UNLEASHING THE POTENTIAL OF OUR YOUTH:A STUDENT PERFORMANCE ANALYSIS

TEAM ID:LTVIP2023TMID08028

TEAM SIZE:4

TEAM LEADER : MALLAREDDY RAMYA

TEAM MEMBERS : NEELAGIRI VS PRAVALLIKA

PERADA NIKHITA

MAVURI KEERTHAN

INTRODUCTION

Abstract— Almost every university has its own management system for keeping track of students' records.

Currently, despite the fact that University Malaysia Sarawak (UNIMAS) has a student management system that controls students' records, no permission is granted for lecturers to access the system. This is due to the privacy settings, which limit access to top management such as Deans and Deputy Deans of Undergraduate and Student Development. As a result, this project offers a system called Student Performance Analysis System (SPAS) to track student performance at the Faculty of Computer Science and Information Technology (FCSIT). The suggested system provides a predictive system that can forecast students' success in the course "TMC1013 System Analysis and Design," hence assisting the instructor.

KEYWORDS:

Student performance ; data mining; student performance analysis ; classification; prediction; system

STUDENT PERFORMANCE :

Students are a university's most valuable asset.

Universities and students play a significant role in developing high-quality graduates with their academic programmes.

performance achievement. Academic achievement

The level of achievement of the students is referred to as achievement.

educational objective that can be measured and tested

examinations, assessments, and other types of evaluations

measurements. Academic performance, on the other hand

Achievement varies depending on the type of student.

various level of performance achievement.

Academic performance of students is typically saved in

student management system, available in several formats such as

Formats include files, documents, records, photographs, and others. These

Data from available pupils could be extracted to create

This is very good information. However, the growing number of

Students' data becomes difficult to analyse while using conventional statistical approaches and database administration.

There are a few objectives that are identified during the development of this system:

i.   To develop a system for students' performance analysis.
ii.   ii. To assist the IS lecturers in analysing and predicting student performance in course "TMC1013 System Analysis and Design" by using data mining technique in the proposed system.
iii.   To identify the factors that affect the students' performance in course "TMC1013 System Analysis and Design"
iv.   To assist lecturers in keeping track of the students' progress throughout the semester.

DATA MINING:

In Minds helps University Malaysia Sarawak (UNIMAS) to monitor the performance of various areas in every UNIMAS's departments [2]. The system enables top and mid-management in UNIMAS to have a clear look on the areas that needed attention by looking at the figures, revenues and risks. The features, ease of use and flexibility provided by the system makes the performance analysis in UNIMAS to be performed in an ideal solution. Charts are provided by the system for ease of student performance's interpretation. From the reviews on these existing systems, useful techniques and features could be applied into the proposed system for a better system's performance. The WEKA is chosen as a tool for data mining because it is open source software.

There are a few features from the existing systems that are employed during the design and implementation phase of the proposed system. These features and functionalities include the user interface, students' performance prediction, illustration displays and report generation. A good user interface provides an user-friendly interface as it is easy to be navigate and not complicated. Meanwhile, the students' performance prediction is included into the proposed system to make sure the objectives are achieved. Furthermore, the generation of reports in Portable Document Format (PDF) and illustration display such as charts in PDF makes student performance analysis easier. From these features found in proposed system, all the user requirements would be fulfilled. The user requirements collected from lecturers of FCSIT during the system analysis phase are as follows:
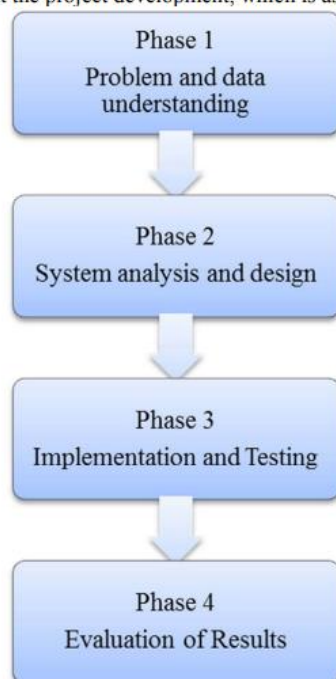
i.   Able to help lecturers to automatically predict students' performance in course "TMC1013 System Analysis and Design"

ii.   Able to keep track and retrieve students' performance in a particular course and
        semester
iii.   Able to view the factors that affect the students' prediction result
iv.   Able to generate students' reports
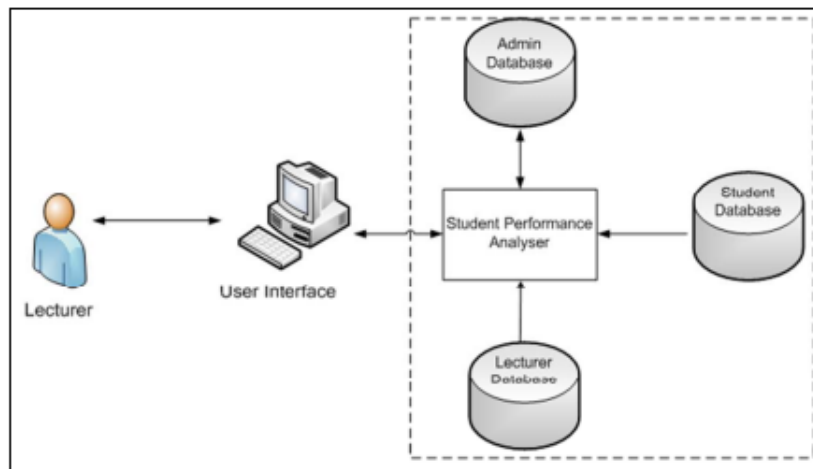
METHODOLOGY:

There are several phases of methodology used throughout the project development, which is as follows:

- Phase 1 — Problem and data understanding
- Phase 2 — System analysis and design
- Phase 3 — Implementation and Testing
- Phase 4 — Evaluation of Results

1.Problem and data understanding:

The problem and data understanding is critical in determining the success of the Student Performance Analysis system. Before the system development, problems and data understanding is identified to define the project goal and objectives.

2.System analysis and design:

In this phase, the overall flow of the system is planned, analysed and designed. The system and user requirements are analysed and listed in table format. Data flow diagram is used to chart the input, processes and output of the system. Data flow diagram from the context diagram up to the first level is analysed and drawn. The hardware requirement in this phase is a computer for analysis and design. Other than that, Microsoft Office Visio 2007 is needed to draw the ERD and data flow diagrams.

