



Business Intelligence Challenge

Japanese Restaurants



JAPANESE RESTAURANTS DATA



3 csv files

13,507 records

49,138 restaurant visitors



39

restaurants



17

months



14

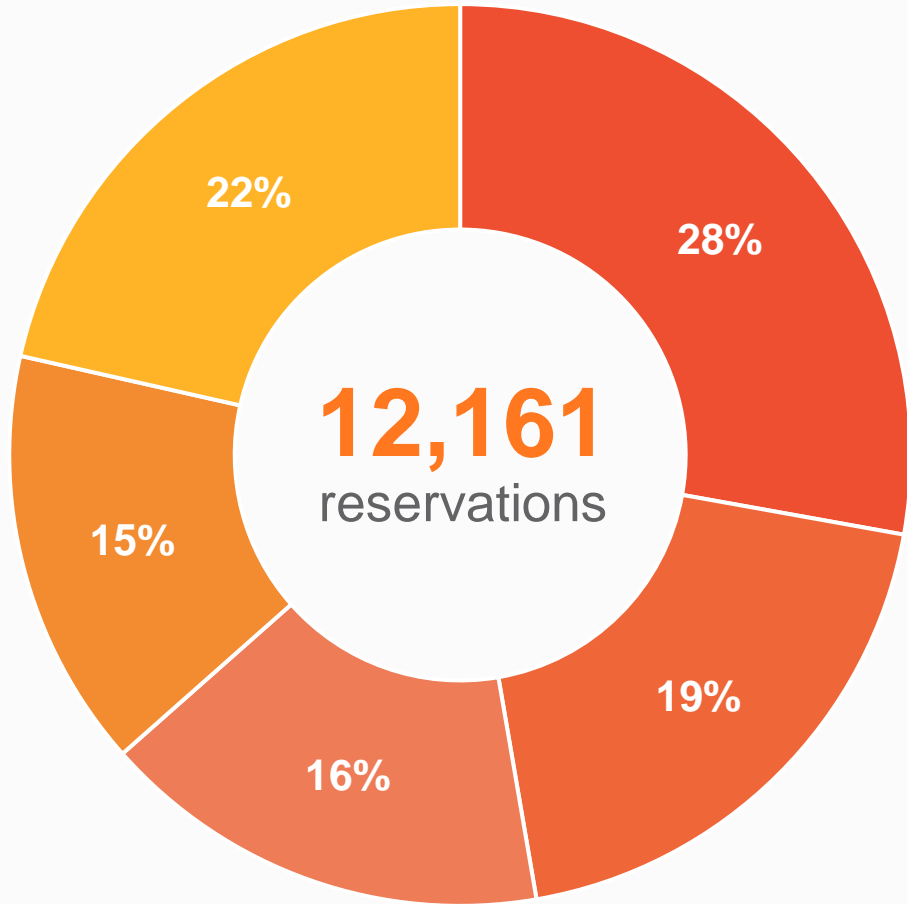
genres



9

prefectures

“Izakaya” is the genre of restaurant with the most reservations. The genre with fewer reservations is “Bar/Cocktail”



