Predict 498 Capstone
Milestone 1 - Project Goals
Project: Table for Four
Team: WildC  TTS Analytics

Catherine Tolley

Tamara Williams

Tom Alig

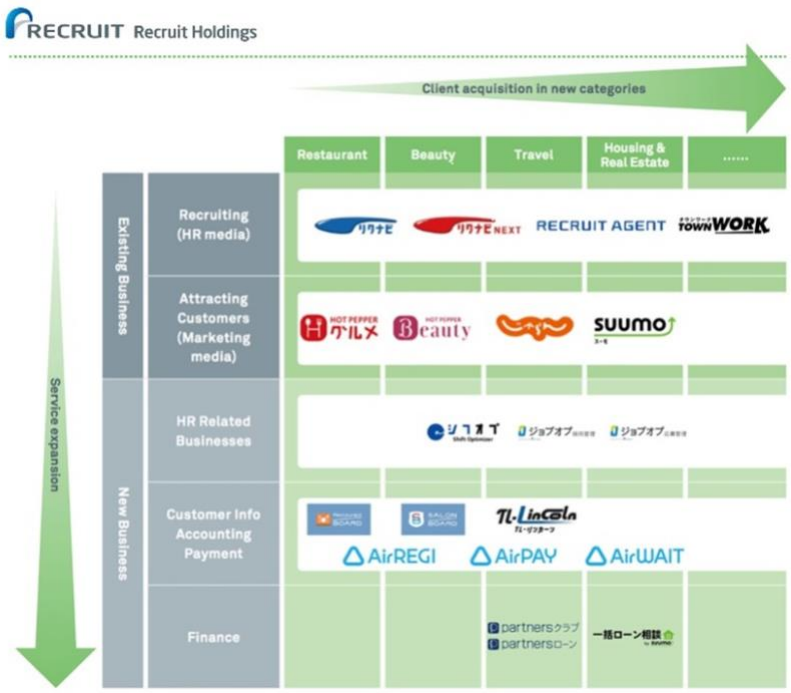
Sheela Rao

April 21, 2018

Contents

Project Background

Figure 1



Recruit Holdings Co., Ltd. is a global holding company, whose core business is divided into 3 primary areas: HR Technology, Media and Solutions (advertising), and Staffing.

With the world's ninth-largest population (127 million) (Wikipedia, n.d.), Japan has approximately 107 million mobile phone users (Statista, n.d.). Adding in roughly 24 million foreign tourists with their smartphones, these mobile phone users represent a marketing opportunity for the travel, restaurant, and entertainment business.

Boasting over $16 billion USD in annual revenue, Recruit Holdings collects information and provides services in the housing, bridal, educational, automobile, travel, dining, and human resources industries, among others.

Business Case

For this engagement, we will be targeting the Media and Solutions vertical. We will focus on the company's offerings in the restaurant sector. Recruit Holdings' AirREGI point-of-sale (POS) product currently offers cash register and table management services via iOS tablet and mobile devices for restaurants in Japan. A related Recruit Holdings business, Hot Pepper Gourmet (HPG), a restaurant marketing website and mobile app, provides Japanese diners with a searchable interface to locate restaurants and obtain promotional coupons. The company has targeted these two products for continued growth in 2018. Recruit Holdings has hired WildCATTS Analytics to further its vision of connecting customers with businesses, specifically in the Japan restaurant market. The proposed project will provide new business insights and drive shareholder value.

Benefits to Recruit Holdings

This project will tap into Recruit Holdings breadth of data collection in the Japan restaurant market. Achieving success in the project goals will enable Recruit Holdings to adopt the same capabilities across its multiple product channels worldwide.

AirREGI Enhancements

Restaurants lose 4-10% of food inventory due to overproduction, spoilage, and expiration (National Restaurant Association, n.d.). Both understaffing and overstaffing relative to restaurant volume negatively impact restaurant revenues. Promotional offerings must be used tactically to avoid cutting into profits.

By operating at or near capacity, by having the number of diners match the restaurant occupancy, restaurants may maximize revenue. Maximizing revenue can lead to optimizing profits.

What if a restaurant owner knew how many diners were expected on Saturday night? She can use her knowledge and experience to:

- ✓ Adjust staffing to anticipated volumes, ensuring a balance of customer satisfaction and cost control.
- ✓ Offer targeted promotions only when they are needed to drive volume and avoid promotions at times when volumes are already sufficient.
- ✓ Offer a limited menu when low volumes are beyond the restaurant's control due to holidays and off-peak seasons
- ✓ Limit or expand hours.

The Table for Four project's daily restaurant volume prediction will enable AirREGI clients to optimally manage their bottom line in the context of their business. This enhancement creates a best-in-class feature differentiator for the AirREGI product, creating an uplift in AirREGI POS adoption by restaurants.

