Forecasting the revenue generated by ATM / CARD / CASH for EZTABLE to identify the potential / actual revenue from different payment systems.

Team #6
Member


Sean, Xie
103078517


Sam, Wang
103078502


Lydia, Jiang
103078511


Suneel
10307882

Summary

The primary stakeholder is EZTABLE, an ecommerce based platform for reservations for different restaurants in Taiwan**.** Currently, there is a mismatch between the number of bookings and the amount of revenue generated by the bookings. The problem is the current booking system which enable the user to book reservations even without prior payment. Hence the actual revenue falls short of the potential revenue due to the existing system. The business goal of our project is to let our client realize the potential revenue and the need for changing the strategies. At the same time, the mode of payment is changing its trend from CASH to ATM. It gives a clear indication that the consumer habits are changing rapidly. Undoubtedly, this is the right time for our client to take some strategic decisions which can help it to achieve the potential revenue.

The data consists of the purchases information for EZTABLE. The time period is from 2012 January to 2014 September and the number of data is about 220,000. We handle the missing value and outlier value. We aggregate data into monthly one and partition it. All we do before forecasting is to make our forecasting more accurate. Since the revenue has both trend and seasonality, we have used the methods which can take care of the both.

We choose the multiple linear regression to do the forecast because the data has polynomial trend and seasonality. The input values are T, T square, seasonal index and the output value is payments. Our client can use the forecasting model to predict the actual revenue and potential revenue.

We suggest EZTABLE to re-design the payment systems which is more convenient for customer to pay. To make sure the reservation is successful, our client can do targeted marketing such as if they use the booking then they might receive some offer, etc. In this way, EZTABLE can catch up potential revenue by increasing its actual revenue.

Business goal

The primary stakeholder is EZTABLE, an ecommerce based platform for reservations for different restaurants in Taiwan**.** There are two issues associated with the extant approach of our client EZTABLE. First is that actual revenue is not equal to total revenue mentioned in the purchasing data. Because we found column status including three types of value which are the paid / canceled / new from the EZTABLE purchase row data. The actual revenue that EZTABLE may earn is only from the status of paid. And there has the gap between the number of booking and number of booking which are really paid. Then we also found the revenue from the CASH is going down and revenue from the CARD and ATM is going up very fast. It means that the customer consuming habits might change. We thought it is the opportunity of our client to doing some strategies.

Forecasting goal

The forecasting goal of our project is to forecast the potential revenues and actual revenues generating by different payment systems for the next quarter (three month period). To be specific, the actual revenue would consists of the revenue which consists of the paid bookings only. The potential revenue would consists of the paid booking, new booking, and canceled booking. The three month forecast horizon gives enough time for our client to make a better strategy.

Data description

The data comes from our client - EZTABLE, and we got it from Professor. We choose the purchasing table which contain the purchase of customers though system from 2012 January to 2014 September.

In the table, there is a column named “status” including new, paid, and canceled. We consulted the company directly to understand the meaning of the “status” field. It means when the order is made, the system create a record with status “new”. If the customer doesn’t pay for it, the status always be new. Therefore, the new and the canceled ones don’t generate the revenue for EZTABLE on fact. We separate it into the “paid” time series and “all” time series including new, paid, and canceled to forecast the actual and potential revenue in several.

We observe that there are missing data in the table like the data on 2014 September. We want to generate the monthly forecast value, but the table doesn’t have completed data. We delete the data on last month because it will affect our forecasting result.

Data preparation

We put data on Tableau and see that CASH, CARD, and ATM are the most popular payment and they have enough period for forecasting (figure 1). EZTABLE also generates lots of revenue by yahoo payment, but it’s the new way to pay, so it has little period of data. We choose CARD, CASH, and ATM as our forecasting target eventually.

First, we find that it have also some extreme data which revenue are NTD 99,999 or NTD 570,035. We delete it after confirming that they are just the testing value by asking the employee working in EZTABLE.

