International University
School of Computer Science and Engineering

VIETNAM NATIONAL UNIVERSITY – HO CHI MINH CITY
INTERNATIONAL UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DATA SCIENCE & DATA VISUALIZATION

PROJECT REPORT

Course by Dr. Tran Thanh Tung

Topic: Analyzing the 2020 High School Graduation Exam Scores in Ho Chi Minh City through Web-Based Interactive Visualization

TEAM MEMBERS

STT	Name	ID	Email
1	Phạm Đức Đạt	ITITIU20184	ducdatit2002@gmail.com
2	Phạm Lê Thanh Nhàn	ITDSIU20073	phamlethanhnhan0707@gmail.com
3	Huỳnh Lam Đạt	ITITIU20364	huynh.lam.dat.77@gmail.com
4	Lê Phước Đẫu	ITITIU20185	lephuocdau298@gmail.com

GitHub repository: https://github.com/ducdatit2002/datavisualization-project
GitHub deploy product: https://ducdatit2002.github.io/datavisualization-project/

Table of Contents

TABLES OF FIGURES

I. INTRODUCTION

1. Abstract

The high school exam is a crucial component of the Vietnamese education system as it assesses students' knowledge and learning abilities before they graduate from high school. This exam is regulated by the Ministry of Education and Training and includes subjects like Literature, Mathematics, and Foreign Language. The scores obtained in this exam are used by universities and colleges for admissions and ranking purposes. [1]

The high school exam includes main subjects such as Literature, Mathematics, Foreign Language (usually English) and a few other subjects depending on the curriculum of each grade and the regulations of the Ministry of Education and Training [2]. The exam is divided into written and multiple-choice sections. Universities and colleges use this exam score to admit and rank students. High scores on the high school exam can create opportunities for students to access high-quality university education and attractive study programs.

Figure 1 Photos of the 2020 National High School Exam

In this project, the aim is to visualize the trend of the 2020 High School Graduation Exam Scores in Ho Chi Minh City using web-based interactive visualization tools. The visualization will help parents, students, and educators understand the score distribution in each subject and identify notable achievements and patterns in student performance. The data for visualization is collected from the official website of the Department of Education and Training of Ho Chi Minh City and is pre-processed to ensure data validity.

Visual designs include bar graphs that show the number of scores across subjects, and graphs that compare scores by demographics using line graphs that show the average score per subject of the total number of test takers. born by month of birth and dashboard showing each contestant's score and performance.

These designs provide features such as line charts, bar charts, interactive tooltips, and search functionality to enhance user experience and support data analysis. The project also includes must-have features such as data security, user-friendly interface, and optional features such as interface customization, multilingual support, and sharing and saving capabilities. reserve. Overall, the project aims to provide stakeholders with a comprehensive and accessible understanding of high school test scores to support decision-making and improve educational outcomes.

2. Overview

Based on the outcomes of the 2020 High School Graduation Exam, we present an overview of three learning performance assessment system designs in this report. Every design has significance. unique and fulfills a number of functions, including helping parents and students organize their academic schedules and giving information to colleges and universities, educational researchers, and the general public. Design 1 forecasts trends in college enrollment while concentrating on supplying parents and students with information. By investigating the connection between academic achievement and birth month, Design 2 offers fresh insights for school counseling and educational research. And Design 3 presents a comparative overview of the number of students with above-average scores in six academic subjects, categorized by their birth month periods. A customized dashboard called Design 4 assists students in evaluating their own performance and informs decisions about what to learn. And Each design helps to shape educational strategies for the future in addition to providing a thorough picture of current learning performance. When these three systems are used together, they help build a strong foundation that will advance educational advancement and support students' holistic growth.

3. Goals

- Construct a robust, interactive dashboard to display difficult-to-find insights from raw data.
- Use interactive exercises and charts in theory classes.
- Go through the refactoring and project management processes.
- Examine whether additional features can be added to the foundational code to make the dashboard more succinct.

4. Techniques and tool used:

- Web Development Stack: HTML/ CSS/ JavaScript (using D3.js)
- Demo UI Dashboard: Figma
- Code version management: GitHub
- Collecting and cleaning data: Python Jupiter Notebook by Google Colab

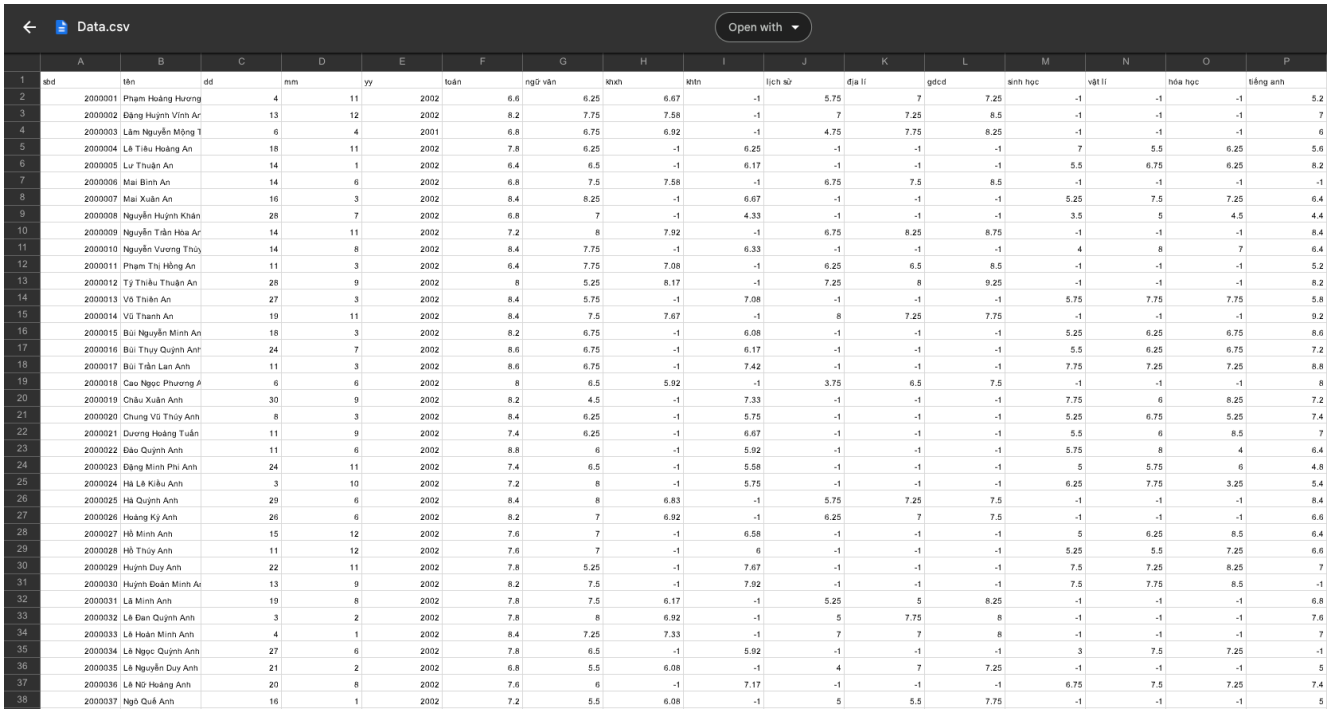
5. Data source collection:

This project is made possible by the dataset uploaded on diemthi.hcm.edu.vn. The Raw data includes students' information and the examination’s results and is saved to a CSV file.

