PREDICTIVE ANALYTICS/RFM FOR THE GAMING INDUSTRY

JANUARY 2014

Applications for Predictive Analytics

- Cross-sell/Up-sell – 47%
- Campaign Management – 46%
- Customer Acquisition – 41%
- Budgeting and Forecasting – 41%
- Attrition/churn/retention – 40%
- Fraud Detection – 32%
- Promotions – 31%
- Pricing – 30%
- Demand Pricing – 30%
- Customer Service – 26%
- Quality Improvement – 25%

*Based on 167 respondents who have implemented predictive analytics solutions. Respondents could select multiple answers.

Predictive Analytics can help to:

- Identify – the casino's most valuable patrons.
- Predict a patron's future worth and/or his or her future behavior.
- Plan the timing and placement of advertising campaigns.
- Create personalized advertisements.
- Define which market segments are growing most rapidly.
- Segment patrons into groups based on their behaviors and then create marketing campaigns to exploit those behaviors.
- Determine a patron's level of gambling skill.
- Identify patrons who come together.
- Identify the likelihood a patron will respond to an offer
- Identify the offer(s) to which patrons are most likely to respond to.
- Predict when a patron is likely to return.

Manipulating Customer Behavior

Successful marketing is about reaching a consumer with an interesting offer when he or she is primed to accept it. Knowing what might interest a patron is half the battle to making a sale and this is where customer intelligence and predictive analytics comes in.

Customer analytics has evolved from simply reporting customer behavior to segmenting customers based on their profitability to predicting that profitability, to improving those predictions, to *actually manipulating customer behavior* with target-specific promotional offers and marketing campaigns.

Predictive Analytics – Survival or Duration Analysis:

A branch of statistics involves the modeling of time to event data; in this context, death or failure is considered an 'event' in the survival analysis literature – traditionally only a single event occurs, after which the organism or mechanism is dead or broken. Survival Analysis is the study of lifetimes and their distributions. It usually involves one or more of the following objectives:

- To explore the behavior of the distribution of a lifetime.
- To model the distribution of a lifetime.
- To test for differences between the distributions of two or more lifetimes.
- To model the impact of one or more explanatory variables on a lifetime distribution.

By applying survival analysis to revenue management models, casino operators can gain a truer picture of their table games revenue.

Predictive Analytics – Marketing Campaigns

By utilizing data from past campaigns and measures generated by the predictive modeling process, casino operators can track actual campaign responses versus expected campaign responses, which can often prove wildly divergent.

Additionally, casino operators can generate upper and lower 'control' limits that can be used to automatically alert campaign managers when a campaign is over or underperforming, letting them focus on campaigns that specifically require attention.

Predictive Analytics – Pattern Discovery
Cluster Analysis

- **Behavioural segmentation**: look at the behaviour of customers (can be across different metrics/landmarks) and attempt to group "like" customers together.
- **Occasion segmentation**: interested in identifying the needs and desires of customers on various occasions. Could have potential for seasonality, day of week analysis.
- **Benefits segmentation**: customers are segmented according to what kind of benefits they are seeking. This may have application in a loyalty/rewards sense for investigating the casino customer's behaviour.

Predictive Analytics – Applications of Segmentation Models

Some applications include target marketing current customers, retention, up sell, cross sell, acquiring new customers but it is certainly not limited to this.

- **Global customer segmentation model:** use of metrics including demographic (e.g. age, income), monetary (spend within the casino), other behavioural metrics associated with a casino (e.g. Visits, tenure, average time between visits…..etc. These could be put into a clustering algorithm to derive like groups of customers.
- **RFM model:** metrics around recency, frequency and monetary value of customers. Traditional segmentation model used as an attempt to assign value to a customer.
- **Specific segmentation models** that identify clear customer behaviour. Some examples of this may be game preference, price point (budget, mainstream, premium) on f&b on ship, cruise type preference, lead time for booking (e.g. last minute, planned….). These could be further used to create cohorts across different metrics.

The Steps in the Analytic Workflow

- Define the Analytic Objective
- Select Cases
- Extract, Validate, Repair & Transform Input Data
- Apply Analysis
- Generate Deployment Methods
- Integrate Deployment
- Gather & Assess Results

The Steps in the Analytic Workflow
1. Define the Analytic Objective

Critical to success of analysis. Without this, can't know if we've got to where we want to be, or in fact how to get there.

