**Christina Huffaker**

**Sujoy Datta Choudhury**

**Zaoxian Liu**

**DSC 611: Data Visualization**

**Final Project Report**

**Project Title: Economical and Effective Sports Sponsorships by Oculytica**

**Abstract**

**Introduction**

Significant competitors provide branded team apparel to high-school and college athletic teams, particularly successful teams with a record of wins (Jensen, Wakefield, Cobbs, & Turner, 2016 p. 289-90). This approach is overly costly based on projected ROI, wasting valuable resources with this model, and an alternative model for athlete sponsorships is urgently needed. Oculytica, LLC will outline which Olympic teams to target for sponsorship, using different criteria to engage high-achieving athletes early in their careers, and expose the brand to a wide potential customer base.

**Related Work**

Jensen et al. (2016) writes on the ROI for sporting goods companies on large investments in sponsorships, determining that most companies are unwisely paying more than expected ROI. Wang & Kaplanidou (2013) describe the impact of emotion-lifting on spectators of sporting events, and that even negative emotions induced by sports (such as a team loss) can be leveraged to generate purchases. Research by Mazodier, Corsi and Quester (2016) shows that advertising messages typical of an event are effective in transferring associations of an event to a brand. Taken together, this research presents an alternative, and not yet popular, tactic for sporting good companies’ marketing strategies. In review of literature, no studies were identified in which a company synthesized the aforementioned works. Oculytica Analytics, in using these works to inform and guide this analysis, hopes to generate a new vein of scholarship on sports sponsorship.

**Data Description**

The Kaggle Dataset “120 Years of Olympic History: Athletes and Results” (rgriffin, 2018) provides 136,000 observations on 15 variables. In preparatory data processing, the following variables were selected: Sex, Age, Height, Weight, Team, Year, Season, City, Sport, Medal, Country, and Continent. The data was provided in .csv format, allowing it to be easily imported into R and Tableau without modifying the format. Each row in the data corresponds to an individual athlete competing in an individual event (rgriffin, 2018). After data cleaning was performed, a new dataset containing 85,256 rows with 11 features Name, Sex, Age, Height, Weight, Country, Year, Season, NOC, Sport, and Medal was created. NOC column was joined with a separate file with Country Code for plotting WorldMap.

**Methods**

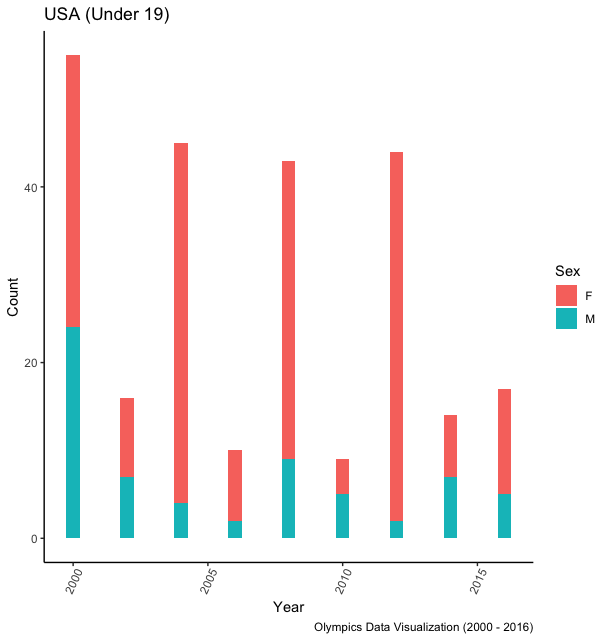
For this analysis, the data was first retrieved from the Kaggle website. Originally, the data was scraped from the website Sports Reference ([https://www.sports-reference.com](https://www.sports-reference.com/)) by the user rgriffin in May 2018 (rgriffin, 2018). Directly downloaded by the Oculytica team from Kaggle, the data required pre-processing to prepare for the analysis.

Following the data cleaning and pre-processing, several preliminary analyses were performed, which were used to inform the most useful focus for the final analysis. The concept to focus on athletes under age 19 arose from the literature. Examples of more high-profile teams were generally assumed to be comprised of older athletes, and sponsorship of these athletes more costly. Examples of more inexpensive sponsorships, which would generate a more cost-effective ROI, tended to center on younger teams, mainly in high-school, or in college.

**Results**

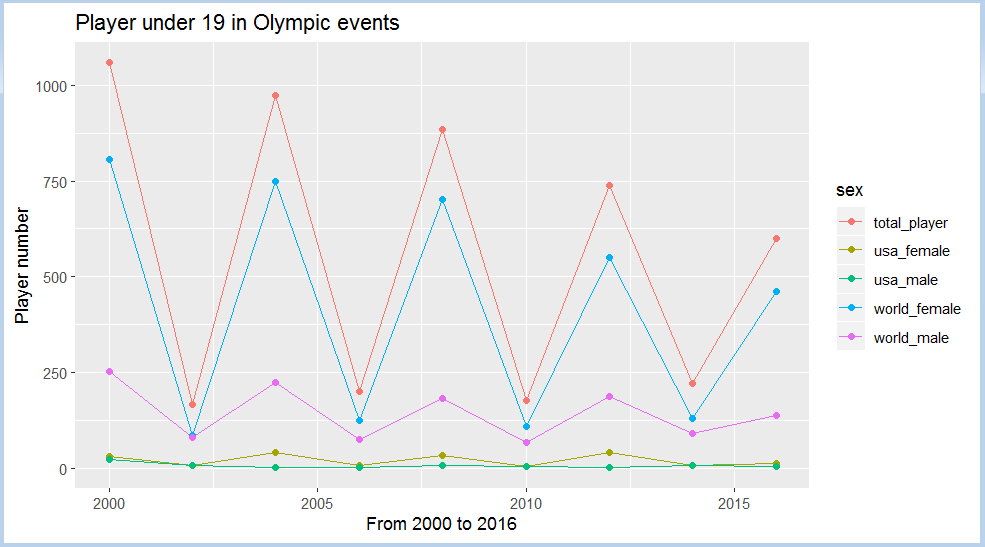
In preliminary analyses, the demographics of athletes under age 19 were explored.

