

Comparative Analysis of University Spending and Enrollment

Marc Leonti*

Raed Binsaeed†

Nick Hu‡

Northeastern University

ABSTRACT

This paper will discuss a web-based data visualization tool we have developed that can help school administrators compare and explore the spending patterns of post-secondary institutions in an attempt to better understand their impacts on enrollment. Our work provides visualizations that compare time series data related to enrollments from 2004 to 2018, along with the amount each school spent in eight major expense categories, for a basket of up to ten schools. These direct comparisons can help universities develop theories as to which spending patterns lead to higher enrollment, and which lead to lower enrollment. Visit the GitHub for this project by clicking on [this link](https://neu-ds-4200-f20.github.io/project-group15-university_enrollment-spending/) or entering the following address in any browser: https://neu-ds-4200-f20.github.io/project-group15-university_enrollment-spending/

1 INTRODUCTION

Public education in Massachusetts is suffering from decreased enrollment and increasing annual budget deficits. Many capital investments have been made in recent years to attract students, but enrollment continues to decline while spending has increased.

This issue is of great importance, not only to students who need a lower-cost alternative to private school tuition but also to higher education unions, and the local communities that will suffer if the state school system collapses.

We have developed a tool to visualize changes in financial data and enrollment over time for post-secondary institutions. By comparing enrollment against the eight major expense categories between similar schools we believe trends will be uncovered that correlate with higher enrollment and lead to balanced annual budgets.

2 RELATED WORK

[1] This report has many visualizations that measure metrics related to our topic such as revenue sources of universities and briefly touches on their spending. This can be used to inspire some of the visualizations we are going to make.

[2] This source provides guidance on a readily available analysis tool; we plan to expand on it.

[3] The paper shows trends related to revenues and proves how government and state spending on education is declining. We could explore how university enrollment is impacted.

[4] This publication offers substantial information on overcoming spreadsheet errors through visualizations, the concepts of which will be important in this project.

[5] This source is the starting point of our project as we will be expanding on the work of IPEDS. The information has been collected by them as raw data and we will leverage it to create interactive visualizations.

[6] This paper demonstrates the challenges of developing visualizations used to analyze financial statement data of public companies. Some concepts can be applied to our project.

[7] This paper describes the proper financial metrics that should be used to determine financial strength of schools. This might be very beneficial for our comparative analysis.

[8] This article shows the metrics analysts use to formulate opinions concerning higher education financial stability. It also offers a perspective on the evolving business environment that affects financial stability.

[9] The paper is not directly related to post-secondary education, but showcases important methodologies applicable to all financial visualizations, especially when it is related to trend analysis.

[10] This article explores the downfall of a failed merger, a helpful source to justify and handle some of the outliers we may find in our data, which will be helpful to know that when exploring the dataset.

3 PARTNER

Salem State University is doing particularly poorly when compared to its eight sister schools. Enrollment is declining faster, and as a result, revenue is rapidly declining. The President of Salem State University has formed The Sustainable Path Forward Task Force, to make proposals for a strategic plan by the end of this year to propose ways to boost enrollment and reduce expenses. The Task Force currently does not have any comparative tools to identify differences between schools that may explain what the major causes of the increasing deficit are. An interactive visualization of the data would allow several hypotheses to be explored and support further exploration and discovery. We have partnered with Dr. Kurt von Seekamm, an assistant professor of economics at Salem State University and a member of the Task Force, who has provided guidance and technical support. After conducting an initial interview, the team had the following thoughts:

Raed: "The interview went extremely well, and we learned a lot from it. The most important takeaway was narrowing our focus. Initially, our idea was over-complicated, but the interviewee helped direct us towards very specific achievable goals. One thing that surprises me is that Salem State University is in a critical situation, needing to resolve a \$15 million deficit, and our work will be used to help drive those savings. The interview motivated me to explore which other universities might need a similar tool and how our project can also cater to their needs in the future. From the meeting I understood that such a tool is not a luxury, but a necessity."