Data pre-processing: Strict text processing methods were used to remove unnecessary spaces, duplicate characters, and line breaks from the raw text data that was taken from the page source. A comprehensive de-duplication process was implemented to eliminate superfluous HTML/XML tags and blank lines that could potentially obscure the data's meaning. With Unicode encoding, special consideration was given to handling unique characters, guaranteeing the data's consistency and compatibility across various platforms and systems.

Processing-data: During data processing, we ensured the validity of the data by checking for completeness and correcting any invalid values. Duplicate entries were identified and resolved to guarantee dataset accuracy. Birthdates were reformatted into a standardized Date Time format, with separate dd, mm, yy columns for day, month, and year, respectively. For subjects not attempted by students, the corresponding score cells were marked with "-1" to indicate absence, facilitating clear identification during analysis.

Cleaned data: dataset was saved in a CSV format, comprising 16 columns and 74,444 rows, correlating to the attributes and entries of the exam participants. It was structured in ascending order based on the students' exam IDs. The attributes include text-based entries for student IDs and names, numeric entries for day (dd), month (mm), and year (yy) of birth, and numeric scores for two subject combinations (social and natural sciences), as well as nine individual subject scores, each rounded to one decimal place.


Data.csv — Open with

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	sbd	tên	dd	mm	yy	toán	ngữ văn	khxh	khtn	lịch sử	địa lí	gdcd	sinh học	vật lí	hóa học	tiếng anh
2	2000001	Phạm Hoàng Hương	4	11	2002	6.6	6.25	6.67	-1	5.75	7	7.25	-1	-1	-1	5.2
3	2000002	Đặng Huỳnh Vĩnh An	13	12	2002	8.2	7.75	7.58	-1	7	7.25	8.5	-1	-1	-1	7
4	2000003	Lâm Nguyễn Mộng T	6	4	2001	6.8	6.75	6.92	-1	4.75	7.75	8.25	-1	-1	-1	6
5	2000004	Lê Tiêu Hoàng An	18	11	2002	7.8	6.25	-1	6.25	-1	-1	-1	7	5.5	6.25	5.6
6	2000005	Lư Thuận An	14	1	2002	6.4	6.5	-1	6.17	-1	-1	-1	5.5	6.75	6.25	8.2
7	2000006	Mai Bình An	14	6	2002	6.8	7.5	7.58	-1	6.75	7.5	8.5	-1	-1	-1	-1
8	2000007	Mai Xuân An	16	3	2002	8.4	8.25	-1	6.67	-1	-1	-1	5.25	7.5	7.25	6.4
9	2000008	Nguyễn Huỳnh Khán	28	7	2002	6.8	7	-1	4.33	-1	-1	-1	3.5	5	4.5	4.4
10	2000009	Nguyễn Trần Hòa An	14	11	2002	7.2	8	7.92	-1	6.75	8.25	8.75	-1	-1	-1	8.4
11	2000010	Nguyễn Vương Thúy	14	8	2002	8.4	7.75	-1	6.33	-1	-1	-1	4	8	7	6.4
12	2000011	Phạm Thị Hồng An	11	3	2002	6.4	7.75	7.08	-1	6.25	6.5	8.5	-1	-1	-1	5.2
13	2000012	Tỷ Thiều Thuận An	28	9	2002	8	5.25	8.17	-1	7.25	8	9.25	-1	-1	-1	8.2
14	2000013	Võ Thiên An	27	3	2002	8.4	5.75	-1	7.08	-1	-1	-1	5.75	7.75	7.75	5.8
15	2000014	Vũ Thanh An	19	11	2002	8.4	7.5	7.67	-1	8	7.25	7.75	-1	-1	-1	9.2
16	2000015	Bùi Nguyễn Minh An	18	3	2002	8.2	6.75	-1	6.08	-1	-1	-1	5.25	6.25	6.75	8.6
17	2000016	Bùi Thụy Quỳnh Anh	24	7	2002	8.6	6.75	-1	6.17	-1	-1	-1	5.5	6.25	6.75	7.2
18	2000017	Bùi Trần Lan Anh	11	3	2002	8.6	6.75	-1	7.42	-1	-1	-1	7.75	7.25	7.25	8.8
19	2000018	Cao Ngọc Phương A	6	6	2002	8	6.5	5.92	-1	3.75	6.5	7.5	-1	-1	-1	8
20	2000019	Châu Xuân Anh	30	9	2002	8.2	4.5	-1	7.33	-1	-1	-1	7.75	6	8.25	7.2
21	2000020	Chung Vũ Thúy Anh	8	3	2002	8.4	6.25	-1	5.75	-1	-1	-1	5.25	6.75	5.25	7.4
22	2000021	Dương Hoàng Tuấn	11	9	2002	7.4	6.25	-1	6.67	-1	-1	-1	5.5	6	8.5	7
23	2000022	Đào Quỳnh Anh	11	6	2002	8.8	6	-1	5.92	-1	-1	-1	5.75	8	4	6.4
24	2000023	Đặng Minh Phi Anh	24	11	2002	7.4	6.5	-1	5.58	-1	-1	-1	5	5.75	6	4.8
25	2000024	Hà Lê Kiều Anh	3	10	2002	7.2	8	-1	5.75	-1	-1	-1	6.25	7.75	3.25	5.4
26	2000025	Hà Quỳnh Anh	29	6	2002	8.4	8	6.83	-1	5.75	7.25	7.5	-1	-1	-1	8.4
27	2000026	Hoàng Kỳ Anh	26	6	2002	8.2	7	6.92	-1	6.25	7	7.5	-1	-1	-1	6.6
28	2000027	Hồ Minh Anh	15	12	2002	7.6	7	-1	6.58	-1	-1	-1	5	6.25	8.5	6.4
29	2000028	Hồ Thúy Anh	11	12	2002	7.6	7	-1	6	-1	-1	-1	5.25	5.5	7.25	6.6
30	2000029	Huỳnh Duy Anh	22	11	2002	7.8	5.25	-1	7.67	-1	-1	-1	7.5	7.25	8.25	7
31	2000030	Huỳnh Đoàn Minh An	13	9	2002	8.2	7.5	-1	7.92	-1	-1	-1	7.5	7.75	8.5	-1
32	2000031	Lã Minh Anh	19	8	2002	7.8	7.5	6.17	-1	5.25	5	8.25	-1	-1	-1	6.8
33	2000032	Lê Đan Quỳnh Anh	3	2	2002	7.8	8	6.92	-1	5	7.75	8	-1	-1	-1	7.6
34	2000033	Lê Hoàn Minh Anh	4	1	2002	8.4	7.25	7.33	-1	7	7	8	-1	-1	-1	7
35	2000034	Lê Ngọc Quỳnh Anh	27	6	2002	7.8	6.5	-1	5.92	-1	-1	-1	3	7.5	7.25	-1
36	2000035	Lê Nguyễn Duy Anh	21	2	2002	6.8	5.5	6.08	-1	4	7	7.25	-1	-1	-1	5
37	2000036	Lê Nữ Hoàng Anh	20	8	2002	7.6	6	-1	7.17	-1	-1	-1	6.75	7.5	7.25	7.4
38	2000037	Ngô Quế Anh	16	1	2002	7.2	5.5	6.08	-1	5	5.5	7.75	-1	-1	-1	5

