

## Project 2 W200

**Team Members:** Matt Whittaker, Charlie Boatwright, Matt Pribadi

**Github:** [https://github.com/UC-Berkeley-I-School/Project2\\_Pribadi\\_Boatwright\\_Whittaker](https://github.com/UC-Berkeley-I-School/Project2_Pribadi_Boatwright_Whittaker)

### Datasets:

1. **Primary:** The [College Athletics Financial Information Database](#) details the annual sources of revenue (such as ticket sales) and expenses (such as coaches' compensation) for hundreds of schools, based on information self-reported to the NCAA and federal government
2. **Supplemental:** For college football, we will use [this api](#) to create a data set of win/loss records, joining on school name from the primary dataset

**Overall question:** What is the relationship between financial expenditures on athletics and athletic success, defined by wins and losses?

### Variables types and exploratory analysis:

- **Expenses:** Some interesting expense types (i.e., columns) included in the athletic expenditures include athletic student aid, coaches compensation, recruiting, medical expenses, and facility expenses
- **Revenues:** Revenue types are segmented by institutional/government support, donor contributions, student fees, NCAA distributions, and ticket sales
- **School classification:** Schools are grouped by conferences, with certain conferences likely to have much higher expenditure for high priority sports - for example, the Southeastern Conference (SEC) will have a higher expenditure than the Sunbelt Conference, but will also have higher revenues
- **Additional school reference data**
  - Region
  - Undergraduate size
  - Number of sports teams
  - Carnegie Classification (e.g., research institution, master's colleges, baccalaureate colleges)

### Final Report Outline and Organization (8-10 pages):

1. Executive summary with key findings and conclusions (1-2 pages)
2. Overall questions (1 page)
3. Data cleaning and sanity checks (1 page)
4. Text, data stories, and figures to answer key questions (5 - 6 pages)

**Figures / Tables:** The Southeastern Conference (SEC) has won 11 of the last 15 NCAA national football championships. Schools in the SEC spend more on athletics than any other conference, with a median of almost \$135mm per school (figure 1). While the SEC spends more, they also receive more funding from donors (figure 2). Please refer to the next page for additional detail.

**Figure 1: Median athletic spending by conference in 2018**

Conference	Median Athletic Expenses
Southeastern Conference (SEC)	\$134,965,406 (see figure 2)
Big Ten Conference	\$131,467,820
Pacific-12 Conference (Pac-12)	\$106,944,047
Atlantic Coast Conference (ACC)	\$104,678,737
Big 12 Conference	\$90,378,001
American Athletic Conference (AAC)	\$57,106,913
Mountain West Conference	\$48,079,095
Conference USA	\$35,149,886
Mid-American Conference (MAC)	\$33,158,489
Sun Belt Conference	\$32,428,520

**Figure 2: \$ received from donors, by conference**