Hot Pepper Gourmet Benefits

HPG currently displays restaurant choices, menus, and a standard promotional coupon to Japanese diners. Restaurant owners may choose to respond to predicted volumes with ***targeted and limited-time promotional offerings*** to potential diners on the HPG website and app. Streamlining access to HPG promotions for restaurant owners increases the product's perceived relevance and driving downloads.

Project Goals

1. Develop a robust predictive model to predict daily restaurant volume up to 30 days in advance.
2. Develop a dashboard for the AirREGI POS solution to deliver predictions in context, enabling restaurant managers to take appropriate actions, including pushing customized promotions to the HPG marketing product.

3. Develop a prototype of a mobile experience for restaurant clients to interact with and potentially respond to volume predictions.

4. Maximize insights from Recruit Holdings' valuable restaurant sector data with meaningful data visualizations.
5. Recommend strategy to monetize prediction-enhanced AirREGI Point of Sale premium-subscription solution.

Potential Recommendations

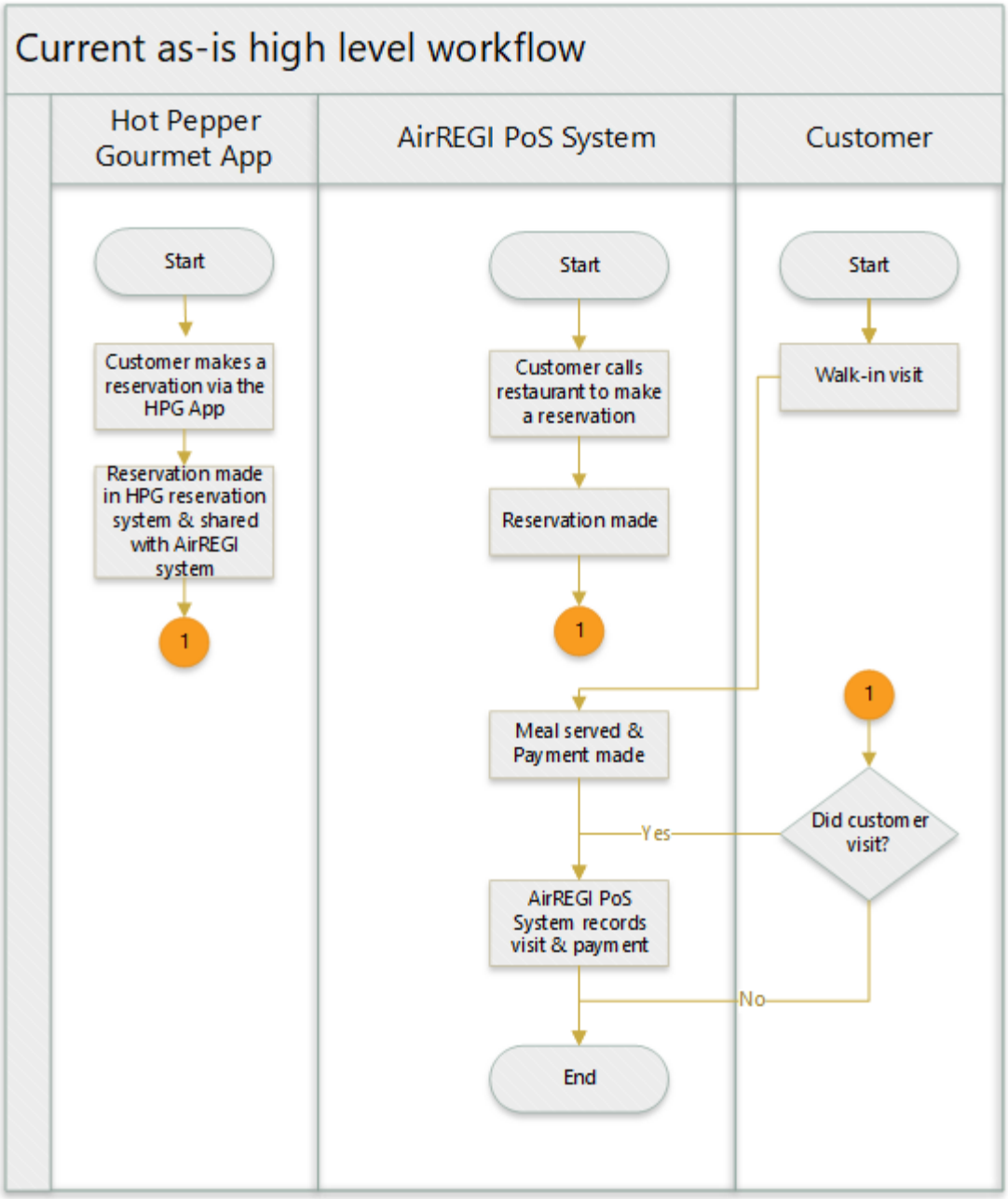
- Deliver specific recommendations to AirREGI clients, including but not limited to
  - Just-in-time, weekend and happy hour promotions to HPG product based on volume predictions.
  - Optimize inventory and resource allocation.

