Visualizing Korean Tourism

Kim Do Yeon, Lee Hyeonjeong and Huiyeon Kim

Abstract – The tourism industry in South Korea has always been of great importance to the Korean economy. Under the Ministry of Culture and Tourism of South Korea, Korean Tourism Organisation (KTO) currently oversees and promotes Korea as a vibrant tourism destination. In order to address the lack of an in-depth analysis platform, this research project aims to enhance the current implementations of KTO through an interactive and comprehensive visualization that provides insight on seasonal arrival patterns, trends in tourism attraction visits, as well as demographics and purpose in visit. Other than sourcing for relevant datasets, substantial data preparation, planning dashboards and implementation of visualization has been done. Key features include cycle plots, bar charts and geomap. With these in hand, KTO can derive various insights in order to identify areas that require further attention and investment to achieve its planned goals in the future.

Keywords — South Korea, Tourism, Visitors, Arrival Patterns, Demographics

I. INTRODUCTION

The Korean Tourism Industry has been rapidly growing since over ten years ago, with 12 million international visitors having reached in 2013, and over 17 million having reached in 2017[1][1]. Furthermore, it has accounted for around 4.7% GDP in 2017. The Korea Tourism Organization (KTO), under the Ministry of Culture and Tourism in South Korea currently works to develop the tourism sector as the growth engine for an enhanced national economy and quality of life through various strategic assignments. As mentioned in its *Goals by 2022*, KTO aims to generate 25 billion USD to boost the national economy and attract over 340 millions of visitors to Korean tourist attractions. In order to achieve this, a better understanding of Korean inbound tourism data in the form of seasonal arrival

patterns, trends in various tourist attractions across Korea, and demographic insight is strongly needed.

Currently, KTO provides detailed text reports per annum as well as datasets on yearly arrivals by nationality and purpose of visit. However, the usage of such data and information is limited to simple descriptive analytics such as understanding the growth in South Korean visitor arrivals and basic demographics. Hence, we present NO.G.A.D.A – a dynamic visualization platform specially designed for users.

Consisting of nine main sections, this research paper documents detailed steps taken to propose, design and implement the interactive visualization that aids in KTO's analysis of tourists' arrival patterns and demographic information. Section I, II provides an introduction, motivation and objectives that will be addressed in the visuals. This is followed by an overview of background works or related surveys that has aided us in proposing our design. Section IV and V details our design considerations and data exploration required to implement the visualization. Section VI, a walkthrough of the application is documented. In section VII, this research paper addresses the main key findings of our research. Lastly, in sections VIII and IX, the paper highlights the main technologies used for the application and concludes with possible factors to consider in the future direction of this project.

II. MOTIVATION & OBJECTIVES

This research is motivated by a lack of web-based visualization provided to analyse detailed inbound arrival trends and visitor profiles. Moving forward from text reports and static info-graphs, this project aims to provide KTO with a comprehensive idea of how to allocate resources and investments depending on tourist sites' popularity and visitors' profiles, with the goal to drive tourism demand in the future. Specifically, it seeks to address the following:

Kim Do Yeon is an undergraduate in School of Information Systems, Singapore Management University (email: doyeon.kim.2016@sis.smu.edu.sg)

Lee Hyeonjeong is an undergraduate in Lee Kong Chain School of Business, Singapore Management University (email: hj.lee.2016@business.smu.edu.sg)

Huiyeon Kim is an undergraduate in School of Information Systems, Singapore Management University (email: Huiyeon.kim.2016@sis.smu.edu.sg)

[1] https://www.seoulsolution.kr/en/node/6558

1) To be able to understand seasonal and monthly visitors' arrival patterns by nationality;
2) To provide an extensive visitors' profile dashboard with features such as gender, age group and purpose of arrival;
3) And provide a map visualization framework with tourist attraction sites and trends by foreigners and locals.

III. RELATED WORKS

Previously, we can see background surveys of related works with similar or identical datasets with the aim to visualize South Korea's inbound tourism. Referring to *Figure 1*, we can identify that there are relevant monthly trends that should not be ignored when performing arrival analysis. However, this chart merely displays 2016 data; the insights given is insufficient when we aim to see the *trend* in monthly arrivals.

Figure 1: Monthly Arrival Patterns in 2016

One way to overcome such limitation is to provide both yearly and monthly filters so that it is possible to understand patterns throughout the years.

