**Christina Huffaker**

**Sujoy Datta Choudhury**

**Zaoxian Liu**

**DSC 611: Data Visualization**

**Final Project Milestone # 1**

**Introduction**

The Oculytica project will allow companies like Nike, Adidas, and Under Armour to capitalize on specific customer demographics, in addition to reaching previously underused athlete populations.

**Statement of the Problem**

Competitors have demonstrated that high-school and college-aged athletes are essential resources used widely to increase visibility and brand loyalty in young athletes by providing branded team apparel to as many high-school and college teams as possible, particularly successful teams with a record of wins. But research shows that this approach is overly costly based on projected ROI, and companies are wasting valuable resources with this model [1]. Our solution will increase immediate and future revenue.

**Solution**

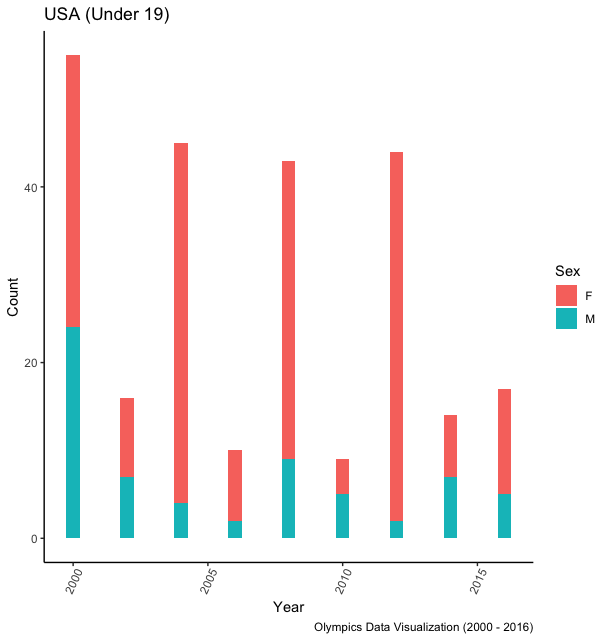
Oculytica provides reports outlining which players, age groups, and sports to advertise in specific locations, using different criteria than competitors. Targeting high-school and college athletes in countries which typically produce Olympians in these age-groups will engage high-achieving athletes from an early point in their potential careers, in addition to exposing the brand to wide potential customer base.

**Data**

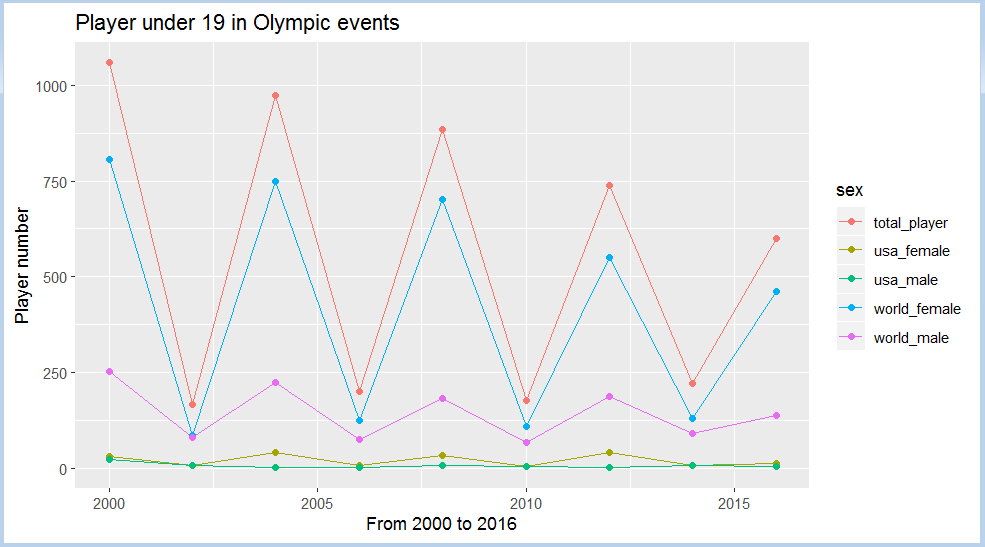
The Kaggle Dataset “120 Years of Olympic History: Athletes and Results” [2] provides 136K observations on 15 variables. In preparatory data processing, the following variables were selected for analysis: Sex, Age, Height, Weight, Team, Year, Season, City, Sport, Medal, Country, and Continent.

**Methods / Preliminary Results**

Preliminary bar chart results indicate number of medal winners based on demographics.



The chart to the left indicates a decrease in the number of U.S. athletes under 19 participating in the 2016 Olympics. It also 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.



Participation of U.S. male and female athletes under 19 look more similar, but female athletes in the world overall participate far more than male athletes. The drop in 2016 is present here, too.

Next, research relevant to the project will be reviewed, as well as a deeper analysis with additional visualizations to identify more patterns in the data. Market analysis to evaluate competitors’ strategies will inform which demographics will be the primary focus.

**References**

[1] Ratten, V., Madichie, N., 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.

[2] 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>

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

**Appendix A**

