A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

BY

CHEAH CHENG QIAN

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION SYSTEMS

Faculty of Information and Communication Technology

(Kampar Campus)

JUNE 2022

A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

BY

CHEAH CHENG QIAN

A REPORT

SUBMITTED TO

Universiti Tunku Abdul Rahman

in partial fulfillment of the requirements

for the degree of

BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION SYSTEMS

Faculty of Information and Communication Technology

(Kampar Campus)

JUNE 2022

UNIVERSITI TUNKU ABDUL RAHMAN

REPORT STATUS DECLARATION FORM

Title: A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

Academic Session: 2022

I CHEAH CHENG QIAN
(CAPITAL LETTER)

declare that I allow this Final Year Project Report to be kept in Universiti Tunku Abdul Rahman Library subject to the regulations as follows:

1. The dissertation is a property of the Library.
2. The Library is allowed to make copies of this dissertation for academic purposes.

Verified by,

(Author's signature)

(Supervisor's signature)

Address:
No. 32 Laluan Pinji Wani,
Taman Pinji Wani
31650 Ipoh Perak

Ts Dr Chan Lee Kwun
Supervisor's name

Date: 02 September 2022

Date: 02 September 2022

FACULTY/INSTITUTE* OF INFORMATION AND COMMUNICATION TECHNOLOGY

UNIVERSITI TUNKU ABDUL RAHMAN

Date: 02 September 2022

SUBMISSION OF FINAL YEAR PROJECT /DISSERTATION/THESIS

It is hereby certified that Cheah Cheng Qian (ID No: 17ACB04662) has completed this final year project/ dissertation/ thesis* entitled "A Visualized Analysis on Student Performance by Using Cloud Business Intelligence" under the supervision of Ts Dr Chan Lee Kwun (Supervisor) from the Department of Information Systems, Faculty/Institute* of Information and Communication Technology , and _____ (Co-Supervisor)* from the Department of _____, Faculty/Institute* of _____.

I understand that University will upload softcopy of my final year project / dissertation/ thesis* in pdf format into UTAR Institutional Repository, which may be made accessible to UTAR community and public.

Yours truly,

_____

(*CHEAH CHENG QIAN*)

*Delete whichever not applicable

DECLARATION OF ORIGINALITY

I declare that this report entitled "**A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE**" is my own work except as cited in the references. The report has not been accepted for any degree and is not being submitted concurrently in candidature for any degree or other award.

Signature : _____

Name : Cheah Cheng Qian

Date : 02 September 2022

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to several individuals and organizations for supporting me throughout my Graduate study. First, I wish to express my sincere gratitude to my supervisor, Ts Dr Chan Lee Kwun, for her enthusiasm, patience, insightful comments, helpful information, practical advice, and unceasing ideas that have always helped me tremendously in my research and writing of this thesis. Her immense knowledge, profound experience, and professional expertise in the field of Business Intelligence has enabled me to complete this research successfully. Without her support and guidance, this project would not have been possible. I could not have imagined having a better supervisor in my study.

To a very special person in my life, Yeoh Pei Ying, for her patience, unconditional support, and love, and for standing by my side during hard times. Finally, I must say thanks to my parents and my family for their love, support, and continuous encouragement throughout the course.

Finally, a big thanks to Universiti Tunku Abdul Rahman especially to the Faculty of Information Technology for giving me a chance to study here.

ABSTRACT

During in this state of global pandemic, it is claimed that there is a rise in trend of poor academic performance among student in certain courses or other specific skills. In this matter, due to this uprising yet concerning issue, it is crucial that the deteriorations of students' academic performance seek a proper solution. Previous research has primarily relied on the traditional and simple method to analyze students' performance. However, the results and outcomes are not too accurate due to lack of limited data as well as analysis of students' CLO performance. Therefore, possible factors have examined to have certain degree of relevance in affecting students' performance in certain skills or courses. The aim of this project is to gain deeper understanding and insights to find out which skills or courses where most students can't perform well and to propose a solution into solving the issue by implementing analysis tools using software tools like Python and MS Power BI. Hence, a comparison of business intelligence tools will be carried out to choose an appropriate BI tool to conduct the analysis due to vast amount of software which available in the IT market. Next, a simple prototype of a system is developed to conduct the analysis alongside a business intelligence tool to graphically display and visualize the analysis results along with graphs on the existing dashboard. To analyze students' performance, a sample data of students' results is created to check its CLO through specific and measurable statements in which defines students' knowledge, skills, and attitudes by demonstrating by the completion of a course. Cognitive, Practical and Affective acts as an indicator to see which student did not perform well in which components in a course and indicates what a student should understand and be able to do after learning. The dashboards developed will be used as a module to analyze students' performance. After analyzing the findings through graphical objects, formulations and recommendations, the results and outcomes of the data analysis are shown and presented and then described based on the finding. Recommendations to overcome the issue is proposed and conclusions are drawn for future research opportunities as well as data collection. The results of this project will help us gain a better insight and understanding on the academic performance of students as well as knowing which area of weaknesses are they into so that students in higher education could be able to have a huge leap in academic performance.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

BI	Business Intelligence
CLO	Course Learning Outcome
DW	Data Warehouse
ETL	Extraction, Transformation and Loading
OLAP	On-line Data Processing
CRM	Customer Relationship Management
DSS	Decision Support Systems
GIS	Geographic Information Systems
MD	Master Data
DQM	Data Quality Management
IoT	Internet of Things
IA	Information System Engineering
CS	Computer Science
IB	Business Information Systems
IT	Information Technology
SQL	Structured Query Language
NPL	Natural Language Processing
IBM	International Business Machine

Chapter 1
Introduction

Cloud computing

Cloud computing has emerged as one of the most significant technologies in our current technological era. As a result, cloud computing is defined as a computing sharing technology that provides available computing resources, with cloud computing encompassing storage, compute control, and application delivery via the internet as a service [5]. Aside from that, Business Intelligence (BI) is characterized as a critical idea that has sparked widespread interest in both academia and industry, with BI systems being used mostly in areas of business that require decision-making to obtain value and insights. Business intelligence also refers to the use of specialized tools and data to make decisions in a variety of organizational settings. Due to the expanding relevance of BI systems, they are still the most widely and widely utilized IT solution today. The overall goal of a business intelligence system is to help people make better decisions.

Numerous studies have proven that higher education is critical to a country's economic progress and prosperity. As a result, data is critical for major businesses, particularly higher education, to make informed decisions. Because educational institutions have a high number of students, employees, and activities, BI tools can be used to evaluate a vast amount of data to forecast students' academic behavior, performance, and shortcomings. These statistics, for example, are crucial throughout the decision-making process because they may be utilized to communicate the need for change and improvement, which helps to improve the academic institution of a particular institution's students [6]. BI solutions such as Power BI and Tableau give services and assistance that aid top management in completely analyzing complex business circumstances and streamlining decision-making processes.

With the availability of 'big data' and breakthroughs in machine intelligence, Business Intelligence has become an exceptionally crucial term in this modern era of technological growth [1]. BI systems had already generated significant attention in industry and academia [2], and they've been extensively applied in many aspects of industry that require making beneficial choices. Also, practitioners as well as academics required, however, get a better

understanding of the mechanisms through which businesses might derives its value from BI in addition to helping it reach its full potential. Studies that have investigated BI utilizing a variety of hypotheses, investigation methods, and scientific data so far. Business intelligence (BI) refers to the technologies, tools, and procedures that are used to collect, combine, evaluate, and report business data. Aside from that, the basic goal of business intelligence is to allow people to make better decision. Organizations are under mounting pressure to collect, analyze, and to use their data to enhance decision-making as well as enhance company operations.

Business Intelligence

Business intelligence (BI) is a framework that enables a company to present the data, make sense of the information, and knowledge into insight [3]. "A combination of techniques, processes, systems, technologies, and tools that turn data into useful and valuable information intended to provide greater strategic planning, tactical, as well as operational insights and decision-making," as according to Forrester Research. Aside from that, there are also numerous benefits of implementing business intelligence in our modern world as business intelligence is more than just a software. One of the main benefits of business intelligence is faster analysis alongside with intuitive dashboards. BI solutions are built to handle large amounts of data processing in the cloud on the company's infrastructure. Also, BI tools collect data from many sources and store it in a data warehouse, where it is then analyzed using user requests, drag-and-drop analytics, and dashboards. In other words, Business intelligence dashboards make the data analysis easier and simpler for non-technical people, allowing them to build narratives with data without attempting to learn coding. It was also found that business intelligence increased organizational efficiency. Leaders may use BI to access the information and get a comprehensive perspective of their activities, as well as measure outcomes against the rest of the company. Leaders may uncover regions of potential by taking a comprehensive picture of the organization. As companies spend less time analyzing data and generating analyses, they have much more opportunities to develop on new initiatives and services for the company [4].

Other than that, business intelligence provides better data driven business decisions as well as improving employee satisfaction. IT staff and analysts devote fewer time to response to requests from corporate users. Departments that previously couldn't view the personal data without consulting analysts or IT may either do so with minimal training. BI is built to be

scalable, so it can deliver data solutions to organizations that really need them as well as people that desire it. Hence, business intelligence delivers trusted and governed data. Data organization and analysis are aided by BI systems. Various departments' data is compartmentalized in conventional data analysis, and consumers must visit many databases to address specific information requirements. Sophisticated business intelligence solutions could now merge every one of these internal databases with external sources like client data, social data, as well as even weather data information into a single data warehouse. At the same time, departments from across a company may view the existing data [4].

Cloud business intelligence

Cloud-based business intelligence, often also known as cloud BI, is defined as the process of translating data into usable information in a cloud infrastructure, either partially or completely. Without the cost or trouble of physical gear, cloud BI provides organizations with the knowledge businesses ought to execute data-driven choices. Therefore, Cloud BI encompasses all SaaS-based business intelligence [7]. In this advanced world, majority of business users and consumers are adopting Cloud BI and there are several benefits to implement it. Presently, it is very tough to locate a company which doesn't employ at least one cloud-based tool, such as customer relationship management (CRM) software, online collaboration software, online data storage, as well as certain help desk software. It's reasonable that businesses will want a cloud-based bi solution as well. The transformation of data into valuable insights is at the heart of business intelligence. It all boils down to having the appropriate information to the correct individual time. The cloud is an excellent complement to this procedure. It allows you to process, retain, send, and access that information in a quick and scalable manner [7].

It's crucial to understand the constraints of desktop BI solutions before considering the advantages of cloud BI. If business users transfer data items from the cloud or client system to additional devices via desktop BI, IT loses transparency and controls, which could also lead to certification breaches and making version tracking challenging. With the help Cloud BI, all these issues could be solved as well as providing other benefits too. One of the benefits is that it is easy to deploy. Because cloud BI solutions are web-based, they are easy to set up and doesn't require the help of IT specialists. Also, Cloud BI is mobile friendly, and it can be accessed through mobile. Remote access is essential with the development of remote work. With cloud BI, users are no longer restricted to on-site access and desktop PCs. Instead, its

business intelligence (BI) products, such as Domo, can be accessed from any location or on any device. Your customized dashboards are always available to you [7].

1.1 Problem Statement and Motivation

BI systems have been an important component of enterprises for decades, but there is a lack of expertise and resources on how to successfully maintain these systems beyond the development stage [25]. As a result, BI deployment is frequently beset by several complex processes, obstacles, problems, risks, and downsides, making BI benefits difficult to realize. Integration of data from several sources or systems into an organization's data warehouse, for example, is a significant difficulty in BI implementation because it necessitates a large amount of effort. Apart from the benefits of BI, academics and practitioners have discovered that BI systems are costly, complex, resource-intensive, and difficult to manage. As a result, BI system deployment involves more than just buying hardware and software; it is a complex project requiring proper resources and infrastructure for a long period of time to stay up with management decision-making [26]. According to some reports, BI systems use billions of dollars each year. Despite this, more than half of these BI efforts are found to be ineffective [27]. [28] asserts that the benefits of a BI system can only be realized if it is well deployed. Now, several studies have revealed that understanding the factors for BI system implementation is critical, as they believe and believe that the high rate of BI failure is due to a lack of understanding of the factors for BI system implementation, and that these factors subsidies the realization of BI system benefits [29].

Many firms struggle to manage their BI projects due to their complexity and ambiguity [30]. One of the problems with deploying a BI system, according to [31] is the organization's readiness for BI adoption and implementation. Another challenge with BI deployment is the lack of information in the organization regarding the benefits and prospects of BI systems, as well as BI success criteria [32]. Furthermore, [32] stated that, despite the burgeoning market for business intelligence and the complexity associated with its deployment, the beneficial outcome determinants of BI system implementation remain poorly recognized. A few studies on BI implementation were undertaken, and it was discovered that most BI system implementation failures are due to a lack of understanding and preparedness to use BI systems in many businesses [33].

Another problem for BI deployment, according to [29], is how professionals evaluate them, because experts have varied perspectives on BI. Experts in data warehouses, for example, may regard BI as an add-on solution. Meanwhile, statisticians may consider BI to be a predictive analysis-based tool, and data mining professionals may consider BI to be an advanced support system that uses data mining techniques and algorithms. Organizations are hesitant to invest in BI due to differing opinions of experts. As a result, there are various external and internal aspects that influence the outcome of BI solutions, such as senior management support, technical resource capabilities, the sorts of industries in which the firms compete, data source quality, and investment funds [35]. Furthermore, integrating data from several sources or systems into the organization's data warehouse is a key challenge in deploying BI because it necessitates a great amount of effort [29]. As a result, managing the data warehouse is a difficult task. Despite the many advantages of a business intelligence system, there are still several issues with its deployment and use.

The aim of this thesis is to find out which components in a course where the students didn't do very well in terms of their CLO and cognitive C, practical P, and affective A as an indicator. In this matter, students' weaknesses in specific skills are analyzed to understand what are causes of the underlying issue. At these times, many students from higher educational institutions are dropping in terms of academic result. This project is essential due to vast number of students who are starting to deteriorate in academic performance during in times of pandemic. However, previous works on the existing project is not so accurate into analyzing student performance as students' weaknesses in certain skills or courses are remain unsolved or lack of recommendations to overcome the issue. In this matter, it is crucial that this project to be carried out to analyze deeply students' academic performance with an improved method to accurately find out which area of weaknesses categories are they into.

1.2 Objectives

First of all, the main objective of this project is to find out which area of weaknesses the students based on the CLO and bloom taxonomy. In fact, the objective is also to identify the problems related to students performance, which reported to have been deteriorating in this current pandemic where all students are required to conduct their learnings at home remotely. Students results are then analyzed by implementing the data analysis method to further propose

a discussion result to find out what are their backgrounds or education which contributes to the performance of students.

Secondly, the aim of the thesis is to study the software of business intelligence tools and to carry out a comparison to seek for an appropriate tool for the project requirements. Due to high availability of BI tools in the market, it is essential to find out which of these tools are suits the needs of the project. To pick the right BI solutions on the huge amount of alternatives on the current market, an exploration is required and find out which of the features are needed and what are the features which are not necessary for the project's needs. This comparison is vital as BI tools come with many different features and complexity like tools with basic capabilities and tools with advanced platform with more specialized features. Hence, the comparison will enable us to find out which of the listed features should be used in which the solution is adaptable easily, easy to handle, kind of support of functionalities being used, how good it can handle security matters, as well as most vitally the cost of the solution. Therefore, after finding out what is the right tool, the sample data of the students are collected and integrated into the BI tool to visualize the student data by drag and drop queries to form a graphical representation in a dashboard.

Third, the purpose of this research is also to analyse the results and outcome by collecting, modelling, and analysing the data with the purpose to extract insights and information which then provide a decision making to solve the matter. This is because when students' data are visualized in a form of graph, pie chart or bar chart, it makes the data easier to understand and pull insights from by highlighting the trend and outliers as well as removing noise from data and only focus on the important and useful information. This will enable us to find out what are the weaknesses of students in particular field of a certain subject and propose a recommendation to overcome the problem.

Lastly, the purpose of the research is to create a system of a business intelligence tool using Python with integrated matplotlib to make use the business intelligence tools to carry out the analysis. While using an existing tool to conduct the analysis, a simple system is developed to automate and optimize the tiring process, create customized visualization as well as creating machine learning modules and visualization of the predicted outcomes. In this case, the system is created to enhance the data visualization as it helps us to tell stories by curating data in a form which is easier to understand ad highlighting the trends as well as outliers.

1.3 Project Scope and Direction

In this section, the project scope is to explore business intelligence tools which are available in the market as a solution to overcome the issue of student academic performance as well as creating a system alongside to do the analysis. This outcome of this project will enable us to identify which area of skills or courses where students can't perform well, and the CLO of the students will be analysed, and a solution will be proposed to recommend suggestions that how to overcome this issue. As due to the covid pandemic, many students are struggling to keep up with their studies and studies have found out that students' academic performance is deteriorating and hence, the result of this project can guide us to monitor the performance of each CLO as well as forwarding towards improvement in the course which is taught.

1.4 Contributions

The development of this project will enable lecturers or tutors from higher educational institutions to gain better insight and understanding on students' main area of weaknesses by finding out if their factors of academic performance deterioration have a relation with their educational background. Hence, the outcome of project will enable lecturers of tutors to know what the knowledge or skills are required by end of the semester. Throughout this way, lecturers may change the method or the way of teaching in class where it will be easier and feasible way to convey vital information to students. In that case, more students will have a huge leap in academic performance especially in these times of critical covid situation in our country. In fact, the end of this project will reveal how much of most of the students could not perform satisfyingly and the factors that contribute to such outcome are to be studied and recommendations to be proposed to overcome this issue. Other than that, this project will show the significance of applying both Power BI and Python in educational field and how it would help in solving real world problems.

1.5 Report Organization

The details of this research are shown in the following chapters. In Chapter 1, the introduction will briefly explain what business intelligence is as well as identifying the underlying issues which is significant to solve. Therefore, the details in Chapter 1 will include the key details regarding the project with the purpose of conveying the main message of the project scope.

Secondly, Chapter 2 will conduct a literature review which enables us to find out what has been done and what has not been done by other publishers in the field of business intelligence before carrying out to propose a novel solution for the project. Furthermore, Chapter 3 will talk about system methodology for the development based project where it will include methodologies involved where it will be performed in several stages which are data collection and preparation, Business Intelligence software selection and system development, data analysis and findings, and then formulation of conclusions and recommendations. On chapter 4, it will further discuss the structure of the business intelligence system design in the form of system block diagram. This will provide us a better view of how the process of data visualization takes place in several stages. On Chapter 5, system implementation will takes place where the stages of data visualization will take place throughout the process. Aside from that, Python will be used a programming tool to develop the existing result generated by Power BI with the integrated platform, Jupyter Notebook with installed data analysis libraries like Matplotlib and Pandas. On Chapter 6, the results of the data visualization are discussed analyzed based on the graphical representation to find out which area of weaknesses the students are into, as well as proposing a recommendation to solve the problems identified based on data analysis. In Chapter 7, a conclusion is formed to summarize the results of the entire report and states the novelties of this project, as well as the contribution of this project.

Chapter 2

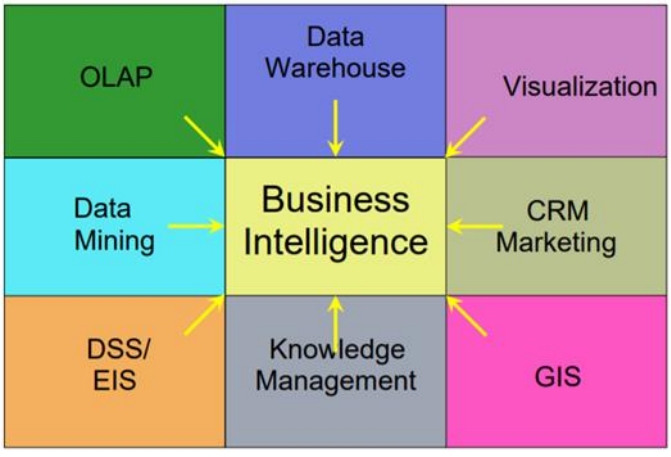
Literature Review

2.1 Previous works on business intelligence

2.1.1 Business Intellignce systems

Business intelligence (BI) is defined as the translation of information contained in knowledge that enables us to provide appropriate information to a specific user at the right time to support real-time decision making. As a result, BI brings together a variety of tools and technologies that enable data collection, integration, analysis, and visualization.

To deploy a BI platform, various intermediate processes, such as the creation of a Data Warehouse, must be completed, which is common in the development of this type of software tool (DW) [15]. In this case, the Kimball process was used to design, develop, and deploy the DW. As a result, several of these steps, such as task planning and expected results, describing the architecture that will follow the BI system, selecting and installing the most appropriate BI tool, building the Data Warehouse dimensional model, the Extraction, Transformation, and Loading (ETL) process, and finally developing the BI application, are known to be critical for the successful implementation of a BI system [15].



where: OLAP = on-line data processing, CRM=customer relationship management,

DSS= decision support systems, GIS = geographic information systems

Figure 2.1 BI system

In addition, selecting the software that is most appropriate for attaining the intended results is critical when developing a BI application. For example, it is vital to examine most of the software on the market and select the kind that gives the required and desired resources.

2.1.2 Business Intelligence Systems for Analyzing University Student Data

The research work is presented in few stages starting from collection of data together with preparation, selecting an appropriate Business Intelligence software and development of system, analyzation of data and findings and then formulation of recommendations to tackle the issue.

Throughout the first stage of the strategy, students' data is collected for data analysis. The data comes from the lecturer's records who led exercise sessions in one of the University of National and World Economy's university courses. The data is gathered using Excel software. The dataset used for the study included 575 records of student performance in the course's exercise classes over three academic years, from 2011 to 2013. During the exercise classes, the researchers use two Microsoft Office software products: Excel (a software tool for creating electronic tables) and Access (a software tool for creating relational databases), and they must pass two assessments (on Excel and Access) over the course of the two semesters. The average of the students' scores from the two tests is used to compute the student's overall score for the exercise classes.

In the second step of the strategy, the software tools for developing Business Intelligence applications are selected. Today's IT market offers a plethora of Business Intelligence tool packages. Unfortunately, many of them are commercial items that are quite expensive and accessing them for the purpose of research is normally prohibited. Fortunately, several of the industry's largest IT manufacturers have already launched their own academic programmes and are offering academic licenses to lecturers and professors who want to use the software for research and education. QlikTech International was chosen as the business intelligence solution due to its user-friendliness, ease of understanding and usage, and good graphical presentation of the processed data and information. The developed QlikView application is divided into three dashboards, each of which displays the study's most important data in the form of various

graphical elements such as bar and pie charts, performance indicators, tables, and list boxes (used as data filters on the various attributes).

The findings of the data analysis are presented and reported, conclusions are reached, and recommendations are made for future research possibilities and data collecting during the project's final step [35].

2.1.3 Business Intelligence for Educational Institution: A Literature Review

Research has also been done by managing organizations data to improve decision making. [36] studies that certain relevant literature in the field of business intelligence in higher institution. The methodology which was implemented were identification, selection, and feasibility. The very first step were by searching through sites like IEEE, SpringerLink as well as Google Scholar. On the other hand, the keywords which were used to search were 'Business Intelligence', 'Education in Business Intelligence', 'Higher institution business intelligence'. After the previous is done, literature selection was performed by selecting based on title and year of publication. The purpose of literature title selection is to identify whether it fits the field of business intelligence in higher institutions or not. At the same time, a condition in a range 2000 to 2020 is limited for the selection of literature. Once the literature was identified and selected, the leftover literatures are performed a review. Sustainability of according to research were examined as well as the availability of literatures in complete text. [36] also stated that chosen literature will be studied and there will be 14 articles to be discussed.

2.2 Strength on the previous works

First and foremost, one of the main strengths analyzed from the proposed previous work is that they made good use with the adoption of business intelligence tool for analyzing variable data could contribute to taking better and informed decisions. Other than that, the results of the data analysis are well presented in detailed in different forms of graphical objects which includes bar and pie charts, performance indicators, tables, and list boxes as well. Therefore, the decision making is based on dashboards, scorecards, and reports, which allows graphical sharing of vital data with the stakeholders. In this matter, a good BI dashboard allows you to monitor and measure business performance and metrics. BI dashboard tools offer real-time centralized

access to users enabling them to interact with and evaluate information, assisting them to make smarter, data-driven decisions.

Hence, an appropriate business intelligence tool is chosen which is QlikTech where it brings data and analytics together seamlessly with the only end-to-end, real-time analytics data pipeline.

2.3 Limitation of the previous work

Based on my findings on the proposed previous work, it is found that there are several limitations or weaknesses of implementation of BI tools in the education industry. One of the very main issues is that CLO (Course Learning Outcomes) of students is not analyzed. CLO is very crucial in terms of student academic performance as it observe or measure results that are expected after a learning experience rather than seeing from academic results. While a student's result does not accurately indicate the understanding of student in certain courses, CLO must be analyzed to gain insights of what the student have learnt throughout the semesters.

Furthermore, just a little amount of data is collected, resulting in low data quality. Due to the restricted amount of data available, this could have an impact on decision-making, as data correctness relates to how well it describes the real-world situations it is intended to describe. Inaccurate data will cause problems because it will lead you to make the wrong judgement. Because those findings and justifications are founded on erroneous data, the actions, and decisions you make based on them may not have the desired results. For example, data may lead a marketer to believe that the bulk of their clients are females in their twenties. Because their consumers are largely males in their 40s, and their data is inadequate and erroneous, they will wind up targeting the wrong demographic with their marketing, resulting in poor decision-making.

