Design and Implementation of a Personalized Data Visualization System for Higher Education

Olta Llaha

Logos University College, Faculty of Applied Sciences, Tirana, Albania, olta.llaha@kulogos.edu.al

Esmeralda Kadena

Logos University College, Faculty of Applied Sciences, Tirana, Albania, esmeralda.kadena@kulogos.edu.al

Obuda University, Bánki Donát Faculty of Mechanical and Safety Engineering, Budapest, Hungary, kadena.esmeralda@uni-obuda.hu

Eleni Papapano

Logos University College, Faculty of Applied Sciences, Tirana, Albania, eleni.papapano@kulogos.edu.al

Abstract: This paper outlines the design of a data visualization solution, intended for use in higher education organizations. The data visualization tool allows for effective interaction with complex data through a customized manner that ensures that all representations of data are understandable. The pedagogues, administration, and students can utilize the system in their operations in order to make well-informed decisions.

The system will ensure its relevancy by integrating the feedback of the users and making visualization adaptable based on different roles. This helps in making the system sustainable. Among the important parts of the solution include the use of dynamic dashboards, which provide up-to-date information on different activities, prediction of different patterns and data processing capabilities.

Such solutions should help the users in making decisions based on relevant data.

Keywords: Data Visuazilation, Higher Education, Decision Making

1 Introduction

Assisting people in processing large amounts of information is the primary purpose of data visualization, as human cognitive capacities are limited and data volumes are large [1]. To simplify visualization and help people make effective decisions, data visualization should be produced in a way that allows domain experts to clearly see all problems and solutions [2]. Visualization techniques, in general, are excellent tools for helping people understand information and can facilitate visual thinking by making complex relationships more easily understood [3]. To effectively forecast academic achievement in higher education, providing a clear definition of what constitutes academic success is crucial. In a case study conducted by [4], the utilization of learning analytics dashboard (LAD) interventions by undergraduate students was examined at a public university located in the mid-Atlantic region of the United States. The study found that student understanding is linked to data relevance, accuracy, and context. Our study has led to the conclusion that changes made to the curriculum have affected students' understanding, and this has been facilitated through data visualization. Other researchers conducted a study on HEI data management practices and identified domain policy-makers' knowledge preferences [5]. Choosing appropriate visualization, interaction, and analytical techniques depends on the dimensions of the data. For our study, we focused on numerical data related to student assessments, graduation, and enrollment in different academic years. Visualization techniques that leverage visual perception are useful for converting tracks into visual information [6]. We employed a chart visualization technique, as also used in the study of Zotov et al. (2021), to graphically present information about student evaluations in various subjects. Data visualization can facilitate better analysis of student assessments and lead to informed decisions to improve their performance, as well as to effectively inform students of their performance in different subjects and academic years through up-to-date assessments, comparisons with the class average, and projections of future outcomes. This could potentially lead to better student behavior, based on observations of the graphs.

2 The Visualization Model for Higher Education Data

Information visualization enhances the understanding of information. It addresses the challenge of representing diverse data types for easy transmission and interpretation. The primary objective of any visualization is to assist users in analyzing vast amounts of information. Data can take various forms (e.g., one-dimensional, two-dimensional, multidimensional, hierarchical, and graphical), while visualization can manifest in diverse formats using a range of techniques (such as standard 2D/3D displays, geometrically transformed displays, icon-based displays, dense pixel displays, and stacked displays).

In the initial case of the study, we possess data concerning student evaluations across various subjects and their graduation years, with a specific emphasis on exploring the correlation between these variables. In the subsequent case of the study, data is organized based on questionnaires concerning the influence of data visualization technology on decision-making processes.

