|  |
| --- |
| Leveraging ML at Casino Woodbine |
|  |



### **Executive Summary**

This report will analyze the work conducted by JARVIS Consulting Group in relation to leveraging unsupervised ML techniques at Casino Woodbine. As a background, Casino Woodbine was looking for a new way to build a more personalized experience for its customer base. Currently the Casino leverages a one-dimensional tiering system based on Gold. Silver and Rewards tiers. These tiers are then used for planning purposes by the Casino to schedule and develop marketing campaigns. However, in today's competitive environment, customers want more relevant content that suits their needs and preferences. As a result, unsupervised ML techniques were leveraged using data from Woodbine’s loyalty program to identify and build seven customer personas based on a hierarchical clustering algorithm. These personas would serve as an internal segmentation for Woodbine that would be leveraged for more personalized targeted campaigns.

The data used for the clustering algorithm included customer ID level data. Features of the dataset comprised of a wide range of types from categorical to numerical ones. These included demographic data points such as gender, city and marketing consent preferences, to numerical features such as monetary KPIs, frequency metrics and marketing comp data. Using feature engineering, additional features were generated to provide better behavioral indicators. For example, a lapse flag was created to identify players that churned over the course of the previous 12 months. Using pre-existing features provided by Woodbine in addition to engineered ones, the dataset was run through a hierarchical based clustering algorithm to determine similar clusters within the dataset. The resulting optimal number of clusters of seven paved the way for the development of seven personas.

These seven personas shared vast differences between one another in terms of demographic, monetary and frequency based KPIs. It also indicated the need for Casino Woodbine to move away from the tiered model and instead focus on an internal segmentation/personification of its customer base when it came to developing marketing campaigns. The personas also enabled Casino Woodbine to identify their “true” high value base, customers that naturally enjoyed gambling (sure things), customers that could be ignored (lost causes) and customers that should be comped (persuadable). Finally, recommendations were made in terms of potential marketing campaigns that could work for each persona.

### **Overview**

The Canadian gaming industry is a large and mature industry that is present in every region of the country and provides meaningful economic returns to Canadians (Canadian Gaming Association, 2017). In recent years, we have witnessed a shift from traditional gambling activity to online gambling activity. Casinos, especially, have witnessed decreasing revenues as the online sector becomes more prevalent (Wheeler, 2019). As a result, Casinos are now under increased pressure to adapt to the challenge of better serving their customers and players.

From a marketing perspective, engagement and loyalty have been as entrenched in Casino operations as any other program. Loyalty has been around for decades primarily delivering benefits based on amount spent and value gained through progressive elite status (Watson, 2015). As the trends continue to shift, the evolution of loyalty has broadened consumer expectations such that the traditional approach is not as effective as it used to be, posing incredible challenges for marketers.

Diving into Casino Woodbine, we are faced with campaign challenge to better understand the segments amongst their players by capturing additional layers to build complete player profiles. Casino Woodbine implements a tiered rewards model where players are assigned a level based on their activity within the previous rolling 12 months; as Gold, Silver, and Rewards levels.

This model does work to understand player value but has room for improvement through the addition of other underutilized metrics obtained by the Casino. Similarly, rewarding purely the actions of spending money and time is crucial but is only one piece of the puzzle. The current heuristic that has been in place for several years is no longer meeting the demands of more personalized content that their clients require. Through a more sophisticated use of analytics, personalization is now as simple as amassing player practices and preferences to tailor each experience and keep them relevant (Watson, 2015). With a refined focus on data, Casino Woodbine would be able to zone in on what is important to a player and grant them the ability to respond quicker in improving engagement.

Our report below highlights an assessment and recommendations on ways to develop stronger realizations of Casino Woodbine’s customer segments to improve marketing spend efficiencies and build a more personalized experience for their customer base.

### **Data Description**

The data obtained from Casino Woodbine included customer level data captured on a per ID basis. The data was captured by the Casino’s loyalty program which encourages players to use their membership card when playing slot/table games to earn tier points. When a player does use their membership card, Casino Woodbine is able to capture demographic, financial and comp related data on their “known” player base. The dataset included a total of 35 variables ranging from demographic, financial and marketing spend type variables. A data dictionary of these variables has been provided at the end of the section.

**Demographic Data**

Comprised of demographic data such as gender, city, country, tier, mail consent, email consent and age on a per customer ID basis.

**Financial Data**

Comprised of financial data related to customer spend. Variables include total revenue per customer, spend per visit and number of times a customer visited Casino Woodbine on a per customer ID basis.

**Marketing Spend Data**

Comprised of marketing spend data related to the number of comps earned by a customer, the value of comps earned by a customer and the value of comps taken (redeemed) by the customer on a per customer ID basis.

**Timeline**

The financial and marketing spend data above are based on the previous 12, 6, 3 and 1 month time intervals. As a result, all customers in the dataset had some sort of activity at Casino Woodbine in the previous 12-month period, with customers “churning” at each interval thereafter. If a customer did not have any activity in the time interval in question, then a 0 would be assigned for that KPI.

