Sports Analytics 2020

MSBA Atlanta Sports Survey Assignment

Market Research Proposal

Susie Bai, Ricky Chen, Frank Fan, Mayur Kumar, Daniel Weinberg & Yaping Zhang

**1. Survey Overview**

In the five sections of last year’s survey, the theme for each section was redundant; many similar questions appeared in different sections. Here is a possible summary for each section:

1. Team preferences, media consumption and sports-related product consumption
2. Team Evaluation, Reason of fandom and game experience
3. Media consumption and sports-related product consumption
4. Esports and Gambling
5. Demographic Information

For the new revised version of the survey, each section should be clearer about its particular topic. The newly designed sections should be as follows:

1. Sports preferences, media & game consumption (Information about a sports fan)
2. Star Power related question
3. Generational Fandom
4. Gambling
5. Demographic Information

We combined the original 1st and 3rd section into one section, and used different themes to distinguish this section from the others.. We restructured the original 2nd section into questions regarding Star Power and Generational Fandom. We also adjusted the focus of the 4th section to elucidate more relevant and useful details regarding sports gamblers. Finally,, we kept the 5th section relatively the same in order to understand the the various backgrounds of those who will take our survey.

**2. Survey Questions**

**a. Star Power**

1. On a scale of 1 to 5, with 5 being highly important, to what magnitude does the team of the opposing star influence your likelihood to attend a basketball game?
   1. Home Team
   2. Opposing Team
2. How many games have you attended where one of the major reasons to attend was to watch a particular star athlete?
3. On a scale of 1 to 5, with 5 being the highest, please rate your experiences watching star athletes in terms of how likely you are to attend a future game with that same star?
   1. How often would you attend a game if there is a star player playing?
4. How many star athletes in a game would make you attend the game?
5. Do you own any apparel (jersey, poster, etc.) of any star athletes?
6. What would increase the likelihood of you purchasing team-related merchandise for any of the professional sports teams in Atlanta? (Select all that apply)
   1. Star Player
   2. Team playing an important game
   3. Performance of the team in the league till that point or in the last seasons
   4. Stronger sense of community around that team
   5. Improved team uniform Discount on merchandise at the time of purchase

**b. Generational Fandom**

1. Which of the following sports do your relatives watch? (select all that apply)?
   1. Basketball (college or professional)
   2. Baseball (college or professional)
   3. Boxing/UFC
   4. Esports
   5. Football (college or professional)
   6. Golf
   7. Hockey
   8. Other
   9. Soccer
2. In any given week, how often does you and your relatives watch sportsTV or other online streaming services?
   1. Never
   2. Only big events
   3. 1-3
   4. 3-4
   5. 5-6
   6. 7+
3. In a year, how often does you attend sporting events with your relatives?
   1. Never
   2. Rarely
   3. Sometimes
   4. Frequently
   5. All the time
4. How much money would you estimate your relatives spend on sports-related products (e.g. tickets, jersey)?
   1. Less than $100
   2. $100 to $400
   3. Greater than $400
5. Do you root for different teams than the ones your family roots for?
   1. Yes
   2. No
6. Is your favorite league different from the one your relatives like the most?
   1. Yes
   2. No
7. Do you have children?
   1. Yes
   2. No

If yes to #8, answer questions 8-10

1. Would you like to take your kids with you when heading to a game?
2. Have you ever attended a sporting event with one or more of your children?
   1. Yes
   2. No
3. How would you generally describe your children’s attitude toward sports, relative to your own fandom?
   1. Less interested
   2. About the same
   3. More interested
4. Please select all of the following factors that encourage you and your family to root for a team?
   1. Hometown team
   2. # of winning games
   3. # of titles
   4. # of fame hall members
   5. # of all-star players
   6. Connection to a specific player or coach

**c. Gambling**

1. Have you bet on sports within the last 12 months?
   1. Yes
   2. No
2. Have you bet on sports prior to the last 12 months?
   1. Yes
   2. No
3. Do you have friends or relatives who have gambled within the last 12 months?
   1. Yes
   2. No
4. Do you live within 50 miles of a casino?
   1. Yes
   2. No
5. Do you consider gambling to be immoral?
   1. Yes
   2. No
6. Do you believe gambling should be illegal?
   1. Yes
   2. No

If yes to #3, answer questions 9 through 16

1. On a scale of 1 to 5, with 1 being almost never and 5 being daily, please describe your gambling activity in the past year?
2. What platform(s)/methods do you use to gamble? (select all that apply)
   1. Online betting with a bookie
   2. Daily Fantasy Sports
   3. Seasonal Fantasy Sports
   4. One-time events (ex: super bowl boxes)
   5. Offline betting with a friend/family/colleague
   6. Other
3. What is your annual income?
   1. Less than $29,999
   2. $30,000 to $69,999
   3. $70,000 to $99,999
   4. $100,000 to $199,999
   5. $200,000 to $499,999
   6. Greater than $500,000
4. If you bet online with a bookie, how much money is your typical unit bet?
   1. $20-$50
   2. $50-$100
   3. $100-$200
   4. $200-$500
   5. $500-$1,000
   6. $1,000-$3,000
   7. $3,000-$10,000
   8. $10,000-$50,000
   9. $50,000-$100,000
   10. $100,000-$1,000,000
   11. $1,000,000+
5. What do you gamble on? (select all that apply)
   1. Basketball (college or professional)
   2. Baseball (college or professional)
   3. Boxing/UFC
   4. Esports
   5. Football (college or professional)
   6. Golf
   7. Hockey
   8. Other
   9. Soccer
6. On a scale of 1 to 5, with 1 being “I would struggle” and 5 being “Anytime”, please describe how you would respond to the statement: “I can walk away from gambling at any time.”
7. Why do you choose to gamble? (select all that apply)
   1. Interest in sports/fandom
   2. Financial reasons
   3. For the thrill/enjoyment
   4. Escape from problems/depression/boredom
   5. Other
8. Do you consider gambling to be a source of income?
   1. Yes
   2. No
9. Have you ever had to ask somebody for money to pay off a gambling debt?
   1. Yes
   2. No
10. On a scale of 1 to 5, with 1 being “Strongly Disagree” and 5 being “Strongly Agree”, please describe how you would respond to the statement: “My chances of winning are better after I lose.”