Marc: "The interview we had with Dr. von Seekamm was very informative and helped us understand the goals his task force is trying to accomplish, the challenges they face, and how a data visualization interface will make their analysis much easier. They plan to examine comparable universities in Massachusetts and the surrounding region, comparing the changes in spending and enrollment at Salem State University to the spending and enrollment of a basket of comparable schools. The task force will use this information to guide further research to determine if there is causation between certain types of capital investments and increasing enrollment. The challenge they face is there is no way to easily explore the data

*e-mail:leonti.m@northeastern.edu

†e-mail:binsaeed.r@northeastern.edu

‡e-mail:hu.r@northeastern.edu

without a visualization tool that can summarize the data. Our project will allow the task force to select a number of data fields that can be displayed in direct comparison to the same data for Salem State University. Dr. von Seekamm advised us the goal of this project is to enable discovery and exploration, while determining causation is the responsibility of the task force and beyond the scope of our project. For this reason, design simplicity and ease of use were of primary importance. Dr. von Seekamm explained that if this tool is effective it should be easily adapted to be used by any school to drive revenue and spending policy changes to help increase enrollments.

Nick: "The interview has provided us plenty of information about what we should be working on either in data visualizations or the set of requirements we must accomplish. Specifically, we are being asked to design a data visualization that can help them identify and compare annual spending patterns and enrollment rates to determine where finances can be saved or better utilized. The interview motivated me to design a useful data visualization to help solve a real-world problem."

4 TASK ANALYSIS

Our focus has been on a on high level of discovery as well as deriving the insights from our data sets. The main purpose of this project has been to design a data visualization in order to help school officials comparatively examine the spending patterns of the similar colleges (Table 1). A data visualization that explores the data to help discover patterns is crucial for our clients.

The primary consumer of our visualization is Salem State University and its decision makers. Additionally, other universities that are looking to achieve similar goals can be potential users as well with very few modifications.

5 DATA

IPEDS (Integrated Post-secondary Education Data System) is a large, publicly available data set maintained by the U.S. Department of Education. It contains vital information on nearly all colleges and universities on admissions, enrollment, finance activities, human resources, and more.

Using this data, visualizations have been created that will be used by the Task Force to compare Salem State University to its sister schools, each of the University of Massachusetts campuses, and other similar schools in the region. The data is clean, available in CSV format, and includes all the necessary documentation needed for comprehensive analysis. The only modifications we have made to the original data was to include meaningful column headers, as the original headers are numeric codes, and created columns that calculate spending per student and percentages of each expense category as a portion of the school's total expenses. The data for over 6,000 schools is freely downloadable from nces.ed.gov/ipeds as CSV files, and a simple Python program automates this process.

6 EXECUTION & DESIGN PROCESS

Our plan began with an exploratory interview with Dr. Kurt von Seekamm for information gathering. We sought to understand the specific questions that they would like to answer so that we create the visualizations needed to answer such questions. We created initial hand-drawn sketches of the visualizations and then recreated them using Tableau, then reported back to Dr. von Seekamm to see if there were any additional requirements. As for technologies, we are applying what we learned in HTML, CSS, JavaScript, and D3 to create a website that can be used by Dr. von Seekamm and his colleagues on the Task Force.

The visualizations use brushing and linking to show the interaction between the data sets and the visualizations. The user will be able to discover which types of spending each university has made and whether they have affected enrollment rates. The data visualizations will enumerate a list of predetermined colleges for SSU to explore the spending patterns and financial distributions. Starting with an overview of enrollment metrics, the user can compare changes in enrollment for each school. Additional details are available on demand for each data point. Moving into an overview of spending, the user can quickly see which categories of spending have the most variance and quickly identify any outliers. Once the user has identified an expense category of interest, the user can review each school's annual spending in this category over the entire time frame. The list of schools can be filtered, and more details are again available on demand for each data point.