Taking a look at the side-by-side bar chart as per *Figure 2* below, we can identify foreign tourists visiting Korea's various tourist attraction sites by year (2003 – 2015). Providing clear legends next to the graph and having data labels next to the bars effectively display both the popularity of attraction sites and the visits to each site by year, which can also show the trends over the three years.

However, if there were over ten tourist attractions to consider, it may be difficult to effectively compare over the years. To improve this model, it would be useful to provide a geomap where users can see information in one glance.

Figure 2: Arrivals to Korea's Tourist Attraction by Year (2013 - 15)

As seen in *Figure 3* below, an extensive info-graphic on most visited tourist attractions is provided in a map feature. Compared to *Figure 2*, potential patterns by region location can be spotted, meaning that it is useful to provide a map framework to understand which region is most visited by tourists.

Figure 3- Most Visited Tourist Attractions in 2015

However, when a "time" attribute is involved, there is a limitation to see such changes as this info-graphic is static. Furthermore, there should be a better overview before drilling down to the exact data labels, such as a chloropeth feature that can immediately let users understand arrivals by region. Overall, these visualizations in no doubt serve to aid in our brainstorming of design and a firm understanding of current measures. We aim to take a step further to improve on such models given with the ability to analyse comprehensively based on highlights, filters and effective use of tooltips. With inspirations gained from these related works, we attempt to apply the Schneiderman's mantra by

providing the *overview, followed by a zoom in and filter, then details on demand[2]*[2].

III. VISUALIZATION APPROACH

Followed by problem identification and addressing objectives, three main steps were taken to develop the visualization: 1) data exploration and preparation, 2) brainstorming and prototyping the dashboards, and 3) implementing the visualization.

IV. DATA EXPLORATION & PREPARATION

It is necessary to gather representative and meaningful data to aid in KTO's actions to enhance tourism demand and quality experience. Data has been gathered from official websites such as the Korean Tourism Organization (KTO) website itself (http://kto.visitkorea.or.kr/eng.kto), as well as the Korean Tourism Statistics Board website (http://know.tour.go.kr).

The datasets were selected based on representative data as per announced in the KTO website. The key statistics in KTO displays that 11 million foreign tourists reached Korea in 2012, and 17 million reached Korea in 2016[3][3]. For brief testing and validation, Tableau was used to prove that the sum of visitors in 2012 and 2016 aligns with the reports on the official website.

Furthermore, KTO and the Korean Tourism Statistics Board provide visitors' arrival information ranging from years back in 1975 to recent years and months in 2018. For the purpose of this project, the selected focus time range was chosen to be 2008 – 2018. The reasoning behind this range is that it not only is free-of-error and has the most comprehensive data in both years and months but also serves as a meaningful range to identify relevant trends.

V. DESIGN CONSIDERATIONS

For design, we researched various visualization ideas and technologies through websites such as D3.js and other external databases. Considerations taken when planning the storyboard were clarity of visuals, interactivity that can be enhanced through

appropriate use of technological library and aesthetics for the users' perspective.

In the start of the proposition, sketching on pen or whiteboard was done as a team to briefly lay out the visualization design. After which, such ideas were further inspired and improved based on feedback from Prof. Kam Tin Seong and a revision in our objectives. Throughout the data exploration process, however, the team members were able to suggest innovative approaches that could be taken to further improve the model. Overall, the steps taken to design was an iterative process.

A. Cycle Plots
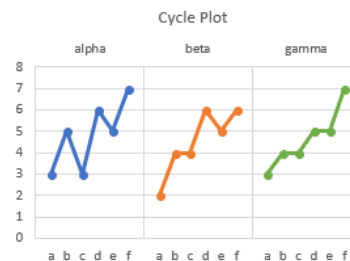
The cycle plot framework is a unique and important feature that displays the cycle or trend and the "month-of-the-year" aspect, first introduced by Cleveland, Dunn and Terpenning in 1978[4][4]. This would be used to analyse seasonal patterns of visitor arrivals. There will be reference line for each pane to show the average value, which will be easier to compare among months when analysing high and low arrival months.

