

# **Casino Feasibility Study Visitor Segmentation Phase**

**A Case Study**



*William M. Bailey, Ph.D.*

---

<http://home.earthlink.net/~statmanz>

## Table of Contents

Introduction.....	3
Observations .....	4
Findings .....	5
Classification Model for Visitation.....	5
Visitor Segment Profile .....	9
Factors for Success .....	10



## Introduction

Owners of a new resort/casino commissioned the author to "profile" the most probable user. Further, whether or not there is a relationship between the resort's location and probable visitor residence.

Several multivariate methodologies are applied to accomplish this effort. The first objective is to "classify" likely versus unlikely visitors based on respondent impressions and opinions from select survey questions. Secondly, the segmentation phase profiles/characterizes the most likely user group.<sup>1</sup>

**Discriminant Analysis:** The goal of discriminant analysis is to classify cases into one of several mutually exclusive groups based on their values for a set of predictor variables. In the analysis phase, a classification rule is developed using cases for which group membership is known. In the classification phase, the rule is used to classify cases for which group membership is not known. This analysis is applied to identify and explain two visitation segments, those that are somewhat to very likely to visit The Resort versus those who are unlikely.

**CHAID:** *CHI*-square Automated Interaction Detector is a tree-based classification system that aids in segmentation research and exploratory data analysis. It is used here to identify homogeneous segments that include most likely visitors to the planned resort applying the results from the discriminant classification model.

**Factor Analysis:** Factor Analysis is concerned with the study of interrelationships among a set of variables, none of which is given the special status of a criterion variable. The methodology is used to identify viable groupings. It is a statistical technique that provides for data reduction and subsequent interpretation. The result is to identify a small set of variables that can be used to represent interrelationships among many variables. The Kaiser-Meyer-Olkin statistic (KMO) is an index for comparing the magnitudes of the observed correlation coefficients to the partial correlation coefficients. Small KMO values indicate that a factor analysis may not be a good idea since correlations between pairs of variables cannot be explained by the other variables. A KMO of at least .81 is considered "meritorious." The KMO for the four-factor model is .815.

**Reliability Analysis:** This algorithm tests the viability of each factor using Cronbach's Alpha ( $\alpha$ ). Cronbach's Alpha is based on the "internal consistency" of a test. That is, it is based on the average correlation of items within a test. The Alpha for the preliminary four-factor model ranges from .886 (Factor 1) to .625 (Factor 4). Subsequently, two factors are retained for descriptive purposes.

---

<sup>1</sup> All procedures use SPSS' advanced applications modules. CHAID is derived from SPSS AnswerTree.™



### Observations

This segmentation study finds that the more "probable" visitor will be in the upper income group, slightly older, tend to play the slots and/or table games, have lived in their current residence three to ten years, and should visit the resort three to five times a month.

Moreover, the likely visitor will "demand" good customer service, seek out restaurants, lounges, movie theaters, and a resort that is conveniently located. While convenient location to residence is an important factor, it (along with "closer to residence") is subservient to the services and amenities provided.

Further, though the sample pool includes known gamblers, this study finds that from 65% to 75% of these will actually visit the resort.

The following findings, specific to the response set most likely to visit the resort, support the study's sub-objectives:

- ◆ To rank order the individual respondent's top 3 activities. [Page 9]

Golf, hiking, and exercise stand out. This most likely visitation segment appears to be above average, relative to the overall response base, "active oriented people."

- ◆ To determine the likelihood that the respondent will visit the resort. [Page 6 footnote and page 8 Classification Matrix]

The Segmentation Study estimates that 65% to 75% of the sample representing known gamblers will actually visit. This compares to the near 90% who say that they are at least "somewhat likely" to visit the resort based on the aided visitation question (Q68).