- Deliver recommendations to Recruit Holdings:
  - Stimulate quarterly revenue growth and expanding services to more potential client restaurants.
  - Promote the new subscription fee premium AirREGI POS system.

Business Understanding

Reservations are made via the HPG app or calling the restaurant directly. The current system collects restaurant reservation data and daily volume.

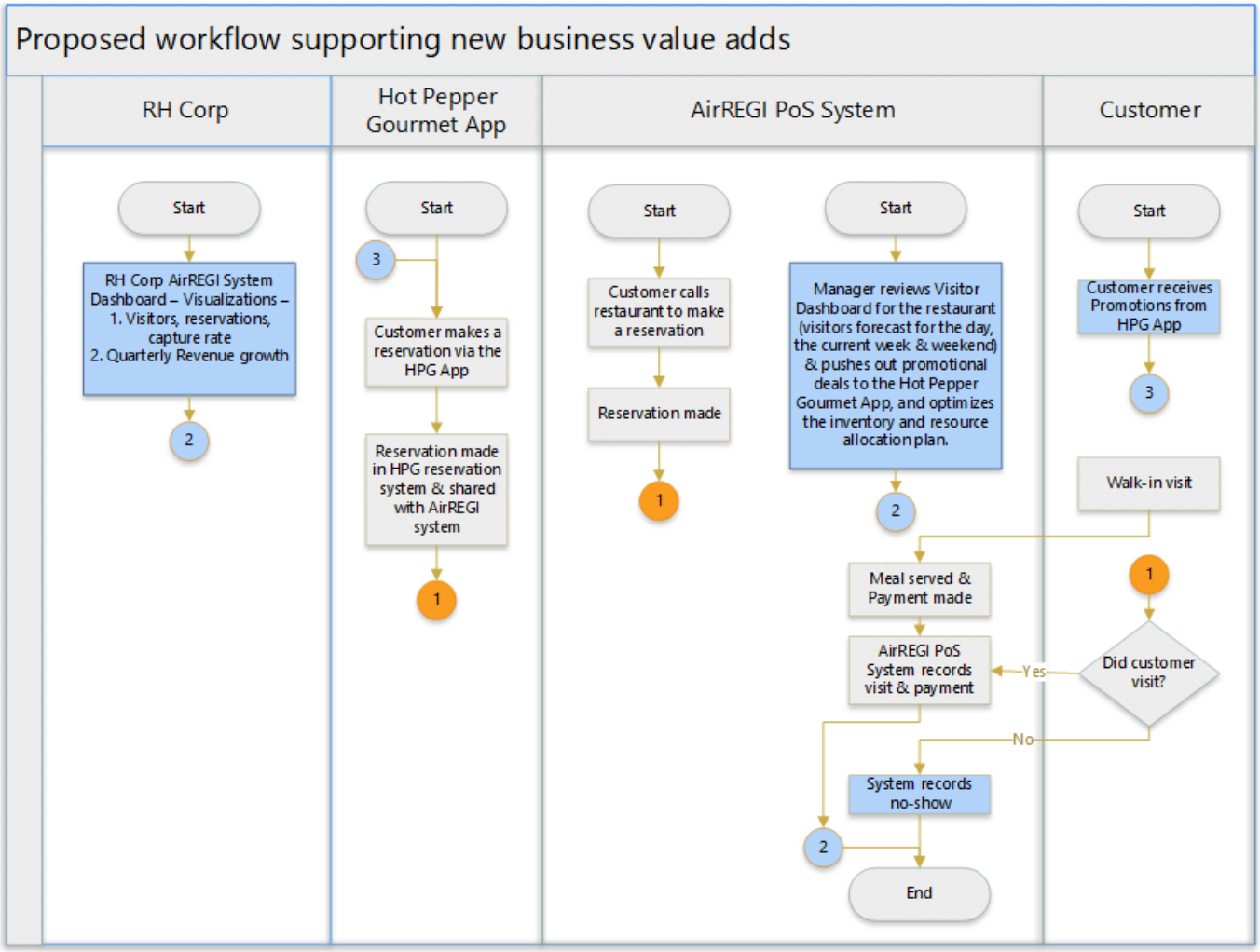
Figure 2


Current as-is high level workflow
Hot Pepper Gourmet App
AirREGI PoS System
Customer
Start
Customer makes a reservation via the HPG App
Reservation made in HPG reservation system & shared with AirREGI system
1
Start
Customer calls restaurant to make a reservation
Reservation made
1
Start
Walk-in visit
1
Meal served & Payment made
Did customer visit?
Yes
AirREGI PoS System records visit & payment
No
End

Proposed Changes to Current Workflow

- WildCATTS Analytics proposes to use existing reservation data, in conjunction with external data, to increase the ability of restaurants to be at capacity.
- The proposed system will provide the individual restaurant owner with a prediction of volume for his/her restaurant for a given date, up to 30 days in advance.
- The proposed system will enable the manager to push mobile notifications to HPG for promotions to increase the number of diners.
- The proposed system will provide data to Recruit Holdings, which will enable management to visualize the big picture, including number of diners, reservation capture rate, and quarterly revenue growth.