B. Map Visualization

The regional map visualization enables us to display geographical locations of major tourist attraction sites per region and convey the arrival in such sites by local/ foreigner attribute. With an overview of the arrivals of visitors marked by intensity of colour, users can zoom in to particular regions' attraction sites and explore detailed trends in arrivals.
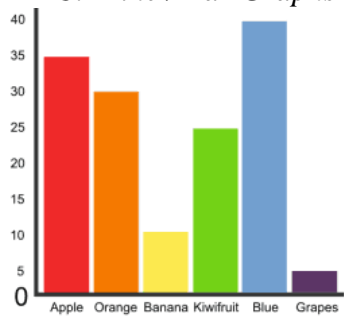
---

[2] http://www.ifp.illinois.edu/nabhcs/abstracts/shneiderman.html
[3] https://kto.visitkorea.or.kr/eng/overview/About/history.kto

[4]
https://www.perceptualedge.com/articles/guests/intro_to_cycle_plots.pdf

C. Line / Bar Graphs

The bar chart and the line chart may be of basic display, however, it is considered one of the most informative charts to exist. With that in mind, these charts will be used to create informative charts on Yearly trend, categorical analysis, etc.

VI. DATA VISUALIZATION WALKTHROUGH

For a convenient platform to view all the designed visualization, a website has been created as seen below.

A. View of Home Page

Figure 8- No.G.A.D.A Landing page

The design of the web page was inspired by dynamic South Korean tourist attractions as a background. The use of a dynamic background is rationalized by having users become exposed to the various attractions in South Korea. Scrolling down further, users can see the problem & motivation and objectives of this project.

B. Storyboard of Visitor Arrivals
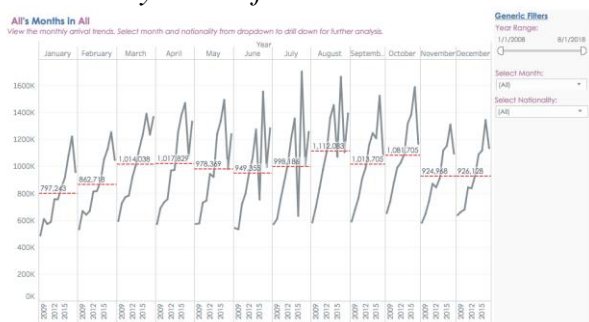
Figure 9- Visitor Arrivals to South Korea (2008 - 2018)

The Visitor Arrivals Storyboard consists of a yearly time-series graph, bar chart, and cycle plot as seen in *Figure 4*. Users can select year, month, and nationality as a generic filter on the right to explore in-depth analysis. Overall, the storyboard consists

of 4 dashboards, starting from a general trend in South Korean tourism, followed by a drilldown such as by the trend in arrivals by nationality, seasonal, and monthly arrivals.

C. Demographics and Purpose of Visit
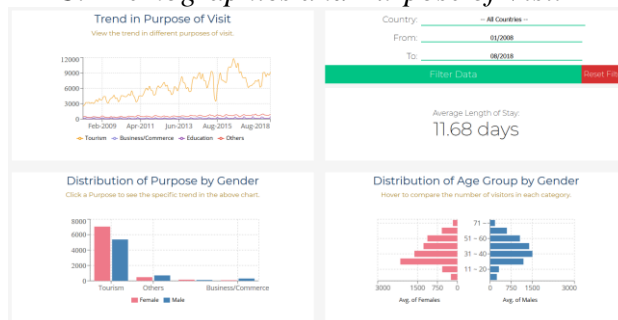
Figure 10- No.G.A.D.A Demographic page

The Demographics and Purpose of Visit dashboard, as seen in *Figure10*, visualizes the visitors' profile consisting of gender, age group, length of stay and purpose of visit. Users can select the country and time range in the top right filter for further analysis. The line chart enables users to view the trend in purpose of visit, and the side-by-side bar chart and age-sex pyramid chart helps to understand patterns by gender.

D. Arrivals to Visitor Attractions

This dashboard shows choropleth and the line graph of local and foreign tourists in tourists' attraction. The line graph is interactive with the map which means the line graph can be filtered by the province and tourists' attraction.
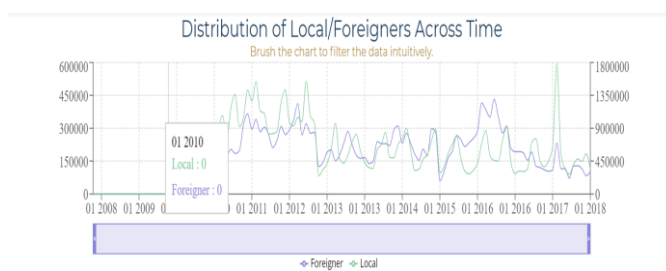

Figure 11- Distribution of Local/Foreigners

Here, the user will be able to change the timeline flexibly by brushing the timeline below the graph. Thus, the line graph can be shown in both yearly and monthly.