- ◆ To determine the relationship between distance from The Resort and the stated likelihood of visiting the casino. [Page 6]

There is a relationship between distance from a gambler's residence and the likelihood of visiting a casino. Six attributes can identify likely visitors to the resort, of these, "closer to your [the respondent's] residence" ranks third, distance to the respondent's favorite casino is fourth with distance to the resort sixth. The most probable visitor will live within 6.8 miles of their favorite casino and 3.6 miles from the casino.

- ◆ To determine how the distance between competing resorts and the respondent and stated visitation are related. [Page 9 Profile table and footnote]

Persons most likely to visit are willing to travel 1.6 miles further from their residence to the preferred casino than those less likely to visit the resort. Moreover, they will tend to live one mile closer than probable non-visitors. Obviously, their attendance also is dictated on being provided the services expected as shown in the table on page 11.



## Findings

### *Classification Model for Visitation*

The classification model applies Discriminant Analysis and is based on the Visitation Likelihood question (question 68)...

["The Resort will be the first luxury golf and spa resort located off the Strip. The Resort offers elegant surroundings that include two hotels, a world-class luxury spa, golf, beverages and fine dining at 11 different restaurants and pubs, shopping at ten different shops and boutiques, a conference and banquet center for meetings and wedding receptions, and a casino. The spa offers a variety of hydrotherapy and body treatments from around the world, a boutique, exercise areas and personal fitness training, consultation on nutrition and stress management, and a salon for hair, nail and facial services. It is also bordered by three championship golf courses and offers tee times and shuttle service to six area golf courses. Based on what you now know about The Resort, how likely are you to go there when it opens - would you say..."]

Next, the series of questions relating to the importance of several decision-making characteristics are introduced (questions 8-20) ...

["Now, I'm going to read a short list of characteristics you may consider when deciding which casino you will visit. After I read each, please tell me if you feel that characteristic is very important, important, somewhat important, not very important, or not at all important in helping you decide which casino you will visit. When selecting a casino to visit, how important is it for the casino you select to..."]

Finally, with the addition of respondent demographics, the response base is segmented into two components: Likely to visit and Not likely to visit. The correlation's matrix below identifies an initial set variables that have a strong relationship with likely intent to visit the resort.

Correlation Matrix  
[Likelihood to visit]

Attribute	Pearson Correlation	Significance (2-tailed)	N
Distance to Favorite Casino [updated]	.202 **	.000	375
12. Provide 'comps' for 'preferred' customers	.201 **	.000	395
Distance to The Resort [reversal variable]	-.197 **	.000	396
18. World-class spa and health club	.189 **	.000	397
13. Have 'loose' slot machines that let you win	.186 **	.000	397
1998 Gross or total household income	.179 **	.001	363
10. Type of slot machines	.178 **	.000	397
11. Casino club card program	.175 **	.000	396
14. Good customer service	.150 **	.003	396
20. Convenient location	.135 **	.007	397
19. Closer to your residence	.127 *	.012	395
17. Restaurants, lounges, movie theaters...	.113 *	.024	397
8. Offer blackjack, craps, and other table games	.098	.051	395
9. High table limits of \$5000 or more	.084	.096	394
16. Paycheck cashing service	.078	.123	396
15. Sports books	.059	.241	396
Respondent's age	.031	.531	397

\* Correlation is significant at the 0.05 level (2-tailed)

\*\* Correlation is significant at the 0.01 level (2-tailed)

Note: The sign of the correlation for Distance denotes that the further away they are the less likely they will go.



## Casino Segmentation Study

---

Regarding distance, analysis finds that the mean distance [5% trim to reduce the affect of extreme values] from respondents' residence is 3.6 miles. On the other hand, the mean distance for those least likely to visit is 4.6 miles.

The series of questions relating to respondent attitudes toward h/her favorite casino are not used to distinguish visitation likelihood since they are directly related to a specific casino and not "generic" opinions. These are, however, important elements and useful in the profiling phase that follows.