Figure 3



Our Approach to Analysis

Tools

The tools we will use on this project include, but are not limited to:

- R and Python - for coding models, exploratory analysis and data enrichment
- Tableau Desktop - for data visualization
- Tableau Server/Tableau Mobile - for data visualization on mobile devices

Modelling

We will build one or more predictive models; we may also build descriptive models. The potential methods we may use include:

Predictive

- Time series forecasting - ARIMA, ETS, external regressors models
- Regression - Random Forest, Generalized Linear Regression
- Artificial Neural Networks (ANN)

Descriptive

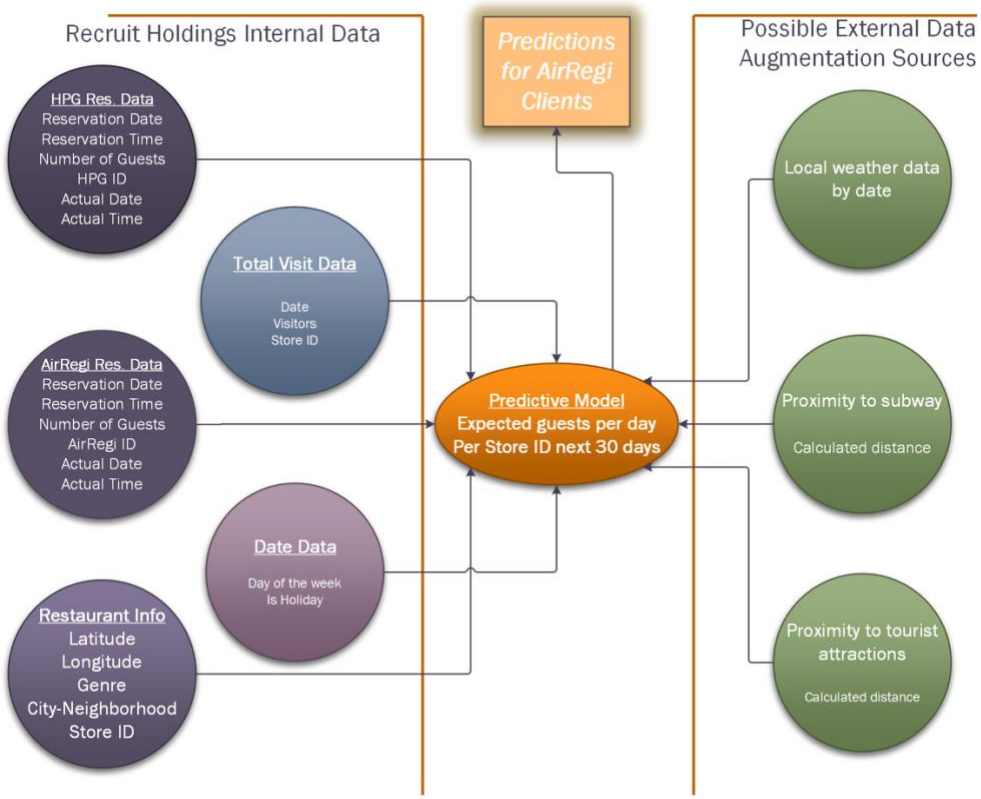
- Cluster Analysis

Data Understanding

This engagement relies on the following data, illustrated in Figure 4. All Recruit Holdings data come from a Kaggle competition which ended in February, 2018.

- Reservations made through the HPG app for the time period we are modeling
- Reservations made thru the AirREGI POS system for the time period we are modeling
- Total daily volumes for 829 AirREGI client restaurants
- Individual restaurant details such as type of cuisine served, and location data
- Mapping data to relate AirREGI IDs to HPG IDs
- Day-of-week and holiday data for the time period we are modeling
- Possibly, additional publically available data, like weather, to enhance the model