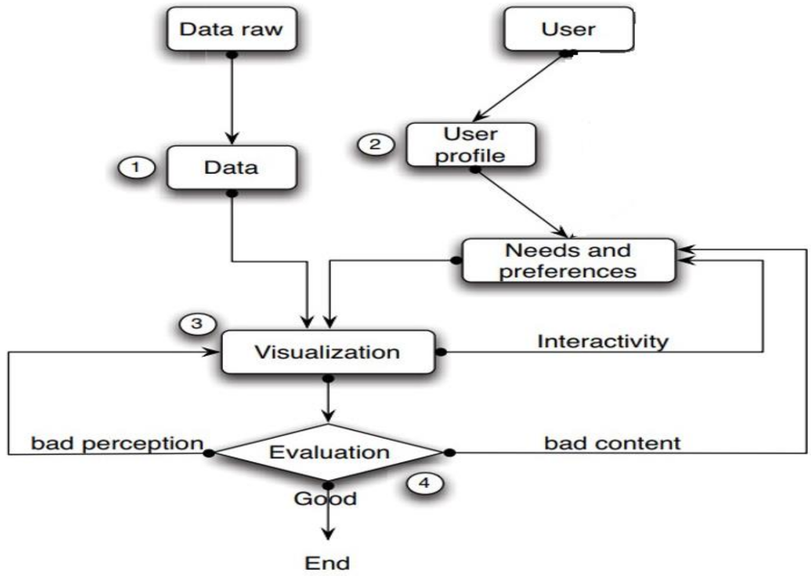

Figure 1

The Visualization Model for Higher Education Data

Figure 1 represents the visualization generation process in four stages that include: 1) Data Preparation based on data collected from raw data sources; 2) User profiling by establishing the user's need and preference; 3) Visualization Generation according to the need and preference of the user; 4) Evaluation of Visualizations to enhance them.

Figure 1 shows the data-based visualization generation and improvement process according to the needs of users. Let us now take a closer look at each part of the figure.

• Raw Data Preparation:

The process consists of cleaning, transforming, and structuring the data in such a way that it becomes useful for the purposes of visualization.

• From Raw Data to Usable Data (Stage 1):

Raw data implies data that has not been processed yet. The data becomes subject to processing in the data preparation block; thus, it becomes useful for visualization. It is the first stage when raw data is transformed into useful information.

• User and User Profile (Step 2):

Users' profiles contain some explicit data such as preferences, goals, or interactions with other users. Such profiles become an input to the "Needs and Preferences" block, where all data are combined and used for personalizing visualization in accordance with user's needs.

• Visualization Generation (Step 3):

The system will generate interactive visualizations after processing the data and understanding what is required by the user. The purpose of such visualizations is to provide data clarity while satisfying the preferences of users.

• Feedback Loops (Step 4):

As soon as visualizations become operational, feedback loops will be able to fine-tune the content and visual clarity.

When bad perception occurs: If the user is dissatisfied with the visualization and finds it difficult to interpret, the feedback loop will correct the visualization for better visual clarity.

When bad content arises: When the content fails to meet the user's requirements, the feedback loop will reassess the preferences and data of the user to fine-tune the content.

If the visualization meets expectations, then the process ends in success.

• Interactivity:

The ability for users to interact with visualizations enables them to engage with data more deeply. Interaction plays an integral role in continuous improvement of visualizations through feedback.

• Continuous Improvement:

Continuous improvement is at the heart of the process in question through feedback cycles. Visualizations will be continually improved, especially if the output does not meet the requirements of the users, in terms of its content and presentation.

To sum up, the picture above provides a framework that illustrates how data visualization should be undertaken, taking into account continuous improvement, data transformation, customization, among other considerations.

3 Data Visualizer Application

The Data Visualizer is a data visualization software allowing users to import data in different formats including .csv, .xlsx, and then visualize data in the form of interactive graphics like line chart, histogram, pie chart, and many others. This software provides customization of graphics and a user-friendly interface suitable for everyone from beginner to expert.

This software acts as a tool for data visualization and file conversion. It was built using the Streamlit framework and allows users to upload CSV files for visualization, apply dynamic filters on them, and customize plots. Moreover, this software converts XLSX files to CSV files.

