Recruit Restaurant Visitor Forecasting

Project Goals

28th January, 2018

Foodie Analytics

Shipra Sethi
Dong Bing
Daniel Colvin
Subba Muthurangan
Erik Platt

Engagement Letter

Foodie Analytics
401 N. Michigan Ave.
Chicago, IL 60611

January 28, 2018

Dr. Donald Wedding, CEO
Recruit Holdings Co., Ltd.
Chuo-ku, Tokyo 104-0061 Japan

Dear Dr. Wedding,

We at Foodie Analytics would like to thank you for giving us the opportunity to bring in our 15+ years of analytics experience in the hospitality industry and build predictive models that will serve your needs. We understand running a thriving local restaurant is not as charming and straightforward as it appears at first. To run a successful restaurant, it is important to know how many customers to expect each day to effectively plan the purchase of ingredients and schedule the restaurant staff. Our expert team of data scientists will utilize the reservation and customer visitation data that your company owns to not only predict the future customer visits to a restaurant, but also define the competitive environment for an existing restaurant, and the most optimal location for a new restaurant. The product will also provide the flexibility to utilize the same predictive model for other players in the industry such as hotels, casinos, and bars, if similar data becomes available.

Our team brings broad experience in predictive modeling, data visualization and project management. We plan to evaluate different programming solutions and data visualization tools and choose the ones that best fit the needs of this project. At the end, we will deliver a final time series forecasting model, data visualizations based on our modeling solution, and provide our recommendation on the best tools to use for future modeling needs. As needed, we can also provide training on these software packages and tools to your company's analytical team.

The following report outlines the project goals, our approach to accomplish these goals, and preliminary insights into the data. This will be a ten-week engagement where we will provide a presentation on initial findings by the end of week 6 and a final presentation at the end of week 10. Please let us know at your earliest convenience if you have any concerns about the approach and/or the project deliverables.

Sincerely,
Foodie Analytics

Table of Contents

Introduction

Recruit Holdings has unique access to key data sources on the restaurant industry in Japan. Specifically, Recruit Holdings owns Hot Pepper Gourmet (a restaurant review service), AirREGI (a restaurant point of sales service), and Restaurant Board (reservation log management software). These data sources track reservation and visitation data for any participating restaurant and has amassed a significant amount of historical data over time. If properly mined, this data would not only offer insights on the number of customers to expect for any existing or new restaurant owner, but also provide demographic customer flow on the restaurant industry as a whole.

To take advantage of this wealth of data, Recruit Holdings has appointed Foodie Analytics, a leading data consulting firm headquartered in Chicago, IL to develop a forecasting model. The primary goal of the model is to make automated predictions on the number of visitors to a restaurant for a future date. By knowing how many visitors to expect in advance, the restaurant owner will:

1. Reduce operation uncertainty and ensure more efficient resource allocation
2. Cut costs by procuring the proper amount of ingredients and adequate staffing
3. Facilitate growth by identifying the most optimal area to expand
4. Better understand the impact of visitation seasonality and holidays on sales
5. Gain competitive intelligence through benchmarking

Project Goals

Our project seeks to improve restaurant operational efficiency by answering the following questions:

1. Customer traffic flow: How many customers are expected to visit a restaurant on any given day? What factors (such as visitation seasonality and holidays) impact number of visitors?
2. Location: Where should the next restaurant be opened to generate the highest amount of foot traffic? Which restaurants should be scaled back given the lack of demand?
3. Competition: What's the competitive landscape for any existing or new restaurant?
4. Preference: What type of restaurants are most popular? Is there any trend indicating shifts in customer dining preference?
5. Delivery Platform: How can data insights be delivered on the go?

