45:00 194.51 Kronos Mart 9551 2014-01-18 03:48:00 150.36 Kronos Mart 8332 Showing 1 to 5 of 5 entries Previous 1 Next

2.4 Bipartite Plot

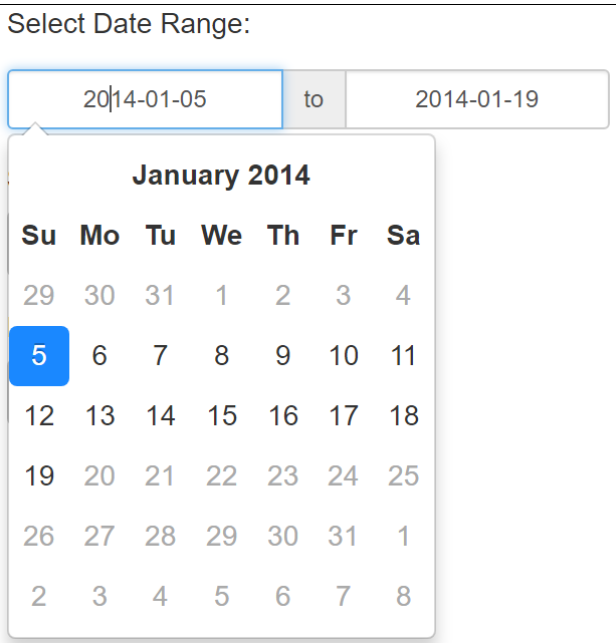
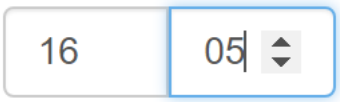
When modelling relations between two different classes of objects, bipartite graphs very often arise naturally. Hence, we used bipartite graph to check whether there is cross used between credit card and loyalty card.

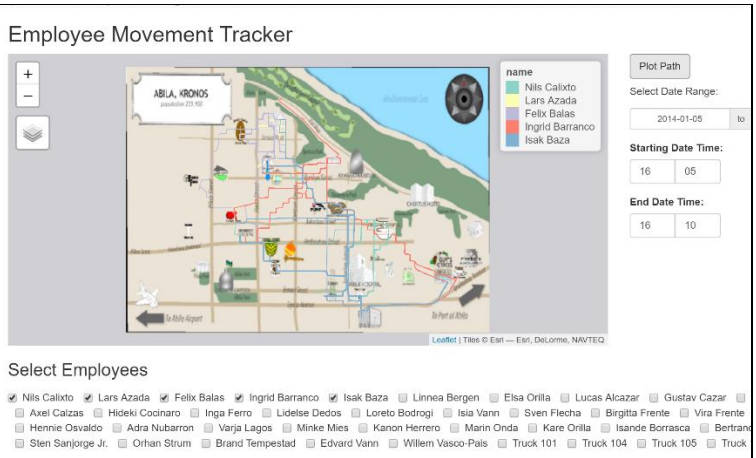


3. Geospatial Analysis

We used maps extensively to uncover all suspicious relationships and movements. In this view, movements data records will be represented as route.

No.	Step	notes

1	Select the desired date range, the dates available for plotting are colored black. Note : if planning to plot a path that does not change days, select the same date for both boxes	
2	Select the starting time by either clicking on the arrows for the hour and minute box or type the desired number, this input will determine the hour for the starting date. Note: when typing the time, keep in mind that minutes can only be selected by 5-minute intervals	
3	Do the same for the end time, this input will determine the hour for the end date	
	E.G. For the screenshots attached, it can be seen that the date range is between January 5th and January 19th while the start time and end time is 16:05 and 16:10,respectively. This means that the path that will be plotted will be travels done between January 5th 16:05 until January 19th 16:10.	
4	Select the employees whose gps data will be plotted by checking the box of their name	
5	Click on the “Plot Path” button on the top right	

6	The paths of the selected employees at the selected timeframe will be generated	

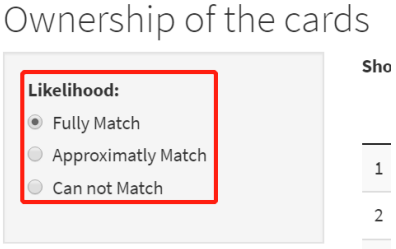
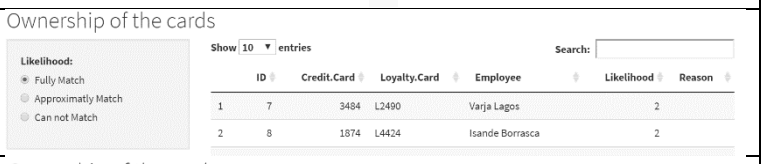
4. Card Ownership

This is an additional table for users to check the owner of the credit card and loyalty card. Column "Likelihood" shows the credibility of the information.

0 = Cannot give any confidence of identifying a specific employee as an owner of a card

1 = Positional data or time data may not be in-sync, could be person was in the vicinity or at the exact place but outside of the time window by a few minutes without anyone else being a possible answer.

2 = Cross-check of transaction data and positional data shows person was the only possible person to use this card at that time

No.	Step	notes
1	Click different button on the side bar will changed to different matching table	
2	There are three kinds of table. The first one shows the owner that we can fully match	
3	The second one shows the owner we can only approximately match because of some reasons we listed	
4	The third one shows owner we cannot match