Figure 2 Quick view of dataset

II. TASK TIMELINE

1. *Contribution*

NAME	RESPONSIBILITY	CONTRIBUTION
Phạm Đức Đạt	Team leader, Chart designer, Write Report, Code for design 01, Code for design 03, PowerPoint	100%
Phạm Lê Thanh Nhàn	Data analyst, Code for design 01, Code for design 03	100%
Huỳnh Lam Đạt	Code for design 02	100%
Lê Phước Đẫu	Code for design 04	100%

Figure 3 Tasks contribution of members

2. Timeline and division

Week	Task
Week 1 - 2	- Conduct project kickoff meetings to establish objectives and allocate tasks. - Set up project documentation, including requirements and design processes. - Gather data from the Department of Education and Training in Ho Chi Minh City.
Week 3 - 4	- Select design prototypes based on team discussions. - Create a proposal that provides an outline and summary of the project. - Begin processing and cleaning the data, removing duplicate records, and handling missing values.
Week 5 - 6	- Design the user interface for the web application. - Select appropriate chart types to visualize the data. - Build the charts and integrate them into the user interface.
Week 7 - 8	- Display the data on the user interface using charts and tables. - Implement interactive features for users, such as subject selection, data filtering, and displaying detailed information on mouse hover.
Week 9 - 10	- Optimize the performance of the application to ensure fast loading and smooth responsiveness. - Conduct testing and bug fixing to ensure stability and reliability of the application.
Week 11 - 12	- Develop user manuals and documentation for the application. - Continue troubleshooting remaining errors or issues.
Week 13 - 14	- Perform final checks and ensure data security. - Collect and analyze feedback from users to fine-tune the app. - Improve user interface and features based on data input from users.

Figure 4 Timeline of the project

III. METHODOLOGY

1. Project structure

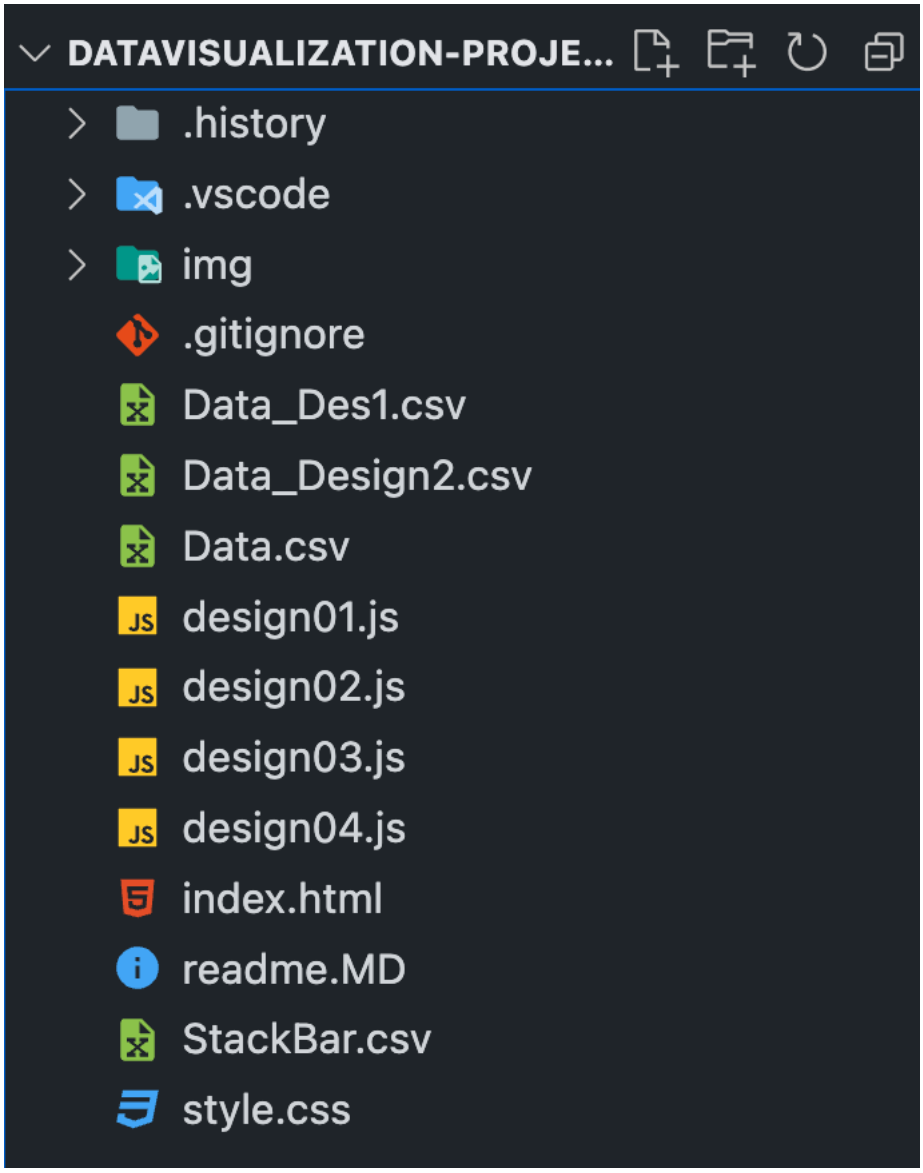


Figure 5 Project structure

The project structure for the dashboard development can be described as follows:

1. img/: This directory is intended to store image files used within the project, potentially for the dashboard interface or documentation purposes.
2. Data Files:
   Data_Des1.csv: Contains the dataset used for the first design of the dashboard, likely including specific attributes and records relevant to that particular visualization.
   Data_Design2.csv: Holds the dataset for the second dashboard design, formatted to suit the requirements of that particular analysis.
   Data.csv: This file could either be the raw data collected for the project or a master dataset that includes comprehensive records used across various aspects of the dashboard.
   Stackbar.csv: Holds the dataset for the third dashboard design, formatted to suit the requirements of that particular analysis.
3. JavaScript Files:

design01.js: The JavaScript file dedicated to the functionality and interactive elements of the first design of the dashboard.
design02.js: Similar to the first, this JavaScript file is tailored for the second design, containing logic and interactions specific to that design.
design03.js: This file contains the scripting necessary for the third dashboard design, encapsulating any unique functionality it may require.
design04.js: This file contains the scripting necessary for the fourth dashboard design, encapsulating any unique functionality it require.

4. HTML Files:
   index.html: The main entry point of the web-based dashboard. This file structures the HTML content and integrates all other resources to present the final user interface.
5. Style Files:
6. style.css: This Cascading Style Sheets file is responsible for the styling of the web content, defining the visual aesthetics like layout, colors, and fonts.
7. Documentation:
   readme.MD: A Markdown file providing an overview, instructions, or documentation for the project. It typically includes details on how to set up, use, and navigate the dashboard, along with any other relevant project information.

2. Chart usage

This section will look at the various charts that are used in the Ho Chi Minh City dashboard for the 2020 high school graduation exam scores. Every graph serves a distinct purpose and provides valuable insights into scores by channel and mark. To help with understanding, we have provided a brief description of each chart's objective, mark, and channel in the table below.

