1. Introduction

1.1 Company's Background

Eatigo is a restaurant reservation platform that was founded in 2013 in Bangkok. Its mission is to connect restaurants with customers by offering time-based deals to users daily at all participating restaurants through its online website and mobile application. Users can choose to dine in or take away from upscale hotels to popular food chains and enjoy time-based discounts of up to 50%, while restaurants get to fill empty seats during off-peak hours.

Eatigo operates on a commission-based revenue model. They would receive commissions according to a tiered system (higher tiers give a higher percentage of commission) based on the restaurant's revenue.

1.2 Problem Statement

With an increasing number of restaurants opening in the Food and Beverage industry, Eatigo aims to expand the number of restaurants to bring onboard their platform. Thus, a model is required to help Eatigo predict and assess future restaurants' potential performance and subsequently decide on the restaurants to onboard. The result of this model is to establish a metric called Lead Score which aggregates and summarizes various information into a single data point to convey the overall quality of the merchant.

This process involves assigning numerical points from a range of 0-6 to potential restaurants that are not currently on Eatigo's platform. This would be based on 4 attributes, namely, tier, restaurant type, cuisine and neighborhood that correlate with revenue. In addition, Quality Index (QI) – relative restaurant revenue compared to average country revenue, is also taken into account as a contributing factor of the Lead Score and Revenue model.

1.3 Objectives

The project seeks to build a predictive Lead Score and Revenue model that will help Eatigo's Sales Team prioritize their leads and identify potential merchants which Eatigo can bring onboard their platform. Ultimately, the model aims to create a more efficient and impactful work process to optimize the operations of Eatigo.

2. Performance Measures & Evaluation

The performance measures are the Lead Score for each restaurant and the potential revenue that Eatigo can generate from each restaurant. This is derived from mathematical computations between revenue and each attribute as well as QI. The relationship between attended reservations and each attribute is not calculated as revenue and attended reservations are already strongly correlated.

The Lead Score formula for each restaurant equals SUMPRODUCT (all attributes point and QI: all attributes and QI weight). A maximum Lead Score of 6 is capped as Total QI can go to infinity based on the revenue, thus making the values more readable and understandable. Different Lead Score range corresponds to a different status as shown in Table 1, but users have the flexibility to adjust accordingly based on their business needs:

Lead Score Range	Status
0-1	More research is needed if Eatigo decides to bring the restaurant onboard
2-3	Consider whether to bring the restaurant onboard based on factors like having a good mix of restaurant attributes
4-6	Onboard the restaurant

Table 1: Classification of Lead Score

The consequence variables used in the computation of Lead Score are shown in Table 2 and the weights for 4 attributes and QI were adjusted to total up to 100%:

Variables	Points	Weights
Tier Attribute	• Attribute points are calculated based on the relative relationship between the standard deviation (S.D.) of revenue values • S.D. is used instead of average to gauge the variability in the raw data • The largest S.D. is to be set as the base for points allocation because a large S.D. has data that is more spread out, providing a more comprehensive analysis	• Dummy variable used to convert categorical variables (e.g., FTR, STR, TTR) in raw data to numerical values, that take on the value of either 0 or 1 for regression analysis • k-1 dummy variables are created where k = total number of categories and the baseline is decided from the previous points calculations

		• After which, correlation is obtained from R and converted to a percentage as the weight between Tier and Revenue
Restaurant Type Attribute	• Same steps as those for Tier attribute	• Same steps as Tier attribute
Cuisine Attribute	• As regression analysis in Excel has a limit of inputting 16 variables, and there are a total of 44 cuisine types in the raw data, reclassification of cuisine is needed • Cuisine types with lesser than 150 counts of restaurants are considered small sample size and are reclassified under "Others" • After which the subsequent steps are the same as those for Tier attribute	• Same steps as Tier attribute
Neighborhood Attribute	• As regression analysis in Excel has a limit of inputting 16 variables, and there are a total of 138 neighborhood areas in the raw data, reclassification of neighborhood is needed • With reference to the Urban Redevelopment Authority map, neighborhood areas are reclassified into 6 zones namely Central, West, East, South, North and North-East • After which the subsequent steps are the same as those for Tier attribute	• Same steps as Tier attribute
QI	• Used as an index to determine the relative revenue of a restaurant compared to the country revenue (sum	• Dummy variables not used as data are already in numerical values • Same steps to obtain R

	of restaurants revenue) onboard Eatigo's platform in the country
	• The total QI point for each restaurant is calculated based on the average QI for each restaurant obtained from raw data (Refer to "Lead Score & Revenue ALL" sheet)

