## **Group 4: Analyzing FAFSA Trends and Institutional Impact Yuetong Yin, Meixu Chen, Guanyu Qu**

**Link to our github repository:** [**https://github.com/guanyuq03/DS4200\_Group4\_FAFSA\_Project**](https://github.com/guanyuq03/DS4200_Group4_FAFSA_Project)

**Link to our website:**[**https://guanyuq03.github.io/DS4200\_Group4\_FAFSA\_Project/**](https://guanyuq03.github.io/DS4200_Group4_FAFSA_Project/) **FAFSA Application Trends Dashboard**

Our FAFSA dashboard provides an interactive, user-friendly platform to explore how FAFSA applications vary across institutions, states, and student demographics in the U.S. education system. The website allows users to switch between five quarterly periods (Q1–Q5), giving temporal insight into financial aid patterns. We integrated D3.js, Plotly, and Altair to present diverse visualization types—including histograms, bar charts, maps, scatter plots, and boxplots—each highlighting a different facet of the dataset.

The site’s layout emphasizes clarity and visual balance. The color scheme blends blues and whites to signal professionalism and trust, aligning with the topic of financial aid. Smooth transitions, responsive tabs, and hover-enabled tooltips enhance interactivity and accessibility. Altogether, our goal is to support deeper understanding for policymakers, institutional planners, and students by revealing how FAFSA applications reflect institutional type, dependency status, state-level demand, and quarterly trends.

**1. Distribution of FAFSA Applications (Histogram)**

This Altair histogram illustrates how FAFSA application counts are distributed across all schools during a selected quarter (Q1–Q5). The majority of institutions cluster toward the lower end of the distribution, while a minority receive significantly higher numbers of applications. This skewed distribution suggests that most FAFSA activity is concentrated in a few large universities, likely due to their enrollment scale, affordability, and public funding.

From a design standpoint, the histogram uses binning and responsive axes to create a clear, high-level view of application density. Each bin groups institutions by application count ranges, making it easy to identify whether patterns change over quarters. Interactivity is handled by Vega-Lite and updates dynamically when a new quarter is selected. By using tooltips and standardized sizing, we ensure that even institutions with relatively small values are not visually neglected. The consistent width and height improve user experience across devices.

This visualization encourages users to think about resource allocation, institutional scale, and the systemic concentration of FAFSA demand, which can inform debates about federal aid reform or student outreach programs.

**2. Top 10 Institutions by FAFSA Applications (Bar Chart)**

This D3 bar chart presents the top 10 institutions that received the highest number of FAFSA applications in the selected quarter. It shows that public universities dominate the list, reflecting their accessibility, affordability, and enrollment size. Users can instantly see how much more these institutions receive compared to others, and this visualization supports comparisons across quarters to detect changes.

To improve readability, especially for long school names, the x-axis labels are rotated and spaced out. The y-axis is scaled dynamically to emphasize variation in application volume. Tooltip interactions provide precise values for each bar, and color changes on hover help draw attention during exploration. We also added significant bottom margin space to prevent label overlap.

From a design perspective, this chart was chosen because of its ability to highlight outliers and make clear comparisons across institutions. It answers questions like: Which schools have the highest FAFSA demand? Do the same schools lead every quarter? It encourages institutional benchmarking and gives insight into public vs. private reach in financial aid distribution.

**3. FAFSA Applications by State (Map)**

This choropleth map (built with Plotly) visualizes total FAFSA applications by U.S. state in the selected quarter. States with higher application counts are shaded darker, with a logarithmic color scale to prevent data skew. California, Texas, and New York typically stand out, demonstrating their large populations and number of institutions.

We implemented a logarithmic scale for better visibility of lower-application states, avoiding a flat visual where small states would otherwise be indistinguishable. Hover tooltips reveal actual application counts for each state, supporting quick comparisons. The use of a `YlGnBu` color palette aligns with educational themes while providing intuitive gradients for quantity.

Design-wise, this map was chosen to give a national perspective on FAFSA distribution. It bridges geographic and institutional data, enabling users to consider how regional demographics or education policies affect FAFSA activity. For example, users may explore how states with fewer schools or lower population densities still manage aid applications. Overall, the choropleth complements institution-level graphs by aggregating insights at the state level.

**4. Dependent vs. Independent Applicants (Scatter Plot)**

This D3 scatter plot compares the number of dependent vs independent applicants across institutions in the selected quarter. Each point represents a school, with the x-axis showing dependent students and the y-axis showing independent students. A regression line is calculated to show correlation.

This graph allows for state filtering, so users can zoom into specific regions and explore how institutional dependency trends vary locally. The inclusion of a regression line provides a statistical anchor to visualize the relationship strength between the two types of applicants.

Design choices include dynamic axis scaling and tooltips for individual school data. By offering a quadrant-style view, the plot lets users identify outliers—such as schools that serve mostly independent students (often older, returning learners or students without parental support).

The goal of this plot is to reveal demographic trends in financial aid need. For example, a high number of independent applicants might indicate institutions serving working adults or non-traditional students. This plot fosters more nuanced policy insights into how different student groups interact with financial aid.

**5. Top 10 States by FAFSA Applications: Dependent vs. Independent (Grouped Bar Chart)**

This grouped bar chart (D3) ranks the top 10 states by total FAFSA applications, broken down into dependent and independent applicants. Each state has two adjacent bars, allowing clear side-by-side comparison.

The visualization highlights regional differences in student demographics. For instance, some states show a higher proportion of independent students—possibly due to economic conditions or education system design—while others are dominated by traditional dependent applicants.

From a design standpoint, we used contrasting colors (blue and orange) to distinguish applicant types and added a legend. Bars are sorted by total volume to guide user attention from most to least impactful states. We also included axis label rotation and tooltips for detailed readability.

This visualization helps users explore not just the volume of FAFSA applications, but the type of student applying. It’s especially useful for comparing how social, economic, or policy factors influence student independence across states.

**6. FAFSA Applications by Institution Type (Altair Boxplot)**

This final chart uses a Vega-Lite boxplot to show the distribution of FAFSA application counts by institution type—Public, Private, Proprietary, and Unknown—across all five quarters. Each column represents a different quarter, allowing users to track changes in median and variance over time.

Boxplots were chosen for their ability to represent data spread, outliers, and central tendency. This format quickly shows that public institutions consistently serve the largest applicant pools, while private institutions show wider variance and proprietary schools receive far fewer applications.

The design benefits from Altair’s clean automatic layout: evenly spaced columns, responsive sizing, and consistent styling. Color is used to differentiate institution types, but the emphasis remains on comparing medians and ranges.

This visualization is especially insightful for institutional researchers and policymakers. It answers questions like: Are public schools carrying most of the aid burden? How much variability exists within each category? Are private schools consistently polarized between elite and low-volume institutions?