Next, we aggregate the data into monthly data. The reason is that the monthly data is more helpful for EZTABLE to do marketing and some strategies, and it can remove the missing data on some days which don't have any transaction by specified payment that day.
Then, our forecast horizon is 3 months, so we partition the data and set 3 months for validation period to verify the model to choose the best one to forecast.
Eventually, we can find our time series have level, noise, trend and yearly seasonality by observing the plots. To observe the seasonality, we split it into three year. We take "CARD" time series for example (figure 2). Although the plot doesn't show the obvious seasonality, we can see that there are the low points on February and November and the peak on May. The peaks and low points are on the same month. Maybe the reason is that people celebrate the mother's day on May. We should choose the method which has ability catching both trend and seasonality.

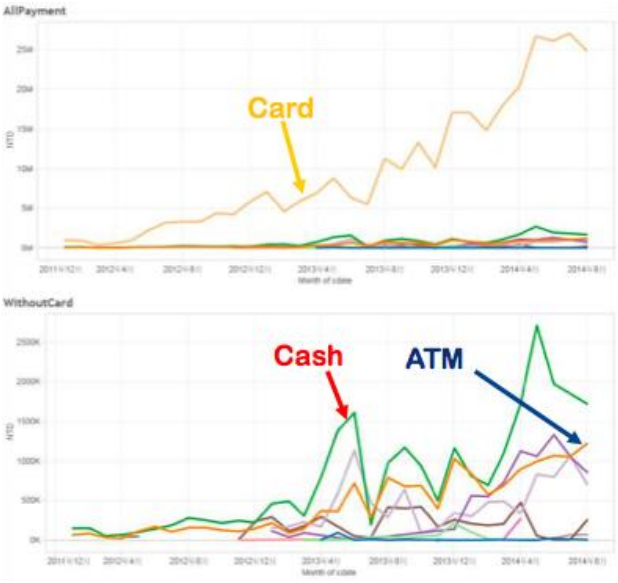

Figure 1 - sum of revenues

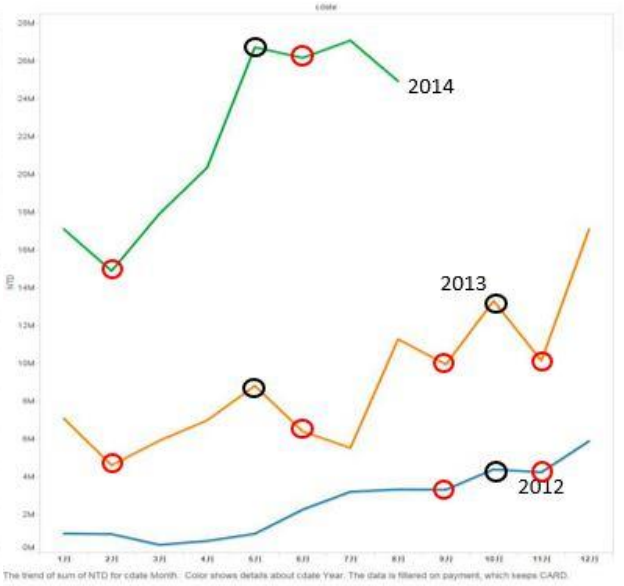

Figure 2 – CARD's yearly

Forecasting solution

We consider that every payment have trend and seasonality. So we first create some derived variables like T, T square, seasonal indexes, dummies and log. Then we partitioned each series into training data and validation data. We made last 3 months as our validation data and the remaining data as training data. And we used several methods by XLMiner. For examples, seasonal naive, holt winter (additive trend), neural network, multiple linear regression (log, seasonal index), multiple linear regression (T, seasonal index), multiple linear regression (T, T square, seasonal index).The chart of comparing residuals and the Actual v.s. Forecast is shown below.