Chart name	Purpose	Channel	Mark	Description
Bar chart	The design provides a detailed visual representation of score distributions across various subjects in the 2020 High School Graduation Exam. This allows for a comparative view that can inform students, parents, and educators about performance trends.	The design is implemented as an interactive web-based dashboard, utilizing technologies such as HTML, CSS, and JavaScript, particularly leveraging the D3.js library for dynamic data visualization.	The central feature of the design is a bar chart that marks the frequency of exam scores for a selected subject. It specifically highlights the score categories with a distinct color coding for certain score ranges, enhancing the visual differentiation	The score distribution provided by the design is in the form of a histogram, with the number of examinees with each score indicated on the y-axis and the score values represented on the x-axis. Quick visual cues regarding score distribution and concentration are provided by the color-coded bars. Users can choose various subjects for a customized analysis thanks to the dashboard's interactive

			of performance levels.	features, which update in real time. Tooltips are an additional feature that enhances the user experience by offering a quick reference for particular data points. The dashboard is made both aesthetically pleasing and informative by the addition of auxiliary elements such as interactive buttons, chart titles, and axis titles.
Line chart	The design examines the correlation between students' birth months and their academic performance across various subjects in the 2020 high school graduation exam. This analysis can unveil patterns and provide insights for educational planning and student counseling.	This analysis is presented through a web-based interactive line chart, developed with the D3.js library. It forms part of a dashboard that educators, students, and researchers can access to explore the data dynamically.	The key visual elements are the lines in the chart, each representing the average scores for a subject across different birth months. These lines allow for immediate visual comparison of trends and variances in academic performance.	The chart displays a series of lines, each corresponding to a different subject, plotted against the x-axis that represents birth months and the y-axis indicating average scores. The colors are distinct for each subject, enhancing the ability to track and compare them. The inclusion of interactive features, such as tooltips and hover effects, allows users to engage with the data more deeply by revealing specific values as the cursor moves across the chart. Additionally, the legend enables users to focus on individual subjects by highlighting or dimming the lines. This line chart not only serves as an analytical tool but also as a means to communicate complex

				data in a digestible format for a broad audience.
Stacked bar chart	The purpose of this chart is to visualize the distribution of students with above-average scores in six different subjects (Biography, Physics, Chemistry, History, Geography, and Civic), based on their birth quarter (January-April, May-August, September-December). This visualization helps to identify any potential patterns or correlations between the students' birth months and their academic performance across various disciplines.	This analysis is rendered through an interactive stacked bar chart on a web-based dashboard, developed with the D3.js library. It enables stakeholders, such as educators, students, and analysts, to dynamically engage with and interpret the educational data.	Stacked bars represent the key visual elements in this chart, with each bar's segment indicating the number of students achieving above-average scores in a particular subject within a specific birth quarter.	The chart arranges stacked bars along the x-axis, which represents birth quarters, against the y-axis that quantifies the number of students. Varied colors differentiate the subjects within the bars, facilitating visual distinction and comparison. Interactive tooltips provide detailed information on demand, enriching the user's interaction with the chart. A legend on the side assists in identifying the subjects corresponding to each color. This visual tool is crafted not just for analysis but also for effective communication of complex data to a broad audience.
Dashboard	The Student Performance Comparison Dashboard is tailored to provide a personalized overview of a student's scores across different subjects taken in the 2020 High School Graduation Exam. It highlights individual strengths and weaknesses, aiding in personal assessment and targeted learning.	The dashboard is web-based, interactive, and user-driven, relying on D3.js and JavaScript to dynamically present the data. It is designed to be accessed by students, parents, teachers, and education managers for various analytical and educational purposes.	The central visual mark of this dashboard is the table of student scores, complemented by textual indicators of a student's strongest and weakest subjects. This visual presentation enables a straightforward and focused view of	Students see an easy-to-read table of their exam results after entering their identification number. The "STRENGTHS" and "WEAKNESS" categories on the dashboard are where the highest and lowest scoring subjects are immediately analyzed and displayed. A customized experience is made possible by interactive features like a submit button, and error handling

			performance across multiple subjects.	offers assistance in the event that a student's data cannot be located. This dashboard is used as a planning tool for upcoming instructional strategies in addition to being a performance tracker.

Figure 6 Chart usage table

3. ***Purpose of the design***

a) The bar chart of analyzing the scores of subjects in the 2020 High School Graduation Exam:

- Support for Students and Parents: This analysis helps students and parents better understand score trends, thereby making more effective study and review plans.
- Forecasting College Admissions Trends: Universities and colleges can use this information to forecast the expected quality and quantity of applicants, helping them better prepare for the admissions process.
- Educational Research and Development: Educational researchers use this data to analyze trends, characteristics and problems in the education system, thereby recommending solutions for improvement.
- Providing Information to the Public: Scorecards also provide important information to the public about the quality of education, helping them better understand the learning and testing situation at high schools.

b) The line chart shows the average score of the subjects in the 2020 high school graduation exam by month of birth:

- Trend analysis: To see if there is any relationship between birth month and academic performance. Some studies have shown that birth month can affect academic achievement due to differences in age and psycho-social development among students.
- Educational research: Provides data for educational researchers to better understand the influence of factors such as age and date of birth on learning outcomes.
- School and career counseling: Providing information to school counselors so they can advise, and support students based on their birth month characteristics, which can help students choose a suitable study path.
- Comparison between groups: To compare educational outcomes between groups of students born in different months and determine if there are any significant differences.

c) The stacked bar chart represents number of Students with Above-Average Scores by Birth Month Period:

- Correlation Exploration: Taking into account research indicating birth month can influence age-related development and academic success, this study aims to determine whether students' birth months and their accomplishments in a range of academic subjects are correlated.

- Academic Strategy Development: To make use of the information to plan individualized lesson plans that address the needs of students who were born in a different month as well as to strategically approach education.
- Teachers' Perspective: To provide teachers with a better understanding of how birth month can impact learning outcomes, possibly influencing classroom dynamics and teaching strategies.
- Counselor Resource: to give school counselors information that could help them steer students toward academic pathways that correspond with the birth month's trends.

d) The dashboard is designed to personalize the display of student scores in each subject and identify the subjects in which the student is strongest and weakest in the 2020 High School graduation exam:

- Personal analysis: Students can view their scores to self-assess their learning performance in each subject.
- Study guide: Based on identified strengths and weaknesses, students can make a study plan to improve skills in necessary subjects.
- Parents and teachers: Parents and teachers can use information from this dashboard to support and guide students in their studies and prepare for future exams.
- Compare performance: Students can compare their scores to an average standard or to the scores of their friends to determine their position in the group or class.
- Evaluate education planning: Education managers can use aggregated data from dashboards like this to evaluate and improve the quality of teaching and curriculum.

4. Interactive usage

The interactive nature of the score dashboard is critical to raising user inquiry and engagement. By adding interactive elements, users can engage with the charts and dynamically alter the data to gain deeper understanding.