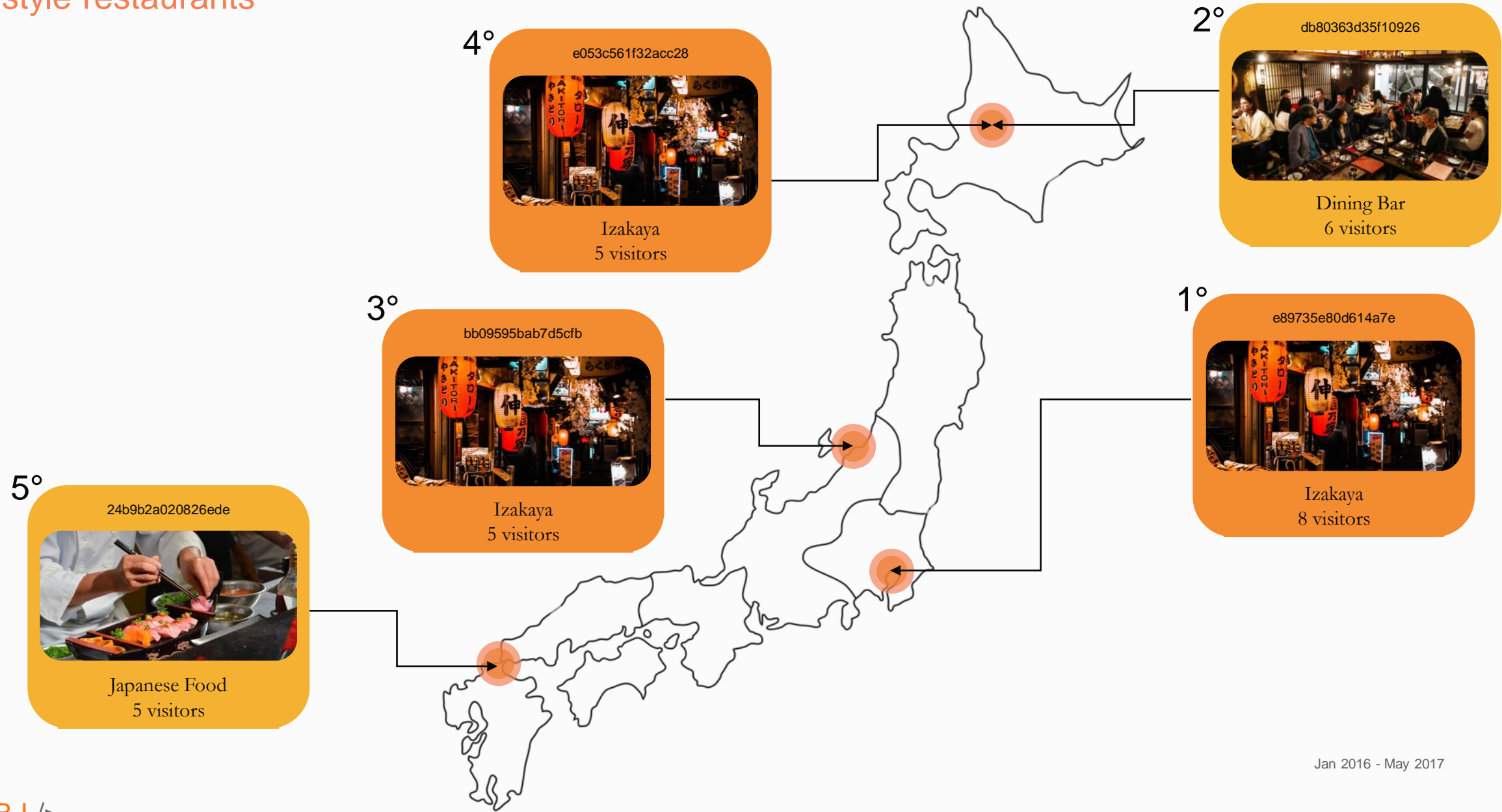
Jan 2016 - May 2017

■ Izakaya ■ Italian/French ■ Japanese Food ■ Western Food ■ Other



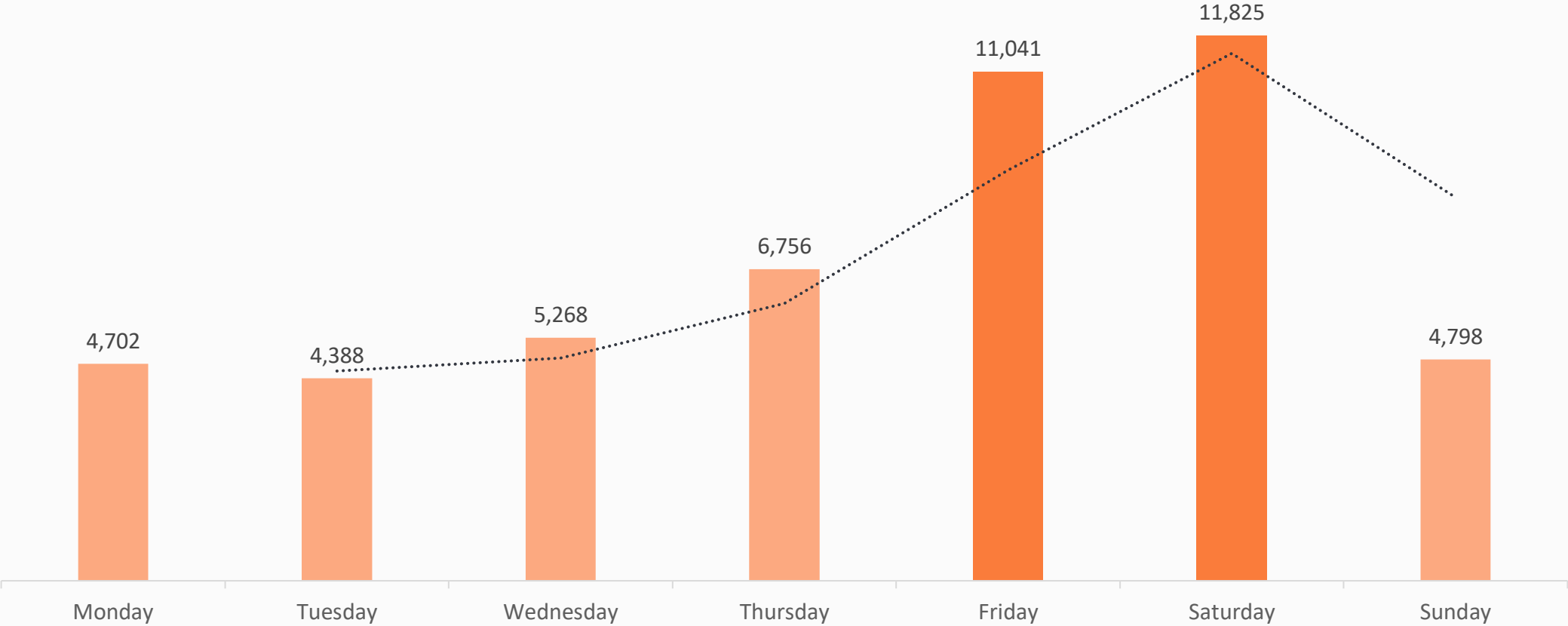
Izakaya (居酒屋) are casual drinking establishments. They are one of the most common types of restaurants in Japan and can be easily found around train stations and entertainment districts.

Among the **top 5** restaurants with the highest average number of visitors on holidays, **3 are Izakaya-style restaurants**