Applying the selection decision and demographic variables to the discriminant model, the analysis is able to correctly classified 71% of original grouped cases. That is, on average, seven out of ten people who provide answers to six questions can be correctly identified as a likely or unlikely visitor of the resort.<sup>2</sup>

The first table below displays the "impact" rankings of the resulting variables in the classification model: Of the resulting key classification variables, "Provide "comps" for "preferred" customers?" is the more discriminating.<sup>3</sup>

Standardized Discriminant Function Coefficients

Variable	Variable "Impact"
World-class spa and health club	.472
Provide 'comps' for 'preferred' customers	.444
Closer to your residence	.375
Distance to favorite casino	.366
1998 Household Income	.358
Distance to the resort/casino	-.450 *

The sign here indicated that the distance function for this variable is skewed to the left of its mean.

---

<sup>2</sup> Logistic regression modeling found an even tighter 77% classification, however it is deemed less stable for classification purposes.

<sup>3</sup> For purposes of segmentation, a rating of 1, 2 or 3 in the question 8-20 series are classified as "unimportant," whereas a rating of 4 or 5 is "important" based on the 5-point Likert scale used in the survey. For likelihood (question 68), a ratings of 1 or 2 is defined as "likely," a 3, 4 or 5 "unlikely."



## Casino Segmentation Study

---

The actual discriminant models for classification purposes are...

Classification Function Coefficients

Variable	Likelihood to go	
	Somewhat to Very	Not Likely
Provide 'comps' for 'preferred' customers	1.338	1.027
World-class spa and health club	1.159	.793
Closer to your residence	2.301	2.026
1998 Household Income	2.429	2.119
Distance to favorite casino	1.008	.895
Distance to the resort/casino	1.298	1.529
<b>(Constant)</b>	-18.374	-15.076

Fisher's linear discriminant functions

Each model is applied to each case and the result with the highest score is the group, likely or not likely, to which the case is assigned. The models are as follows...

**Likely Score** =  $-18.374 + (1.338 * comps) + (1.159 * world-class\ spa) + (2.301 * closer\ to\ residence) + (2.429 * HH\ Income) + (1.008 * distance\ to\ favorite\ casino) + (1.298 * distance\ to\ resort/casino)$

and

**Not Likely Score** =  $-15.076 + (1.027 * comps) + (0.793 * world-class\ spa) + (2.026 * closer\ to\ residence) + (2.119 * HH\ Income) + (0.895 * distance\ to\ favorite\ casino) + (1.529 * distance\ to\ resort/casino)$



## Casino Segmentation Study

---

The following table shows the cross-classification matrix for the resulting discriminant model. As seen, the six-variable model correctly classifies 71% of those who say that they are likely to very likely to go The Resort. In other words, based on respondents' initial response to the likelihood to visit question, their responses to the set of decision questions, household income, and distance to their favorite casino, two out of three are correctly classified.

On the other hand, the model classifies 29% of respondents who say that they are likely to very likely to go to the proposed resort as probable or tentative "non visitors." This does not imply a conflict; it suggests that this group's responses to the likelihood to visit question and the set of decision questions are "inconsistent" with those in the "profiled" group.

Classification Matrix

Likelihood to go	Predicted Group Membership	
	Likely to Very Likely	Somewhat to Not at All Likely
Likely to Very Likely	71%	29%
Somewhat to Not at All Likely	29%	71%



# Casino Segmentation Study

## Visitor Segment Profile

Based on the classification model, the visitor "profile" is ...

Visitor Profile  
[Discriminant Model, Likely visitors]