**Initial Insights from Data Dump (High Level)**

* Total of 14,427 unique customers with activity at Casino Woodbine in the previous 12 months
* Average Age of known player base is 58 years old
* 78% of customer base has opted into receiving marketing communication via mail
* 43% of customer base has opted into receiving marketing communication via email
* 45% of customer base are males vs 55% are females
* 72% customers are in the Gold Tier (Highest Tier) whereas 28% customers are either in Rewards (Lowest Tier) or Silver Tiers (Second highest)

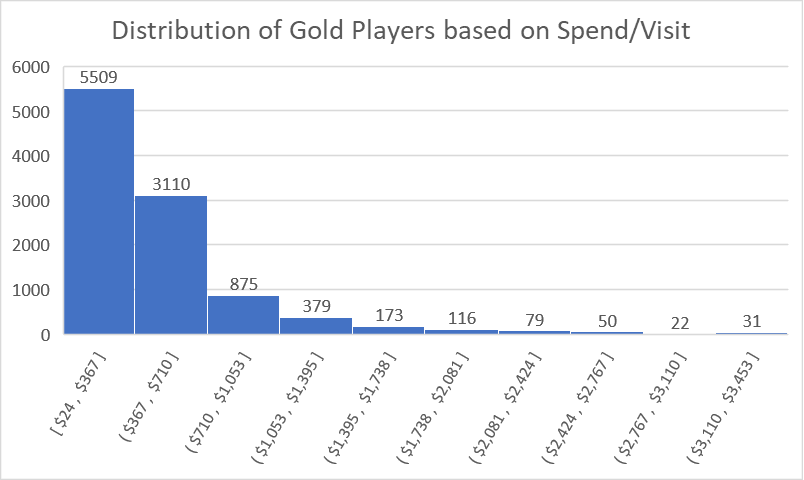
**Descriptive Analytics**

Based on the preliminary look at the dataset, there is clear evidence to show the need for Casino Woodbine to move from mass marketing campaigns to more **targeted campaigns** based on player profiles/personas.

* )

Current player base is left skewed indicating an aging population and potential long-term issues.

Of the total player base, 41% of players have not had any Casino activity in the previous month.

****

Inequality of gold tier players apparent in the above histogram. Although all players are classified as “Gold Tier” there is a huge discrepancy in their spend per visit. These values range from $24 to $3,453.

Huge discrepancy between comps earned vs comps taken especially in the gold tier. The lack of redeemed comps provide evidence that players are not engaged with the offers by Casino Woodbine. This has led to inefficient spending of marketing dollars. The difference in comps earned vs taken is known as “net comps”

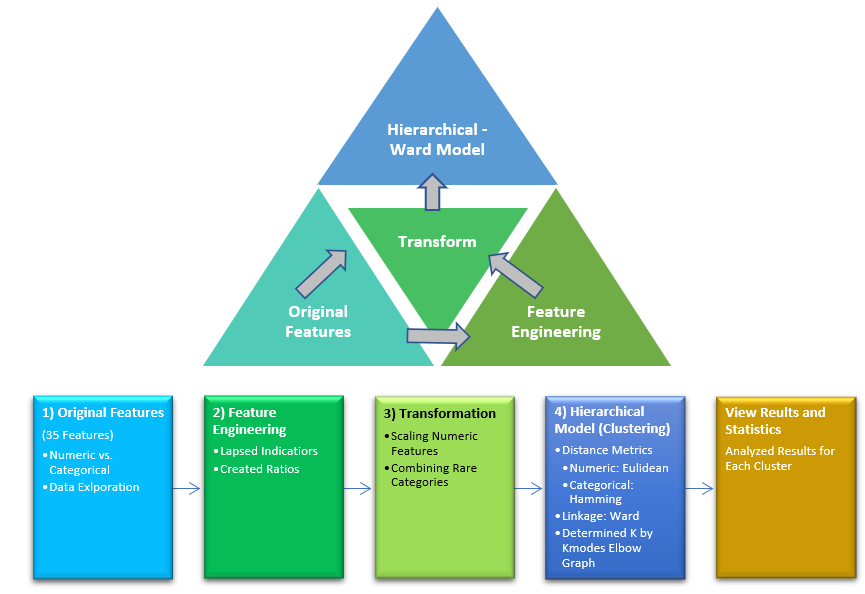
**Data**

**Data Dictionary**



### **Model Development**

The approach that was taken to provide Casino Woodbine with more sophisticated player segments involved unsupervised ML techniques based on the demographic, financial and marketing data mentioned above. This approach was selected because there was no target variable (labeled Y data) to help train the model. Instead, we wanted the algorithm to help determine similarities and differences of underlying customer segments that could be leveraged for more targeted campaigns. Thus, a Hierarchical-Ward clustering algorithm was selected to provide results that would dig deeper into the tiering system that Casino Woodbine currently leverages. The model development phase of this project consisted of 5 major steps, with the last one being the development of personas based on the results of the clustering algorithm.



**Step 1 Original Features Dataset**

The Python Pandas Profiling package was used to create exploratory analysis of the data (EDA), which contained dataset information, variable types, correlations, missing value detection and univariate statistics. The basic EDA would also provide analytical insights for variable transformations and feature engineering.