Task ID #	Domain Task	Analytic Task (low-level)	Search Task (mid-level)	Analyze Task (high-level)
1	\$15 Million in cost savings or additional revenue for Salem State University. "Where can we cut costs without significantly affecting the student experience?"	Identify	Locate	Consume & Produce
2	Compare Salem State to other public universities in MA to find areas they significantly over or under spending in.	Compare	Explore	Consume & Produce
3	Identify areas of spending that increase enrollment to determine where to spend money to increase enrollment rates and revenue."	Identify	Locate	Consume & Produce
4	Convince decision-makers at the University to change spending patterns based on the findings.	Identify	Look up	Consume & Produce
5	Understanding the trends within the local public universities to see how Salem State's annual revenue and spending has changed compared to other schools	Compare	Explore	Consume & Produce

Table 1: Domain task and abstract tasks.

Our partner has identified eight schools which are similar to Salem State University and would be appropriate for comparison based on region (Massachusetts), sector (four-year public institutions), and size (more than 1,000 but less than 15,000 students).

- Bridgewater State University
- Fitchburg State University
- Framingham State University
- Westfield State University
- Worcester State University
- University of Massachusetts: Boston
- University of Massachusetts: Dartmouth
- University of Massachusetts: Lowell

The data for the three University of Massachusetts schools have been averaged and is represented as one entity (gray). The average of the MA State Schools, excluding SSU, is also displayed (pink).

The task force expressed they would like to be able to see the rate of growth of enrollments expressed as the number of Full-Time Equivalent Students, distribution of spending as a percentage of total spending, and time-series information showing the year over year differences in spending calculated per Full-Time Equivalent Student.

7 VISUALIZATION DESIGN

This has been accomplished with the use of a line graph to show time-series enrollment data, a violin plot to show the distribution of spending in eight major categories, and individual line graphs to show the time-series detail of spending in each of those eight categories. In each of the line graphs, additional details are displayed when hovering over any point with the pointer.

The first line graph shows the number of Full Time Equivalent (FTE) Students reported by each school, which is a measurement equivalent to one student enrolled full time for one academic year. This is the best way to measure enrollments when making comparisons between universities, as it captures full-time, part-time, and partial-year enrollments by calculating, based on the aggregate total number of credit hours, how many students would have attended if they were all enrolled full-time for an entire academic year.

In this line graph, hovering over any data point will reveal more data about that specific point. The user can click any colored dot in the legend below to filter by school. Continuing to click the colored dots will add or remove schools from all of the line graphs. The user can click anywhere in the violin plot area to reset the graph to show all schools.

Below the legend, on the left side of the screen, the violin plots show the distribution of the data from all the schools in each of the eight major expense categories. Each colored dot represents one school's spending in that category for a particular year, as a percent of their total budget. A gray violin body gives an alternate representation of the distribution, read similar to a sideways histogram. The height of each violin body represents the range of the values. The width of each violin body represents the frequency of the values, with many values in a narrow range resulting in a wide violin body. The violin plot, addresses the end-user's most important goals by allowing the user to explore the data from the comparison schools in several spending categories by generating violin plots that show the distribution of the values. This helps the user visualize areas of spending that are different to assess for statistical significance and possibly provide insight into where spending should be redirected. Any group of data points outside of the main body of data should be investigated for correlation with changes in enrollment.

On the right side of the screen, line graphs show each school's spending trends per category over time. Each graph corresponds with one violin from the violin plots, with the top line graph representing the same data as the far left violin, and the bottom line graph representing the same data as the far right violin. The data displayed are the dollar amounts spent annually per Full-Time Equivalent Student. This is a powerful metric allowing for direct comparisons regardless of the number of students or total budget size.

The line graphs are interactive based on the user's interactions with any violin plot. Again, details are available on demand by hovering over any plot, and brushing and linking is enabled between these line graphs and the legend, to filter schools, and the violin plots, to highlight specific data points. The line graphs show time-series data for any institution's data that contributed to the generation of the violin. The institutions displayed will correlate with the data points represented in that portion of the violin. The data displayed shows the dollar amounts spent annually per Full-Time Equivalent Student. This is a powerful metric giving the user the ability to make direct comparisons regardless of total budget size. The data points highlighted in the violin plots will also be highlighted in the line graphs to identify trends over time.