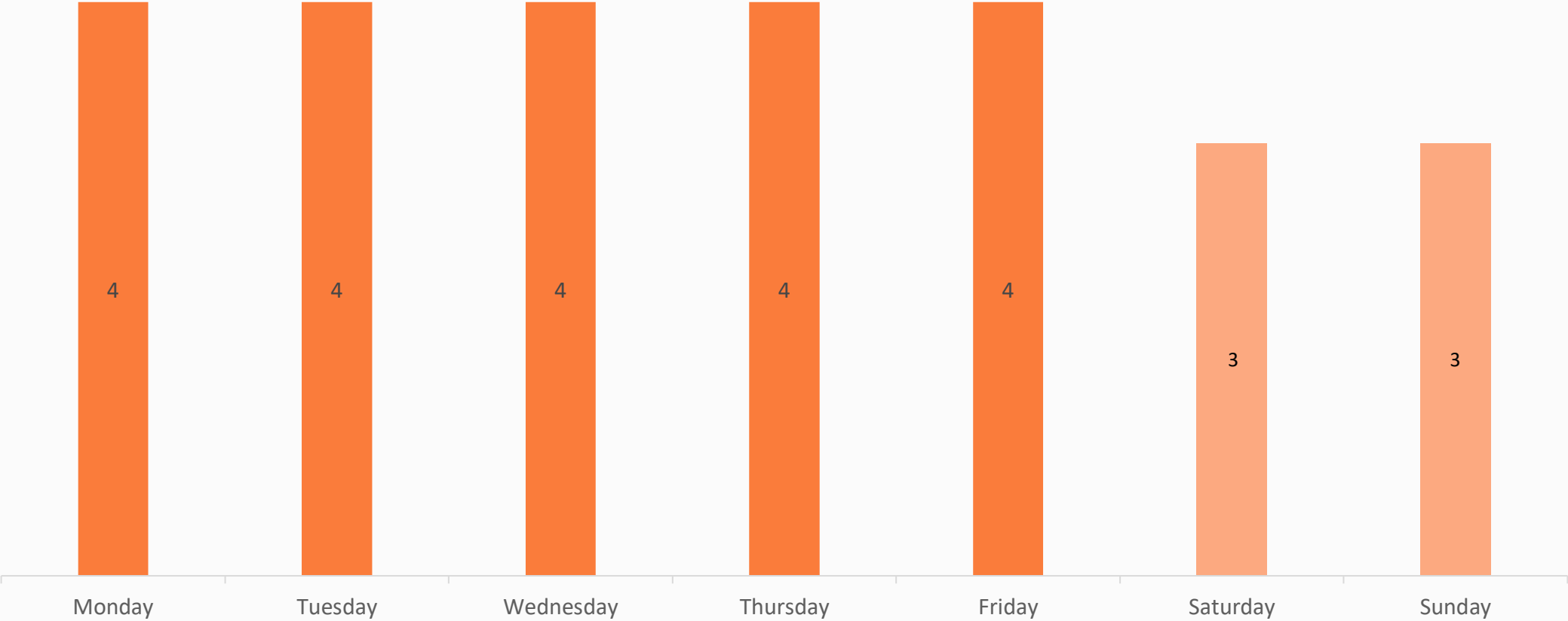
Jan 2016 - May 2017

Fridays and Saturdays have the highest total number of visitors...



