

Graduate Earnings Across Institutions and Programs

Abdul Bholat, Albert Carreno, Andrew Darwin, Anish Deshpande, Caleb Williams, Naren Prakash



Outline

01

Problem Statement & EDA

02

UCLA vs. Other Institutions

03

Cohort Size & Earnings Trends

04

Earnings Benchmarks

05

Institution Level Analysis

06

Questions

Abstract

This project examines how graduate earnings vary across academic programs and institutions, with a particular focus on UCLA as a benchmark for comparison. Specifically, we seek to answer two questions: (1) How does the number of graduates from UCLA's degree programs compare to peer institutions? and (2) How does program scale affect post-graduation earnings? To address these questions, we compiled and merged datasets from the College Scorecard and IPEDS, filtering for recent and complete institutional and program-level data.

Our statistical approach involved imputation of missing values using linear regression for small gaps, and K-nearest neighbors for categorical variables. We then performed clustering on institutions based on control type, graduation rate, size, and selectivity to identify UCLA's peer groups. Regression analysis was used to evaluate the relationship between program scale (i.e., number of graduates) and median earnings, while controlling for institution type and field of study. Preliminary results suggest that scale is positively but weakly associated with earnings in fields such as Business and Computer Science ($p < 0.05$), but not in Engineering or Social Sciences ($p > 0.1$), indicating statistical significance only in some contexts.

A key challenge was addressing missing or inconsistent reporting in program-level earnings data, especially for smaller institutions or niche programs. Our next steps include incorporating cost-of-living adjustments to account for geographic wage differences and refining our regression models with interaction terms to capture nonlinear scale effects. These improvements will strengthen the causal interpretation of program scale on earnings and inform institutional decision-making.

Problem Statement

How do UCLA grads compare to peers from other institutions across degree types, fields of study, and institutional characteristics (program size, school type, region/state, tuition, etc.)?

EDA

- College Scorecard data contains aggregated metrics for variables, mostly handling missing values and ensuring numeric formatting
- Look at just UCLA first, then compare with other programs
- Evaluate how UCLA compares to:
 - Ivies / top ranked colleges
 - All CA schools
 - Similarly ranked colleges
 - National Averages

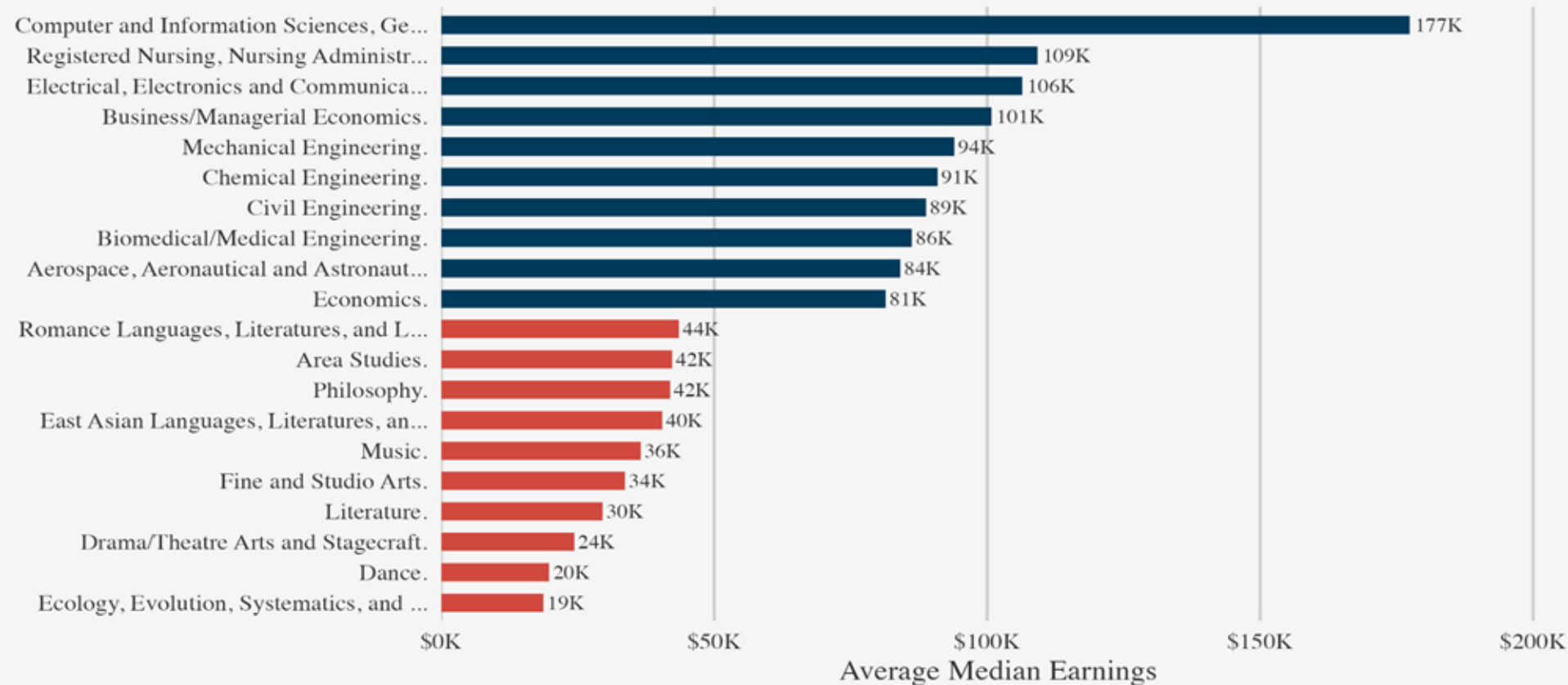
Variables

- Institution Names and ID's
 - CIP codes and descriptions
 - Median Earnings (1yr and 5yr)
 - Program-Specific
 - Debt After Graduation
 - Cohort Size per program
- Institution Data includes longer term outcomes
 - Program-specific data focuses on short term

Top and Bottom 10 UCLA Majors by Avg. Median Earnings (1YR & 5YR)

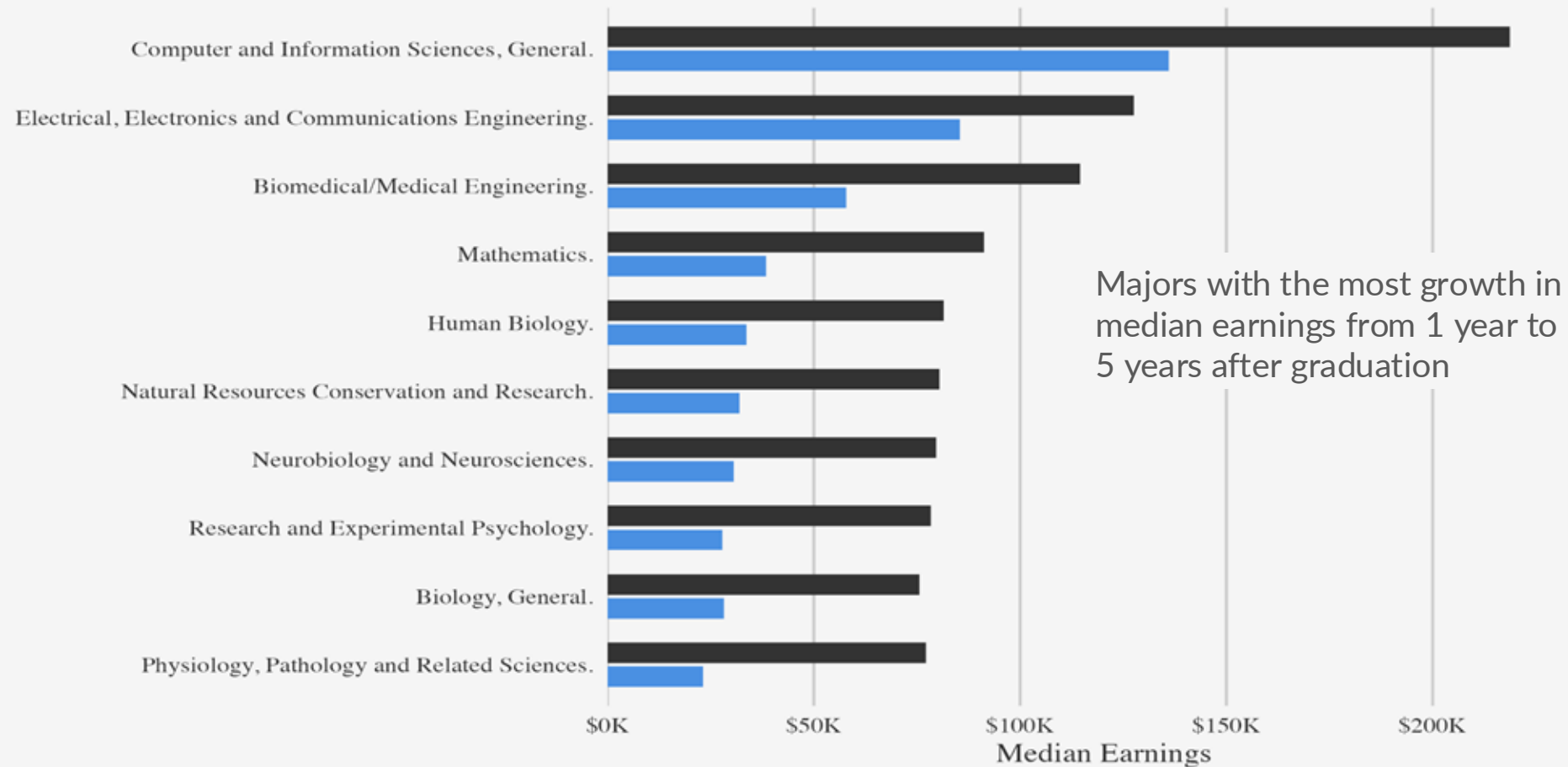
Average of 1-Year and 5-Year Median Earnings Post-Graduation

Category Bottom 10 Top 10



Median Earnings: 1-Year vs 5-Year After Graduation

Time 1 Year 5 Year



Debt vs. Early Earnings by Program

Each point represents a UCLA undergraduate program

Median Earnings (1 Year After Graduation)

Plot of Median Debt at graduation vs.
median earnings 1 year after
graduation

- Top right point is Computer science, bottom right is Theater arts
- See linked dashboard for full labeling