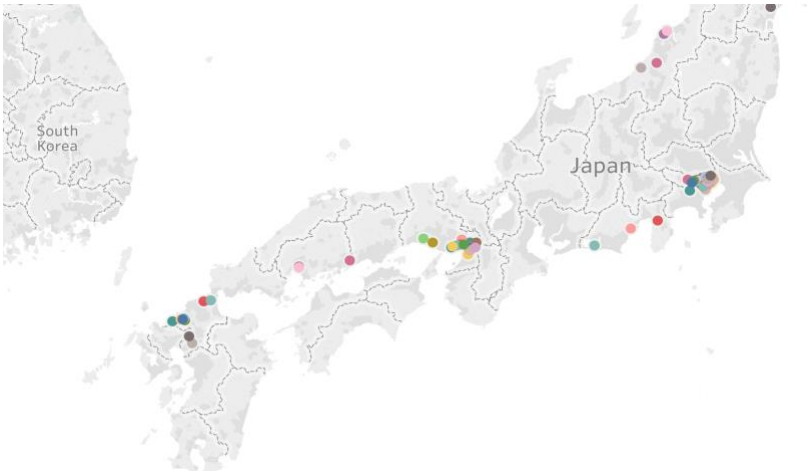
Figure 4



Initial EDA

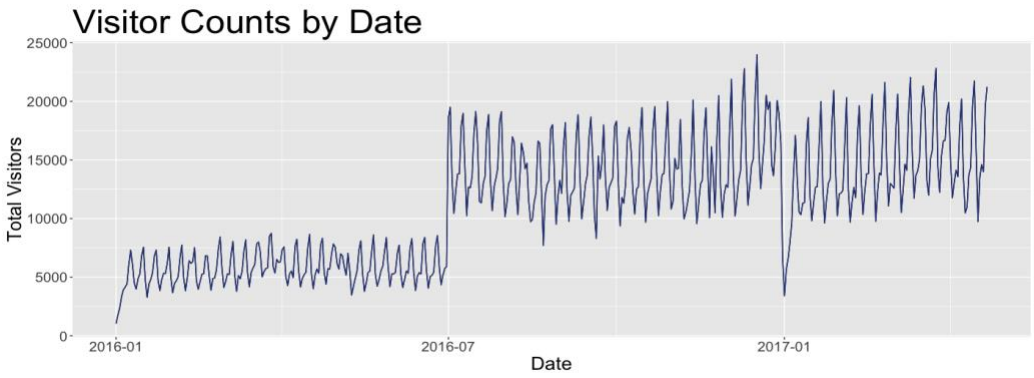
As seen in Figure 5, data are from restaurants in multiple locations across Japan. City sizes vary, which means data will need to be normalized when making city-to-city comparisons.

Figure 5



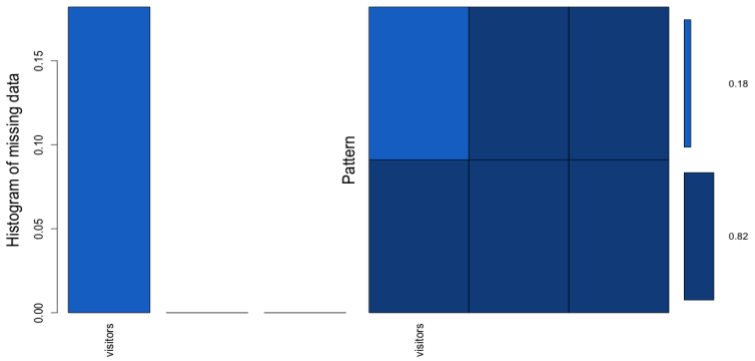
The set of data for training models covers more than a year, allowing us to explore seasonal patterns. As seen in Figure 6, the data reveal seasonal impact of “Golden Week”.

Figure 6



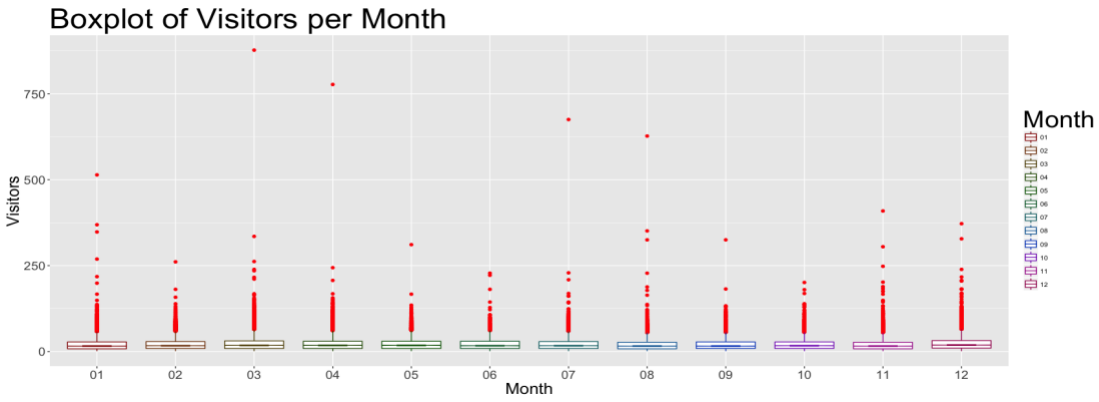
An analysis of the data reveals there are missing data. In Figure 7, missing data are shown in the light blue color All missing values are from the Visitors field, indicating either the restaurant did not report visits, or there were no visits due to closure, like a holiday. There are a variety of techniques which can infer missing data. Additional investigation into the most appropriate choice will be needed.