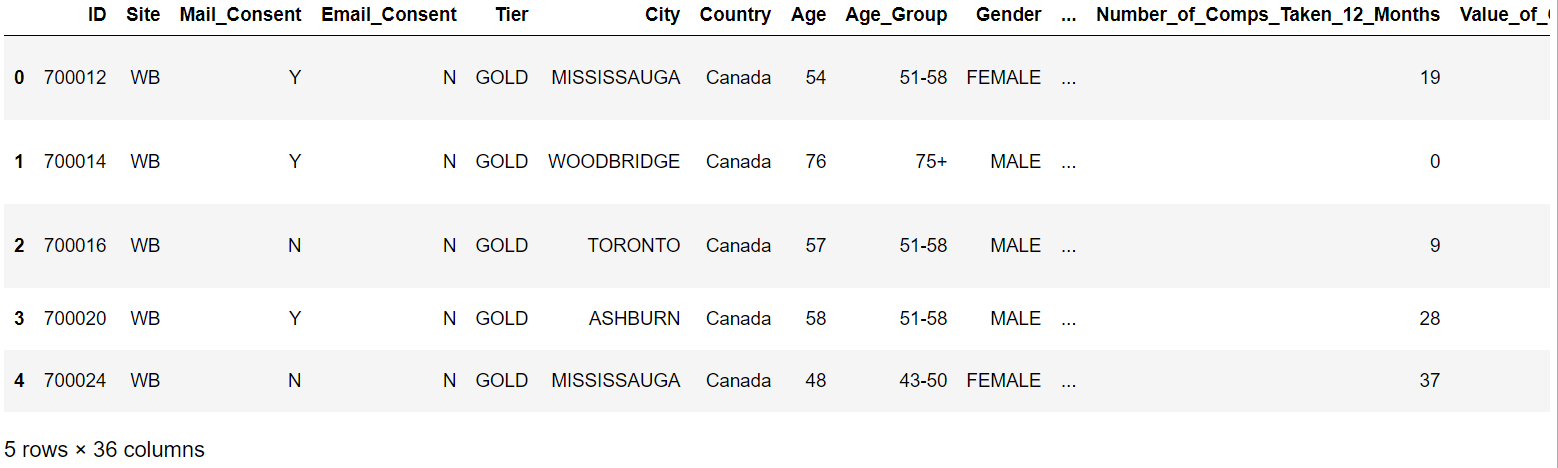
From the original source dataset, it could be observed that the composition of the Woodbine Casino data was made up of the 14,428 instances, with 35 features.

**Feature Composition**

* Numeric Features: consisted of 28 features.
* Categorical Features: consisted of 7 features.

|  |  |  |
| --- | --- | --- |
|  | **Numeric Features** | **Categorical Features** |
| Features | Age | Email\_Consent |
| Twelve\_Month\_ADT | Tier |
| Six\_Month\_ADT | City |
| Three\_Month\_ADT | Country |
| One\_Month\_ADT | Age\_Group |
| Twelve\_Month\_Theo\_NW | Gender |
| Six\_Month\_Theo\_NW |  |
| Three\_Month\_Theo\_NW |
| One\_Month\_Theo\_NW |
| Twelve\_Month\_Actual\_NW |
| Six\_Month\_Actual\_NW |
| Three\_Month\_Actual\_NW |
| One\_Month\_Actual\_NW |
| Twelve\_Month\_Trips |
| Six\_Month\_Trips |
| Three\_Month\_Trips |
| One\_Month\_Trips |
| Number\_of\_Comps\_Taken\_12\_Months |
| Value\_of\_Comps\_Taken\_12\_Months |
| Twelve\_Month\_Net\_Comp |
| Number\_of\_Comps\_Taken\_3\_Months |
| Value\_of\_Comps\_Taken\_3\_Months |
| Three\_Month\_Net\_Comp |
| Number\_of\_Comps\_Taken\_1\_Month |
| Value\_of\_Comps\_Taken\_1\_Month |
| One\_Month\_Net\_Comp |

**Original Data Features**



**Original Sample Data**

**Step 2 – Feature Engineering**

The following features were created to further capture player churn patterns in the dataset:

**Lapsed Indicators**

From the 1, 3, 6, & 12-month trip features, a lapsed indicator was engineered to capture the number of months since a member had visited the Casino. This was created by observing if a member had visited in each of the respectful trip month intervals. A categorical feature was then created for each instance

* If a member did not have data in “One\_Month\_Trip”, then the Lapse feature would show “1 Month Lapsed”.
* If a member did not have data in “Three\_Month\_Trip”, then the Lapse feature would show “3 Month Lapsed”.
* If a member did not have data in “Six\_Month\_Trip”, then the Lapse feature would show “6 Month Lapsed”
* If a member had data in all trip features, then the Lapse feature would show “No Lapsed”.

**Ratios**

Ratios were engineered to obtain a member’s 1, 3 and 6-month period ratios against 12- month measures of Theoretical Net Win, Actual Net Win and Trips to evaluate their behavior pattern by time intervals. Ratios consisted of Theoretical Net Win Ratios, Actual Net Win Ratio, Visit/Trip Ratios.

* **Theoretical Net Win Ratios:**

One Month Theo NW Ratio = One Month Theo NW / Twelve Month Theo NW

Three Month Theo NW = Three Month Theo NW / Twelve Month Theo NW

Six Month Theo NW = Six Month Theo NW / Twelve Month Theo NW

* **Actual Net Win Ratios:**

One Month Act NW = One Month Actual NW / Twelve Month Actual NW

Three Month Act NW = Three Month Actual NW / Twelve Month Actual NW

Six Month Act NW = Six Month Actual NW / Twelve Month Actual NW

* **Visit/Trip Ratios:**

One Month Trips = One Month Trips / Twelve Month Trips

Three Month Trips = Three Month Trips / Twelve Month Trips

Six Month Trips = Six Month Trips / Twelve Month Trips

**Step 3 Transformations**

**Scaling Numeric Features**

In this data preprocessing step, scaling, standardizing, and transformation were essential for numeric features which were used to treat skewed distributed features and scale them to a normalized range for modeling. We standardized independent numeric features present in the data in a fixed range but ensured to still maintain the same characteristics and distribution (outliers were important to retain in the data).