\$100K

\$50K

\$10K

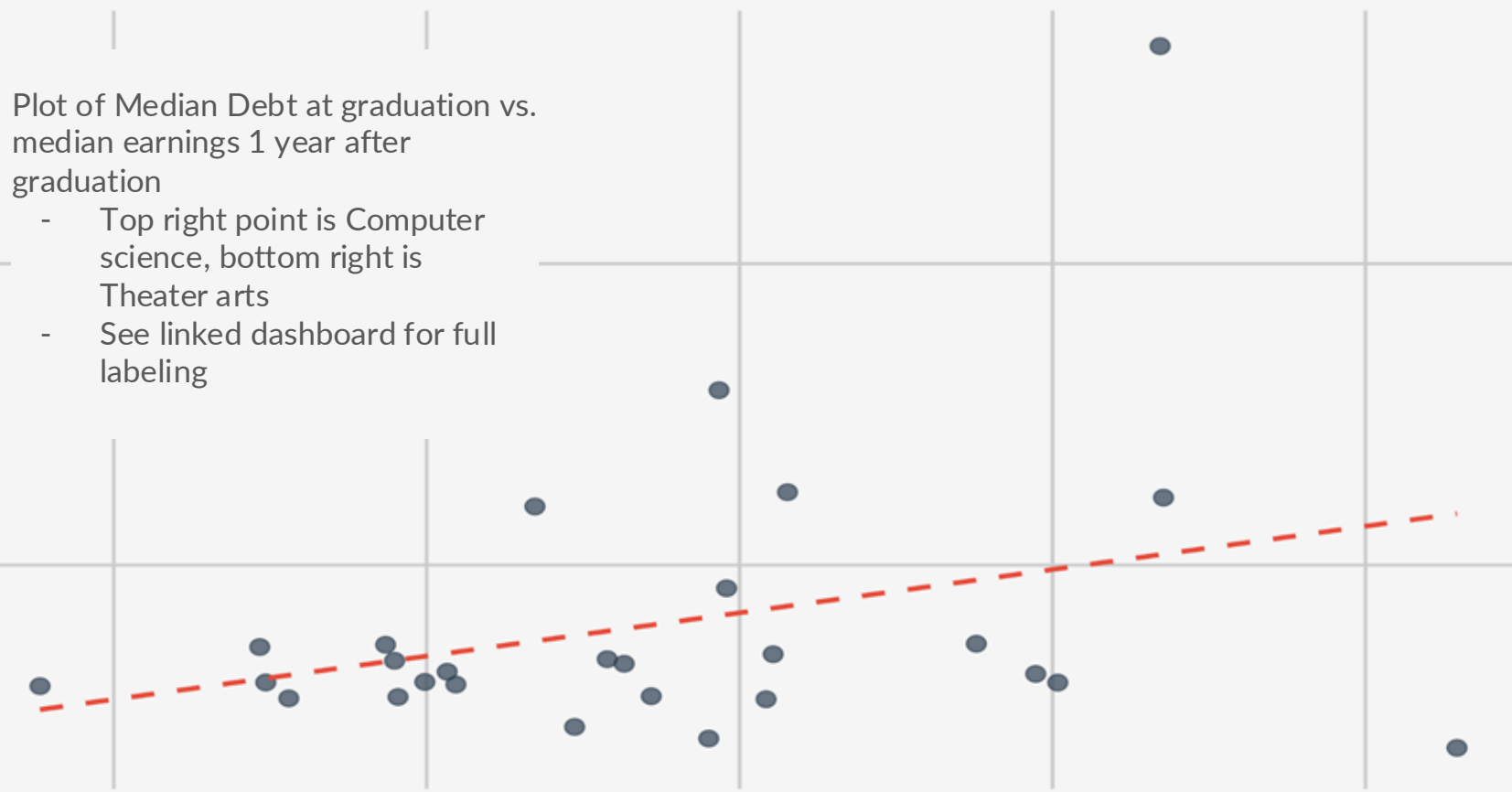
\$20K

\$30K

\$40K

\$50K

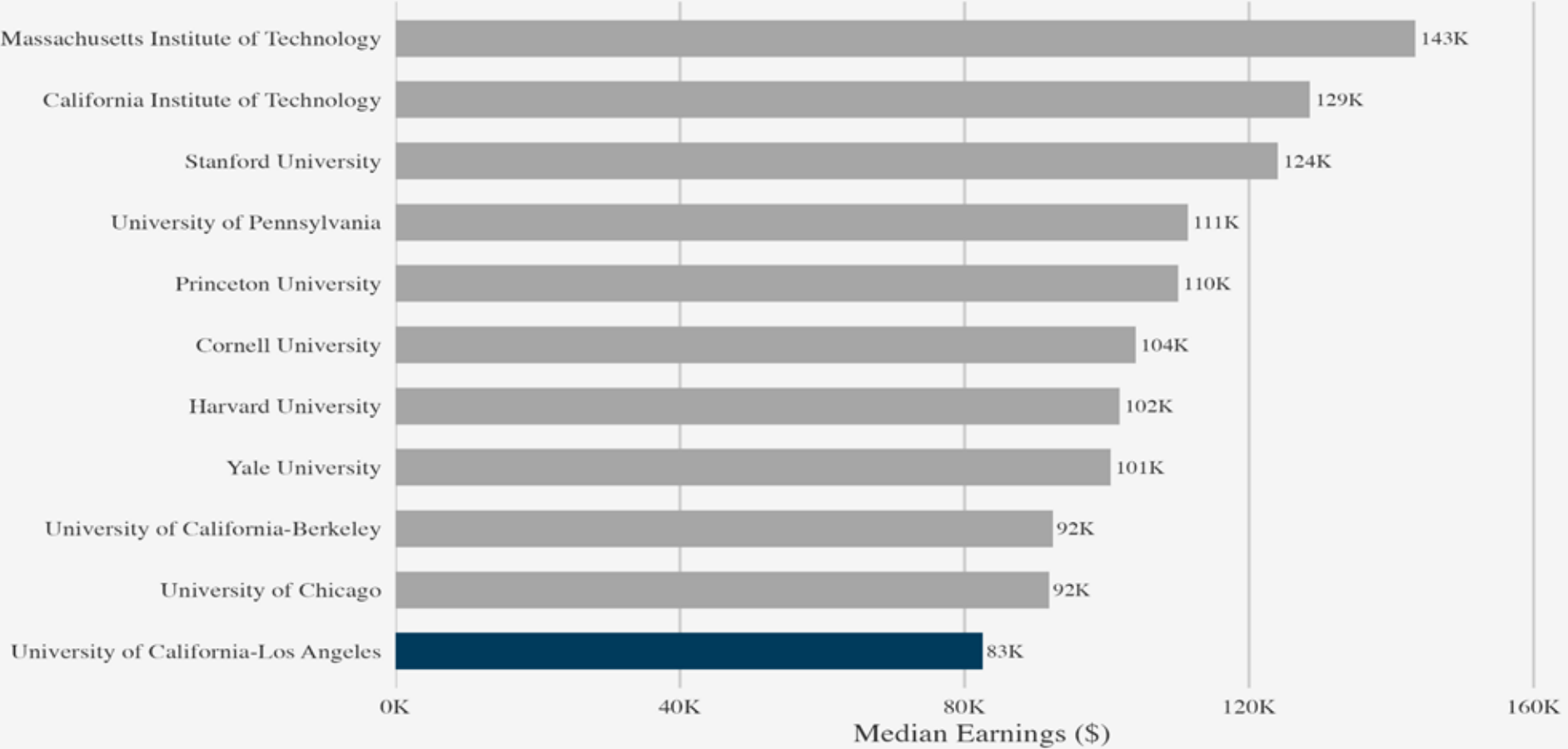
Median Debt at Graduation



Comparing UCLA to Other Institutions

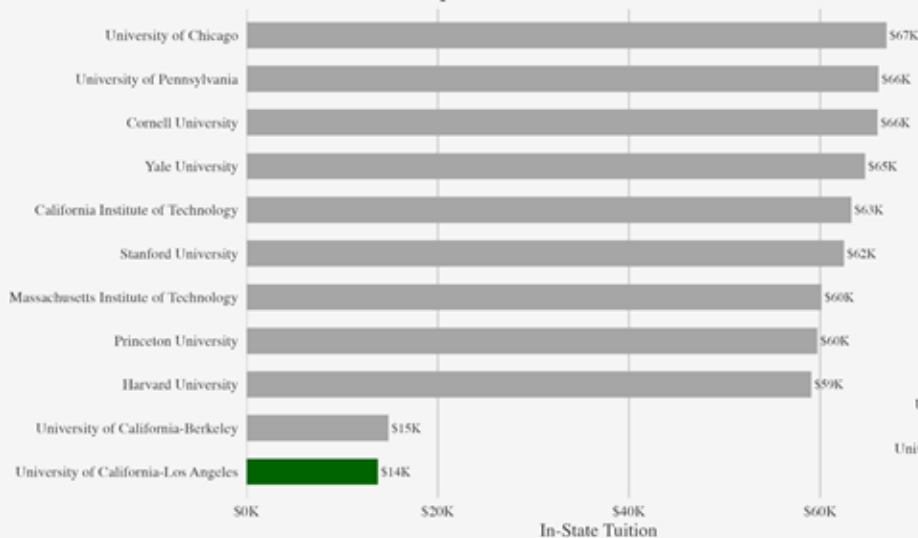
UCLA vs. Top 10 Nationwide Schools

Median Earnings 10 Years After Entry

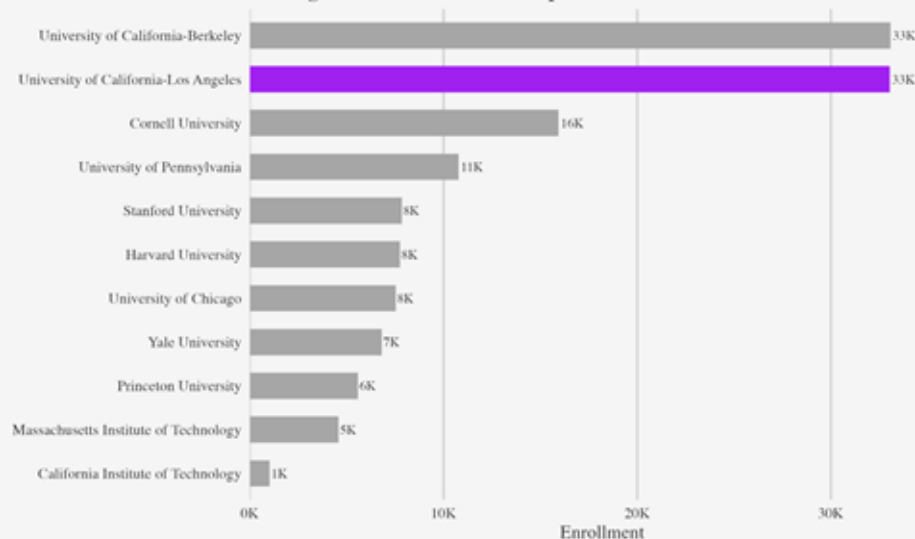


Other Comparisons to Top 10 Schools

In-State Tuition Comparison



Undergraduate Enrollment Comparison