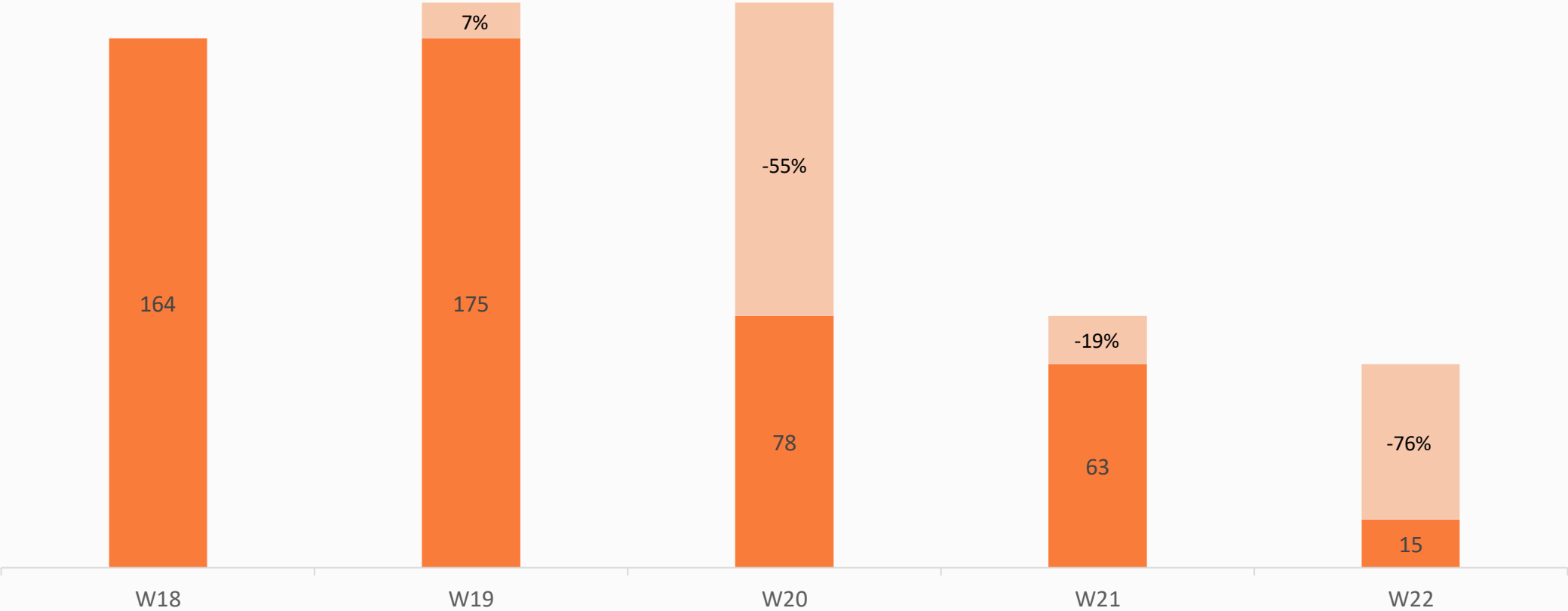
Jan 2016 - May 2017

However, Weekdays have Higher Average Visitors than Weekends



Jan 2016 - May 2017

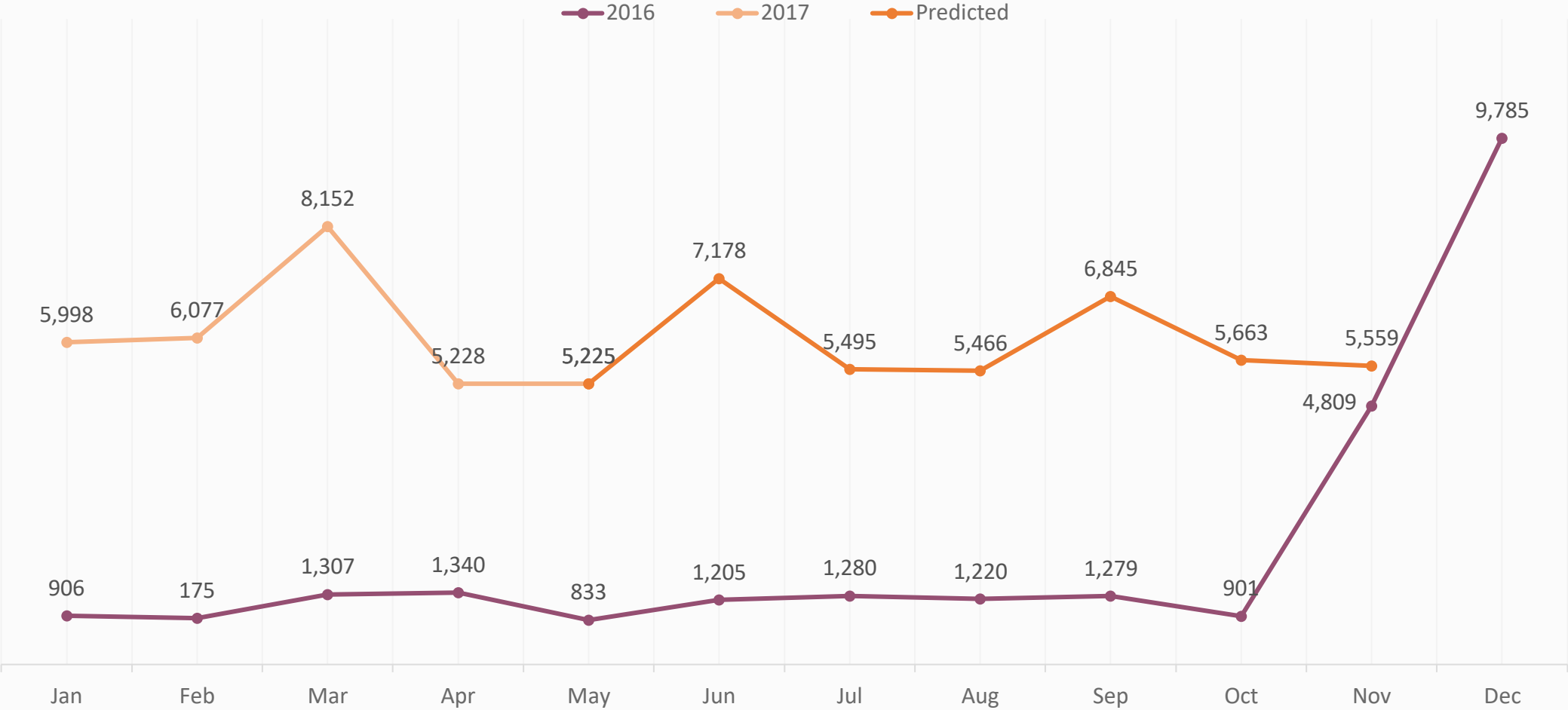
The percentage growth of the past 4 weeks has been decreasing...



May 2017

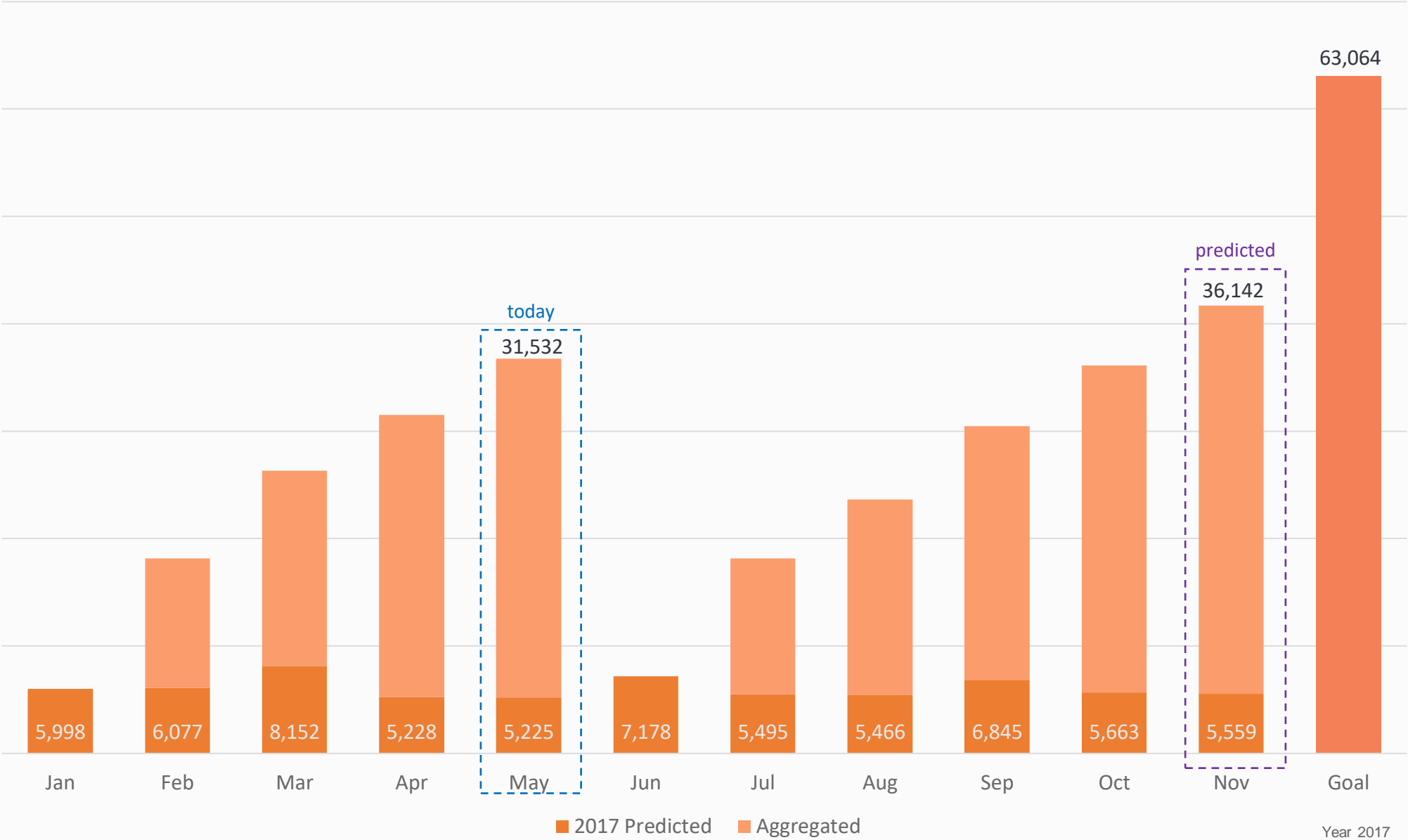
However, the sum of visitors is expected to grow for next month (June) and to average around 5,900 visitors for the remainder of the year.

