

RESEARCH ARTICLE

Approach of CHAID Decision Tree Model to Sport Market: A case study on an Arena Football League franchise

Dr. S. Roger Park

Associate Professor

Co-Director of Graduate Program, Department of Sport & PE.

School of Education, Gonzaga University, WA, USA



Article Info

Article

Received: 25/11/2014

Revised from: 05/12/14

Accepted on: 07/12/14

Available online: 08/12/14

ABSTRACT

The main purpose of this current research is to gather data on current and potential customers/fans (i. e., demographics and psychographics information) and aid business decision in reducing the risks involved in the process of critical business decisions for Spokane-based Arena Football League (AFL) franchise, Spokane Shock. Furthermore, it is to identify and confirm the effectiveness of ticket type-based segmentation as a relevant tool in order to increase the operational income for Shock through the sequential tree analysis among the variables of sport fandom, fans' budgets/expenditures at the venue, family income, and ticket types. A total of 31 survey crews volunteered (from the research method class) for this current research and successfully collected surveys (a paper and pencil survey) from 191 Shock fans from two Shock home games. Each survey respondent was given a small rubber ball with Spokane Shock log on. The completed surveys were inserted into statistics programs of SPSS and Decision Trees 21.0 for the statistical analyses. The CHAID decision trees model suggested that fans making more than \$25,000 a year were more likely to spend at the venue than those making less than \$25,000 a year. In other words, Shock needs to diversify ticket options such as 3-game tickets, 5-game tickets, and 10-game tickets and other entertainment options so that light consumers can be escalated to the avid fans with increased sport fandom. Secondly, the decision trees found that season ticket holders and even complimentary ticket holders partially contributed to the operational revenues through parking, concessions, and merchandise sales.

©KY Publications

In 2005, entrepreneur Brady Nelson submitted a proposal to bring an arena football team to Spokane and league commissioner approved it (http://www.spokaneshock.com/team/staff/?staff_id=2). In fact, Spokane Shock began its inaugural season in 2006 at "Arenafootball2" (AF2), sub-division of Arena Football League (AFL) and became one of three expansion teams with Everett Hawks and Stockton Lightning. Spokane Shock accomplished both on-field and off-field success by winning division titles in all four seasons and Arena Cups in 2006 (<http://www.nationmaster.com/encyclopedia/Spokane-Shock>) and off-field success and by achieving 25 sell-outs of 32 games from 2006 to 2009 while playing at AF2. After 2009 season, AF2 filed a bankruptcy heavily due to its economic downturn and financial difficulty that made Shock become a member of AFL since 2010. However, Shock was not as competitive in AFL as they used to be both from on-field by being ranked 5th, 3rd, and 6th in 2010, 2011, and 2012, respectively and off-field sides by selling out only 1 game of 26 games from 2010 to 2012 (http://www.arenafan.com/teams/Spokane_Shock-149/history).

The Roles of Marketing Research

According to Bellenger (1979), marketing research serves as a critical component not only in decision-making process, but also in achieving the organizational goals and objectives. Tarka (2012) also asserted that marketing research helps the organization's management on the righteousness and correctness of their decisions to be undertaken at present or in the future. Even though the main purpose of marketing research varies depending on the needs and wants of the organization, it can be conducted in the following two ways. First of all, the organization collects and analyzes the external environmental factors such as political factors, economic climate, demographics, demand trends, technological factors, and competitor analysis (Beech & Chadwick, 2007). Secondly, the organization analyzes the internal factors such as human resources issues, financial status, and internal culture in order to increase the effectiveness and efficiency in a highly competitive market conditions. As mentioned above, both academic and business industry agree on that marketing research is helpful to the organization in their decision making process by minimizing the risk of inappropriate decisions. However, not every company conducts marketing research due to the following challenges. Firstly, marketing research cost money and the decision makers are not willing to spend money unless they have to. Secondly, the decision makers may not want to wait until the marketing research is complete and apply the results of marketing research into their decision-making process. Thirdly, the lack of understanding of the decision makers in statistics serve as a barrier and lastly, status quo mindset of the practitioners are willing to ignore the importance and value of the marketing research.