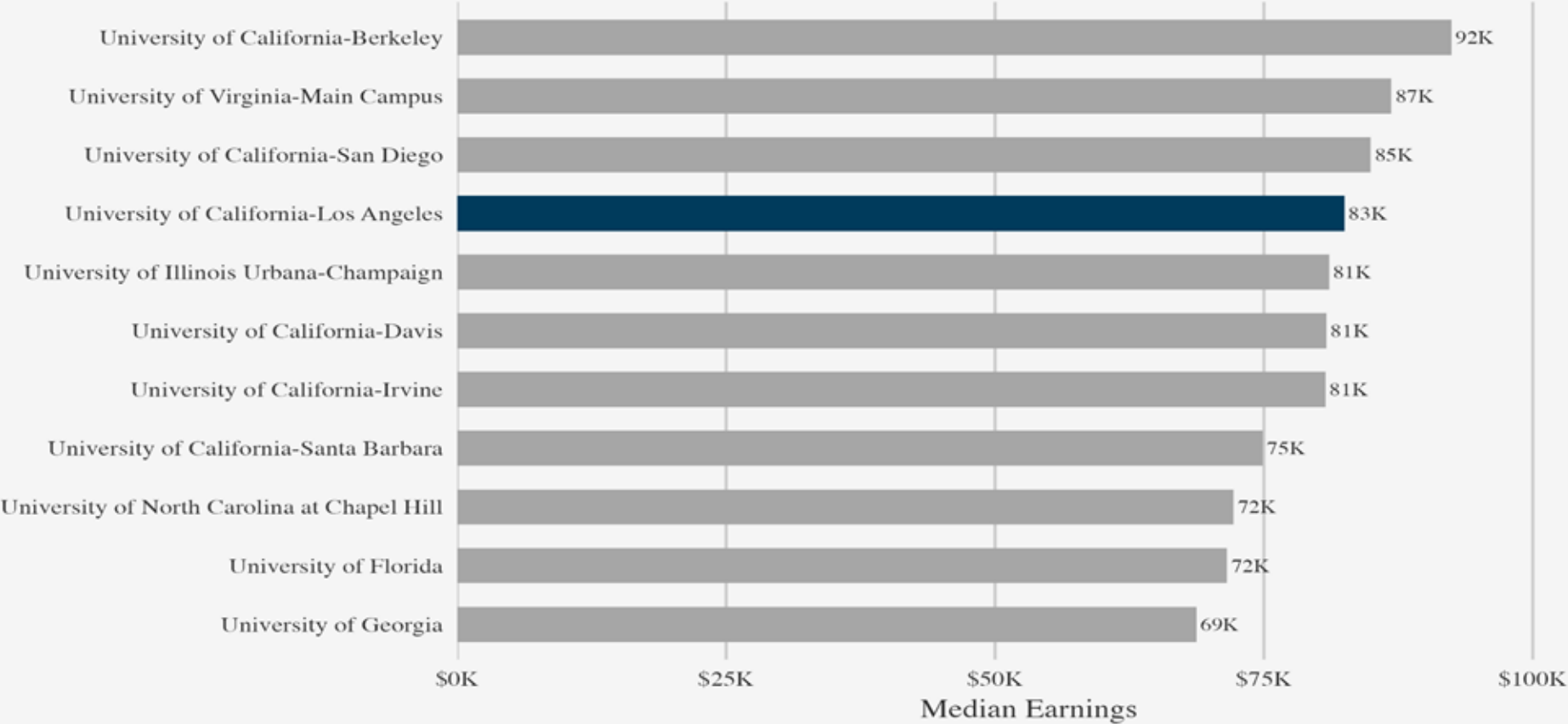
https://www.usnews.com/best-colleges/rankings/national-universities/top-public?_sort=rank&_sortDirection=asc

<https://www.topuniversities.com/world-university-rankings>

UCLA vs. Other Top Public Schools

UCLA vs. Top Public Universities

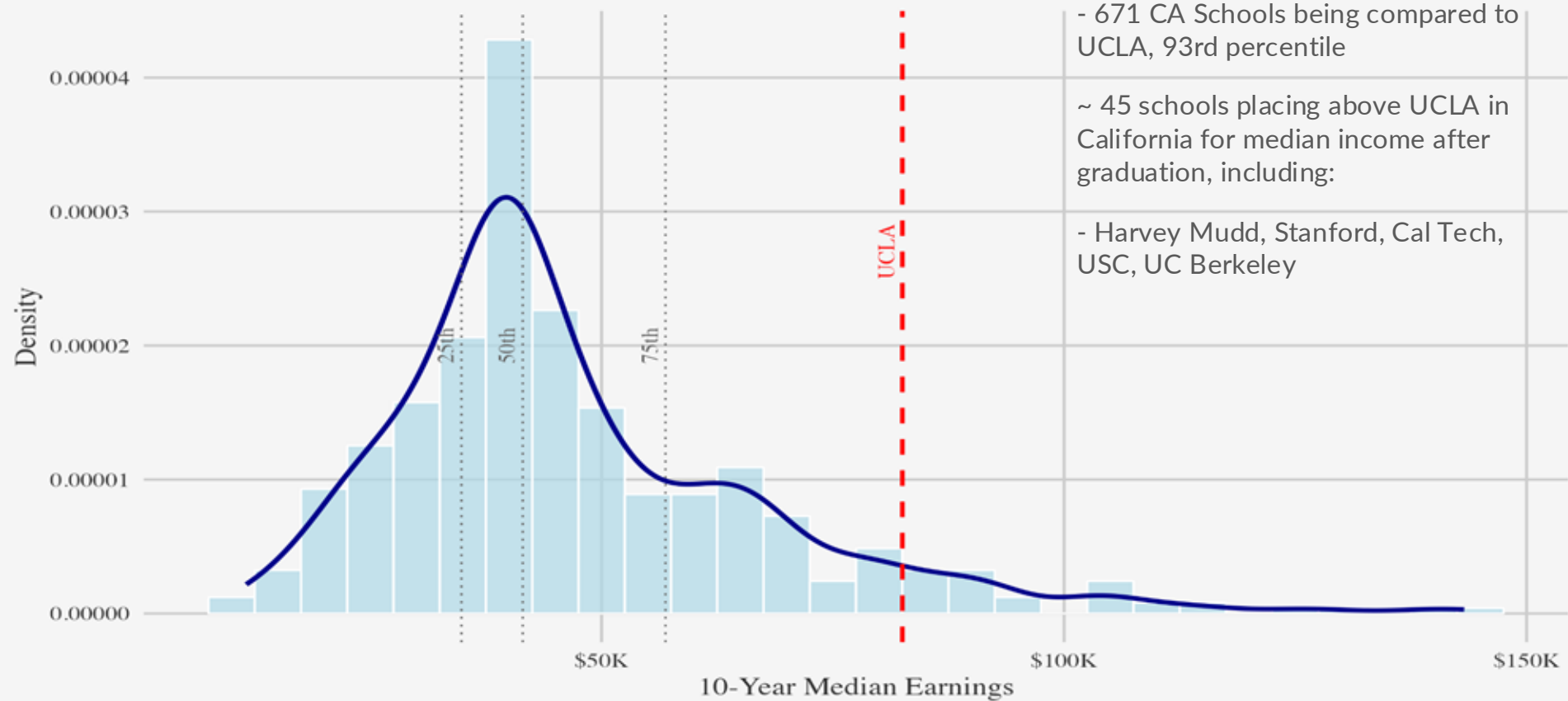
10-Year Median Earnings After Entry



UCLA vs. California Schools

UCLA Earnings Compared to CA Colleges (10-Year Median)