In fact, I have also found out that the proposed work did not do a precise comparison of BI tools available in the market. It is crucial to carry out an analysis of the comparison of business intelligence tools to find out which are the necessary and which ones are the insignificant ones. It is also important to know what exactly the organization needs or requires. By doing a comparison, you will know BI tools will fits your requirements. Hence, it's always vital that the BI tools available in the market listed out alongside with its functionalities or features and

carry out an analysis so that an appropriate tool will be chosen to do better decision making effectively.

Finally, it is discovered that the proposed earlier work does not include python for data analysis. Using Python and a business intelligence platform, we can generate stunning, personalized visualizations, allowing us to create machine learning modules as well as visualizations based on predicted outcomes. It can swiftly develop and manage data structures thanks to the use of Python. Pandas, for example, offer a variety of tools for manipulating, analyzing, and representing data structures and complex datasets. As a result, time series and more complex data structures such as merging, pivoting, and slicing tables will be used to develop new perspectives and views on current data sets.

2.4 Proposed solution

To begin, I propose that the problem be solved by selecting and executing the appropriate business intelligence software tool. To pick the best BI tool among many available solutions on the market, you must first determine what kind of features are necessary or compulsory needed and which are not critical to the company's demands. So, should we go with a basic solution for everyday use or go for a more advanced platform with highly specialized and advanced features? The solution is not direct, and towards the completion of this article, I will expand on and justify a few of the options.

Second, most today's BI solutions (>90 percent) provide basic (standard) capabilities such analyzation of data, ad-hoc reports, dashboards, visualization of data, performance metrics, ad-hoc query, ad-hoc analysis, and key performance indicators (KPIs). When selecting a BI solution, we must keep in mind that certain complex platforms with more advanced or specialized capabilities may not have all the basic features. As a result, there are various compromises between advanced functionality and the breadth of business intelligence tools. It's not unheard of for a product to make use of basic or advanced capabilities from other solution providers.

First and foremost, I will offer a full inventory of advanced features and functionalities in Business Intelligence systems, from which users should anticipate or demand the best solution available. The following are some of the advanced features:

- **Master Data & Data Quality Management** – To make accurate choices, solely reliable data will be used. MD/importance DQMs are based on the demand for high-quality, full, and accurate data. In this context, BI solution vendors have placed a strong emphasis on master data management as well as data quality, and the quality of data cycle. The data quality cycle encompasses every stages of offering high-quality data to organizational users, including metric identification, evaluation, data mending as well as cleaning, classification, archival, discovery, as well as evaluation [8]. Users could track and data quality handled through a single, integrated source that contains the master data, rather than by many, disjointed databases.

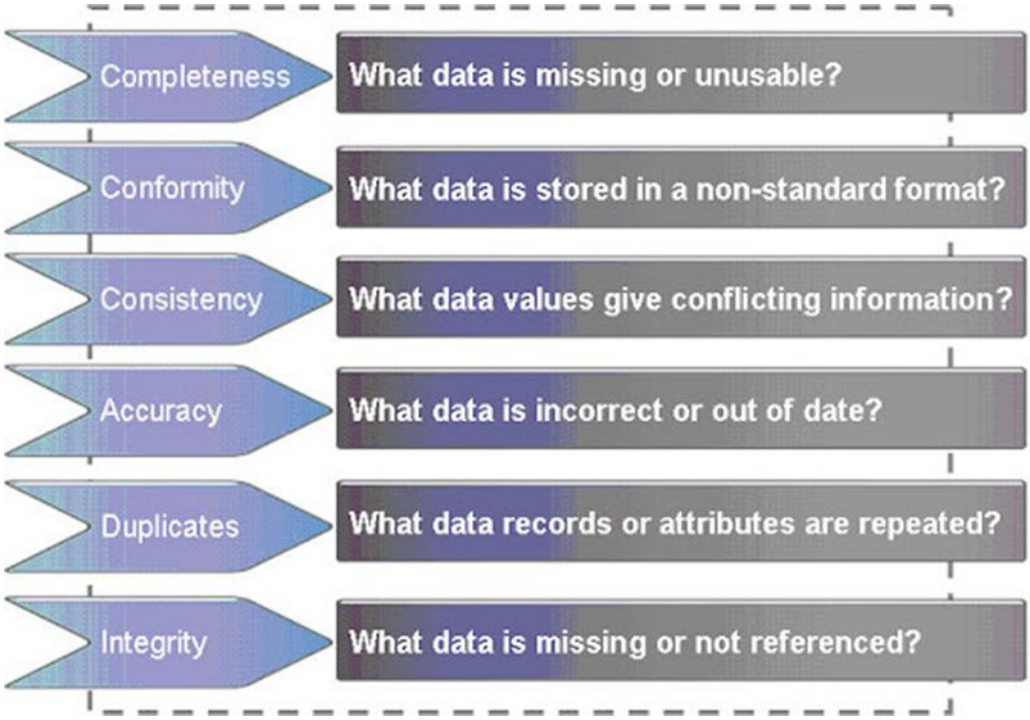
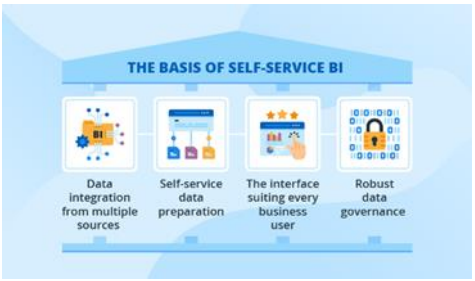


Figure 2.2 Master Data and discovery

- **Data Visualization** - Whenever it comes to data preparation for decision-making, BI systems with extensive data discovery as well as visualization capabilities can help eliminate the need for data scientists. The use of exploratory data discovery and visualization together can help you overcome challenging data analysis barriers. Hence, machine learning is also being merged into data discovery technologies in purpose for assisting business analysts with every stages of the process, along-with devising to investigation to presentation. Tools for integrating numerous sources, filtering,

improving, and altering data to build novel data sets for use in analysis visualization or up to date analytics should all be included in the solution.

- **Self-Service BI** – Consumers formulate and maintain reports and analyses within the context of a well-tested and managed structure and tool set [9]. As according to Gartner, business customers with self-service analytics capabilities will outperform experienced data scientists. It will also be able to assist corporate users in learning ways to apply and gain advantage from effective analytics as well as business intelligence technologies, resulting in great business outcomes [9]. Self-service BI increases efficiency and minimize the time to discovery, but it should not jeopardize the quality or effectiveness of the outcomes. It must accomplish several objectives, namely maintaining that greater speed does not threaten data trust and facilitating access to the data and understanding. Improved data security and governance should assure data accuracy and reliability, and the amendment should include all stakeholders and parties responsible. So, it is crucial to fulfil architecture and requirements have been met while maintaining a balance of flexibility and efficiency [10].



Self Service BI

Figure 2.3

- **Data Governance** – This technique guarantees that data satisfies specific criteria as well as business needs when entering data into the database. Countless businesses have endured massive losses in the recent because of a lack of data security [11]. Data governance comprises designing a data framework that takes regulating policies and procedures, evaluating, and maintaining data capital while involving people, stages, as

well as technology into account to avoid these mistakes. Data governance that adheres to regulations such as the Data Protection Regulation (GDPR) is also a great feature for BI solutions [12].

Figure 2.4 Data Governance

- **Cloud BI/Data Management** - As stated by Gartner, numerous companies are going to shift an important portion of its data undertakings to the cloud before 2021 [13]. As a result, most Business Intelligence and data management vendors currently provide a cloud-based service. Despite cloud Business Intelligence as well as data management have much of the same functional capabilities as its on-premises versions, they usually come at a lesser cost and put fewer load on IT departments

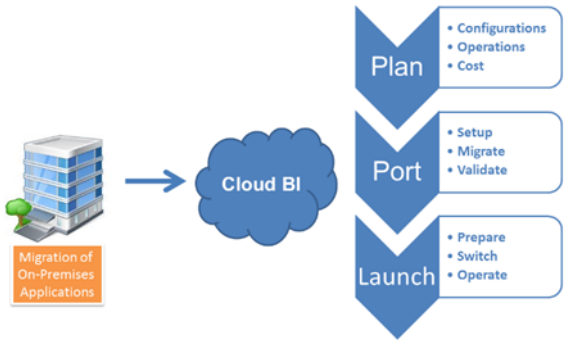


Figure 2.5 Cloud BI

- **Mobile BI** - The advantages of letting employees or supervisors to make choices regardless of where they are situated are considered by organizations. Because of its volume to give powerful and precise data visualization within structure of graphs including charts, dashboards, and performance measures, notepad and mobile devices are improving effectiveness of mobile Business Intelligence, which had seen significant commercial adoption [14].

- **Deep Learning-Powered Analytics** – Furthermore, Deep Learning is defined as a sort of machine learning in which a neural model is trained into doing human-like activities like speech recognition, responding natural language inquiries, and distinguishing things and activities in pictures and videos. As per according to Gartner, deep learning would be the dominant option strategy for data applications before 2023. [9].

- **Real-Time Analytics** – Data is evaluated using logic and mathematics to generate insights that assist individuals in making improved major decisions. In certain circumstances, real-time analytics implies that now the analysis has been conducted within seconds or minutes of fresh data being received [9]. In several firms, faster data reporting and analysis is a roadblock. To enable quicker and more fact-based organization's decision, companies are increasingly depending on data from transactional systems. Real-time analytics in BI may augment a company's existing Business Intelligence strategy by offering novel, crucial data findings [15]. Enhanced change data capture (CDC) technologies are being used by organizations to identify as

well as gather data modification and data structures when they occur, as well as to alert users of such changes.

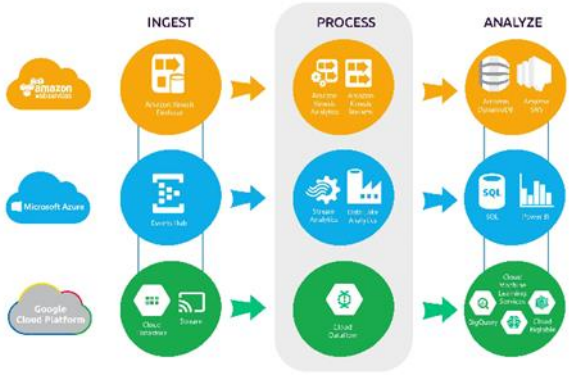


Figure 2.6 Real time analytics

- **Agile BI Development** – Iterative development and information storage commoditization are valued in this flexible and efficient system. This allows companies to easily respond to different business requirements and minimizing total cost of ownership. Agile BI involves fast prototype cooperation between business and IT, allowing businesses to boost production speed whilst adapting to business requirements more swiftly. The agile BI development process is complemented by project administration, where it encompasses arrangement, requirements collecting, development, and iterative functional, regression, as well as usability evaluation [15].

- **Data Warehouse Modernization** - New technical and business problems keep businesses on their toes. They see the potential in using different data warehousing architecture methodologies along with integrating extra technology options including in-memory processing, cloud storage, including data storage automation instruments. While competing with novel and less expensive implementation choices coming from external service providers, IT must address evolving analytical requirements. Multidisciplinary strategies must always be employed to suit the rise of requirements of the market to optimize the business value of data. Therefore, it is time to assess how new hardware and technology will improve and simplify corporate operations by comparing obsolete data warehouses to present regulations [15]. Conventional data

warehouses aren't developed to withstand the fast growth of data and the variety of data kinds in which it generates, including big data. As a result, they aren't built to continue alongside the changing needs of end users as well as the programmes which depend on them.

- **Data-Driven Culture** - Modern manpower is so much more capable than ever before. Therefore, the stages of education, engagement, as well as knowledge was at an all-time high. A data-driven culture allows corporations to obtain every necessary data and continue to exploit their values. Also, data-driven culture assists to elevate company culture to the next level of performance by obtaining relevant KPIs rooted in business. Also, transparency of data obtained KPIs is regarded as the part of aspect of the data-driven civilization method of an organization [36].

- **Data Preparation for Business Users** - Data cleaning, structuring, and enriching are called the stages of cleaning, structuring, and enriching data. These are the steps for exploratory and advanced analytics. Hence, preparation of data is all about giving people the tools they need to make their data satisfy their agenda without having to use IT. Technology that is easy for people to use and understand, as well as smart coaching and quick outcomes, were needed with the purpose to get data preparation to business people [15].

- **Integrated Platforms for BI/Performance Management** - Many businesses want to use the same tool for both BI and project management. The trend has led to one of the most common and long-term things in the business intelligence market, as well [15].

- **Embedded BI and Analytics** – In operational applications, intelligence is being applied. Embedded business intelligence and analytics refers to the incorporation of business intelligence software aspects (such as dashboard outlining, visualizing data, as well as analytics instruments) into non-business intelligence applications. Customers gain from a far more pleasant and user-friendly user experience when using integrated business intelligence, which is also its key benefit over solutions which require two distinct platforms [16].

- **Data Storytelling** – Data visualizations, infographics, dashboards, data presentations, and a range of other tasks are all handled by data storytelling. It's not only about making aesthetically appealing graphs. Data storytelling is described as a method of expressing data insights that incorporates three key components: data, visuals, as well as narrative. By integrating the proper visuals and narrative alongside the correct information, a data story which able to affect as well as drive change may be generated [17].

- **Using External/Open Data -** Consumer, market, climatic, regional, and demographic data, as well as similar analytical findings, may all give significant information. These and many other kinds of data may be provided to firms by other Business Intelligence generalists, specialist service providers, or data exchange platforms. Also, open data is always utilized to create models that are based on focused research [18].

- **Analytics Teams/Data Labs** – Described as independent business segments developed specifically to get a company's data science programme off the ground. They'll need to invest in novel technology to keep, process, as well analyses information. When analytics improves, so does the importance of solution implementation and efficiency. It's a novel stumbling barrier for software solution providers, and then it involves the creation of new organization structure to integrate data labs, IT departments, and business divisions [19].

- **Visual Design Standards** – Described as a method of presenting significant data in a manner that is both effective and efficient. The tendency of visual design qualities has evolved in the past three years because of the rising requirement to analyses massive volume of data to remain competitive and give its most direct results. Support for visual design standards is seen as a need that BI vendors must meet in software selection procedures [19].

- **IoT Analytics -** Because as internet of things (IoT) advances with novel sensor, device, as well as wireless technologies, significant economic value must be identified in analytics instead than new hardware. Sellers are beginning to provide these services to their customers, as well as extending their service portfolios into new business sectors. IoT data necessitates real-time data processing. In addition, the variety of IoT data

mandates the use of new architecture, tools, and procedures to effectively process, retain, as well as evaluate IoT data [37].

- **Big Data Analytics** – gives the capacity to examine large, varied, and velocity data sets obtained from different sources, such as textual, sensory, geographical, and clickstream data. Inside these arrangements, big data analytics should be able to analyses massive datasets in real-time or near real-time, which includes modelling, visualization, forecast, and improvement. Businesses are turning to big data analytics to help them make better decisions and improve their processes [20]. Structured information analytics, sentiment analysis, web analytics, audiovisual analytics, social network business intelligence, and systems analysis are all examples of big data analytics.

- **Data Lake** – Based on James Dixon's original proposal, a huge storage of data in its inherent state of unstructured and structured data [21]. Without any purification, standardization, remodeling, or modification, it is not feasible to keep and analyses data in its raw and original form, direct from the data sources. Data lakes enable for ad hoc queries, data analysis, and findings analytic searches since information management and architecture might well be adjusted on the fly during runtime. [38]. The goal of a data lake is to keep all information on the same system, encompassing relational, non-relational, including big data. [22].

- **Edge Computing and NLP** Its purpose is to make NLP computation more responsive to user demands. BI is paying attention to efforts to understand user behavior, attitudes, and emotions [23]. The usage of NLP in combination with DL is required to comprehend verbal and nonverbal messages. AI has made significant progress in this direction recently, and even more BI solutions that include speech or language understanding technologies are expected to develop soon [24].

2.5 Analysis and discussion

After listing the features of BI tools, we'll conduct an examination of open-source business intelligence solutions available in the market and compare those features as well as

functionality. We've whittled down the number of business intelligence products to three. Using methodology to enable organized study and comparison of BI solutions will provide greater insights to the present status of industry as well as assist in the selection of a business intelligence tool that includes all the aspects. Furthermore, the research conducted allows us to detect current trends in BI solution development. This study can be used as a road map for selecting a business intelligence solution that meets the needs of the enterprise. As a result, we produced a feature-to-BI solution cross-comparison table and a feature pyramid to reflect the latest trends in BI solution development, with a focus on unique and future features.

Table 1 shows three different BI tools which are currently on the market, alongside a link to each of the descriptions. Most of these tools are currently the most popular business intelligence solutions. As a result, according to user reviews, these solutions are the top business intelligence and analytics software of all time. Essentially, the list comprises simple-to-use systems that enable a wide range of analytic work capabilities without requiring much IT participation, allowing users to quickly embrace BI tools. For example, tools can pre-define data models as a requirement to analysis, and in some situations, they can even generate a reusable data model automatically. As a result, we will assess the functionality of the BI tools listed in Table 1 by comparing them to the proposed basic and advanced features of BI tools in the next section.

Business Intelligence tools	Web Page
Power BI	https://powerbi.microsoft.com/en-us/
Tableau	https://www.tableau.com/
IBM Cognos Analytics	https://www.ibm.com/products

Table 2.1: BI tools chosen for the comparison

2.5.1 Overview of Power BI

Microsoft Power BI is a set of tools, software services, and connectors that work together to transform disparate data into visually appealing and interactive insights. Power BI can operate

with both simple data sources like as Microsoft Excel and more complex ones such as cloud-based or on-premises hybrid data warehouses. Power BI also allows you to connect to your data sources, visualize, share, and publish your findings to whomever you want.

Besides, Power BI is simple and fast enough to connect to an Excel workbook or a local database. Hence, it can also be robust and enterprise-grade in which it's ready for extensive modeling as well as real time analytics. This means it can be implemented in various environments from a personal report and visualization tool to the analytics and decision engine behind group projects, divisions, or entire corporations.

Due to Power BI is a Microsoft product and has built in connections to Excel, there are lots of functions that will be familiar to an Excel user.

Figure 2.7 Power BI

2.5.2 Overview of Tableau

Tableau was established in 2003 because of a computer science project at Stanford that aimed to enhance the flow of analysis and allow the data more accessible to people through visualization. Co-founders Chris Stolte, Pat Hanrahan, and Christian Chabot developed and patented Tableau's foundational technology, VizQL where it is visually expresses data by translating drag-and-drop actions into data queries via an intuitive interface. Since their foundation, they have continuously invested in research and development at an unrivaled pace, developing solutions to assist anyone working with data to receive faster and find out unanticipated insights.

Furthermore, making machine learning, statistics, natural language, and smart data prep more usable and dependable to assist human ingenuity in analysis is part of this. They also offer not only a comprehensive, integrated analytics platform, but also proven enablement resources to help companies implement and scale a data-driven culture that delivers resilience and value through compelling outcomes. In 2019, Tableau was acquired by Salesforce, but its aim remains the same to help people see and understand their data. Currently, companies of all sizes, from non-profits to multinational corporations, and from every industry and department are using Tableau to empower and encourage their employees to make data-driven decisions.

Figure 2.8 Tableau

2.5.3 Overview of IBM Cognos Analytics

IBM, a web-based reporting, and analytic tool is designated as Cognos Business Intelligence. Its goal is to do data aggregation and generate thorough reports that are easy to understand. Graphs, several pages, different tabs, and interactive prompts can all be stored in reports. Web browsers or mobile devices such as tablets and smartphones can be used to access these reports.

Aside from that, Cognos gives you the option of exporting the report in XML or PDF format or viewing the report in XML format. As a result, you may schedule the report to run in the background at a specific time, which saves time in terms of monitoring the report on a regular basis because you don't have to run it all the time.

Also, IBM Cognos has a wide range of capabilities and may be considered enterprise software because it provides a versatile reporting environment that is suitable for both large and small

businesses. Power Users, Analysts, Business Managers, and Company Executives will find it useful. Analysts and power users want to be able to produce ad hoc reports and different views of current data. Business executives demand dashboard layouts, cross tabs, and visualizations to examine summarized data. Cognos makes both options available to all users.

Figure 2.9 IBM Cognos Analytics

2.6 Feature to BI solution comparison table

At this part, 3 business intelligence tools are evaluated against their functionalities using the product/functionality matrix produced in Table 2 to create a product/feature matrix. As a result, the green checkmarks will show tools that are integrated, while the red 'x' will represent tools that are lacking. Furthermore, these aspects enable us to examine the features of a given BI solution and compare them to those of other systems. In this case, it will provide us an understanding of the present state of the BI market and assist us in identifying solutions that make use of advanced features.

Based on the analysis, it is revealed that over 90% of business intelligence tools provides data analysis, ad hoc analysis, dashboards, ad hoc query tools, KPI’s, and performance metrics that belong to the basic group of features. Other features include data management, data exploration, self-service BI, saas BI, mobile BI, deep learning, and big data analytics. It's important to note that, while most suppliers offer the entire range of capabilities in their products, they isn't quite as practical or high-quality as capabilities found in other solutions.

The research also revealed that most solutions allow for the merging of certain functionalities by other companies. Advanced visualization capabilities of one solution, for example, able to be combined to different solutions, purpose for delivering the correct visualization solution to their consumers. Instead of developing their own solutions, sellers are expected to integrate the greatest-rated solutions from other providers to seller's products.

Functionalities	**Power BI**	**Tableau**	**IBM Cognos**	**My proposed solution**
Master Data	✔	✔	✔	✔
Data Discovery	✔	✔	✔	✔
Self-service BI	✔	✔	✔	✔
Data governance	✔	✔	✔	✔
Cloud BI	✔	✔	✔	✔
Mobile BI	✔	✔	✔	
Real Time Analytics	✔	✔	✔	✔
Agile BI development	✔	✔	✔	✔
Data warehouse	✔	✔	✔	✔
Data driven culture	✔	✔	✔	✔
Data preparation for business users	✔	✔	✔	✔
Integrated platforms for BI	✔	✔	✔	✔

Apps Integration	✔	✔	✔	✔
Embedded BI and analytics	✔	✔	✔	✔
Data Storytelling	✔	✔	✔	✔
Using external/open data	✔	✔	✔	✔
Analytics team	✔	✔	✔	✔
Visual design standards	✔	✔	✔	✔
Big data analytics	✔	✔	✔	✔
Data Lake	✔	X	✔	✔
IoT analytics	✔	X	✔	✔
Augmented analytics	✔	X	✔	✔
Deep learning powered	✔	X	✔	✔
Edge computing and NLP	X	X	X	X

Table 2.2: Business Intelligence solutions and the capabilities they offer

2.7 BI feature pyramid

In this section, we'll go over the feature pyramid, which demonstrates BI functionalities via those frequencies with which they appear in BI solutions (see Figure 2.7). The most fundamental features, which are implemented in most Business Intelligence tools, are placed at the base of the pyramid, whereas rare functionalities are at the highest peak.

The base layer, first and foremost, embraces elements that we classed as fundamental, and which are found in 60 percent of BI systems today. The second tier includes augmented analytics, IoT analytics, as well as data lakes, where they were employed in around a quarter of business intelligence solutions. Also, nearly 40% of the evaluated solutions said organizations had began using augmented analytics in a more advanced form. Solutions gives business decision makers innovative visual access to complex data, backed by clever recommendation algorithms that help them comprehend hidden and useful insights. As a result of the use of automation and embedded intelligence, the time required for data preparation has been drastically reduced. The investigation, however, reveals that every of the solutions are strongly focused on augmented analytics. Meanwhile, deep learning-powered analytics make up the third layer. The report shows that about a quarter of the analyzed Business Intelligence solutions provides alternatives of deep learning analytics.

Finally, developing capabilities such as edge computing and natural language processing (NLP) enable Business Intelligence solutions to acknowledge human perceptions, behavior, as well as feelings. Even though none of the systems we examined give such possibilities (0%). However, based on current patterns, we can confidently predict that these features will be included in future BI editions.

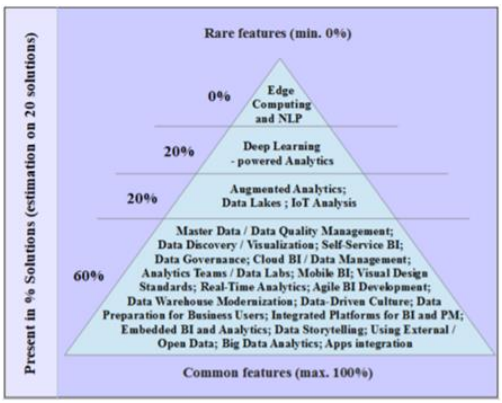


Figure 2.10 The pyramid of BI tools features

2.8 Business Intelligence Tools – Assessment Results

After analyzing the business intelligence tools based on Table 2, it could be concluded that the most appropriate tool to analyze student performance would **Power BI**. Throughout my findings, I found out that Power BI is an easy-to-use tool for data analysis and visualization due to the user-friendly interface unlike other tools which adopts highly complex layouts. Due to this matter, it's safe to say that this tool is a perfect starter choice for beginners who has little or no experience in data analytics. In fact, this tool performs at great speed and better in terms of performance when the data volume is limited. While this tool tends to drag slowly while handling large data, it fits anyone who just need it analyze small amount of data. As for my project, no bulky data is required, therefore this tool suits perfectly for simple visualization. Comparing to Tableau which able to handle bulky volumes of data, at the same time it has high complexity layout and normally made for professionals to use. In this case, it wouldn't be a beginner friendly tool for starters who begin to adopt business intelligence in the market.

As previously said, Power BI's user interface is easy, allowing it to readily interact with other Microsoft products such as MS Access, MS Excel, and others. The Power BI interface is simple to understand and grasp, even for people with no prior expertise or experience in the field of business intelligence, and it is user-friendly, allowing you to use more efficiently. The report, data and the model view are all available on the left side of the canvas in Power BI desktop.