Customer Segmentation and CHAID Decision Trees

According to Fullerton (2007), consumers are classified into sub-groups based on the following factors of age, gender, occupation, income, attributes, and geographic location in order for the company to build a profile of consumers and thus serve customers better. Among several benefits to gain from the market segmentation, the most critical one can be to make the company choose its own level of marketing investment according to the expected return from each customer after being aware of the profitability of new customers as a result of the market segmentation. Secondly, once the main value of the customers is available to the marketers, the company will be able to convert a customer into a member of a higher profitability segment. This tactic can only be achieved after completely valuation and analysis of the features and behaviors of the customers for optimizing marketing investments (Tirenni, Kaiser, & Herrmann, 2007).

The marketers are quite satisfied with the results of the decision trees over other methods used in the academia such as neural networks and logistic regression. Particularly, from the practitioners' perspectives, the most valuable point is that decision trees model can help them easily categorize the customers and thus implement profit-maximizing marketing strategies. Therefore, decision trees model was chosen as an appropriate method for this current study since it was conducted as a form of research collaboration between the academia and industry sponsor.

Tirenni, et al. (2007) summarized the benefits of decision trees related to high flexibility over other methods. First, decision trees are nonparametric method-meaning no distributions or functional forms need to be specified. Secondly, pre-selection of variables are not required for decision trees and rather a robust stepwise selection process can be used. Thirdly, both continuous and categorical variables can be easily handled by regression trees because cases are partitioned and each group can be analyzed separately. Fourthly, regression trees are robust to the effects of outliers, and finally, decision trees can generate useful results even though data cannot generally be fitted well. Most importantly, it explains, "single rules can still perform extremely well and can greatly assist decision makers in improving the precision of their forecasts when the average performance of the output classification rules is poor (p. 133)."

The main purpose of this current research is to gather data on current and potential customers/fans (i. e., demographics and psychographics information) and aid business decision in reducing the risks involved in the process of critical business decisions for Spokane-based Arena Football League (AFL) franchise, Spokane Shock. Furthermore, it is to identify and confirm the effectiveness of ticket type-based segmentation as a

relevant tool in order to increase the operational income for Shock through the sequential tree analysis among the variables of sport fandom, fans' budgets/expenditures at the venue, family income, and ticket types.

METHOD

Participants and Procedures

The authors of this current study had an introductory meeting with Director of marketing of Spokane Shock to discuss the needs of business marketing research and the business research process and logistics such as research timelines, game scheduling for data collection, selection of survey crews, and budgets, etc. After the initial meeting, the research questions were created and developed which included the questionnaires of demographics and psychographics, sport fandom and the perceived satisfaction level of Shock promotions activities and events. The survey questions were submitted to Institutional Review Board (IRB) and approved by the primary author's institution. Once the survey instruments were finalized, the researchers recruited survey crews from three sport management classes on a voluntary basis (The survey crews were given the free game tickets of Spokane Shock) and the orientation was held to inform them with their duties and responsibilities.

A total of 31 survey crews volunteered (from the research method class) for this current research and successfully collected surveys (a paper and pencil survey) from 191 Shock fans from two Shock home games (Table 1). Each survey respondent was given a small rubber ball with Spokane Shock log on. The completed surveys were inserted into statistics programs of SPSS and Decision Trees 21.0 for the statistical analyses.

Instruments

The survey instruments consist of three parts: (1) demographics with 6 items and psychographics information with 14 items, (2) sport fandom questionnaire with 5 items developed by Wann (2002), and (3) other promotion-related questionnaires with 10 items. A panel of experts ($n = 6$), including marketing director and manager of Spokane Shock, three research assistants, and one professor in sport management conducted a test of content validity in terms of item relevance and clarity. Each of these items was phrased into a statement anchored by a Likert 5-point scale (1 = strongly disagree to 5 strongly agree).