Red line = UCLA | Dotted lines = 25th, 50th, 75th percentiles

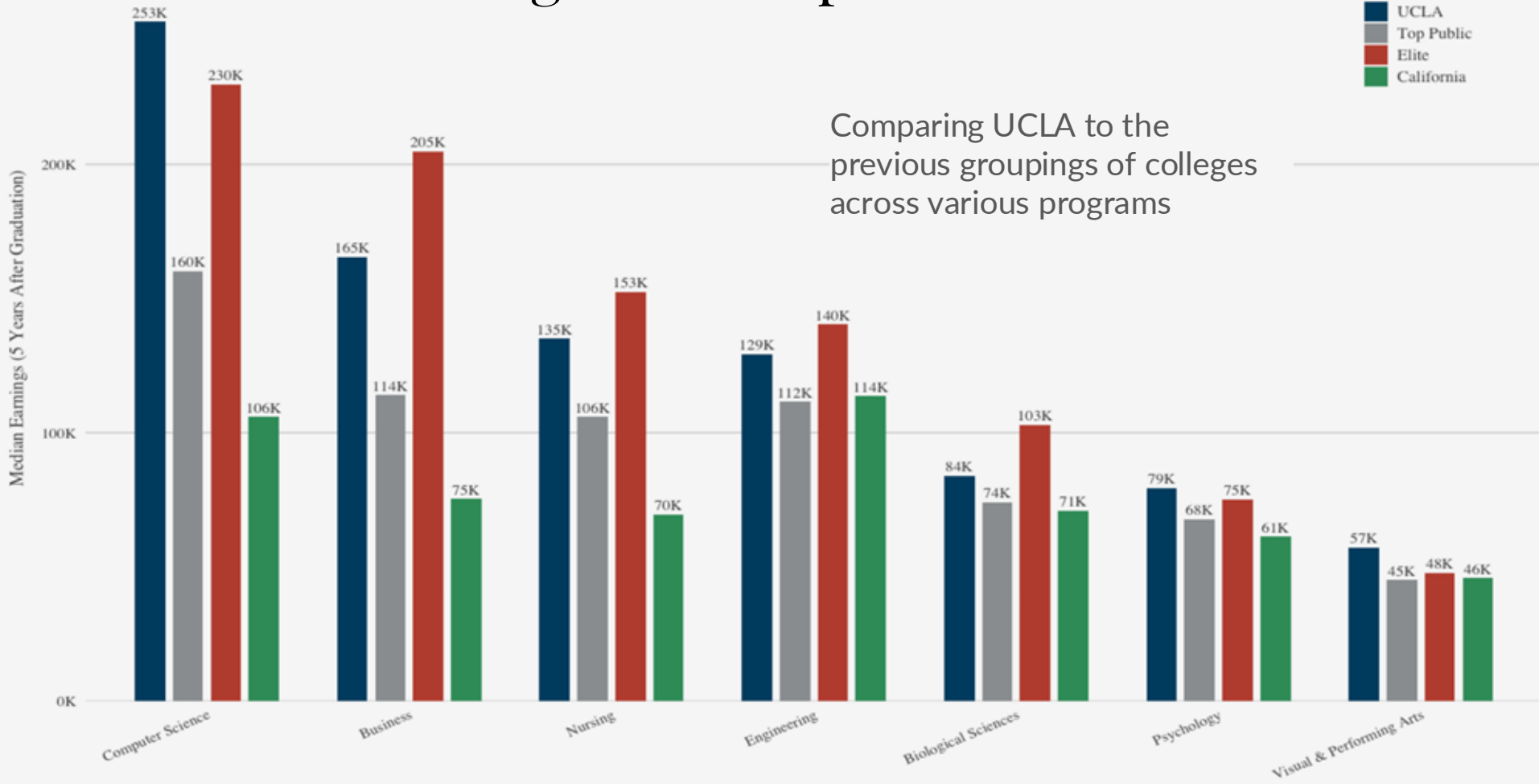


Program Comparisons

Institution Group

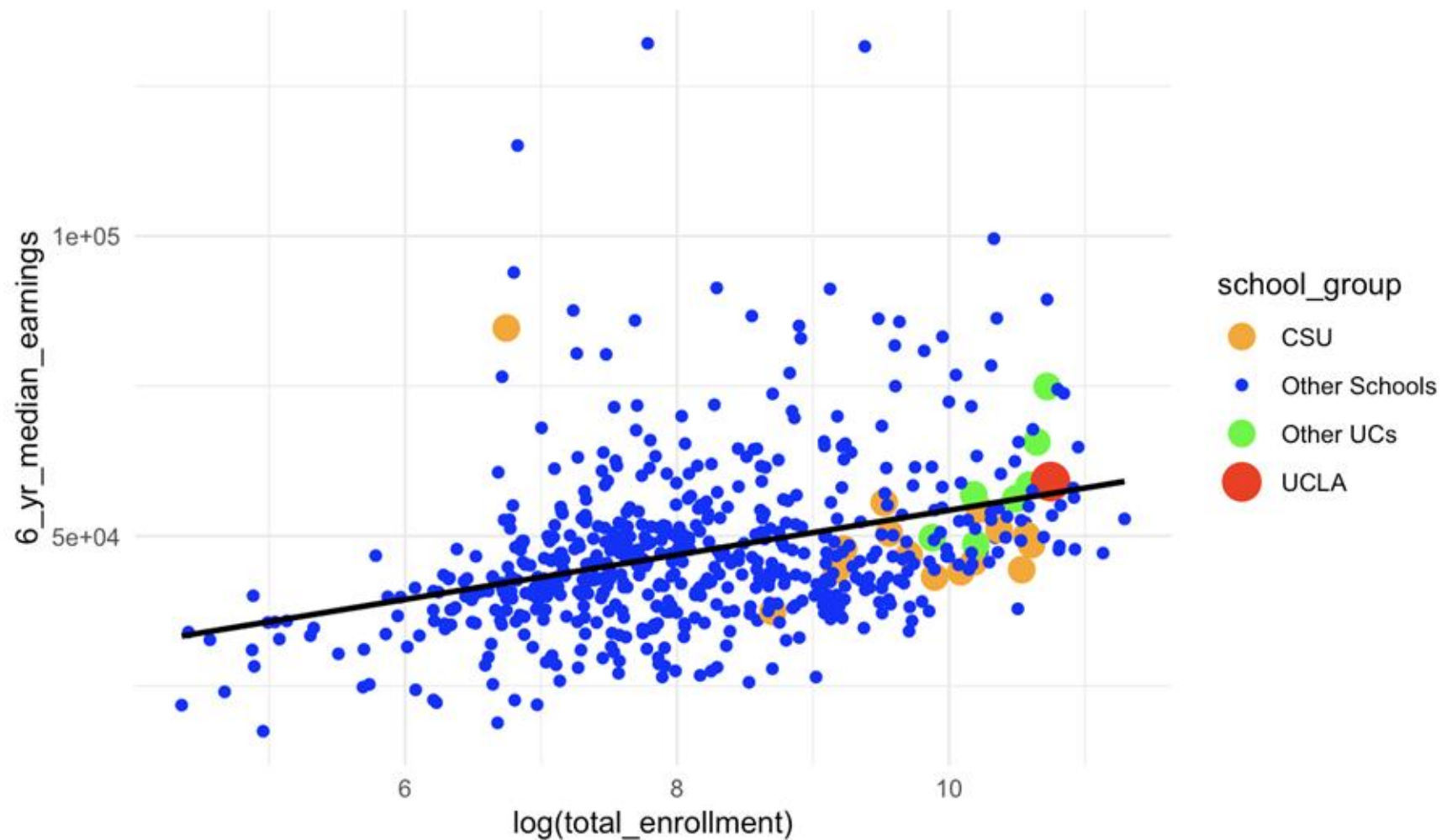


Comparing UCLA to the previous groupings of colleges across various programs

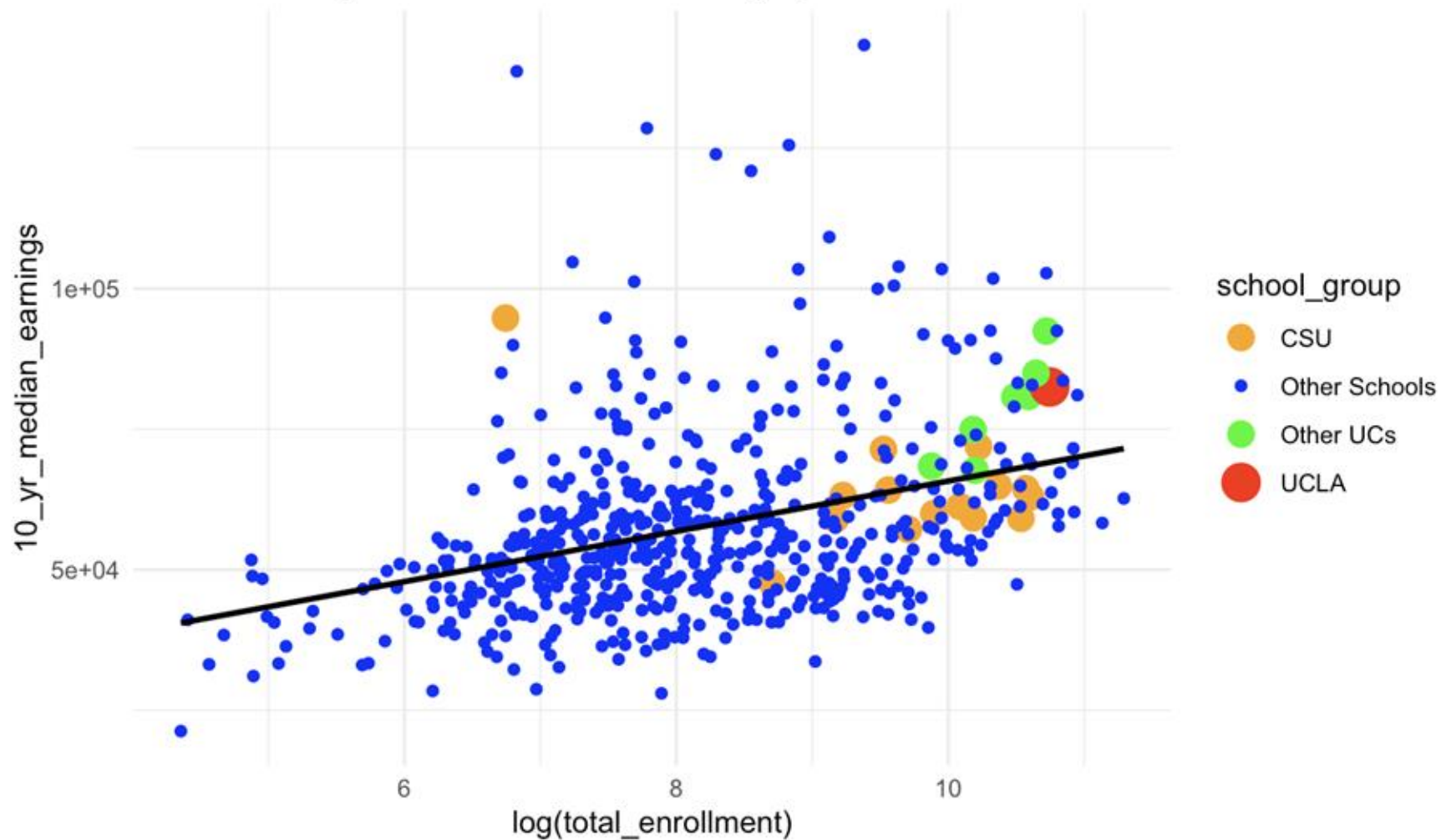


Graduate Cohort Size and Earnings Trends

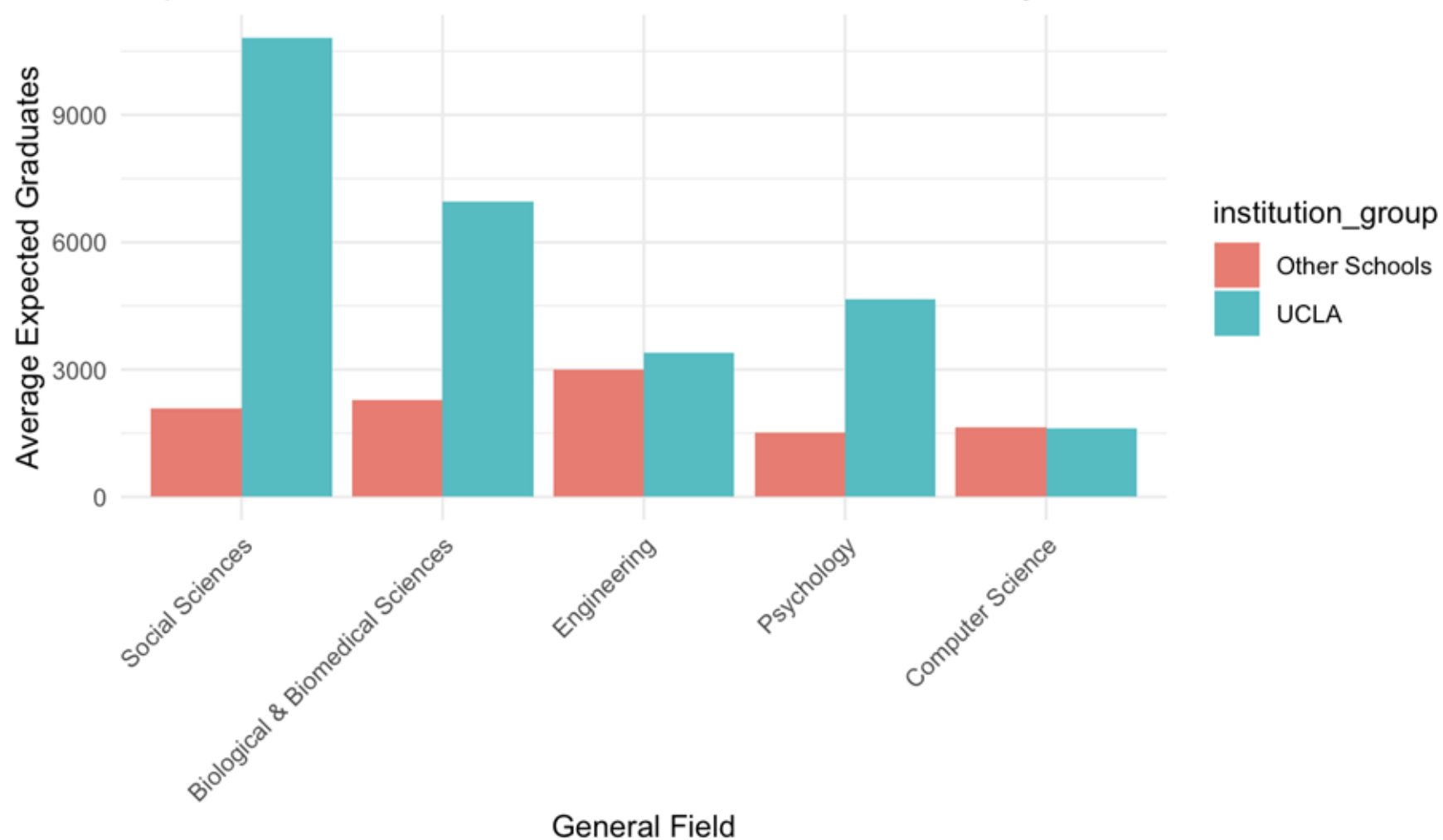
Enrollment by 6-Year Median Earnings per Institution