Proposed Solution

Foodie Analytics will deliver data insight through a predictive modeling solution. Our analysts will identify the best performing forecasting model to predict the number of visitors at any given time. Through this model we will be able to provide following recommendations:

1. Specific restaurant predictions will provide direct insight into foot traffic and fluctuations in visitation over time
2. Predictions based on city name will provide insight into the local competition as well as identifying the most popular dining locations
3. Predictions based on restaurant genre will allow owners to see how their restaurant is performing (in terms of visitation) compared to other similar restaurants

Our Approach

The team will utilize R as the primary analytic tool for Exploratory Data Analysis (EDA) and model development. R is an excellent tool for EDA as it is well supported and documented, produces powerful visualizations, and is open source. Other tools under consideration include SAS, Python, and Tableau. We will conduct our analysis for the following purposes:

1. Maximize insight into the data sets
2. Uncover underlying structure
3. Extract important variables
4. Detect outliers and anomalies
5. Test underlying assumptions
6. Develop parsimonious models
7. Determine optimal factor settings

Once we validate our assumptions and identify patterns that can better inform our understanding of the task at hand, Foodie Analytics will develop and deploy a forecasting model utilizing a variety of different techniques. These techniques may include ARIMA, XGBoost, Holt-Winters seasonal method, Neural Networks, or a hybrid model that utilizes external regressors. Ultimately, the model we select will be the one that returns the highest accuracy in the validation dataset for deployment.

The primary delivery channel will be a web-based dashboard, providing a thorough understanding of how to transform our data insights into actionable decisions. Additionally, Foodie Analytics has found much utility in mobile platforms. Ensuring that clients can access data while away from their desk has proven to be quite valuable. As such, an interface will be developed for on-the-go analytics.

Data Overview

The primary datasets that will be used for insight generation and modeling purposes come from two separate data sources:

- Hot Pepper Gourmet (HPG): Similar to Yelp, users can search restaurants and make a reservation online. This data source contains 1 file for visitors and reservation data.
- AirREGI / Restaurant Board (Air): Similar to Square, a reservation control and cash register system. This data source contains 2 separate files for visitors and reservation data.

Both Air and HPG data sources provide nearly one and a half years worth of reservation data, spanning from January 2016 until April 2017. Each platform tracks the Store ID (which is the proxy for the particular restaurant), the time a reservation was created, the time that the reservation was created for, and the number of guests on the reservation. Additionally, the data sources provide a directory of all concerned restaurants, translating the Store ID into hierarchical descriptors such as Genre Name (what type of food is served at the restaurant), Area Name (location of the restaurant), and the latitude and longitude coordinates. Store ID has been anonymized and generalized to prevent any tampering with the efficacy of the model.

Since the Air platform is a cash register system as well as a reservation system, the Air datasource also provides visits--tracking the number of clients to visit a particular restaurant on any given day through the period. Using the Store ID Relation dataset, restaurants appearing in

both datasets can be mapped to one another. However, it is important to note that a HPG store may not necessarily appear in the Air dataset, and vice versa. Lastly, a calendar dataset is provided marking which particular days are celebrated as a holiday in Japan. Ultimately, Foodie Analytics will use these datasets to forecast future restaurant visitor totals for a given date.

Preliminary Findings

Foodie Analytics began the analysis by examining visitor patterns and restaurant genres in Japan from AirREGI (Air) and Hot Pepper Gourmet (HPG) data sources.

- 252,108 observations and 3 variables found in Air visitors data
- 92,378 observations and 3 variables found in Air reserve data
- 2,000,320 observations and 4 variables found in HPG visitors and reserve data
- Preliminary analysis took into account only the following major cities for genre, visitor trends, reservation and holiday flag:

System	City
Air	Tokyo
Air	Fukuoka
Air	Osaka
HPG	Tokyo
HPG	Hyogo
HPG	Osaka
HPG	Hiroshima

Table 1: Air and HPG data sources with major cities

- The focus of the preliminary analysis was on genre, area of restaurant, and visitor trends
- Challenges in the data quickly became apparent; the restaurant ID in the Air and HPG data sources is not consistent. The plan is to merge the two data sources together.

The first step was to gain data understanding. Figure 1 and Figure 2 below show a summary of the number of restaurant genres within Air and HPG data sources.