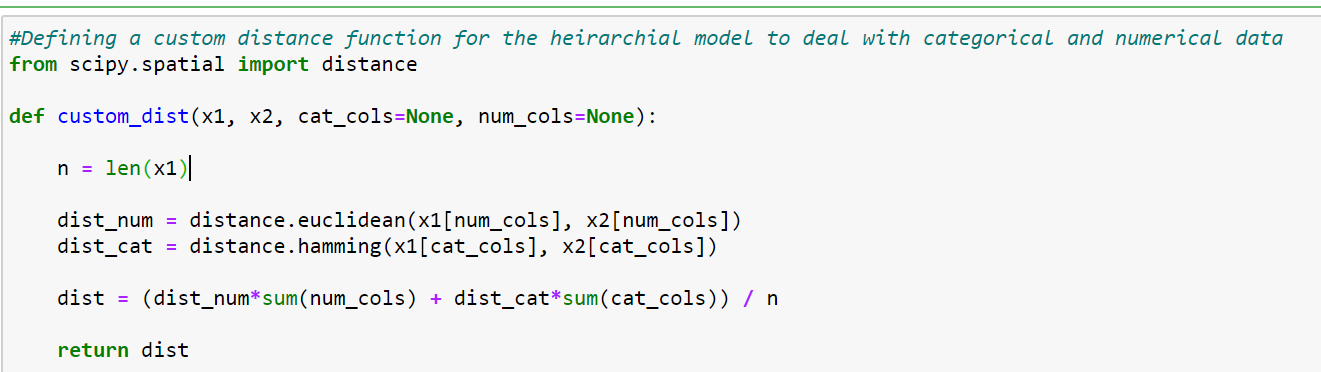
**Combining Rare Categories**

The dataset also contained features with high number of levels. For example, the “City” feature contained many across Ontario. As a result, we combined lower frequency cities into a new attribute called “Other\_City”. All levels in the “City” categorical feature that were below the decided cut-off of 27% were placed into the “Other\_City” variable. This cut-off threshold represented the percentile distribution of the different levels. What was observed was that the “Other\_City” variable were areas outside the GTA which was interesting to examine. By grouping these rare cities into one category, it decreased the level of cardinality within the “City” feature and helped increase model performance. It also allowed for a more thorough understanding of where customers were coming from (urban, suburban, or rural areas).

**Step 4 Hierarchical Model Development**

By using a hierarchical model, categorical features did not need to undergo any numeric encoding or scaling. The only hyperparameters that were needed for the hierarchal based clustering included a defined distance metric and link function. A distance matrix needed to be defined for both numerical and categorical features, which measured how far apart two instances were from one another.

**Custom Distance Metric**

We used the Euclidean distance metric on numeric features and the Hamming distance metric on categorical features. We then then combined both distance metrics into a custom defined function via a weighted average for use in the hierarchical model.

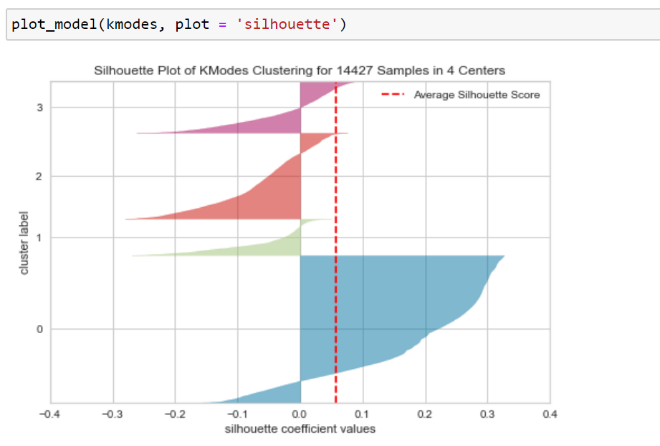
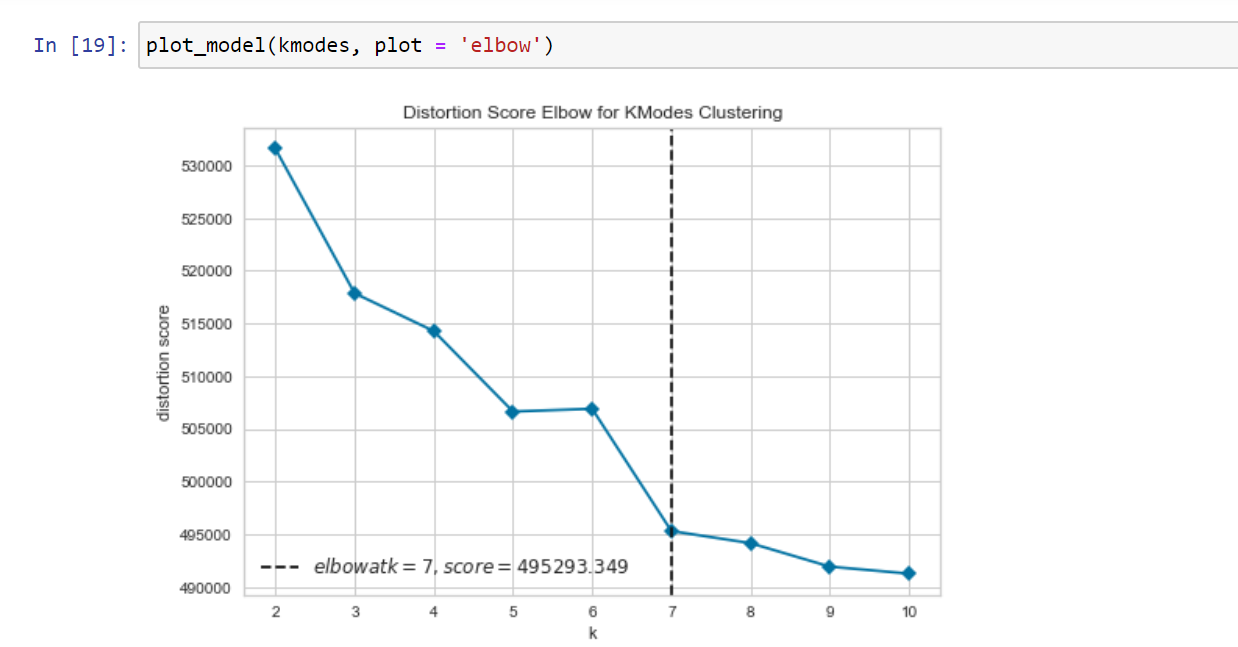
**Linkage**