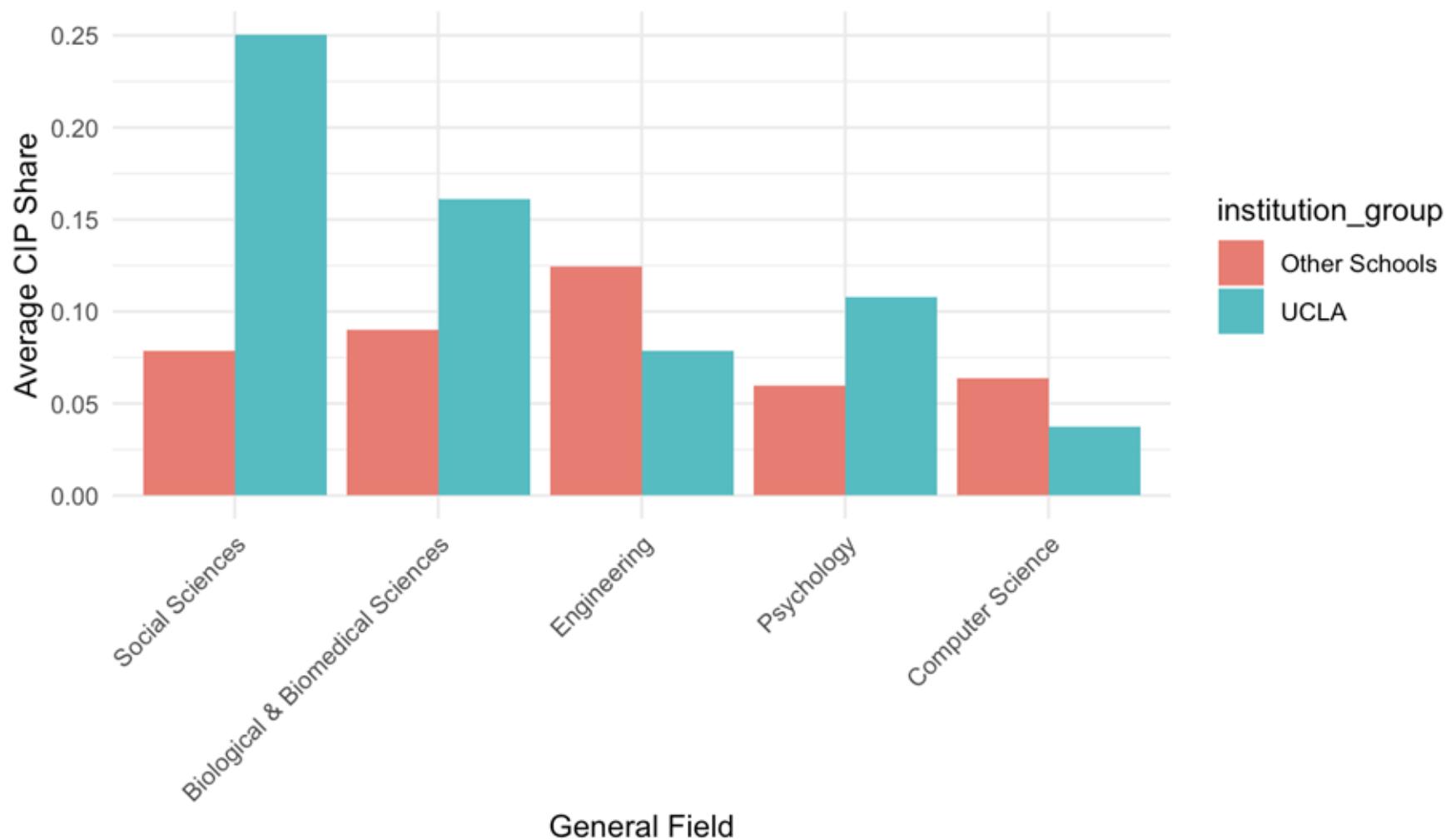
Enrollment by 10-Year Median Earnings per Institution

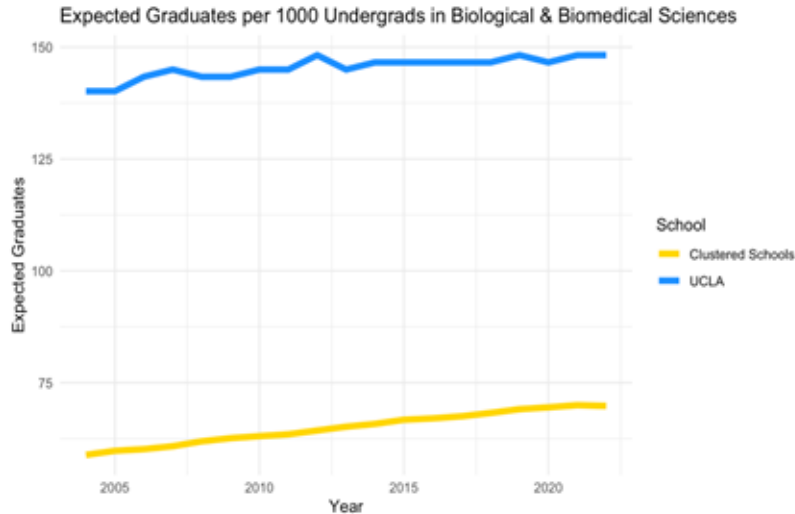
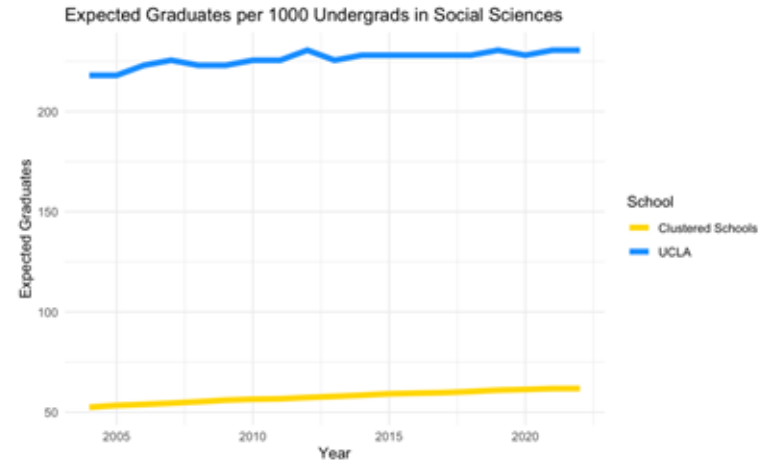
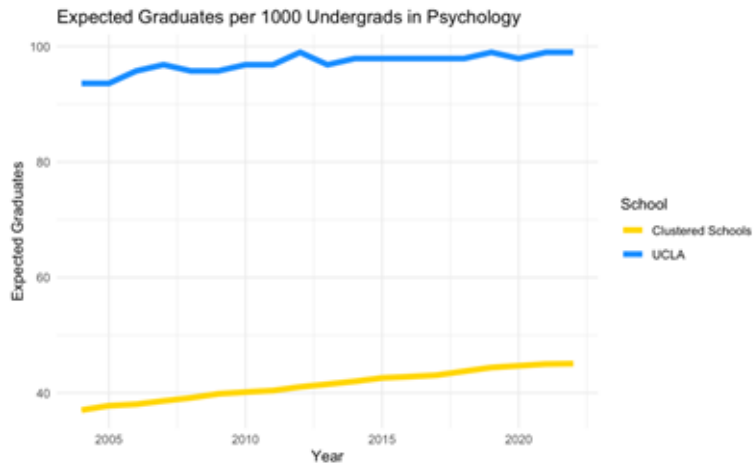


Expected Graduates: UCLA vs Other Similar Schools by CIP Code

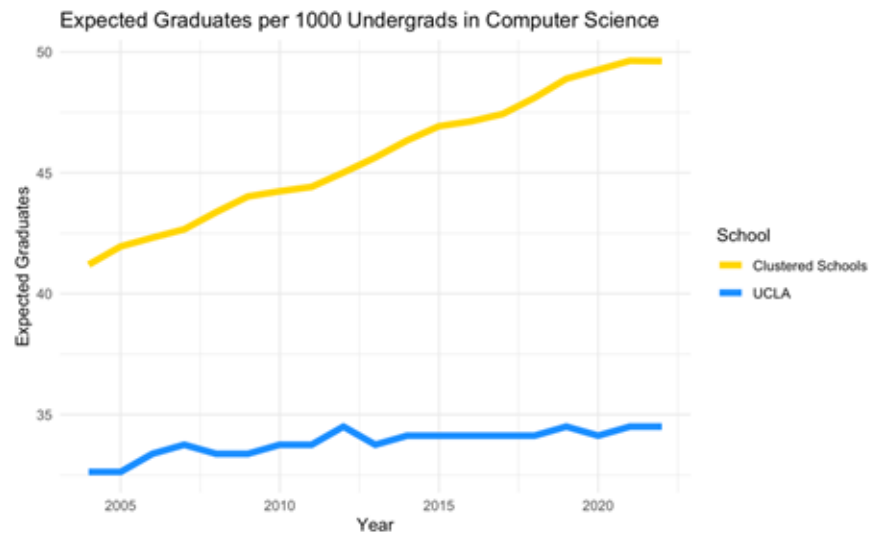
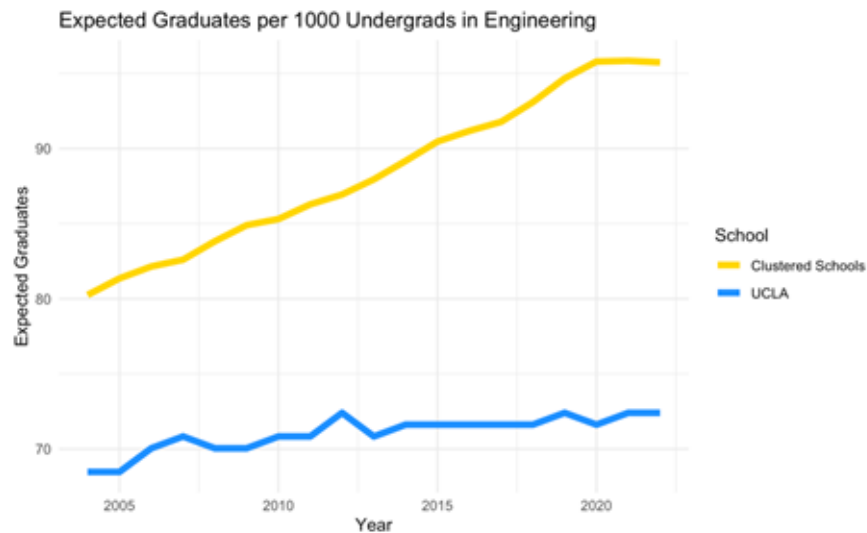