Preliminary bar chart results indicate number of medal winners based on gender demographics, shown in Figure 1. This figure shows two noteworthy features. First, the gender demographic, dividing the athletes into categories of males and females, reveal that there have been more female participants under age 19 in nearly every year displayed in Figure 1, but during the Winter Olympics, which occurred in 2002, 2006, 2010, and 2014, the disparity is much smaller than during the Summer Games, which occurred in 2000, 2004, 2008, 2012, and 2016. It is also important to note the decline in total number of medal-winning athletes under 19 from 2012 to 2016. Though the direct cause of this decline is unclear, at this point, it could indicate a lack of participants in this age category, medal-winning or non-medal-winning alike.



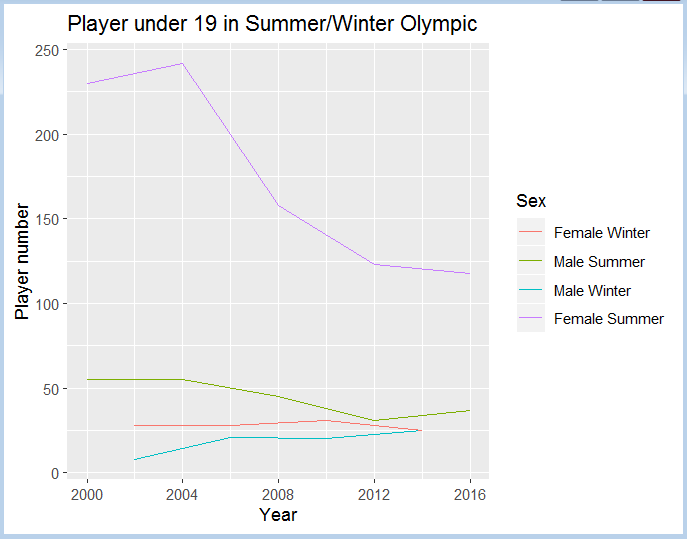
**Figure 1.** Medal-winning Olympic athletes under age 19 from years 2000-2016, displaying gender for each year

The chart in Figure 2 echoes the same information displayed in Figure 1, as the decrease in the number of U.S. athletes under 19 participating in the 2016 Olympics is clear. However, some additional information can be gathered from Figure 2, which displays not only the total gender demographics, but those compared to athletes from the US, and the rest of the world. Figure 2 indicates that there is a greater number of U.S. females under 19 participating than males in the same demographic. Athletes of this age group could be under-resourced, and good candidates for sponsorships.



**Figure 2.** Olympic athletes under age 19, total, from the US, and from the rest of the world, from years 2000-2016, displaying gender for each year

Participation of U.S. male and female athletes under 19 look more similar to each other than in other countries, but female athletes in the world overall participate far more than male athletes. The sharp decline in the number of Olympic athletes under age 19, particularly during the Summer games, is visible in Figure 3, as well.



**Figure 3.** Male & female athletes in recent Summer and Winter Olympic Games

**Discussion / conclusions**

Initial sponsorships for this demographic will ensure that athletes are provided with resources, likely increasing numbers of competitors in this age group and enhancing brand visibility. While both male and female groups will benefit from sponsorships, investing more in male athletes will likely increase numbers of male participants in future Olympics, similarly enhancing brand visibility. Focusing on under-represented groups will minimize the cost of sponsorships to the company (Jensen, et al., 2016 p. 289-90), in turn creating greater ROI, but not negatively impact branding if emotion-lifting tactics are employed (Wang & Kaplanidou, 2013, p. 9).

**References**

Jensen, J., Wakefield, L., Cobbs, J., & Turner, B. (2016). Forecasting

sponsorship costs: Marketing intelligence in the athletic apparel industry. *Marketing Intelligence & Planning,* *34*(2), 281-298.

Mazodier, M., Corsi, A., & Quester, P. (2018). Advertisement typicality: A longitudinal

experiment - can sponsors transfer the image of a sporting event to their brand? (what we know about sports sponsorships). *Journal of Advertising Research,* *58*(3), 268.

rgriffin. (2018). 120 years of Olympic history: athletes and results [dataset]. Retrieved from

<https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results>

Wang, R., & Kaplanidou, K. (2013). I want to buy more because I feel good: The effect of sport

induced emotion on sponsorship. *International Journal of Sports Marketing &*

*Sponsorship,* *15*(1).

Appendix A

GitHub Project Link: <https://github.com/sujoydc/DS-611-Project>

**Appendix A**

