Analysis of International Food Chain Taipei Stores Performance using Visual Analytics

Kelvin Chia Sen Wei, Linus Cheng Xin Wei, Wang Xuze

Abstract — The rapid growth of data in the context of business and the increasing importance of data analytics in business decisions provide us with a great chance to utilize visualization techniques to discover insightful business opportunities. The explosive growth of information collected by business provides a plethora of data in various forms such as text, numbers, geospatial etc. Good visualizations help to provide understanding, by dissecting the data into insights used to drive decisions. This paper aims to explain the ways and methodologies to utilize various visualization tools and technologies that the business could use on their data to generate actionable insights and possible business values. We demonstrate the potential of our work to be applied for business of different scales and with various forms of valuable data.

Index Terms — Visual analytics, Business intelligence, Geospatial information system, Web Application

1. INTRODUCTION

As the whole business landscape is expanding toward a more digitized and dynamic environment, there are more research development regards to navigating its data meaningfully. This is particularly true multinational companies extend their global reach, creating an urgent need for effective management. This raises a challenge for effective visual analytics technologies to be placed by providing an interactive application that is customizable for each specific use case. This paper presents a solution and research direction for visual analytics for business intelligence and aims to address the business visualization challenges of our research target client International Food Chain (IFC).

It consists of six sections. Section 1 provides a general introduction to the paper. Section 2 provides an overview of the motivation and objectives of our research effort. This is followed by a review of the existing visualization techniques for visualising and analysing point data. The application architecture and description are discussed in section 3. In section 4, we present our application usage for IFC as well as our sales analysis. Future direction of our research is highlighted in section 5. Lastly, the conclusion in section 6.

2. BACKGROUND

2.1 Motivations and objectives

Our research and development efforts were motivated by the general lack of effective and easy to use web-enabled visual analytics tool for discovering business opportunities and patterns in massive existing data. It aims to provide decision makers with an analytical tool enabling interaction in time and space scales. Specifically, the design and implement attempt to support the following analysis requirements:

1) To enable Business profiling of the company's store to determine Points-Of-Interests (POIs) that can generate insights such as: Highest earning stores, relative performance of stores, store's profile patterns and item sales information;

2) To create an easy and intuitive dynamic visualisation to make informed business decisions, such as determining locations for new store openings with matching POIs of high sales stores, in both macro and micro views;

3) To digitize each chain's trade and delivery area;

4) To build a scalable application which incorporates both historical and future data, as well as display detail records on-demand.

2.2 Data collection

The data used to develop this research is provided by IFC. These include Taiwan zoning areas, POI locations, individual IFC stores in Shapefiles (SHP), along with Trade Area and Daily delivery sales data in a flat file format. The data undergoes extraction of relevant POIs, digitization of polygons denoting trade areas, aggregation of all POI and sales data, as well as a combination of all existing data into a GeoPackage utilizing two software, QGIS (contains python console) and Microsoft Excel.

The research work and exploration focus on final data sets includes geospatial knowledge of IFC stores information, the POIs found in each store's trade areas, the delivery order's duration and the daily delivery sales.

2.3 Existing visualization tools

There has been work done on the visualization of business performance data through Off-the-shelf Enterprise software like Tableau and PowerBI. By and large, these solutions enable visualization in various ways and mostly customizable. Tableau enables businesses to connect to multiple data sources and easily blend the data, with dashboards showing both aggregated and individual data pointers. [1] However, this off-the-shelf software has limitations and lack the flexibility when it comes to raw file processing geospatial data formats like SHPs or GeoPackages. Moreover, when a dataset has too many entries or features, such visualization software become inadequate as it becomes too slow and memory intensive.

One way to overcome said difficulties is to work with scripting languages like R. R offers a huge support library that allows for easy manipulation of SHPs, as well as allowing direct access to the contents of GeoPackages.