CIP Share: UCLA vs Other Similar Schools by General Field





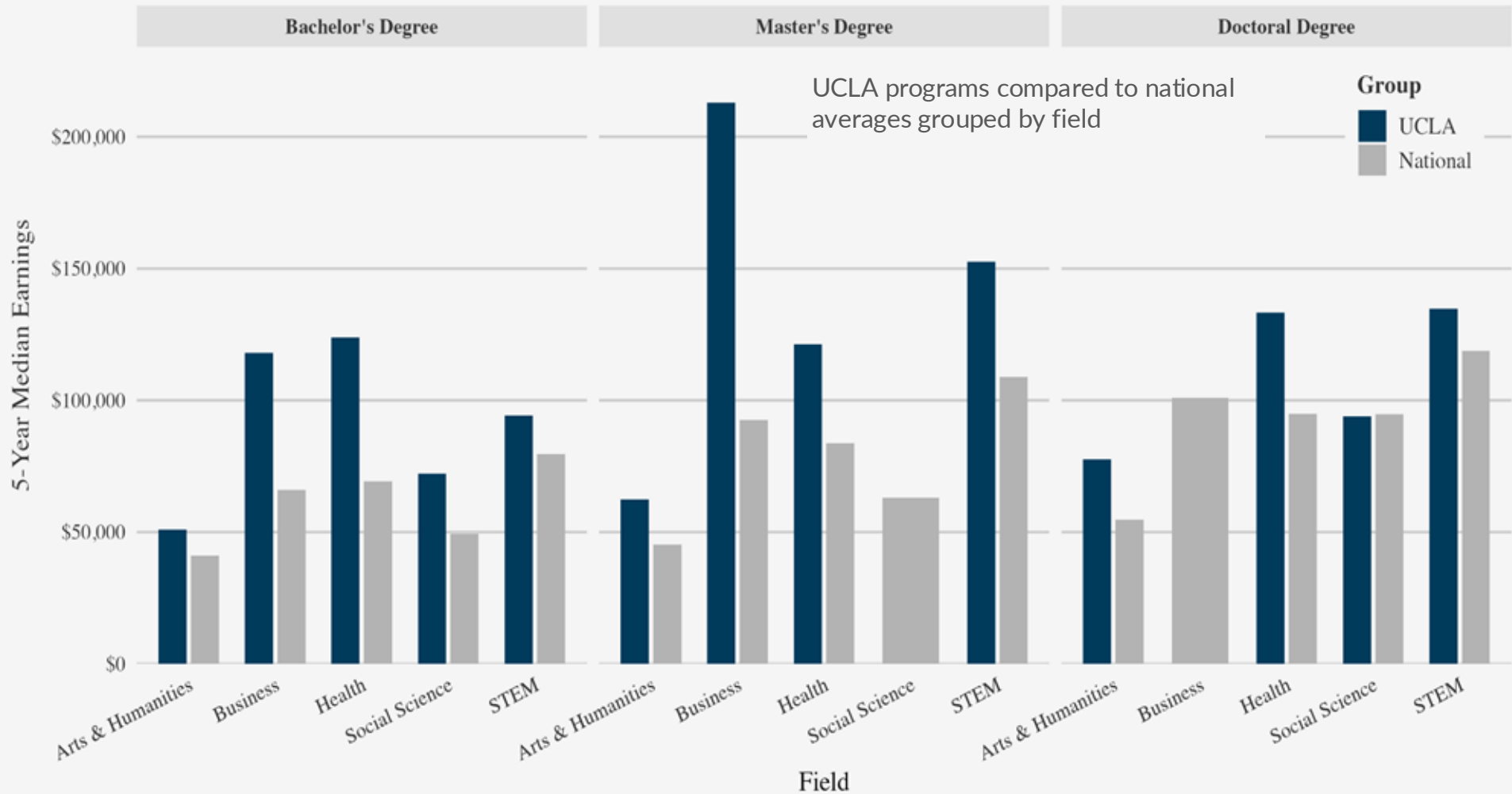
- These plots help answer “for every 100 undergraduate students at a school, how many of them do we expect to graduate in a specific field?”
- We’re also able to assess how this rate of expected graduates in these fields have changed over the last 20 years



- Widening gap between UCLA and clustered schools over time in these major STEM fields
- What might we attribute this to?
 - Selectivity(~5% for Samuelli)
 - Direct admission
 - Student demographics
 - Physical capacity(UCLA has added less STEM space than other schools)

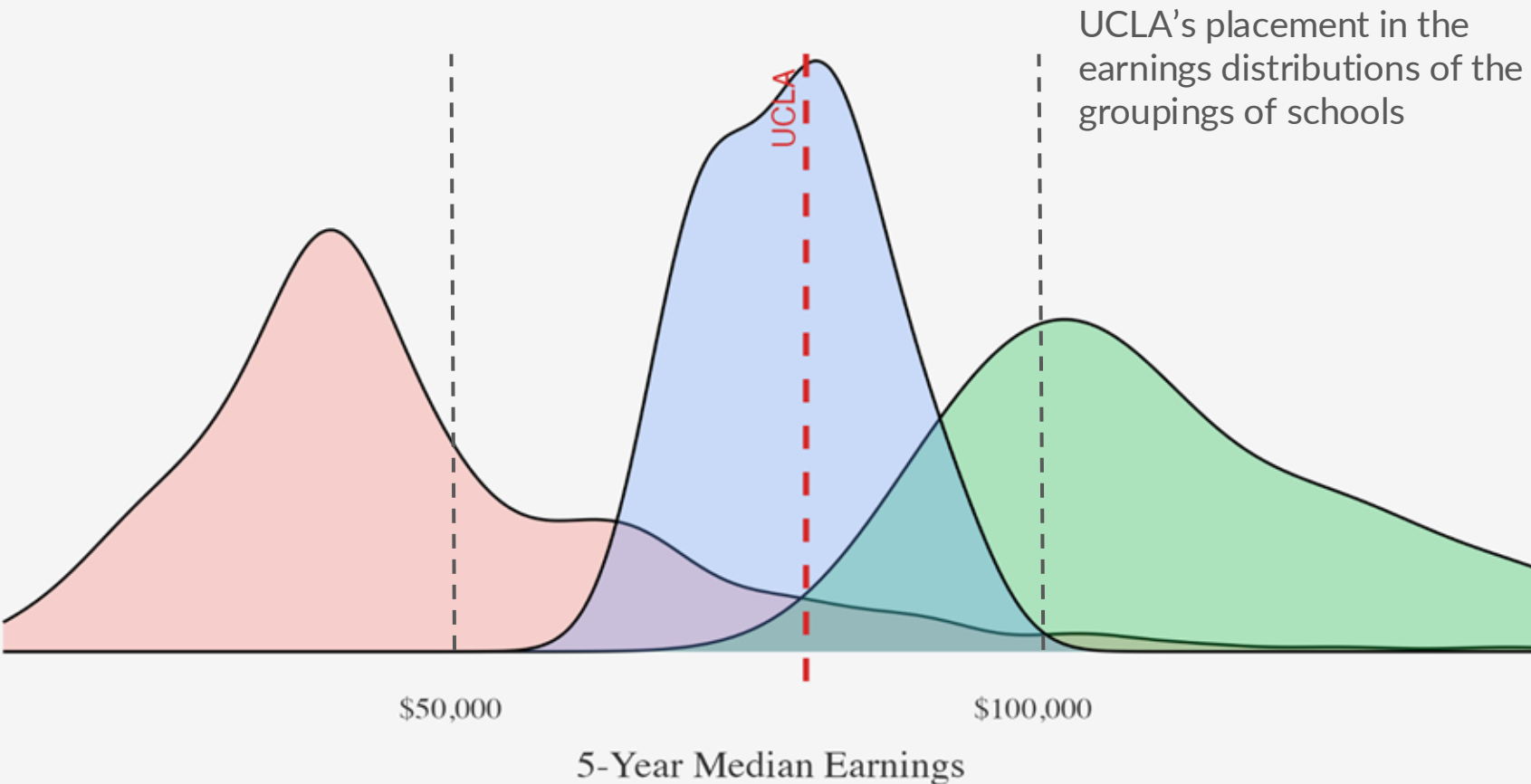
Earnings Benchmarks and Distributions

UCLA vs National Median Earnings by Field and Degree Type



5-Year Median Earnings Distribution by Institution Group

Institution Group California Elite Top Public



UCLA 5-Year Earnings Percentile Rank by Program Category			
Program Category	UCLA Earnings (5 Yr)	National Median	Percentile Rank
Computer & Information Sciences	\$253,100	\$74,585	99.6%
Business & Management	\$165,434	\$65,279	99.2%
Education & Public Service	\$106,889	\$54,386	97.1%
Engineering	\$129,308	\$89,865	94.8%
Humanities & Languages	\$69,776	\$46,146	93.4%
Arts, Design & Media	\$57,193	\$40,407	90.4%
Biological & Health Sciences	\$107,535	\$59,032	89.7%
Math & Physical Sciences	\$101,418	\$69,032	88.7%
Social Sciences & Psychology	\$76,132	\$55,161	86.1%

Implications

- UCLA consistently above 85th percentile
 - Reaches 95-99th percentiles in competitive fields
- Contends with even top private institutions
 - UCLA offers a comparable education and outcome to ivies for significantly lower price and to more students
- Continue to monitor and report outcomes to ensure UCLA continues this pattern

Earnings and Graduate Outcomes Product: Dashboard

- <https://ucla-earnings-dashboard.onrender.com/>

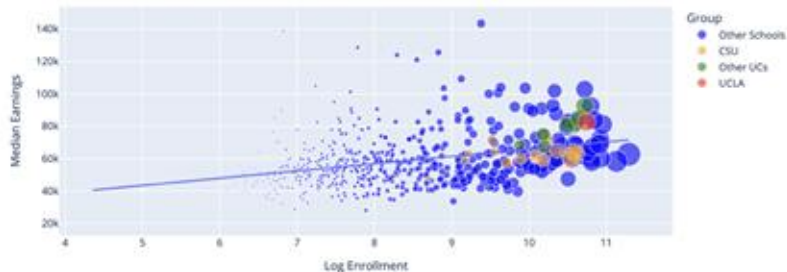
UCLA Graduate Earnings Dashboard

Earnings vs. Enrollment

Earnings Horizon:

☐ 6-Year ☒ 10-Year

Log Enrollment vs 10 Yr Median Earnings



Enrollment vs. Other Schools

Select Another Institution:

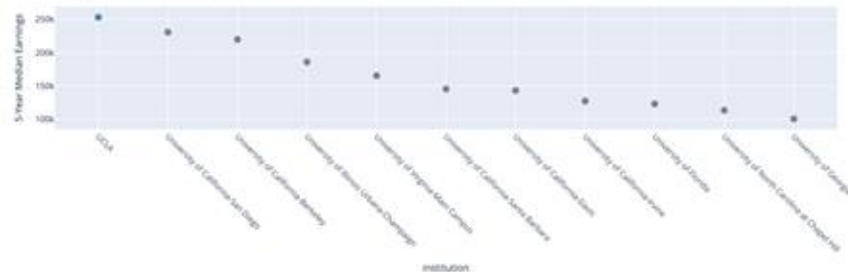
University of California-Berkeley

Earnings by Field of Study: UCLA vs Other Institutions

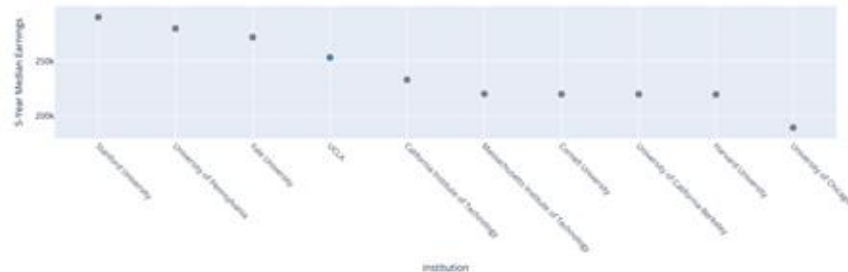
Select a Program:

Computer Science

Computer Science: UCLA vs Top Public Schools

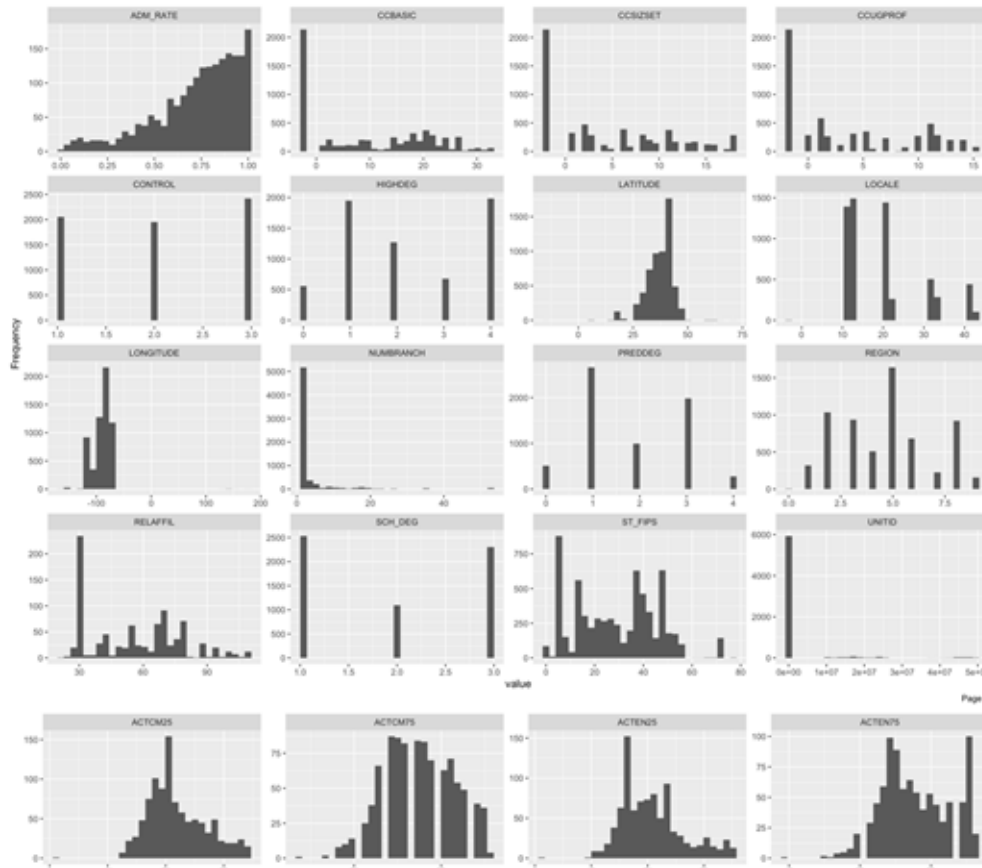


Computer Science: UCLA vs Elite Universities



Institution Level Analysis

Data exploration: Distributions



Defining “Peer Institutions”

Methodology:

1. Variable selection for institutional characteristics related to earnings
 - a. Mixture of missingness analysis and looking at papers
2. Non-parametric imputation
 - a. Using MICE and RF
3. HDBSCAN clustering
 - a. Finding the “true clusters” in the data
4. Variable importance
 - a. What created these clusters?
5. Defining a metric
 - a. How do we quantify student income?
6. Institutional factor analysis and comparisons
 - a. What causes differences in financial success and how does UCLA stack up?

Variable Selection

Feature Name	Description
id	Institution ID
ownership	Institution Type (Public/Private/For-profit)
carnegie_basic	Carnegie Classification (Basic)
carnegie_undergrad	Carnegie Classification (Undergraduate Focus)
under_investigation	Under Federal Investigation
degrees_awarded.highest	Highest Degree Awarded
admission_rate.overall	Overall Admission Rate
demographics.race_ethnicity.white	% White Students
demographics.race_ethnicity.black	% Black Students
demographics.race_ethnicity.hispanic	% Hispanic Students
demographics.race_ethnicity.asian	% Asian Students
completion_rate_4yr_150nt	4-Year Grad Rate (150% of normal time)
retention_rate.four_year.full_time	Full-time First-Year Retention Rate
usnews.median_rank	Median U.S. News Ranking (Lower is Better)
living.wage	Living Wage in Institution's County
minimum.wage	Minimum Wage in Institution's State
annual.income_pretax	Estimated Annual Income (Pre-Tax)

Note: Wage data sourced from the MIT Living Wage Calculator (livingwage.mit.edu).

Imputation and Clustering

- Using MICE and Random Forest
- Non-parametric imputation to account for the lack of normality amongst the distributions of many of the selected variables
- Initial clustering:
 - Peers defined as
 - Cal State Campuses
 - UC Campuses
 - Caltech

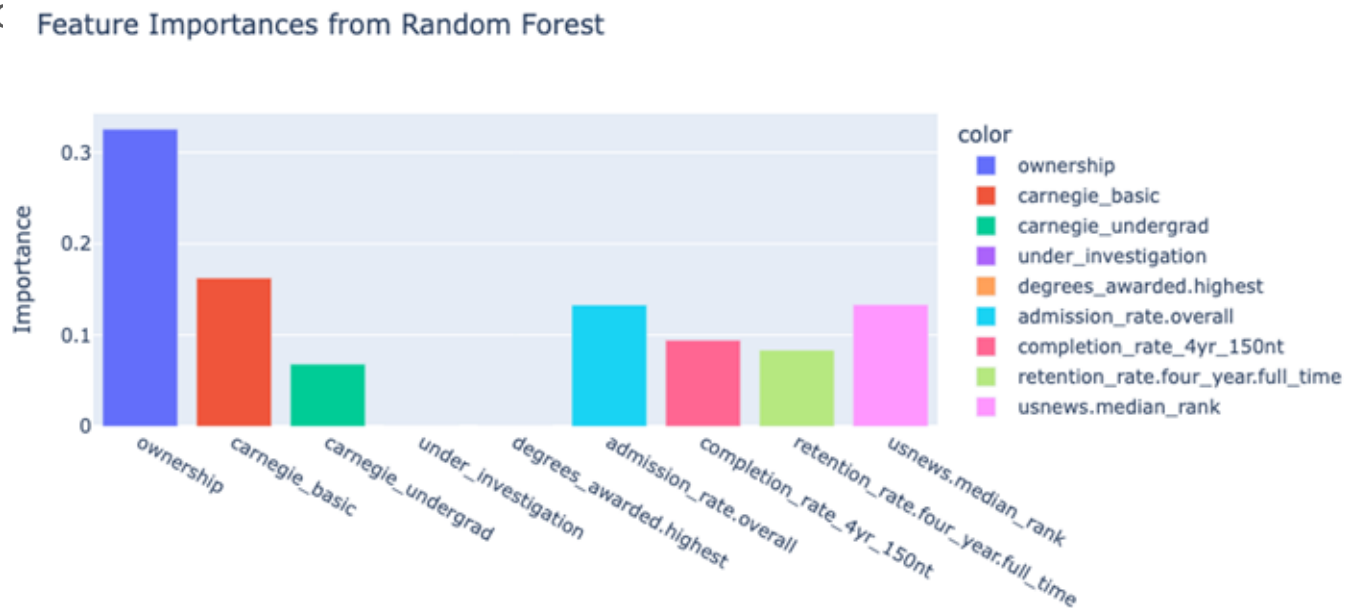
New Clustering

- Clustering **WITHOUT** demographic variables
 - Including aggregated college rankings in the equation
 - Aiming for less geographical relationships
 - More institutional similarity
- Defining peer institutions nationally:



New Clustering

- What determined these clusters? (using a RF model fit on the cluster outputs)

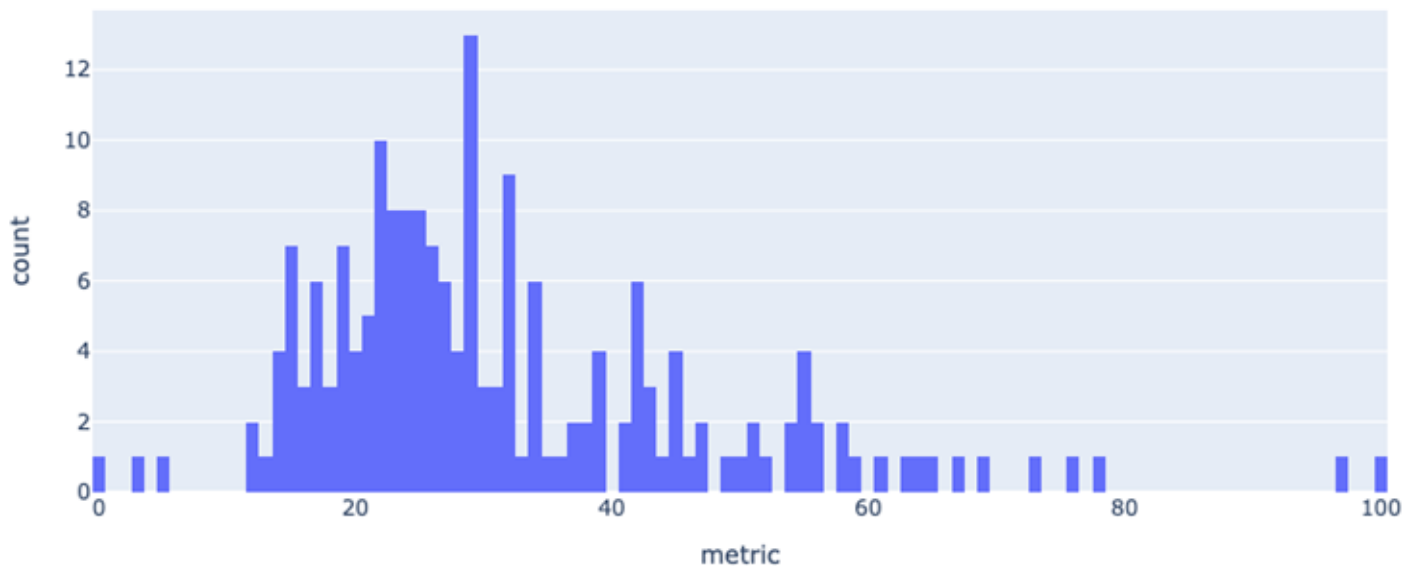


Defining a Financial Metric

- Using early career (defined as 1-5 years) median earnings from each university
 - All universities in the US News Rankings data
 - 1 year post-grad
 - 4 years post-grad
 - 5 years post grad
- Adjusting for location and cost of living
 - MIT living wage calculator data for livable wages, annual pretax income information
 - Cost of living used due to statistics mentioning a large tendency for many students to remain within 200 miles of university locations (note: this is based on reports from individual colleges and is considered an assumption here)
 - Creating a ratio between pretax income and livable wages to represent how much “better” than average salaries were