Furthermore, compared to Tableau, Power BI offers a variety of data sources but has limited access to other databases and servers. Microsoft Excel, Text/CSV, Folders, MS SQL Server, Access DB, Oracle Database, IBM DB2, MySQL database, PostgreSQL database, and others are among the data sources. Because the only database I needed for my project was a MySQL data source, this programme will provide enough capabilities and functionality to complete my project. When it comes to simplicity of use, Power BI has an advantage because its user interface is built on Microsoft 365, which most of the users are familiarize with.

Furthermore, Power BI has a drag-and-drop capability that is simple to utilize. It has features that make data more aesthetically appealing. It has many detailed and appealing visualizations that can be used to generate reports and dashboards. You may ask questions about your data with Power BI, and it will provide you with useful insights into where the benefits are found. Power BI, on the other hand, has the advantages of Microsoft Business Analytics, which includes platforms like Azure Machine Learning, SQL Server-based Analysis Services, real-

time data streaming, and many Azure databases. This aids in the comprehension of data as well as the analysis of trends and patterns in the data.

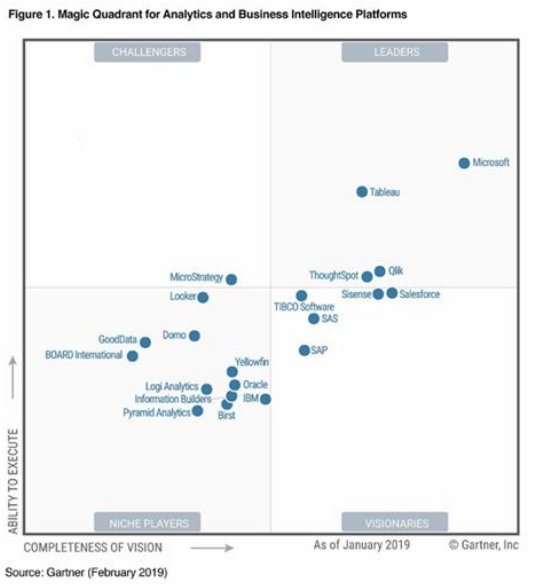


Figure 2.11 Magic Quadrant for Analytics and Business Intelligence Platforms

Although Power BI is still new in the business intelligence field, the magic quadrant shows that it is quite competitive in the market when compared to Tableau or IBM Cognos Analytics. Power BI is, in fact, frequently used by corporate executives and organizations. Within the Leaders' quadrant, Microsoft Power BI is positioned furthest to the right in terms of completeness of vision and furthest up in terms of capacity to execute, according to my study. One of the reasons is that it is cloud mature, relying on one of the fastest growing BI clouds. In fact, Power BI Mobile is without a doubt the best-in-class mobile experience, allowing end-users to easily connect to and interact with both cloud and on-premises data, and take it with them wherever they go.

To summaries, all the standards listed in Table 2 were chosen with the education industry in mind. As a result, Power BI was chosen to construct a BI application and demonstrate the efficacy of Business Intelligence (BI) tools in analyzing student performance, based on these requirements. Power BI Suite covers all the standards and offers more benefits than other BI products like as Tableau and IBM Cognos Analytics. Furthermore, it is a tool with a good and better interface, as well as a range of distinct representations. However, we must keep in mind

that the tools discussed and reviewed in this section are continually expanding and changing, and functionality that they currently lack may become available soon.

Chapter 3
System Methodology

Proposed Method/Approach

This section will include the methodologies involved where it will be performed in several stages which are data collection and preparation, Business Intelligence software selection and system development, data analysis and findings, and then formulation of conclusions and recommendations.

During the first stage, data collection and preparation, the sample data will be created and prepared for analysis. In this phase, it starts off with the identification of the research domain. UCCD9999 Website Design and Development methodology module was picked for the research domain and the course learning outcomes (CLO) for this module is studied. In short, the course has an objective to teach students regarding the expert system development methodology by utilizing an expert system development life cycle. Therefore, the developments of the CLO for this course are based on the Bloom Taxonomy which is as shown in Table 3. According to the Bloom taxonomy, the three domains are consisting of cognitive domain, affective domain, and psychomotor domain in which they are applied to CLOs of the course. Thus, a few evaluation methods are implemented to test the students' understanding on the materials taught throughout the semester. The assessments are consisting of midterm (20%), assignment (20%), practical test (10%) and final examination (50%). Aside from that, this subject will give a brief introduction to the constructing web application which is progressive as well as working web application with the implementation of knowledge and skills in HTML, JavaScript, Bootstrap, CSS, and event handing programming. The table below demonstrates the CLO of students at the end of the semester.

CLO	Course Learning Outcome	Bloom taxonomy
CLO 1	Create static, dynamic, and interactive web pages using different technologies	Cognitive domain
CLO 2	Apply knowledge of event-driven and object based programming concepts in client-side web development	Cognitive domain
CLO 3	Apply emerging front-end web technologies in the design and development of web applications	Cognitive domain
CLO 4	Able to develop responsive as well as progressive web applications	Psychomotor domain

Table 3.1 CLO for UCCD9999 Website Design and Development

The software which will be used to store the database as data collection is Microsoft Excel. The dataset will include a total 239 number of students who will be enrolling in this course. In this stage, the marks of the students are collected based on their respective coursework (midterm, practical test, and assignment) totaled up 50% and a final exam which is 50%. The following results of students will be imported into Power BI to undergo interactive data visualization and analytics. By doing this, it allows us to visualize data into meaningful insights to explore solutions to overcome the issue faced which to find out what area of weaknesses of students in. Aside from that, alternative software, Python will be implemented as well to visualize data into graphical presentations with pre-installed matplotlib and pandas.

During the second stage, software tool is selected for the Business Intelligence to analyze the student performance. Nowadays, there are lots of Business Intelligence tools available in the IT market. Unfortunately, most of them are very costly and very limited access for research purposes. Luckily, there is an alternative solution where Power BI will be chosen as the appropriate tool for this project. It is very easy to use and navigate and it packs with limited features which are enough to utilize for basic usage. Similarly, to other BI tools, it offers data preparation and discovery, amazing and interactive dashboards, and rich visualizations and most importantly it is free for the basic features. Due to its simplicity, the data can be visualized in just a few clicks, and it has a user-friendly interface. Currently, Power BI is one of the best tools available on the market which adopted by many business leaders around the world and widely used by people. Based on the above-mentioned reasons, it is concluded that Power BI will be used a business intelligence tool. Hence, a simple system will be developed as well to perform the data analysis other than using the existing tool. In this development of the system, Python will be top pick for data analysis due to its simplicity and readability and at the same time offering plethora of useful options for data analyst. Hence, it is a very good programming language for novice or beginners to start off with, where you can use it its relatively easy syntax to form effective solutions for complex situations by using a few lines of codes. Therefore, Python will integrate with Power BI to write codes of script to visualize the graphical visualization on the BI tool.

In the next stage, data analysis and findings, Power BI and Python will be used as the business intelligence tool to conduct the visualization of student results. Anaconda is downloaded to use Jupiter Notebook as it has integrated Python 3.6 with pre-installed matplotlib. The results of the students are analyzed and organized in several dashboards in which the most essential information from the data analysis is shown in the form of different graphical objects such as bar and pie charts, performance indicators, tables as well as list boxes.

In the last stage, formulations and recommendations, the results and outcomes of the data analysis are shown and presented and then described based on the finding. Weaknesses of students in the course are identified and recommendations are drawn to overcome the issue which is to be proposed and conclusions are drawn as well for future research opportunities and ad data collection.

3.1 System Requirement

3.1.1 Hardware

The hardware which will be involved in this project will be a computer. A computer will be issued to carry out the data analysis and to perform data visualization with Power BI and Python. The specifications of the laptop needed to carry out the project is as shown as below:

Description	Specifications
Model	Asus A510U series
Processor	Intel Core i5-8250U
Operating System	Windows 10
Graphic	NVIDIA GeForce GT 940MX
Memory	12GB DDR4 RAM
Storage	1TB SATA HDD

Table 3.2 Specifications of laptop

3.1.2 Software

The software which will be involved in this project will be Power BI. The business intelligence tool is issued to run reports and surface insights based on a data collected. Therefore, the requirements of the software needed to carry out this project is as shown as below:

Description	Requirements
Operating system	Windows 8.1 / Windows Server 2012 R2, or later
Memory (RAM)	At least 2 GB available, 4 GB or more recommended
Display	At least 1440x900 or 1600x900 (16:9) required
CPU	1 gigahertz (GHz) 64-bit (x64) processor or better recommended

Table 3.3 System Requirements of Power BI

The next software which will be included in this project will be Python. The programming software is issued to manipulate, analyze, and represent data structures and complex dataset as well as visualizing the graphical objects. The requirements of the software to perform the tasks is as shown in below:

Description	Requirements
Operating system	Windows 7 or 10
Memory (RAM)	At least 4gb ram
Storage	At least 5GB disk free
CPU	Minimum Intel Atom® processor or Intel® Core™ i3 processor

Table 3.4 System Requirements of Python

3.2 Activity Diagram

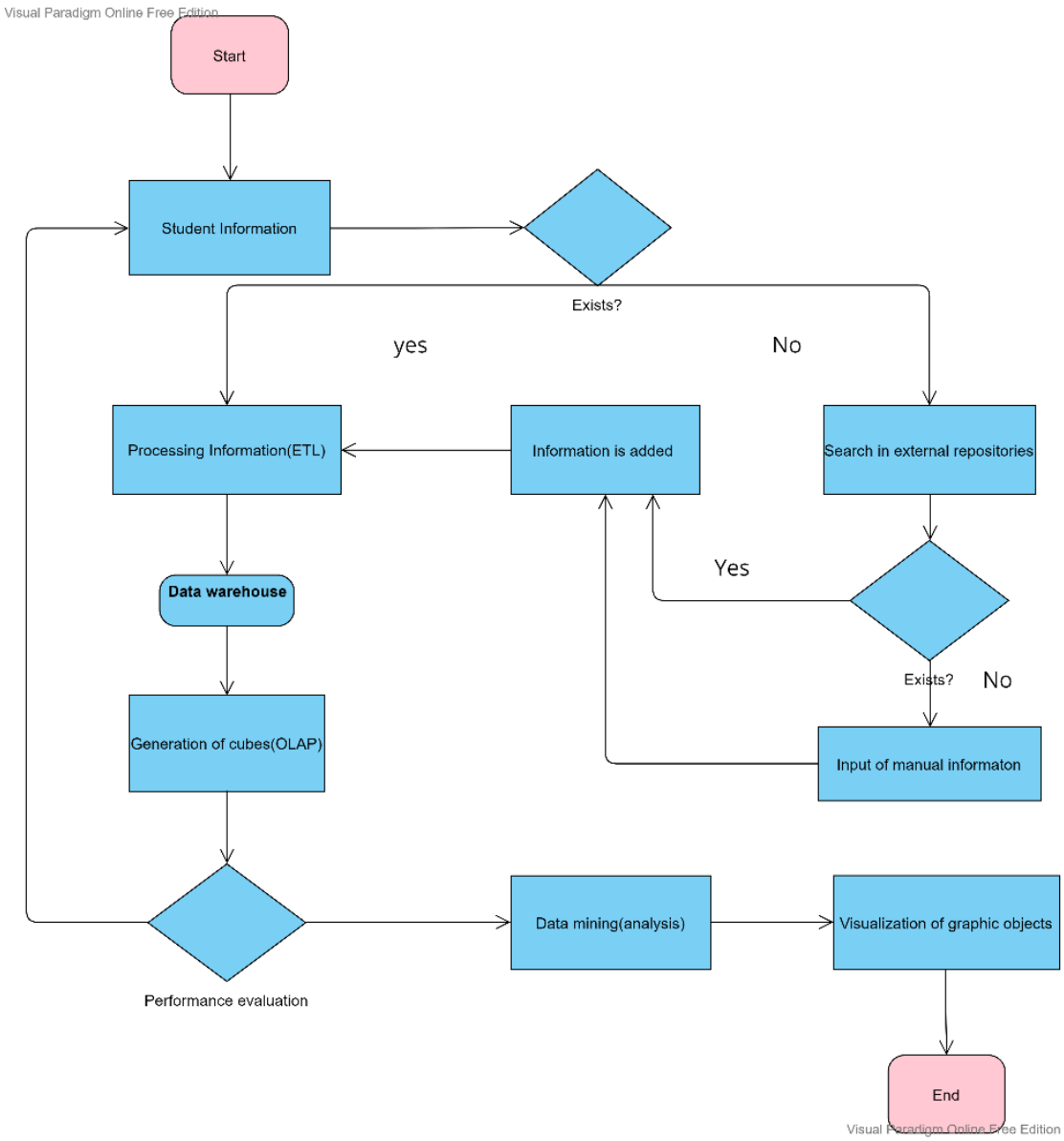


Figure 3.1 Flowchart of BI system

3.3 System Architecture

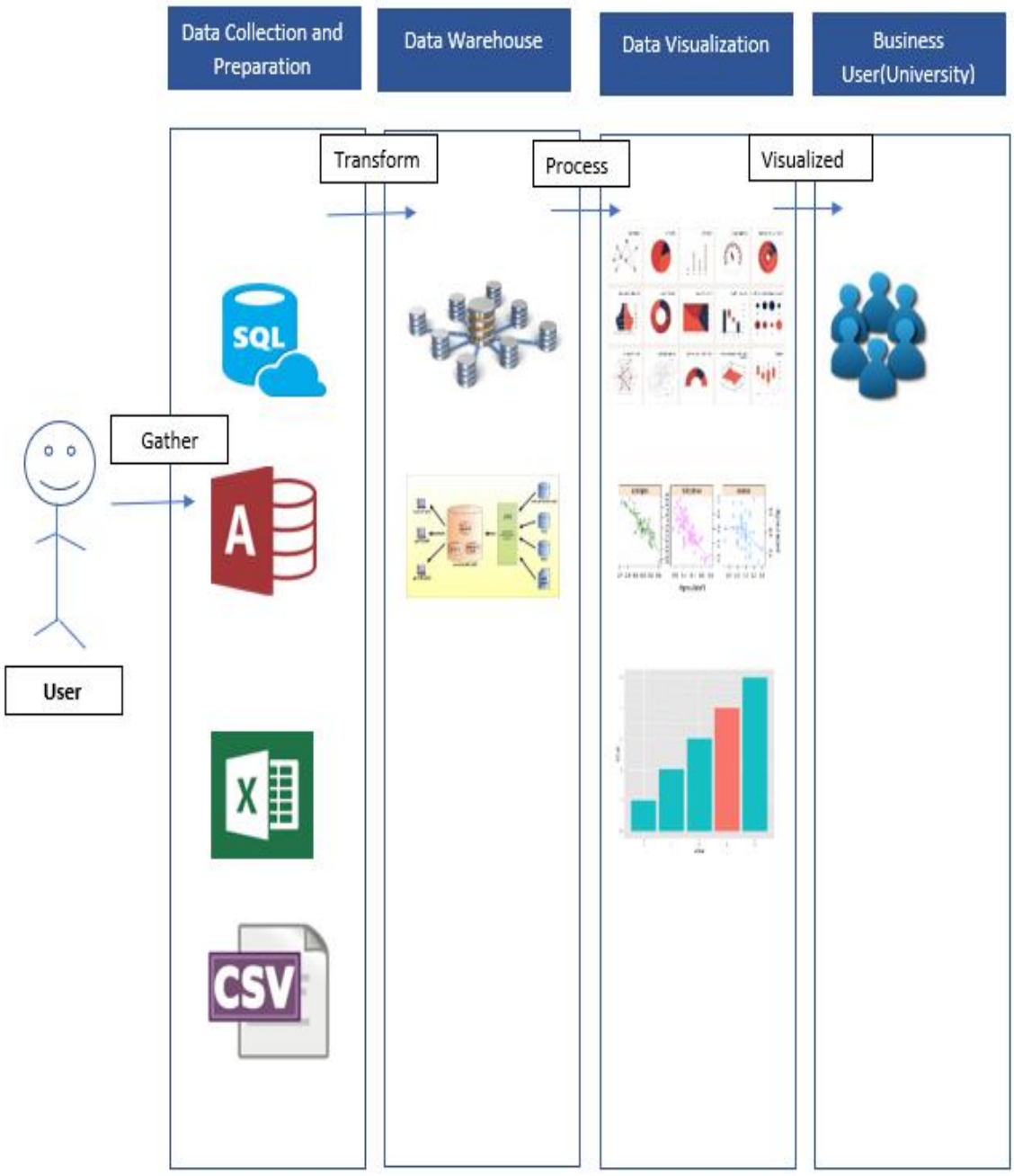


Figure 3.2 System architecture of data analysis

Chapter 4

System Design

4.1 System Block Diagram

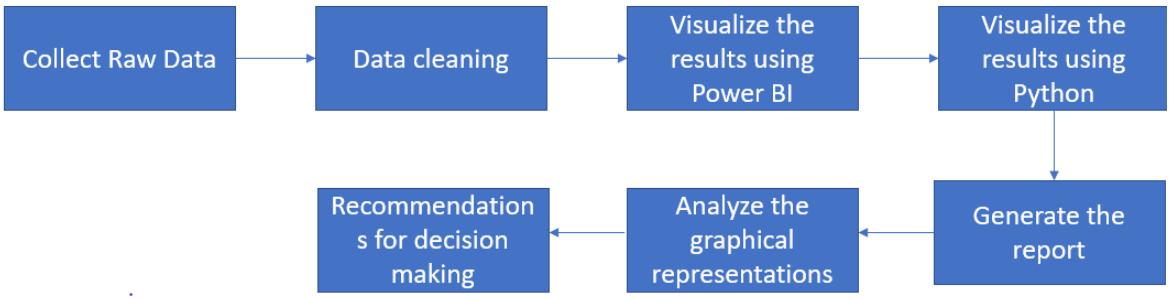


Figure 4.1 Block diagram of data analysis

Firstly, a group of students' data are collected based on their respective scores on midterm, assignment, practical and final exam. The group of students consist of different programmes, Business Information Systems, Information Systems Engineering and Computer Science. The next process will be data cleaning, where it's a process of fixing, incorrect, incomplete or any null numbers within the dataset collected. Aside from that, this is also to ensure that only the most significant data is filtered to ensure accuracy and reliability of data before conducting data analysis. The next step will be visualizing the results by using an open source business intelligence tool which is chosen from the available market. The data is then transferred and integrated into Power BI and the graphical representations are then produced. The existing results are well as developed using Python with installed libraries of matplotlib and pandas for data visualization. Both of the reports are then generated with all the stats of the students' academic performance in the form of different graphical representations. The reports are then analyzed and discussed which then further given recommendations for decision making.

Chapter 5

System Implementation (For Development-Based Project

The goals of preliminary data analysis are to modify the data to make it ready for subsequent study, define the data's essential properties, and summarise the findings. The theoretical and practical techniques to achieving the desired objectives are covered in this chapter.

5.1 Setting up

5.1.1 Software

Before we begin visualizing the data, there are three software needed to be installed and downloaded in my laptop:

1. Power BI
2. Python 3.9
3. Anaconda navigator (with matplotlib installed)

5.2 Steps of data visualization and analysis

The hardware involved in this project is computer alongside with a few software to do the data visualization. A sample data of 239 students with different course background are transferred into raw data before integrating the data into business intelligence software to graphically present the visualization in the form of dashboards as well as visualizing the data using a programming software, Python. Furthermore, an analysis will be conducted based on the results processed.

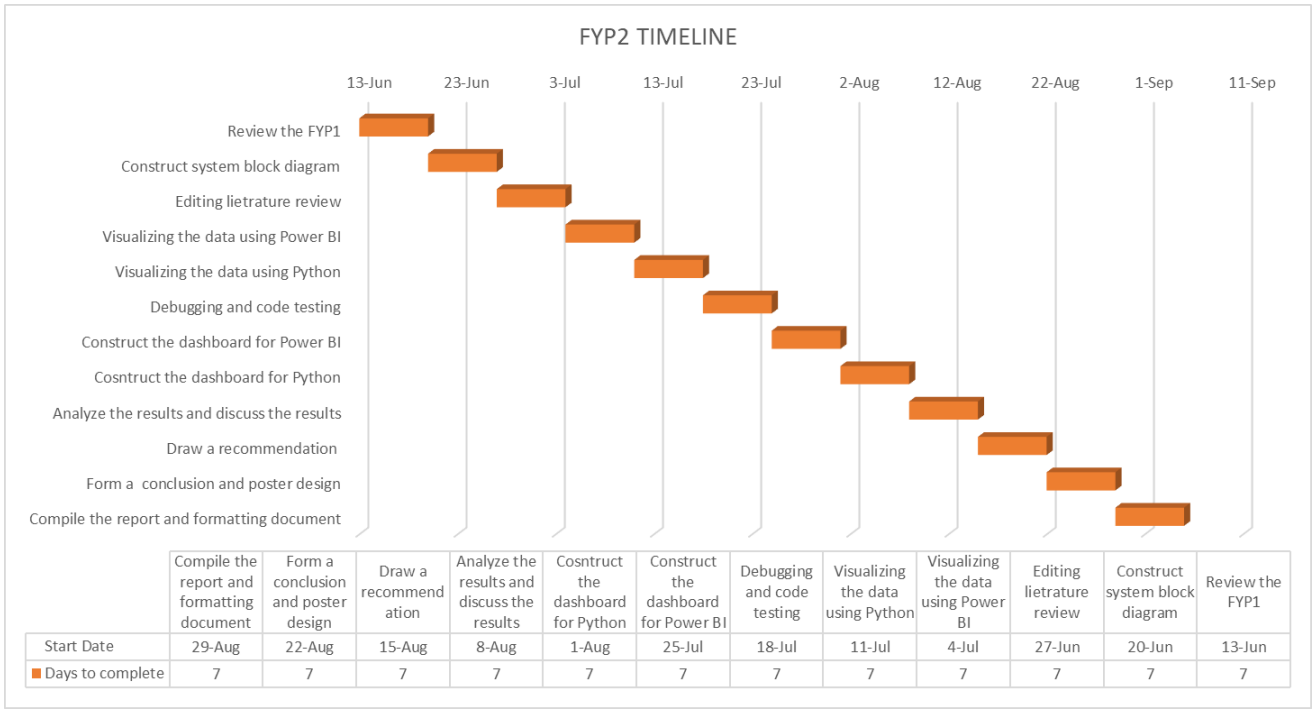
The following are the steps of data visualization process:

i. Data collection and preparation
ii. Visualize the results using Power BI
iii. Visualize the existing results using Python
iv. The graphical presentations are analyzed and discussed

v. Recommendations and formulation to improve students' performance

5.3 Timeline

To organize the tasks in a more systematic way, a gantt chart is created to plan the sub activities which will be useful for planning and scheduling Project 2 for this semester. The gantt below demonstrates what need to be done in a precedence order to avoid confusion as the each task is contingent on the completion of another. Based on the gantt chart below, 12 weeks are given to complete Project 2 where each task shall be done by week 1 till the end of completion. The starting date of the project will be on Week 1, 13 June 2022 till the end of the project ending at Week 12, 29 August 2022.



	Compile the report and formatting document	Form a conclusion and poster design	Draw a recommendation	Analyze the results and discuss the results	Cosntruct the dashboard for Python	Construct the dashboard for Power BI	Debugging and code testing	Visualizing the data using Python	Visualizing the data using Power BI	Editing lietrature review	Construct system block diagram	Review the FYP1
Start Date	29-Aug	22-Aug	15-Aug	8-Aug	1-Aug	25-Jul	18-Jul	11-Jul	4-Jul	27-Jun	20-Jun	13-Jun
Days to complete	7	7	7	7	7	7	7	7	7	7	7	7

Figure 5.1 Gantt Chart of Project

5.4 Data collection and preparation

In this process, we collected a handful of students' marks based on each of their respective course, UCCD9999 Website Design and Development where the coursework will be divided into three parts which are midterm test, practical test, and group assignment and a final exam. The students who took this course consists of different students from different programmes of study which are CS (Computer Science), IB (Business Information Systems), and IA (Information Systems Engineering). Therefore, Table 2 summarises the percentages distribution of each question according to CLO based on the categorisation as shown in below.

No	**Programme learning outcome (PLO)**	**Bloom Taxonomy**	**Delivering Methods**	**Assessment method and marks breakdown**						
				Midterm	**Practical test**	**Assignment**	**Final exam Q1**	**Final exam Q2**	**Final exam Q3**	**Final Exam Q4**
				20%	**10%**	**20%**	**12.5%**	**12.5%**	**12.5%**	**12.5%**
1	CLO1	Cognitive	Lecture, Practical				/	/		
2	CLO2	Cognitive	Lecture, Practical	/					/	
3	CLO3	Cognitive	Lecture, Practical			/				/

4	CLO4	Psychomotor	Lecture, Practical		/					

Table 5.1 Assessment Method and marks breakdown

5.5 Visualization of results using Power BI

In this process, we will be implementing the software, Power BI as our business intelligence tool to visualize the output of the students' results. In this first step, the sample data of students' mark which in the form of Excel format are integrated into Power BI. Therefore, to begin, we get data from the Excel workbook to locate our database which to be integrated into the software.