Chart name	Interactivity	Purpose	Description
Bar chart	Subject Score Filtering	To allow users to view and compare the distribution of scores for a specific subject.	This feature includes buttons or dropdowns that correspond to various topics. The score distribution for each subject is shown in the bar chart when the user selects that subject. Understanding scoring trends within individual subjects is made easier by the interactive data visualization that offers a dynamic approach to display data for individualized analysis.
Bar chart	Hover-Over Data Reveal (Tooltip)	To display detailed information about the score distribution for individual score brackets within a selected subject.	A tooltip that displays the exact score value and the number of examinees who met that score appears when users move their cursor over any bar in the histogram.

			Through the provision of on-demand, comprehensive data without overcrowding the visual representation, this interaction improves the user experience. It makes comparing various score ranges quickly possible and aids in spotting trends or abnormalities in the distribution.
Line chart	Birth Month Trend Line Visualization	To examine the relationship between students' birth months and their average scores across different subjects.	Hovering over the line chart allows users to interact with it; tooltips displaying the average score for a particular birth month appear when this happens. This makes it possible to examine the monthly performance variations in great detail. Additionally, it might have click or tap capabilities on the legend to emphasize or mute specific subject lines, allowing users to concentrate on specific regions of interest within the dataset.
Stacked Bar Chart	Hover-Over Data Reveal (Tooltip)	To explore the correlation between students' birthday and their performance in different academic subjects.	When a user hovers over a section of the stacked bars in the chart, interactive tooltips appear, showing the proportion of students who received above-average marks in a particular subject during a particular birth quarter. A thorough examination of the score distribution among subjects and birth periods is made possible by this interactivity.
Dashboard	Student Score Input and Analysis	To provide a customized view of individual student performance, highlighting their best and worst subjects.	The student's entry of their identification number into a form field and submission of the form constitutes the main interaction. The dashboard then computes the total score, finds the highest and lowest scores, and retrieves and shows their individual scores for each subject. It

			provides a tailored assessment of one's strengths and shortcomings, providing practical insights into areas that have been accomplished and those that still need work.

Figure 7 Interactive usage table

5. *Implementation with D3.js*

Following a thorough evaluation of the D3.js library's official documentation and the course materials provided, we have decided to implement D3.js version 4 in our project. This decision was influenced by the compatibility of this version with the guidelines highlighted in the instructional materials and our practical experience gained during laboratory sessions. To begin utilizing the library in our project, we must first incorporate it into our HTML file in the following manner:

```
<script src="https://d3js.org/d3.v4.min.js"></script>
```

The d3.csv() function works similarly to the main function for importing datasets and drawing graphs

```
function updateChart(subject) {
    opt = subject;
    d3.csv("Data_Des1.csv", function(data) {
        var mappedData = data.map(function(d){
            return {
                Subject: d["Subject"],
                Score: parseFloat(d.Score),
                ExamineeNumber: parseInt(d.ExamineeNumber)
            };
        });
        var filterData = filterDataBySubject(opt, mappedData);
        renderChart(filterData); // Call renderChart function to update the chart
    });
}
```

Function using d3 to render the chart in Bar chart

```
function renderChart(filterData) {
    d3.select("#design1").selectAll("*").remove();

    var margin = { top: 100, right: 100, bottom: 20, left: 50 },
        width = 1200 - margin.left - margin.right,
        height = 400 - margin.top - margin.bottom;

    var svg = d3.select("#design1")
        .append("svg")
        .attr("width", width + margin.left + margin.right)
        .attr("height", height + margin.top + margin.bottom)
        .append("g")
        .attr("transform", "translate(" + margin.left + "," + margin.top + ")");

    var xScale = d3.scaleLinear()
        .domain([0, 10])
        .range([0, width]);

    var yScale = d3.scaleLinear()
        .domain([0, d3.max(filterData, function(d) { return d.ExamineeNumber; })])
        .range([height, 0])
        .nice();

    var barWidth = width / filterData.length;
```

```
var bars = svg.selectAll(".bar")
    .data(filterData)
    .enter()
    .append("rect")
    .attr("class", "bar")
    .attr("fill", function(d) {
        if ([1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0].includes(d.Score)) {
            return "#6EC3C9"; // Change color for specified x values
        } else {
            return "#d0d0d0"; // Default color
        }
    })
    .attr("x", function(d) { return xScale(d.Score); })
    .attr("y", function(d) { return yScale(d.ExamineeNumber); })
    .attr("width", barWidth)
    .attr("height", function(d) { return height - yScale(d.ExamineeNumber); });

// Add a title to the chart
svg.append("text")
    .attr("class", "chart-title")
    .attr("x", width / 2)
    .attr("y", -margin.top / 2)
    .attr("text-anchor", "middle")
    .attr("font-family", "Roboto Slab")
    .attr("font-weight", "bold")
    .style("font-size", "25px")
    .style("fill", "#009298")
    .text("SCORE SPECTRUM: " + opt.toUpperCase());
```

```
var labels = svg.selectAll(".label")
    .data(filterData)
    .enter()
    .append("text")
    .attr("class", "label")
    .attr("transform", function(d) {
        return "translate(" + (xScale(d.Score) + barWidth / 2) + "," + (yScale(d.ExamineeNumber) - 15) + ")
        rotate(-90)";})
    .attr("dy", "0.35em")
    .attr("font-family", "Roboto Slab")
    .style("font-size", "8px")
    .style("text-anchor", "middle")
    .text(function(d) {
        if (opt === "Literature") {
            svg.selectAll(".label")
                .filter(function(d) { return d.ExamineeNumber > 20; })
                .text(function(d) {
                    if (d.ExamineeNumber !== 0) {
                        return d.ExamineeNumber.toLocaleString();
                    }
                    return "";
                });
        }
        else {
            if (d.ExamineeNumber !== 0) {
                return d.ExamineeNumber.toLocaleString();
            }
            return "";
        }
    });
```

```
    svg.append("g")
        .attr("transform", "translate(0," + height + ")")
        .call(d3.axisBottom(xScale).ticks(40));

    svg.append("g")
        .call(d3.axisLeft(yScale));

    var tooltip = d3.select("#tooltip_des1"); // Select the tooltip div

    // Add X-axis position line
    var xAxisLine = svg.append("line")
        .attr("class", "x-axis-line")
        .style("stroke", "#777777") // Color of the X-axis line
        .style("stroke-width", 1)
        .style("stroke-dasharray", "4"); // Optional dashed line style

    // Add Y-axis position line
    var yAxisLine = svg.append("line")
        .attr("class", "y-axis-line")
        .style("stroke", "#777777") // Color of the Y-axis line
        .style("stroke-width", 1)
        .style("stroke-dasharray", "4"); // Optional dashed line style
```

```
    bars.on("mouseover", function(d) {
    // Show the existing tooltip with Score and ExamineeNumber values
    tooltip.style("display", "block")
        .style("width", "200px")
        .html("Score: " + d.Score + "<br>Number of Examinees: " + d.ExamineeNumber);

    // Get mouse position relative to the current bar
    var mousePos = d3.mouse(this);
    var xPos = mousePos[0];
    var yPos = mousePos[1];

    // Update the X-axis position line
    xAxisLine.transition()
        .duration(100)
        .attr("x1", xPos)
        .attr("y1", 0)
        .attr("x2", xPos)
        .attr("y2", height);

    // Update the Y-axis position line
    yAxisLine.transition()
        .duration(100)
        .attr("x1", 0)      Pham Le Thanh Nhan, last week • design1
        .attr("y1", yPos)
        .attr("x2", width)
        .attr("y2", yPos);
```

```
        // Calculate the tooltip's position relative to the mouse cursor
        var tooltipXPos = d3.event.pageX + 10;
        var tooltipYPos = d3.event.pageY - 10;

        // Set the tooltip's position
        tooltip.style("left", tooltipXPos + "px")
            .style("top", tooltipYPos + "px");
        })
}
```