**d. Media and game consumption**

1. Through which channel do you often watch sports games?
   1. TV
   2. On site
   3. Website
2. What kinds of items would you purchase when you are watching sports games on site?
   1. Sports equipment
   2. Clothes
   3. Artware
   4. Other
3. How much time do you spend watching games per week?
   1. Less than 1 hour
   2. 1-10 hours
   3. 10-30 hours
   4. More than 30 hours
4. How much money do you spend on sports per year?
   1. $20-$50
   2. $50-$100
   3. $100-$200
   4. $200-$500
   5. $500-$1,000
   6. $1,000-$3,000
   7. $3,000-$10,000
   8. $10,000-$50,000
   9. $50,000-$100,000
   10. $100,000-$1,000,000
   11. $1,000,000+
5. How much money do you spend on food in an on-site game?
   1. $0
   2. Less Than $10
   3. $10-$20
   4. $20-$50
   5. $50-$100
   6. $100-$200
   7. $200-$500

**3. Analysis Approach**

1. **Star Power Analysis**

* Goal: understand the linkage between star players and fan engagement
* Questions Related: ‘Star power’ Section questions 17 and questions about purchasing behavior from the original question set
* Analysis Design:
  + Regression analysis: we want to use star power, as measured by the survey questions, as a predictor in our regression analysis, using the probability of purchasing behavior as the target variable. While controlling for other factors such as income, we will be able to analyze how much star power contributes to purchasing decisions. One possible modelling approach is a logistic regression model.

1. **Generational Fandom Analysis**

* Goal: understand how generational fandom influences individual fandom; understand what factors contribute to generational fandom
* Questions Related: ‘Generational fandom’ Section questions 1 through-10
* Analysis Design:
  + Regression analysis: Use the results of questions 2,3 and 4 as main features to indicate the degree of generational fandom, and use the results of question 1,5 and 6 as the target variable. The regression result will indicate the impact and the level of impact of generational fandom on individual fandom. Possible modelling approaches includelinear regression and K-NN.
  + Descriptive analysis: From the pool of users who have the same sport team/league with their relatives, use the survey result of question 10 to visualize and determine what reasons contribute to generational fandom the most.

1. **Gambling**

* Goal: discover the influences of gambling in various aspects related to sports (e.g. ethics); investigate whether the rising trend of gambling activities and possibly the changes/advances in technologies in this field would impact fandom (currently and historically) and in which ways
* Questions Related: “Gambling” section question 1 through 18.
* Analysis Design:
  + Firstly, we consider doing some exploratory data analysis (such as descriptive statistics based on people’s responses in this section), for the sake of figuring out how the current participation of gambling in Atlanta is and what the dynamics of local gambling activities have been so far, as well as extracting some possible insights on the changing landscape of people’s gambling behaviors and channels/tools/technologies used to gamble from a historical perspective. This is basically to have a big picture of the topic of Atlanta’s gamling and to facilitate deeper understanding afterwards;
  + Then we may be interested in discovering the relationship between gambling and some social or individual phenomenon in relevant critical topics, such as ethics, and socioeconomic issues
  + Lastly and importantly, we want to inspect the impact of this kind of “gambling” events and its many evolutions on local people’s cognition and expression of “fandom” nowadays, and thus on the sports business:
    - Unsupervised learning: for example, clustering of sports consumers with different levels & types (characteristics) of gambling behavior; possible approach: k-means
    - Supervised learning:
      * Classification: whether certain event/behavior of a consumer would happen/in what categories of certain event/behavior would happen (binary or multi-class target variable) given this added factor of “he/she gambles” (e.g. “whether a customer would be on-site to watch a game if he/she engages in online betting with a bookie”)
      * Regression: predict the value of certain continuous outcome variable given the presence of this factor of gambling (e.g. “how much would a customer still be willing to pay for team souvenir clothes if he/she lost $1000+ on gambling for a game”)
    - Statistical analysis:
      * E.g. How significant would a specific aspect of gambling behavior be affecting a customer’s loyalty to a team

1. **Media and Game Consumption**

* Goal: understand the current and changing ways & media through which Atlanta people consume sports games; the characteristics, impacts (especially on the business side, such as profitability, or competitiveness) and development of different kinds of sports media, and what it means for the transformation and current conditions of “fandom”
* Questions Related: “Media and game consumption” section question 1 through 5
* Analysis Design:
  + Summary Statistics based on survey responses and Descriptive Analysis: to understand people’s consumptions styles, changing preferences and habits with regard to sports media in a collective and further break-down manner
  + Predictive Analysis:
    - Regression: e.g. use one’s consumption level/participation/money spent as a target variable, and use people’s habitual media for consumption as one of the indicators, make predictions based on past data
    - Classification: e.g. predict whether a customer would buy ticket for a certain game or not if that game is arranged to be live on Youtube/Facebook/any other sport-fan go-to website (as one factor)

Some possible models to employ are linear regression, logistic regression, decision tree and Naive Bayes.

* + Statistical Analysis: e.g. explain the statistical effect of spending time watching games via a specific sports medium on a person’s interest in a particular category of sports(e.g. Basketball, baseball, etc.)