The Ward linkage calculated the distance between cluster centroids, which was good for globular data and getting evenly sized clusters.

**Hierarchical Model**

We performed hierarchical clustering using the custom-built distance metric and Ward linkage above as the hyperparameters for the scipy.cluster.hierarchy.linkage() function and plotted the dendrogram.

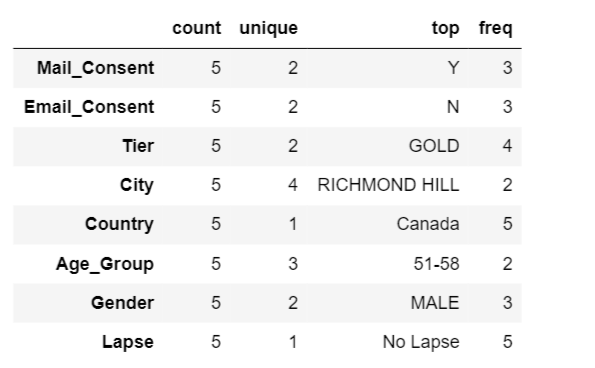
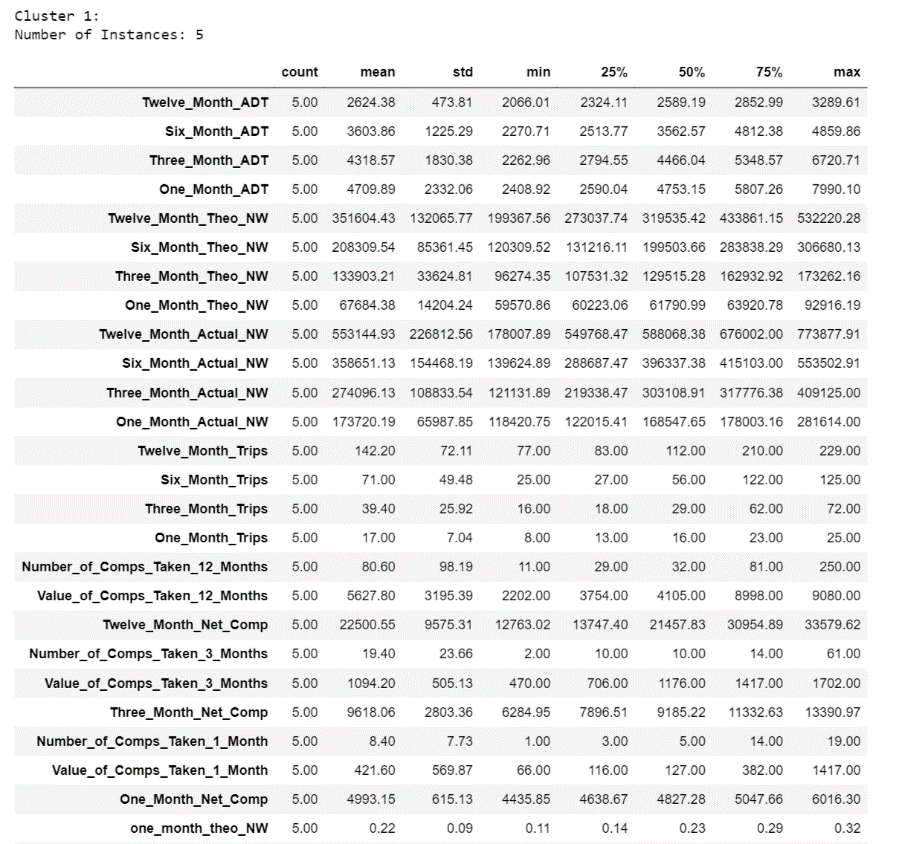
The ideal cut for the dendrogram was determined using the ideal K-value (optimal number of clusters) based on the elbow chart below. This was done via K-Modes to determine the ideal number of clusters. As shown below, the ideal number of clusters was seven. The last step involved the usage of the **scipy.cluster.hierarchy.fcluster() function** to flatten the dendrogram, thereby allowing us to assign an appropriate cluster to each customer in the dataset.



**Step 5 Analyzed Results for Each Cluster via Summary Statistics**

Using numeric and categorical features from the varying clusters, we analyzed summary statistics and counts for categorical variables. These statistics were used for evaluating cluster purchasing, demographic and marketing spend behaviours to create the personas shown below.