VII. KEY FINDINGS & OBSERVATIONS

Through this visualization, certain key findings and insights were obtained as follows:

1. Objective 1: To be able to understand seasonal and monthly visitors' arrival patterns by nationality

Overall, there is a generally increasing trend in the international visitor arrivals to South Korea from 2008 to 2018. On average, the top 10 countries are (in descending order) China, Japan, U.S.A, Taiwan, Hong Kong, Philippines, Thailand, Others, Malaysia, and Russia, with China and Japan accounting for most of the total number of visitors each time period. Although Japan has once been the top arrival country from 2008 to 2012, China currently stands to be the top arrival country. Given that these countries have remained rather stable over the ten years, KTO should continue allocating investments for these countries, with increased marketing efforts and innovative tourism programs to further drive up the demand

In terms of seasonal arrivals, in general summer tends to be the season for most arrivals, with winter as the lowest. August is the month with the highest number of visitor arrivals and January is the lowest.

2. Objective 2: To better understand demographics of tourists in terms of their purpose, gender and age.

This dashboard displays four demographic information.

I)   Trend purpose of visit
II)   Distribution of Purpose by gender
III)   Distribution of Age group by gender
IV)   Average Length of stay

Firstly, there are 4 purposes of the visit. Tourism, Education, Business and others. Tourism is the most prominent factor of all.
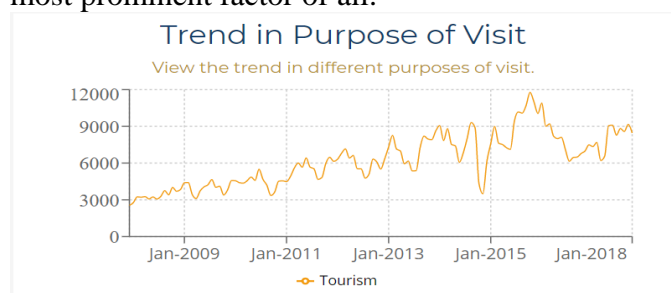

Figure 12- Trend in purpose of visit(Tourism)

According to the graph shown above, there has been general increase in tourism sector in korea from 2008 to 2018. However, there was steep decrease in both jan 2015 and in between 2016 and 2017.
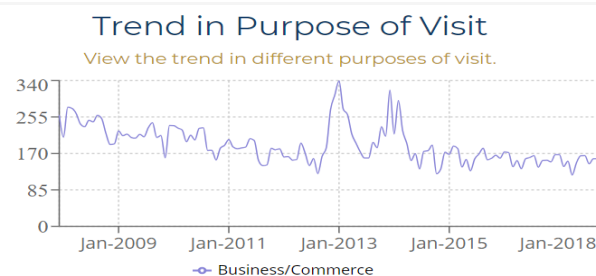

Figure 13- Trend in purpose of visit(Business)

For Business, the number of people visiting has been generally decreased from 2008 to 2018.There was a steep increase in between 2013 to 2014. However, it has been consistent since 2014 till 2018.
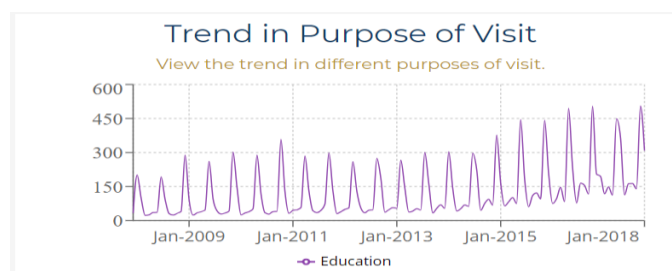

Figure 14- Trend in purpose of visit(Education)