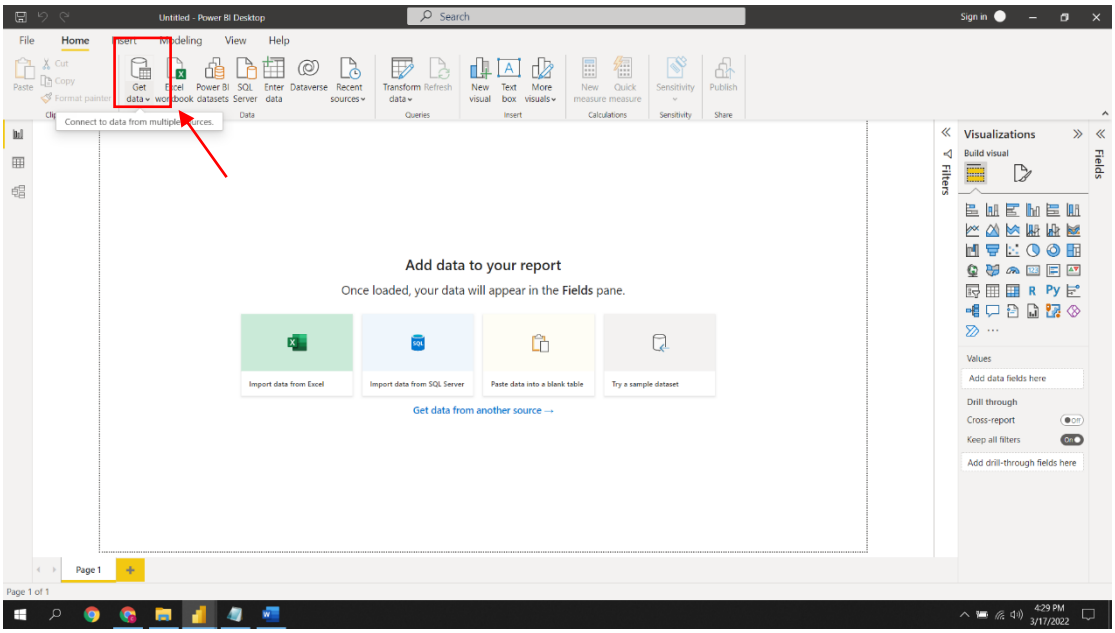


Figure 5.2 Step 1 of data visualization

In our next step, before we load the data, we would transform the data first to do some modification to the queries like renaming the column names, removing null fields, and filter out only the significant information we need for analysis. This is done so that we could consolidate them into a useful and meaningful query.

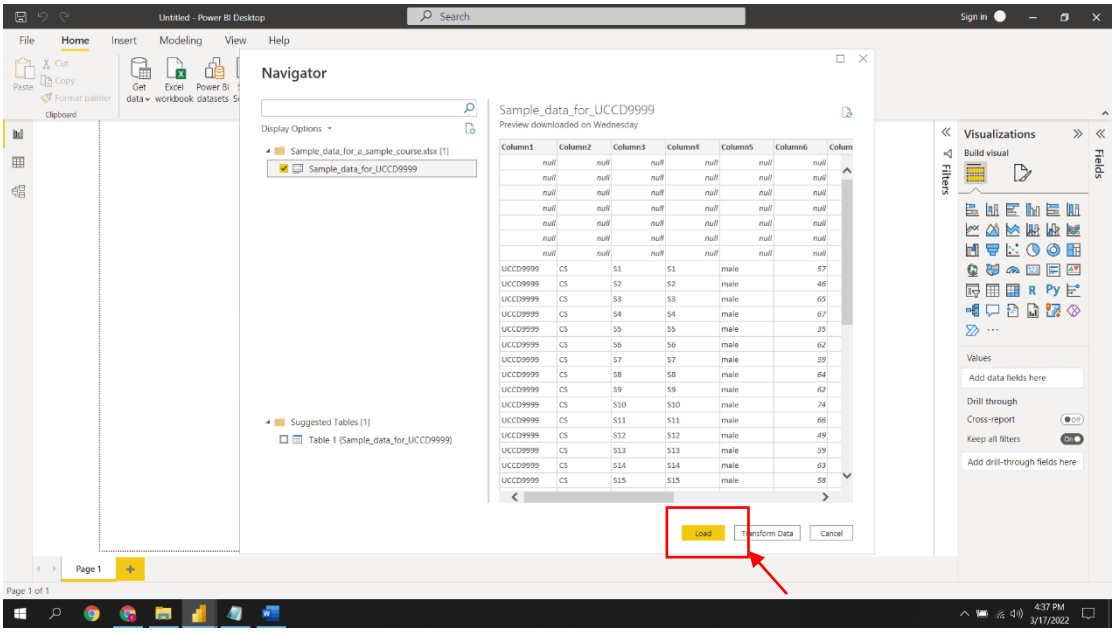


Figure 5.3 Step 2 of visualization

In our third step, we had renamed and filtered the rows and it is ready to be loaded into the platform by clicking “Close & Apply” to save the changes made to the queries.

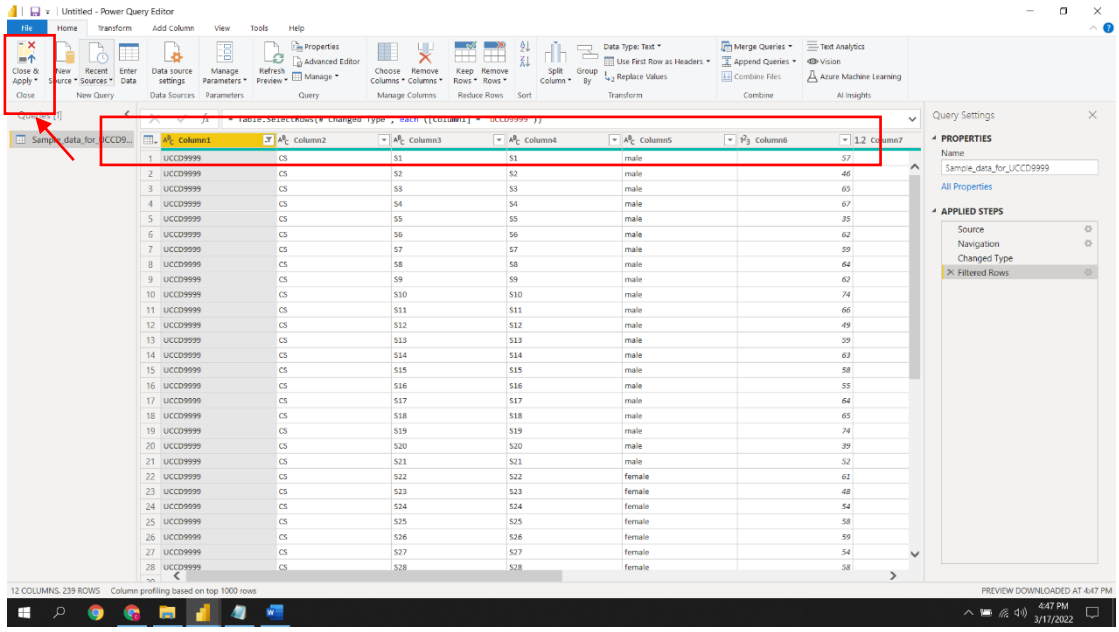


Figure 5.4 Step 3 of visualization

In our final step, we visualize the data in the form of graphical representations like stacked column chart, pie-chart, table, funnel, clustered bar chart, area chart, Tree map, and line chart by dragging the data on the right side into the visualization fields. The first dashboard, which presented in Figure 4.5 provides the brief information of midterm results according to average scores, programme, gender, and midterm results. and student id.

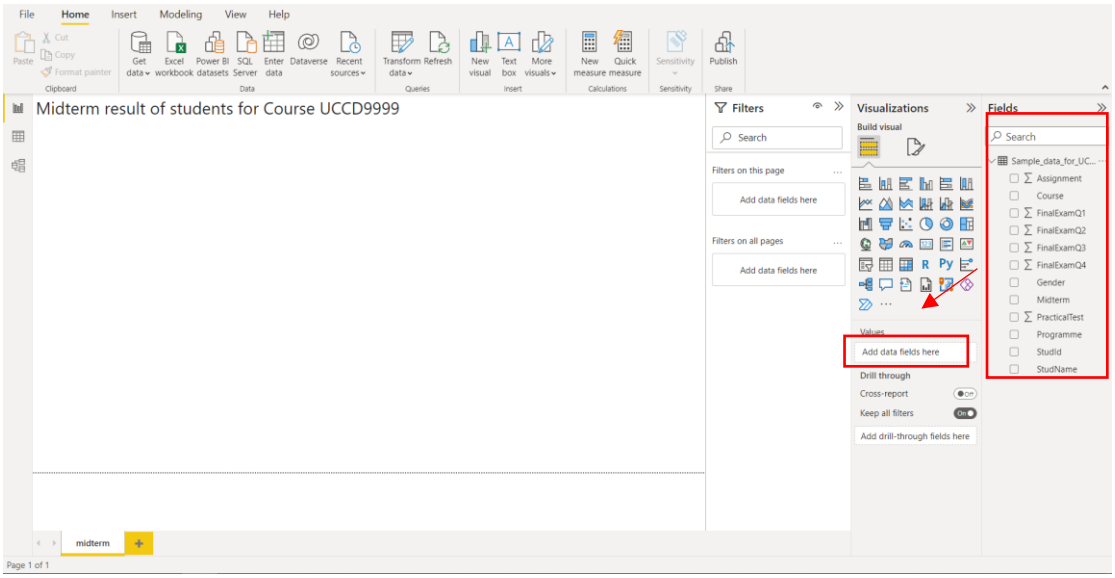


Figure 5.5 Step 4 of visualization

The Power BI application consists of five dashboards where the **first dashboard**, presented in Figure 4.5 conveys the overall information regarding the midterm test results of students according to average marks of midterm by all students, average marks by programme, average marks by gender, average marks by programme and gender, number of students who pass(marks =>50), number of students who failed(marks < 50), students gender who take the course and programme of students who took the midterm.

The **second dashboard** provides the information average of practical test by student and programme, average by programme, average by gender, maximum of student marks for each programme, median of student marks by programme, number of students who pass (marks =>50) and fail (marks <50), maximum by programme and maximum practical test by gender.

The **third dashboard** provides the information average of assignment by student, average by programme, average by gender, maximum of student marks for each programme, median of assignment by programme, number of students who pass (marks =>50), maximum by programme, minimum assignment by programme and maximum assignment by gender.

The **fourth dashboard** provides the information of final score by students, average of finals score by programme, the number of students who passed(=>40), average of finals score by gender, average of final exam Q1, Q2, Q3, Q4, average of final exam Q1 by programme, average of final exam Q2 by programme, average of final exam Q3 by programme, and average of final exam Q4 by programme.

The **fifth dashboard** provides the information final result(100%) by students, average of final exam by programme, number of students who passed(=>50), number of students who failed(50<), average of final result by gender, max of final result by programme, median of final result by programme, and average of final exam and coursework.

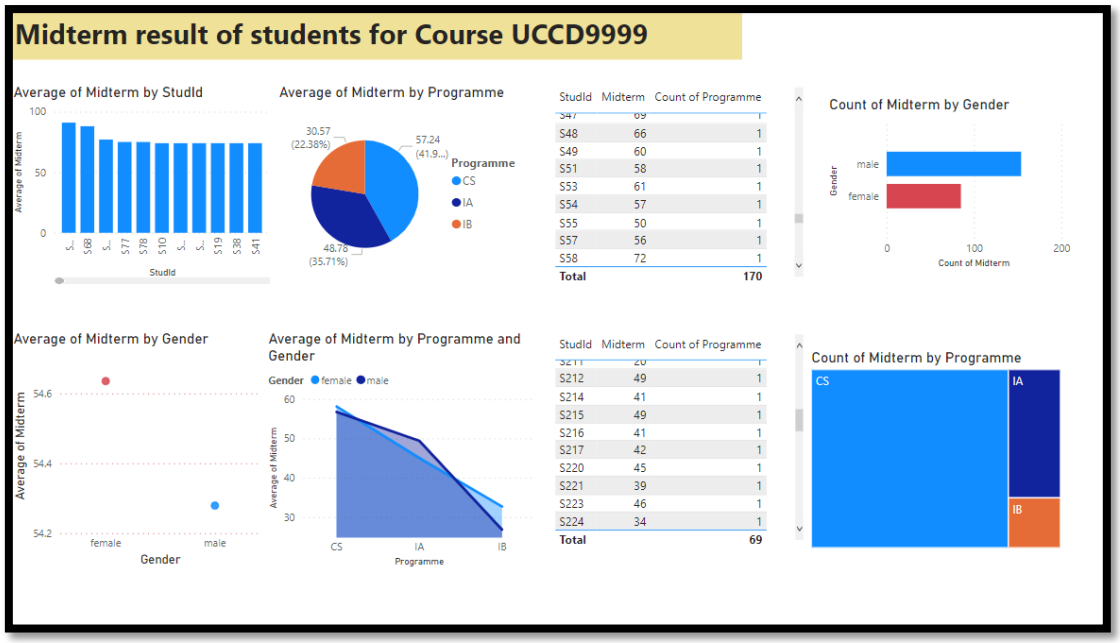


Figure 5.6 Dashboard overview of midterm result using Power BI

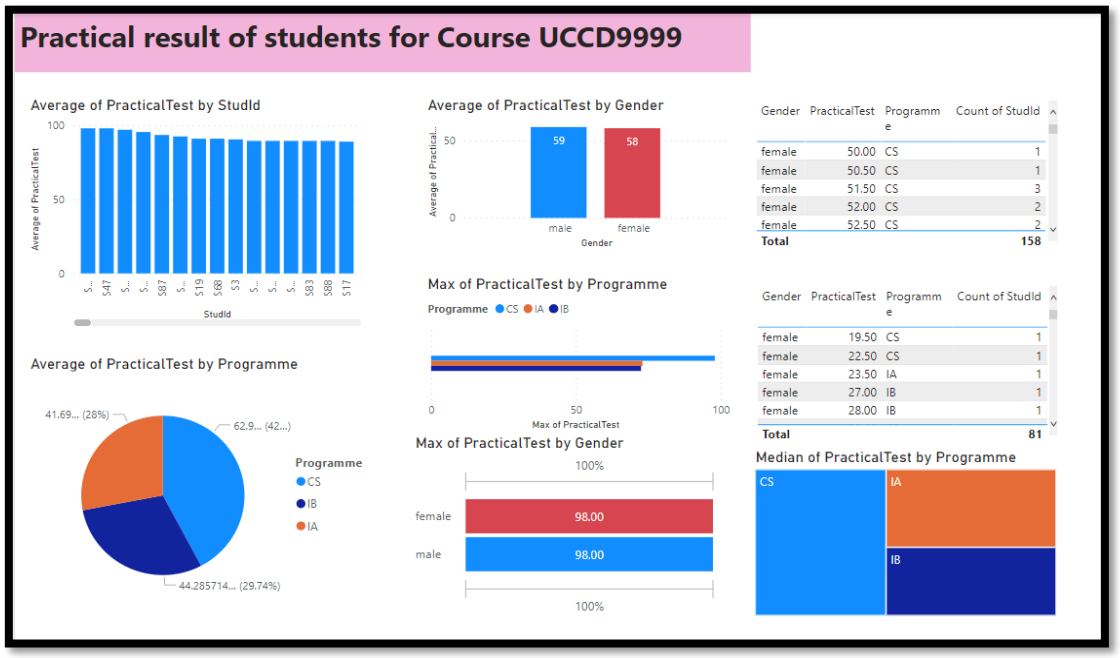


Figure 5.7 Dashboard overview of practical test result using Power BI

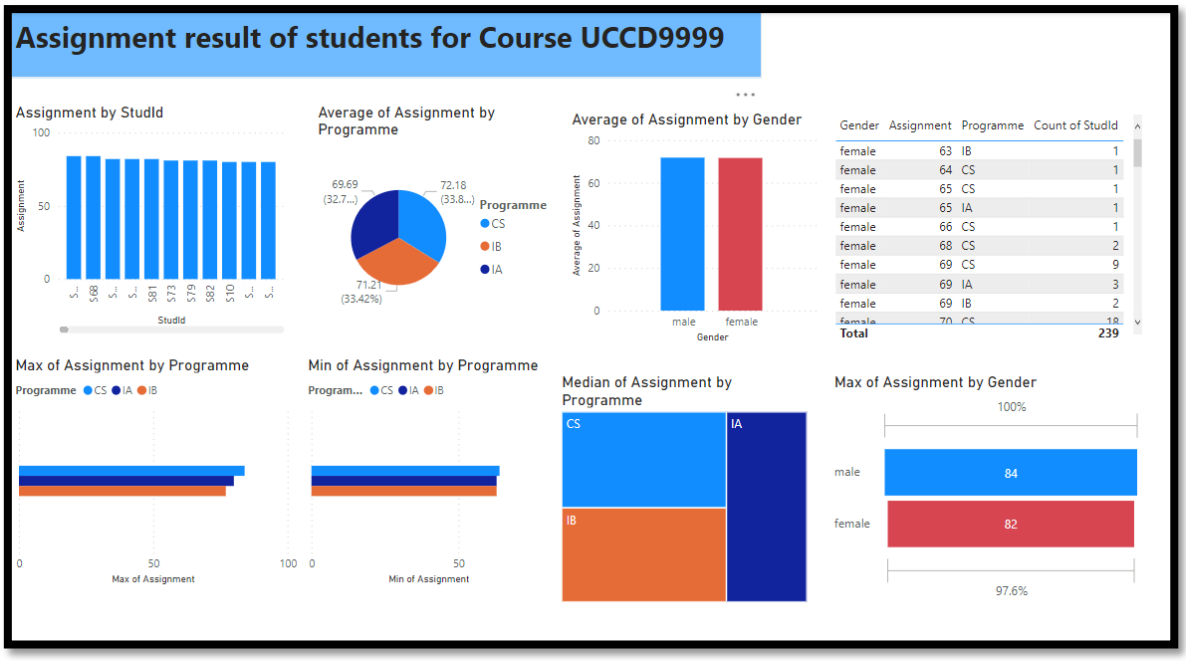


Figure 5.8 Dashboard overview of assignment result using Power BI

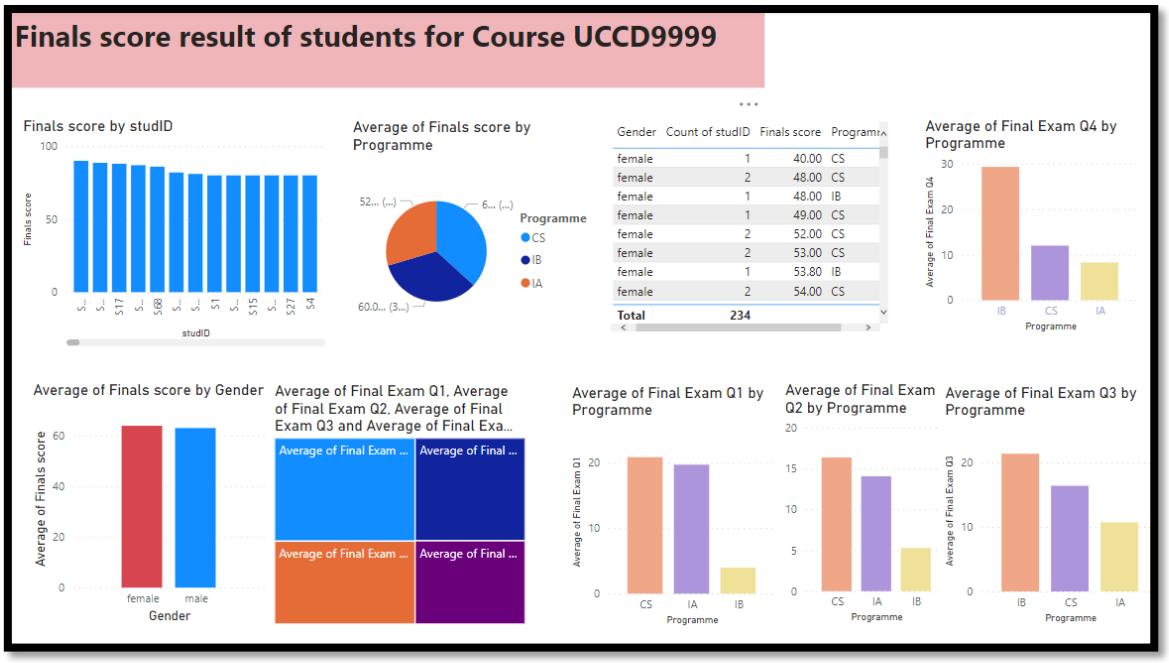


Figure 5.9 Dashboard of final score result using Power BI

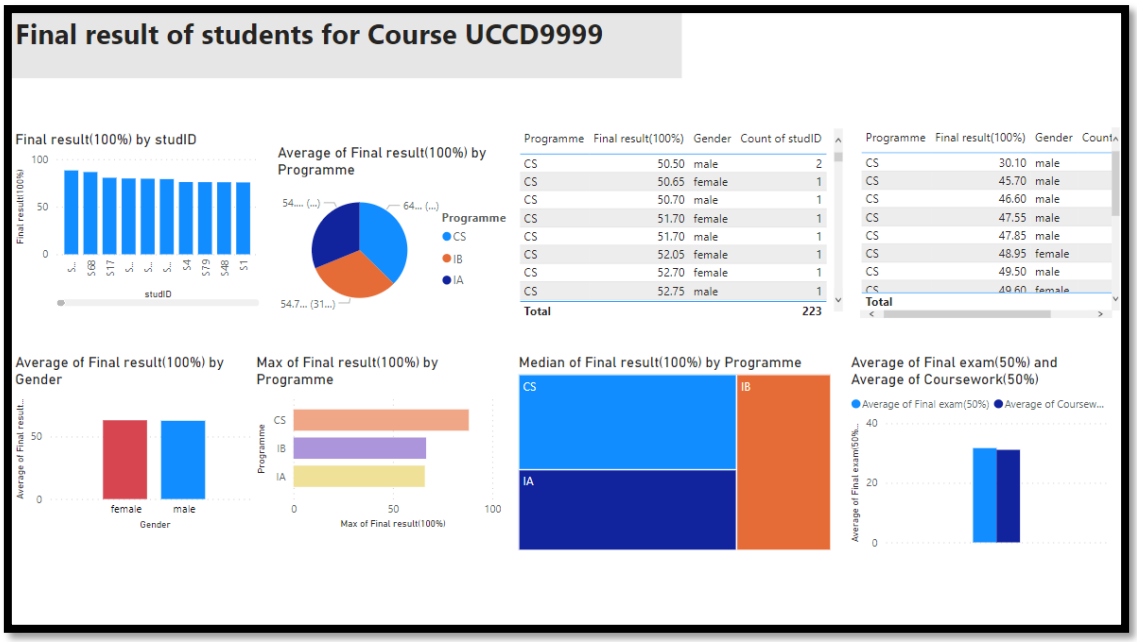


Figure 5.10 Dashboard of final results using Power BI

5.6 Visualize the existing results using Python

In this process, similar results of Power BI output are achieved with the use of Python with integrated matplotlib and panda. In this matter, Anaconda navigator is downloaded into system which enables us to launch application as well as manage conda packages, environments, and channels without the use of command line commands. Once the application is successfully installed, launch the Jupyter Notebook which is an open-source web application for the purpose of scientific computation, data visualizations and data processing. In short, this web application is a useful tool for writing as well as iterating on Python data analysis code. With the usage of this tool, you can create lines of code and execute them at the same time instead of creating and rewriting a full programme.

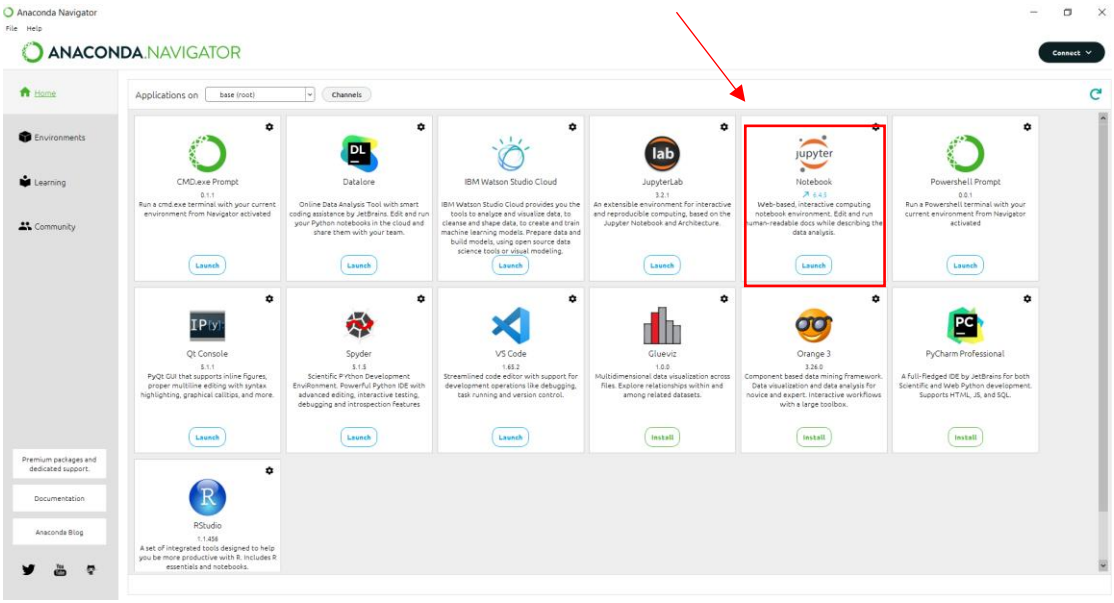


Figure 5.11 Launch Jupyter Notebook in Anaconda Navigator

In our next step, locate the database saved in the folder in Desktop

Figure 5.12 Locate the database saved

Modify the data source by data cleaning so that it is free of errors and anomalies as clean and consistent data allows us to perform data visualization easier.

