

# DiDi Challenge: Bi Analyst

Juan Guillermo Brewster



## Strategies to increase sales by two times.

- Opening/Expanding popular food genres Bars/Cocktail, Dining bar, Japanese Food and Izakaya into the zones with the most average visitors like Hokkaidō Asahikawa-shi 6 Jōdōri, Fukuoka-ken Kitakyūshū-shi Ōtemachi and Hokkaidō Sapporo-shi Minami 3 Jōnishi. These zones concentrate roughly half of the total visitors and have an average of 4 and higher.
- Implementing offers for new clients like discounts or points and create loyalty programs for frequent customers. Also market this programs through the biggest social media platforms in Japan (LINE, Instagram, Twitter). Creating special events for Saturday and Sunday which are the lowest average day with visitors and offering extra discounts on Bars and which represent the highest average of visitors on weekends. While on holidays boosting the Sweet type of restaurants.

## Extra data to have a more in depth detail.

- Customer demographics like age, gender and income, this way we can go into a higher level of granularity within the data. This will be helpful when creating custom marketing campaigns since we know the target audience better.
- Transportation data to better understand the flow of people through the city and know which specific spots are more prone to gather people.
- More info on local events e.g. festivals, concerts, conventions.
- Social media data to understand review sentiments and track any unwanted situations.
- Survey data from specialized vendors like Qualtrics/Medallia for increasing customer experience.

# DiDi channels and their cost/quality.

- Social media platforms like Facebook, Instagram, Twitter, Youtube. I'd measure the quality by the reviews, engagement and sentiment analysis, costs include ads and headcount to manage these platforms.
- App stores. The quality is measured by the reviews and number of downloads. There are more engagement metrics that can be tested with Hypothesis testing. App store fees and new developments by software devs.
- Email. Conversion rates (from opening the e-mail to actually downloading the app) and the costs are associated with content creation done by graphic designers.
- Whatsapp / Text message. Just like e-mail, conversion rates and costs are more related to server processing/storage and API consumption to send the messages.
- Referral programs. Can be measured by the ratio of referrals that actually downloaded the app and registered by the amount of referrals. Costs associated with incentives.
- Non-digital ads like billboards. These can be tracked by surveys. The costs are associated with ad rent and materials.

# Model to predict customer churn

The target variable for the classification model would be “passangers with no trips in the past 4 weeks”. This is considered a passager that churned. It could have the value 1.

The features to look at are:

1. User demographics like Age, Gender and Address
2. User activity like number of trips taken this month and last weeks, time since last trip and account age, most frequent trip locations, trip distance, trip time, promotions used and trip cost.
3. User experience interactions like reviews and reported issues with taxi driver, payment or app.
4. External factors like competition and seasonality.

Models to test with the following tecniques: (ROC Curves, AUC Score, Recall and Precision, Cross Validation)

- “Explainable methods” like: Logistic Regression and Decision trees.
- “ Black box methods” like: SVM’s, ensamble methods like XGBoost and Random forests.