

## **Mini Project Idea 2**

### **Research Question:**

Can scholarship receipt and net price predict post-graduation outcomes, and how does this relationship vary across institutions with different tuition levels?

### **Motivation:**

As the cost of higher education continues to rise, students increasingly rely on scholarships and financial aid to access college. However, scholarship receipt is not random; it often reflects both financial need and institutional assessments of student preparedness. At the same time, net price after aid provides a more accurate measure of the true cost students face than sticker tuition alone. This project is motivated by the need to understand whether financial aid and affordability are associated with better post-graduation outcomes, and whether these relationships differ between high-tuition and low-tuition institutions. Examining these patterns can shed light on how effectively financial aid supports student success and educational return on investment.

### **Proposed Data Sources:**

This project will use data from the Integrated Postsecondary Education Data System (IPEDS), which provides institution-level information on tuition and fees, graduation rates, retention rates, undergraduate outcome measures, and institutional characteristics such as public or private status, region, and degree level. To better capture affordability and student need, the analysis will incorporate data from TuitionTracker, including net price after financial aid and measures of grant or scholarship receipt. Net price data for full-time, first-time undergraduates will be used to move beyond published tuition and approximate the actual financial burden faced by students. These datasets will be merged at the institutional level.

### **Methodology**

The analysis uses descriptive statistics, regression models, and unsupervised clustering. Institutions are first divided into high- and low-tuition groups based on tuition percentiles, and differences in graduation rates, retention rates, and outcome measures are compared. Regression models then examine whether scholarship receipt and net price predict post-graduation outcomes while controlling for institution type and region. Net price is measured for full-time, first-time undergraduates from low-income families to capture the actual cost of attendance beyond sticker tuition. Finally, clustering is used to identify patterns in affordability, financial aid, and outcomes across institutions.

## **Challenges**

A key limitation is that student ability and job outcomes are not directly observed at the individual level; scholarship receipt and net price serve only as imperfect proxies. Post-graduation outcomes available in IPEDS are limited and may not fully capture long-term labor market success. Additionally, selection bias may influence results, as students self-select into institutions with different tuition and aid structures. As a result, findings will be interpreted as correlational rather than causal.