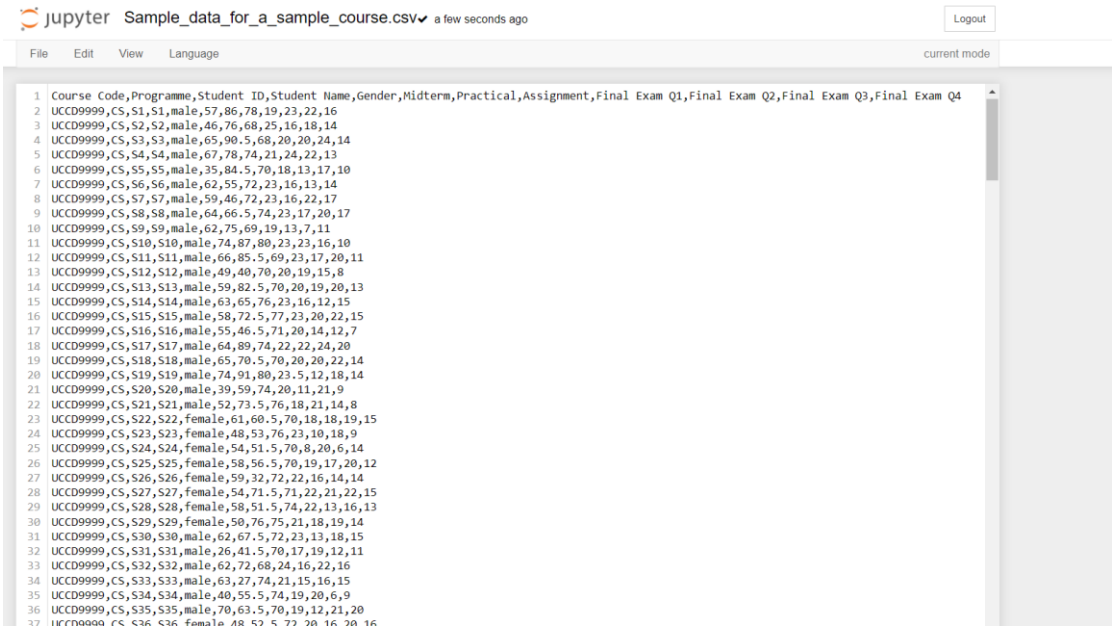
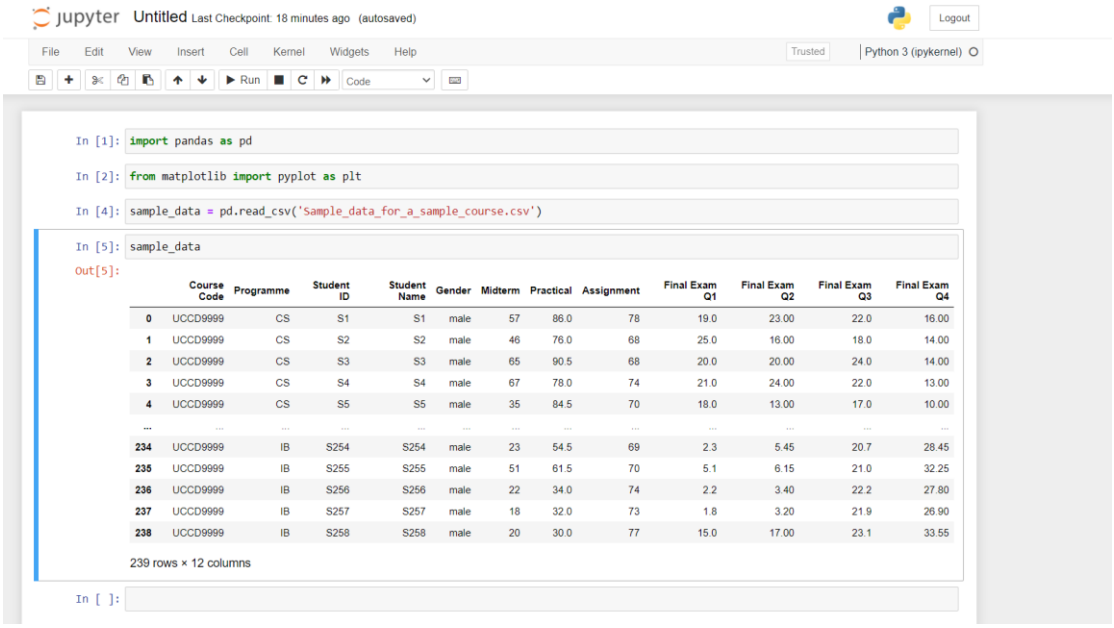


Figure 5.13 Data Cleaning

In our fourth step, click New and select 'Python 3 (ipykernel)' to start visualizing data.

Figure 5.14 Select 'Python 3(ipykernel)'

Next, import pandas and matplotlib library with the purpose to create static, animated, and interactive visualizations. Then transport the sample data from located file into the web application and display the data which needed to be visualize later.



	Course Code	Programme	Student ID	Student Name	Gender	Midterm	Practical	Assignment	Final Exam Q1	Final Exam Q2	Final Exam Q3	Final Exam Q4
0	UCCD9999	CS	S1	S1	male	57	86.0	78	19.0	23.00	22.0	16.00
1	UCCD9999	CS	S2	S2	male	46	76.0	68	25.0	16.00	18.0	14.00
2	UCCD9999	CS	S3	S3	male	65	90.5	68	20.0	20.00	24.0	14.00
3	UCCD9999	CS	S4	S4	male	67	78.0	74	21.0	24.00	22.0	13.00
4	UCCD9999	CS	S5	S5	male	35	84.5	70	18.0	13.00	17.0	10.00
...	...	...	...	...	...	...	...	...	...	...	...	...
234	UCCD9999	IB	S254	S254	male	23	54.5	69	2.3	5.45	20.7	28.45
235	UCCD9999	IB	S255	S255	male	51	61.5	70	5.1	6.15	21.0	32.25
236	UCCD9999	IB	S256	S256	male	22	34.0	74	2.2	3.40	22.2	27.80
237	UCCD9999	IB	S257	S257	male	18	32.0	73	1.8	3.20	21.9	26.90
238	UCCD9999	IB	S258	S258	male	20	30.0	77	15.0	17.00	23.1	33.55

Figure 5.15 Read the data from the located file

Visualize the data with the implementation of matplotlib and panda's library to produce the clustered bar chart with sorted descending order. The bar chart below is visualized based on the average of midterm(y-axis) and student ID(x-axis) alongside with title and legend.

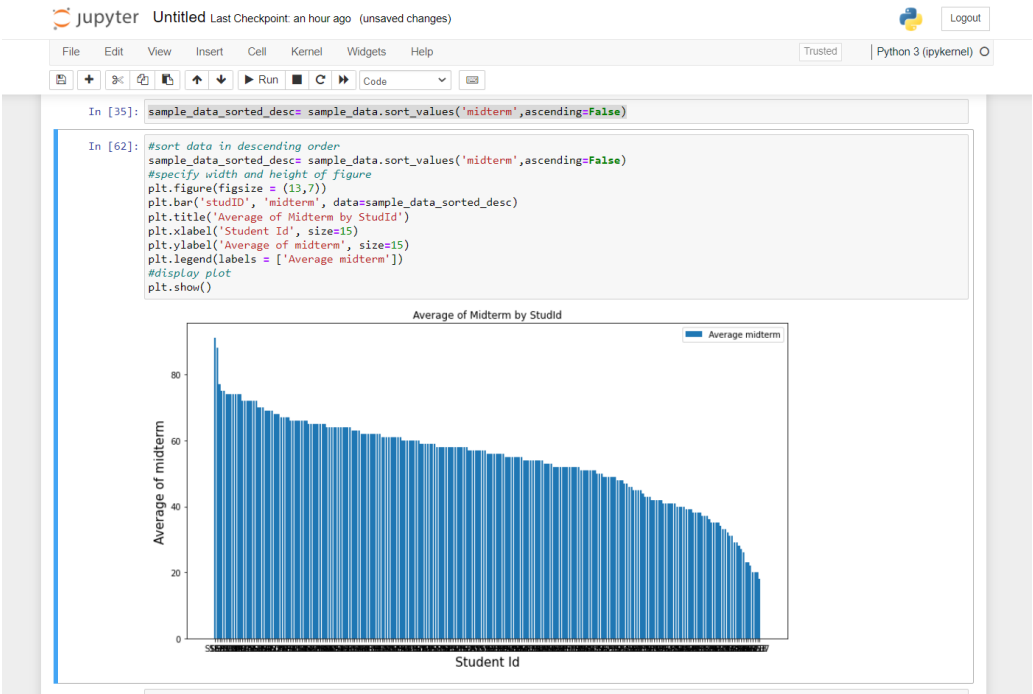


Figure 5.16 A bar graph output of midterm

Similar results are achieved with the use of Python with integrated matplotlib and pandas. The first and second dashboard is shown below as the data analysis of midterm and practical result of students for the course UCCD9999.

Figure 5.17 Dashboard overview of midterm using Python

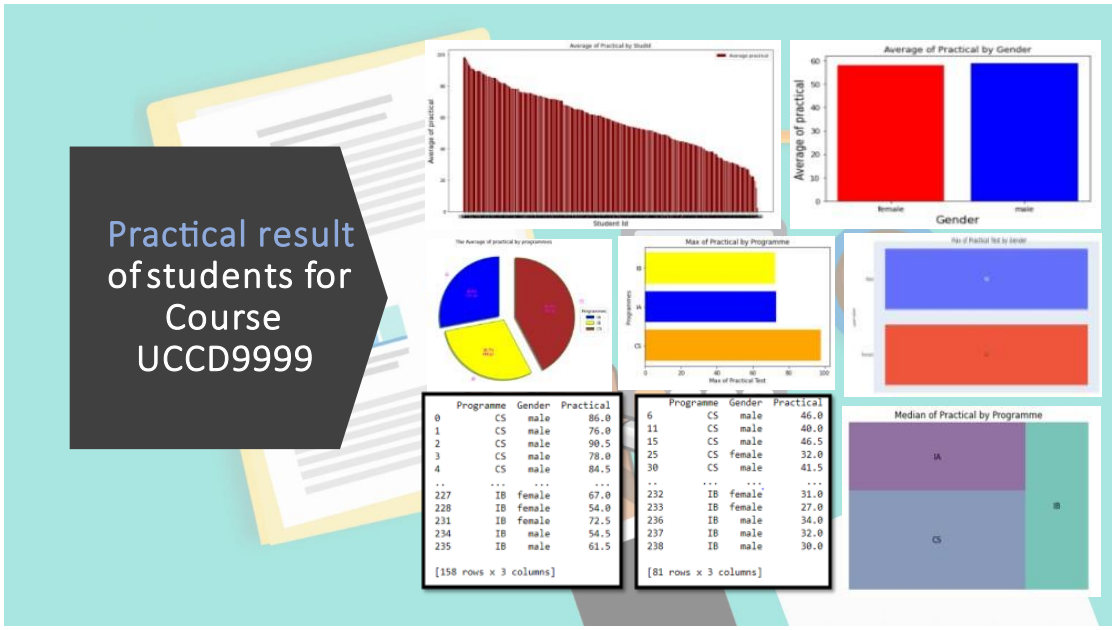


Figure 5.18 Dashboard overview of practical using Python

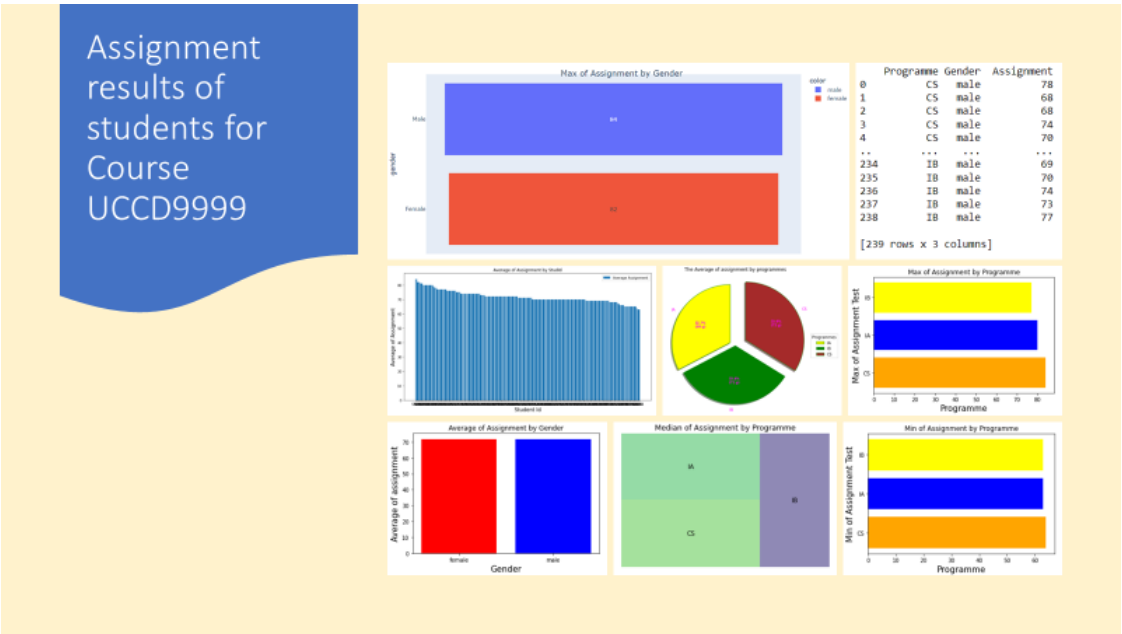


Figure 5.19 Dashboard overview of assignment using Python

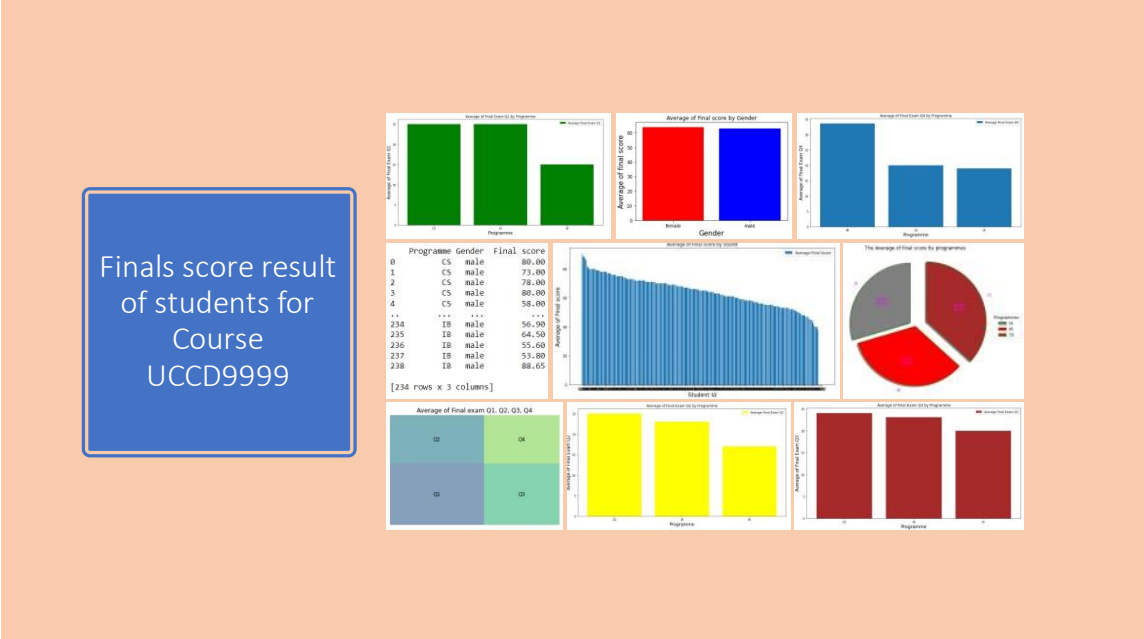


Figure 5.20 Dashboard overview of final score using Python

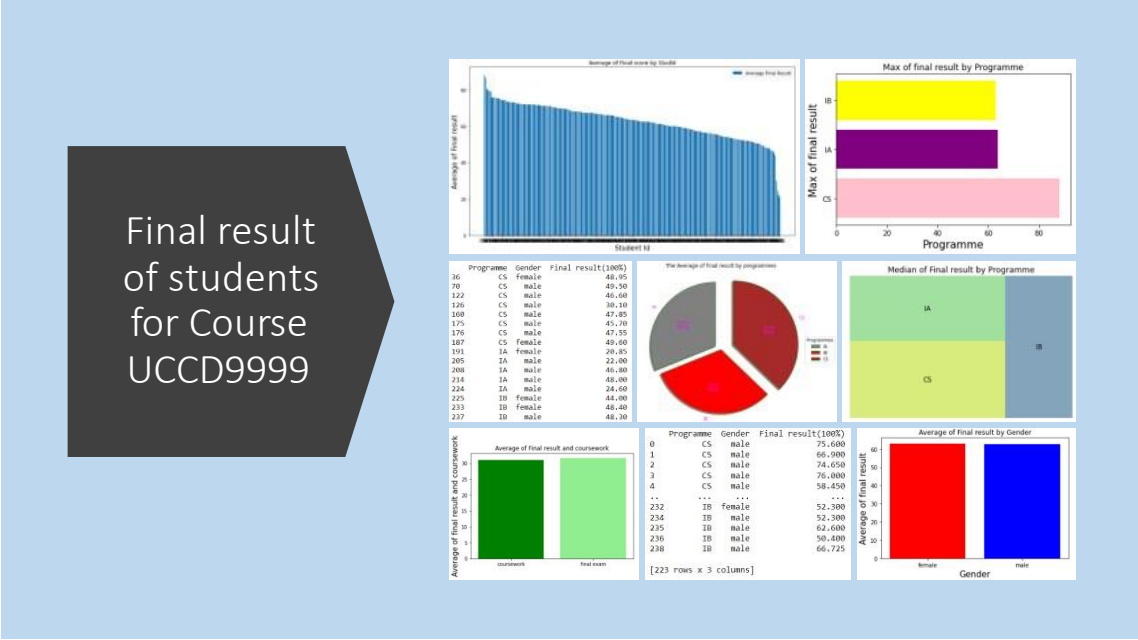


Figure 5.21 Dashboard overview of final results using Python

Chapter 6

System Evaluation and Discussion

6.1 Analysis and discussion

6.1.1 Analysis of dashboard 1 (Midterm)

According to the dashboard of bar chart, representing the average of midterm by Student ID, it was speculated that the highest score was S168, who scored 91 out of 100 in the midterm test whereas the lowest score was S257, who scored 18 out of 100 in the midterm test. Hence, the data also revealed that S168 is a computer science undergraduate student while S257 is a business information system undergraduate student. To add on, only 5 out of 239 students managed to achieve excellent mark (marks >= 75) which once again from a computer science background which covers only 0.02% while the other 234 students (99.8%) scored below 75. Although students from information systems engineering background didn't manage to surpass most of the computer science students, their results are doing quite well at satisfactory level where the highest IA student(S236) scored is 72 marks which is 19 marks below the highest CS student who got 91. On the other hand, IB student(S250) who scored the highest (53 marks) in his programme was apparently the lowest comparing to the other students from IA and CS.

Based on the pie chart according to the average of midterm by programme, it was observed that CS students scored the highest average which was 57.24 marks followed by IA (48.78) and IB (30.57). This is also fair to justify that student from CS programme covers the highest percentage based on the data which is 41.91% while IA and IB covers 35.71% and 22.38% respectively. Hence, the observed table data shows that 170 students of all programmes have passed the midterm by scoring 50 and above. This also means that 71.13% of all students managed to pass the midterm where the highest recorded was 91 marks and the lowest was 50 marks. Although majority of the students may have passed, it was also found out that only 2 students from IB (S250 & S255) managed to pass by scoring 53 and 51 respectively. In this matter, 2 out of 14 students (14%) from IB programme were scoring at satisfactory level while majority of CS students managed to pass with students totalling 151 out of 189 which makes it 79% of the CS programme. Surprisingly, less than half of students from IA programme

managed to score above 50 where only 17 out 36 students passed. This also means only 47% students have satisfactory scores.

Based on the table data showing students scoring below 50, 69 students from all programmes have failed the midterm test. This means that 28.87% of all students have failed the midterm. Data shows that 38 students from CS, 19 students from IA, and 12 students from IB have performed below satisfactory rate. IB students were performing at a bad state as majority of the students have failed which covers 85% of the programme. Students from CS on the other hand were doing fine as only 20% of the students in the programme have failed while IA students were also not doing quite well as 52% have failed the midterm, which exceeds a little more than half. The horizontal bar chart stating the number of genders who took the course shows that 154 students who were males and 85 students who were females out of the 239 students. This is equivalent to 64% of males and 36% of females are taking the course where there are more males than females by 28%. Next, the scatter chart stating average of midterm shows that the average score by female is 54.64 and male is 54.28. This generally means that the average of female scored higher than the average of male by 0.36 marks. Although there are more male students who took this course than number of females, it was surprised that the average of 85 females score higher than the average of 154 males.

Hence, according to the data from the line chart based on the average of midterm by gender and programme, it shows that the average of female and male students from CS scored the highest among all the programmes, 58.10 and 56.73 respectively. On the other hand, the average of female and male students from IA scored averagely between CS and IB, 45 and 49.39 respectively. However, the average of female and male students from IB scored the lowest among all programmes, averaging at 32.67 and 26.80 respectively where the students scored below satisfactory level regardless of gender. Last, the tree map data shows that 189 CS students, 36 IA students and 14 IB students out of 239 students are taking this midterm. Generally speaking, CS, IA and IB are made up of 79%, 15%, and 5.8% respectively of the whole programme.

6.1.2 Analysis of dashboard 2 (Practical)

Based on the dashboard of bar chart which represents the average of practical by Student ID, it was speculated that the highest score was S134, who scored 98 out of 100 in the practical test whereas the lowest score was S244, who scored 2 out of 100 in the practical test. Hence, the data also revealed that S134 is a computer science undergraduate student while S244 is an information system engineering undergraduate student. To add on, 56 out of 239 students score excellent marks which is 75 and above. This is a huge leap compared to the midterm where only 5 students managed to score excellent marks. Also, all the 56 students are made from CS students which is equivalent to 23% of the total of students all programmes. The highest marks scored by an IB student(S251) was 72.50 which is slightly better than the previous result in midterm while the highest marks score by an IA student was 73 marks which is also slightly better than the previous midterm. Although IB student has the highest of 72.50, CS student remains the all-time highest of 98 follows by IA which is 73.

According to the pie chart, stating the average of practical by programme, students from CS programme scored the highest average, which is 62.91 followed by IB and IA, 44.29 and 41.69 respectively. This is equivalent to say that CS covers 42.25%, IB at 29.74% and IA at 28%. It was analysed that the average of IA in practical test wasn't performing as good as in midterm which has the lowest average among all programmes. Next, the bar chart according to average of practical test by gender shows that the average of male scored slightly higher than female with scoring 59 for male while 58 for female. This is analysed that the average of 154 male students score higher by 1 mark than the average of 85 females. Also, the horizontal bar chart of max practical test by programme shows that CS student leads the highest score, which is 98 followed by IA, 73 while IB, 72.50. This justifies that CS students was far ahead than students from the other programme. Hence, the funnel, according to the max practical test by gender shows that the highest score by female and male student have the equal score which is 98. In fact, both students were also students from CS programme.

Looking at the table data based on the number of students who passed the practical by scoring 50 and above, it was analysed that 158 students managed to pass. This means that 66% of the students from all programmes are performing at a satisfactory result. Majority of the students are once again from CS programme where it made out from 140 students while 12 students

from IA while 6 students from IB. It was analysed that most of the IA and IB students didn't scored 50 and above as both measuring at 33% and 42% respectively which is not more than half of the students in the same programme. Based on the table data showing students scoring below 50, 81 students from all programmes have failed the midterm test which has higher failure rate than in midterm. This means that 33% of all students have failed the midterm. Data shows that 49 students from CS, 24 students from IA, and 8 students from IB have performed below satisfactory rate. IA students were performing at a bad state as majority of the students have failed which covers 66% of the programme. Students from CS on the other hand were doing fine as only 25% of the students in the programme have failed while IB students were also not doing quite well as 57% have failed the midterm, which exceeds a little more than half. Lastly, based on the tree map data stating the median of practical by programme, CS remain the highest median score of 63.50 followed by IA, 44 and IB, 38.50.

6.1.3 Analysis of dashboard 3 (Assignment)

Based on the bar chart of assignment by students, it was found out that S168 scored the highest which is 84 while the S245 scored the lowest which is 63. S168 appeared to be from CS while S245 is from IB. Looking on the pie chart data of average of assignment by programme, the average of CS students obtained the highest score at 72.18 followed by IB students at 71.21 and IA students at 69.69. Next the stacked column chart shows the average of assignment result by gender, male and female. It was analysed that the score of male is slightly higher than female, in which the score of male at 71.81 while 71.65. This show the difference in 0.16 between both genders. Next, the table shows the data of students who passed the assignment(=>50). It was discovered that none of the students failed the assignment with passing rate of 100% regardless of programme or gender. Other than that, the clustered bar chart shows the max of assignment scored by programme. The highest reported is 84 by CS programme followed by IA at 80 and IB at 77. Meanwhile the minimum score of assignment is 63 by IB programme followed by IA, which is also at 63 and followed by CS which is at 64. The tree map shows the median of assignment by programme. The highest median scored was at 72 by CS, followed by IB at 71 and lastly followed by IA at 70. Lastly, the max of assignment shows that male scored slightly highly than female, in which male scored at 84 while female scoring at 82.