Figure 7



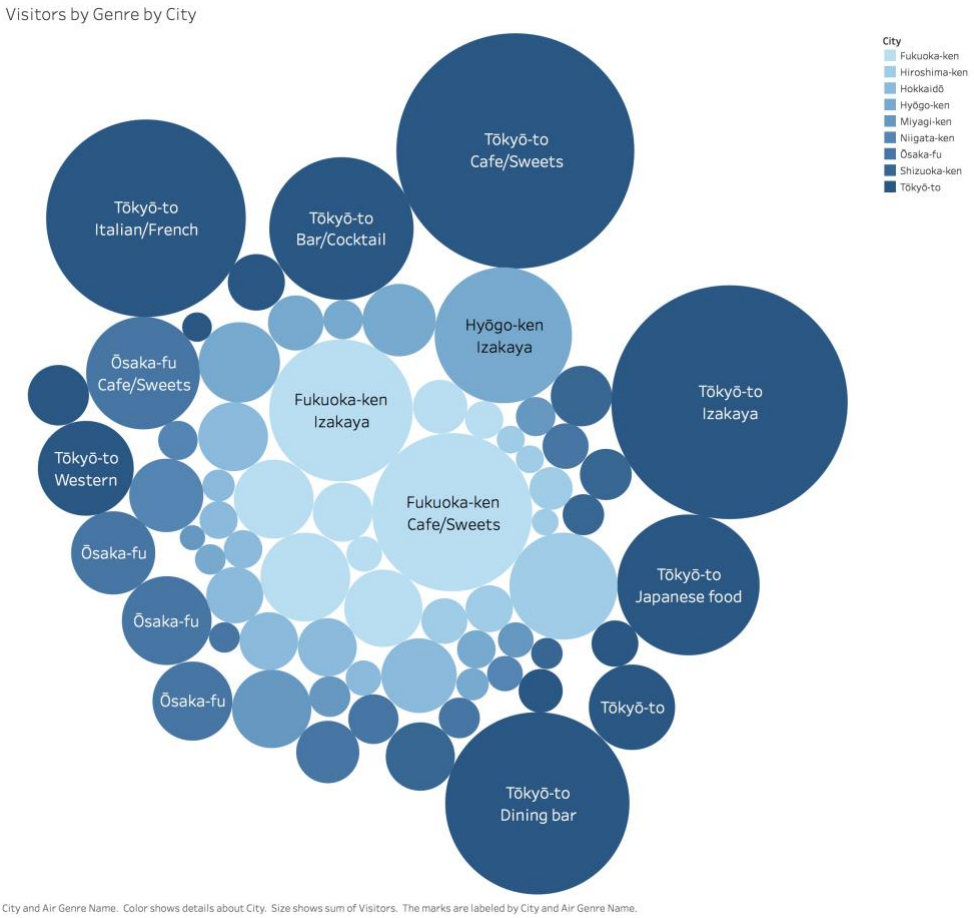
The view of total visitors per month (Figure 8) shows evidence of outliers (in red) which may impact the models. Additional investigation will be needed to determine if, and how, these should be handled.

Figure 8



The visitors by city and genre view of the data in Figure 9 suggests there may be clusters in the data worth exploring as part of our modelling process.

Figure 9



Additionally, there are day of the week (Figure 10) and monthly (Figure 11) variations which suggest cycles and will need to be explored more fully.

