Family Income, School Ranking, and Acceptance Rate
Analyzing the Three Variables in U.S. University Admissions Processes

Yitong "Tony" Zhao

AP Statistics Final Project

Princeton International School of Mathematics and Science

Table of Contents

1 Introduction

The college admissions process in the United States is a complex and multifaceted system that involves various factors influencing students' acceptance into universities, like academic performance and extracurricular activities. However, family income remains a significant factor in determining students' chances of admission, as it may correlate with the quality of education and resources available to students. To address this issue of socioeconomic inequality, many top institutions claim to implement policies that help students from all backgrounds access similar opportunities.

To investigate the current state of family income's impact on college admissions, we conducted this study using data from the 2020 A2C census online survey [1], which was posted on r/ApplyingToCollege [2] (the largest U.S. college-admissions subreddit, with 1.2 million subscribers). We also hypothesize that the relationship between income and acceptance rate differs by institutional selectivity. To test this, we incorporate the WSJ/Times Higher Education 2020 U.S. ranking [3] to see whether higher-ranked schools exhibit a weaker or stronger income effect.

2 Data Sources

2.1 A2C Survey

The A2C census was administered on Reddit's r/ApplyingToCollege subreddit in 2020 (class of 2024). Survey participants self-reported demographic information (family income, race/ethnicity, gender, state of residence, etc.), standardized test scores (SAT/ACT), and admissions outcomes (accepted/waitlisted/rejected) for each school they applied to. Other data—such as intended field of study, high school GPA, and future plans—were collected but not used in this study. Our primary variable is the response to "Approximately how much is your family's total yearly income?" with these categories:

- Less than $20,000
- $20,000 - $45,000
- $45,000 - $70,000
- $70,000 - $100,000
- $100,000 - $150,000
- $150,000 - $200,000
- $200,000 - $250,000
- More than $250,000

To understand how this survey population compares to the broader pool of applicants, we examine key demographics in Figure 1.

Figure 1: Selected demographic information of the A2C census participants [4]

The survey, with 3,882 participants, provides a relatively large sample size, increasing the reliability of statistical analysis. However, from Figure 1, it can be observed that the data source has its limitations in unbiasedly representing a wider population, as the following characteristics are notable:

- There are considerably more male participants than female (57% vs. 40%).
- There are much more Asian participants than national average (42% vs. 6% [5]).
- Politically, the samples are strongly skewed towards the left, with 65% of participants identifying as liberal or progressive, while only 5% identifying as conservative or right-wing.
- More international students than national average (11% vs. 6% [6])
- The participants have significantly higher academic performance than the national average, as:
  - Half of the student have received GPA in the range of 3.75 - 4.00, whereas the national average is 3.0 [7].
  - The average SAT score of the participants is 1470, while the national average is 1050 [8].
  - The average ACT score of the participants is 33, while the national average is 19.4 [9].
- The participants are financially more privileged than the national average, as 24% of the them reported a family income of more than $250,000, while the national median is $80,610 in 2023 [10].

Possibly, these skewed demographics are due to the fact that the survey was posted on a subreddit that is primarily used by high-achieving students who spent a lot of time on college admissions.

2.2 School Ranking

Among major ranking systems, the WSJ/Times Higher Education list uniquely integrates both liberal arts colleges (LACs) and large universities—aligning with the A2C data, which includes both types of institutions. By combining LACs and universities into one ordered list, WSJ/THE lets us group every respondent's school without dropping any observations, maximizing sample size. In contrast, US News & World Report, while popular for U.S. schools, separates LACs and universities, and the QS World University Rankings focus on global research metrics but largely omit U.S. liberal arts colleges. In this study, top 200 schools from the WSJ/THE 2020 U.S. ranking are used.