6.1.4 Analysis on dashboard 4 (Final score result)

Based on the clustered column chart on the finals score by StudID, the highest score obtained is S191, scoring 90 while the lowest score obtained is S37, scoring 38. The four students who scored 0 might have withdrawn from the course or might be absent for the final exam. Next, the pie chart shows the average of finals score by programme. CS students have the highest average score at 65.59, while IA scoring at 60.02 and IB scoring at 52.78. The table data which shows students who passed the finals with a score(=>40). In total, 234 students have passed the exam with a minimum score of 40 while 5 students have failed the exam, where one of the students scored 38 while the others got 0. The clustered column chart shows the average of final score by gender. It was examined that female students scored slightly higher than male where female category scoring at 63.92 whereas male scoring at 63.01. The next tree map shows the average of final exam Q1, Q2, Q3, Q4. The highest scored obtained is in final exam Q1 at 19.71, followed by final exam Q3 at 15.82, final exam Q2 at 15.36 and final exam Q4 at 12.43. The clustered column chart of average of final exam Q1 by programme shows that the highest average scored is by CS at 20.88 followed by IA at 19.72 and followed by IB at 3.99. Next, the clustered column chart of average of final exam Q2 by programme shows that the highest average scored is by CS at 16.35, followed by IA at 14.06 and IB at 5.29. The next clustered column chart of average of final exam Q3 by programme shows that the highest average scored is by IB at 21.36 followed by CS at 16.39 and IA at 10.72. The final clustered column chart of average of final exam Q4 by programme shows the highest average scored by IB at 29.29 followed by CS at 11.97 and AI at 8.28.

6.1.5 Analysis on dashboard 5 (Final result)

Based on the clustered column chart, it was analysed that the highest score obtained was by student S168 at 88.20, from CS programme while the lowest obtained was student S211 at 20.85, from IA programme. The pie chart shows the average of final result by programme, where the highest average score is 64.97 by CS programme followed by IB at 54.79 and IA at 54.25. Next, the table which shows the amount of students who passed this subjects shown that 223 students have passed while the other 16 students have failed the subject, scoring below 50. This also means that 93% of the students have passed while the other 7% students have failed. It was also found out that males have a higher failing rate compared to females. The clustered column chart showing the average of final result by gender where female scored a higher

average at 63.03 compared to male, scoring an average at 62.61. Next the clustered bar chart shows that the maximum achieved was 88.20 by CS, followed by IB at 66.73 and IA at 66.05. The tree map shows the median of final results based on programme achieved where the highest is still obtained by CS students following IA and IB. Lastly, the clustered column chart shows the comparison of final exam and coursework. It was analysed that students tend to score higher at final exam, scoring at an average of 31.67 while the coursework mark at 31.09.

6.2 Discussion of results

Based on the findings after the performed analysis on the dashboard, it was found out that students from CS background generally performed better than students from other programmes followed by IA and IB. One of the very main reasons is students who studied computer science are smarter as they do enjoy math and problem-solving skills where these are the requirements for a subject that needs coding. In fact, computer science programme is not an easy path as it requires much continuous coding practices every day. Also, many people believe computer science to be amongst the most difficult subjects to study. They feel that Computer Science is a challenging subject to study and there is no denying it. Caused by a variety of circumstances, studying, and understanding computer science may be difficult. Numerous factors like some of the most challenging areas in the discipline include computer technology, software, and statistical techniques. So, meaning to say that website development wasn't difficult to most CS students as they had experienced in more harder subjects which require complicated coding structures like C++ language. In this matter, website development is considered easier for them who has deep experience in other programming languages. Besides, it was also found out that students from CS have exposure to coding since foundation in science, before entering to degree field. On the other hand, IA students did not perform as good as CS students since IA has a narrower scope while CS has a wider scope. For instance, IA requires less details of mathematical theories and less complexity compared to CS although it still requires maths and problem-solving skills. Unfortunately, students from IB didn't perform as its peak comparing to CS and IA. One of the main factors is that the students never expose much in the field of coding as it was also a business-related programme. In fact, students have less experience in coding where what is learnt throughout the programme was broad and only basics of programming was learnt, and with a mixture of business knowledge as well. This is also fair to say that programming language which IB students are learning is the tip of the iceberg. Besides,

IB students have no coding knowledge or experience as most of them enrolled into foundation in arts leading to accountancy and management where no subjects required coding. In this matter, it is studied that the students basically have zero exposure and very novel to them. It was also studied that students from all programmes generally performed better in practical test compared to midterm where the average score in midterm was 54.41 while average of practical score is 58.63, which is 4.22 marks higher. It is also safer to say that most students managed to pass both tests. Based on the bloom taxonomy, it was studied that students performed better in psychomotor domain than in cognitive domain. This means that students are better at hands on coding physically at computer lab rather than memorising or knowledge based. This is because most students have hands on practice their coding during practical class session every week and more time are spent on coding rather than reading. As midterm test will require knowledge and understanding on the theoretical information, students didn't manage to score higher as it require much studying and reading to recall the information absorbed. Other than that, it was also observed that the average female students scored higher than average of male in midterm while average male students scored higher than average of female in practical. This also means that females performed better in cognitive domain whereas males performed better in psychomotor. One of the factors is that females are basically good at reading and writing while males are better in terms of logical and problem solving. Cognitive tasks like sitting for a midterm test, where memory derived from, grows, and develop more rapidly in females that in males. As a result, this brings an impact to reading as well as writing skills. In terms of psychomotor, where it requires logical thinking, this is where males come to place as males are studied to have better logical reasoning when it comes to problem solving in practical. Hence, it was analysed that there are more failures in practical test than in midterm test. This also means that most of the students are performed better in cognitive than in psychomotor domain although the average of practical is slightly higher than midterm. This explains that majority of the students are reasonably good in knowledge-based task as most of the exams encountered by students throughout the stages of education are mostly cognitive which includes quiz, midterm, and finals and assignment. These covers 80% of the total course where cognitive domain is involved. Another factor which leads to more failure rate in practical is that students have shallow understanding of the programming concept through the act of personal experience like practicing coding everyday instead of theoretical learning. This also justified that student didn't spend much time on coding practices which able to improve engagement as well as knowledge retention and this is necessary in every subject that need coding skills. On the other hand, it is analysed that there are 100% passing rate for the assignment task. This also means

that most of the students have better cognitive skills by encouraging discussions of what is being taught in class and to explore and understand of how the principles are to be applied into assignment tasks. Besides, the final exam score(50%) also shows that there 98% passing rate where only 5 students failed the test, out of 239 students. As for the final result, it was analysed that 223 students have passed the course while the other 16 students have failed the course. This also means that majority of the students have strong grasph of knowledge and skills about this subject. Lastly, it was also found out that the final exam overall score of students were slightly higher than their coursework marks. This is proven that students are better in terms of cognitive domain than psychomotor domain.

6.3 Recommendations and formulations

First and foremost, one of the recommended solutions is that higher educational institution should implement another foundation programme, leading to Information and Communication Technology so that students are able to expose in the field of technological world before enrolling into an IT course. It is vital to equip students with the basics of programming skills which will be required and relevant during degree studies so that students won't have troubles or difficulties when learning the concepts of programming. It is because most students who performed badly are majority from IB, where they have no certain knowledge of discipline in its degree. By introducing a new IT programme in foundation, the subjects will tailor towards enabling students to prepare for their undergraduate computing studies as students have to get used to principles and concepts of the Information Technology industry. On the other hand, the higher educational institution could also introduce a few subjects related to programming language like Introduction to C++ or Java in the foundation of arts (Accountancy and Management) programme to allow students to put an emphasis on basics of computing skills, especially IB or IA students before enrolling into degree of IT. Furthermore, it is also recommended that lecturers or tutors pay more attention to the students who lacks capabilities in learning and performed below satisfactory level regardless of degree programme or gender by asking the students in meeting to solve the questions of a particular problem instead of implementing spoon feeding culture as it tends to create 'learned helplessness' among students. This condition will results students into becoming reliant on others instead of solving by themselves. When students can't solve it, most of them gave up and apparently lose interest in

the subject. Learning how to program will be difficult when there is no interest in it. And because of this, helpless students will take very little responsibility in their own learning and eventually lose interest in the subject and ended up failing the exam. In fact, listening in class and reading concepts of coding is not enough to adopt the skills as coding requires mass number of practices, effort, time, and self-experimenting. Besides, a development of programming courseware is also a suggested solution to improve students' weak performance. As most students begin as a novel starter, they are encouraged to learn HTML and CSS before moving on to harder programming languages like JavaScript as it don't need to have any kind of programming experience and less complexity. Common languages like JavaScript might pose difficulties to certain students due to its multi-paradigm nature and more advanced syntax. For instance, the courseware will be offered to every student throughout the semester regardless of weak or smart students. Also, this programming courseware will have different sets of lessons covering different chapters, exercises, examples, and quizzes. In every chapter, each student must complete the exercises and score within a set target by the lecturers to prove the understanding on the chapters. This will allow lecturers to identify which area of weaknesses are the student into. Another recommendation is by establishing a coding club in the faculty of information and technology where it allows students from all IT programmes to develop applications together as a team and discover latest software. Lecturers could step in as the project managers by assisting students to guide them on the right track as well as diagnosing bugs and fixes. This club would also enable students to think out of the box and exchanging ideas on how to code an application as programming is more than coding itself. Lastly, lecturers are also recommended to implement live coding teaching instead of using slides to explain the concepts and principles. During live coding, lecturers should construct programs in front of students and its much more effective for many factors. First, students get to watch how the lecturer solves the problem and correct the mistakes. By doing this, students know what is right to do and what is wrong to do so. This will help students into preventing mistakes when coding. For example, lecturer demonstrates the execution of the program by adopting "what-if" to test out what would be the output if different methods are used. At this point, students can learn quicker, knowing the rules of thumb to avoid unanticipated bugs and errors. Live coding will encourage students to make predictions the outcomes of the output for all programs run by the lecturer. As a result, more students able to understand what is conveyed instead of using traditional teaching method where lecturer will continuously be teaching without noticing which students are left behind to complete the course syllabus. In a matter of fact, students who can't catch up with the syllabus will performed badly.

Chapter 7
Conclusion and Recommendation

7.1 Project review and discussion

As a conclusion, the findings based on performed data analysis on the results of students have been identified as well as the area of weaknesses of students. It is proven that business intelligence plays a very significant role in education especially into improving students' performance. By undergoing data analysis using BI tools, higher educational institution able to use the collected data gathered and visualized them into a graphical representation to extract valuable insights into solving the problem. And by having this information extracted, it enables universities to perform better at important decision makings. The study findings given in this research show that using Business Intelligence tools to analyse existing data and derive valuable information for decision making has a lot to promise. Students are also expected to improve their ways of learning instead of relying on the notes and learning materials provided by lecturer. The materials itself are insufficient to gain full understanding on the subject. More practice and self-experimenting on coding is required to equip with the knowledge to place themselves more competitive. Furthermore, the bloom taxonomy contributed much beneficial outcomes by outlining the learning objectives for students and lectures to further understand the purpose of the course. With it, students' weaknesses are justified based on the domains and able to find out which area they poorly performed at. Besides, by getting additional data to the original dataset, the student performance analysis might be significantly enhanced. For instance, it will be informative to see and compare the results of how other students performed in the other university courses they take during the semester to enhance the further analysis and discovering greater insights. This will aid in a greater understanding of the students' performance and learning capabilities. Finally, leading educational institutions throughout the world must employ innovative methods to tackle the problems and new possibilities that have arisen. Also, the overall research of this project has proven that gender, programme studied, background experience have an significant impact on the students' performance. The implementation of advanced analytical techniques, such as business intelligence (BI) systems and data mining tools and techniques, allows for more substantial use and analysis of obtainable university data, resulting in much more effective and efficient

performance, efficient governance, and informed choices focusing on clean, precise, stable, and consistent data.

7.2 Novelties and Contributions

First of all, one of the novelties of project is the **implementation of python** along with the use of anaconda navigator and Jupyter Notebook as my platform to perform my data visualization. This is to make use of programming tool to solve the problem rather than using in-built business intelligence tools, which other projects have already been done by others before. One of the very reasons why Python is the preferred choice for data analysis and visualization is due to its basic and simple syntax, tons of frameworks, and cutting the time needed to write a line of code, in which makes it an excellent programming tool. On the other hand, it provides support for data science libraries like Keras, Matplotlib and TenserFlow as they are ideally made for data gathering, analysis, visualization and data modelling. In other words, Python is an extremely popular programming tool which is a current trending language for data science and widely used globally. Besides, Jupyter Notebook was used as well as the platform for Python due its great environment where to develop code and also interacting with the results. It doesn't only work as an integrated development environment(IDE) but also a form of presentation as well.

Secondly, **bloom taxonomy** is used to in this project as an indicator to find out which area of weaknesses are the students into. The Bloom taxonomy is classified into the different domains, like cognitive, affective and psychomotor. For instance, cognitive domain is mainly focused on skills like intellectual abilities including analytical analysis, conflict resolution, and knowledge building. The affective domain emphasizes students' emotions, beliefs, passions, and admiration. Its linked hierarchy starts with taking in there and paying attention to information and progresses to description, or integrating values as well as continuously upholding them. It aims to assist students in comprehending their personal values and how they've evolved. While the includes all aspects of a student's physical capabilities for task completion, mobility, plus skill performance. These would be a great way to evaluate student's performance in their academics, which what most of the projects lacked off. Data visualization itself is not enough to achieve the results as it may not be accurately find out which area of weaknesses are the students into.

Lastly, one of the novelties is **comparison of available business intelligence tools on the markets.** It might be difficult to compare BI tools. Particularly considering the abundance of products currently available that all claim to offer a wide range of business advantages while also providing solutions to a wide range of issues. With all the available tools on the market, we must explore what is the current trending tool being adopted by most companies or organization for decision making, business intelligence most important criteria's, and to study what are the latest features and the significance of them to the project. Therefore, it is vital to do a comparison to find which tools may suit to solve the current problem.

The contributions of this project will assist lecturers in higher institution gain a better insight and understanding on the academic performance of students as well as knowing which area of weaknesses are they into so that students in higher education could be able to have a huge leap in academic performance.

7.3 Future work

In the future work, more complex business intelligence tools will be used like Tableau as it is a growing and trending tool implemented in the organizations for decisional making. Also, there will be more novel technologies and features introduced like data governance, self-service BI, collaborative and integrative BI and so many more. Therefore, to remain competitive in the data and analytics effort, these trends must be adopted and embraced and strengthening today's competitive business world. Aside from that, existing sample data from other higher institutions will be analysed as more meaningful data provides a better and concrete decision making. This will includes data analysis on different subjects as well to find out more justifiable information on the area of weaknesses the students are into. Also, Microsoft SQL Server would be considered to be implemented in the future as the database because it is well known and trending among data science industry and also extremely compatible and suited with Microsoft Azure as well as Microsoft's Business Intelligence tools. Hence, this database is designed to help data analysts who need to access enormous datasets quickly and effectively and as large amount of data will be needed in future project, this would be the right tool to use. SQL Server has the capability to handle a variety of datasets, particularly non-relational but also unstructured data, whereas most systems concentrate on managing organized and hierarchical information.

References

[1] R. Agarwal, V. Dhar, “big data, data science, and analytics: The opportunity and challenge for IS research, Information Systems Research”, 2014, pp. 443-448.

[2] A. Shollo, K. Kautz, “Towards an understanding of business intelligence, In the Proceedings of the 21st Australian Conference in Information Systems (ACIS)”, 2010.

[3] C.Carlos, M.Steven, R.Peter, “Database Systems: Design, Implementation, and Management, Course Technology”, 2014, ch.13

[4] “The top 7 Benefits of Business Intelligence (BI)”. [Online] Available: https://www.tableau.com/learn/articles/business-intelligence/enterprise-business-intelligence/benefits Accessed: 11 March 2022

[5] Arpaci, I, “Understanding and predicting students’ intention to use mobile cloud storage services”, 2016, pp.150–157.

[6] D. Apraxine and E. Stylianou, “Business Intelligence in a Higher Educational,” 2017, pp. 1735–1746.

[7] “What is Cloud BI,” Accessed on Mar. 11, 2022. [Online]. Available: https://www.domo.com/glossary/what-is-cloud Accessed: 23 November 2021

[8] Debattista, J. Londoño, S. Lange, C. Auer, “A Framework for Linked Data Quality Assessment”, 2016. [Online] Available: https://www.researchgate.net/publication/269418019_Luzzu_- Accessed: 23 November 2021

[9] Gartner. (Apparently of 2018b), “IT Glossary”, 2016. [Online] Available: https://www.gartner.com/it-glossary/self-service-business-intelligence Accessed: 24 November 2021

[10] Informatec, “Work more effectively: With Self-service Business Intelligence”, 2018. [Online] Available: https://www.informatec.com/en/effective-selfservice-bi Accessed: 24 November 2021

[11] Norman, Patricia M, "Are Your Secrets Safe? Knowledge Protection in Strategic Alliances." ,2001. [Online] Available: https://www.sciencedirect.com/science/article/pii/S0007681301800732 Accessed: 14 November 2021

[12] Marelli, L. Testa, G, “Scrutinizing the EU General Data Protection Regulation”, 2018. [Online] Available: https://www.researchgate.net/publication/324936977_Scrutinizing_the_EU_General_Data_Protection_Regulation Accessed: 24 February 2022

[13] Gartner. (January 2017), “Cloud Computing Enters its Second Decade”, 2017. [Online] Available: https://www.gartner.com/smarterwithgartner/cloudcomputing-enters-its-second-decade/ Accessed: 24 November 2021

[14] García, J./Technology Evaluation Centers, “Mobile BI: Features, Challenges, and Opportunities”, 2010. [Online] Available: https://www.whitepapers.em360tech.com/wp Accessed: 24 November 2021

[15] BARC, “The Business Application Research Center. (November 2018a). BI Trend Monitor 2019”. [Online] Available: http://barcresearch.com/research/bi-trend-monitor Accessed: 24 November 2021

[16] Bitner, S./Sisense, “Does the Future Lie with Embedded BI”. [Online] Available: https://www.sisense.com/blog/future-lie-embedded-bi/ Accessed: 24 November 2021

[17] Dykes, B./Forbes, “Data Storytelling: The Essential Data Science Skill Everyone Needs”, 2016 [Online] Available: https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-theessential-data-science-skill-everyone-needs/#42e8023752ad Accessed: 24 November 2021

[18] Micek, B./CARTO, “How Entrepreneurs are Using Open Data to Start Businesses” , 2017 [Online] Available: https://carto.com/blog/howentrepreneurs-using-open-data-start-businesses/ Accessed: 21 November 2021

[19] BARC, “The Business Application Research Center. (Apparently of 2018c). Data Quality and Master Data Management: How to Improve Your Data Quality”, 2016 [Online] Available: https://bi survey.com/data-quality-master-data-management Accessed: 21 November 2021

[20] Galetto, M./NGDATA, “What is Big Data Analytics?”, 2016, [Online] Available: https://www.ngdata.com/what-is-big-data-analytics/ Accessed: 23 November 2021

[21] Rajesh, & Ramesh, “An introduction to data lake. I-Manager's Journal on Information Technology, 5(2), 1”, 2016 [Online] Available: http://www.splunk.com/web_assets/pdfs/secure/Splunk_for_BigData.pdf Accessed: 21 November 2021

[22] Shepherd, A. Kesa, C. Cooper, J. Onema, J. Kovacs, P, “Opportunities and challenges associated with implementing Data lakes for enterprise decisionmaking”, 2018 [Online] Available: http://www.iacis.org/iis/2018/1_iis_2018_48-57.pdf Accessed: 21 November 2021

[23] Cambria, Erik, “*Affective computing and sentiment analysis. IEEE Intelligent Systems*”, 2016, sec 31, pp.102-107

[24] Chandrayan, P, “*A Guide To NLP : A Confluence Of AI And Linguistics*", 2017 [Online] Available: https://codeburst.io/a-guide-to-nlp-a-confluence-of-aiand-linguistics 2786c56c0749 Accessed: 21 November 2021

[25] B. Wieder , M.-L. Ossimitz, “The Impact of Business Intelligence on the Quality of Decision Making – A Mediation Model,” Procedia Comput, vol. 64, 2015, pp. 1163–1171

[26] D. Arnott, F. Lizama, and Y. Song, “Patterns of business intelligence systems use in organizations,” Decis. Support Syst., vol. 97, 2017, pp. 58–68.

[27] D. Apraxine and E. Stylianou, “Business Intelligence in a Higher Educational”, 2017, pp. 1735–1746.

[28] A. B. Sangar and N. B. a. Iahad, “Critical Factors that Affect the Success of Business Intelligence Systems (BIS) Implementation in an Organization” Int. J. Sci. Technol. Res., vol. 2, no. 2, pp. 176–180.

[29] N. M. A. and M. A. M. Ahmad Fadhil Yusof, Suraya Miskon, Norasnita Ahmad, Rose Alinda Alias, Haslina Hashim, Norris Syed Abdullah, D 2015, “Implementation issues affecting the business intelligence adoption in public university,” ARPN J. Eng. Appl. Sci., vol. 10,no. 23.

[30] P. Wanda and S. Stian, “The Secret of my Success: An exploratory study of Business Intelligence management in the Norwegian Industry,” Procedia Comput. Sci., vol. 64, 2015, pp. 240–247.

[31] V. Trieu, “Getting value from Business Intelligence systems : A review and research agenda,” Decis. Support Syst., vol. 93, 2017, pp. 111–124

[32] A. I. Magaireah, H. Sulaiman, and N. Ali, “Theoretical framework of critical success factors (CSFs) for Business Intelligence (BI) System,” in *2017 8th International Conference on Information Technology (ICIT)*, 2017, pp. 455–463.

[33] A. H. Anjariny, A. M. Zeki, and A. I. Abubakar, “The Mediating Effect of Teamwork toward Organizational Readiness for Business Intelligence (BI) Implementation,”, 2015, pp. 176–181.

[34] J.-P. Dawson, L., Van Belle, “Critical success factors for business intelligence in the South African financial services sector,” SA J. Inf. Manag., vol. 15, no. 1, 2013, pp. 1–12,

[35] Kartika.M, Prismahardi.R, Tresna.F,Kristya.M, “Business Intelligence For Educational Institution : A Literature Review”, 2020, pp.22-25 [Online] Available: https://www.ijconsist.org/index.php/ijconsist/article/view/32 Accessed: 21 November 2021

[36] *DataDrivenCulture.,* [Online] Available from https://www.datadrivenculture.dk/ .Accessed: 21 November 2021

[37] Harris, D./Software Advice. (n.d.), “4 Emerging Use Cases for IoT Data Analytics”. [Online] Available: https://www.softwareadvice.com/resources/iot-dataanalytics-use-cases/ Accessed: 23 November 2021

[38] TDWI. (n.d.), “TDWI GLOSSARY: What is a data lake?”.[Online] Available: https://tdwi.org/portals/what-is-a-data-lake-definition.aspx Accessed: 23 November 2021