3. APPLICATION

3.1 Application Architecture
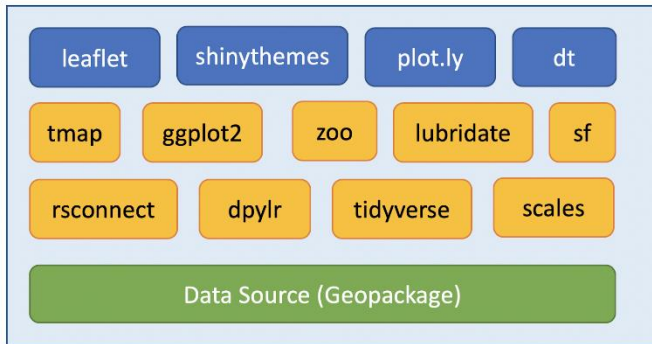
3.1.1 R Shiny Framework

Fig. 1. R Shiny Framework

The IFC Application is fully developed in R, utilizing the Shiny framework. Users of the application can interactively navigate through data, utilizing packages highlighted in blue in Fig.1. The data that is displayed in our application is prepared on the backend, likewise highlighted in yellow. Our raw data source for the application has been taken from our sister SMT 201 project, in the form of a GeoPackage which is directly loaded into R.

3.2 Application Description

We used a complementary colour scheme to make our application feel fresh through the contrast of different colours, while ensuring that the overall contrast is not excessively jarring. The colours utilized are orange, blue, black and white of varying shades. The dashboards also utilise the 12-Column Grid System to ensure a sense of uniformity across various display resolutions.

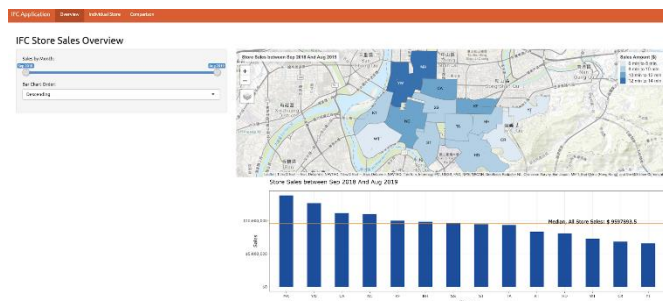#### 3.2.1 IFC Store Sales Overview (Main Dashboard)

Fig. 2. Main Dashboard

The map view of the main dashboard is implemented using the R Leaflet package on CRAN and users will be able to toggle between map layers using the layers filter on the left of the map. The map also limits the zoom to prevent users from heading out of bounds, as well as having each store's name displayed for their respective trade area.

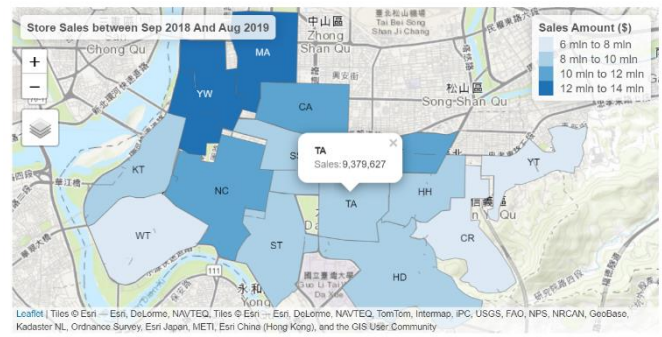The zoom limit for each map is limited between value of 13 and 15 with the default at 14.