Profile Attribute	Segment - Most Likely to Visit
<u>Demographics</u>	
• Household Income	\$50,000 and over (68%*)
• Length of residence	3 to 20 years (68%), 2 to 5 years (41%)
• Age	35-54 (45%), 45 + (44%), 55 or over (23%)
• Gender	Indecisive
• Length of Residence	3 to 10 (46%), 3 to 20 years (63%)
• Distance to Favorite Casino	6.8 miles [5% trim for outliers]
• Distance to resort/casino	3.6 miles [5% trim for outliers]
"Typical" Casino Spending	\$36 and up (80%), \$100+ (24%)
Visits per month	3 to 5 times (28%), 3 times or more (35%)
Main Reason go to Casinos	Gamble - slots, table games, etc (26%)
<u>Decision Characteristics, Casino Selection **</u>	
• Good Customer Service	94%
• Restaurants, Lounges, Movie Theaters	90%
• Convenient Location	73%
• Have "loose" Slot Machines	67%
• Provide "comps" for "preferred" Customers	59%
• Type of Slot Machines	55%
• Closer to Residence	50%
<u>Recreational Activity(ies) Enjoy Most</u>	
• Playing Golf	12%
• Hiking	11%
• Exercising or Working Out	10%

"\*" Shows the percentage in segment attribute, proportion is higher than not likely segment

"\*\*" Proportion denoting the attribute being an important characteristic, at least 50% noted



### *Factors for Success*

The table on the next page displays the results from the Factor Analysis on a series of questions relating to the respondent's "favored" casino. The result identifies two "critical" elements listed in order of their relationship to each other: Customer Service and "Physical Appearance" and "Perks." [The reader is reminded that Factor Analysis does not offer any cause/affect relationship.]<sup>4</sup>

As seen, the key attributes relate to the way that the visitor is treated and the benefits provided. Both are consistent with the previous profile. Factors having a score of .70 or higher should be considered essential to meeting customer acceptance.

---

<sup>4</sup> The Kaiser-Meyer-Olkin Measure of Sampling Adequacy measures the degree of inter-correlation among the variables and the appropriateness of Factor Analysis. The measure can be interpreted with the following guidelines: .90 and above, marvelous; .80 or above, meritorious; .70 or above, middling; .60 or above, mediocre; .50 or above, miserable; and below .50, unacceptable. The overall Adequacy Measure for the model is .862.

While a four-factor model is initially identified, the Reliability Tests using Cronbach's Alpha ( $\alpha$ ) finds that the three noted are the better descriptive of the four (Alpha scores of .900, .784 and .708, respectively followed by .534.)



# Casino Segmentation Study

---

## Factor Analysis [Discriminant Model, Likely visitors]

Statement	Factor Score **
<u>Customer Service</u>	
Employees At Favorite Casino Are Always Willing To Help	.801
Employees At Favorite Casino Strive To Give Customers Individual Attention	.741
Employees At Favorite Casino Try Best To Solve Customer Problems	.733
Employees At Favorite Casino Are Never To Busy To Respond To Request	.724
Employees At Favorite Casino Are Consistently Courteous And Friendly	.722
Service At Favorite Casino Is Consistently Good	.718
Employees At Favorite Casino Provide Prompt Service	.716
Employees At Favorite Casino Correct Mistakes Quickly	.661
Employees At Favorite Casino Seem Competent	.655
Management At Favorite Casino Is Responsive If Player Has Problem	.465
I Get Food Value For The Money I Spend At Favorite Casino	.451
Security At Favorite Casino Is Excellent	.432
Food is always good	.405
Employees At Favorite Casino Greet Me By Name	.334
Adequate amount of convenient parking	.264
<u>Appearance</u>	
Exterior is attractive & inviting	.717
Kept clean & well maintained	.670
Interior décor is attractive & comfortable	.656
Signs are easy to follow	.485
Generally have a good time	.484
Has the most up-to-date slot machines	.457
Seats are comfortable	.439
Clothing worn by service people is attractive	.396
<u>"Perks"</u>	
Comp Policies At Favorite Casino Are Well-Known	.718
It Is Easy To Get Comps At Favorite Casino	.717
It Is Easy To Understand Benefits Of Favorite Casino Club Card	.681
Comps And Club Card Program More Generous At Favorite Casino	.544
Favorite Casino Is [Not] Very Noisy *	.467
Favorite Casino Is [Not] Very Crowded *	.396

\*\*\* Question altered to adjust for reversal

\*\*\*\* Scores represent the attribute's relative weight, the larger the score the more "significant" the respondent's opinion.