Sample data of students for the course UCCD9999

	Sample data for a course											
	https://educerecentre.com/what-are-the-three-domains-of-blooms-ta				Coursework 50%			Final exam 50%				
			Contribution to overall marks		20%	10%	20%	25%	25%	25%	25%	
					100 marks	100 marks	100 marks	25 marks	25 marks	25 marks	25 marks	
					cognitive domain	psychomotor domain	cognitive domain	cognitive domain	cognitive domain	cognitive domain	cognitive domain	
					Learning outcome 2	Learning outcome 4	Learning outcome 3	Learning outcome 1	Learning outcome 1	Learning outcome 2	Learning outcome 3	
1	Course	Programm	StudId	StudName	Midterm Test	Practical Test	Assignment	Final exam Q1	Final exam Q2	Final exam Q3	Final exam Q4	
2	UCCD9999	CS	S1	S1	57	86	78	19	23	22	16	
3	UCCD9999	CS	S2	S2	46	76	68	25	16	18	14	
4	UCCD9999	CS	S3	S3	65	90.5	68	20	20	24	14	
5	UCCD9999	CS	S4	S4	67	78	74	21	24	22	13	
6	UCCD9999	CS	S5	S5	35	84.5	70	18	13	17	10	
7	UCCD9999	CS	S6	S6	62	55	72	23	16	13	14	
8	UCCD9999	CS	S7	S7	59	46	72	23	16	22	17	
9	UCCD9999	CS	S8	S8	64	66.5	74	23	17	20	17	
10	UCCD9999	CS	S9	S9	62	75	69	19	13	7	11	
11	UCCD9999	CS	S10	S10	74	87	80	23	23	16	10	
12	UCCD9999	CS	S11	S11	66	85.5	69	23	17	20	11	
13	UCCD9999	CS	S12	S12	49	40	70	20	19	15	8	
14	UCCD9999	CS	S13	S13	59	82.5	70	20	19	20	13	
15	UCCD9999	CS	S14	S14	63	65	76	23	16	12	15	
16	UCCD9999	CS	S15	S15	58	72.5	77	23	20	22	15	
17	UCCD9999	CS	S16	S16	55	46.5	71	20	14	12	7	
18	UCCD9999	CS	S17	S17	64	89	74	22	22	24	20	
19	UCCD9999	CS	S18	S18	65	70.5	70	20	20	22	14	
20	UCCD9999	CS	S19	S19	74	91	80	23.5	12	18	14	
21	UCCD9999	CS	S20	S20	39	59	74	20	11	21	9	
22	UCCD9999	CS	S21	S21	52	73.5	76	18	21	14	8	
23	UCCD9999	CS	S22	S22	61	60.5	70	18	18	19	15	
24	UCCD9999	CS	S23	S23	48	53	76	23	10	18	9	
25	UCCD9999	CS	S24	S24	54	51.5	70	8	20	6	14	
26	UCCD9999	CS	S25	S25	58	56.5	70	19	17	20	12	
27	UCCD9999	CS	S26	S26	59	32	72	22	16	14	14	
28	UCCD9999	CS	S27	S27	54	71.5	71	22	21	22	15	
29	UCCD9999	CS	S28	S28	58	51.5	74	22	13	16	13	
30	UCCD9999	CS	S29	S29	50	76	75	21	18	19	14	
31	UCCD9999	CS	S30	S30	62	67.5	72	23	13	18	15	
32	UCCD9999	CS	S31	S31	26	41.5	70	17	19	12	11	
33	UCCD9999	CS	S32	S32	62	72	68	24	16	22	16	
34	UCCD9999	CS	S33	S33	63	27	74	21	15	16	15	
35	UCCD9999	CS	S34	S34	40	55.5	74	19	20	6	9	
36	UCCD9999	CS	S35	S35	70	63.5	70	19	12	21	20	
37	UCCD9999	CS	S36	S36	48	52.5	72	20	16	20	16	
38	UCCD9999	CS	S37	S37	52	55.5	70	16	13	0	9	
39	UCCD9999	CS	S38	S38	74	75.5	70	21	20	19	16	
40	UCCD9999	CS	S39	S39	52	50	68	14	20	20	13	
41	UCCD9999	CS	S40	S40	38	45	72	23	20	12	0	
42	UCCD9999	CS	S41	S41	74	86	74	22	10	21	16	
43	UCCD9999	CS	S42	S42	51	74.5	70	20	18	22	13	
44	UCCD9999	CS	S43	S43	54	48.5	71	20	19	16	6	
45	UCCD9999	CS	S44	S44	62	59	78	21	13	20	18	
46	UCCD9999	CS	S45	S45	69	78.5	70	20	16	20	14	
47	UCCD9999	CS	S46	S46	64	87.5	69	19	10	11	12	
48	UCCD9999	CS	S47	S47	69	98	74	18	9	16	14	
49	UCCD9999	CS	S48	S48	66	82	72	20	24	18	18	
50	UCCD9999	CS	S49	S49	60	64.5	72	25	19	21	10	
51	UCCD9999	CS	S50	S50	45	67	77	21	15	20	10	
52	UCCD9999	CS	S51	S51	58	86	76	23	12	22	16	
53	UCCD9999	CS	S52	S52	42	57.5	70	23	19	5	9	
54	UCCD9999	CS	S53	S53	61	48	71	18	16	14	13	
55	UCCD9999	CS	S54	S54	57	53.5	70	24	19	7	10	
56	UCCD9999	CS	S55	S55	50	46	65	19	6	22	5	
57	UCCD9999	CS	S56	S56	38	67.5	74	23	22	19	14	
58	UCCD9999	CS	S57	S57	56	68	65	19	16	7	11	
59	UCCD9999	CS	S58	S58	72	64.5	70	23	9	12	13	
60	UCCD9999	CS	S59	S59	56	44	80	18	12	12	5	
61	UCCD9999	CS	S60	S60	64	85	76	23	13	22	14	
62	UCCD9999	CS	S61	S61	45	63	70	19	24	8	2	
63	UCCD9999	CS	S62	S62	52	76.5	75	22	12	18	9	
64	UCCD9999	CS	S63	S63	63	45.5	71	21	22	20	15	
65	UCCD9999	CS	S64	S64	35	46.5	70	23	17	17	13	
66	UCCD9999	CS	S65	S65	49	43	71	20	11	12	7	
67	UCCD9999	CS	S66	S66	58	54	65	19	17	22	10	
68	UCCD9999	CS	S67	S67	49	51.5	77	20	10	12	8	
69	UCCD9999	CS	S68	S68	88	91	84	25	18	23	20	
70	UCCD9999	CS	S69	S69	64	44.5	65	23	14	13	12	
71	UCCD9999	CS	S70	S70	54	71	75	21	13	22	20	
72	UCCD9999	CS	S71	S71	48	28	73	22	10	10	3	
73	UCCD9999	CS	S72	S72	68	75.5	75	20	16	20	15	
74	UCCD9999	CS	S73	S73	65	81.5	81	23	16	12	18	
75	UCCD9999	CS	S74	S74	31	73.5	73	15	13	16	11	
76	UCCD9999	CS	S75	S75	57	19.5	70	25	13	6	11	
77	UCCD9999	CS	S76	S76	40	48.5	69	17	19	20	2	
78	UCCD9999	CS	S77	S77	75	74.5	72	25	18	19	9	
79	UCCD9999	CS	S78	S78	75	52.5	70	22	6	21	15	

80	UCCD9999	CS	S79	S79	69	74	81	23	18	20	16
81	UCCD9999	CS	S80	S80	56	43.5	70	25	17	14	15
82	UCCD9999	CS	S81	S81	58	83	82	21	17	14	11
83	UCCD9999	CS	S82	S82	67	71	81	22	18	14	16
84	UCCD9999	CS	S83	S83	74	89.5	79	23	16	15	18
85	UCCD9999	CS	S84	S84	64	60	80	18	21	22	9
86	UCCD9999	CS	S85	S85	52	49.5	74	23	19	2	10
87	UCCD9999	CS	S86	S86	52	74	72	23	16	14	15
88	UCCD9999	CS	S87	S87	55	93.5	74	22	12	22	18
89	UCCD9999	CS	S88	S88	53	89.5	72	18	20	22	16
90	UCCD9999	CS	S89	S89	54	65	77	23	15	22	9
91	UCCD9999	CS	S90	S90	51	70.5	72	23	21	20	9
92	UCCD9999	CS	S91	S91	60	38	76	23	22	14	4
93	UCCD9999	CS	S92	S92	57	75.5	69	23	15	17	14
94	UCCD9999	CS	S93	S93	52	75.5	71	18	16	18	13
95	UCCD9999	CS	S94	S94	42	31	65	17	9	14	12
96	UCCD9999	CS	S95	S95	53	65	71	23	14	18	10
97	UCCD9999	CS	S96	S96	56	78	78	23	20	13	14
98	UCCD9999	CS	S97	S97	58	41	72	22	14	16	8
99	UCCD9999	CS	S98	S98	62	85	74	14	10	22	16
100	UCCD9999	CS	S99	S99	64	50.5	72	17	16	20	17
101	UCCD9999	CS	S100	S100	52	43	72	23	18	18	3
102	UCCD9999	CS	S101	S101	43	31	70	17	15	18	2
103	UCCD9999	CS	S102	S102	47	30	69	20	19	8	6
104	UCCD9999	CS	S103	S103	33	62.5	69	19	11	10	8
105	UCCD9999	CS	S104	S104	67	78	80	20	12	20	16
106	UCCD9999	CS	S105	S105	59	53.5	82	25	18	16	16
107	UCCD9999	CS	S106	S106	28	38.5	70	22	12	16	10
108	UCCD9999	CS	S107	S107	70	81	69	23	16	16	17
109	UCCD9999	CS	S108	S108	52	47.5	66	20	14	14	9
110	UCCD9999	CS	S109	S109	61	85	65	25	19	20	11
111	UCCD9999	CS	S110	S110	38	36.5	69	20	17	10	7
112	UCCD9999	CS	S111	S111	57	92.5	71	22	17	18	16
113	UCCD9999	CS	S112	S112	74	95.5	72	23	21	22	16
114	UCCD9999	CS	S113	S113	60	87	82	19	25	20	10
115	UCCD9999	CS	S114	S114	66	61.5	70	21	20	21	18
116	UCCD9999	CS	S115	S115	57	71	69	25	22	12	13
117	UCCD9999	CS	S116	S116	58	61	72	19	23	12	2
118	UCCD9999	CS	S117	S117	74	71.5	69	23	20	21	11
119	UCCD9999	CS	S118	S118	65	80.5	73	21	19	22	13
120	UCCD9999	CS	S119	S119	72	56.5	72	15	21	20	14
121	UCCD9999	CS	S120	S120	55	49	70	21	15	18	10

122	UCCD9999	CS	S121	S121	50	32	69	20	10	10	7
123	UCCD9999	CS	S122	S122	66	79.5	74	22	17	14	17
124	UCCD9999	CS	S123	S123	41	39	70	18	18	0	5
125	UCCD9999	CS	S124	S124	51	52.5	80	23	18	12	10
126	UCCD9999	CS	S125	S125	60	72	70	20	18	16	12
127	UCCD9999	CS	S126	S126	58	71.5	71	23	14	16	13
128	UCCD9999	CS	S127	S127	52	61	68	0	0	0	0
129	UCCD9999	CS	S128	S128	43	38	70	21	17	12	3
130	UCCD9999	CS	S129	S129	35	38	67	9	14	22	8
131	UCCD9999	CS	S130	S130	56	60	75	22	19	20	14
132	UCCD9999	CS	S131	S131	38	49.5	77	23	19	18	9
133	UCCD9999	CS	S132	S132	58	58	72	25	17	12	11
134	UCCD9999	CS	S133	S133	52	89.5	74	20	19	23	16
135	UCCD9999	CS	S134	S134	72	98	77	23	22	18	16
136	UCCD9999	CS	S135	S135	41	65	70	18	23	14	10
137	UCCD9999	CS	S136	S136	66	57	72	23	18	9	9
138	UCCD9999	CS	S137	S137	61	66	70	23	22	13	11
139	UCCD9999	CS	S138	S138	37	51.5	70	18	18	7	10
140	UCCD9999	CS	S139	S139	64	52	72	22	19	20	16
141	UCCD9999	CS	S140	S140	55	40	70	22	19	20	15
142	UCCD9999	CS	S141	S141	65	41	68	20	15	18	19
143	UCCD9999	CS	S142	S142	65	61	70	23	23	12	15
144	UCCD9999	CS	S143	S143	72	54.5	74	12	14	23	14
145	UCCD9999	CS	S144	S144	60	67.5	70	22	16	20	13
146	UCCD9999	CS	S155	S155	51	50.5	70	23	15	16	14
147	UCCD9999	CS	S156	S156	66	63	76	18	25	14	9
148	UCCD9999	CS	S157	S157	55	76	72	23	13	12	17
149	UCCD9999	CS	S158	S158	59	42	72	20	8	16	10
150	UCCD9999	CS	S159	S159	62	54	71	23	23	20	15
151	UCCD9999	CS	S160	S160	63	73	76	23	21	14	15
152	UCCD9999	CS	S161	S161	72	76	72	23	14	18	16
153	UCCD9999	CS	S162	S162	72	78.5	72	23	14	20	13
154	UCCD9999	CS	S163	S163	58	71.5	69	23	13	20	17
155	UCCD9999	CS	S164	S164	62	43.5	80	22	8	15	3
156	UCCD9999	CS	S165	S165	55	73.5	72	21	12	21	5
157	UCCD9999	CS	S166	S166	35	38	72	13	16	14	10
158	UCCD9999	CS	S167	S167	60	49	70	18	14	14	15
159	UCCD9999	CS	S168	S168	91	97	84	25	21	22	19
160	UCCD9999	CS	S169	S169	54	53.5	72	23	17	15	16
161	UCCD9999	CS	S170	S170	61	78	70	19	11	16	6
162	UCCD9999	CS	S171	S171	52	44.5	65	19	9	4	8
163	UCCD9999	CS	S176	S176	61	62	65	23	12	15	11

158	UCCD9999	CS	S167	S167	60	49	70	18	14	14	15
159	UCCD9999	CS	S168	S168	91	97	84	25	21	22	19
160	UCCD9999	CS	S169	S169	54	53.5	72	23	17	15	16
161	UCCD9999	CS	S170	S170	61	78	70	19	11	16	6
162	UCCD9999	CS	S171	S171	52	44.5	65	19	9	4	8
163	UCCD9999	CS	S176	S176	61	62	65	23	12	15	11
164	UCCD9999	CS	S177	S177	77	89.5	69	21	11	23	11
165	UCCD9999	CS	S178	S178	60	75.5	72	21	20	22	16
166	UCCD9999	CS	S179	S179	43	34.5	76	21	15	16	9
167	UCCD9999	CS	S180	S180	60	65.5	70	23	21	22	13
168	UCCD9999	CS	S181	S181	62	62	70	23	13	18	13
169	UCCD9999	CS	S182	S182	66	59.5	74	25	21	16	16
170	UCCD9999	CS	S183	S183	64	58.5	72	21	19	21	18
171	UCCD9999	CS	S189	S189	61	64	72	19	20	21	11
172	UCCD9999	CS	S190	S190	56	71.5	65	25	15	14	9
173	UCCD9999	CS	S191	S191	66	65	74	25	23	24	18
174	UCCD9999	CS	S192	S192	58	84	72	23	20	15	17
175	UCCD9999	CS	S193	S193	49	80	70	21	23	14	13
176	UCCD9999	CS	S194	S194	42	42.5	68	20	14	14	5
177	UCCD9999	CS	S195	S195	44	29	70	19	18	0	3
178	UCCD9999	CS	S196	S196	42	36.5	70	23	8	6	6
179	UCCD9999	CS	S197	S197	61	81.5	73	25	18	20	6
180	UCCD9999	CS	S198	S198	56	73.5	69	22	18	20	10
181	UCCD9999	CS	S199	S199	55	89.5	72	23	20	20	17
182	UCCD9999	CS	S200	S200	64	37	69	22	9	18	16
183	UCCD9999	CS	S201	S201	70	22.5	72	19	16	14	6
184	UCCD9999	CS	S202	S202	41	42	77	24	16	20	11
185	UCCD9999	CS	S203	S203	51	57	72	23	17	20	14
186	UCCD9999	CS	S204	S204	68	52	69	23	10	23	14
187	UCCD9999	CS	S205	S205	64	88	76	24	17	22	14
188	UCCD9999	CS	S206	S206	65	82	71	20.5	17	18	10
189	UCCD9999	CS	S207	S207	57	54	64	23	8	4	5
190	UCCD9999	CS	S208	S208	65	78.5	71	7	14	13	15
191	UCCD9999	IA	S209	S209	39	23.5	69	22	13	14	7
192	UCCD9999	IA	S210	S210	59	59.5	65	23	15	14	15
193	UCCD9999	IA	S211	S211	20	30.5	69	0	0	0	0
194	UCCD9999	IA	S212	S212	49	49	69	19	14	15	11
195	UCCD9999	IA	S213	S213	58	38	70	23	17	14	10
196	UCCD9999	IA	S214	S214	41	28	66	24	6	15	11
197	UCCD9999	IA	S215	S215	49	22.5	66	23	16	12	7
198	UCCD9999	IA	S216	S216	41	31	70	22	12	18	6
199	UCCD9999	IA	S217	S217	42	44.5	70	18	23	6	4

200	UCCD9999	IA	S218	S218	54	44.5	66	24	7	14	11
201	UCCD9999	IA	S219	S219	67	61.5	70	24	15	16	10
202	UCCD9999	IA	S220	S220	45	28	80	22	21	14	6
203	UCCD9999	IA	S221	S221	39	56	70	19	12	13	5
204	UCCD9999	IA	S222	S222	56	30.5	65	22	13	16	7
205	UCCD9999	IA	S223	S223	46	49	70	21	15	12	10
206	UCCD9999	IA	S224	S224	34	34	63	25	14	13	15
207	UCCD9999	IA	S225	S225	36	22	63	0	0	0	0
208	UCCD9999	IA	S226	S226	68	50.5	69	21	12	16	6
209	UCCD9999	IA	S227	S227	69	57.5	65	19	19	12	15
210	UCCD9999	IA	S228	S228	29	53	66	18	17	2	8
211	UCCD9999	IA	S229	S229	41	73	80	16	19	1	10
212	UCCD9999	IA	S230	S230	54	52	80	21	18	10	10
213	UCCD9999	IA	S231	S231	58	61	68	21	16	0	8
214	UCCD9999	IA	S232	S232	55	44	70	20	18	12	11
215	UCCD9999	IA	S233	S233	40	61	80	23	17	6	8
216	UCCD9999	IA	S234	S234	31	29	67	21	13	10	7
217	UCCD9999	IA	S235	S235	45	45	70	25	21	11	7
218	UCCD9999	IA	S236	S236	72	44	65	20	16	10	9
219	UCCD9999	IA	S237	S237	51	33	70	21	12	6	8
220	UCCD9999	IA	S238	S238	37	41.5	70	24	19	14	4
221	UCCD9999	IA	S239	S239	54	34.5	77	23	14	16	5
222	UCCD9999	IA	S240	S240	59	70.5	68	18	19	6	12
223	UCCD9999	IA	S241	S241	61	15	70	24	15	12	6
224	UCCD9999	IA	S242	S242	57	56	74	24	18	16	10
225	UCCD9999	IA	S243	S243	47	26.5	70	20	10	20	19
226	UCCD9999	IA	S244	S244	53	2	69	0	0	0	0
227	UCCD9999	IB	S245	S245	23	28	63	2.3	2.8	18.9	24
228	UCCD9999	IB	S246	S246	27	53	71	2.7	5.3	21.3	29.3
229	UCCD9999	IB	S247	S247	40	67	71	4	6.7	21.3	32
230	UCCD9999	IB	S248	S248	29	54	69	2.9	5.4	20.7	29
231	UCCD9999	IB	S249	S249	33	32.5	70	3.3	3.25	21	27.55
232	UCCD9999	IB	S250	S250	53	43	69	5.3	4.3	20.7	30.3
233	UCCD9999	IB	S251	S251	32	72.5	77	3.2	7.25	23.1	33.55
234	UCCD9999	IB	S252	S252	37	31	70	3.7	3.1	21	27.8
235	UCCD9999	IB	S253	S253	20	27	74	2	2.7	22.2	26.9
236	UCCD9999	IB	S254	S254	23	54.5	69	2.3	5.45	20.7	28.45
237	UCCD9999	IB	S255	S255	51	61.5	70	5.1	6.15	21	32.25
238	UCCD9999	IB	S256	S256	22	34	74	2.2	3.4	22.2	27.8
239	UCCD9999	IB	S257	S257	18	32	73	1.8	3.2	21.9	26.9
240	UCCD9999	IB	S258	S258	20	30	77	15	17	23.1	33.55

Code (dashboard 1 - midterm)

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('midterm',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('studID', 'midterm', data=sample_data_sorted_desc)
plt.title('Average of Midterm by StudId')
plt.xlabel('Student Id', size=15)
plt.ylabel('Average of midterm', size=15)
plt.legend(labels = ['Average midterm'])
#display plot
plt.show()

#piechart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['midterm'].mean()
groupby_mean1 = sample_data.groupby(['Programme']).mean()
print ('Mean midterm: ' + str(mean1))
Programme_data = sample_data["Programme"]
midterm_data = sample_data["midterm"]
plt.show()
print(groupby_mean1)
#Output the figure of piechart
# Import libraries
import numpy as np
import matplotlib.pyplot as plt
# Creating dataset
```

```
programme = ['IA', 'IB', 'CS']
data = [48.77, 30.57, 57.24]
# Creating explode data
explode = (0.1, 0.0, 0.2)
# Creating color parameters
colors = ( "orange", "cyan", "brown")
# Wedge properties
wp = { 'linewidth' : 1, 'edgecolor' : "green" }
# Creating autocpt arguments
def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.1f}%\n({:d} g)".format(pct, absolute)
# Creating plot
fig, ax = plt.subplots(figsize =(10, 7))
wedges, texts, autotexts = ax.pie(data,
                                  autopct = lambda pct: func(pct, data),
                                  explode = explode,
                                  labels = programme,
                                  shadow = True,
                                  colors = colors,
                                  startangle = 90,
                                  wedgeprops = wp,
                                  textprops = dict(color ="magenta"))
# Adding legend
ax.legend(wedges, programme,
          title ="Programmes",
          loc ="center left",
          bbox_to_anchor =(1, 0, 0.5, 1))

plt.setp(autotexts, size = 8, weight ="bold")
ax.set_title("The Average of midterm by programmes")

# show plot
plt.show()
```

```
#studentaboveOrEqualTo50
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols =
['studID','midterm', 'Programme'], low_memory = False)
dataframe.head()
print(
  dataframe[
     (dataframe['midterm'] >= 50)
  ]
)

#studentbelow50
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols =
['studID','midterm', 'Programme'], low_memory = False)
dataframe.head()
print(
  dataframe[
     (dataframe['midterm'] < 50)
  ]

)

#Count of midterm by gender
import pandas as pd
data = pd.read_csv('Sample_data_for_a_sample_course.csv')
data.info()
data['Gender'].value_counts(ascending=False)
Gender = ['Male','Female']
Midterm = [154, 85]
plt.barh(Gender, Midterm, color=['blue', 'red'])
plt.title('Count of midterm by gender')
plt.ylabel('Gender')
```

```
plt.xlabel('Count of midterm')
plt.show()

#Scatter chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['midterm'].mean()
groupby_mean1 = sample_data.groupby(['Gender']).mean()
print ('Mean midterm: ' + str(mean1))
Programme_data = sample_data["Gender"]
midterm_data = sample_data["midterm"]
plt.show()
print(groupby_mean1)
x = ['female', 'male']
y = [54.63, 54.27]
plt.title('Average of Midterm by Gender')
plt.xlabel('Gender', size=15)
plt.ylabel('Average of midterm', size=15)
plt.scatter(x, y, color=['red', 'blue'])
plt.show()

#Area chart
import matplotlib.pyplot as plt
# line 1 points
x1 = ['CS', 'IA', 'IB']
y1 = [56.73, 49.39, 26.80]
# plotting the line 1 points
plt.plot(x1, y1, label = "Male")
# line 2 points
x2 = ['CS', 'IA', 'IB']
y2 = [58.10,45,32.67]
# plotting the line 2 points
plt.plot(x2, y2, label = "Female")
```

```
plt.xlabel('Programme')
# Set the y axis label of the current axis.
plt.ylabel('Average of midterm')
# Set a title of the current axes.
plt.title('Average of midterm by programme and gender')
# show a legend on the plot
plt.legend()
# Display a figure.
plt.show()

#Treemap
import pandas as pd
data = pd.read_csv('Sample_data_for_a_sample_course.csv')
data.info()
data['Programme'].value_counts(ascending=False)
pip install squarify
import matplotlib.pyplot as plt
import squarify
sizes=[189, 36, 14]
label=["CS", "IA", "IB"]
squarify.plot(sizes=sizes, label=label, alpha=0.6 )
plt.title('Average of Midterm by Programme')
plt.axis('off')
plt.show()
```

Code (dashboard 2 - practical)

```
#Bar chart
import pandas as pd
from matplotlib import pyplot as plt
```

```
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#clustered bar chart
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Practical',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('studID', 'Practical', data=sample_data_sorted_desc, color = ['maroon'])
plt.title('Average of Practical by StudId')
plt.xlabel('Student Id', size=15)
plt.ylabel('Average of practical', size=15)
plt.legend(labels = ['Average practical'])
#display plot
plt.show()

#Pie chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Practical'].mean()
groupby_mean1 = sample_data.groupby(['Programme']).mean()
print ('Mean practical: ' + str(mean1))
Programme_data = sample_data["Programme"]
practical_data = sample_data["Practical"]
plt.show()
print(groupby_mean1)
#Output the figure of piechart
# Import libraries
import numpy as np
import matplotlib.pyplot as plt
# Creating dataset
programme = ['IA', 'IB', 'CS']
data = [41.69, 44.28, 62.91]

# Creating explode data
```

```
explode = (0.1, 0.0, 0.2)
# Creating color parameters
colors = ( "blue", "yellow", "brown")
# Wedge properties
wp = { 'linewidth' : 1, 'edgecolor' : "green" }
# Creating autocpt arguments
def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.1f}%\n({:d} g)".format(pct, absolute)
# Creating plot
fig, ax = plt.subplots(figsize =(10, 7))
wedges, texts, autotexts = ax.pie(data,
                                  autopct = lambda pct: func(pct, data),
                                  explode = explode,
                                  labels = programme,
                                  shadow = True,
                                  colors = colors,
                                  startangle = 90,
                                  wedgeprops = wp,
                                  textprops = dict(color ="magenta"))
# Adding legend
ax.legend(wedges, programme,
          title ="Programmes",
          loc ="center left",
          bbox_to_anchor =(1, 0, 0.5, 1))

plt.setp(autotexts, size = 8, weight ="bold")
ax.set_title("The Average of practical by programmes")

# show plot
plt.show()



#Bar chart for practical by gender
```

```
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Practical'].mean()
groupby_mean1 = sample_data.groupby(['Gender']).mean()
print ('Mean midterm: ' + str(mean1))
Programme_data = sample_data["Gender"]
midterm_data = sample_data["Practical"]
plt.show()
print(groupby_mean1)
x = ['female', 'male']
y = [58, 59]
plt.title('Average of Practical by Gender')
plt.xlabel('Gender', size=15)
plt.ylabel('Average of practical', size=15)
plt.bar(x, y, color=['red', 'blue'])
plt.show()

#Horizontal bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
max1 = sample_data['Practical'].max()
groupby_max1 = sample_data.groupby(['Programme']).max()
print ('Max practical: ' + str(max1))
Programme_data = sample_data["Programme"]
practical_data = sample_data["Practical"]
plt.show()
print(groupby_max1)
x = ['CS', 'IA', 'IB']
y = [98, 73, 72.5]
plt.title('Max of Practical by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Max of Practical Test', size=15)
```

```
plt.barh(x, y, color=['orange', 'blue', 'yellow'])
plt.show()

#Funnel
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
max1 = sample_data['Practical'].max()
groupby_max1 = sample_data.groupby(['Gender']).max()
print ('Max practical: ' + str(max1))
Gender_data = sample_data["Gender"]
practical_data = sample_data["Practical"]
plt.show()
print(groupby_max1)

!pip install plotly
import plotly.express as px
data = dict(
    practical=[98, 98],
    gender=["Male", "Female"])
fig = px.funnel(data, x='practical', y='gender', color = ['blue', 'red'])
fig.update_layout(
    title={
    'text' : "Max of Practical Test by Gender",
    'y': 0.9,
    'x': 0.5,
    'xanchor': 'center',
    'yanchor': 'top'})
fig.show()