RESULTS

Descriptive Statistics Information

A total of 191 Shock fans (male = 107; female = 68; no answer = 16) completed the survey as shown in Table 1. The majority of fans were White (86.8%) and African American and Asians were 4% each. Sixty-seven fans were, season tickets, 73 fans were single game ticket holders, and 34 fans were complementary ticket holders. Forty-nine fans' family incomes were more than \$75,000 a year followed by \$50,000-\$75,000 and \$25,000-\$50,000. The average budget on parking, concession, and merchandise was fairly spread out from "less than \$10" to "more than \$40."

Table 1. Descriptive statistics for background variables of spectators of Spokane Shock

Variable	Category	N	%
Gender	Male	107	61.1%
	Female	68	38.8%
Ethnicity	African American	7	4%
	White	152	86.8%
	Hispanic	2	1.1%
	Asian	7	4%
	Indian	3	1.7%
	Other	4	2.2%
Ticket Type	Season Ticket	67	38.5%
	Single Game Ticket	73	41.9%
	Complimentary Ticket	34	19.5%
Family Income	Less than \$15,000	25	15.5%

	\$15,000 - \$25,000	16	9.9%
	\$25,000 - \$50,000	28	17.3%
	\$50,000 - \$75,000	43	26.7%
	\$75,000 or more	49	30.4%
Average Budget	Less than \$10	25	14.2%
	\$10 - \$20	40	22.8%
	\$20 - \$30	45	25.7%
	\$30 - \$40	28	16%
	\$40 or more	37	21.1%
Shock Game Attendance	0-1 game	19	10.7%
	2-5 games	36	20.4%
	6-10 games	35	19.8%
	11-20 games	47	26.7%
	21 games or more	39	22.1%
Follow Shock Facebook	Yes	88	47.1%
	No	99	52.9%
Follow Shock Twitter	Yes	34	18.3%
	No	152	81.7%
Have one or more pets at home	Yes	117	62.6%
	No	70	37.4%

As you can see Table 2, the majority of fans were satisfied [by combining “agree” and “strongly agree”] with Shock’s promotional activities performed at the venue such as volume of music (76.3%), pregame instructions (66.1%), half-time shows (64.5%), Shock dance team (83.8%), game announcer (86.1%), and mascot (80.3%).

Table 2. Marketing and Promotion-Related Data

Variable	Category	N	%
Volume of Music at the Arena	Too Loud	27	14.5%
	Just Right	142	76.3%
	Too Soft	17	9.1%
Pregame Instructions were well made and delivered	Strongly Disagree	4	2.4%
	Disagree	4	2.4%
	Neutral	49	29.2%
	Agree	52	31%
	Strongly Agree	59	35.1%
Half-time Show was great	Strongly Disagree	4	2.4%
	Disagree	5	3%
	Neutral	51	30.2%
	Agree	45	26.6%
	Strongly Agree	64	37.9%
Shock Dance team performed well	Strongly Disagree	4	2.3%
	Disagree	4	2.3%
	Neutral	20	11.6%
	Agree	60	34.7%
	Strongly Agree	85	49.1%
Game Announcer performed well	Strongly Disagree	4	2.3%
	Disagree	3	1.7%

Mascot performed well	Neutral	17	9.8%
	Agree	48	27.7%
	Strongly Agree	101	58.4%
	Strongly Disagree	3	1.7%
	Disagree	5	2.9%
	Neutral	26	15%
	Agree	50	28.9%
	Strongly Agree	89	51.4%

The first research question of this current study was to identify if avid Shock fans are more likely to have season tickets than single game and complimentary tickets.

Analysis of variances (ANOVA) was conducted to identify if sport fandom level was different by the ticket type (Season ticket holders, single ticket holders, or complimentary ticket holders). The Post-hoc Analysis (Tukey HSD) showed that season ticket holders were more likely to have significantly higher levels of sport fandom toward Spokane Shock than single game ticket and complimentary ticket holders, $F(2, 163) = 9.46, p < .001$. No other differences were found on other variables.