3.Implementation and testing:

It focuses on understanding the project goals and requirements form a business point of view, then converting this information into a data mining problem afterward a preliminary plan designed to accomplish the target. It uses various measures of statistical validity to determine whether there are problems in the data or in the model.
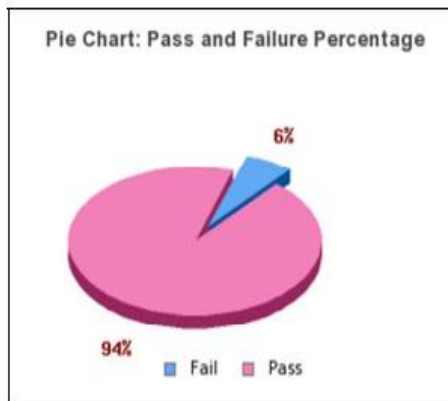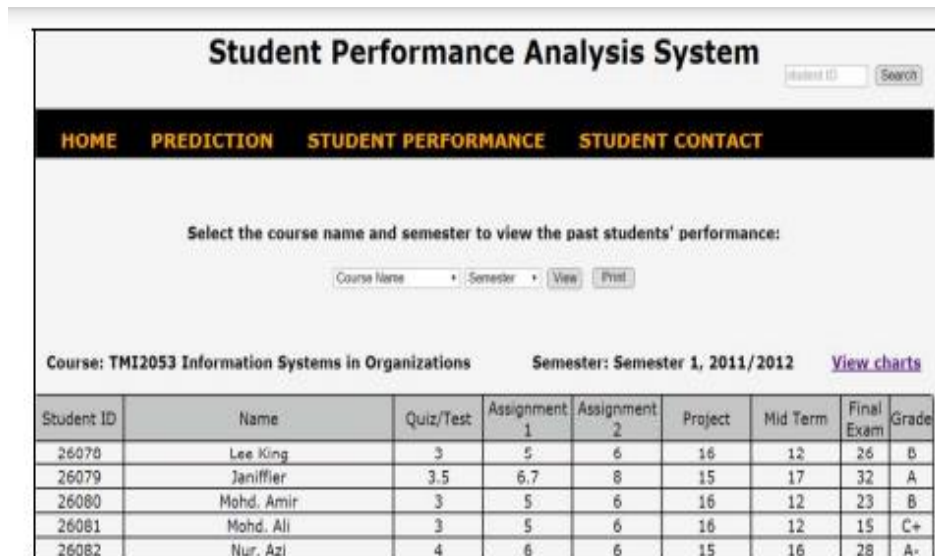
4.Evaluation of results:

Evaluation of System For the evaluation of the system, five end-users are requested to evaluate the usability of Student Analysis Performance system. This is to ensure the objectives of the proposed system are achieved as well as to ensure the ease of navigation across the interfaces of the proposed system. Moreover, the evaluation is performed to ensure the high effectiveness of the proposed system is achieved. From the evaluation performed, a list of users' recommendation is stated as shown below:

i.   Apply the students' results prediction to other courses offered by FCSIT.
ii.   Enable the viewing of all semesters' performance when search for a student's performance.
iii.   Provide a flexible rules when perform students' results prediction.

Student Performance Analyser (SPA):

SPA is existing secure online web-based software that enables educators to view the students' performance and keep track of the school's data. The SPA is a tool designed for analysing, displaying, storing, and getting feedback of student assessment data [3]. It is a powerful analyser tool used by schools worldwide to perform analysis and displays the analysis data once raw student data is uploaded to the system. The analysis is done by tracking the student or class to get the overall performance of student or class. It helps to identify the students' performance which is below the expected level, at expected level or above the expected level. This would allow the educators or staffs to identify the current students' performance easily. Other than that, it enables various kinds of students' performance report such as progress report and achievement report to be generated.

Student Performance Analysis System

Student ID	Name	Quiz/Test	Assignment 1	Assignment 2	Project	Mid Term	Final Exam	Grade
26078	Lee King	3	5	6	16	12	26	B
26079	Janiffier	3.5	6.7	8	15	17	32	A
26080	Mohd. Amir	3	5	6	16	12	23	B
26081	Mohd. Ali	3	5	6	16	12	15	C+
26082	Nur. Azi	4	6	6	15	16	28	A-
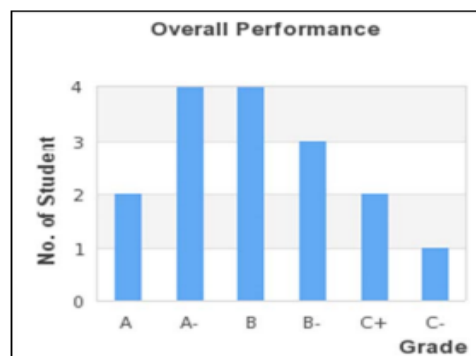
In addition, there are two bar charts and one pie chart implemented to illustrate the overall students' performance in a course. Above figures shows the charts displayed in the system to keep track of students' performance.

After the system is built, the unit testing, system testing and user acceptance testing are needed for errors detection before the system is distributed and used by IS lecturers. This is to ensure the performance of the system is in its optimal state. Besides, the errors and bugs that are detected during testing of the proposed system can be fixed. The unit and system testing will be tested by developer while the user acceptance testing will be tested by a few end-users to ensure the functionalities of system are working as expected.

Classification and Predication:

There are two forms of data analysis that can be used to extract models describing important classes or predict future data trends. These two forms are as follows:

1. Classification
2. Prediction

We use classification and prediction to extract a model, representing the data classes to predict future data trends. Classification predicts the categorical labels of data with the prediction models. This analysis provides us with the best understanding of the data at a large scale.

Classification models predict categorical class labels, and prediction models predict continuous-valued functions. For example, we can build a classification model to categorize bank loan applications as either safe or risky or a prediction model to predict the expenditures in dollars of potential customers on computer equipment given their income and occupation.

Classification and Prediction Process
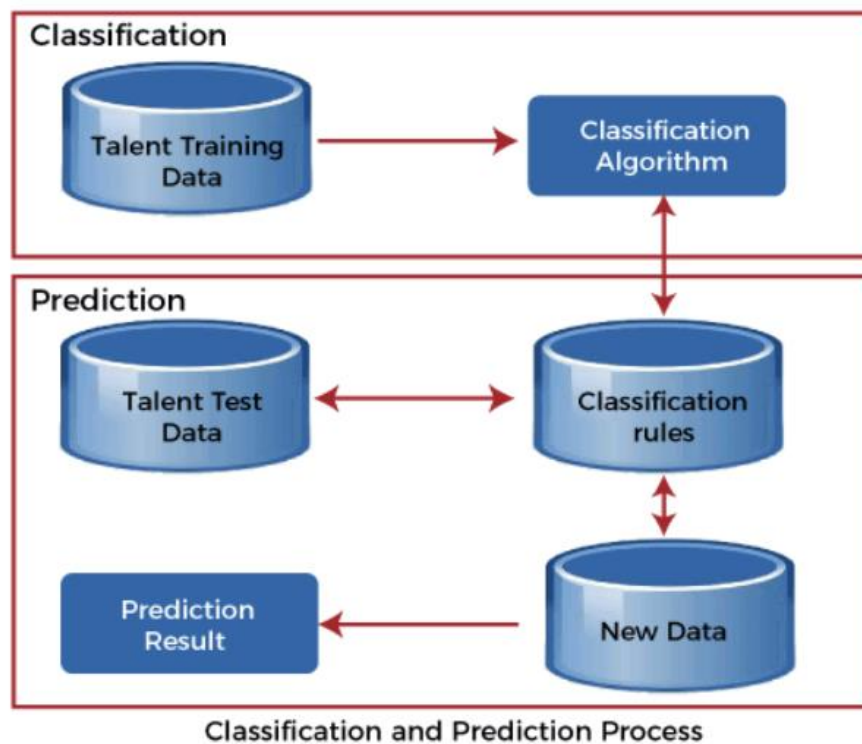
What is Classification?

Classification is to identify the category or the class label of a new observation. First, a set of data is used as training data. The set of input data and the corresponding outputs are given to the algorithm. So, the training data set includes the input data and their associated class labels. Using the training dataset, the algorithm derives a model or the classifier. The derived model can be a decision tree, mathematical formula, or a neural network. In classification, when unlabeled data is given to the model, it should find the class to which it belongs. The new data provided to the model is the test data set.