Would be defined in light of applications outlined above that are relevant to the casino and the needs of the business.

What are the objectives for the casino?

Case study/Example of Game Preference

The Steps in the Analytic Workflow
2. Select Cases

- Important to identify the customers who are going to be included in the analysis. This would mean rules defined that exclude immature customers as their behavior is volatile.
- Records in the DB need to be unique. Cleaning should ensure duplicate records for customers do not exist.
- Decide whether model should be built on a sample of customers due to volume of data and computational efficiency.
- What is the size of the DB and volume of transactions?

The Steps in the Analytic Workflow
3. Extract Input Data

Inputs need to be:

- meaningful to the analysis: any modelling problem should exclude irrelevant and redundant information.
- limited in number: too many inputs can result in too many clusters making interpretation difficult.
- symmetric in their distribution: prevents orphan clusters resulting due to outlier/influential observations.
- relatively independent: ensures best chance of homogenous groups.
- interval in measurement: requirement for SAS EM.

The Steps in the Analytic Workflow
4. Validate Input Data

This may or may not have been done in the data audit/quality exercise. It is important that missing or erroneous values are not present in the data as this will lead to fantasy results built on false data.

It would also be necessary at this stage to check that the input data meets the requirements from above. This may also involve preparation for step 6. if a different view of the customer is required as part of the segmentation objective.

The Steps in the Analytic Workflow
5. Repair Input Data

Depending on results from 4., some repairing may be required to ensure we have complete case coverage across the metrics. Any cases that do not contain information as required by the selected inputs would be omitted from the analysis.

The Steps in the Analytic Workflow
6. Transform Input Data

If irregular distributions exist across the inputs, it would be appropriate to transform them into a more symmetric and regular distribution. As was previously mentioned, we'd be ensuring that orphan clusters don't result due to highly unusual data values. It might also be necessary here to apply some transformation and manipulation to the raw data.

If a model is required that identifies a clear customer characteristic, e.g. game preference, the idea would be to aggregate customer spend by game type, and transpose the customer behaviour so that a proportion can be extracted for each game. This would then form the inputs for the clustering algorithm.

One major advantage of this approach is that clear delineation amongst customers and their preferences result. It also allows for preferences that don't look obvious to be extracted. As an example, if the population proportion for a particular game is 4%, and a customer has a 9% spend on this game they should be identified as preferential to that game. This approach would do that.

The Steps in the Analytic Workflow
7. Apply Analysis

It is anticipated that a hierarchical clustering approach will be used (k-means) to derive an appropriate number of segments. It should be noted that the statistically best result may not be the best result from a practical standpoint. For instance, the business may have requirements of only seeing x clusters, but the algorithm has come up with x+6 as the optimal solution. Furthermore, there it is a necessity that the analysis produces segments which are meaningful to the analysis objective. Whilst rare in practice, theoretically there wants to be perfect boundaries between the segments identifying unique behaviour. The segments should also be relatively similar in magnitude, unless there are clear reasons and behavioural features that have produced outlier (small) clusters.

The Steps in the Analytic Workflow
8. Generate Deployment Methods

Once the casino is happy with the segments that have been produced and the analysis suggests their needs are in sync with what is derived, the stability of the model would be validated on the holdout sample. Similar distributions across both the modelling and hold out sample ensure that the behaviour that has been identified is of a general nature and the model can be deployed.

This would depend on whether a sample was taken for model building. If so, the remainder of the DB would need to be scored using the segmentation algorithm. If not, each customer (except those omitted due to rules) would be grouped into a segment.

The decision would have to be made as to what the best way forward was for deployment i.e. in the SAS environment or elsewhere.

The Steps in the Analytic Workflow
9. Integrate Deployment

How the deployment is chosen to be done by the casino will have to include a decision of how best to integrate the results into the business.

The Steps in the Analytic Workflow
10. Gather & Assess Results

It is crucial with any analysis to gather results so that the validity of the model can be regularly checked and understood. All models go stale with time as the underlying assumption is that stationarily of the data and its relationships remain consistent. Clearly this is not the case, behaviour changes, as does other external influences and factors. A good segmentation model should see very little perturbation of clusters over short periods of time. It is not a practical result to see customers switching segments month on month. This indicates the behaviour that is being captured is not stable. It is anticipated that the casino may re-score their customers every month and track movement amongst segments. In time, there may be material shifts in the data that require the process to be re-visited and the clustering algorithm updated.