The second research question was to see if Shock fans' family income level and their ticket type would predict their average budget on parking, concession, and merchandise at the venue. Market segmentation was analyzed through the classification tree method (Breiman et al, 1984), using the software Decision Tree (SPSS version 21.0). The average expenditure of Shock fans on parking, concession, and merchandises at the venue was defined as the target variable, and annual family income and a type of ticket (single game ticket, complimentary ticket, or season ticket) formed the set of potential predictor variables. Chi-squared Automatic Interaction Detection (CHAID) was used in order to produce homogeneous subgroups or segments (nodes) of respondents with similar values on the target variable (Soldic-Aleksic, 2012). The target and predictor variables were composed of any mixture of nominal, ordinal and interval scale variables. At each level of the tree, the sample was grouped by the predictor variable on the basis of Gini coefficient. The measure takes the value 0 if the distribution is concentrated on a single category while it takes its maximum value when the distribution of the target variable is uniform, with equal proportions in each category.

The main specifications of the applied CHAID model are presented in the Table 4 and an appropriate decision tree is presented in the Figure 1. Table 3 presents the main components of the CHAID model specifications and results. More specifically, information used on the tree model specifications includes variables used in the analysis and maximum tree depth and minimum cases in parent and child nodes while the results section contains such information as the independent variables included in the model, number of final and terminal nodes, and depth of the trees. Two of four independent variables, family income and ticket type, were specified and this can be interpreted as these two variables make a significant contribution to the created tree model.

Table 3. CHAID model summary

Specifications	Growing Method	CHAID
	Dependent Variable	Average budget to spend at the venue
	Independent Variables	Family income, ticket type, gender, and education
	Validation	
	Maximum Tree Depth	3
	Minimum Cases in Parent Node	100
Results	Minimum Cases in Child Node	50
	Independent Variables Included	Family income and ticket type
	Number of Nodes	5

Number of Terminal Nodes	3
Depth	2

The sample was split first by the predictor variable *annual family income*, into respondents making “less than \$25,000 a year” and those making \$25,000 - \$75,000 a year. The first level of Tree Model explained that 38.2% of fans making “less than \$25,000 a year” are more likely to spend between \$10 and \$20 at the venue for concession and merchandises while 29.7% of fan making between \$25,000 and 75,000 a year are more likely to spend “more than \$41” at the venue, $\chi^2 = 37$, *Adjusted p* = .001.

The fans making between \$25,000 and 75,000 were again split into two groups of single game ticket (including complimentary ticket holders) holders and season ticket holders. According to the second level of Tree Model (see Figure 1), 28.3% of single game or complimentary ticket holders have an average expenditure range of \$21-30 while 39.2% of season ticket holders are willing to spend more than \$41 at the venue that is significantly higher than the single game or complimentary ticket holders (21.7%).

