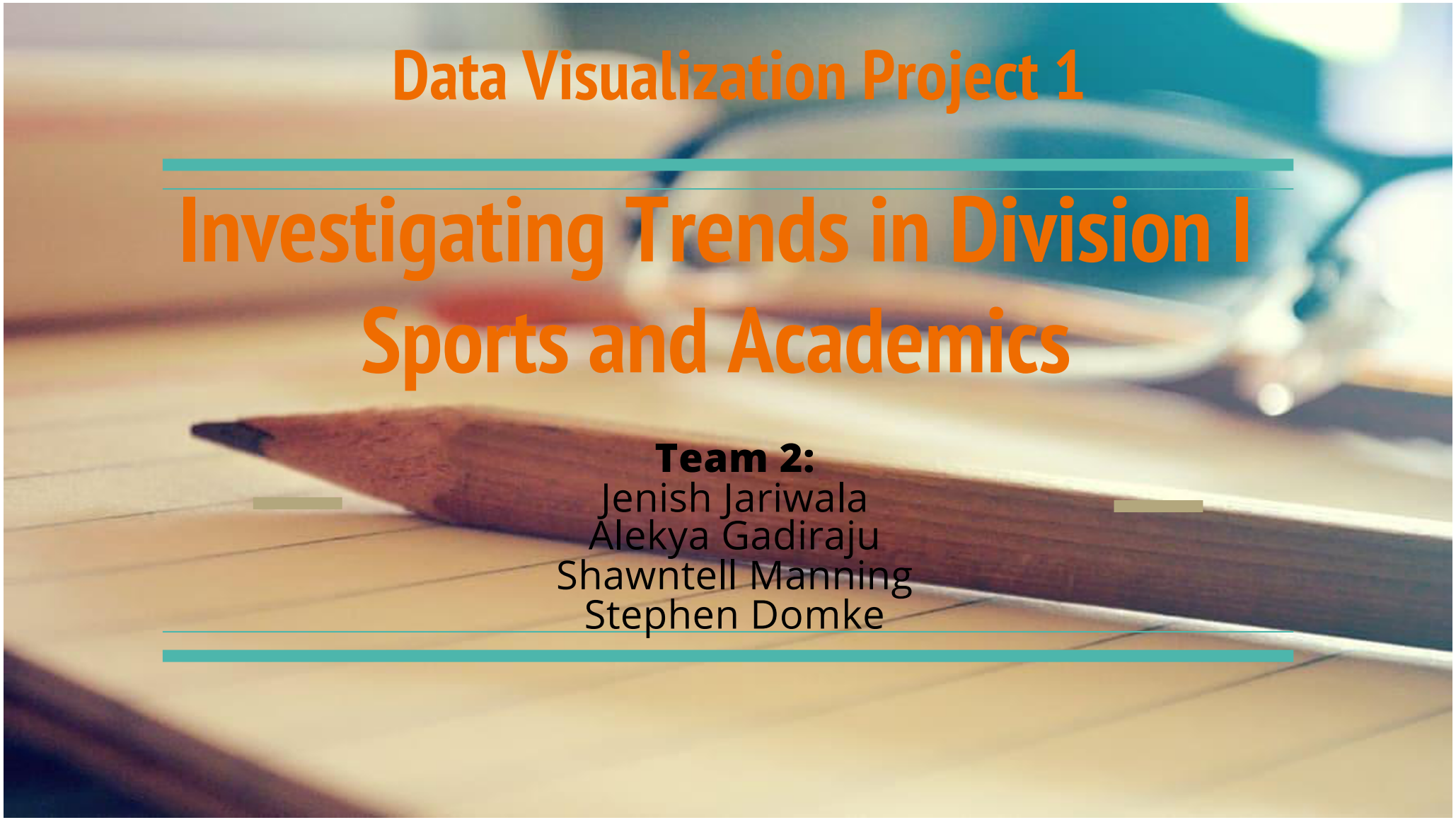


Data Visualization Project 1

Investigating Trends in Division I Sports and Academics

Team 2:
Jenish Jariwala
Alekya Gadiraju
Shawntell Manning
Stephen Domke

Motivation and Summary

- **Hypothesis:** Student-athletes perform better academically compared to the greater student body.
- **Investigation:** Trends and correlations tied to sports involvement and academic performance.
- **Source:** NCAA Division I Colleges
- **Subjects:** Student Athletes vs. Student Body
- **Geography:** US National scope