Figure 10

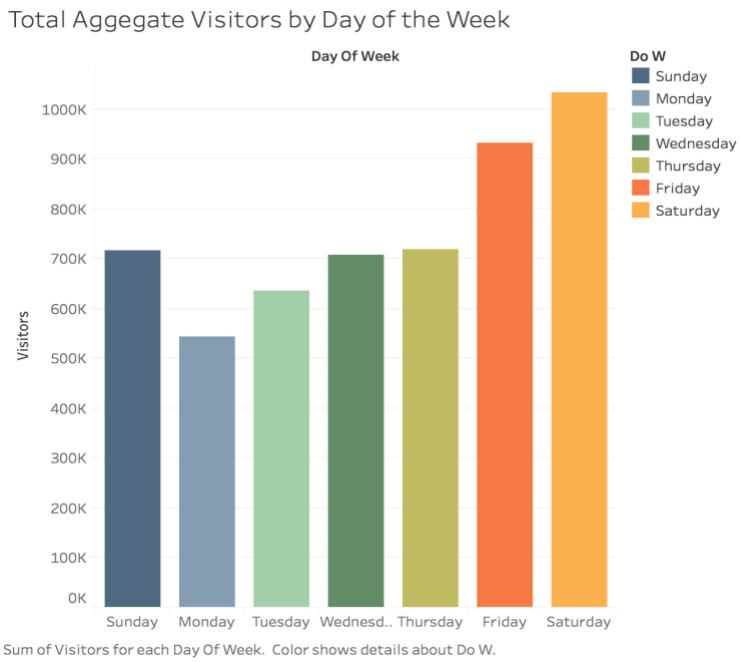
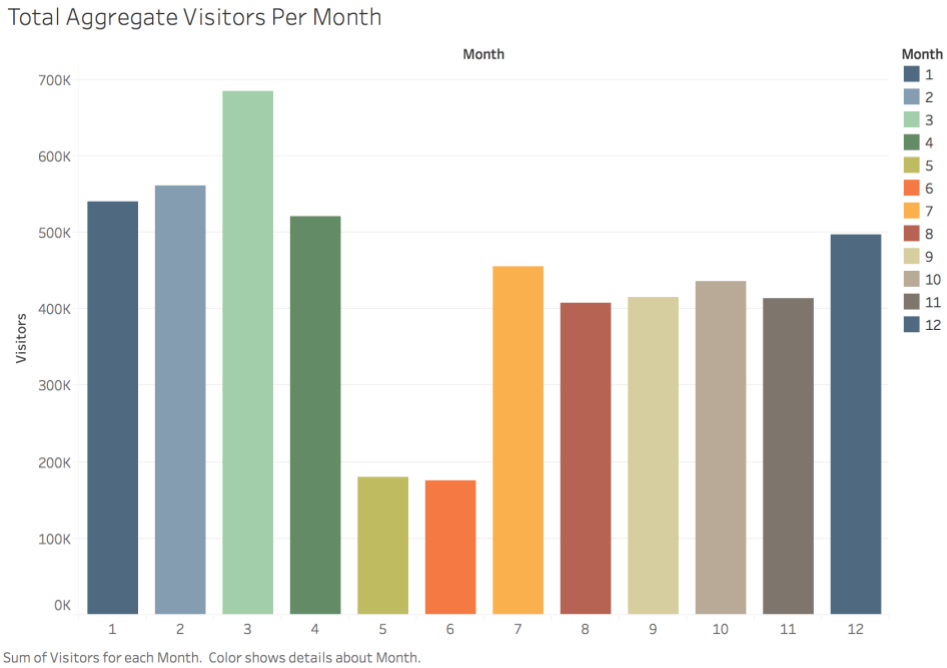
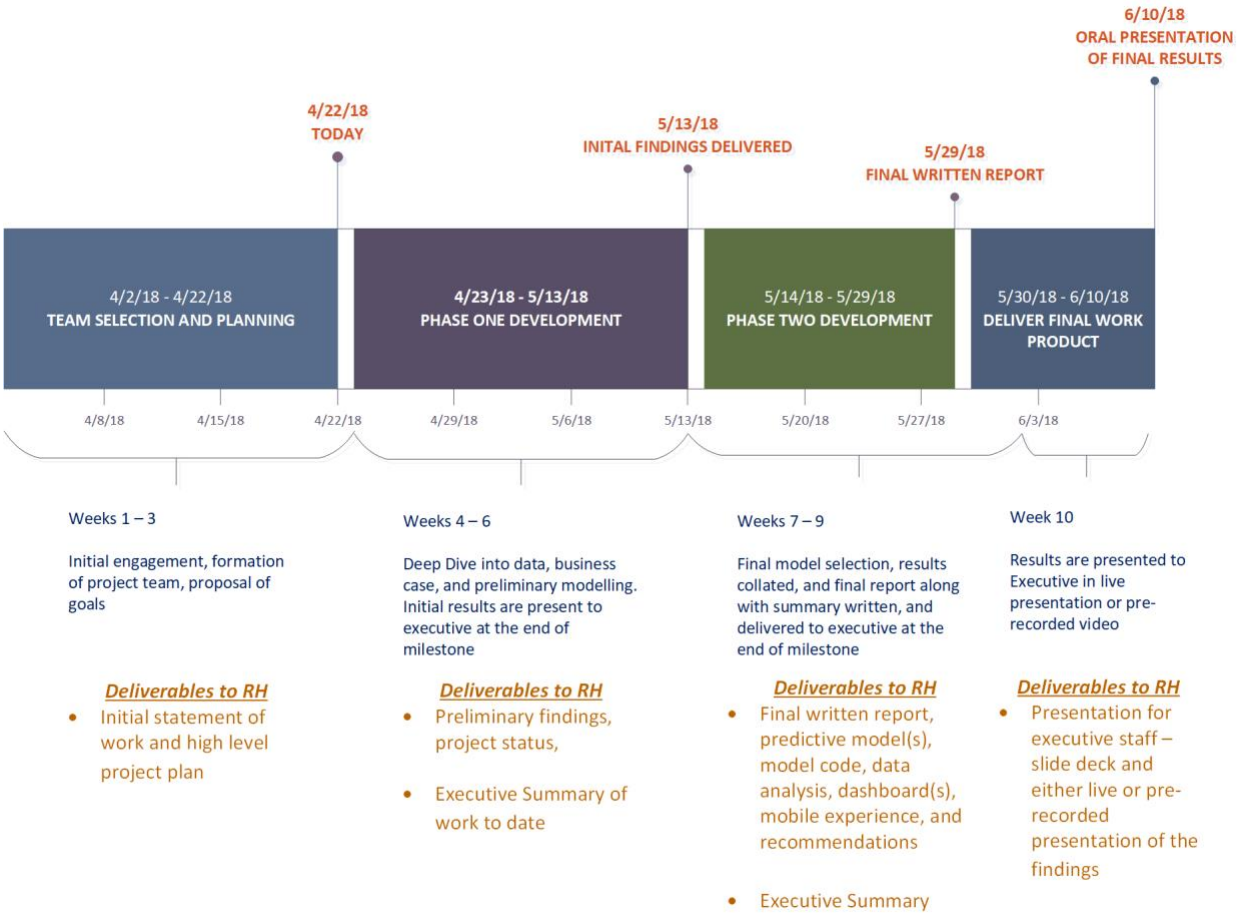