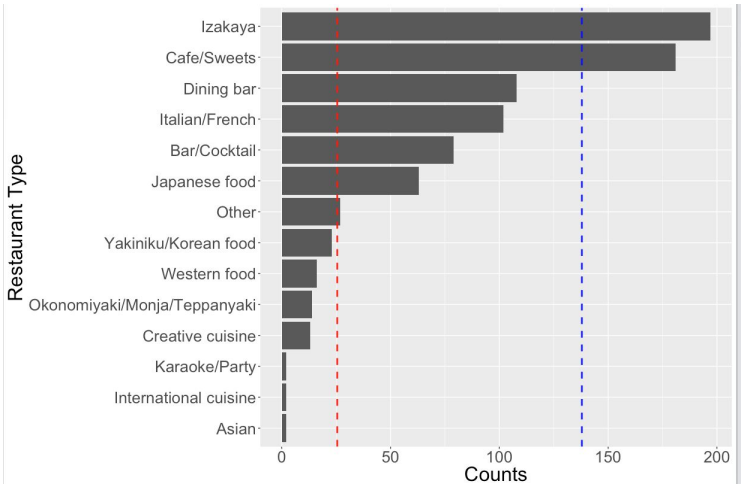

Figure 1: Air data - genres by number

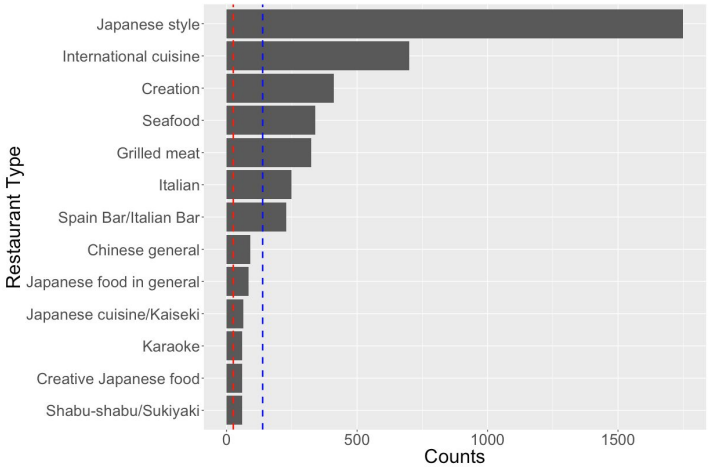

Figure 2. HPG data - genres by number

Based on the above data, we can quickly understand that Air data present greater diversity of genres, but HPG data are concentrated on Japanese styles and International cuisine.

We examined the data further to find the location of restaurants from Air and HPG data sources. Figure 3 and Figure 4 presents the city names with restaurant counts.