3 Methodology

3.1 Overall Trend Across All Schools

We first examined the relationship between family income bracket and acceptance rate when combining all 200 schools, a line plot of which is shown in Figure 2.
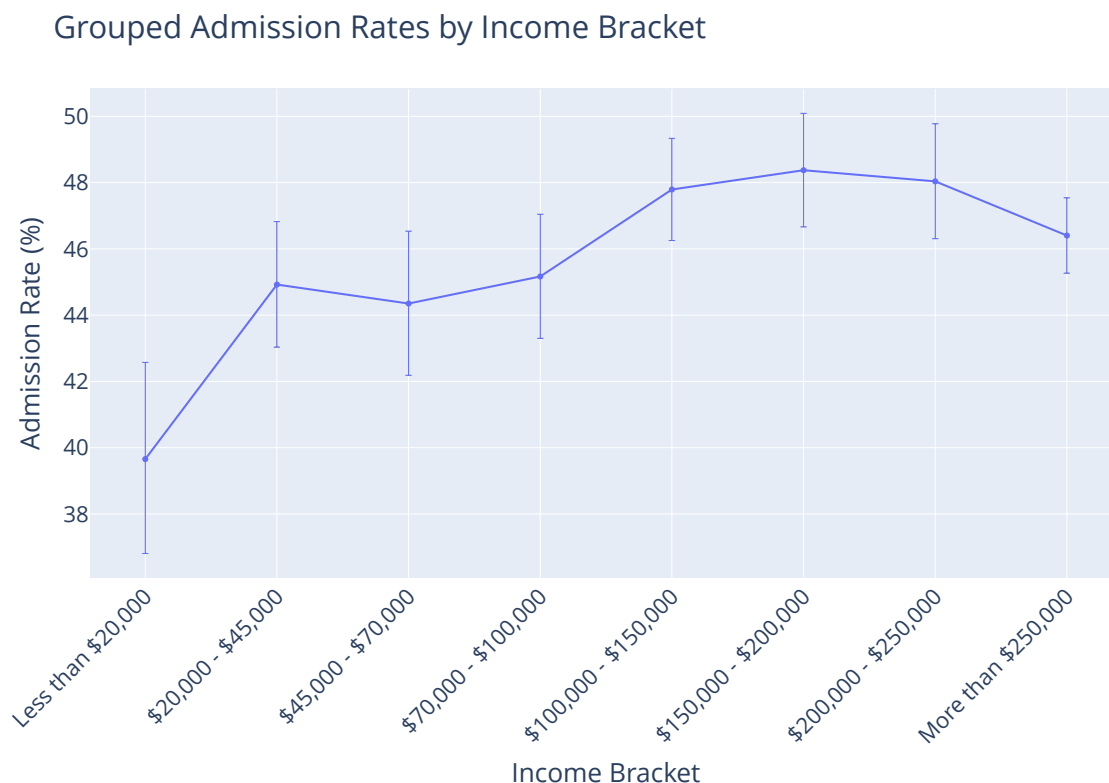
Figure 2: Overall acceptance rate by family income bracket across all schools. The error bars are 95% confidence intervals calculated using 1-proportion z-intervals.

From Figure 2, the following general trends can be observed:

- Acceptance rate initially increases as income rises.
- It reaches a peak around the $150k - $200k bracket.
- It then declines slightly in last bracket ($\geq$ $250k).
- The 95 % CI for the lowest-income bracket (less than $20k) is very wide due to a small sample size.
- The CI for the highest bracket (more than $250k) is comparatively narrow because of a larger sample.

3.2 School Grouping

Since we hypothesize that family income may exert different effects at institutions with varying selectivity—and because individual schools (especially those ranked lower) often have too few A2C respondents to yield stable estimates—we group schools into ranked tiers rather than analyze each separately. By sorting all 200 institutions into ranking-based cohorts, we ensure sufficient sample size in each group while capturing differences in factors like competitiveness, financial aid policies (need-blind vs. need-aware), public versus private status, and overall institutional wealth.

The first proposed grouping method is based on fixed interval, where schools are divided into 20 groups of 10 schools each. This approach is illustrated in Figure 3.
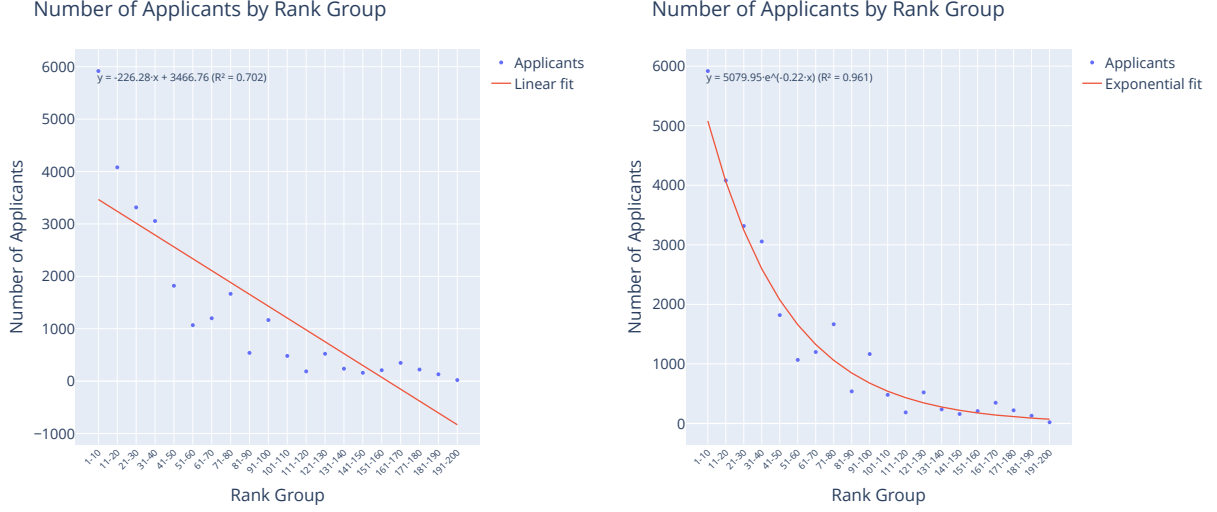
Figure 3: Applicant counts by school ranking group using fixed intervals. The left plot shows a linear fit, while the right plot shows an exponential fit.