Function for set the ranges in Line charts.

```
//NEST : group data by country
var nest = d3
  .nest()
  .key(function (d) {
    return d.Subject;
  })
  .entries(data);
console.log(nest);

// Set the ranges
var x = d3.scaleLinear().domain([0, 12]).range([0, width]);

var y = d3.scaleLinear().domain([0, 10]).range([height, 0]);
// color
var res = nest.map(function (d) {
  return d.key;
});
console.log(nest);

var color = d3
  .scaleOrdinal()
  .domain(res)
  .range(["blue", "green", "orange", "red", "purple", "#637E76", "#DF826C", "black","cyan"]);
```

```
//Hightlight the country which is hovered
var highlight = function (d) {
  d3.selectAll(".line").transition().duration(300).style("opacity", 0.05);

  d3.select("." + d.key)

    .transition()
    .duration(300)
    .style("opacity", "1")
    .style("stroke-width", "4");
};
```

Function for search and display score in dashboard

```
function searchAndDisplayStudent(studentId) {
    d3.csv("Data.csv", function(error, data) {
        if (error) throw error;

        var student = data.find(function(s) { return s.sbd === studentId; });
        if (student) {
            displayStudentDetails(student);
            document.getElementById('student-details').style.display = 'block';
        } else {
            alert('Student not found');
            document.getElementById('student-details').style.display = 'none';
        }
    });
}
```

IV. DEMO AND RESULT

DESIGN 1: COMPARATIVE SCORE ANALYSIS

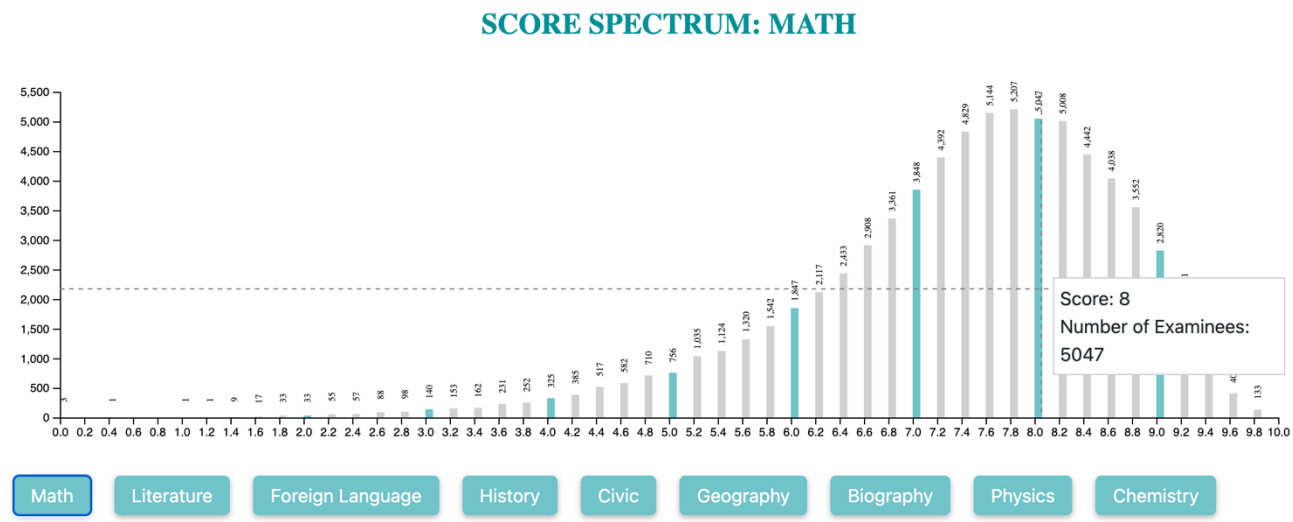


Figure 8 The bar chart of analyzing the scores of Math in the 2020 High School Graduation Exam

DESIGN 1: COMPARATIVE SCORE ANALYSIS

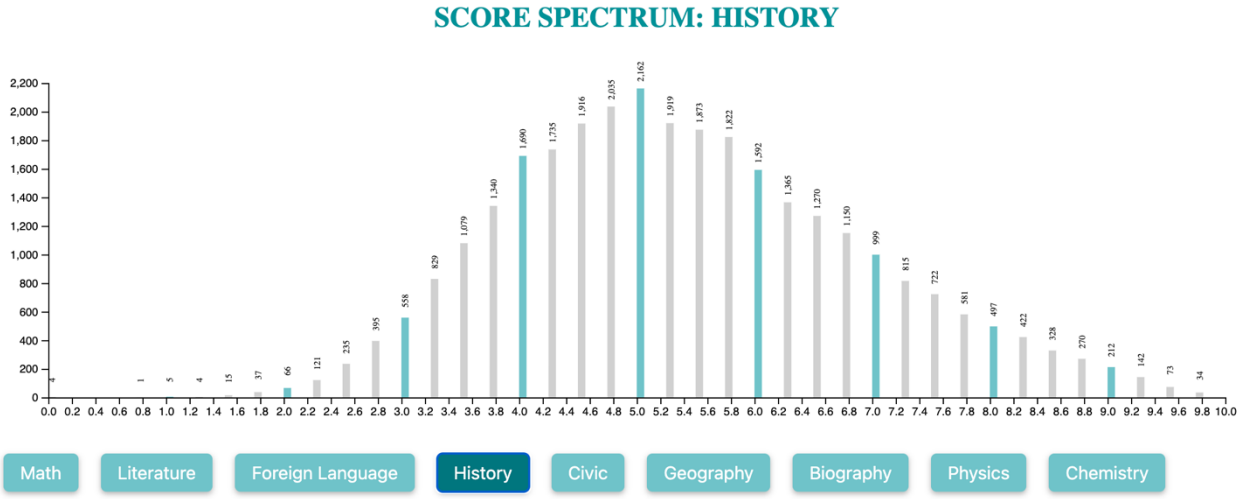


Figure 9 The bar chart of analyzing the scores of History in the 2020 High School Graduation Exam