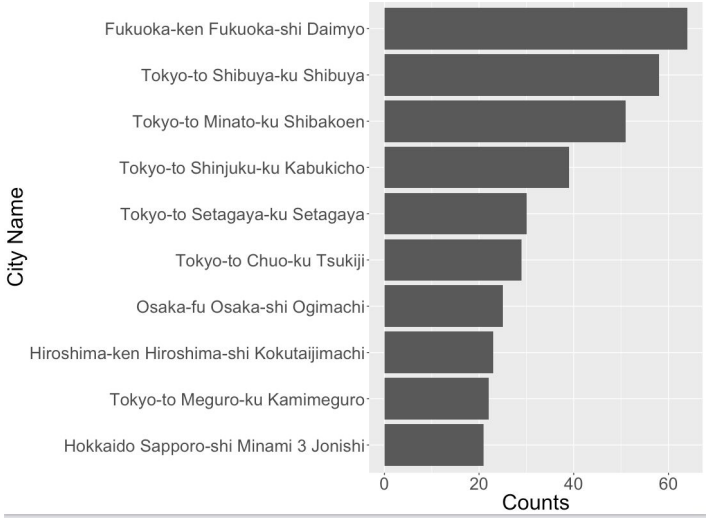

Figure 3: Air stores with location data

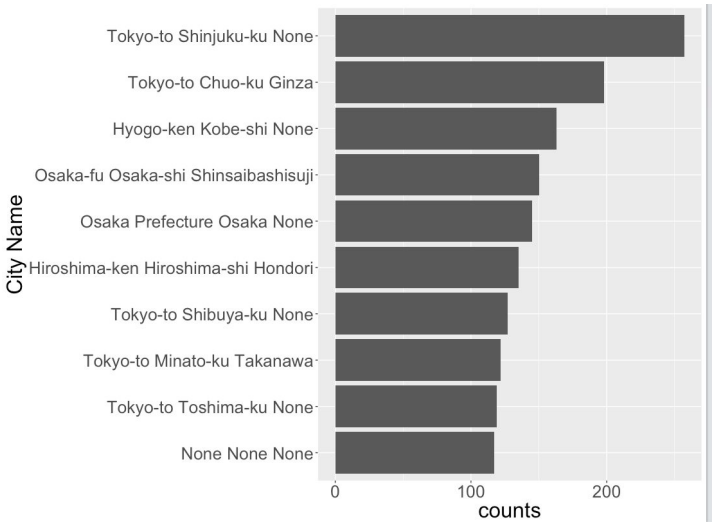

Figure 4: HPG stores with location data

Air stores are concentrated primarily in Fukuoka, Osaka, and Tokyo. HPG stores are concentrated primarily in Hiroshima, Hyogo, Osaka, and Tokyo.

We then ran a restaurant genre comparison based on the above locations to see which genre is most popular based on each particular location. See Figures 5 and 6 below.