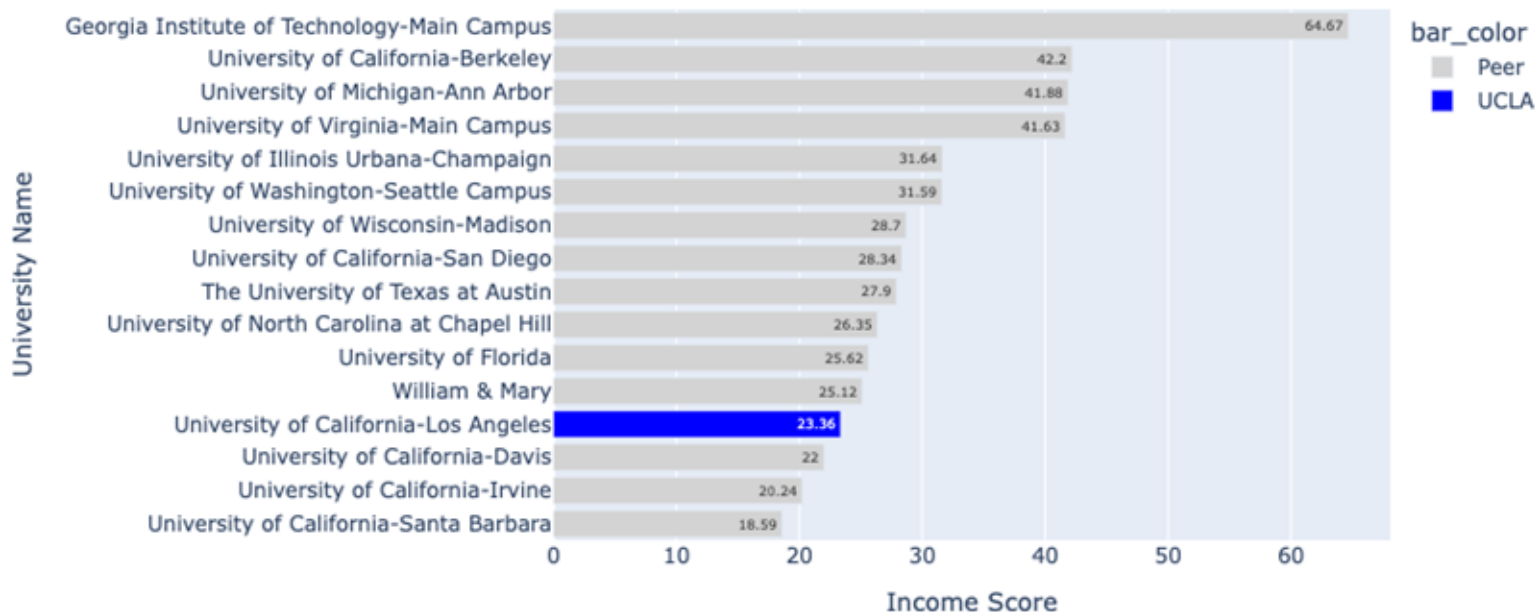
Defining a Financial Metric

Distribution of Income Metric $\text{metric} = (0.5 \cdot \text{earn}_{1\text{yr}} + 0.3 \cdot \text{earn}_{4\text{yr}} + 0.2 \cdot \text{earn}_{5\text{yr}}) \cdot \left(\frac{\text{income}_{\text{pretax}}}{\text{living wage}} \right) \rightarrow \text{rescaled to } [0, 100]$

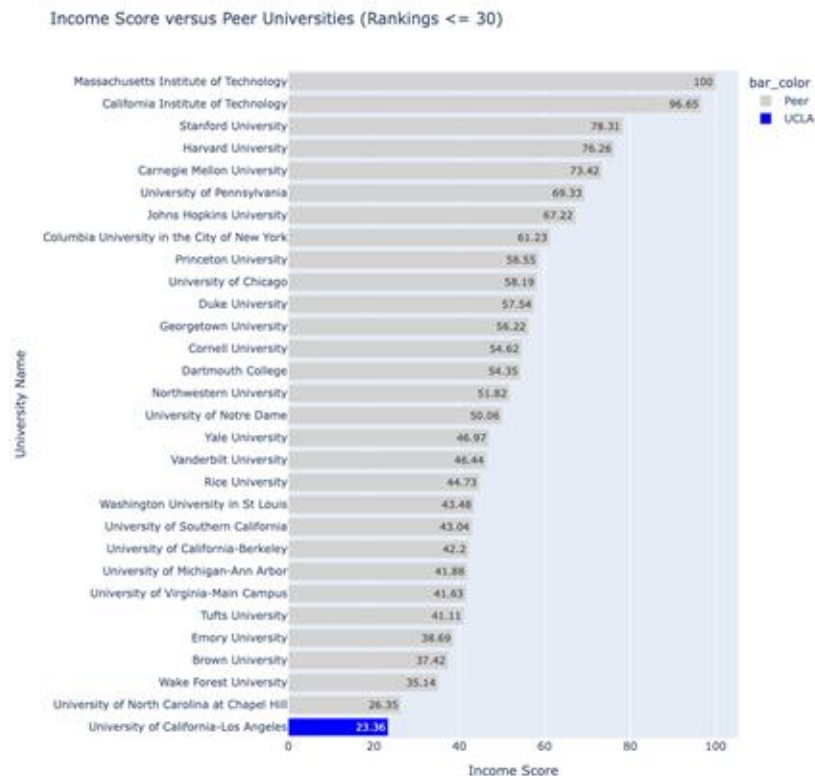


How does UCLA compare?

Income Score versus Peer Universities (Clustering)



How does UCLA compare?



What factors affect income?

Using all of our features from variable selection and predicting our financial metric

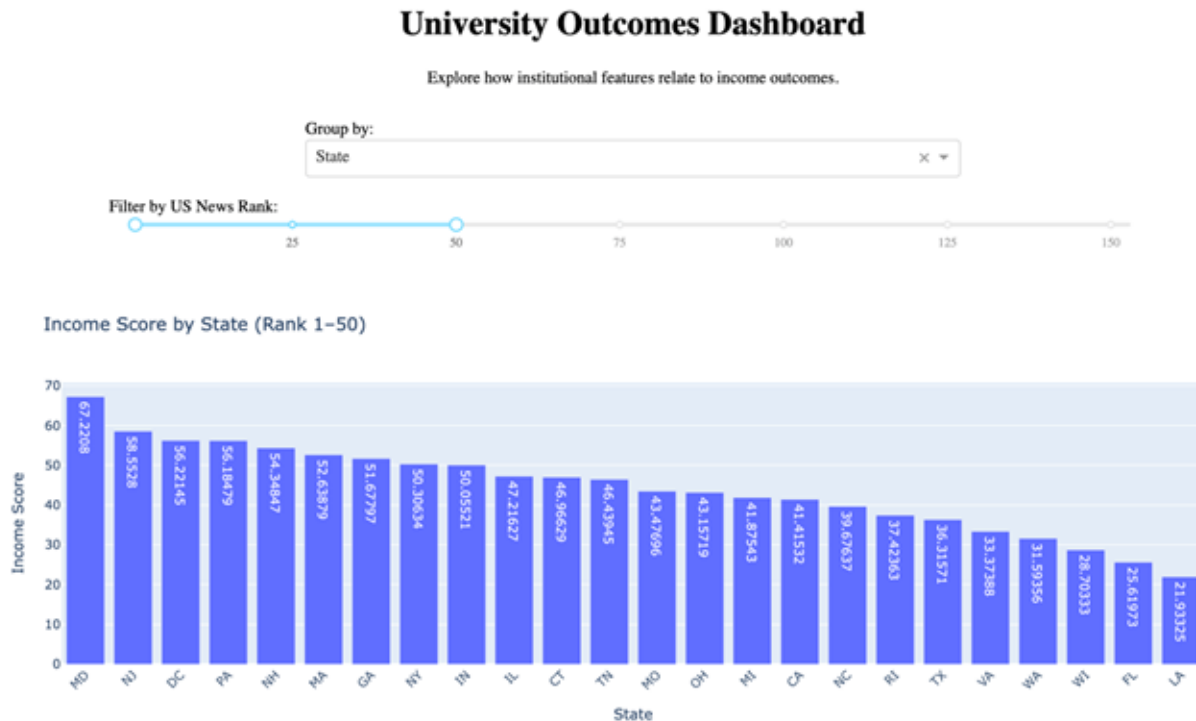
Feature Importances from Lasso (Nonzero Coefficients)



Institutional Factor Product: Dashboard

An interactive look at the effect of institutional factors on our financial metric

- <https://one41xp-income-dashboard.onrender.com/>



Suggestions for the Future (Prompt 1)

Challenges:

- Missing & Privacy Suppressed data
- Uneven program coverage across schools
 - Ivies especially do not offer as many low earning majors, skewing institutional comparisons in UCLA's disfavor
- Ambiguity in matching "similar programs"
 - CIP prefixes are extremely broad

Suggestions:

- Gather more data from a source outside of Scorecard and IPEDS, especially for graduate level programs
- Analyze across wider range of programs and investigate specific programs in more detail (4 digit CIP codes)
- Explore alumni career outcomes from non-monetary sources (ie: LinkedIn Scraping)

Suggestions for the Future (Prompt 2)

- Challenges:

- Gathering data was difficult due to not having access to readily available CSVs for certain variables
- Had to go through various APIs and many different web sources to find the right data
- Creating new variables derived from known variables (ex: estimating expected graduates from total enrollment and graduation rate)

- Suggestions:

- Further analyze the relationship between median earnings and predictors like enrollment and graduation rate through a more sophisticated model than linear regression
- Conduct a more thorough analysis that considers additional institutional factors such as rankings and economic variables(unemployment, tax rate, etc.)

Suggestions for the Future (Prompt 3)

Challenges:

- High missingness for Master's and Doctoral Programs
- Raw CIP data is difficult to visualize and communicate without grouping into broad categories

Suggestions:

- Utilize quantile regression to explore more in depth how UCLA compares to low and high earning ends instead of just average
- Examine longer term longitudinal outcomes beyond 5 and 10 years after graduation

Suggestions for the Future (Prompt 4)

Challenges:

- Uniqueness of UC institutional characteristics
 - The demographics, student selection criteria, and subject focus of the University of California system lead to a misleadingly low financial metric score
- Limited set of institutional factors

Suggestions:

- Incorporation of subject area based weighting for financial metric
- Further quantification of institutional characteristics to capture more aspects
 - Looking at policy for instance
 - Feature engineering