Jan 2016 - May 2017**



* The prediction was made using an ARIMA model.
** This data has been transformed using Feature Engineering

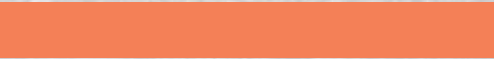
Increasing Visitors: Aiming To Double Total Within The Next 6 Months



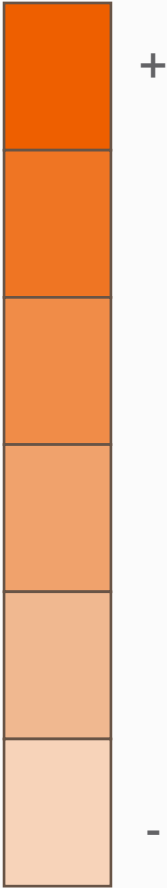
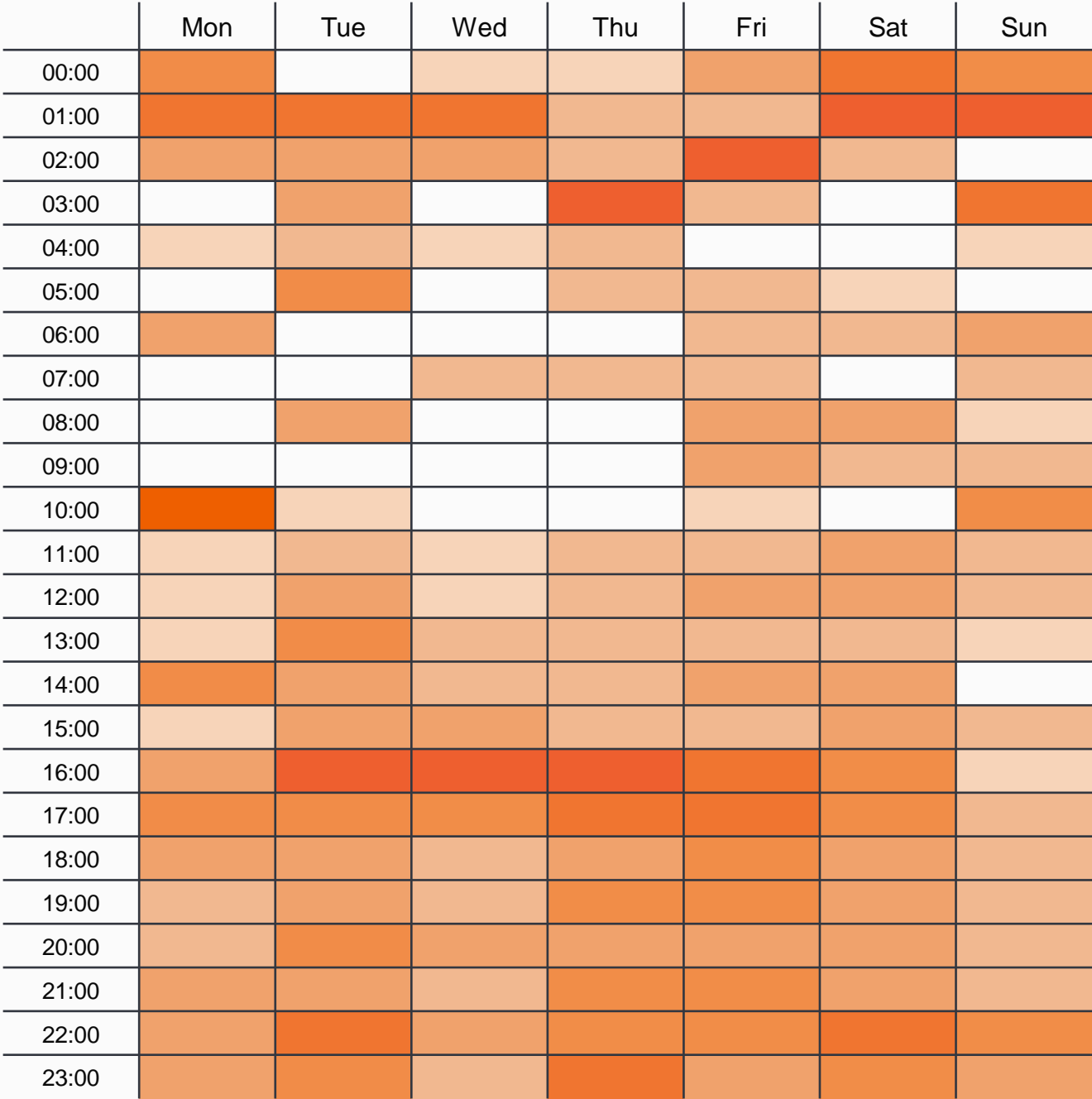