Possible Use of a Segmentation Model
Profiling Customers Using Game Preference

The casino would identify the games of interest to them from a marketing/profiling/profitability point of view. It might be that all games are included, or instead a subset might be chosen with the remainder excluded or a catch all group included.

Historical information for customers would be extracted e.g. Last 24 months and spend aggregated by game. For each customer, this would then be transposed into a single record and proportions derived, as shown below:

Customer	Baccarat	Roulette	Blackjack	Slot Machine	Total
xxxx	$12,000	$1,100	$150	$2,286	$15,536

Customer	Baccarat	Roulette	Blackjack	Slot Machine
xxxx	77.2%	7.1%	1.0%	14.7%

Possible Use of a Segmentation Model
Profiling Customers Using Game Preference

One feature of using this approach is that as many segments/clusters as games are likely to be the optimal result so that a distinct segment relates to each game. This can have positive ramifications for the business as it makes it very easy to target a particular game/segment.
It would be expected that some games would be more profitable to the house than others. If customers from the profitable games are to be encouraged a target marketing campaign(s) could be easily activated. It may be that incentives can be attached to the campaign to entice the customer, knowing that the expected return will outweigh the cost of the incentive.
Similarly, players on games with a lower margin may not need to be encouraged with attractive offers as they are deemed not overly worth retaining/encouraging, alternatively, incentives to spend elsewhere might be offered by way of discounts, vouchers.

Possible Use of a Segmentation Model
Profiling Customers Using Game Preference

It is always necessary to profile the segments that result from the analysis. This would obviously be done on the metrics that have been used as inputs but should also include demographic information as well as other business metrics deemed important. This gives deeper insight into what type of customers exist in each segment, aside from what has been used to derive them. Indexing can be performed to find out how much more/less likely a particular feature is in a segment when compared with the population.

Reporting can be done as part of the uses of the model to ensure the behaviour and change in behaviour is being monitored and analysed. Customer movement might be an important driver for analysis. This could be used to manipulate customer behaviour by trying to push customers into more profitable sectors.

Possible Use of a Segmentation Model for Profiling Customers Using Game Preference

The learning's and results of unsupervised models can be used in a prediction problem. They may be used as inputs to facilitate dimension reduction (i.e. the reduction of input variables into a model).

Alternatively, the results of the segmentation model might in itself become a prediction problem.

If the casino is interested in customer acquisition, the demographic, and public domain information can be used to try and predict segment membership. This information can be used to find out which factors are important.

An acquisition drive could be undertaken that targets customers not currently on the books who exhibit the characteristics required.

Success of PA in the Industry

- Hotel attendance forecasting: predictions are strongly considered when setting park hours and performing other strategic planning.
- Labor demand planning system, which generates transaction forecasts for every 15-minute period at many locations throughout the property, including park entry turnstiles, quick-service restaurants and merchandise locations. These forecasts help the resort plan labor effectively to ensure guest service standards are met.
- Customize offerings and experiences that better match resort guests' desires. This analysis, coupled with optimization models, allows the company's Web site and call center agents to present offers that provide a more customized vacation planning experience.

Success of PA in the Industry

Analytics were utilized throughout the planning process, including organizing the logistics around training the nearly 1,700 new crew members and planning crew rotations on and off the Disney ship.

- Restaurant table-seating optimization: helps the company understand the patterns around party sizes, arrival times and table turn times. This knowledge is incorporated into mathematical models that determine the right mix of tables to best meet guest demand.
- Streamline back-of-house operations: an on-site textiles facility handles nearly 300,000 pounds of laundry every day, servicing costumes from across the operation as well as linens from resort hotels.
- The resort leverages computer simulation to recreate the facility in a virtual environment. Simulation offers many benefits prior to making physical changes, including identifying potential bottlenecks and testing new concepts or designs that can increase the overall capacity of the facility.

Success of PA in the Industry

- Personalization – "With the knowledge of our guest preferences through our data, we can better target our promotional offers based on their historical behavior and their preferred amenities. It is going to get more competitive and we have to be as efficient as possible and target our guests as best as possible. - - Foxwoods looked at the property's shows, to see if country and western acts were bringing in more than hard rock.
- Identify the best customers geographically – allowed the property to target messages in those areas.
- Slot optimization
- Revenue management – helps making pricing decisions. Selling the right room at the right time at the right price.