The key elements of this system are the possibility to represent CSV data in interactive graphs and the dynamic filtering of data through the sidebar. Furthermore, this system has the capability of converting XLSX data into CSV files, as well as modifying the datasets and downloading them in their new form. One of the capabilities of the system is the creation of different kinds of graphs, including bar, line, scatter, and pie graphs. Additionally, users have the possibility to adjust the visualization in terms of its colors, scale, labeling, and other parameters. The system also allows exporting the visualized graphs into PNG files, in addition to offering an intuitive interface.

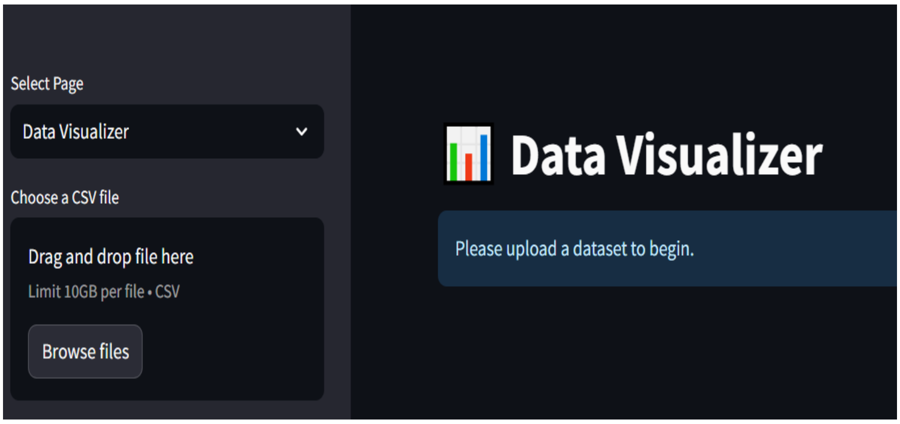

Figure 1
Home Page of the Data Visualizer Application

The important libraries that have been implemented in the development process of this project include Streamlit, which creates a web interface, and Pandas, which manipulates and analyzes the data. Chardet has been used to determine the file encoding, while Plotly has been used to create dynamic charts. Finally, Xlsx2csv has been used to convert files in XLSX format to CSV format.

Plot Type	Data Type	Description
Histogram	Numeric	Displays the distribution of numeric data by grouping into bins.
Box Plot	Numeric	Shows the distribution of data based on five summary statistics(min, Q1, median, Q3, max).
Boxen Plot	Numeric	A variation of a box plot, showing more details about the tails of thedistribution.
Heatmap	Numeric	Displays the data values using color scales in a matrix format.
Bubble Chart	Numeric	A scatter plot with an additional size variable to represent anotherdimension.
Density Plot	Numeric	Shows the distribution of numeric data using a density contour plot.
Violin Plot	Numeric	Combines a box plot with a kernel density estimate, showing datadistribution.
3D Scatter Plot	Numeric	A scatter plot in three dimensions (X, Y, Z axes).
Bar Plot	Numeric	Displays data using bars, where the length of each bar isproportional to the value.
Bar Chart	Categorical	Displays data using grouped bars for each category.
Pie Chart	Categorical	Shows parts of a whole as slices of a pie.
Sunburst	Categorical	Represents hierarchical categories in a circular diagram.
Dot Plot	Categorical	Plots dots along an axis for each category to show frequency or value.
Heatmap	Categorical	Displays categorical relationships in a matrix using color.
Treemap	Categorical	Displays hierarchical data as nested rectangles.

Table 1

Supported plot types

The table represents an overview of plot types utilized in data visualization, data type, and functions. Plots represent the way visual information is created; hence, there are two main types of plots – for numerical and categorical data. The first column contains plot names, the second one specifies what type of data they work with, either numerical or categorical, and the third one contains descriptions of those plots. As can be seen from the table, almost all plots work with numerical data showing distribution, comparisons, and relations between data. Thus, the histogram is used to show the distribution of data in bins (intervals of values), while box plots show the distribution of data statistically (the minimum value, lower and upper quartiles, and maximum value). Box plots give an opportunity to get more detailed information about the tails of the distribution. In turn, the bar chart is utilized for categorical data to make comparisons. Various plots can be used depending on what kind of data is analyzed, which allows analyzing distributions, making comparisons, and revealing connections between data. Visualizing data is an important step that makes its analysis easier.