Strategy to Double Total Visitors

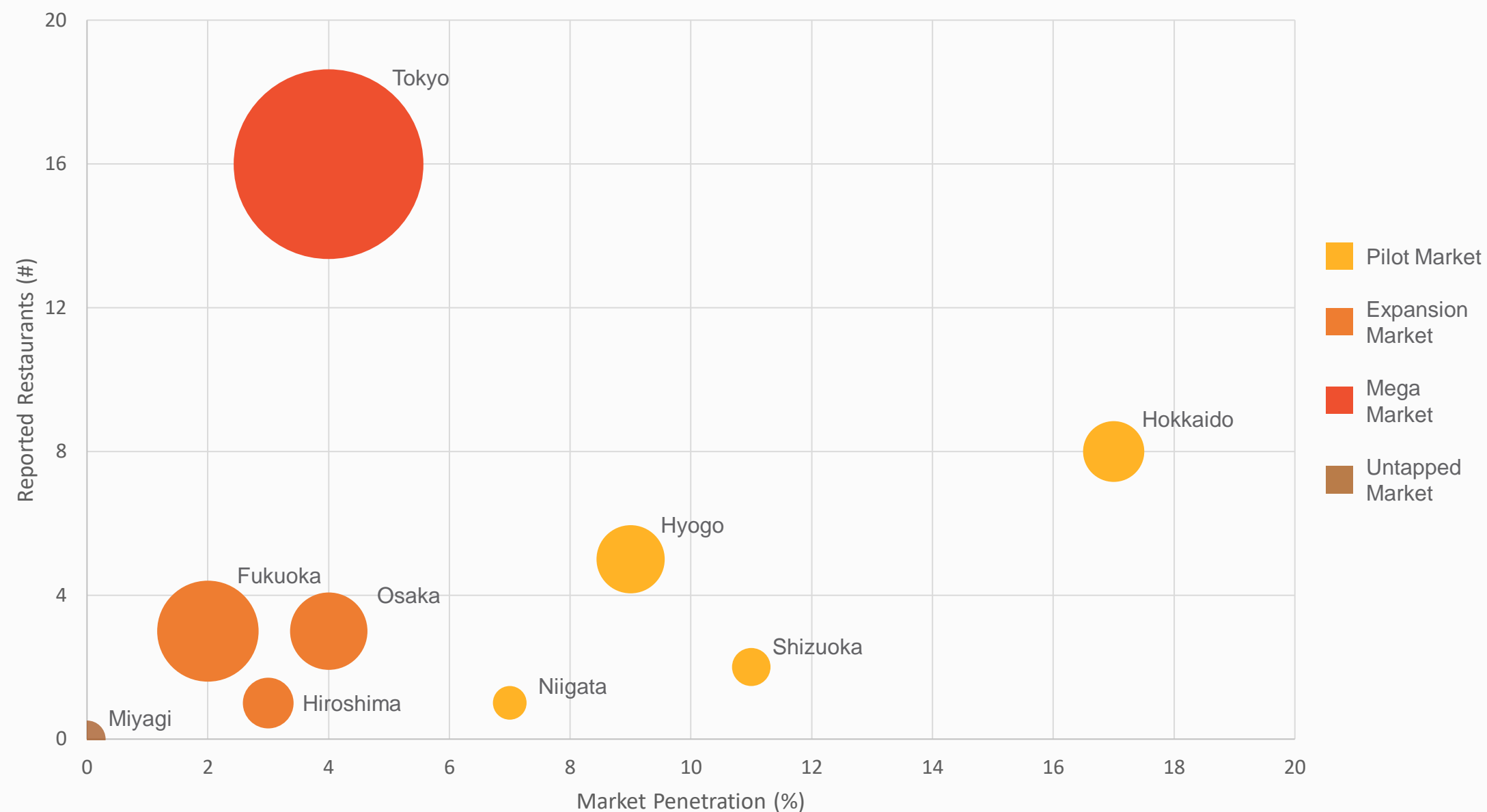
Japanese Restaurants



Average Visitors per Hour

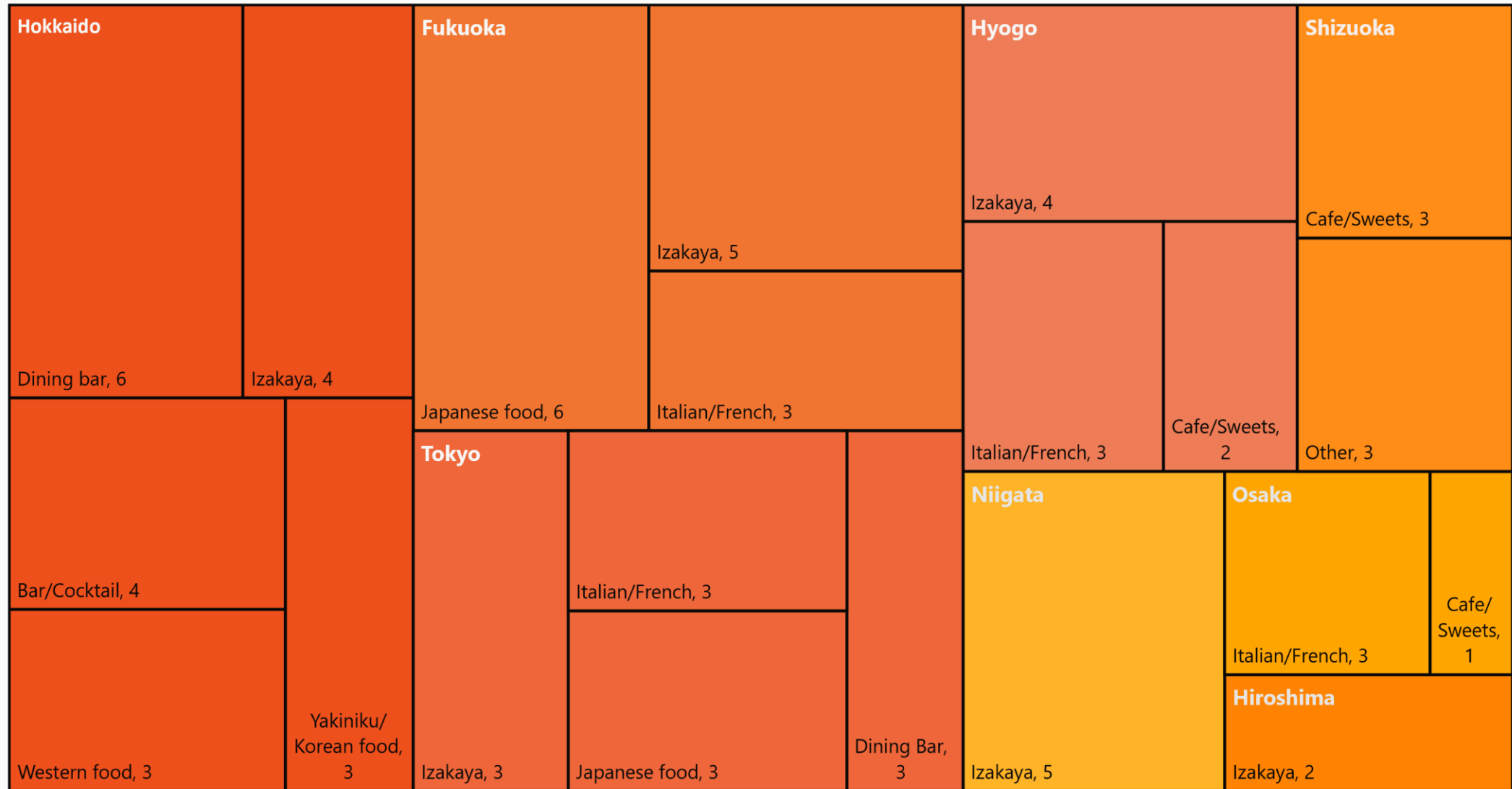


Market Segmentation across Japanese Prefectures



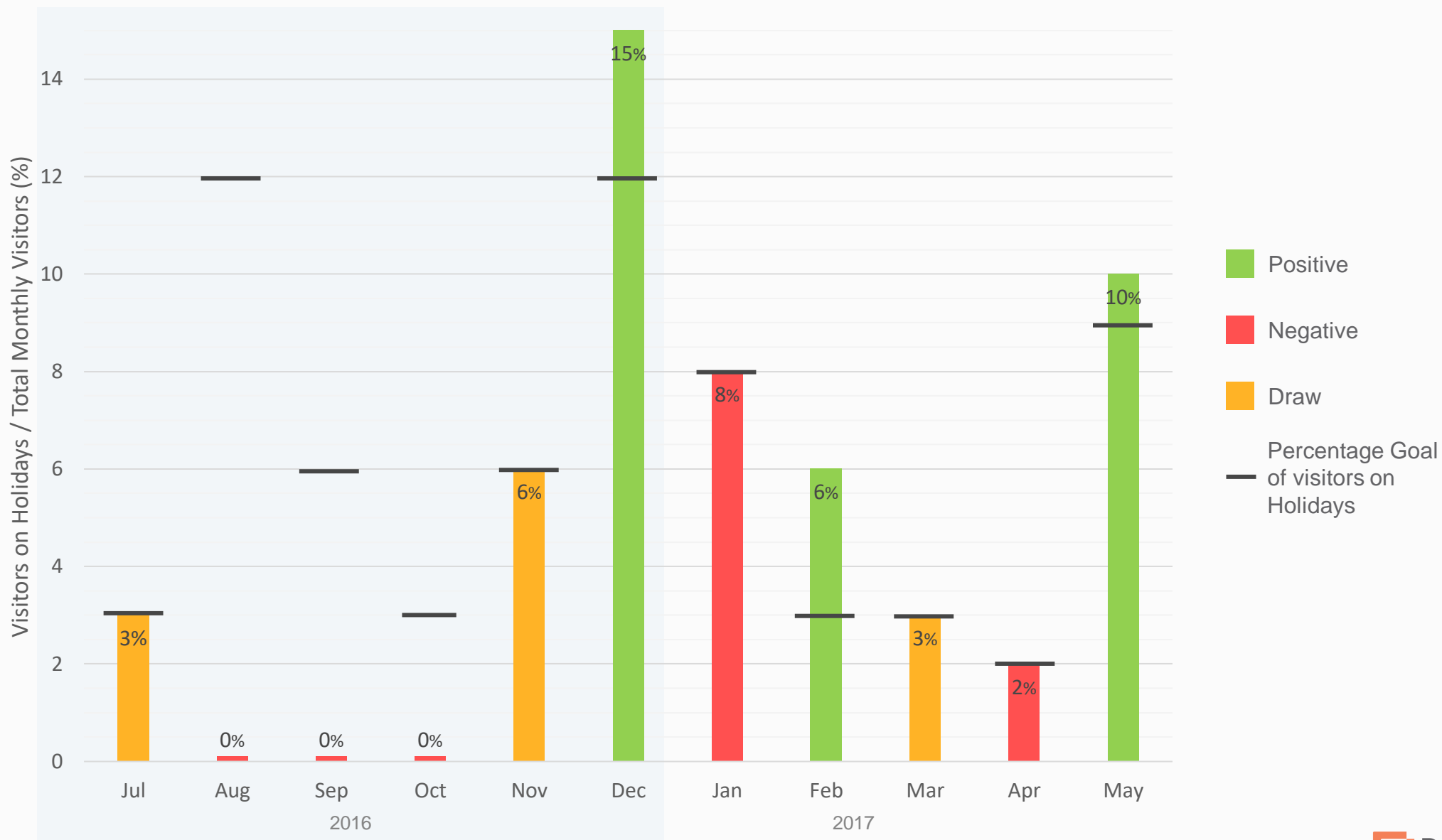
Jan 2016 - May 2017

Average Visitors per Genre across Japanese Prefectures



Jan 2016 - May 2017

Only 3 out of the last 11 months have capitalized on the holidays



PROJECT DOUBLE DINERS

initiatives



Happy Hour Promo on Weekdays (17:00-18:00)

- Sustaining momentum beyond Peak Hours (16:00)
- Capitalizing on Weekdays footfall by optimizing table turnover
- Transforming After-Office visits into routine habit



Themed Nights in Smaller Markets

- Attraction of new customer base
- Distinguishing identity amidst competitors
- Strategic testing ground for wider expansion
- Cultivation of exclusivity and anticipation



Holiday Footfall Boost in Off-Peak Months

- Elevated customer engagement and excitement
- Strategic utilization of the holiday spirit for growth
- Generation of positive word-of-mouth referrals through holiday appeal

Beta (Jun-Jul)

Growth (Aug-Sep)

Refinement (Oct-Nov)

Objective

Gather performance data of initiatives in Pilot Markets

Key Activities

- Test different promotions and themes.
- Monitor visitors' response and engagement.
- Analyze data and identify key success factors and areas of opportunity.

Goal

8,500 visitors / month

EVALUATION

Objective

Accelerate visitors acquisition in Expansion Markets

Key Activities

- Tailor initiatives to adapt to the market's characteristics and preferences.
- Track customer engagements and acquisition rates.
- Compare customer's behavior in bigger markets.