Classification is the process of classifying a record. One simple example of classification is to check whether it is raining or not. The answer can either be yes or no. So, there is a particular number of choices. Sometimes there can be more than two classes to classify. That is called *multiclass classification*
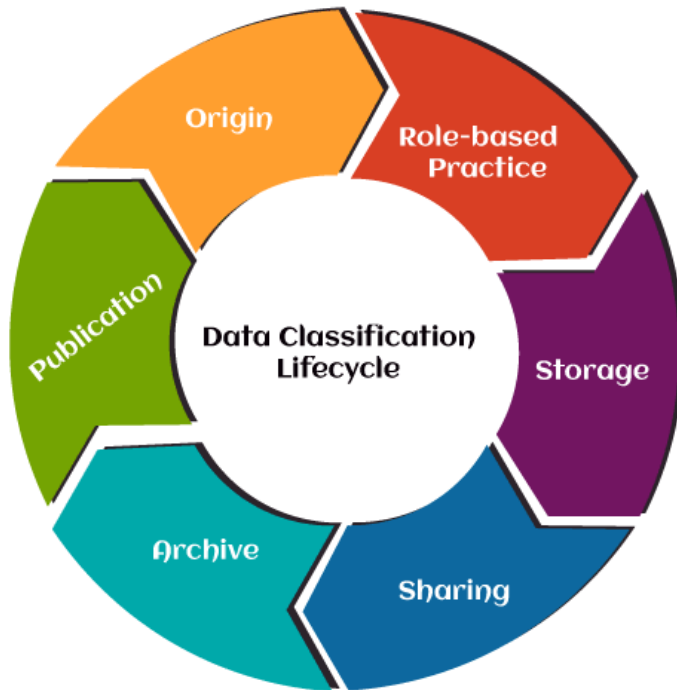
1. **Origin:** It produces sensitive data in various formats, with emails, Excel, Word, Google documents, social media, and websites.

2. **Role-based practice:** Role-based security restrictions apply to all delicate data by tagging based on in-house protection policies and agreement rules.

3. **Storage:** Here, we have the obtained data, including access controls and encryption.

4. **Sharing:** Data is continually distributed among agents, consumers, and co-workers from various devices and platforms.

5. **Archive:** Here, data is eventually archived within an industry's storage systems.

6. **Publication:** Through the publication of data, it can reach customers. They can then view and download in the form of dashboards.

What is Prediction?

Another process of data analysis is prediction. It is used to find a numerical output. Same as in classification, the training dataset contains the inputs and corresponding numerical output values. The algorithm derives the model or a predictor according to the training dataset. The model should find a numerical output when the new data is given. Unlike in classification, this method does not have a class label. The model predicts a continuous-valued function or ordered value.

Regression is generally used for prediction. Predicting the value of a house depending on the facts such as the number of rooms, the total area, etc., is an example for prediction.
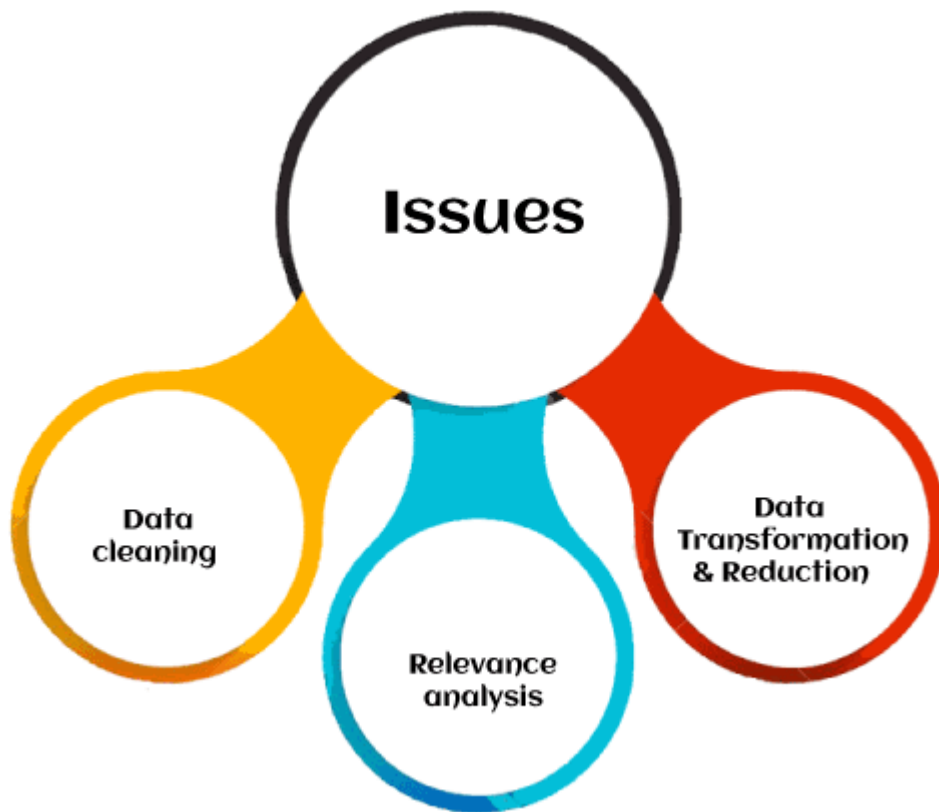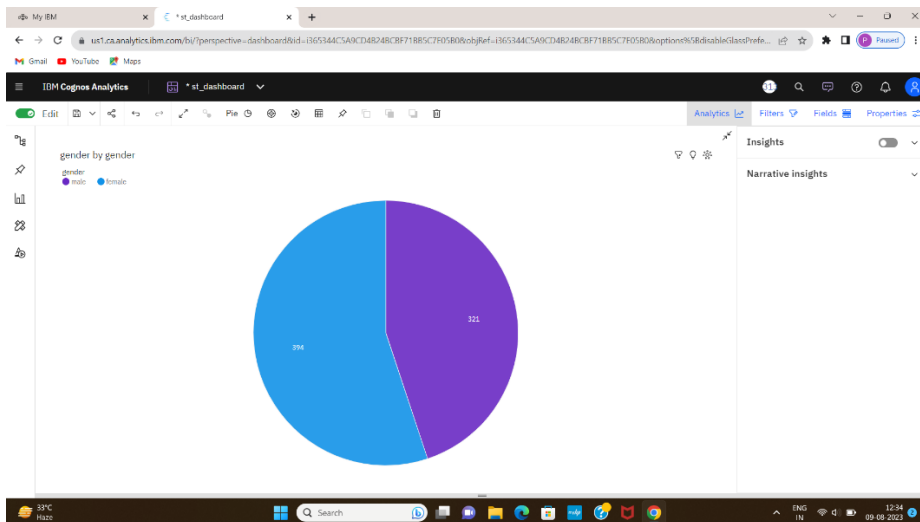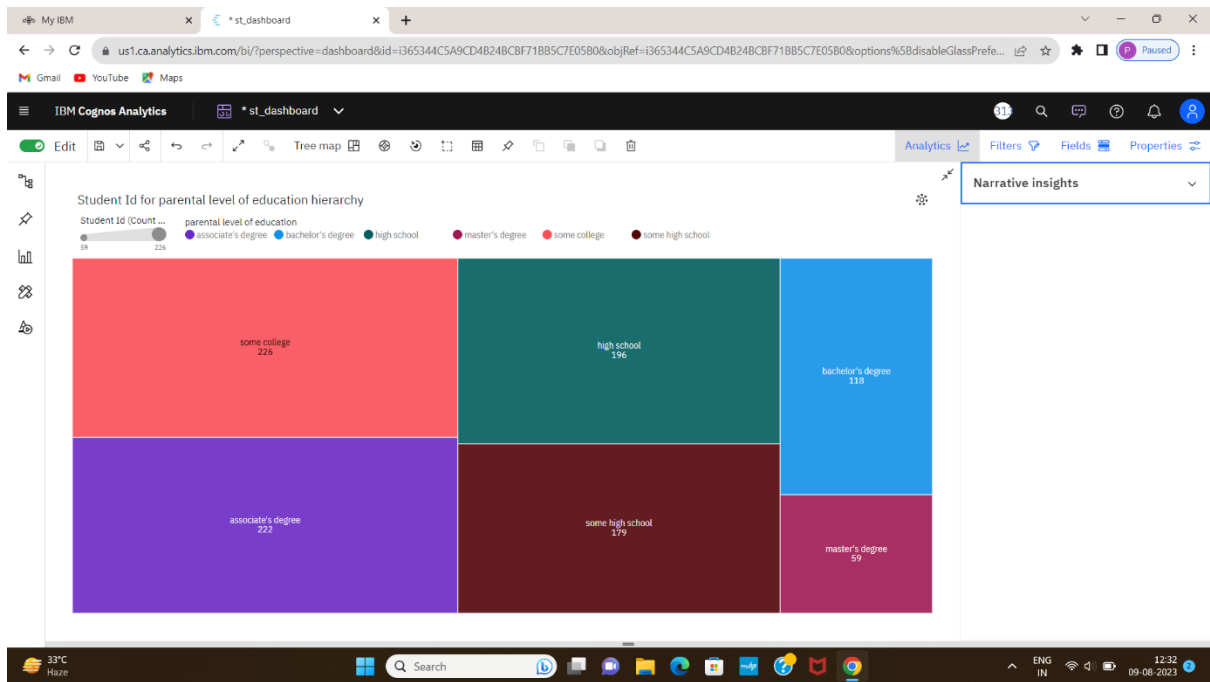
1. **Data Cleaning:** Data cleaning involves removing the noise and treatment of missing values. The noise is removed by applying smoothing techniques, and the problem of missing values is solved by replacing a missing value with the most commonly occurring value for that attribute.