**Example Results of Cluster 1 (Similar approach was taken on all 7 clusters):**



|  |  |  |
| --- | --- | --- |
| Male Gambler Stock Illustrations – 341 Male Gambler Stock ... |  | **Persona 1:** The diamond 5  Cluster 1 |
| **Profile**  *“Trust me, I know what I’m doing”*  *“See you tomorrow!”*  *“Hobby? I do this for a living”*  **summary**   * PRofessional Gamblers * high spend * high frequencY * male dominate * DisinterEsted IN COMPS * **sure things**   **KPIs**   * **SIZE OF CLUSTER: 5 PLAYERS** * 40% OF CLUSTER OVER THE AGE OF 51 * ALL LIVE WITHIN GTA * ALL GOLD PLAYERS * HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 1** | **Player Base Average** | | 12 Month ADT | $2,624 | $459 | | 12 Month NW | $553,145 | $14,178 | | 12 Month Theo NW | $351,604 | $14,060 | | 12 Month Net Comp | $22,501 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 1** | **Player Base Average** | | 12 Month Visits | 142 | 34 | | 6 Month Visits | 71 | 17 | | 3 Month Visits | 39 | 9 | | 1 Month Visits | 17 | 3 | | Percent of Lapsed Players | 0% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
| Rich Old Woman Stock Illustrations – 535 Rich Old Woman Stock ... |  | **Persona 2:** RICH RETIRED RACHELS  Cluster 2 |
| **Profile**  *“What better way to spend my time”*  *“Will be back tomorrow!”*  *“I love the excitement of a jackpot!”*  **summary**   * AFFULENT RETIREES * main form of entertainment is the Casino * high spend * high frequencY * Disintersted IN COMPS * **Persuadable**   **KPIs**   * **SIZE OF CLUSTER: 343 PLAYERS** * 67% OF CLUSTER OVER THE AGE OF 51 * MANY LIVE WITHIN GTA * MIX OF GOLD, SILVER AND REWARD TIERS * SECOND HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 2** | **Player Base Average** | | 12 Month ADT | $1,175 | $459 | | 12 Month NW | $120,232 | $14,178 | | 12 Month Theo NW | $118,472 | $14,060 | | 12 Month Net Comp | $5,723 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 1** | **Player Base Average** | | 12 Month Visits | 129 | 34 | | 6 Month Visits | 68 | 17 | | 3 Month Visits | 35 | 9 | | 1 Month Visits | 11 | 3 | | Percent of Lapsed Players | 3% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
|  |  | **Persona 3:** TREPIDATION TOURIST TREVOR  Cluster 3 |
| **Profile**  *“Great way to pass the time”*  *“Casino isn’t really the first thing that comes to mind when I think of entertainment”*  *“I try to stop by whenever I get a chance”*  **summary**   * CASUAL PLAYERS * LOW spend * LOW frequencY * FEmale dominate * INTERESTED IN THAT OCCASIONAL THRILL * INTERESTED IN COMPS * **LOST CAUSES**   **KPIs**   * **SIZE OF CLUSTER: 4,500 PLAYERS** * 75% OF CLUSTER OVER THE AGE OF 51 * MANY LIVE OUTSIDE GTA * MIX OF GOLD, SILVER AND REWARD PLAYERS * SIXTH HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 3** | **Player Base Average** | | 12 Month ADT | $425 | $459 | | 12 Month NW | $4,156 | $14,178 | | 12 Month Theo NW | $3,923 | $14,060 | | 12 Month Net Comp | $181 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 3** | **Player Base Average** | | 12 Month Visits | 10 | 34 | | 6 Month Visits | 3 | 17 | | 3 Month Visits | 1 | 9 | | 1 Month Visits | 0.2 | 3 | | Percent of Lapsed Players | 85% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
|  |  | **Persona 4:** FRUGAL DREAMERS  Cluster 4 |
| **Profile**  *“Just one more spin..I know I can win”*  *“I do it for the chance to win that big jackpot!”*    *“Imagine what I can do with my winnings”*  **summary**   * CASUAL PLAYERS * LOW spend * MEDIUM frequencY * FEmale dominate * intersted IN COMPS * **PERSUADABLE**   **KPIs**   * **SIZE OF CLUSTER: 4,880 PLAYERS** * 76% OF CLUSTER OVER THE AGE OF 51 * MAJORITY LIVE WITHIN GTA * MIX OF GOLD, SILVER AND REWARD PLAYERS * FIFTH HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 4** | **Player Base Average** | | 12 Month ADT | $404 | $459 | | 12 Month NW | $9,306 | $14,178 | | 12 Month Theo NW | $9,329 | $14,060 | | 12 Month Net Comp | $466 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 4** | **Player Base Average** | | 12 Month Visits | 27 | 34 | | 6 Month Visits | 13 | 17 | | 3 Month Visits | 6 | 9 | | 1 Month Visits | 2 | 3 | | Percent of Lapsed Players | 30% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
| People Restaurant clipart - Restaurant, Food, Hotel, transparent ... |  | **Persona 5:** BUFFET LOVERS  Cluster 5 |
| **Profile**  *“Love the buffet!”*  *“I come for the restaurant, stay for the games”*  *“More of a foody than a hardcore player”*  **summary**   * foodies * LOW spend * LOW frequencY * FEmale dominate * intersted IN COMPS * **sure things**   **KPIs**   * **SIZE OF CLUSTER: 1,339 PLAYERS** * 77% OF CLUSTER OVER THE AGE OF 51 * MAJORITY LIVE WITHIN GTA * MIX OF GOLD, SILVER AND REWARD PLAYERS * LOWEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 5** | **Player Base Average** | | 12 Month ADT | $369 | $459 | | 12 Month NW | $2,608 | $14,178 | | 12 Month Theo NW | $2,751 | $14,060 | | 12 Month Net Comp | $145 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 5** | **Player Base Average** | | 12 Month Visits | 11 | 34 | | 6 Month Visits | 9 | 17 | | 3 Month Visits | 7 | 9 | | 1 Month Visits | 3 | 3 | | Percent of Lapsed Players | 27% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
|  |  | **Persona 6:** HIGH VALUE HOPEFULS  Cluster 6 |
| **Profile**  *“I don’t visit often but when I do, I go all out”*  *“I gamble on a lot of platforms”*  *“Gambling is definitely a hobby of mine”*  **summary**   * HIGH POTENTIAL * high spend * LOW frequencY * BOTH GENDERS EQUALLY REPRESENTED * Disintersted IN COMPS * **sLEEPING DOGS**   **KPIs**   * **SIZE OF CLUSTER: 1,112 PLAYERS** * 70% OF CLUSTER OVER THE AGE OF 51 * MAJORITY LIVE OUTSIDE GTA * MIX OF GOLD, SILVER AND REWARD PLAYERS * THIRD HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 6** | **Player Base Average** | | 12 Month ADT | $1,048 | $459 | | 12 Month NW | $36,240 | $14,178 | | 12 Month Theo NW | $36,952 | $14,060 | | 12 Month Net Comp | $2,179 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 6** | **Player Base Average** | | 12 Month Visits | 44 | 34 | | 6 Month Visits | 22 | 17 | | 3 Month Visits | 12 | 9 | | 1 Month Visits | 4 | 3 | | Percent of Lapsed Players | 12% | 58% |   **DEMOGRAPHICS** |