Fig. 3. Map View of Main Dashboard

Clicking on the store's area displays the sales value for the respective store through a tooltip.
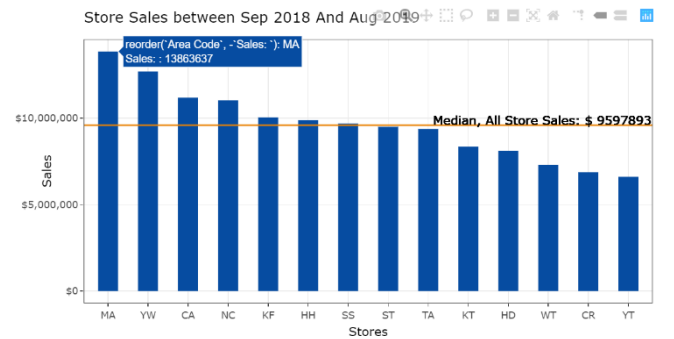

Fig. 4. Line Chart of Main Dashboard

The bar chart shows the top-performing stores along with the median line to gauge each stores' performance. When hovered, it displays the sales amount for the specific store.
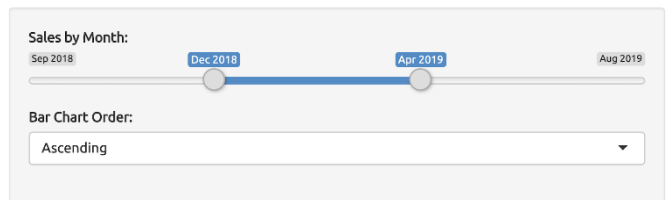

Fig. 5. Left Side Panel of Main Dashboard

On the left side panel, user can select a period they would like to view, and both the map and bar chart will update accordingly to said period. This panel also contains a sorting option for the bar chart.
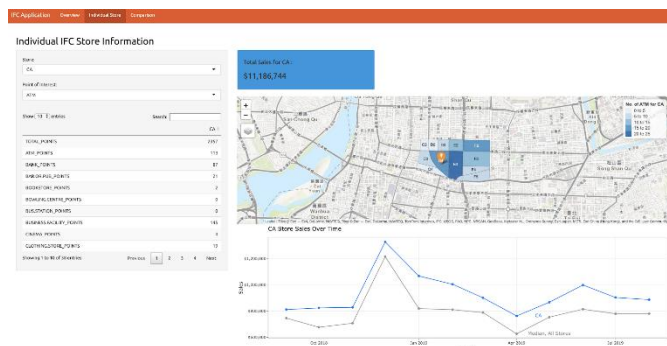
3.2.2 Individual IFC Store Information

Fig. 6. Individual Store View

This dashboard contains a map view, line chart and a data table to provide user with a detailed view for each store. In the map view, the store is labelled with an orange marker along with the number of selected POI for each subzone.

Fig. 7. Map View of Individual Store View

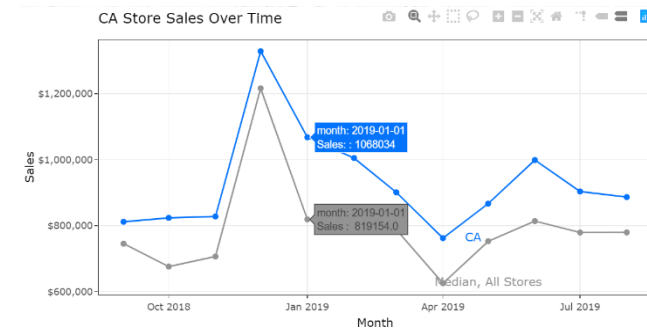When a subzone is clicked on the map, the number of POIs will be displayed above the selected point.