It can be observed from Figure 3 that the number of applicants dropped gradually as the school ranking group increases, given a fixed interval of 10 schools. Both linear and exponential fits are conducted using least squares regression, and the R^2 value for the linear fit is 0.702, while the R^2 value for the exponential fit is 0.961. The high R^2 value of the exponential fit suggests that the number of applicants decreases exponentially as the school ranking group increases.

Due to the uneven distribution of applicants across fixed intervals, our purpose of grouping schools to ensure sufficient sample size is not achieved. We thus proposed a second grouping method based on quantiles, where schools are divided with dynamic intervals to ensure that each group contains approximately the same number of applicants. A simplified Python function to generate such intervals is provided below. It requires the first interval to be specified, which is used to set the target number of applicants for all subsequent intervals. For the actual implementation, please refer to Section 6

```python
def generate_intervals_by_applicants(
    rank_counts: dict[int, int],
    sorted_ranks: list[int],
    first_interval: tuple[int, int],
) -> tuple[list[tuple[int, int]], list[int]]:
    """
    Partition a sorted list of ranks into consecutive intervals so that each
    interval's total applicant count matches (as closely as possible) the
    applicant count of the initial interval.

    Args:
        rank_counts: Mapping from rank → number of applicants at that rank.
        sorted_ranks: Ascending list of all ranks to include.
        first_interval: A (low, high) tuple indicating the first rank interval,
                        whose sum of applicants sets the target for all intervals.

    Returns:
        intervals: List of (low, high) rank-intervals covering all ranks.
        counts:    List of applicant totals for each corresponding interval.
    """
```

```python
    low0, high0 = first_interval
    target = sum(rank_counts.get(r, 0) for r in sorted_ranks if low0 <= r <= high0)

    intervals = [first_interval]
    counts = [target]
    start = high0 + 1
    max_rank = sorted_ranks[-1]

    while start <= max_rank:
        cumulative = 0
        best_diff = float("inf")
        chosen_end = start

        for r in sorted_ranks:
            if r < start:
                continue
            cumulative += rank_counts.get(r, 0)
            diff = abs(cumulative - target)
            if diff < best_diff:
                best_diff = diff
                chosen_end = r
            if cumulative >= target:
                break

        if cumulative < target:
            chosen_end = max_rank

        intervals.append((start, chosen_end))
        counts.append(cumulative)
        start = chosen_end + 1

    return intervals, counts
```

With the dynamic grouping, we obtained 10 intervals with closer applicant counts, as shown in Figure 4.
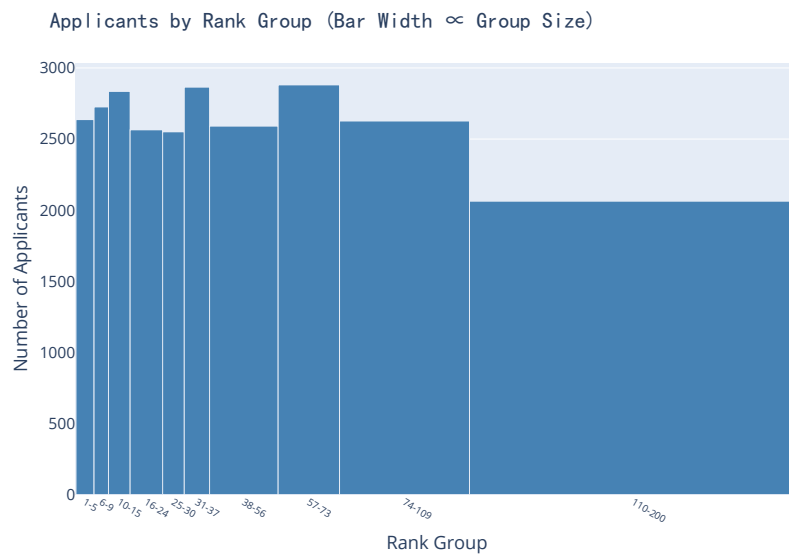

Figure 4: Applicant counts by school ranking group using dynamic intervals.