Education displays an interesting pattern compared to other purposes. For education sector, there is increase during the holiday period(May-August, December), while there is a steep decrease after the holidays.From here, it can be seen that foreigners visit korea during holiday for exchange programme and Summer/Winter school.
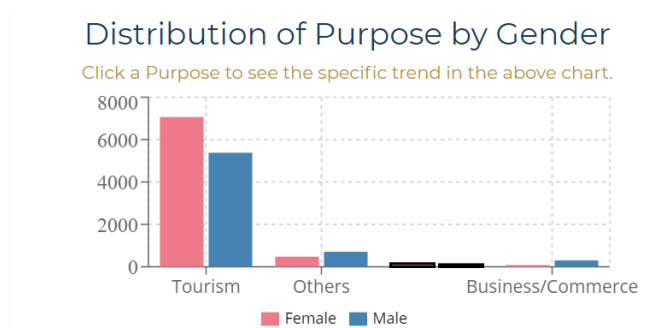

Figure 15- Distribution of Purpose by gender

Demographics information is also shown by comparing purpose by gender. In general, more female has visited korea for tour and education. More male has visited for business purposes.

Distribution of Age Group by Gender

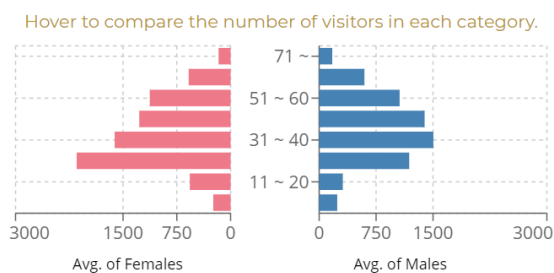Hover to compare the number of visitors in each category.

Figure 16- Distribution of Age group by gender

The dashboard also shows distribution of age group by gender. There are more female tourists coming from age 10 to 30 while there are more male tourists coming for those who are aged in between 41 to 50. In general, there are more female tourists coming to korea.

Not only the features above, we also show average length of stay. In general people have stayed 11.68 days in korea.

3. Objective 3: To better understand the local and foreign tourists' distribution on tourist attractions.

From this dashboard, popular tourists' sites can be spotted. Firstly, choropleth shows the overall tourists who came to the region. The darker the colour, more people visited the area. From the map, it can be seen that Seoul, Gyeonggi, South Chungcheong and South Jeolla is the most popular tourists province in Korea. People have visited to the region that is nearer to the capital of the country.

Figure 17- Choropleth of Korea by Region

Then the tourists' attractions will be shown after the user clicked the province. The user will be able to click the tourists' attraction and see the number of local visitors and foreigners from 2008 to 2018.Thus, number of local visitors and foreign visitors can be compared easily.

VIII. DATA VISUALISATION ELEMENTS

The data visualization elements such as line graph, bar graph, choropleth is created using Tableau, D3.js a JavaScript library. Re-charts and Leaflet library were used to show the data visualization. For the back-end Node/Express.js was used for the server with PostgreSQL for our database. React was used for our front end.
Tableau has an impressive integration with most of the data type and offers integration with other myriad platforms. Not only React but it also integrates with R as well. Re-chart has interactive visualization with smooth transition of the graph using filtering and brushing function. Leaflet shows a great interactivity of map and the data. It uses GeoJSON data where the locations are well shown. Its' marker and highlighting feature makes map more interactive to users.

IX. CONCLUSION

The data visualization in NOGADA enables users to explore and analyse seasonal trend, demographic information and popular tourists' attraction. NOGADA encompasses a potential to be used for other countries' data as an input.

For future research our interactive visualization can explore many potential areas such as

i) Consumption expenditure by nationality/region
ii) Earnings of different tourist attractions
iii) Prediction model for visitor arrivals in the future

Not only that the current feature can be further improved in terms of user interfaces. These features can be selected and adopted in the future.

X. REFERENCES

[1] Kiyong, K. (2018, February 20). Characteristics and Prospects of South Korea's Inbound Tourism Markets. Retrieved from https://www.seoulsolution.kr/en/node/6558

[2] Shneiderman, B. (1997, November 6). A THOUSAND-FOLD INCREASE IN HUMAN CAPABILITIES. Retrieved from http://www.ifp.illinois.edu/nabhcs/abstracts/shneiderman.html

[3]https://kto.visitkorea.or.kr/eng/overview/About/history.kto

[4] Robbins, N. B., Ph.D. (2008, January 1). Introduction to Cycle Plots. Retrieved from https://www.perceptualedge.com/articles/guests/intro_to_cycle_plots.pdf