2. **Relevance Analysis:** The database may also have irrelevant attributes. Correlation analysis is used to know whether any two given attributes are related.

3. **Data Transformation and reduction:** The data can be transformed by any of the following methods.

   o **Normalization:** The data is transformed using normalization. Normalization involves scaling all values for a given attribute to make them fall within a small specified range. Normalization is used when the neural networks or the methods involving measurements are used in the learning step.

   o **Generalization:** The data can also be transformed by generalizing it to the higher concept. For this purpose, we can use the concept hierarchies.

RESULT:

DATA VISUALIZATION**:**
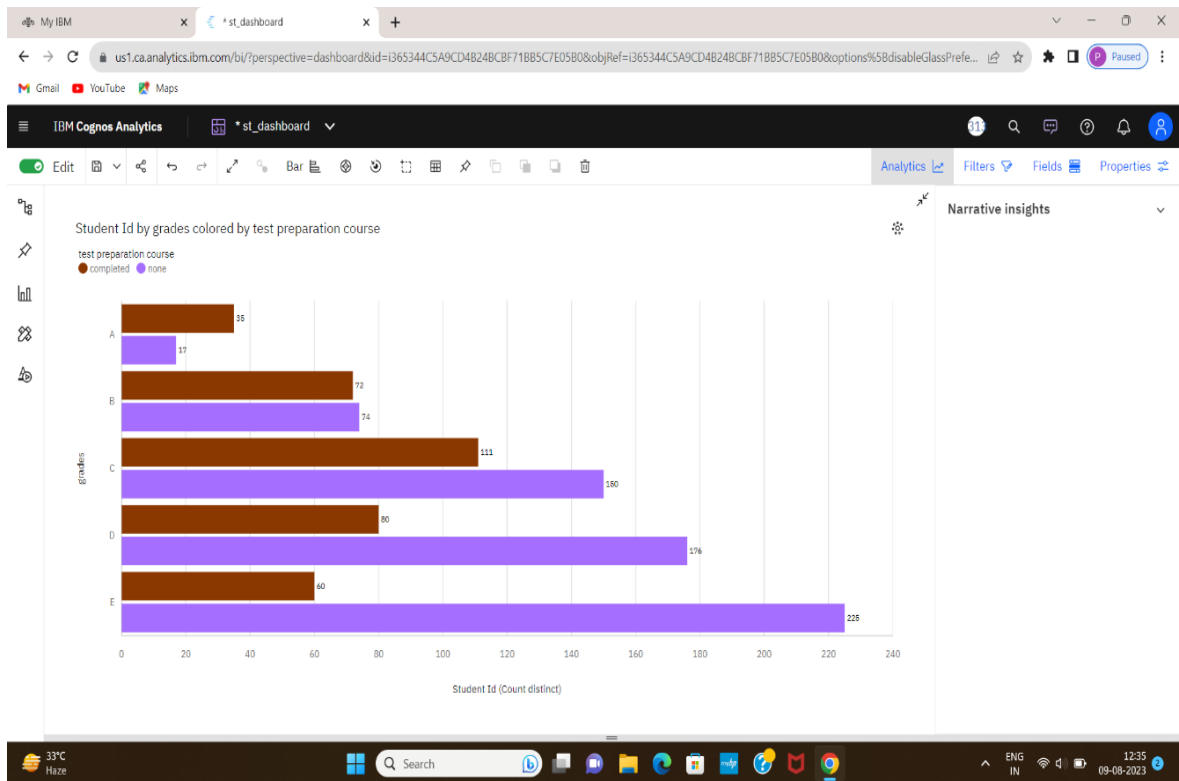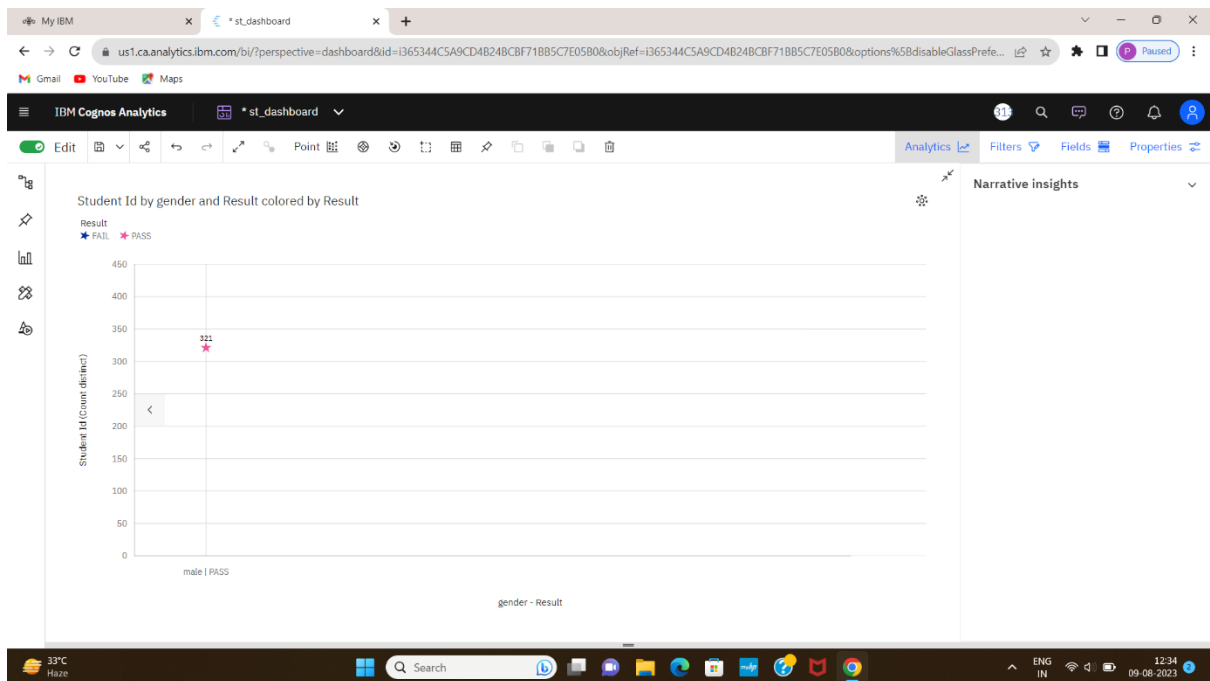
STORY:

Student Id by grades colored by test preparation course

test preparation course
● completed ● none

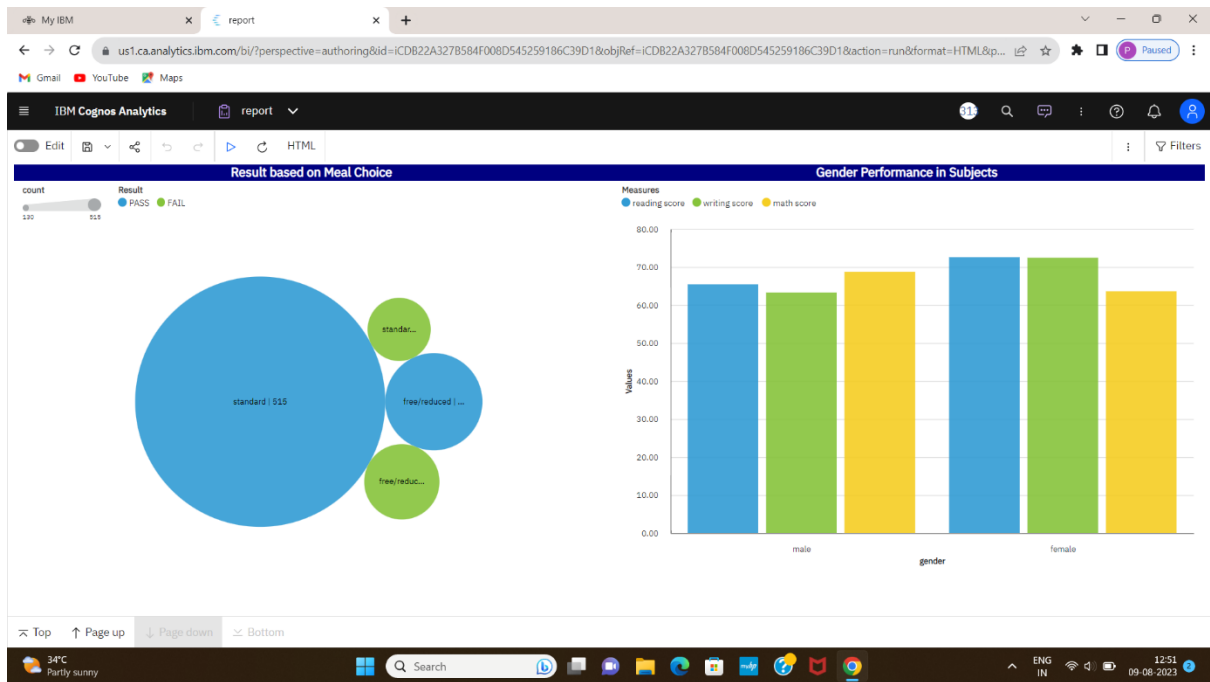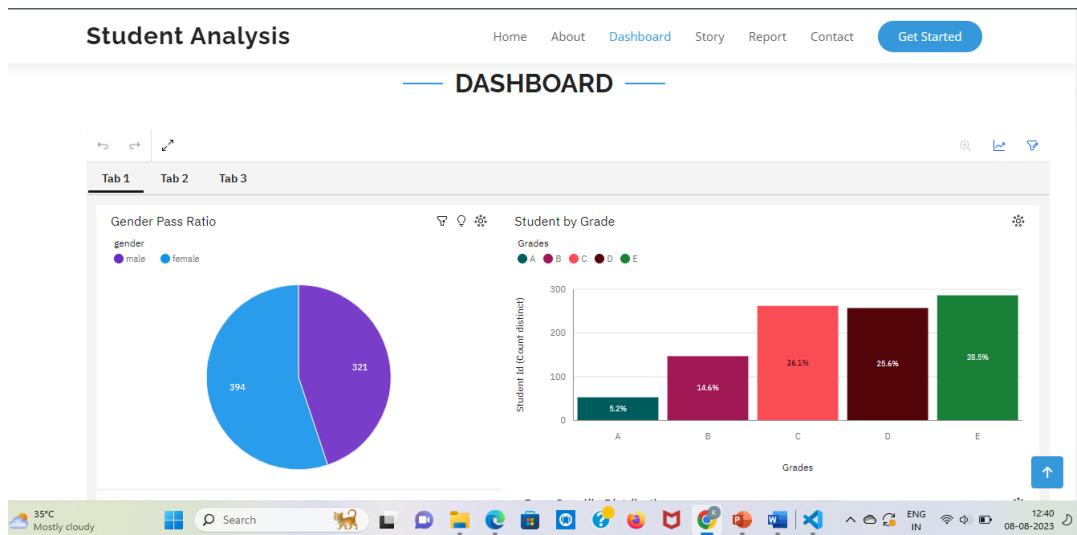Student Id by race/ethnicity colored by Result
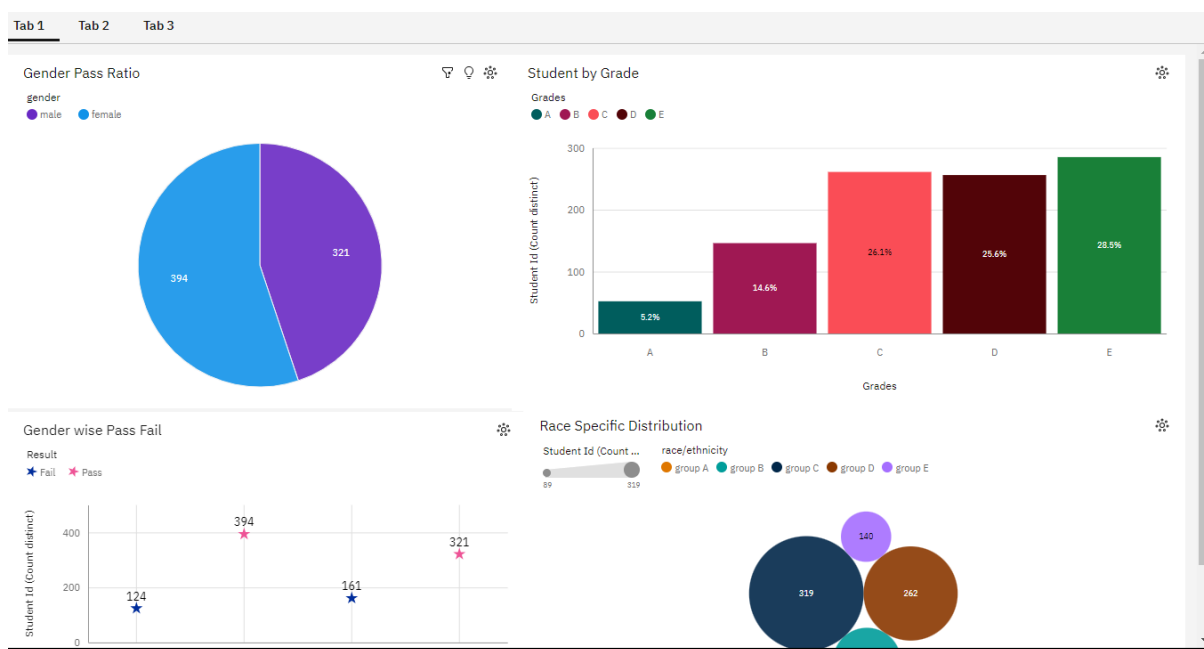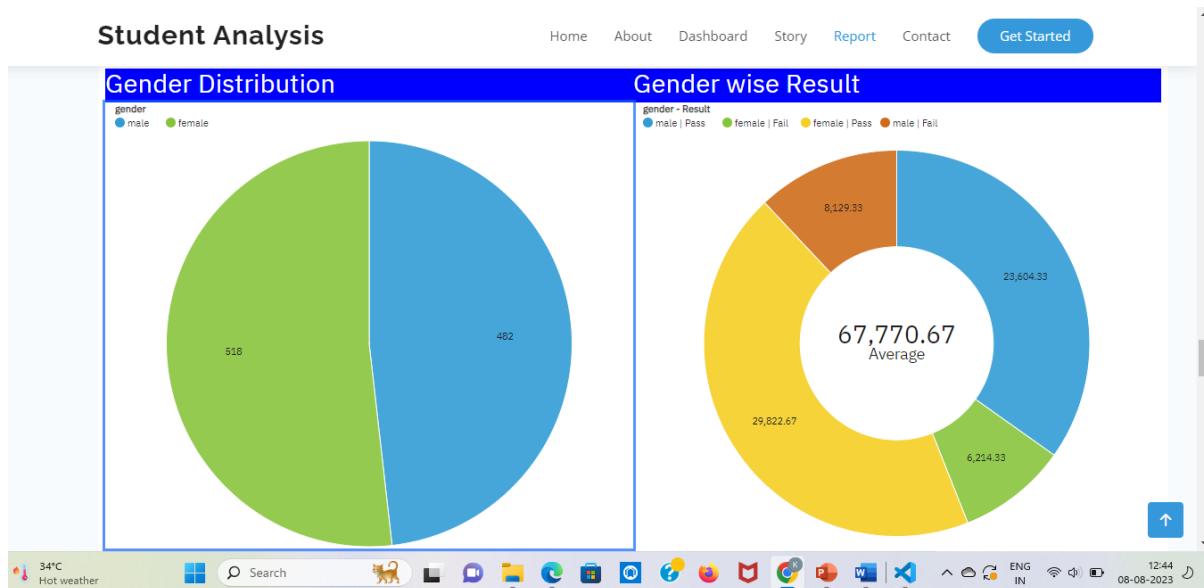
Result
● FAIL ● PASS