The 10 specific groups are as follows:

1-5 (1), 6-9 (2), 10-15 (3), 16-24 (4), 25-30 (5), 31-37 (6), 38-56 (7), 57-73 (8), 74-109 (9), and 110-200 (10). The amount of applicants are mostly within the range of 2500-3000 for each of the 10 groups.

With this grouping, we can obtain the following trends in different school ranking groups, as shown in Figure 5.
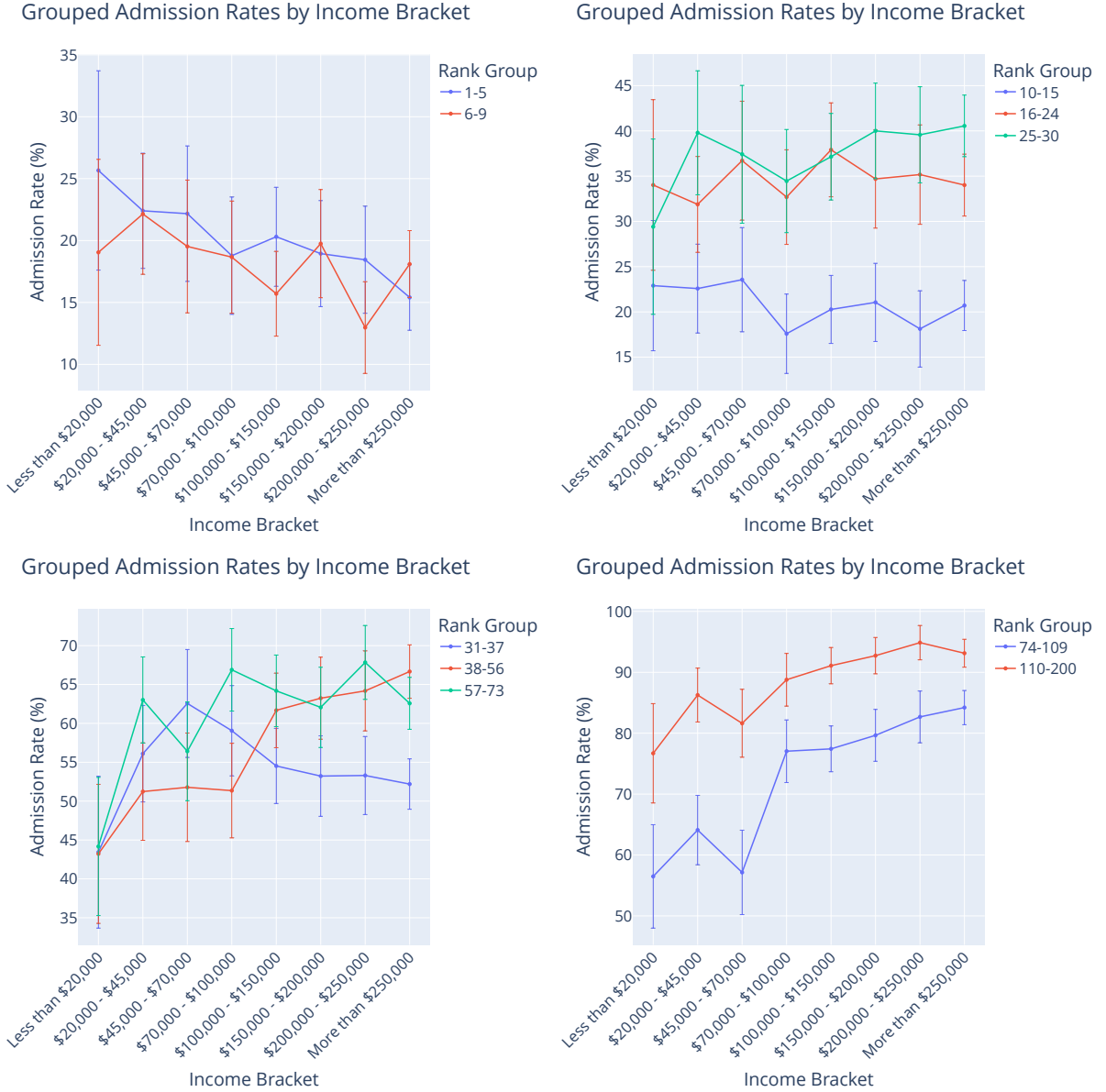

Figure 5: Acceptance rate by family income bracket across different dynamic interval school ranking groups. The error bars are 95% confidence intervals calculated using 1-proportion z-intervals.

