Group1: Alena Zheng, Jiaqi Wu, Peiyuan Li, Qi Qiao, Yanqi Liu

Stat436 Project Milestone 2: Exploring the Design Space

Word Count: 962

ShinyIO Deployment: https://pli233.shinyapps.io/Milestone2/

Published Source Code:

https://github.com/pli233/Working/tree/main/Stat436/Group%20Project/Milestone2

Stat436 Project Milestone 2: Exploring the Design Space

Introduction

In Milestone 1, we uncovered international students' profound contributions and economic impacts on the American higher education system. Building upon the solid theoretical foundation established in Milestone 1, we will adopt the perspective of educational institutions to delve deeper into international student data. We aim to elevate static datasets into dynamic visual narratives, designed to support educational institutions in formulating sophisticated student engagement and recruitment strategies.

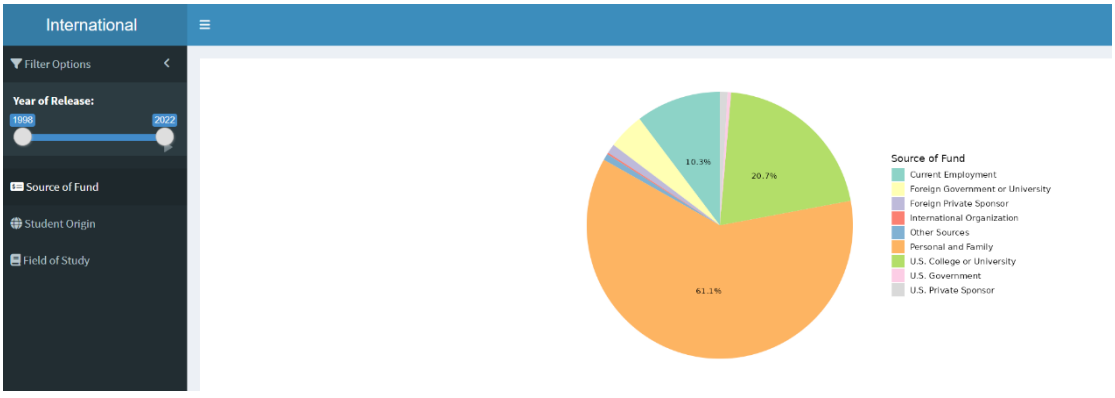
Milestone 2 promises to unveil a comprehensive suite of visualization components. These components are intended to synthesize our research findings and demonstrate the efficacy of our Design Studio approach. Our objective extends beyond merely presenting the outcomes of exploratory design; we also seek to pave

the way for a deeper and more informed understanding of the international education landscape.

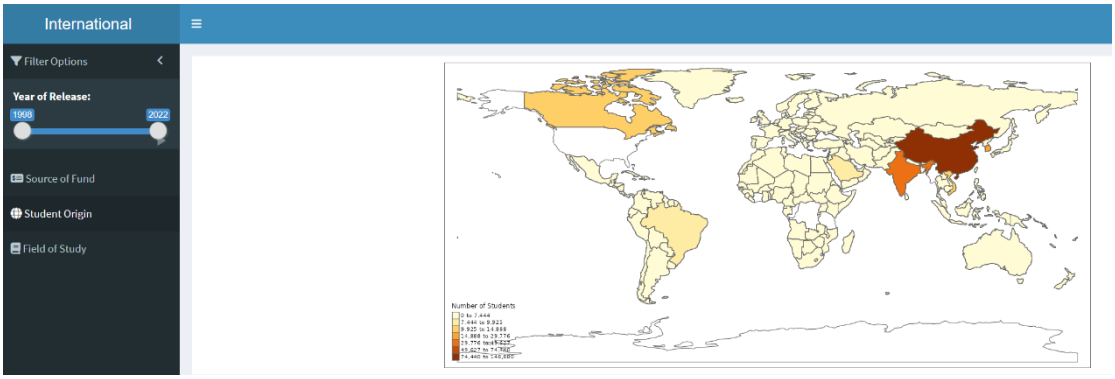
Choice of Design Prototypes

We take the advantage of Shiny web application as the main user interface. It offers a clean, intuitive interface that encourages deep exploration of international student data. The layout, with its clear headers, sidebars, and body sections, is streamlined for easy navigation and comprehensive analysis. With a filtering feature prominently positioned on the left, users are allowed to adjust the time frame of the data under scrutiny, offering a range from 1998 to 2022. Such interactivity is essential, allowing users to hover on specific temporal data segments, enhancing the application's utility for trend analysis and strategic planning.