Fig. 8. Line Chart of Individual Store View

With the aid of the grey median line in the line chart, users can better analyse the performance of the chosen store with minimal visual disruption. The line chart is also interactive and hovering over the points bring up the date and sales of the respective stores. As the line chart is built upon Plotly, there are several options that enhance the interactivity of the chart.
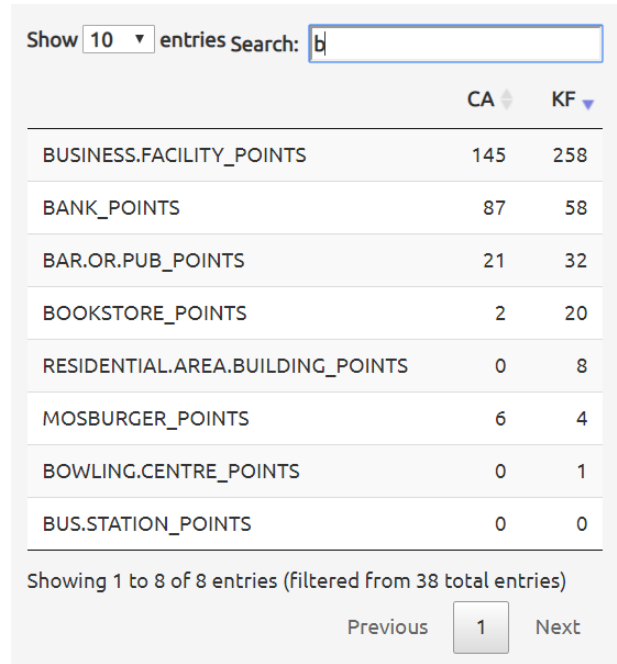

Fig. 9. Data Table of Individual Store View

The data table on the left side panel provides a detailed tabulation of the number of POIs for the chosen store. Users will be able to view any POI information using the search bar to filter unwanted data. The pagination allows the page to be viewable without scrolling.

The user can choose the store and point of interest from the left side panel to filter its selections. When a store is selected, the map view will automatically shift its focus to the map box of the selected store. The line chart and data table will also filter its data to reflect the selected information.

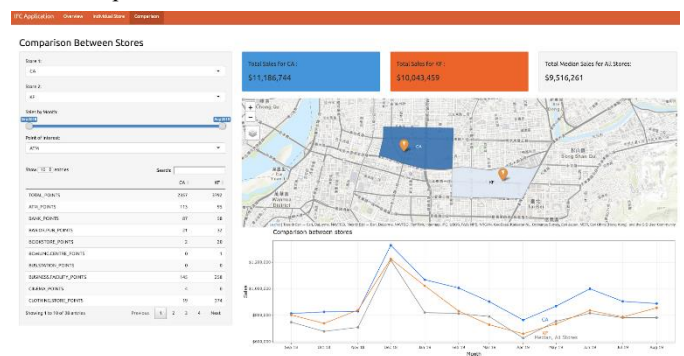### 3.2.3 Comparison Between Stores

Fig. 10. Comparison between stores View

This dashboard presents comparison between two stores, with a similar view as above.

The user can choose two stores and a single POI from the left side panel for comparison. When the two stores are selected, the map view will automatically scale to suit these two stores. The line chart and data table will also filter its data to reflect information about the selected stores.

Fig. 11. Map View of Comparison View

The map view now shows the two store trade areas and the aggregated number of selected POIs is displayed when clicked.
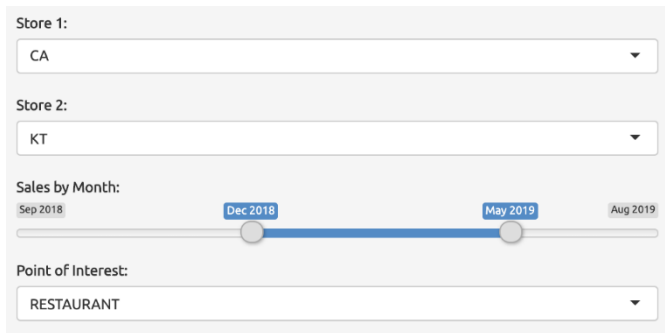

Fig. 12. Left Side Panel of Comparison View

The slide bar allows users to choose the duration for sales analysis. The sales comparison chart is automatically updated according to the selected time range. This provides the business owner with more focused analysis.

The chart includes both stores sales, as well as the median sales of all stores.