Table 2: Variables used in computation of Lead Score

3. Data Collection & Analysis

The raw data provided is historical data from 2019, which is during the pre-COVID period, that can account for a more accurate representation of actual restaurant functioning capacity for the analysis. Bolded attributes under the Raw Data column in Table 3 are important in the computation and analysis of the Lead Score and Revenue Model. The attributes (Neighborhood Zones, Cuisine Category, Tier and Restaurant Type) and Revenue are used to find the attributes' points and weights. QI and Revenue are used to find total QI and QI weights. After which, the points and weights would be used to derive the Lead Score. Reclassification was done for Neighborhood and Cuisine.

Raw Data	Description
Country_code	Country which the restaurant is located in
Month	Date of revenue and reservations
Restaurant_id	Unique restaurant identifier
Restaurant_name	Restaurant name
Address	Restaurant address
Neighborhood	Neighborhood in which the restaurant is located in
Neighborhood Zones	Reclassification of neighborhood into different zones e.g., Central, North etc.
Coordinates	Restaurant latitude and longitude
Cuisine	Type of cuisine
Cuisine Category	Reclassification of cuisine i.e. cuisines with less than 150 counts of restaurants are reclassified to "Others" and the remaining remains as it is
Tier	First/Second/Third tier Restaurants (based on Average Menu Price)
Restaurant_type	Chain vs Independent vs Hotel Restaurant

Gross_reservations	Number of reservations made
Attended_reservations	Number of reservations attended at restaurant
Revenue	Restaurant revenue in USD
Quality Index	Restaurant revenue relative to average country revenue
Atmosphere (Hidden Sheet)	Atmosphere attribute for the restaurant; the same restaurant can have multiple atmospheres therefore multiple rows

Table 3: Raw data required for analysis

4. System Scope & User Functionalities

The contents in the column 'Areas covered by system' represent each sheet in our model.

Areas covered by system	User Functionalities
Introduction	• Purpose of Lead Score and Revenue Model and contents page
Data Description	• Description and purpose of each sheet • Definition and explanation of data in the raw data table. E.g., 'restaurant_type' refers to Chain, Independent or Hotel Restaurant
Raw Table (Insert)	• The uncleaned version of raw data
Raw Data Table 1	• Presented in table form consisting of numerous columns such as date, restaurant details, various attributes, QI etc. with filters applied • Additional columns of recategorized attributes as compared to the uncleaned version
Raw Data Table 2 (hidden)	• Consist of restaurant id, name and atmosphere but hidden as it was not used for analysis
Points Calculation	• Computation of attributes points using raw data obtained from Raw Data Table 1 sheet
Weights Calculation	• Computation of attributes and QI weights using raw data obtained from Raw Data Table 1 sheet
Points & Weights Summary	• Table that summarizes the correlation, adjusted weights and attributes points
Lead Score & Revenue Model	• ***Top 10 Lead Score section:*** Restaurants with the top 10 lead score will be automatically displayed based on the "LARGE" formula, with details as follows:

(Top 10 Restaurants)	➤ Attributes, attributes' points & weights, total QI and QI weights, Lead Score and restaurant's revenue ➤ There are 3 types of status for each restaurant based on the Lead Score, namely "Onboard", "Consider" and "Research". Users have the flexibility to change the Lead Score range based on business needs. ➤ Eatigo's potential revenue is calculated based on the tiered commission percentage. The commission percentage for each tier is an assumption made by the team as information was not revealed to us. Users have the flexibility to adjust the percentage accordingly. *Display of Lead Score for Selected Restaurant section:* Users can choose restaurants from the drop-down list and the respective details will be displayed *Charts section:* Displays 3 charts – 1) Top 10 Lead Score, 2) Selected Restaurants trade-off comparing between Lead Score and Eatigo's revenue, 3) Compare earnings for Eatigo between restaurants *Refresh button* is a Macro set to update the whole Excel when new raw data are added with the following steps: 1. Copy and paste data from Raw Table (Insert) into respective columns/cells into Raw Data Table 1 2. Update pivot table in Points Calculation sheet 3. Filter all pivot tables, taking out (Blanks) and NA for zone areas 4. Update new S.D. value of "Others" under cuisine table 5. Clear all contents in Weights Calculation sheet 6. Prepare all the values needed for regression analysis - Copy and paste revenues from raw data to specific columns in Weights Calculation sheet - Copy and paste other necessary fields such as tier, restaurant type, cuisine, neighborhood, QI from raw data - Prepare dummy values required for categorical regression 7. Goes through regression analysis under Weights Calculation sheet for all 4 attributes and QI against revenue