These capabilities will allow the end user to explore the data set, discovering useful insights, and directing their research and in-depth statistical analysis. Depending on those findings it can be possible to identify potential cost savings, and convince decision-makers in affecting financial policy changes.

8 DISCUSSION

We can see with these visualizations that most of the schools in this analysis have reported positive growth over time, but Salem State University has shown steady negative growth since 2010. Bridgewater State has the most FTE Students of the MA State schools, and the UMass schools have the highest average enrollment among all MA public schools. Both schools have reported steady growth over the timeframe.

Since these schools seem to be getting it right, we would look for similar spending patterns between these schools to find a correlation between certain expenses and enrollment changes. In the following line graphs, interesting results were found that directed the task force to take a closer look at several expense categories, especially the "Other" category. Our visualizations found about \$30 million in expenses that have no explanation, and may be significant to balancing the budget.

9 CONCLUSION

These capabilities will allow the end user to explore the dataset and hopefully discover useful insights. While an in-depth analysis will be required to determine if the findings are statistically significant, and is beyond the scope of this project, we hope this tool will help users direct the focus of their research. Depending on those findings it can be possible to identify which expenses help boost enrollment, which expenses should be reduced, and ultimately convince decision-makers in affecting financial policy changes.

10 ACKNOWLEDGMENTS

I would like to thank everyone who helped make this possible: Dr. Kurt von Seekamm from Salem State University, for your unwavering guidance and comical support; Kyle Langford, for your technical advice and honest critiques; Professor Cody Dunne, for putting things into perspective; And especially our loved ones for their patience and encouragement.

REFERENCES

- [1] American Academy of Arts & Sciences. Public research universities: Understanding the financial model. Web, 2016. <https://shar.es/abBDmX>.
- [2] American Council of Trustees and Alumni. How colleges spend money. Web, 2020. <https://www.howcollegesspendmoney.com/builder>.
- [3] Federal Reserve Bank of Cleveland. Trends in revenues at us colleges and universities, 1987-2013. Web, 2017. <https://www.clevelandfed.org/en/newsroom-and-events/publications/economic-commentary/2017-economic-commentaries/ec-201705-trends-in-revenues-at-us-colleges-and-universities-1987-2013.aspx>.
- [4] Y. Horry. Financial information description language and visualization/analysis tools. *Computer Languages, Systems, and Structures*, 50:31–52, 2017. doi: 10.1016/j.cl.2017.05.005
- [5] Institute of Education Statistics. Integrated postsecondary education data system. Web. <https://nces.ed.gov/ipeds/use-the-data>.
- [6] A. Khalil, A. Reza, P. A. Junaedi, and B. Kanigoro. Data visualization application for analyzing public company financial statement. *Procedia Computer Science*, 59:45–53, 2015. doi: 10.1016/j.procs.2015.07.336
- [7] M. K. Montanaro. *The Influence of Financial Performance on Higher Education Academic Quality*. PhD thesis, St. John Fisher College, Aug 2013. https://fisherpub.sjfc.edu/education_etd/117.
- [8] J. J. Selingo. Despite strong economy, worrying financial signs for higher education. Web, Aug 2018. <https://www.washingtonpost.com/news/grade-point/wp/2018/08/03/despite-strong-economy-worrying-financial-signs-for-higher-education/>.
- [9] E. Sorenso and R. Brath. Financial visualization case study: Correlating financial timeseries and discrete events to support investment decisions. *2013 17th International Conference on Information Visualisation*, pp. 232–238, 2013. doi: 10.1109/IV.2013.31
- [10] K. Woodhouse. Anatomy of a failed merger. Web, 2015. <https://www.insidehighered.com/news/2015/08/05/college-merger-negotiations-are-long-and-complicated>.