Questions and Data

- Is there a correlation between sports involvement and academic performance?
- How does academic performance vary between different sports?
- Do any outliers that exist?
- Are there any significant differences based on gender?
- How does the academic performance of student athletes compare with student bodies as a whole?
- Are there any geographical correlations to discover?

Data Cleanup and Exploration

- Had access to multiple years of data (2002-2019)
- Noticed some missing data up to 2015
- Used latest year data for analysis (2019 for NCAA, 2018 for Federal)
- Unexpected complexity with measurements of academic success.
- Datasets use different measures, based on subject base.
- Includes APR, GSR & Federal GRS

A brief explanation of these measurements follows...

Federal Methodology: Graduation Rate Survey

Measures: Retention as the proportion of students who entered a school and graduated from that same institution within six years.

Limitation: Does not include transfers. Consider them as non-graduates from both the college they left and the one they eventually graduate.

Source: https://www.ncaa.org/about/resources/research/graduation-rates

The US Department of Education uses the Integrated Post-Secondary Data System Graduation Rate Survey (IPEDS-GRS).

The NCAA acquires student-athlete graduation rate data from the IPEDS-GRS and builds their own methodologies to factor in its limitations.

Graduation Success Rate

Measures:

Begins with the Federal (GRS) Methodology group, then tracks groups the Federal methodology omits, including transfers, mid-year enrollees, and non-aid recruited athletes

In 2002, the NCAA introduced a methodology for Division I called the Graduation Success Rate (GSR), which accounts for student-athletes who transfer into or out of a particular college.

Source: https://www.ncaa.org/about/resources/research/why-gsr-better-methodology

Academic Performance Rate

Measures:

- Eligibility (grades)
- Graduation (credits)
- Retention (degree progress)

The APR is an annual scorecard of academic achievement calculated for all Division I sports teams.

Source: https://www.ncaa.org/about/resources/research/division-i-academic-progress-rate-apr

Data Analysis

Fig. 1: Average Rate by Sport (APR)

Findings:

- **Lowest APRs:** Men's Football, Basketball and Track
- **Highest APRs:** Women's Skiing, Lacrosse, Ice Hockey, Gymnastics, Golf, Swimming and Tennis
- Multiple Women's sports exceed the highest men's sport (Golf)