4 Utilization of the Application at Logos University College

An Executive Information System (EIS) for higher education is one of the critical means that enables university and college managers to gather and analyze data about their organization's performance in terms of academic excellence, financial situation, and efficient work of the university. EIS gives managers an overall picture regarding different functions within a university or a college such as student enrollment, academic achievements, financial status, and employees' job performances.

For instance, in order to improve their academic performance as well as their management of operations, the EIS known as Data Visualizer for Strategic Management is used in "Logos" University College. The top managers of the university require quick and simple ways of getting information on their students, faculty members, and financial management of the institution.

Important EIS functions in Logos University College:

1. Student Enrollment Monitoring: Visualization: An interactive display illustrates the number of new enrollments during a year and their percentage among the overall number of enrolled students. This data can be filtered according to the specific faculty or study program.

.

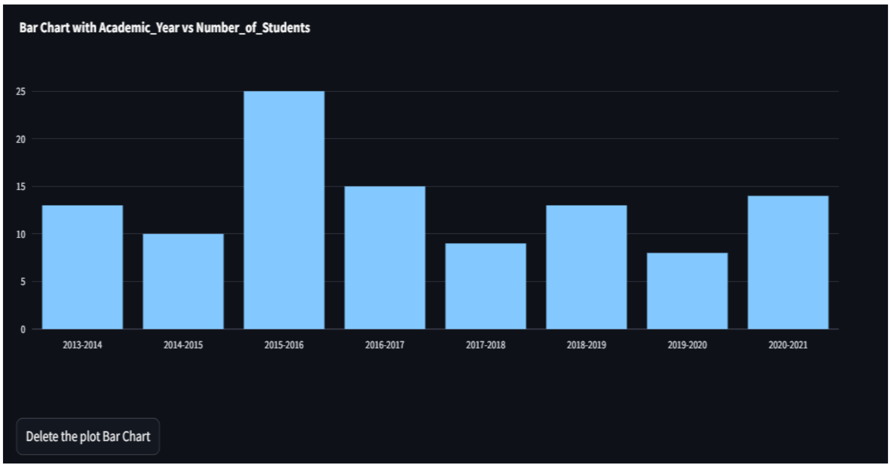

Figure 2
Student Enrollment in the Applied Informatics Study Program

The bar graph shown in figure 2 illustrates an unstable dynamics for the number of students over the years due to considerable changes and lack of long-term linearity. The first decline is seen during 2013-2014 to 2014-2015 years; then, there occurs an abrupt increase at 2015-2016 year. The latter may be associated with different reasons such as growing interest in the program, favorable policies for admission, or other. Then, the chart displays a constant decrease up to 2017-2018 year; in this case, we may say about decreasing demands and impact of some external factors (e.g., migration, competition, or others). The next years show only a small growth, but unstable one because it decreases at 2019-2020 year and rises at 2020-2021 year. Therefore, the analysis of dynamics demonstrates that the admission process is sensitive to external factors.

2. Student Academic Performance: Visualization: Graphs depicting grade distribution in each faculty and department will be available to aid in identifying areas that require further attention. Performance Comparison: Student performance will be compared to previous performances to identify academic trends.