From Figure 5, we confirmed our hypothesis that the relationship between family income and acceptance rate varies across different school ranking groups, which might be different from the overall trend shown in Section 3.1. Specifically, we can observe the following trends:

1. For ranking group 1-5 and 6-9, the acceptance rate decreases as family income increases.

2. For ranking group 10-15 and 16-24, no clear trend can be observed, while for group 25-30, the acceptance rate slightly increases as family income increases.
3. For ranking group 31-37, the acceptance rate increases significantly at first, then slightly decreases as family income increases.
4. For ranking group 38-56, 57-73, 74-109 and 110-200, the acceptance rate clearly increases as family income increases.

3.3 Statistical Tests

To better understand and verify the visual trends observed in Figure 5, we conducted the following two statistical tests.

3.3.1 χ^2 Test for Independence

The χ^2 test for independence is used to determine the significance of the association between two categorical variables. In this case, we used the 8 family income brackets and whether accept or not as the two categorical variables. The same test is conducted for each school ranking group to determine whether the acceptance rate is independent of family income. In this test,

- the null hypothesis H_0 is that family income and acceptance rate are independent.
- and the alternative hypothesis H_a is that family income and acceptance rate are dependent.
- the p-values means the probability of observing the data if the null hypothesis is true.
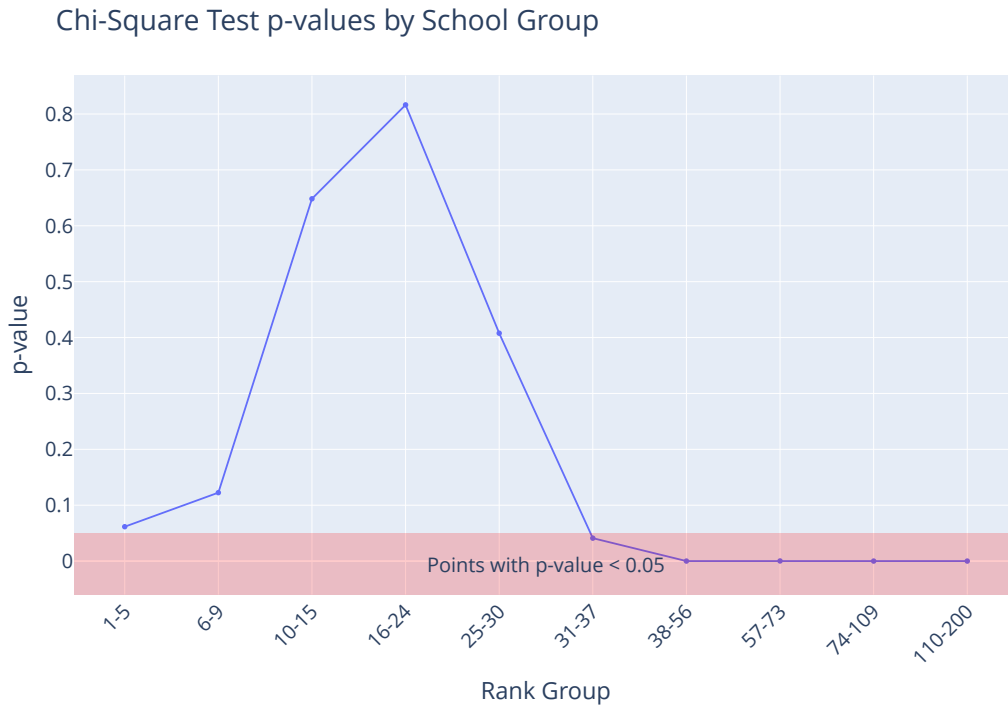

Figure 6: p-values of the χ^2 test for independence for each school ranking group. The red region indicates p-values less than 0.05, indicating a significant association between family income and acceptance rate.

From Figure 6, we can observed that the p-values for ranking groups 1-5, 6-9 are very low at first, then as the ranking group increases, the p-values increase significantly. The p-value, however, dropped again for ranking groups 31-37, 38-56, 57-73, 74-109 and 110-200, meaning there is a significant association.

Note that schools from 31-200 have p-value lower than 0.05, indicating that family income and acceptance rate are dependent given a $\alpha = 0.05$ significance level.