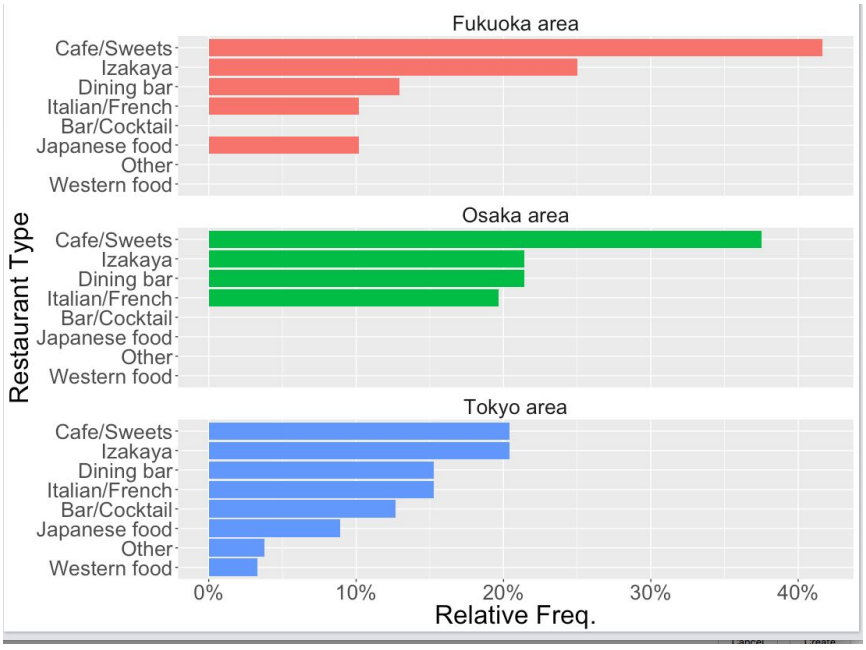

Figure 5: Air stores with genre comparison based on location

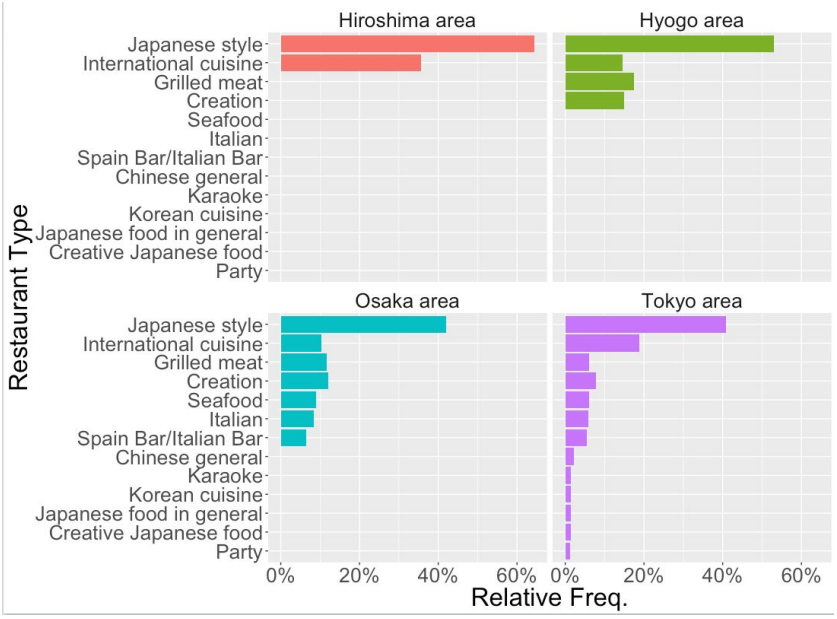

Figure 6: HPG stores with genre comparison based on location

From Air data source, it is evident that Cafe/Sweets, Izakaya and Dining bar are the most popular genres for all three locations. Also, Japanese food ranked 5th consistently in all locations. From the HPG data source, Japanese food ranked higher than any other food genre, with International cuisine coming next.

The next step was to plot the visitor data density map, one of the most important analyses. Both Air and HPG data sources were plotted as shown in Figures 7 and 8 below.

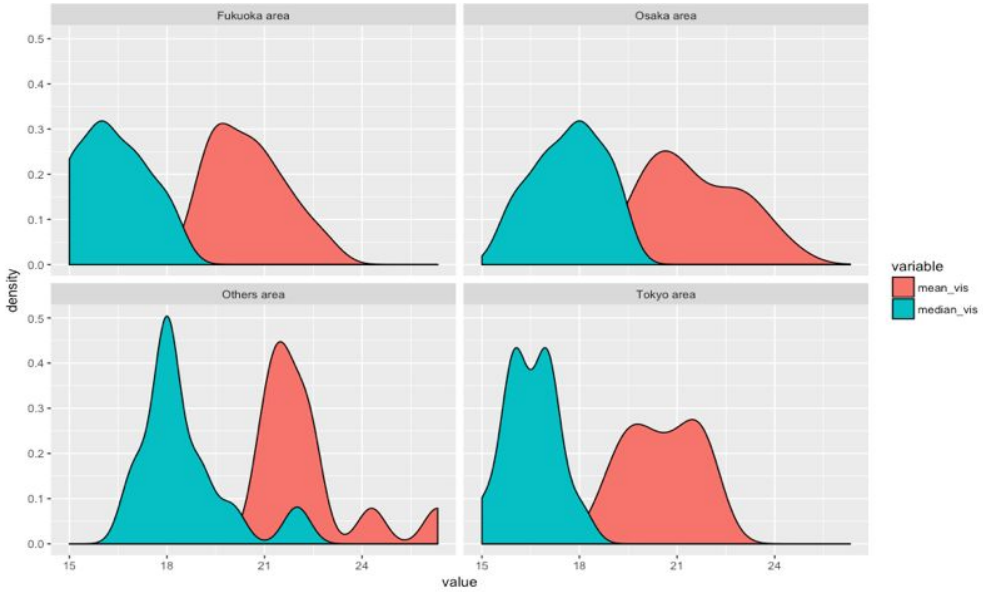

Figure 7: Air stores with visitor density map

The Air density map suggests that visitor data are not consistent across all areas for both mean and median graph. The visitor data spikes in April and June for Tokyo, but for other regions it

increases in April and then decreases.