Goal

10,500 visitors / month

EVALUATION

Objective

Optimize initiatives for continuous growth in Mega Market

Key Activities

- Enhance initiatives based on previous stages' feedback.
- Evaluate customer feedback and make micro-adjustments.
- Analyze overall impact of the strategy.

Goal

12,500 visitors / month

EVALUATION

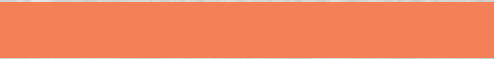
Additional data that should be considered to replicate this strategy in **Monterrey...**





Business Intelligence Challenge

Didi Questions



How many channels can you think of downloading a DiDi Rides APP and how will you estimate the quality and cost of each channel?

| channel | costs | efficiency metric |
|----------------------|-----------------------------------------------------|----------------------------------------------------|
| App Stores | App Store Paid Promotions | App Ranking App Impressions Conversion Rates |
| Social Media Ads | Cost per Click | Click-Through Rate |
| Billboards | Placement Cost | Traffic Volume |
| Email | Email Marketing Platform Subscription | Open Rate Click-Through Rate |
| SMS | SMS Marketing Platform Subscription Cost per SMS | Open Rate Click-Through Rate |
| Users Referral Codes | Incentives Costs | Conversion Rate |
| Influencer Marketing | Influencer Fees | Conversion Rate Engagement Metrics |
| QR Codes | QR Marketing Platform Subscription | Conversion Rate Scan Rates |
| Internet Search | Search Engine Optimization (SEO) | Conversion Rate Search Ranking |
| Physical Events | Rent Fees Operational Spend | Attendance Interaction Rate |

We want to build up a model to predict “Possible Churn Users” for DiDi Rides APP (e.g.: no trips in the past 4 weeks). Please list all features that you can think about and the data mining or machine learning model or other methods you may use for this case.

| Features | Description |
|----------------------|-----------------------------------------------------|
| account_age | Days since the user's registration. |
| last_trip | Days since the user's most recent trip. |
| total_trips | Total number of trips taken. |
| monthly_trips | Average number of trips taken monthly. |
| weekly_trips | Average number of trips taken weekly. |
| trip_duration | Average duration of the trips in minutes. |
| trip_distance | Average distance of the trips in kilometers. |
| fare | Average fare price of the trips. |
| ratings | User's average rating score. |
| location | User's location. |
| payment | User's preferred method of payment. |
| promos | Total number of promotional codes redeemed. |
| support_interactions | Total number of interactions with customer support. |

We want to build up a model to predict “Possible Churn Users” for DiDi Rides APP (e.g.: no trips in the past 4 weeks). Please list all features that you can think about and the data mining or machine learning model or other methods you may use for this case.

Classification Models

Support Vector Machine

Can identify linear and non-linear patterns.

Random Forest

Can handle different data types such as numerical and categorical data

Logistic Regression

Can analyze the direct impact of individual features.

Business Intelligence Challenge

SQL QUERIES

1. Write the SQL queries necessary to generate a list of the five restaurants that have the highest average number of visitors on holidays. The result table should also contain that average per restaurant.

SQL Query

```
SELECT TOP 5 id, genre_name, area_name, AVG(reserve_visitors) AS 'avg_visitors_holidays'
FROM restaurants_visitors
JOIN date_info ON calendar_date = visit_date
JOIN store_info ON store_id = id
WHERE holiday_flg = 1
GROUP BY id, genre_name, area_name
ORDER BY 'avg_visits_holidays' DESC
```

Results

| id | genre_name | area_name | avg_visitors_holidays |
|------------------|---------------|--------------------------------------|-----------------------|
| e89735e80d614a7e | Izakaya | Tōkyō-to Chiyoda-ku Kudanminami | 8 |
| db80363d35f10926 | Dining bar | Hokkaidō Asahikawa-shi 6 Jōdōri | 6 |
| bb09595bab7d5cfb | Izakaya | Niigata-ken Niigata-shi Teraohigashi | 5 |
| e053c561f32acc28 | Izakaya | Hokkaidō Asahikawa-shi 6 Jōdōri | 5 |
| 24b9b2a020826ede | Japanese food | Fukuoka-ken Kitakyūshū-shi Ōtemachi | 5 |

2. Use SQL to discover which day of the week there are usually more visitors on average in restaurants

SQL Query

```
SELECT day_of_week, AVG(reserve_visitors) AS 'avg_visitors'
FROM date_info
JOIN restaurants_visitors ON visit_date = calendar_date
GROUP BY day_of_week
ORDER BY 'avg_visitors' DESC
```

Results

| day_of_week | avg_visitors |
|-------------|--------------|
| Friday | 4 |
| Monday | 4 |
| Thursday | 4 |
| Tuesday | 4 |
| Wednesday | 4 |
| Saturday | 3 |
| Sunday | 3 |

3. How was the percentage of growth of the amount of visitors, week over week, for the last four weeks of the data? Use SQL too

SQL Query

```
WITH CTE AS (  
    SELECT TOP 5 DATEPART(wk, visit_date) AS week_num,  
        SUM(reserve_visitors) AS total_visitors,  
        ROW_NUMBER() OVER (ORDER BY DATEPART(wk, visit_date) DESC) AS rn1  
    FROM restaurants_visitors  
    WHERE YEAR(visit_date) = 2017  
    GROUP BY DATEPART(wk, visit_date))  
  
SELECT T1.week_num, T1.total_visitors, ((T1.total_visitors - T2.total_visitors) * 1.0 / T2.total_visitors) * 100 AS 'percentage_change'  
FROM CTE T1  
LEFT JOIN CTE T2 ON T1.rn1 = T2.rn1 - 1;
```

Results

| day_of_week | avg_visitors | percentage_change |
|-------------|--------------|-------------------|
| 22 | 15 | -76.19 |
| 21 | 63 | -19.23 |
| 20 | 78 | -55.42 |
| 19 | 175 | 6.70 |
| 18 | 164 | NULL |

Business Intelligence Challenge

repository: bit.ly/BI_Challenge