3.3.2 Regression Slope Test

Although the χ^2 test for independence can tell us the depenence between family income and acceptance rate, the direction remains unclear. To further investigate the direction of the relationship, we conducted a regression slope test for each school ranking group.

Specifically, the 8 income brackets are each assigned a index from 1 to 8, which is used as the explainatory variable x in the linear regression model. The acceptance rate is used as the response variable y. The regression model is then fitted using least squares regression, and the slope of the regression line is tested for significance using a t-test. In this test,

- the null hypothesis H_0 is that the slope of the regression line is 0, meaning there is no relationship between family income and acceptance rate.
- the alternative hypothesis H_a is that the slope of the regression line is not 0, meaning there is a relationship between family income and acceptance rate.
- the p-values means the probability of observing the data if the null hypothesis is true.
- the r-value is the correlation coefficient, which indicates the strength and direction of the relationship between family income and acceptance rate.

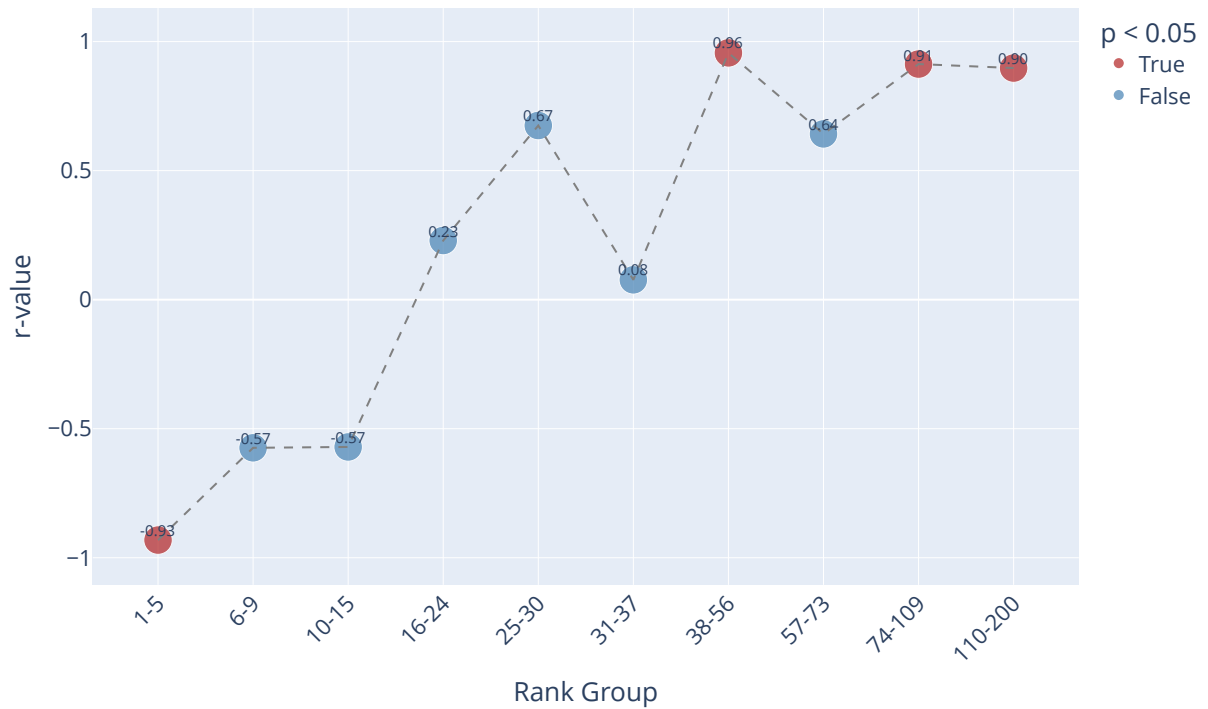## Income-Admit Correlation by School Group

Figure 7: Regression slope test result for each school ranking group. Red dots means that the p-value is less than 0.05, indicating a significant relationship between family income and acceptance rate.

From Figure 7, we can observe that there exist a positive relationship between school ranking group and r-value, meaning that as the school ranking group increases, schools tend to accept students from higher family income brackets more. Note that group 1-5, 6-9, and 10-15 have negative r-values, indicating that students from lower family income brackets are more likely to be accepted. However, among them the p-values for group 6-9 and 10-15 groups are not significant, meaning that the relationship is not statistically significant.

4 Discussion

Overall, combined data reveal an inverted-U income-acceptance curve, but stratification by ranking uncovers that top tiers actually favor lower- to middle-income applicants, mid tiers show mixed effects, and lower tiers steadily favor higher-income applicants.