|  |  |  |
| --- | --- | --- |
|  |  | **Persona 7:** BALLERS ON A BUDGET  Cluster 7 |
| **Profile**  *“What better way to spend my time than at Woodbine”*  *“I enjoy gambling, but also know not to get carried away”*  *“I’ll try again tomorrow!”*  **summary**   * KEEP THEM COMING * LOW spend * high frequencY * FEmale dominate * interEsted IN COMPS * **sure things**   **KPIs**   * **SIZE OF CLUSTER: 2,248 PLAYERS** * 77% OF CLUSTER OVER THE AGE OF 51 * ALL LIVE WITHIN GTA * MIX OF GOLD, SILVER AND REWARD PLAYERS * FOURTH HIGHEST NET COMP |  | **monetary**   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 1** | **Player Base Average** | | 12 Month ADT | $296 | $459 | | 12 Month NW | $23,411 | $14,178 | | 12 Month Theo NW | $23,356 | $14,060 | | 12 Month Net Comp | $676 | $645 |   ***Average per Player:***  **frequency**  ***Average per Player:***   |  |  |  | | --- | --- | --- | | **KPI** | **Cluster 1** | **Player Base Average** | | 12 Month Visits | 94 | 34 | | 6 Month Visits | 49 | 17 | | 3 Month Visits | 25 | 9 | | 1 Month Visits | 9 | 3 | | Percent of Lapsed Players | 3% | 58% |   **DEMOGRAPHICS** |

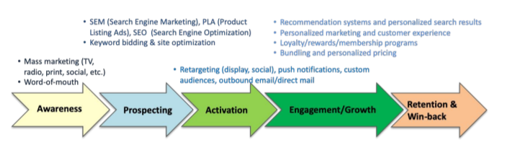
### **Recommendations**

**Problem Area 1**

Throughout the data analysis journey, a few factors stood out as problem areas for Casino Woodbine. As portrayed in the “Data Description” section of the report, one problem area included an aging population. It was clear that the overall known player base of Casino Woodbine skewed much older which could result in financial consequences from a long-term perspective.

**Tactic**

Awareness is key for driving traffic, especially for younger folks. As shown in the diagram below, leveraging more marketing dollars towards mass marketing campaigns is a great way for Woodbine to attract that younger demographic. However, the journey for that customer does not stop there. It is imperative for Woodbine to convert these players from “non-carded” to “carded” (meaning they sign up for the loyalty program). This will allow Woodbine to gain further insights into players and therefore treat them in a way that resonates with that customer.



### **Problem Area 2**

### Currently, Casino Woodbine leverages the tiering system to help in implementing marketing campaigns. For example, based on the tier of a customer (either Gold-Highest, Silver-Middle, Rewards-Low), a marketing campaign will be tailored to that tier. However, this is not the most result driven approach as not all customers are the same.

### **Tactic**

### The tiering system provides simplicity for customers. It is a great tool to leverage from a front facing perspective. A customer can easily understand what is required to further improve upon their current tier. However, a problem with the tiering system is that it is one dimensional and does not consider all data pertaining to a customer. A better approach would be the use of an internal segmentation (which the customer is not aware of) which considers other factors such as demographic, financial, and marketing spend. Woodbine can then leverage this internal segmentation to tailor their marketing efforts and focus more on personalized content.

### Building upon problem area 2, JARVIS Consulting Group has recommended Woodbine to leverage the seven personas mentioned above when it comes to creating more tailored experiences for their customer base. In addition, it also provides glaring differences across the seven segments that Woodbine can capitalize on to efficiently leverage their marketing budget.

### At a high level, JARVIS Consulting Group recommends Woodbine to leverage Segmentation, Targeting and Positioning (STP) tactics.