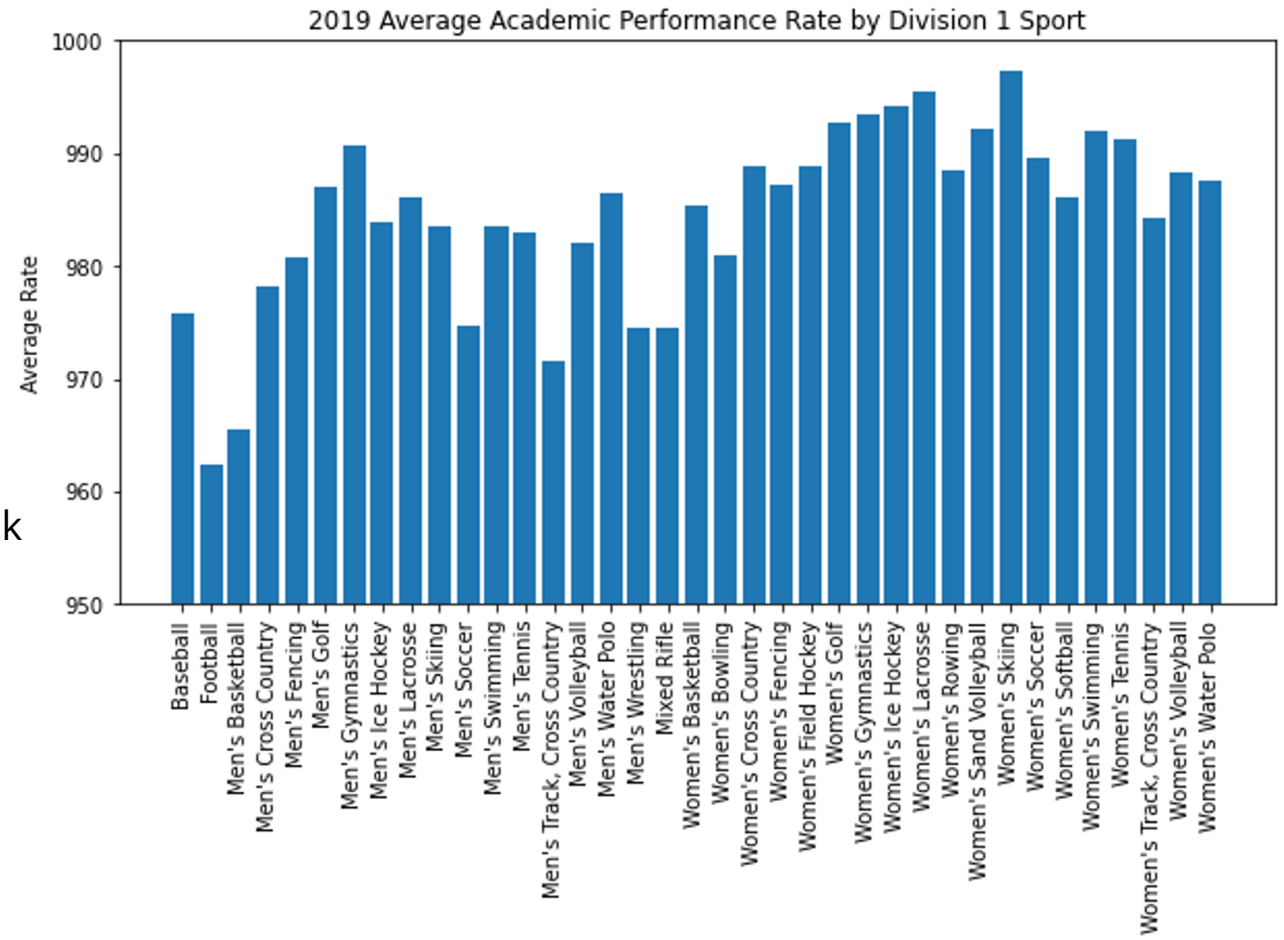


Fig. 2: Comparison by Gender (APR)

Finding:

Female student-athletes show higher academic performance rates on average compared with male student-athletes.