We have proposed the following potential explainations for the aforementioned findings:

4.1 Need-Blind Admissions at Top Institutions

Many of the highest-ranked schools (e.g., Princeton, Harvard, Stanford, MIT, University of Pennsylvania, Cornell) implement need-blind admissions. In fact, all top 10 universities on the WSJ/Times Ranking uses need-blind systems [11]. In need-blind syste, an applicant's ability to pay is not considered during the admission decision; financial aid is arranged afterward if admitted. As a result, low- and middle-income applicants who meet academic and extracurricular criteria face the same admissions standards as wealthier peers.

4.2 Holistic Review Practices

Top universities often practice holistic review, evaluating the entirety of an applicant's background—academic metrics, essays, extracurricular impact, personal essays, and the adversity they have overcome—rather than relying solely on grades and test scores. Holistic review can mitigate the resource gap faced by low-income students who may lack access to expensive test-prep or elite summer programs. In practice, a low-income student with strong leadership in community service or a first-generation background—and who can demonstrate how they overcame financial hardship—may be viewed more favorably than a wealthy student with similar scores but fewer demonstrable challenges. For example, both Princeton and MIT explicitly list "first-generation status" and "work experience" as factors "considered" in their Common Data Sets (CDS)[12], [13].

4.3 Under-Representation of Ultra-High Income Applicants

A *New York Times* article [14] based on large scale research [15] have reported that, given the same test score, students with top 0.1% of parent income are 2.2 times more likely to be admitted to selected ivy-Plus universities than the average value.

Admissions rate at elite colleges among students with the same test scores

TOP 0.1%
2.2x likelier admission

Twice as likely to be admitted

Average admissions rate

Half as likely to be admitted

| 10th | 30th | 50th | 70th | 90th | 99.9 |

← POORER   Parents' income rank   RICHER →

Data is from at least three of the dozen top colleges where the researchers had access to detailed admissions records.
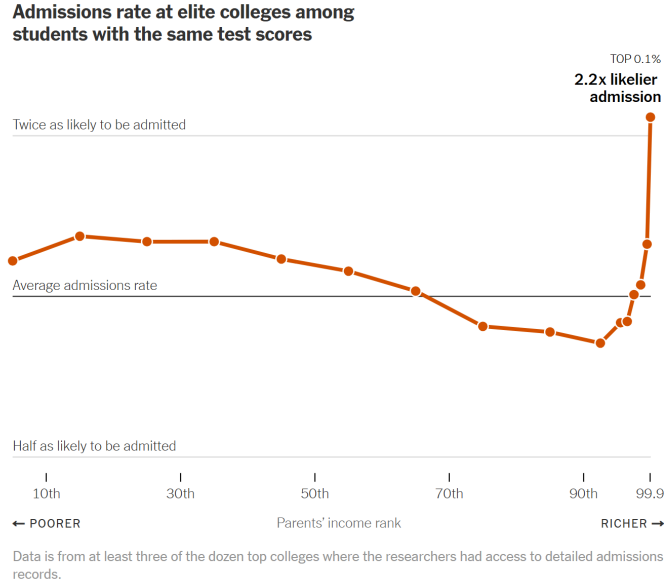
Figure 8: A figure from the *New York Times* article [14], showing the relationship between parent income and admission rate at top universities.

Although the survey we used ([1]) provided 8 different brackets that covers a wide range of family income, it doesn't have a more fine-grained bracket for the ultra-high income group. Althoguh the absolute income value of the top 0.1% is not disclosed in the research[15], according to the IRS, the top 0.1% of U.S. Adjusted Gross Income in 2022 was \$3,271,387 [16], which is way higher than the baseline of the highest bracket in the A2C survey (\$250,000).

Since the income brackets in the survey failed to capture the ultra-high income group, it is possible that the acceptance rate for the ultra-high income group is much higher than the highest bracket in the survey. This could explain why high-tier schools in Figure 5 did not show an increase in the highest income group, as shown in Figure 8.

5 Conclusion

This study demonstrates that family income and school selectivity jointly shape college acceptance rates. When combining all 200 institutions, acceptance follows an inverted-U across income brackets—rising from low incomes to roughly \$150k–\$200k, then declining among the \$250k+ bracket. However, stratifying by WSJ/THE ranking tiers reveals distinct patterns: the most selective schools (ranks 1–5, 6–9) actually admit proportionally more lower- and mid-income applicants (potentially due to need-blind and holistic policies), mid-rank schools show mixed or weak income effects, and lower-rank tiers exhibit steadily increasing acceptance with income. These findings suggest that top institutions' admissions practices can offset some income disparities. However, survey limitations undercount ultra-high-income advantages. Overall, while income remains a significant factor, its impact varies markedly with institutional selectivity and admissions policies.