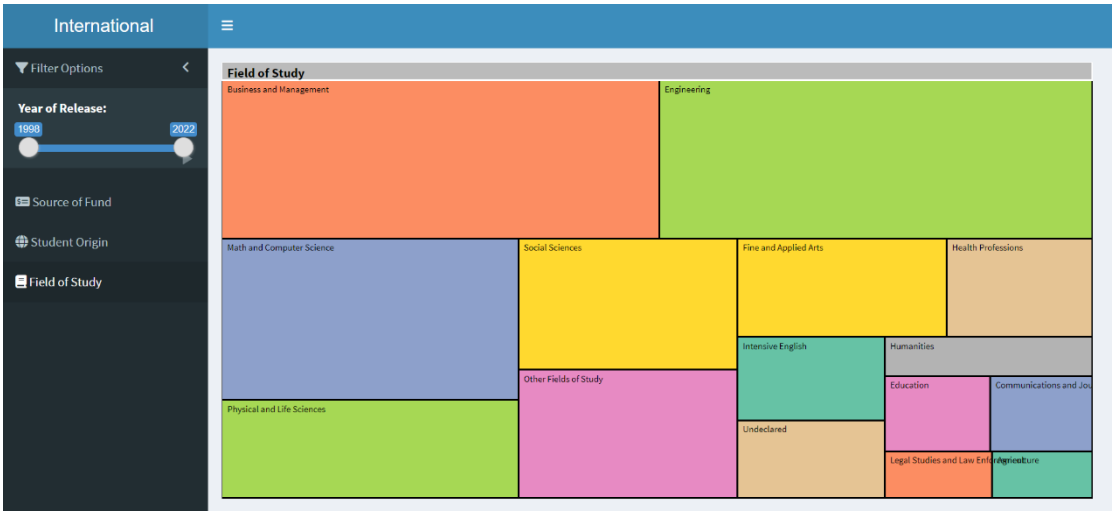
Prototype 1 (Pie Chart for Source of Fund): Selected for its adeptness at representing categorical distributions, the pie chart is an embodiment of simplicity and clarity. It serves as a mirror reflecting the diverse financial support systems of international students, visually partitioning the data to underscore the proportions and influence of each funding source. This pie chart transcends mere aesthetics; it becomes an analytical tool for stakeholders, illuminating the economic currents shaping educational journeys.



Prototype 2 (Choropleth Map for Student Origin): The choropleth map paints a geographical tapestry of student origins, where each color gradient is a narrative thread revealing the density of students from different regions. This visualization capitalizes on spatial patterns, making it an intuitive medium for stakeholders to navigate the global dynamics of international education. It is a testament to our commitment to delivering data-driven insights that are both comprehensive and comprehensible.



Prototype 3 (Treemap for Field of Study): Our selection of a treemap to depict the hierarchical nature of students' academic pursuits is a calculated decision to harness its spatial economy and categorical lucidity. The treemap's vibrant mosaic is informative, revealing the relative popularity of fields of study in a format that promotes easy comparison and instant recognition.



Demonstration and Implementation Discussion

The R code we've developed serves a crucial role in the data pre-processing and visualization phases, ensuring datasets are cleaned and structured effectively for analysis. Notably, functions like preprocess_dataset have been engineered to refine specific aspects, such as the 'year' column, thereby streamlining datasets for more efficient analysis. This preprocessing is integral to our Shiny application, which leverages dynamic server logic to respond to user inputs—like selecting a year range—automatically filtering datasets to refresh visualizations accordingly. This interactivity is a testament to the application's ability to merge analytical rigor with user-friendly interfaces, enhancing the exploration experience.

```r
#2. Create Function to preprocess datasets by separating 'year' column
preprocess_dataset <- function(data) {
  filtered_data <- data %>% separate(year, sep = "/", into = "year", convert = TRUE, extra = "drop")
  return(filtered_data)
}
```

Our implementation strategy extends beyond code to the thoughtful construction of a user interface via the shinydashboard package, emphasizing intuitive navigation and interactive dashboard elements. The application showcases a range of visualizations, from pie charts to choropleth maps and treemaps, each designed to be dynamic and responsive to user interactions. This approach enriches the user experience and ensures that visualizations remain informative, readable, and visually appealing, regardless of the dataset's size or complexity.

Additionally, we have independent methods for drawing graphs for each module. This allows us to handle image details quickly and accurately, as well as add or remove specific templates, during the later stages of the project.

```
#4.1 pie graph generator
pie <- function(df) {···
}

#4.2 map graph generator
map <- function(df) {···
}
```

```
#4.3 tree graph generator
tree <- function(df) {···
}

#4.4 inter graph generator
inter <- function(p) {···
}
```

Overall, the core of our project lies in its clarity, modularity, and adaptability, which facilitates current data visualization needs and future enhancements. Our Shiny application and its underlying code structure are benchmarks for developing scalable and replicable data visualization tools in the educational sector. Through this blend of technical sophistication and strategic design, we aim to transform raw data into powerful narratives and actionable insights, underscoring our commitment to advancing educational analytics.

Critical Evaluation

Upon evaluating our implemented visualization approaches using real data, we identified several trade-offs for improvement. Firstly, although the pie chart effectively displays the distribution of funding sources, it may suffer from overcrowding if there are too many categories, potentially leading to difficulty in interpretation. To address this, we plan to explore interactive features such as tooltips or drill-down capabilities, which allow users to explore specific categories in more detail while maintaining a clear overview.

Secondly, although the choropleth map offers insights into the geographic distribution of student origins, it may lack granularity in densely populated areas. This may obscure some localized trends potentially. To improve this, we intend to incorporate zooming and filtering functionalities, enabling users to focus on specific regions of interest and obtain more detailed information.

Thirdly, while the treemap efficiently represents the hierarchical breakdown of fields of study and majors, it may become cluttered with a large number of nodes and diminish readability. To solve this challenge, we aim to implement collapsible nodes or interactive filtering options. This empowers users to explore different levels of the hierarchy selectively based on their interests.

Moving forward, our plan for the final interface involves refining the visualizations through user feedback and incorporating interactive elements to enhance usability and facilitate deeper data exploration. We will continue to iterate on our designs based on the balance between providing a comprehensive overview and enabling users to interactively explore specific aspects of the data.