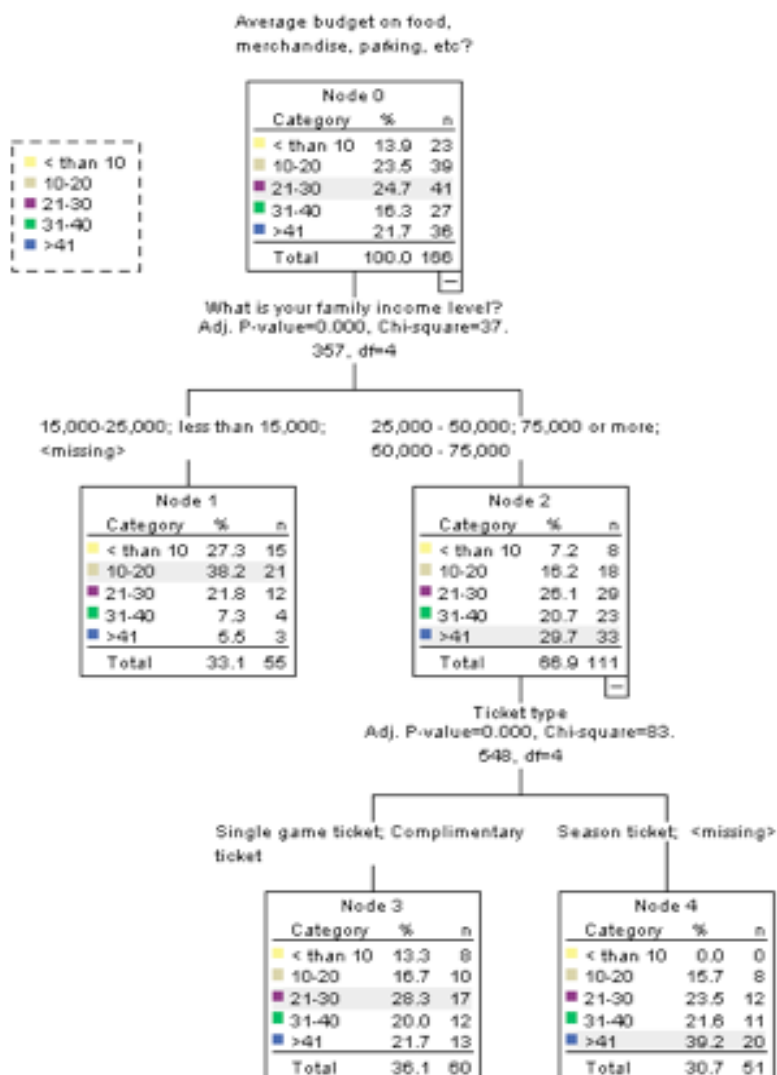


Figure 1. Tree Diagram Illustrating Fan Segmentation

Managerial Implications

According to the results of ANOVA, it is reasonable to conclude that season ticket holders have higher level of sport fandom toward Spokane Shock than single game and complimentary ticket holders. This result indicates that Shock fans that have a higher level of sport fandom are more profitable than lower level of sport fandom holders. Therefore, improving the retention rate of fans with higher level of sport fandom would have a strong leveraging effect on improvement of the whole customer equity. From a practical perspective, Shock should develop the unique and exclusive marketing and promotional programs for a higher level of sport fandom holders to maximize the profit for Shock.

The CHAID decision trees model described in the previous section can help marketers allocate marketing resources on the basis of the predicted profitability levels. Identification process of features and behaviors of new ticket holders must be completed to enable to allocate the marketing resources based on the customer's expected value. For high-valued fans with season ticket holders and high sport fandom can be targeted by appropriate loyalty programs/campaigns. Evidently it is efficient to spend the highest per head amount on these fans. Shock needs to develop very exclusive entertainment options for them such as up-scale dining options, high quality merchandise items, or other unique fan experience programs for the loyal fans.

Secondly, fans making more than \$25,000 a year are more likely to spend at the venue than those making less than \$25,000 a year. In other words, Shock needs to diversify ticket options such as 3-game tickets, 5-game tickets, and 10-game tickets and other entertainment options so that light consumers can be attracted to the loyal fan level with an increased sport fandom. Lastly, the decision trees found that season ticket holders and even complimentary ticket holders partially contributed to the operational revenues through parking, concessions, and merchandise sales. Thus, it confirmed that complimentary ticket holders also proved to be still valuable and meaningful in contributing to the overall revenue.

REFERNCES

- Bellenger, D. N. (1979). The marketing manager's view of marketing research. *Business Horizons*, 22, 59-65.
- Tarka, P. (2012). Modeling marketing research efficiency. *Econometrics*, 1(35), 38-43.
- Beech, J., & Chadwick, S. (2007). *The marketing of sport*. London, England: Prentice Hall.
- Fullerton, S. (2007). *Sports marketing*. New York, NY: McGraw-Hill.
- Tirenni, G., Kaiser, C., & Herrmann, A. (2007). Applying decision trees for value-based customer relations management: Predicting airline customers' future values. *Database Marketing & Customer Strategy Management*, 14(2), 130-142.
- Wann, D. L. (2002). Preliminary validation of a measure for assessing identification as a sport fan: The sport fandom questionnaire. *International Journal of Sport Management*, 3, 103-115.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. I. (1984). *Classification and Regression Trees*. Belmont, Calif.: Wadsworth.
- Soldic-Aleksic, J. (2012). Combined approach of Kohonen Som and CHAID decision tree model to clustering problem: A market segmentation example. *Journal of Economics & Engineering*, 3(1), 20-27.