6 Appendix

The entire project, including raw data, output interactive graphics, presentation and report source file, and code used for data analysis and visualization is available on this repository: https://github.com/ttzytt/ap_stats_final_project

Spefically,

- **Raw data, including ranking and survey responses** is available at: https://github.com/ttzytt/ap_stats_final_project/tree/main/data
- **Interactive graphics in `html` format** is available at: https://github.com/ttzytt/ap_stats_final_project/tree/main/output
- **Presentation source file** is available at: https://github.com/ttzytt/ap_stats_final_project/tree/main/pre
- **The presentation slide itself** can be viewed at: https://ttzytt.com/ap_stats_final_project/pre
- **Source code for data organization** is available at: https://github.com/ttzytt/ap_stats_final_project/blob/main/src/organization.py
- **Source code for data analysis** is available at: https://github.com/ttzytt/ap_stats_final_project/blob/main/src/analysis.py
- **Source code for data visualization** is available at: https://github.com/ttzytt/ap_stats_final_project/blob/main/src/visualization.py

References

[1] u/powereddeath, "A2C 2020 Census Results (Class of 2024)," 2020. https://www.reddit.com/r/ApplyingToCollege/comments/gucq7r/ra2c_2020_census_results_class_of_2024/?utm_source=share&utm_medium=web3x&utm_name=web3xcss&utm_term=1&utm_content=share_button.

[2] "r/ApplyingToCollege." https://www.reddit.com/r/ApplyingToCollege/.

[3] "Wall Street Journal/Times Higher Education College Rankings 2020." https://www.timeshighereducation.com/rankings/united-states/2020.

[4] "r/A2C '19-'20 Census Results." https://drive.google.com/file/d/1YNwyFPiQKvfsd-Ds8AZf_PRMzU8TEYZy/view.

[5] Wikipedia contributors, "Race and ethnicity in the United States — Wikipedia, The Free Encyclopedia," 2025. https://en.wikipedia.org/wiki/Race_and_ethnicity_in_the_United_States#:~:text=East%20Asian%20and,Asian%20(1.74%25).

[6] K. Buchholz, "Where America's International Students Come From." https://www.statista.com/chart/20010/international-enrollment-in-higher-education/#:~:text=International%20students%20make%20up%205.9%20percent%20of%20the%20total%20U.S.%20student%20population%2C.

[7] College Board, "How to Calculate Your GPA on a 4.0 Scale." https://bigfuture.collegeboard.org/plan-for-college/get-started/how-to-calculate-gpa-4.0-scale#:~:text=The%20average%20GPA%20across%20all,all%20U.S.%20colleges%20is%203.1..

[8] College Board, "What is a Good Score?." https://satsuite.collegeboard.org/scores/what-scores-mean/what-is-good-score#:~:text=The%20average%20SAT%20score%20is%20around%201050.

[9] ACT, "Average 2024 ACT Test Scores by State." https://satsuite.collegeboard.org/scores/what-scores-mean/what-is-good-score#:~:text=The%20average%20SAT%20score%20is%20around%201050.

[10] United States Census Bureau, "Income in the United States: 2023." https://www.census.gov/library/publications/2024/demo/p60-282.html#:~:text=Highlights,median%20household%20income%20since%202019..

[11] A. Robinson, "All 116 Need-Blind Colleges in the US: A Complete Guide." https://blog.prepscholar.com/need-blind-colleges-list#:~:text=Csunyo/Unsplash-,A%20Complete%20List%20of%20Need%2DBlind%20Colleges%20in%20the%20United%20States,-Here%20is%20a.

[12] Princeton University, "Common Data Set." https://ir.mit.edu/projects/2023-24-common-data-set/.

[13] MIT Institutional Research, Office of the Provost, "Common Data Set 2022-2023." https://web.mit.edu/ir/cds/cds2223.pdf.

[14] A. Bhatia, C. C. Miller, and J. Katz, "Study of Elite College Admissions Data Suggests Being Very Rich Is Its Own Qualification," *The New York Times*, Jul. 2023, [Online]. Available: https://www.nytimes.com/interactive/2023/07/24/upshot/ivy-league-elite-college-admissions.html.

[15] R. Chetty, D. J. Deming, and J. N. Friedman, "Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges," 2023.

[16] Wikipedia contributors, "Income in the United States." https://en.wikipedia.org/wiki/Income_in_the_United_States#:~:text=%2417%2C855%2C123-,%243%2C271%2C387,-%24663%2C164.