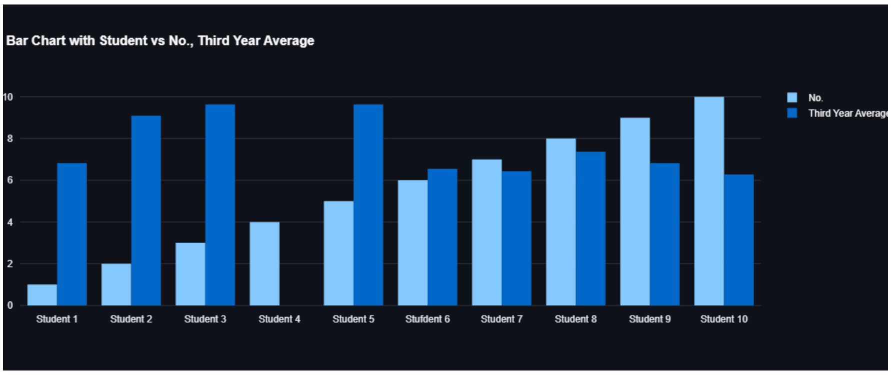

Figure 3
Performance of Applied Informatics Students in Third-Year Courses

Figure 3 shows the average marks for the third year of students, who are assigned numerical codes for anonymity purposes. Based on the analysis performed in terms of the graph above, it can be stated that the students have relatively stable averages and are positioned in a narrow range of performance, which indicates a rather good balance in academic achievements. At the same time, there are some differences in performance: the student with code number 3 demonstrates the best performance in the group, having the highest mark average, while the student with code number 10 has the lowest performance with the worst mark average compared to others.

3. Budget Planning: Visualisation: There will be a visual representation of the sources of funds like tuition fees, government and donor contributions. One will also be able to see where money is spent in terms of salaries, infrastructure, and research. Financial Projections: The system will enable one to conduct "what if" analysis like seeing what the effect would be if one raises tuition fees or invests heavily in infrastructure.
4. Performance Analysis: Visualisation: The system has a dashboard where one sees publication rates of academics, number of research projects undertaken by faculty members and staff, and fund allocation for research activities. Improvement Opportunities: In case of a faculty member not publishing much research work, then the system signals that a mentorship or training programme is required.
5. Placement of Graduates: Visualisation: The system uses a graph to show the percentage of students getting employed within six months after completing their studies.

Conclusions

In the course of the research presented in the article, a data visualization system was developed on the basis of the concept of an Executive Information System (EIS). Namely, the Data Visualizer system was utilized for designing this solution.

The data visualization system introduced in the research showed the well-designed and efficient structure of the platform since various data sources were integrated and represented using visualization tools. In particular, visual panels (dashboards) allowed one to visualize grades' averages, distributions, as well as comparison between students or even groups of students.

The visualizations provided in the article clearly demonstrated how the system functions in practice helping users to identify certain trends, average results, and cases requiring special attention. Thus, it helps improve the decision-making process for both educational and administrative personnel.

Among the major strengths of the system, one can mention high level of flexibility and versatility of the solutions provided.

Concluding, one should say that the utilization of the system discussed in the article proves how important the application of data visualization techniques can be. As the further direction for research, the incorporation of analytical tools, such as predictive modeling or artificial intelligence algorithms, is suggested.

References

[1] Mohd, Maseri & Abdullah, Embong & Mohamad Zain, Jasni. (2010). A Framework of Dashboard System for Higher Education Using Graph-Based Visualization Technique. 87. 55- 69. 10.1007/978-3-642-14292-5_7.

[2] Donohoe, David & Costello, Eamon. (2020). Data Visualisation Literacy in Higher Education: An Exploratory Study of Understanding of a Learning Dashboard Tool. International Journal of Emerging Technologies in Learning (iJET). 15.115.

[3] Shabdin, Nur & Yaacob, Suraya & Sjarif, N.N.A.. (2020). Relationship Types in Visual Analytics. 1-6. 10.1145/3397125.3397127.

[4] Klein, C., Lester, J., Nguyen, T., Justen, A., Rangwala, H., & Johri, A. (2019). Student Sensemaking of Learning Analytics Dashboard Interventions in Higher Education. Journal of Educational Technology Systems, 48(1), 130–154. https://doi.org/10.1177/0047239519859854

[5] Akanmu, Semiu & Jamaludin, Zulikha. (2016). Designing information visualization for higher education institutions: a pre-design study. Journal of Information and Communication Technology. 15. 145 - 163. 10.32890/jict2016.15.1.7.

[6] Zotov, Vladimir & Ibrahim, Iman & Petunina, Irina & Lazareva, Yuliya. (2021). Engagement of Students in Data Visualization for the Purpose of E-Learning Improvement. International Journal of Emerging Technologies in Learning (iJET). 16. 46. 10.3991/ijet.v16i02.18745.