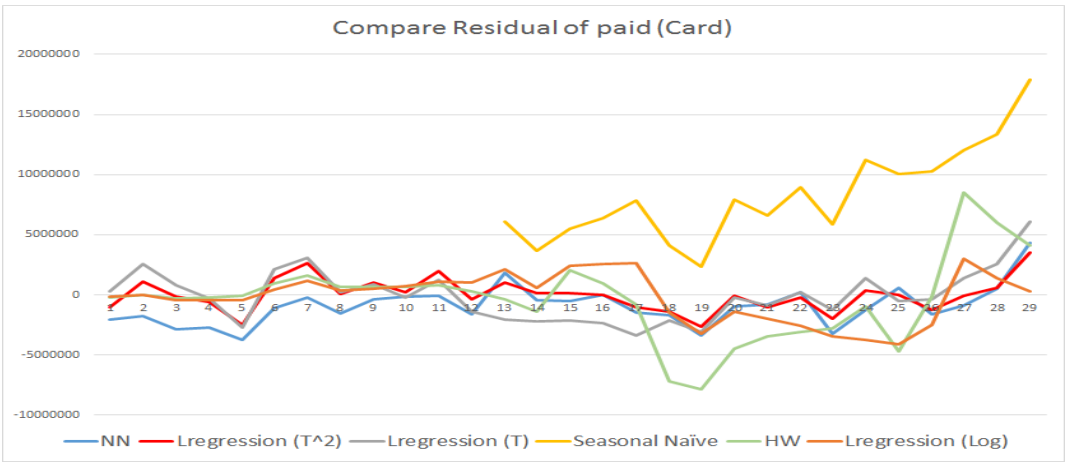

Figure 3

We could see that we can't obviously choose which one is the best with the Actual v.s. Forecast chart (figure 5 in the appendix). But with the residual chart (figure 3, table 1 in appendix), we can figure out that the residual range of the red line (Multiple linear regression d-seasonality d-trend) is the smallest one. So we use the multiple linear regression (T, T square, seasonal index) to be our forecasting method.

And the other two series, CASH/ATM, also used the same way to do the comparison. The result is that we found CASH/ATM have better performance of multiple linear regression (T, T square, seasonal index) also. So we decide to use this method and input value to do the three months ahead forecasting by XLMiner. And the chart of each payment are shown below and all the payments have two series.