### 

### What we have achieved via hierarchical clustering are seven distinct personas that can be used as an internal segmentation by Woodbine and lead to more segmented (or differentiated) marketing tactics.

### **Persona 1: The Diamond 5**

### The **Diamond 5**, as the name implies, are 5 players from the original dataset of 14k that truly outspend the rest in terms of revenue per visit and the average revenue per customer over the course of 12 months. They also have the highest average net comp amount, which as a refresher is the difference between Comps Earned vs Taken. There is a clearly a problem area as the comps that have been assigned to this persona are not being leveraged by these customers. A recommendation would be to build more of a one on one connection with this group instead of opting to go with slot/table offers. The floor manager should ensure that these 5 players are content with the service. Considering this group is small, a survey could also help in understanding more personal preferences.

### **Persona 2: Rich Retired Rachels**

### The **Rich Retired Rachels** persona are other high value customers at Woodbine. Although they do not spend at the same level as the Diamond 5, they are still an integral part in driving future revenue. They can also be classified as sure things. Offers do not matter as much, instead this persona does it strictly for the thrill and excitement that the Casino brings. We can also be confident that they will continue to visit frequently which can be observed by the lapsed rate of 3%. Another interesting insight is that this persona is composed of players from all tiers which again alludes to the fact that Woodbine should leverage an internal segmentation rather than go by tier alone.

### **Persona 3: Trepidation Tourist Trevor**

The **Trepidation Tourist Trevor** are the bottom of the barrel type of segment. They spend little, visit infrequently and do not contribute much to the overall revenue at Woodbine. One interesting insight is that this group consists of 70% Gold players. This again reinforces the notion that not all Gold players are equal as previously mentioned in the Data Description section of the report. Another interesting insight is that this group lives relatively far from Woodbine. Although they may be interested in gambling, distance hinders this group from visiting often. It may be important for Woodbine to consider a digital presence (via a website or app) to engage with those customers that live afar. Until then, these players are more so lost causes due to the low spend and high lapse rate. Mass marketing tactics may be the way to go whenever the Casino plans for a concert, slot tournaments or restaurant specials rather than slot/table free play offers to engage with this group.

### **Persona 4: Frugal Dreamers & Persona 5: Buffet Lovers**

Both personas exhibit similar behavior in terms of frequency. However, what differentiates these segments are their spend behaviors. **Frugal Dreamers** seem to be individuals who are looking for that big jackpot. Although they do not visit the Casino much, whenever they do, more is spent on Casino related activities than other entertainment venues at the Casino. On the contrary, **Buffet Lovers** seem to be individuals who spend more of their time on the entertainment component of a Casino (ie. Drinks, restaurant, events) than the gambling component. As a result, Buffet Lovers may be more interested in communication related to entertainment vs frugal dreamers who may resonate more with slot tournaments.

**Persona 6: High Value Hopefuls**

The **High Value Hopeful** persona is like that of Persona 2, the “Rich Retired Rachels”. The only caveat is that this segment does not visit as often as that of persona 2. It is imperative for Casino Woodbine to focus their attention on this persona as this group has proven themselves to spend a lot at the Casino. One promotional idea may be more frequency-based campaigns (such as visit x times, get x in slot offers) to incentivize that additional visit. Another option would be the development of recommender systems to get players from this group to try different games. Increasing a player’s game profile may result in more frequent or longer stays at Woodbine.

**Persona 7: Ballers on a Budget**

The **Ballers on a Budget** segment rank the lowest in terms of spend per visit across all other clusters. However, they visit the Casino on such a frequent basis that they are in fact one of Casino Woodbine’s most valuable clusters. As a result of this behavior, it seems as if these players have a “budget” for how much they want to spend per trip to Woodbine. In addition, these players can be classified as sure things. They seem generally invested in gambling that they will most likely continue this behavior regardless of any marketing incentives. If Woodbine were to incentivize them, then the target would be to increase their spend per visit. This could be achieved by introducing new games, hosting more frequent slot tournaments, or prolonging their duration at the Casino through lunch/dinner restaurant specials.

**Final Remarks**

From the descriptive analytics phase it was evident that Casino Woodbine had a few areas that needed improvement. The problems ranged from inefficient marketing spend (determined through net comp), inequality of player spend across tiers, to an aging population. Through this initital investigation, it became extremely important for Woodbine to leverage more analytical based solutions in order to assist in the strategy component. Unsuperivised Machine Learning served as the solution as the output of the hierarchical based clustering lead to the development of seven personas. These personas not only showcased vast differences of player preferences across Woodbine’s customer base but also demonstrated the need for an internal segmentation. The hope is that Woodbine can now leverage the seven personas to be more effective in their use of marketing dollars for future endeavours.

### **References**

Canadian Gaming Association. (2017). *The National Economic Benefit of the Candian Gaming Industry.* HLT Advisory.

Watson, P. (2015, September 25). *Engaging loyalty: Creating a loyalty relationship at every point*. Retrieved from Canadian Gaming Business: http://www.canadiangamingbusiness.com/Creatingaloyaltyrelationship.aspx

Wheeler, L. (2019, June 23). *The Gambling Industry in Canada*. Retrieved from 888 Casuno: https://www.888casino.com/blog/casino-tips/gambling-industry-in-canada#:~:text=Canada%20has%20multiple%20provinces%2C%20each,online%20gambling%2C%20and%20horse%20racing.

Mosmer, Masoum (2020,August). *MMA831\_Week3\_POST\_TargetedAdvertising&Promotions*