	8. Deletion of unnecessary data, leaving only regression statistics 9. Update cuisine types in Points & Weights Summary sheet 10. Update restaurant names in Lead Score & Revenue All sheet corresponding to the raw data tab 11. Update the lead rank formula to follow the size of the data 12. Update the rank revenue formula to follow the size of the data 13. Update the correlation formula to follow the size of the data
Lead Score & Revenue Model (All Restaurants)	• Same layout as Lead Score & Revenue Model (Top 10 Restaurants) sheet with 3 additional columns ➢ Adjusted Lead Score (displays as Lead Score in model sheet): Maximum lead score is 6. ➢ Lead Rank: Rank the Lead Score with 1 being the top score ➢ Revenue Rank: Rank Eatigo's revenue generated from each restaurant with 1 being the highest revenue
Overall Dashboard	• Displays charts of Count and Average Revenue of Tier, Neighborhood, Cuisine, Restaurant Type and Lead Score rank vs Revenue rank

Table 4: System scope and user functionalities
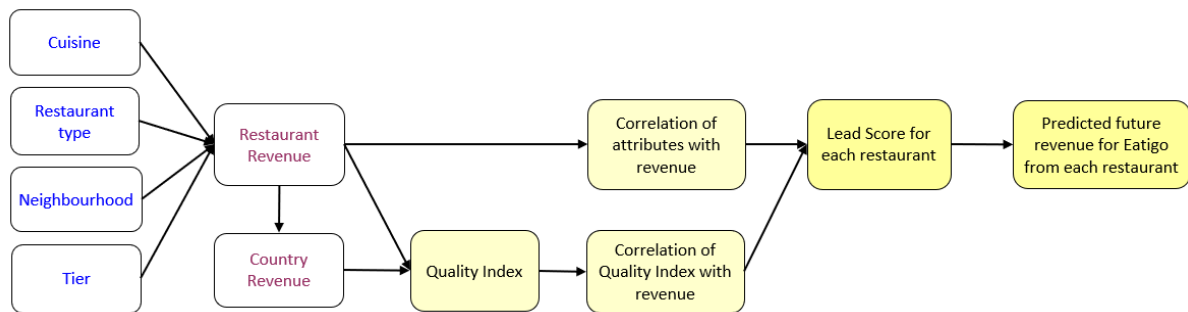
5. Model Sketches

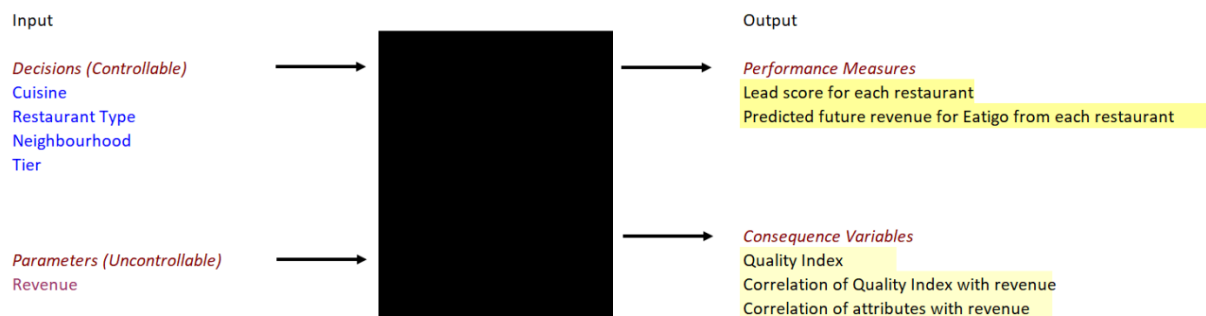

Figure 1: Influence Diagram

Figure 2: Black Box Model

6. Scenarios & Trade-offs

As mentioned above, numerical points ranging from 0 to 6 are assigned to help Eatigo assess restaurants and determine their potential. Hence from the Lead Score, Eatigo will be faced with mainly 3 scenarios and their subsequent courses of action to take.