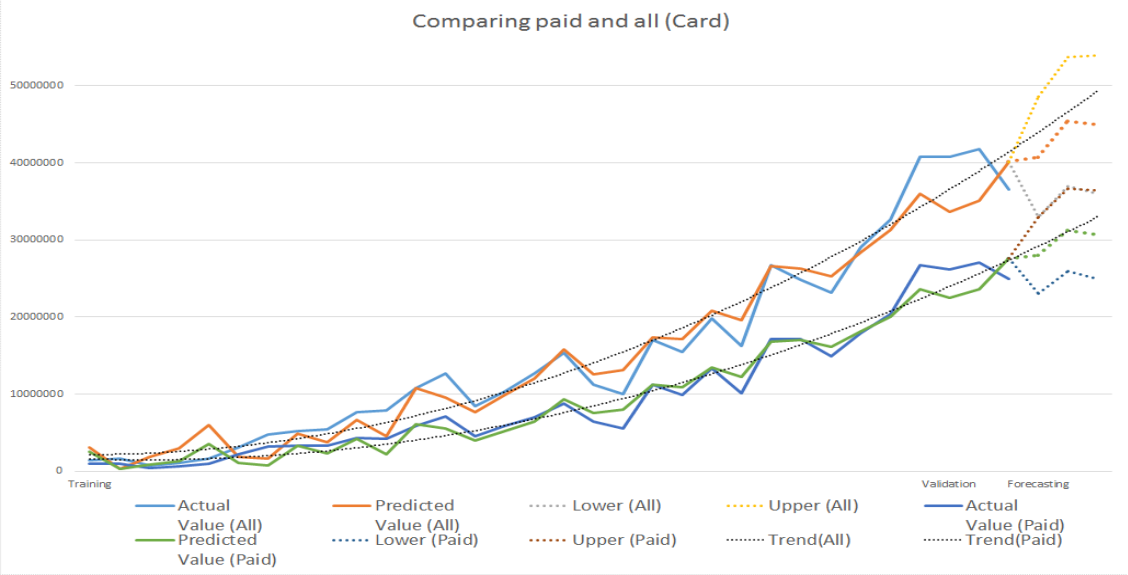

Figure 4

As the graphics, we can figure out that CASH and ATM payments didn't have the obviously gap between all and paid (figure 6, table 2, figure 7, table 3 in appendix). Which means this two payments don't have the problem we just mentioned. And we discover that the gap of CARD payment is growing up every year (figure 4). On the other hand, the revenue of this two payments are relatively small. So we suggest EZTABLE to focus on the CARD payment systems. Thus, the following recommendation is about to make the CARD payment.

Recommendation

EZTABLE can use this model to generate the forecasting of both actual and potential. We suggest EZTABLE can uses these forecasts to make different strategies of promotion. Actual revenue can be used to show investor about their operating condition, and EZTABLE can use it to get the accuracy value to make the right decision. The potential one can be used to persuade restaurants to join them.

Then we find that the different between actual and potential revenue is huge. There are maybe some way to make system more convenient to pay. It can help them make much more revenue.

Appendix

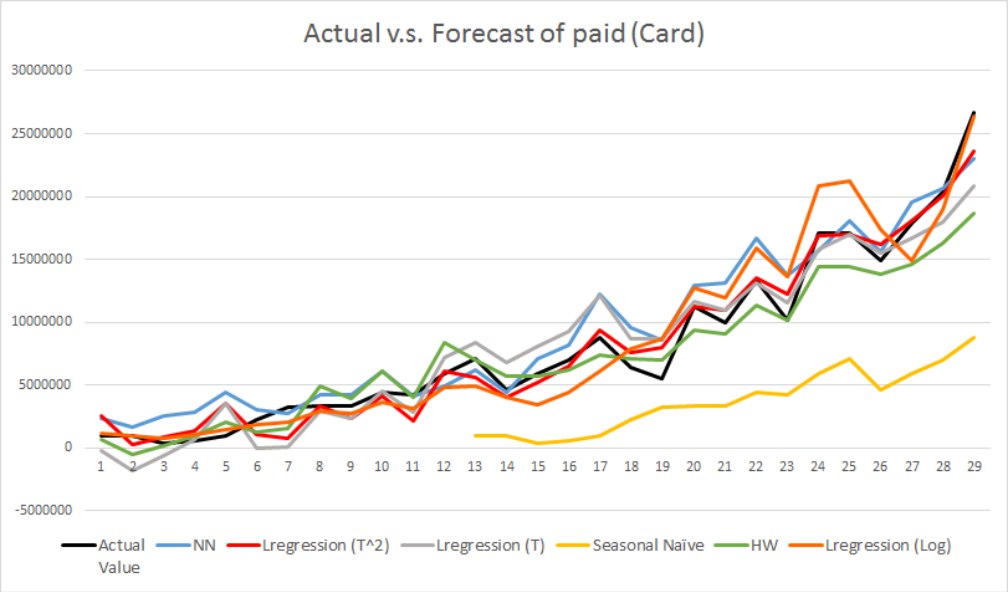

Figure 5

	Total sum of squared errors	RMS Error	Average Error
CARD All	1.25192E+14	2077727.121	1.60573E-09
CARD Paid	5.08679E+13	1324411.476	1.57361E-09