Figure 11



Project timeline & deliverables

Figure 12



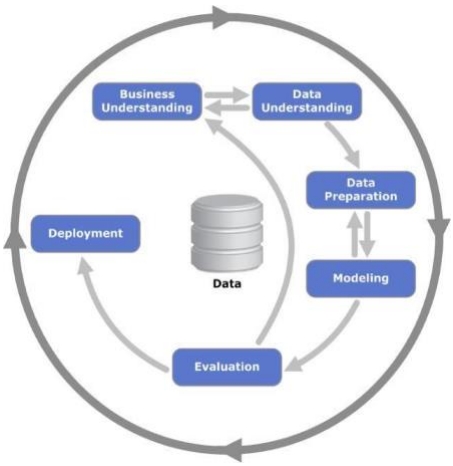
Project Methodology

WildCATTS Analytics utilizes an Agile Scrum management approach, combined with CRISP-DM for organizing and delivering work.

- The overall workflow follows CRISP-DM methodology which is an iterative process shown in Figure 13
- The project Scrum team meets once a week to allocate tasks and set deliverables for the following week

- In each weekly meeting, any incomplete scheduled tasks, are discussed along with time required to complete the task and schedule adjustments are made as needed

Figure 13



- Each team member is highly capable in multiple disciplines; therefore, tasks are allocated based on group capacity, rather than becoming gated by a single role
    - In the event of a task taking long than expected, no one team member becomes a gating factor as there are always multiple people capable of working on each task
    - This also means if a task is running long, and there is a way to split it, multiple people can work to complete it
- The project is divided into 4 phases. Timeline for deliverables in given on the following slide
    - Phase 1 - Team formation, and initial engagement with customer and customer problem
        - Deliverable – Initial statement of work (this document)
    - Phase 2 - Deep dive into the business problem and build first round of models.
        - Deliverable - Initial Findings Report, project status, and Executive Summary of progress to date
    - Phase 3 - Final model section and testing. Generation of final reporting artifacts and recommendations
        - Deliverable - Executive Summary Report, Final Report with all models, code, data analysis, dashboard(s), mobile experience, and recommendations.
    - Phase 4 - Presentation of results
        - Deliverable – Live, or pre-recorded presentation for executive staff of the results and recommendations along and associated slide deck

Team Members

WildCATTS Team

This is a team with diverse backgrounds who are passionate about building business value using Analytics. They are all graduating in MS in Data Science (Predictive Analytics) from Northwestern University this June 2018.

	*Catherine Tolley* Catherine has over 10 years of data analytics experience spanning KPI and dashboard development, strategic analysis, data source development, program evaluation, statistical analysis and multivariate modeling. She holds a B.A. from DePauw University and is a graduate of the Centers for Disease Control's Emerging Infectious Diseases fellowship.
	*Tamara Williams* Tamara's background is in Software Engineering, Software Engineering Management, and Software Quality Assurance. She has more than 20 years of experience working at Microsoft on everything from games to Enterprise software. She holds a B.S. in Mathematics from University of Washington, and an MSE from Seattle University.
	*Tom Alig* Tom's background includes 7 years in various Purchasing roles at General Motors, and 17 years in healthcare. He has a B.A. from the University of Michigan, and an MBA from Michigan State University.
	*Sheela Rao* Sheela is an Engineer (CS) with 20+ years of experience in the Software Industry in software product and services companies. She worked at Microsoft for over 13 years in a variety of Program Management functions. Her last role was that of a Patent Analyst, which got her interested in pursuing a formal education in Data Science.

Works Cited

National Restaurant Association. (n.d.). *Manage My Restaurant*. Retrieved from National Restaurant Association: 1. http://www.restaurant.org/Manage-My-Restaurant/Operations/Back-of-House/Measuring-food-waste-can-bolster-your-bottom-line

Statista. (n.d.). *Number of mobile phone users in Japan from 2013 to 2020 (in millions)*. Retrieved from Statista The Statistics Portal: https://www.statista.com/statistics/274672/forecast-of-mobile-phone-users-in-japan/

Wikipedia. (n.d.). *Japan*. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Japan