Traditional RFM: Option One

Traditional RFM model: metrics defined by the client for a customer's recency of activity (presuming this would be around how long ago last interaction was in days), frequency of interaction (how often customer has been active in certain time period), value of customer (turnover of customer in certain time period). The casino may decide time period is to be days, weeks, months.

The three metrics would be standardised to ensure no one metric dominated the segmentation due to its magnitude, and an algorithm would be employed to ascertain whether pure and stable clusters could be obtained. The best statistical result often isn't the best practical result from a business viewpoint. Ultimately the sole aim is to obtain meaningful clusters that can be used by the business to satisfy their analytic objective.

Less Traditional RFM: Option Two

Less Traditional RFM: Proposal to derive n way metrics for recency, frequency and monetary landmarks. This would be done in consultation with the casino. A set time period would be agreed upon and split into relevant bands e.g. last month, 2-6 months ago, 6-18 months ago. Transactions would be calculated across the bands and a proportion derived to give a customer view of their behaviour on each metric. Three separate segmentation models would be built for each metric, potentially providing the casino with clearer insight into their relevant metrics. Depending on the number of metrics in the n way solution, distinct groups would be obtained in n dimensions. For example, if 3 bins were used for each, there would be 27 distinct groups that result, providing granular view of customer across the metrics. Amalgamation or consolidation of the groups could be performed to make it a manageable and useable output.

RFM: Option Three

More agile segmentation models in line with the casino's objectives and needs: It may be the case that something like recency is not an overly useful metric in terms of business insight. A set of models that structure clients across relevant landmarks similar to option 2. Such metrics might include game (e.g. slots vs. Card games vs. Roulette etc…); bet sizes (create discrete bands of bet sizes and profile customer's behaviour across these bands); visitation metrics (could be done by day of week, frequency etc.). Any number of segmentation models could then be used to derive an "overall" view of a customer in line with these metrics that groups "like" customers together. A method such as a self organised map or a memory based reasoning models (or another clustering algorithm) could be applied over the segments obtained from the individual models to try and ascertain an overarching segment.

RFM

Whichever approach was adopted, profiling would be done on the final results to determine what makes up group membership. Categorical factors such as gender, nationality/locality could be used as well as age (or indeed any other demographic features that are available) to understand the "type" of customer that resides in each group. These factors could be used for each segment and applied against the population metrics to determine how much more or less likely a segment was to exhibit a particular feature or type of behaviour when compared to the customer base as a whole.

Steps to be taken include:

- Ascertaining what the end objective is for the client, which would need to be addressed from the point of view of whether a traditional RFM model is suitable or would a more granular and informative approach be of more use.
- Data that is available including information on amount/volume of history, variables recorded by client at the customer level (including whether it is line by line or summary), number of customers client has in their database.
- Agreed metrics and bins for analysis e.g. recency (visits/bets in last month, 2-6 months, 6+months).
- Agreement on what makes a customer be included in the model….new /immature customers shouldn't be included as behaviour is inconsistent and random in appearance. Other end of the spectrum might be to remove customers who have not been seen in the past x months/years.

Steps to be taken include:

- Provision of raw data that can be used for building models and ensure it can satisfy all objectives. All variables would need to be available (or be able to be derived from raw data). Depending on size of customer base, best practice may be to take a representative sample and build model(s) on sample then apply across remaining customers.
- Derivation of customer views that enable segmentation models to be built including investigation of missing/unusual values, summary metrics for profiling and grouping of customers.
- Building of segmentation model(s) to understand whether consistent homogenous groups result providing best chance of identifying like behaviour. This would also include analysis of how stable the behaviour is across clusters by applying algorithm to hold out sample. This ensures behaviour is consistent and general across customer base rather than specific to data it is applied to.

Steps to be taken include:

- If a sample is used for efficiency purposes, remaining customers will then be scored using the resultant algorithms derived from modelling process.
- Consultation with client to ensure the results are meaningful to the business and also of a nature that is able to be utilised within the business. That is, number of clusters is manageable within current systems and processes.

Contact info:

Andrew Pearson
Qualex Asia Limited
Macau, SAR China
+853 6265-5885
+852 5184-2660
Andrew.pearson@qlx.com
www.qlx.com