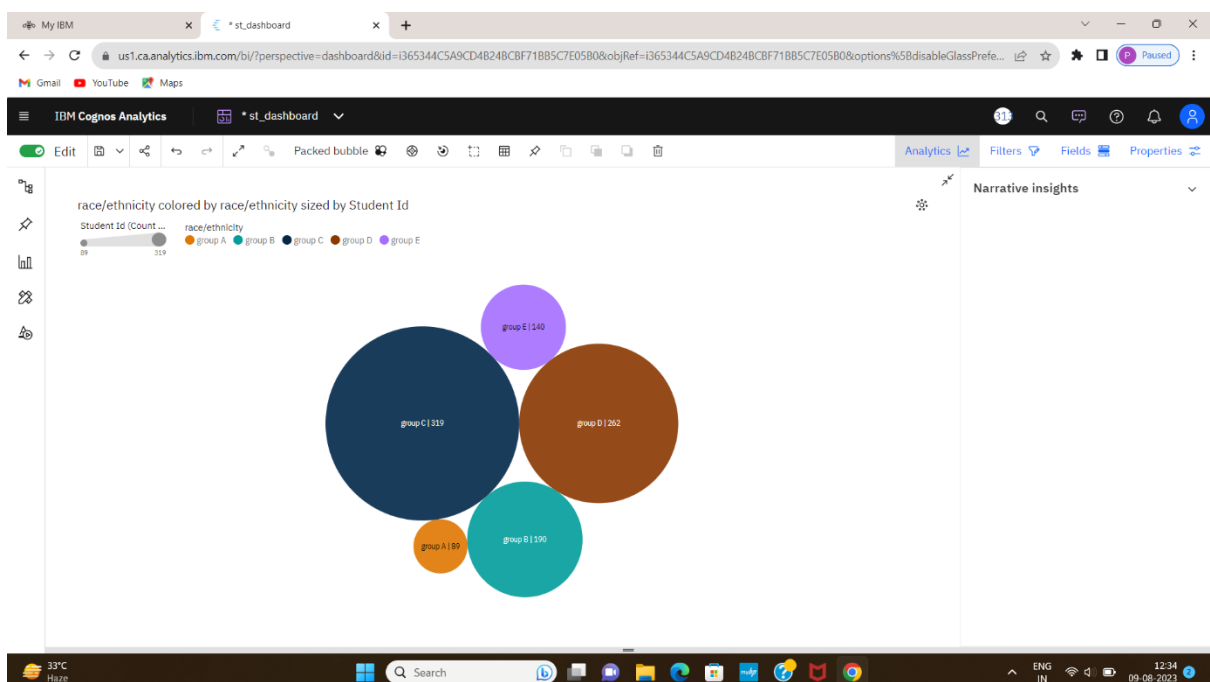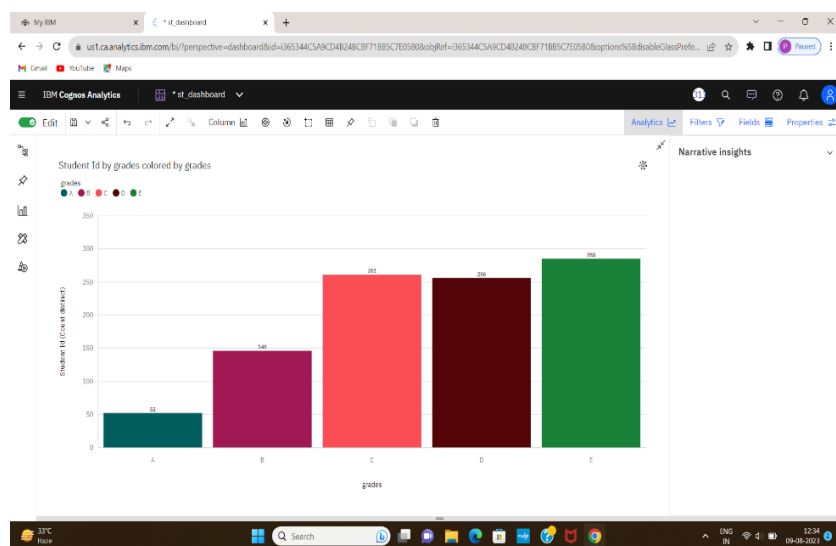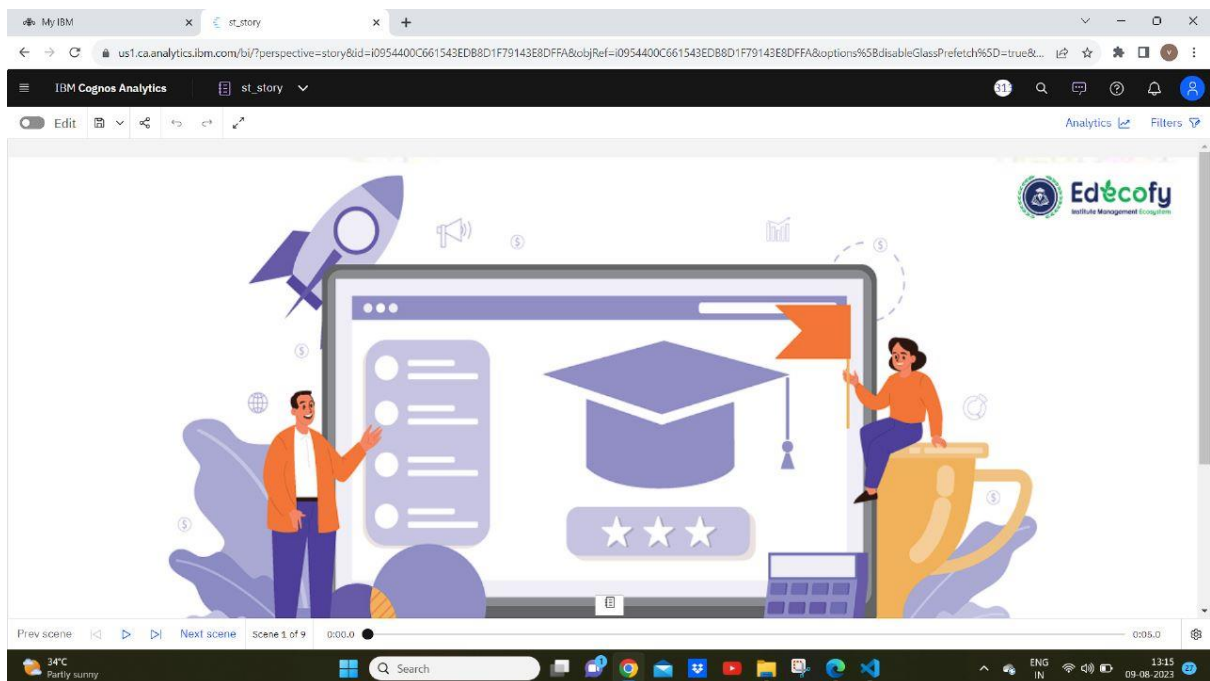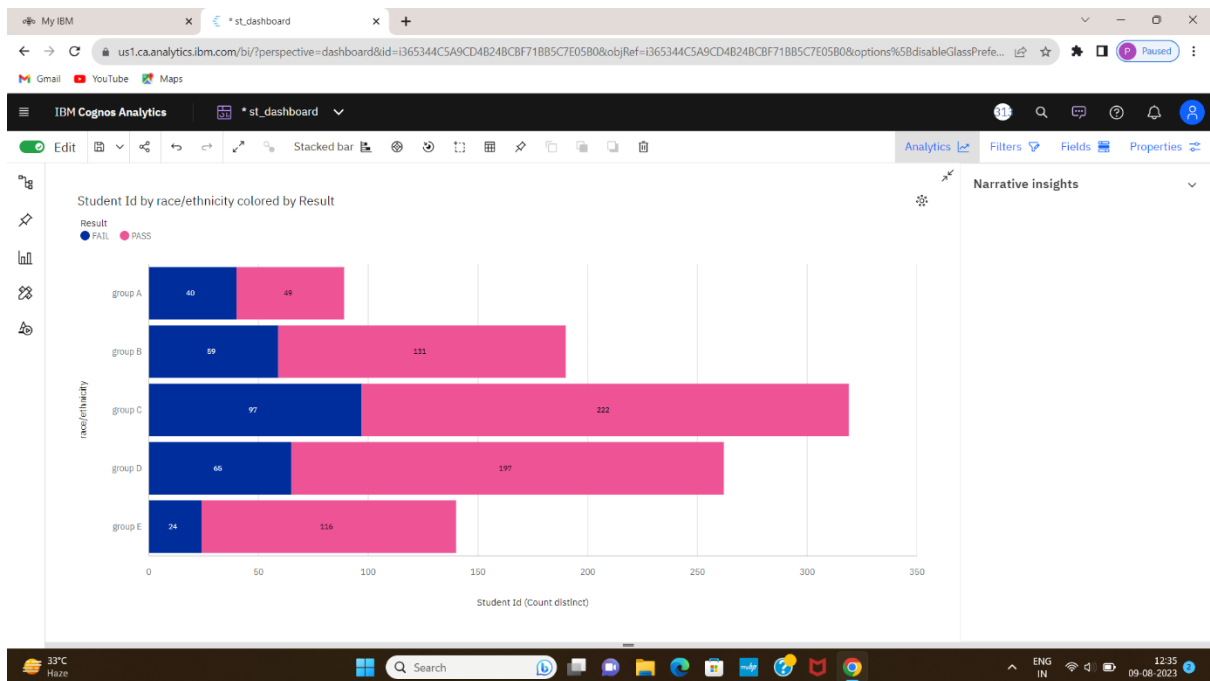
REPORT:
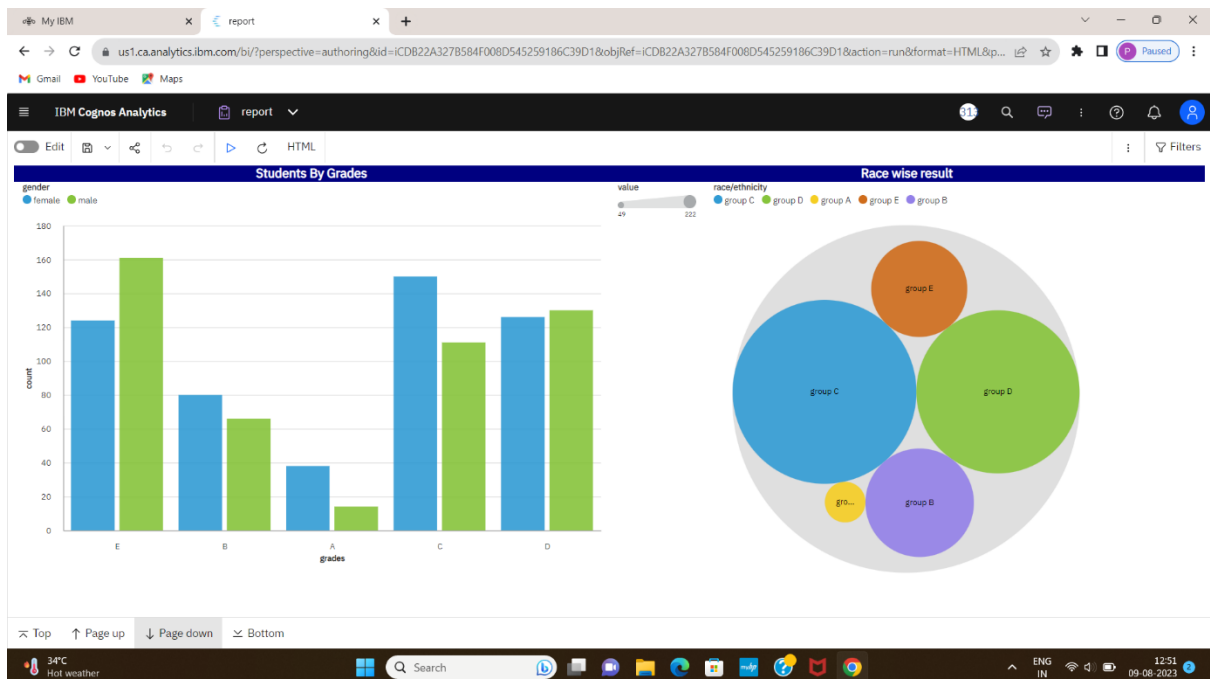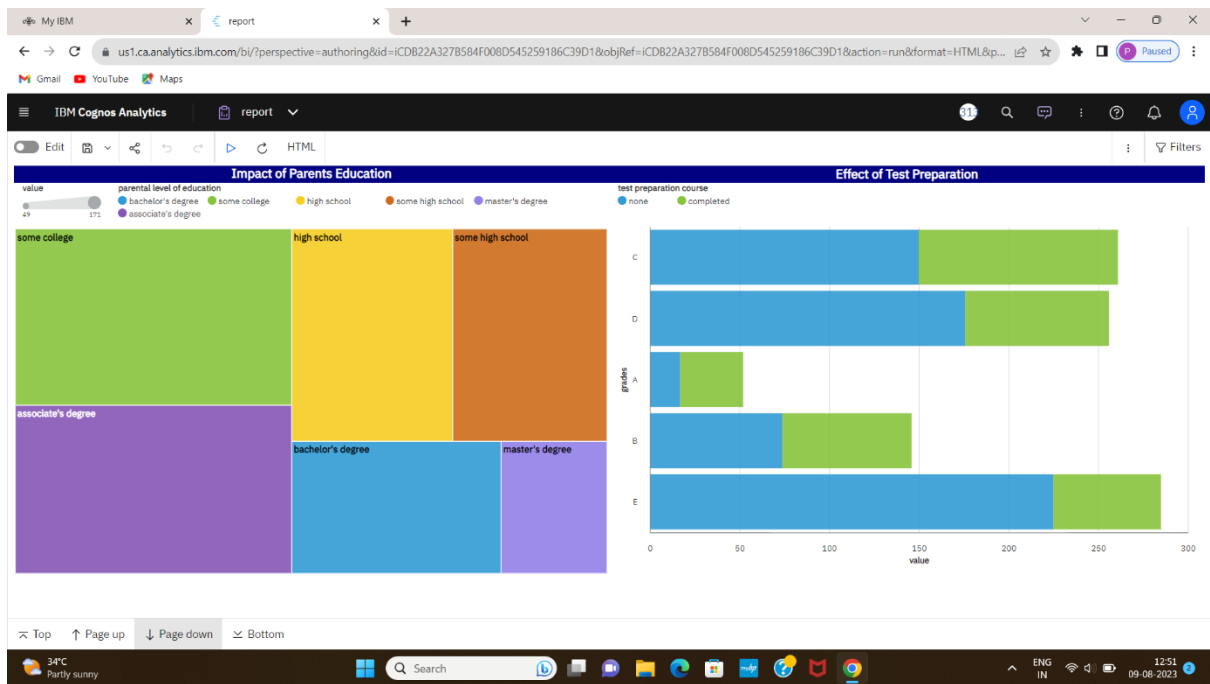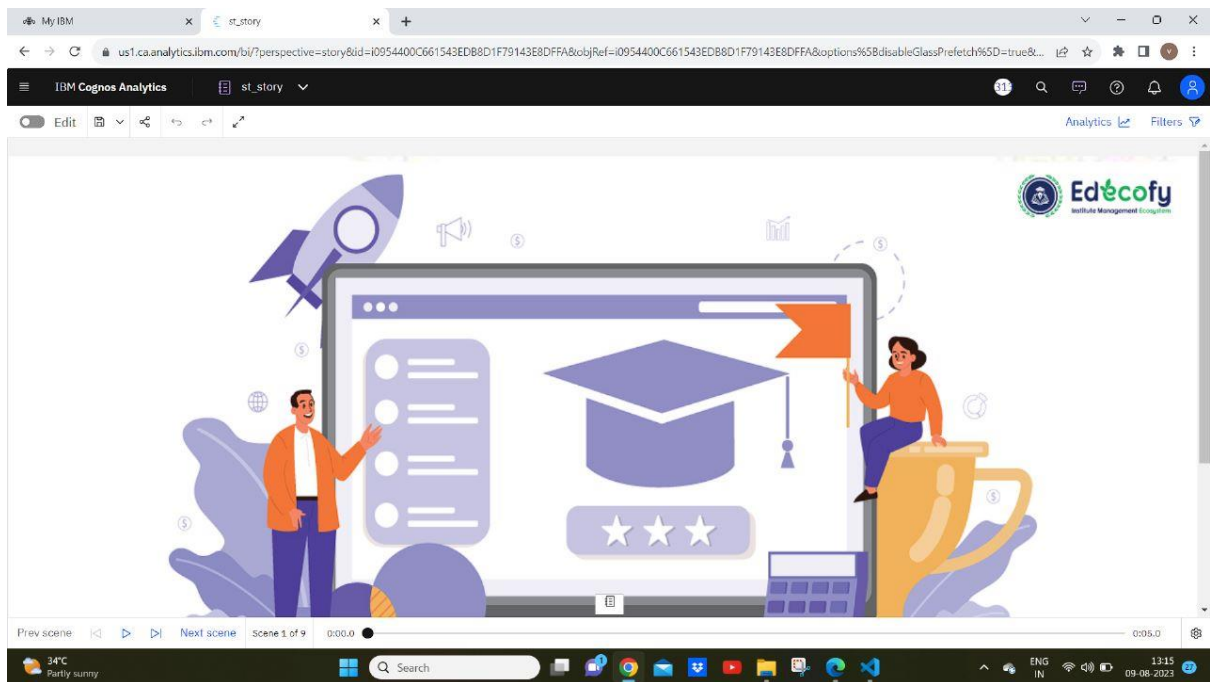
WEB INTEGRATION OUTPUT:

CONCLUSION:

Finally, the project focuses on the creation of a system for analysing student performance. Classification is a data mining approach In this project, an algorithm is used to ensure prediction. of the course "TMC1013 System" students' performance. It is feasible to perform "Analysis and Design." The primary contribution The SPAS helps professors conduct their classes. An examination of student performance. The system helps lecturers. In identifying

kids who are expected to fail the"TMC1013 System Analysis and Design" is a course. The SPAS also assists professors in retrieving information. Their students' progress over the semesters.