4 CASE STUDY: IFC VISUAL ANALYTICS AND BUSINESS INTELLIGENCE

4.1 Scenario

IFC has over 18000 stores worldwide and an ever-growing presence in the global market. The data we used in our research implement include geospatial knowledge of different stores in Taipei and the Point of Interests (POIs) around the area, the delivery time and duration of each store, the delivery sale of each store.

One use-case of the application is utilizing visualization to determine the performance of the store. The number of POIs around a store and the past sales performance could potentially enable us to understand how POIs affect the sales. We have decided to make use of the data collectively and look for correlations among them to discover insights, as following.

Understanding the information around the delivery area and POIs can help the business to evaluate the performance.
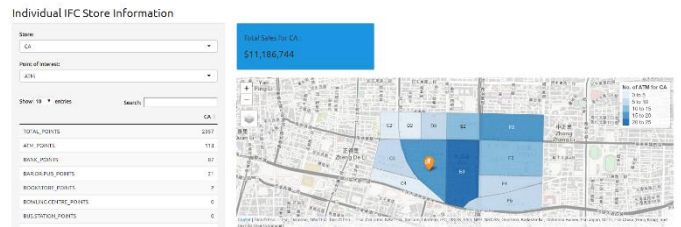

Fig. 13. Individual Store View

A comparative chart visualization of the selected store's sales performance with the median of all stores provides business owner with macro level metrics of the store. This result could be used together with the POIs visualization to evaluate the standing of each store.
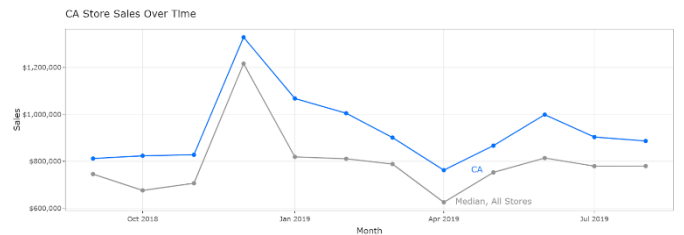

Fig. 14. Line Chart of Individual Store View

A comparative visualization dashboard for two stores provides a micro level analysis of the stores. This allows a direct comparison of both sales and number of POIs between two stores. The time range of the sales could be changed, allowing customization.
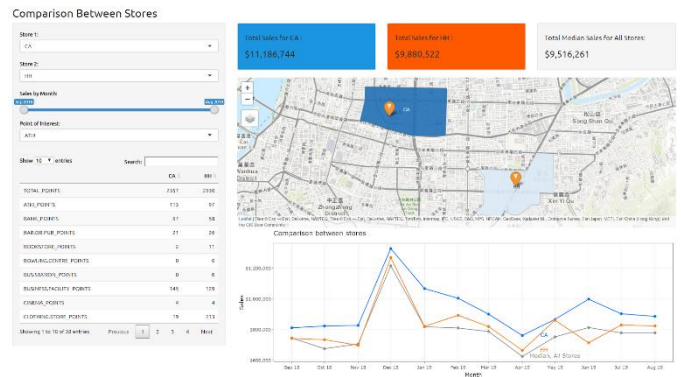

Fig. 15. Comparison between store View

4.2 Results: Sales Analysis

Based on the figure above, during the full period ranging from Sep 2018 to Aug 2019, IFC stores located in Taiwan range from USD$6,809,445 - $13,863,637, with the YT store having the lowest sales, and the MA store having the highest sales. When comparing all stores, the median sales is store SS, with USD$9597893.5. The results are particularly interesting, as both the MA and YT stores are located in popular tourist districts (Zhongshan and Xinyi respectively), although they are of both extremes when comparing sales.
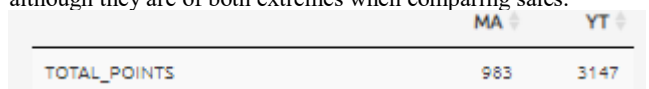
	MA	YT
TOTAL_POINTS	983	3147

Fig. 16. Total Points of Data Table

Looking at the total number of POIs, YT has considerably more when compared to MA. If one were to assume that the presence of different POIs would affect total sales, this could be interpreted as YT does not have the presence of significant POIs, and instead have a huge number of irrelevant POIs in its area.