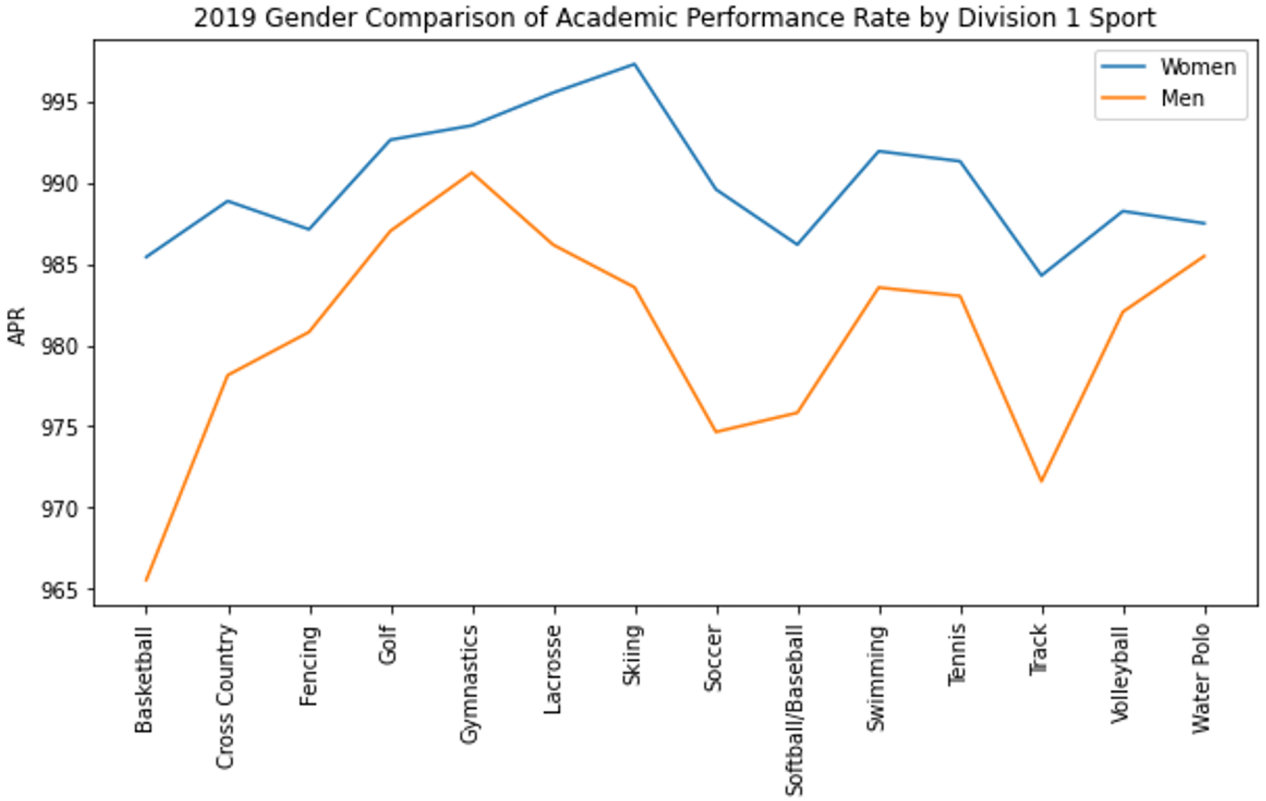


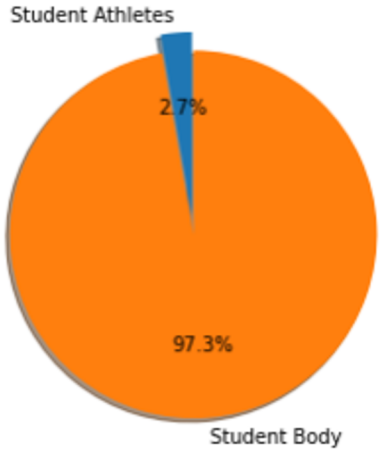
Fig. 3: Comparison by Population/Sample (Federal Graduation Rate Survey)

Finding:
Student-athletes show higher average graduation rates compared to the student body.

Counts:

Student Athletes: 88,736

Student Body: 3,153,916



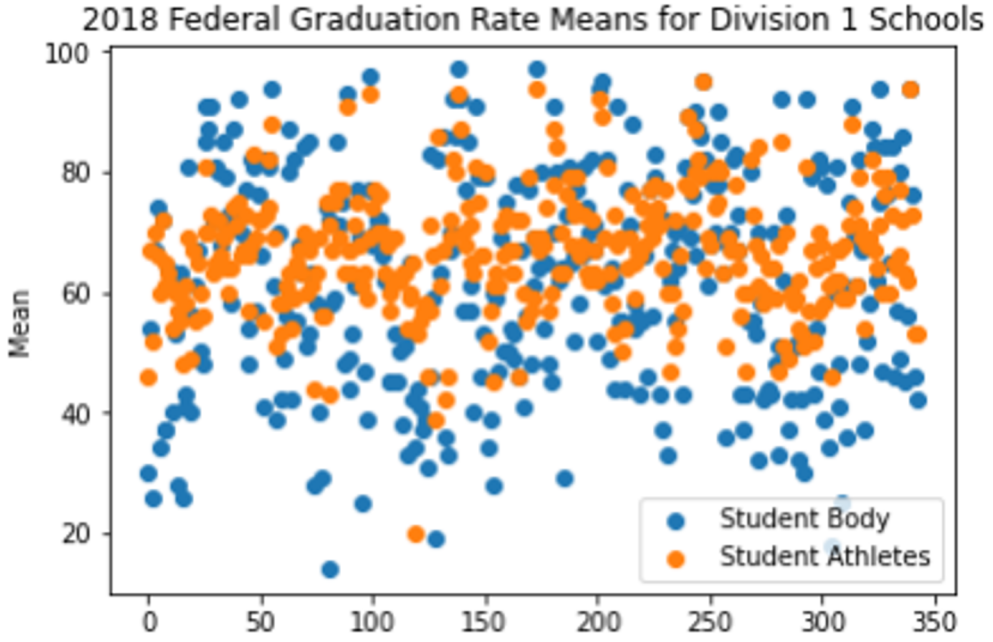
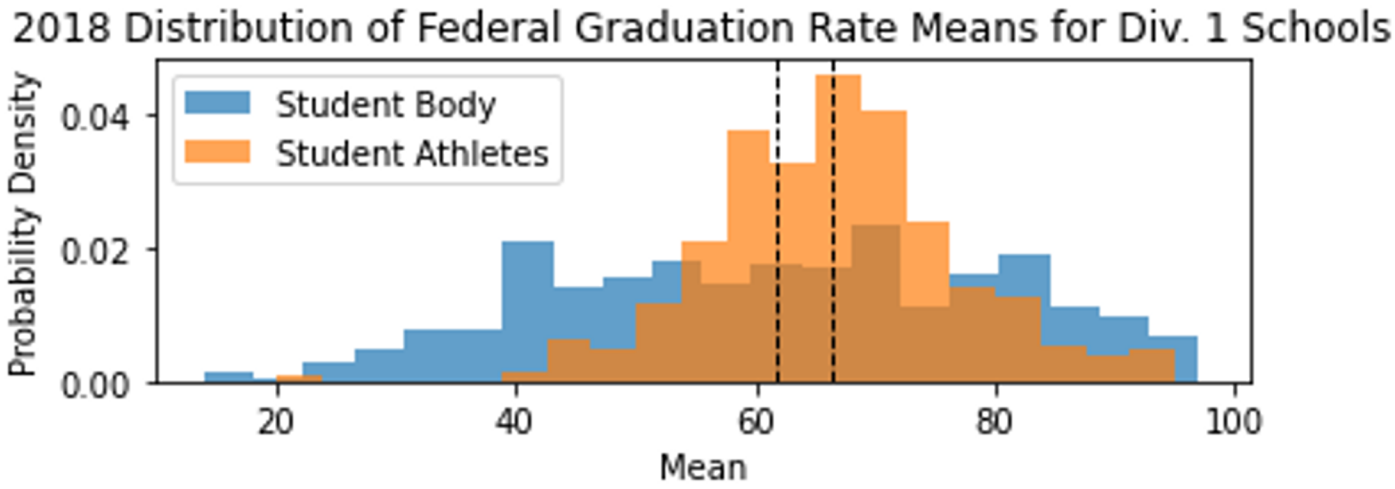


Fig. 4: Distribution of Means (Federal Graduation Rate Survey)

T-test Results:
Statistic = 8.1
P-value = 9.3e-15



Findings: Since p-value is less than 0.05, we can reject the null hypothesis
Conclusion: There is generally more graduation success among student-athletes compared to the general student body.

Discussion

Overall findings aligned with our expectations.

General conclusions are:

- Student-athletes are exposed to habits, characteristics and environments which contribute to academic success.
- Motivational factors embedded in the Div. I collegiate sports programs are helping student-athletes attain higher academic performance in comparison to their non-athletic peers.

Post Mortem

Difficulties Faced:

Initially faced difficulty with finding comparative data to the general student body

Saw a challenge in interpreting the academic measurements across the analysis (there are multiple within different groups)

Cannot use them to compare with different groups.

Additional Questions:

Are there any geographical inferences we can make within this dataset?

Are there certain regions more high performing academically among student athletes?

Questions