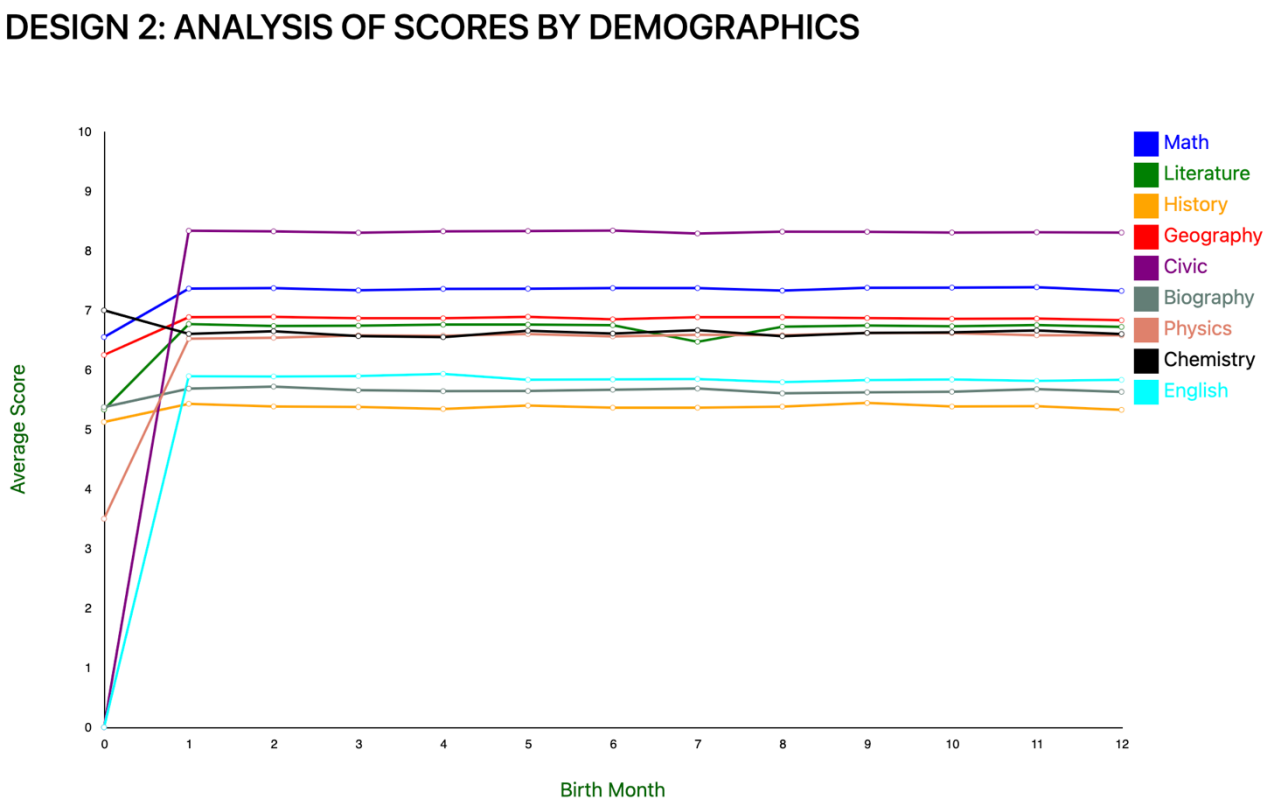


Figure 10 The line chart shows the average score of the subjects in the 2020 high school graduation exam by month of birth.

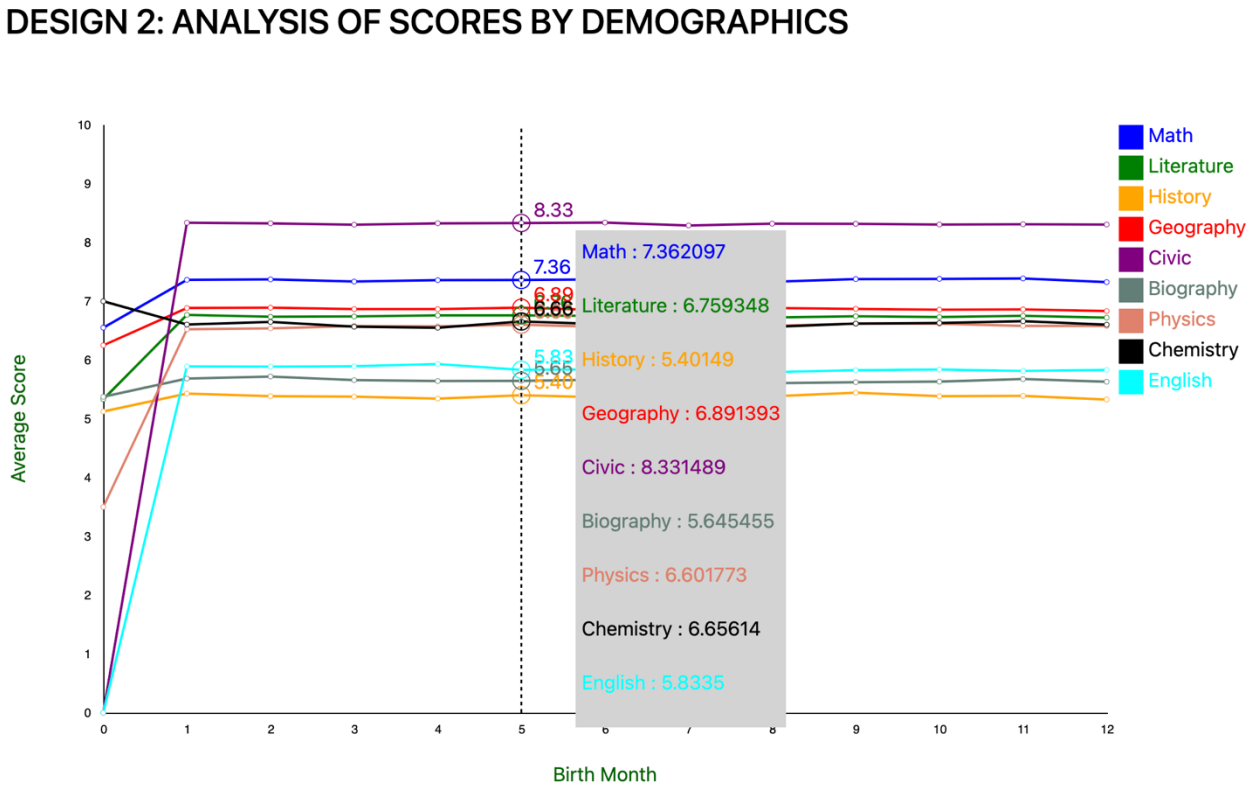


Figure 11 The line chart of analyzing the scores of Math in the 2020 High School Graduation Exam with interactive.

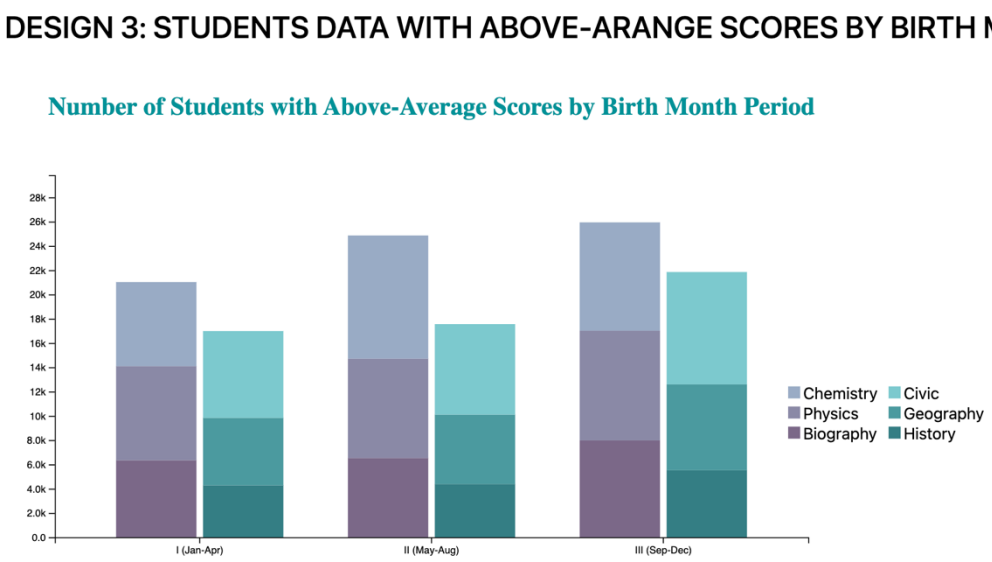


Figure 12 A stacked bar chart represents several students with Above-Average Scores by Birth Month Period.