Fig. 17. School search in Data Table

However, upon further investigation, SCHOOL_POINTS are the only POIs that are more present in MA. Based on pure numeric quantity, this could mean that POIs have little to no correlation with sales, and that other factors affect sales more directly.

Fig. 18. Map View of Comparison View

One possible reason to explain this phenomenon is due to the fact store YT is in a mountainous region, as compared to MA, which is in the middle of the city. By this factor, there might be a lower population density in YT as compared to MA, thus leading to less sales by attribute of having less population.
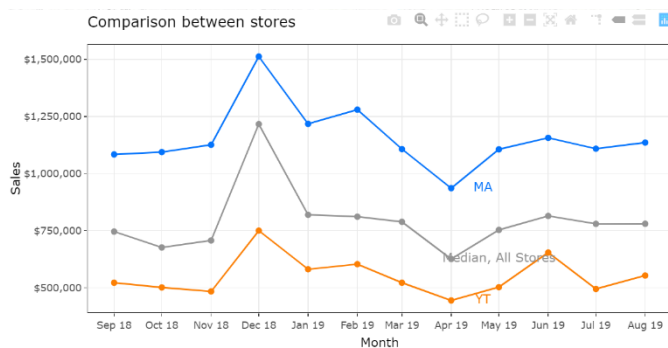

Fig. 19. Line Chart of Comparison View

Another interesting pattern appears when we look at the sales over time. Most stores have their peak sales at Dec 18, and their bottom sales during Apr 19. Some possible reasons to explain the peak sales include having exclusive seasonal items that are available for a short period. As December is a common holiday period for both adults and those in school, this leads to more sales because fast-food chains are usually opened during the holidays as compared to other mom and pop stores, thus increasing overall sales.

However, the dip in sales for the April period is highly unusual and requires more context in order to explain the phenomenon as there are no major events or weather conditions that prevent deliveries.

5 FUTURE WORK

5.1 Integration of other analysis from SMT201 Project

This application was developed in tandem with another project from the module SMT201: Geographic Information Systems for Urban Planning, involving the same client, IFC. Various analysis was performed, such as population density analysis, shortest-path analysis, nearest competitor analysis and regression analysis. However, as the scope of the application focuses mainly on sales data, we have yet to fully integrate these features. We believe these additional analyses serve as a great compliment to our current application, by providing more information and context that attempt to further explain sales patterns.

5.2 Interactive On-the-fly Regression Analysis

In the context of our project, a use for regression analysis would be to identify variables that affect sales of a store. We hope to expand upon our application by running regression whenever the user selects any store or subzone. Through the regression model, the important and significant variables for that subset will be generated then displayed, allowing the user to peak at potentially useful information.

5.3 Visualization of individual coordinates of POIs

Although our application has a table that allow users to filter for specific POIs as well as being able to see the numerical counts, these POIs are not visualized on the geographical maps provided by "Individual Store" and "Comparison" page. In a further iteration, our group hopes to integrate the visualization of each POI geographically, allowing for a more comprehensive application.

6 CONCLUSIONS

As we have presented through our research, visual analytics can help facilitate significant business decisions, especially when it comes to comparing sales. Open-source R-shiny framework enables easy implementation and deployment to provide concrete understanding of business problems. Our vision of the future research direction includes further development of visual analytics with an emphasis on business usage and generation of insights.

REFERENCES

[1] Tableau on Tableau: 5 ways we look at our sales data
https://www.tableau.com/about/blog/2015/9/top-metrics-every-sales-executive-cant-live-without-43746