Table 1

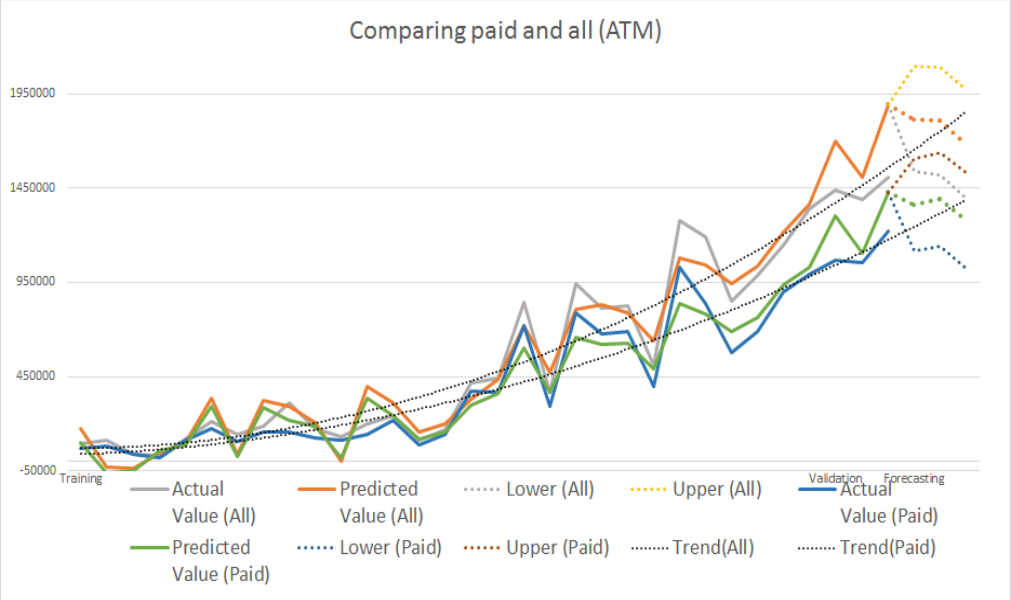

Figure 6

	Total sum of squared errors	RMS Error	Average Error
ATM All	2.86283E+11	99357.07279	4.21504E-11
ATM Paid	2.38889E+11	90760.98	-4.41575E-11

Table 2

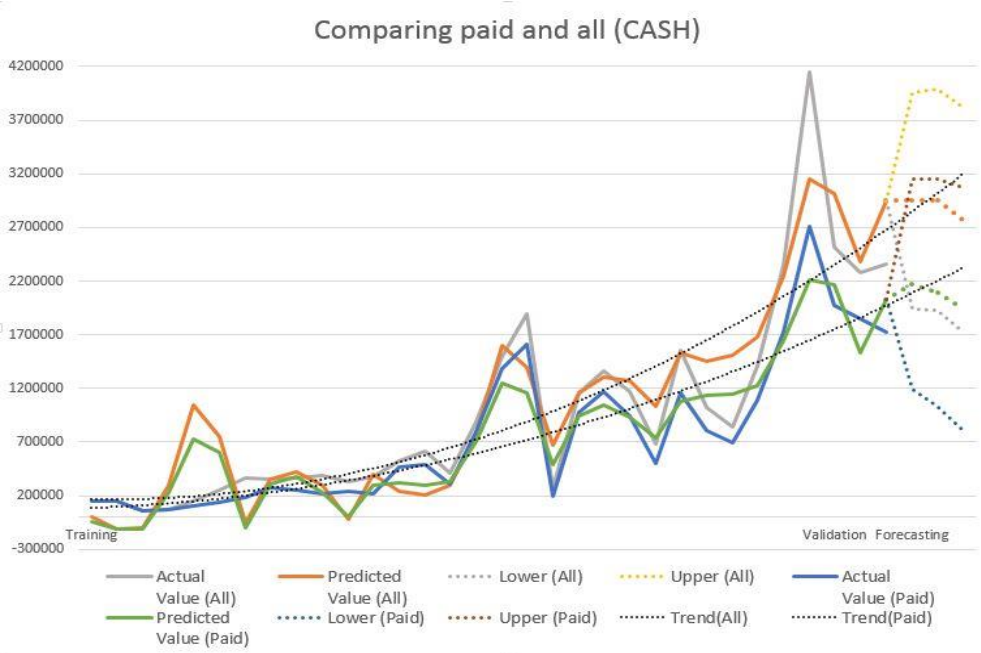

Figure 7

	Total sum of squared errors	RMS Error	Average Error
CASH All	4.07111E+12	374677.3517	0
CASH Paid	1.98596E+12	261689.4087	-4.41575E-11

Table 3