Firstly, restaurants with a Lead Score of 4 to 6 would indicate a high performing restaurant with high revenue. Thus, it would be highly recommended for Eatigo to bring these restaurants onboard their platform as it would be highly profitable to do so. Next for restaurants with a Lead Score of 2 to 3, the revenue earned from the commission would not be as high compared to restaurants in the previous category. Eatigo could further consider whether to onboard the restaurant based on factors apart from revenue such as variety. Finally, it is recommended that restaurants with a low Lead Score of 0 to 1 require additional research from Eatigo to ascertain if they should be brought onboard the platform.

Furthermore, trade-offs are involved when deciding scenarios. For instance, if only restaurants with high Lead Scores are chosen to onboard with Eatigo, there is a risk of Eatigo losing variety. This is due to those restaurants having similar attributes which correlate to higher revenues and therefore, higher Lead Scores. Eventually, this may cause Eatigo to lose customers and profits instead because the options lack diversity, translating to fewer choices for customers, deflecting them away from the platform. In the long run, bringing restaurants with high Lead Scores onboard may not be the optimal decision as introducing those with lower Lead Scores might bring something different to the table. This is a trade-off that the user must bear in mind while making decisions based on the model.

7. Decisions & Results

After analysis of the raw data provided by the team's client sponsor, charts have been designed to provide the company with a summarized version of the data. This is to depict the results for each of the attributes (tier, restaurant type, cuisine, neighborhood), as well as any subsequent results derived from the primary attributes post-analysis. The dashboard aims to provide the client with an overall visualization of the data and provide any crucial information they need at a glance to make more efficient and effective decisions.

There are many ways in which the dashboard can be used by the company. For instance, the platform might be facing the issue of a lack of diversity. How the company can utilize the dashboard to counter this is through the count of the various attributes provided in the dashboard. The count helps the company identify the attributes the platform is lacking in a quick and efficient process. Following this, they can proceed to onboard restaurants with attributes that the platform is lacking.

Figure 3: Dashboard Charts

8. Model Limitations

Firstly, there is an unequal sample count of restaurant data for each attribute type (e.g., FTR vs STR vs TTR) when comparing the average and S.D. of revenue within each attribute. This could have caused unequal/less accurate comparison when deriving the points and weights used for the calculation of Lead Score.

9

Secondly, relating to the first limitation, the data set was not normalized to a common scale before performing the calculation of average and S.D. hence variability is high.

Thirdly, generic categories were used (e.g., zone areas of Central, East, West etc.) instead of specific categories (e.g., neighborhoods of Orchard, Expo, Jurong etc.) as Excel has a limit of 16 variables for regression analysis to find the correlation. This limitation on Excel makes it difficult for a more accurate analysis of specific variables.

Lastly, the model only accounted for 4 attributes (tier, neighborhood, cuisine, restaurant type) and other possible attributes such as online ratings were not accounted for which could have affected the Lead Score.

9. Lessons Learnt

Firstly, problems faced in real-life business environments are dynamic. What we have learnt in this course is only a small portion of how Excel functions can be used to create certain models to solve problems. Many problems out there would require exploration of other functions and techniques (like our project) to find suitable ways of solving them. Additionally, for complex scenarios and more accurate analyses, Excel might not even be the best tool to perform it. It is important to diversify yourself with the tools and software available in the market.

Secondly, another takeaway would be the importance of alignment and communication of expectations towards this project/model between the client sponsor and the team. Although the team has communicated to the client sponsor at the start of this project that this model would be a simple prototype based on the techniques learnt in class, we only realized when nearing the project dateline that the client sponsor had assumed that the team would be well versed with statistical concepts (i.e., S.D., correlations, coefficients, and different types of regressions) and was expecting an elaborate statistical model. The area we could have improved on is to share in greater detail the functions, techniques, and scenarios we have been taught in class so that expectations from the client would be more aligned with our knowledge and skills.

10. Conclusion

Our team has managed to build a predictive Lead Score and Revenue model that seeks to help Eatigo identify potential partners to be onboarded onto their platform despite the challenges faced. Although there are limitations to the model, it still serves the general function of assisting Eatigo in its assessment of restaurants via the Lead Score and Revenue metrics. Through this

project, our team has obtained takeaways such as improved communication skills and exposure to new analysis techniques.

External party contact details: Ms Judy Tan, Co-Founder of Eatigo, jt@eatigo.com