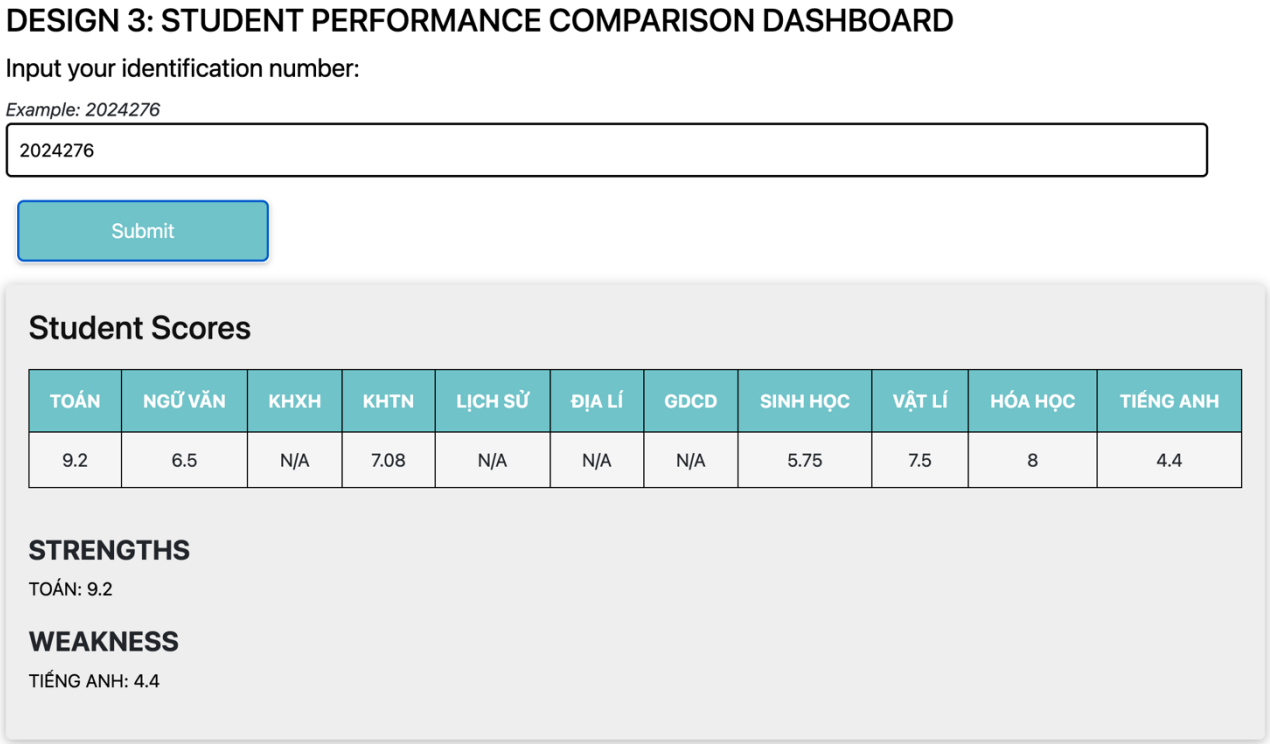


Figure 13 The dashboard displays of student scores in each subject and identify which subject is strongest and weakest in the 2020 High School graduation exam.

V. CONCLUSION

1. Achieved goals.

Following the dashboard project's completion, which examined the 2020 High School Graduation Exam results, we are reporting that the following major goals were achieved.

Enhancement of Data Accessibility: The project has been successful in making exam score data more easily accessible. All parties involved now find the process of data analysis to be more approachable thanks to the development of interactive dashboards.

Amplification of User Engagement: By incorporating interactive features like tooltips and dynamic filtering, users are now much more engaged with the data, leading to a deeper comprehension of the content.

Support for Data-Driven Decision-Making: Our dashboards are now essential for data-driven decision-making procedures. They have made it possible for teachers to identify curriculum areas that require improvement and for students to make well-informed study plans.

Improved Assistance for Students and Parents: One noteworthy achievement is the availability of a tool that helps parents and students identify academic strengths and weaknesses. They now have the tools to identify and rank the areas that require improvement thanks to this feature.

2. **Future work.**

The present dashboard project's success and foundation have been built upon, and the following areas have been determined to be developed further:

Mobile Responsiveness: Make sure users can easily access data across multiple platforms by optimizing the dashboards for use on mobile devices.

Advanced Analytics Integration: Use machine learning algorithms to forecast patterns and offer more in-depth understanding, such as predicting student performance and pinpointing the elements that have the biggest impact on academic achievement.

Collaborative Tools for Students: Provide tools that let students work together with classmates to plan studies and conduct group analyses, promoting a community-driven approach to education.

Enhanced Data Security: Put strong security measures in place to safeguard user information and adhere to data protection laws as the system develops to handle more sensitive data.

Multi-Language Support: To accommodate a diverse user base and overcome language barriers in education, make the dashboard available in multiple languages.

Feedback Mechanisms: Include user input collection tools in the dashboards to gather feedback for the platform's ongoing improvement.

Integration with Educational Institutions: For a more efficient experience, collaborate with colleges and universities to directly integrate the dashboard with their student information systems.

3. **Concluding thoughts.**

We would like to express our profound gratitude to our instructor for the unwavering support and guidance provided throughout our journey. Their expertise and instruction during laboratory sessions have been instrumental in honing the skills necessary for our success. The comprehensive and well-documented materials supplied have been an asset, laying a robust groundwork for both our educational journey and the progression of our project.

Our team deserves commendation for its cohesive collaboration, with each member contributing knowledge diligently and sharing a unified vision. Our swift progress and

the remarkable results achieved can be attributed to our efficient collaboration and well-distributed responsibilities. The fruit of our collective labor is a dynamic and intuitive dashboard that serves the needs of those seeking to delve deeper into the realm of data visualization.

We look forward with anticipation to continuing this path, ready to absorb new knowledge and develop a deeper comprehension in the field.

References

[1] thongtintuyensinh.net, "thongtintuyensinh.net," 2023. [Online]. Available: https://thongtintuyensinh.net/gioi-thieu-chung-ve-ky-thi-trung-hoc-pho-thong-quoc-gia-viet-nam.html.

[2] vietnamnet.vn, "vietnamnet.vn," 09 09 2015. [Online]. Available: https://vietnamnet.vn/bo-giao-duc-chot-phuong-an-thi-quoc-gia-2015-196459.html.