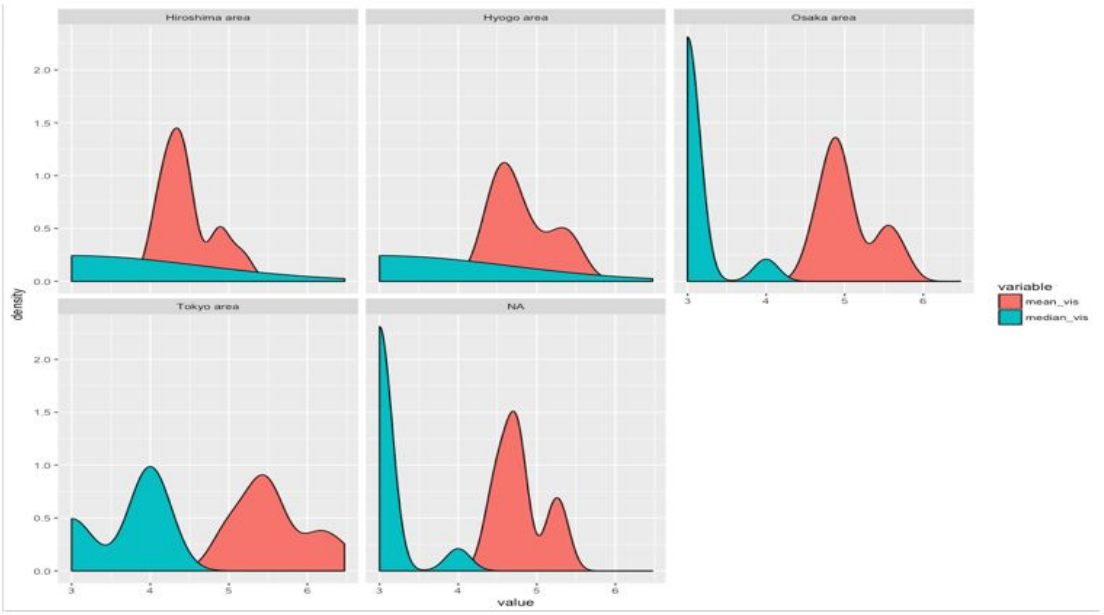

Figure 8: HPG stores with visitor density map

The HPG density map suggests that visitor data are not consistent across all areas for both mean and median graphs. The visitor data spikes in April and then spikes in October. This is because April contains a holiday week in Japan, but we need to investigate further for October spike.

Project Deliverables

At the end of this project, Foodie Analytics will deliver the following products:

- **A time series forecasting model -** This will optimize resource allocation and increase operating efficiency by allowing restaurant owners to predict the number of customers ahead of time.
- **A data visualization platform -** This web portal will give capability to visually explore all relevant data backed by predictive analytics in an interactive environment. Through rich charts and maps restaurant owners can monitor and track daily customer flow as well as spot patterns and unearth hidden opportunities.
- **A mobile application -** An alternative delivery channel to our standard web portal, it will provide the flexibility to utilize our product for those who are constantly on the move.

Team Members

Shipra Sethi - *Engagement Leader with 15 years of analytics and consulting experience in hospitality, regulatory risk, and financial services industry. Adept in using quantitative and qualitative methods to generate actionalble insights for the clients.*

Dong Bing - *Data Architect with over 10 years experience in synthesizing a wide range of data sources and transforming unorganized raw data into actionable strategic knowledge.*

Daniel Colvin - *Data Analytics professional with skills in descriptive and predictive modeling, as well as strengths in customer interaction.*

Subba Muthurangam - *Lead Developer with 20 years of experience in machine learning, data analytics, and applied statistical methods. Proficient in using Python, R, Java, Elasticsearch and NoSQL.*

Erik Platt - *Business Intelligence and Data Analytics professional adept at automation, visualization and dashboard development, as well as a variety of modeling and statistical techniques.*