**Visualization Write-Up**

1. **Video Link to the Visualization:**
   1. <https://youtu.be/wdJmjvMcs1U>
2. **Title of Visualization:**
   1. "Exploring College Metrics: Regional and Financial Insights"
3. **Why This Visualization?**
   1. The visualization is an interactive scatter plot with regional filtering, designed to help prospective college students, parents, and education researchers explore key college metrics. It aims to answer common questions, such as which colleges balance cost, debt, and outcomes, and how these vary by region.
   2. Concept and Audience:
      1. This visualization targets individuals making data-informed decisions about higher education. Its dynamic and interactive design simplifies the exploration of complex datasets.
   3. Questions About the Data:
      1. The visualization focuses on questions such as: How do colleges vary in cost, student debt, and graduate earnings?
      2. Are there correlations between metrics like SAT scores, income, and admissions rates?
      3. How do these attributes differ by region?
   4. Data Type:
      1. The data combines numerical variables (e.g., SAT/ACT scores, admission rates, median family income) and categorical variables (e.g., region). A scatter plot is ideal for visualizing relationships between numerical attributes, while color encoding and filtering enhance regional insights.
   5. Why a scatter plot?
      1. I initially considered a bar chart but found it unsuitable due to the dataset's complexity and clutter. A scatter plot proved better for clearly showing trends, correlations, and outliers.
         1. Ease of Interaction: Scatter plots allow users to select points for details, offering a more seamless exploration than overlapping bars.
         2. Visibility: Each data point is clearly represented as a circle, with region-based colors grouping colleges for better differentiation.
         3. Trend Identification: Scatter plots effectively reveal patterns and correlations, such as relationships between SAT scores and admission rates or family income and median debt, which were harder to discern in bar charts.
4. **Why These Variables and Encodings?**
   1. X and Y Axes:
      1. The dataset contained numerous variables, so I selected those most relevant to prospective students, parents, and researchers. By focusing on metrics like SAT/ACT scores, cost, and income for the X and Y axes, the visualization remains flexible and aligned with key decision-making factors.
      2. X-Axis: Pre-Admission Factors:
         1. The X-axis highlights "before getting into college" metrics, helping students evaluate institutions based on academic preparation and financial circumstances:
         2. Median Family Income: Socioeconomic context of the student body.
         3. % Federal Loans: Financial structure of enrolled students.
         4. SAT Average and ACT Median: Measures of academic readiness.
         5. Average Cost: Reflects the financial burden of attendance.
      3. These variables ensure users can explore colleges based on their personal and academic profiles.
      4. Y-Axis: Post-Admission Outcomes
         1. The Y-axis focuses on "after getting into college" metrics, helping users assess institutional impact and value:
         2. Admission Rate: Likelihood of acceptance.
         3. Mean Earnings 8 Years After Entry: Financial return on investment.
         4. Median Debt: Financial burden post-graduation.
   2. Region Colors:
      1. Regions are assigned distinct colors to highlight differences in college metrics across geographical areas. This ensures clarity and accessibility, helping users identify trends and clusters.
   3. Interactive Sidebar:
      1. The sidebar displays all key attributes for selected colleges, dynamically updating when users interact with the scatter plot. Even when no college is selected, placeholder values maintain consistency, reducing cognitive load.
   4. Filters:
      1. Region-based checkboxes allow users to focus on specific areas, tailoring the data to their needs. This modular design enhances usability without overwhelming the chart.
5. **Insights Learned Through the Visualization**
   1. Correlations and Insights:
      1. An increase in federal loan reliance leads to higher median debt, highlighting financial burdens faced by students at colleges with limited scholarships or grants.
      2. Higher family income correlates with increased student debt due to reduced eligibility for need-based aid and a greater willingness to incur debt for education.
      3. Colleges with higher average costs generally have lower admission rates, reflecting their selectivity and prestige. Conversely, more affordable colleges tend to have higher admission rates, appealing to a broader range of students.
      4. Higher SAT/ACT scores correlate with lower admission rates, showing the selectivity of institutions with competitive applicant pools.