#who pass practical over or equal 50 marks
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols =
['Gender','Practical', 'Programme'], low_memory = False)
```

```
dataframe.head()
print(
  dataframe[
     (dataframe['Practical'] >= 50)
  ]
)

#who fail practical below 50
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols =
['Gender','Practical', 'Programme'], low_memory = False)
dataframe.head()
print(
  dataframe[
     (dataframe['Practical'] < 50)
  ]
)

#Treemap
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
median1 = sample_data['Practical'].median()
groupby_median1 = sample_data.groupby(['Programme']).median()
print ('Median practical: ' + str(median1))
Programme_data = sample_data["Programme"]
practical_data = sample_data["Practical"]
plt.show()
print(groupby_median1)
import matplotlib.pyplot as plt
import squarify
sizes=[63.5, 44, 38.5]
label=["CS", "IA", "IB"]
squarify.plot(sizes=sizes, label=label, alpha=0.6 )
```

```
plt.title('Median of Practical by Programme')
plt.axis('off')
plt.show()
```

Code (dashboard 3 - Assignment)

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Assignment',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('studID', 'Assignment', data=sample_data_sorted_desc)
plt.title('Average of Assignment by StudId')
plt.xlabel('Student Id', size=15)
plt.ylabel('Average of Assignment', size=15)
plt.legend(labels = ['Average Assignment'])
#display plot
plt.show()
```

```
#PieChart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Assignment'].mean()
groupby_mean1 = sample_data.groupby(['Programme']).mean()
print ('Mean assignment: ' + str(mean1))
Programme_data = sample_data["Programme"]
assignment_data = sample_data["Assignment"]
plt.show()
```

```
print(groupby_mean1)
#Output the figure of piechart
# Import libraries
import numpy as np
import matplotlib.pyplot as plt
# Creating dataset
programme = ['IA', 'IB', 'CS']
data = [69.69, 71.21, 72.18]
# Creating explode data
explode = (0.1, 0.0, 0.2)
# Creating color parameters
colors = ( "yellow", "green", "brown")
# Wedge properties
wp = { 'linewidth' : 1, 'edgecolor' : "green" }
# Creating autocpt arguments
def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.1f}%\n({:d} g)".format(pct, absolute)
# Creating plot
fig, ax = plt.subplots(figsize =(10, 7))
wedges, texts, autotexts = ax.pie(data,
                                  autopct = lambda pct: func(pct, data),
                                  explode = explode,
                                  labels = programme,
                                  shadow = True,
                                  colors = colors,
                                  startangle = 90,
                                  wedgeprops = wp,
                                  textprops = dict(color ="magenta"))
# Adding legend
ax.legend(wedges, programme,
          title ="Programmes",
          loc ="center left",
          bbox_to_anchor =(1, 0, 0.5, 1))
```

```
plt.setp(autotexts, size = 8, weight ="bold")
ax.set_title("The Average of assignment by programmes")

# show plot
plt.show()


#vertical bar graph
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Assignment'].mean()
groupby_mean1 = sample_data.groupby(['Gender']).mean()
print ('Mean assignment: ' + str(mean1))
Programme_data = sample_data["Gender"]
assignment_data = sample_data["Assignment"]
plt.show()
print(groupby_mean1)
x = ['female', 'male']
y = [71.65, 71.81]
plt.title('Average of Assignment by Gender')
plt.xlabel('Gender', size=15)
plt.ylabel('Average of assignment', size=15)
plt.bar(x, y, color=['red', 'blue'])
plt.show()


#who pass practical over or equal 50 marks
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols =
['Gender','Assignment', 'Programme'], low_memory = False)
dataframe.head()
print(
```

```
  dataframe[
     (dataframe['Assignment'] >= 50)
  ]
)
```

```
#Horizontal bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
max1 = sample_data['Assignment'].max()
groupby_max1 = sample_data.groupby(['Programme']).max()
print ('Max assignment: ' + str(max1))
Programme_data = sample_data["Programme"]
assignment_data = sample_data["Assignment"]
plt.show()
print(groupby_max1)
x = ['CS', 'IA', 'IB']
y = [84, 80, 77]
plt.title('Max of Assignment by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Max of Assignment Test', size=15)
plt.barh(x, y, color=['orange', 'blue', 'yellow'])
plt.show()
```

```
#Horizontal bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
min1 = sample_data['Assignment'].min()
groupby_min1 = sample_data.groupby(['Programme']).min()
print ('Min assignment: ' + str(min1))
Programme_data = sample_data["Programme"]
```

```
assignment_data = sample_data["Assignment"]
plt.show()
print(groupby_min1)
x = ['CS', 'IA', 'IB']
y = [64, 63, 63]
plt.title('Min of Assignment by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Min of Assignment Test', size=15)
plt.barh(x, y, color=['orange', 'blue', 'yellow'])
plt.show()



#Treemap
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
median1 = sample_data['Assignment'].median()
groupby_median1 = sample_data.groupby(['Programme']).median()
print ('Median practical: ' + str(median1))
Programme_data = sample_data["Programme"]
assignment_data = sample_data["Assignment"]
plt.show()
print(groupby_median1)
import matplotlib.pyplot as plt
import squarify
sizes=[72, 70, 71]
label=["CS", "IA", "IB"]
squarify.plot(sizes=sizes, label=label, alpha=0.6 )
plt.title('Median of Assignment by Programme')
plt.axis('off')
plt.show()



#Funnel
```

```
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
max1 = sample_data['Assignment'].max()
groupby_max1 = sample_data.groupby(['Gender']).max()
print ('Max assignment: ' + str(max1))
Gender_data = sample_data["Gender"]
assignment_data = sample_data["Assignment"]
plt.show()
print(groupby_max1)

!pip install plotly
import plotly.express as px
data = dict(
    Assignment=[84, 82],
    gender=["Male", "Female"])
fig = px.funnel(data, x='Assignment', y='gender', color = ['male', 'female'])
fig.update_layout(
    title={
    'text' : "Max of Assignment by Gender",
    'y': 0.9,
    'x': 0.5,
    'xanchor': 'center',
    'yanchor': 'top'})
fig.show()
```

Code (dashboard 4 - Final score)

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
```

```
sample_data_sorted_desc= sample_data.sort_values('Final score',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('studID', 'Final score', data=sample_data_sorted_desc)
plt.title('Average of Final score by StudId')
plt.xlabel('Student Id', size=15)
plt.ylabel('Average of Final score', size=15)
plt.legend(labels = ['Average Final Score'])
#display plot
plt.show()

#PieChart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final score'].mean()
groupby_mean1 = sample_data.groupby(['Programme']).mean()
print ('Mean final score: ' + str(mean1))
Programme_data = sample_data["Programme"]
finalScore_data = sample_data["Final score"]
plt.show()
print(groupby_mean1)
#Output the figure of piechart
# Import libraries
import numpy as np
import matplotlib.pyplot as plt
# Creating dataset
programme = ['IA', 'IB', 'CS']
data = [52.78, 60.02, 65.59]
# Creating explode data
explode = (0.1, 0.0, 0.2)
# Creating color parameters
colors = ( "grey", "red", "brown")
# Wedge properties
```

```
wp = { 'linewidth' : 1, 'edgecolor' : "green" }
# Creating autocpt arguments
def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.1f}%\n({:d} g)".format(pct, absolute)
# Creating plot
fig, ax = plt.subplots(figsize =(10, 7))
wedges, texts, autotexts = ax.pie(data,
                                  autopct = lambda pct: func(pct, data),
                                  explode = explode,
                                  labels = programme,
                                  shadow = True,
                                  colors = colors,
                                  startangle = 90,
                                  wedgeprops = wp,
                                  textprops = dict(color ="magenta"))
# Adding legend
ax.legend(wedges, programme,
          title ="Programmes",
          loc ="center left",
          bbox_to_anchor =(1, 0, 0.5, 1))

plt.setp(autotexts, size = 8, weight ="bold")
ax.set_title("The Average of final score by programmes")

# show plot
plt.show()



#who pass finals over or equal 40 marks
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols = ['Gender','Final
score', 'Programme'], low_memory = False)
dataframe.head()
```

```
print(
  dataframe[
     (dataframe['Final score'] >= 40)
  ]
)
```

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Final Exam Q4',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('Programme', 'Final Exam Q4', data=sample_data_sorted_desc)
plt.title('Average of Final Exam Q4 by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Average of Final Exam Q4', size=15)
plt.legend(labels = ['Average Final Exam Q4'])
#display plot
plt.show()
```

```
#Treemap
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final Exam Q1'].mean()
mean2 = sample_data['Final Exam Q2'].mean()
mean3 = sample_data['Final Exam Q3'].mean()
mean4 = sample_data['Final Exam Q4'].mean()
print ('Average final exam of Q1,Q2,Q3,Q4: ' + str(mean1) + str(mean2) + str(mean3) +
str(mean4))
```

```
final1_data = sample_data["Final Exam Q1"]
final2_data = sample_data["Final Exam Q2"]
final3_data = sample_data["Final Exam Q3"]
final4_data = sample_data["Final Exam Q4"]
plt.show()
import matplotlib.pyplot as plt
import squarify
sizes=[19.71, 15.36, 15.82, 12.43]
label=["Q1", "Q2", "Q3", "Q4"]
squarify.plot(sizes=sizes, label=label, alpha=0.6 )
plt.title('Average of Final exam Q1, Q2, Q3, Q4')
plt.axis('off')
plt.show()


#vertical bar graph
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final score'].mean()
groupby_mean1 = sample_data.groupby(['Gender']).mean()
print ('Mean Final score: ' + str(mean1))
Programme_data = sample_data["Gender"]
finalScore_data = sample_data["Final score"]
plt.show()
print(groupby_mean1)
x = ['female', 'male']
y = [63.92, 63.01]
plt.title('Average of Final score by Gender')
plt.xlabel('Gender', size=15)
plt.ylabel('Average of final score', size=15)
plt.bar(x, y, color=['red', 'blue'])
plt.show()
```

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Final Exam Q1',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('Programme', 'Final Exam Q1', data=sample_data_sorted_desc, color = ['green'])
plt.title('Average of Final Exam Q1 by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Average of Final Exam Q1', size=15)
plt.legend(labels = ['Average Final Exam Q1'])
#display plot
plt.show()
```

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Final Exam Q2',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('Programme', 'Final Exam Q2', data=sample_data_sorted_desc, color = ['yellow'])
plt.title('Average of Final Exam Q2 by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Average of Final Exam Q2', size=15)
plt.legend(labels = ['Average Final Exam Q2'])
#display plot
plt.show()
```

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Final Exam Q3',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('Programme', 'Final Exam Q3', data=sample_data_sorted_desc, color = ['brown'])
plt.title('Average of Final Exam Q3 by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Average of Final Exam Q3', size=15)
plt.legend(labels = ['Average Final Exam Q3'])
#display plot
plt.show()
```

Code (dashboard 5 – final result)

```
#clustered bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
#sort data in descending order
sample_data_sorted_desc= sample_data.sort_values('Final result(100%)',ascending=False)
#specify width and height of figure
plt.figure(figsize = (13,7))
plt.bar('studID', 'Final result(100%)', data=sample_data_sorted_desc)
plt.title('Average of Final score by StudId')
plt.xlabel('Student Id', size=15)
plt.ylabel('Average of Final result', size=15)
plt.legend(labels = ['Average Final Result'])
#display plot
plt.show()
```

```
#PieChart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final result(100%)'].mean()
groupby_mean1 = sample_data.groupby(['Programme']).mean()
print ('Mean final result: ' + str(mean1))
Programme_data = sample_data["Programme"]
finalScore_data = sample_data["Final result(100%)"]
plt.show()
print(groupby_mean1)
#Output the figure of piechart
# Import libraries
import numpy as np
import matplotlib.pyplot as plt
# Creating dataset
programme = ['IA', 'IB', 'CS']
data = [54.25, 54.79, 64.97]
# Creating explode data
explode = (0.1, 0.0, 0.2)
# Creating color parameters
colors = ( "grey", "red", "brown")
# Wedge properties
wp = { 'linewidth' : 1, 'edgecolor' : "green" }
# Creating autocpt arguments
def func(pct, allvalues):
    absolute = int(pct / 100.*np.sum(allvalues))
    return "{:.1f}%\n({:d} g)".format(pct, absolute)
# Creating plot
fig, ax = plt.subplots(figsize =(10, 7))
wedges, texts, autotexts = ax.pie(data,
                                  autopct = lambda pct: func(pct, data),
```

```
                    explode = explode,
                    labels = programme,
                    shadow = True,
                    colors = colors,
                    startangle = 90,
                    wedgeprops = wp,
                    textprops = dict(color ="magenta"))
# Adding legend
ax.legend(wedges, programme,
        title ="Programmes",
        loc ="center left",
        bbox_to_anchor =(1, 0, 0.5, 1))

plt.setp(autotexts, size = 8, weight ="bold")
ax.set_title("The Average of final result by programmes")

# show plot
plt.show()



#who pass finals over or equal 50 marks
import pandas as pd
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols = ['Gender','Final
result(100%)', 'Programme'], low_memory = False)
dataframe.head()
print(
  dataframe[
     (dataframe['Final result(100%)'] >= 50)
  ]
)



#who fail finals below 50 marks
import pandas as pd
```

```
dataframe = pd.read_csv('Sample_data_for_a_sample_course.csv', usecols = ['Gender','Final result(100%)', 'Programme'], low_memory = False)
dataframe.head()
print(
  dataframe[
     (dataframe['Final result(100%)'] < 50)
  ]
)



#vertical bar graph
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final result(100%)'].mean()
groupby_mean1 = sample_data.groupby(['Gender']).mean()
print ('Mean Final result: ' + str(mean1))
Programme_data = sample_data["Gender"]
finalScore_data = sample_data["Final result(100%)"]
plt.show()
print(groupby_mean1)
x = ['female', 'male']
y = [63.03, 62.61]
plt.title('Average of Final result by Gender')
plt.xlabel('Gender', size=15)
plt.ylabel('Average of final result', size=15)
plt.bar(x, y, color=['red', 'blue'])
plt.show()



#Horizontal bar chart
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
```

```
max1 = sample_data['Final result(100%)'].min()
groupby_max1 = sample_data.groupby(['Programme']).max()
print ('Max final result: ' + str(max1))
Programme_data = sample_data["Programme"]
FinalResult_data = sample_data["Final result(100%)"]
plt.show()
print(groupby_max1)
x = ['CS', 'IA', 'IB']
y = [88.20, 63.73, 63.05]
plt.title('Max of final result by Programme')
plt.xlabel('Programme', size=15)
plt.ylabel('Max of final result', size=15)
plt.barh(x, y, color=['pink', 'purple', 'yellow'])
plt.show()


#Treemap
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
median1 = sample_data['Final result(100%)'].median()
groupby_median1 = sample_data.groupby(['Programme']).median()
print ('Median final result: ' + str(median1))
Programme_data = sample_data["Programme"]
FinalResult_data = sample_data["Final result(100%)"]
plt.show()
print(groupby_median1)
import matplotlib.pyplot as plt
import squarify
sizes=[66.70, 56.38, 53.15]
label=["CS", "IA", "IB"]
squarify.plot(sizes=sizes, label=label, alpha=0.6 )
plt.title('Median of Final result by Programme')
plt.axis('off')
```

```
plt.show()


#vertical bar graph
import pandas as pd
from matplotlib import pyplot as plt
sample_data = pd.read_csv('Sample_data_for_a_sample_course.csv')
mean1 = sample_data['Final result(100%)'].mean()
mean2 = sample_data['Coursework(50%)'].mean()
print ('Mean Final result: ' + str(mean1))
print ('Mean Coursework: ' + str(mean2))
finalScore_data = sample_data["Final result(100%)"]
coursework_data = sample_data["Coursework(50%)"]
plt.show()
x = ['coursework', 'final exam']
y = [31.09, 31.67]
plt.title('Average of Final result and coursework')
plt.ylabel('Average of final result and coursework', size=15)
plt.bar(x, y, color=['green', 'lightgreen'])
plt.show()
```

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 2**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Drawn the block diagram to convey the representation of the data analysis system in the form of blocks

2. WORK TO BE DONE

- Review the literature review and cite articles

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project

4. SELF EVALUATION OF THE PROGRESS

- Just started to revise the previous project from FYP1 and continuing the remaining tasks

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 3**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Do a literature review on the cite some articles related to business intelligence

2. WORK TO BE DONE

- Gather the sample raw data and visualize them into form of graphical representations by integrating into Power BI software

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project

4. SELF EVALUATION OF THE PROGRESS

- Slow progress but at a constant rate

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 4**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Successfully visualized the data in the form of graphical representations

2. WORK TO BE DONE

- Visualize the data by integrating into Python to generate existing results

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project

4. SELF EVALUATION OF THE PROGRESS

- 50% completion of progress and the still in mid-way to finish the project

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 5**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Successfully output the results by using matplotlib and pandas libraries of Python

2. WORK TO BE DONE

- Debugging and testing the code

3. PROBLEMS ENCOUNTERED

- Some codes have minor errors and bugs

4. SELF EVALUATION OF THE PROGRESS

- Slow progress and the completion of the project is at 60%

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 6**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Debugged the codes and ensure all problems are fixed

2. WORK TO BE DONE

- Construct the dashboards of Power BI to provide critical reporting better enhancement of analysis

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Slow but in progress and tasks are done daily

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 7**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Constructed the dashboard needed for Power BI showing the data of assignment result, final exam and overall results.

2. WORK TO BE DONE

- Construct the dashboard for Python to provide critical reporting better enhancement of analysis

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Slow but in still progress and tasks are done daily

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 8**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Constructed the dashboard for Python with all the required information needed

2. WORK TO BE DONE

- Analyze and discuss the results gathered based on the data shown in dashboard

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Project 80% completion and the visualization using Power BI is completed

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 9**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- The data is successfully analyzed with extracted information regarding the performance of student results

2. WORK TO BE DONE

- Provide recommendations to solve the problems identified throughout the analysis

3. PROBLEMS ENCOUNTERED

- Some minor bugs encountered and some errors faced while trying to code out the output.

4. SELF EVALUATION OF THE PROGRESS

- Project almost completed and there are only a few tasks left

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 10**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Recommendations are stated in the purpose to improve students' performance based on the weaknesses identified

2. WORK TO BE DONE

- Conclude the overall results throughout the project

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Project almost completed

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 11**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Conclusion has been completed based on the overview of the project

2. WORK TO BE DONE

- Novelties and contribution of the project

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Project almost completed

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 12**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Proposed the novelties of the project and what is its contribution

2. WORK TO BE DONE

- Design a poster to summarize the entire project. This will include the introduction, objective, methodology, results, analysis, recommendations and conclusion

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Project almost completed and only a few tasks left to go

Supervisor's signature

Student's signature

FINAL YEAR PROJECT WEEKLY REPORT

(Project II)

Trimester, Year: Trimester 3, Year 4	**Study week no.: 13**
Student Name & ID: Cheah Cheng Qian 17ACB04662	
Supervisor: Ts Dr Chan Lee Kwun	
Project Title: A Visualized Analysis on Student Performance by Using Cloud Business Intelligence	

1. WORK DONE

[Please write the details of the work done in the last fortnight.]

- Designed a poster for the project which summarizes the entire content of the project report
- Format the report

2. WORK TO BE DONE

- The project is completed and thus, there are no tasks left to be done.

3. PROBLEMS ENCOUNTERED

- There are no problems encountered so far upon completing the project.

4. SELF EVALUATION OF THE PROGRESS

- Project has been completed

Supervisor's signature

Student's signature

POSTER

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty of Information and Communication Technology

A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

During in this state of global pandemic, it is claimed that there is a rise in trend of poor academic performance among student in certain courses or other specific skills. In this matter, due to this uprising yet concerning issue, it is crucial that the deteriorations of students' academic performance seek a proper solution

Project Developer :
Cheah Cheng Qian

Project Supervisor :
Ts Dr Chan Lee Kwun

INTRODUCTION

Numerous studies have proven that higher education is critical to a country's economic progress and prosperity. As a result, data is critical for major businesses, particularly higher education, to make informed decisions. Because educational institutions have a high number of students, employees, and activities, BI tools can be used to evaluate a vast amount of data to forecast students' academic behavior, performance, and shortcomings. These statistics, for example, are crucial throughout the decision-making process because they may be utilized to communicate the need for change and improvement, which helps to improve the academic institution of a particular institution's student. BI solutions such as Power BI and Tableau give services and assistance that aid top management in completely analyzing complex business circumstances and streamlining decision-making processes.

OBJECTIVE

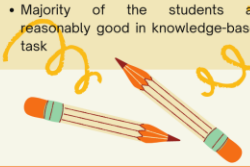
To find out what are the weaknesses of students in particular field of a certain subject and propose a recommendation to overcome the problem

METHODOLOGY

- Data collection and preparation
- Visualize the results using Power BI
- Visualize the existing results using Python
- The graphical presentations are analyzed and discussed
- Recommendations and formulation to improve students' performance

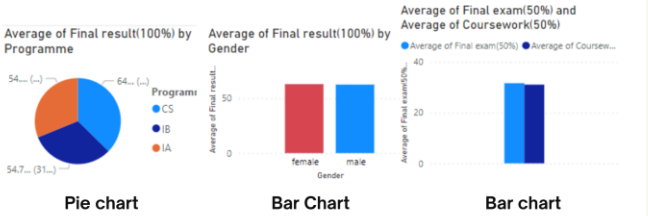
RESULTS

- Students from CS background generally performed better than students from other programmes followed by IA and IB
- Students performed better in psychomotor domain than in cognitive domain
- Females performed better in cognitive domain whereas males performed better in psychomotor
- Majority of the students are reasonably good in knowledge-based task



ANALYSIS

- Based on the pie chart according to the average of final result by programme, it was observed that CS students scored the highest average which was 64.97 marks followed by IB (54.79) and IA (54.25)
- According to the data from the bar chart based on the average of final result by gender, it shows that the average score by female, 63.03 was slightly higher than average score by male, 62.61.
- The bar chart according to average of final exam and coursework shows that the average of final exam slightly higher than average of coursework with scoring 31.67 for finals while 31.09 for coursework



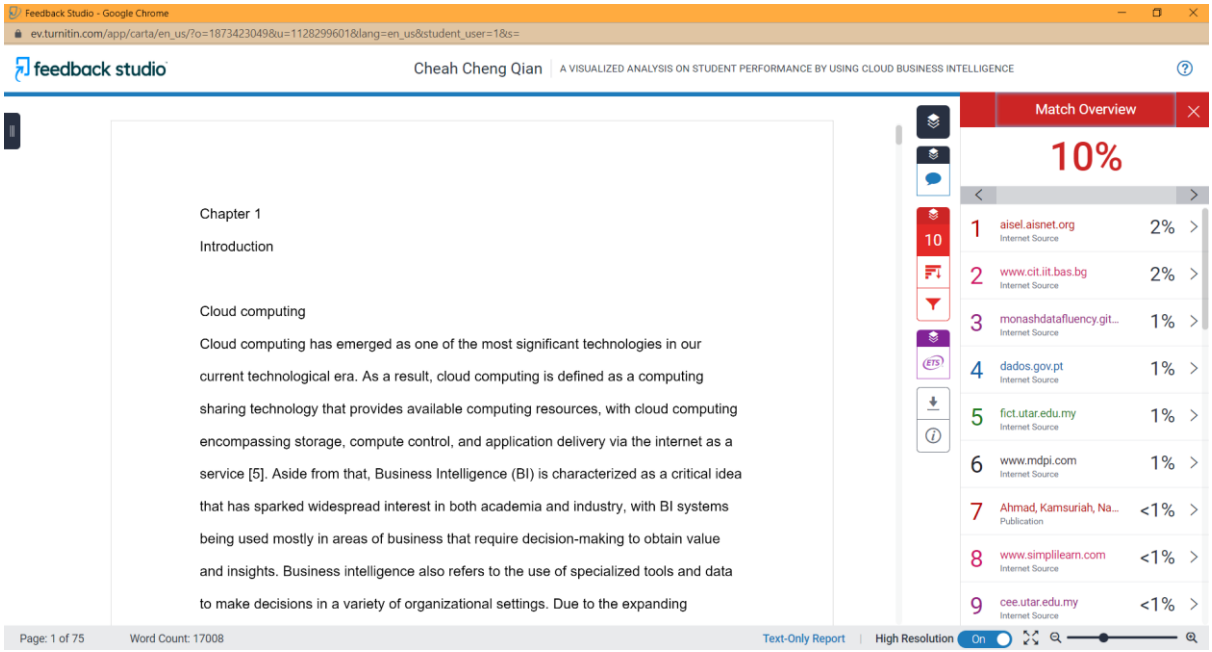
RECOMMENDATIONS

- Higher educational institution should implement another foundation programme, leading to Information and Communication Technology
- Lecturers or tutors pay more attention to the students who lacks capabilities in learning and performed below satisfactory level
- A development of programming courseware
- Establishing a coding club in the faculty of information and technology
- Implement live coding teaching instead of using slides to explain the concepts and principles

CONCLUSION

As a conclusion, the findings based on performed data analysis on the results of students have been identified as well as the area of weaknesses of students. It is proven that business intelligence plays a very significant role in education especially into improving students' performance. By undergoing data analysis using BI tools, higher educational institution able to use the collected data gathered and visualized them into a graphical representation to extract valuable insights into solving the problem. And by having this information extracted, it enables universities to perform better at important decision makings. The study findings given in this research show that using Business Intelligence tools to analyse existing data and derive valuable information for decision making has a lot to promise. Students are also expected to improve their ways of learning by having proper time management in their daily schedules and more to hands on practice at home whenever they have some free time .

PLAGIARISM CHECK RESULT



A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

ORIGINALITY REPORT

10%	9%	4%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	aisel.aisnet.org Internet Source	2%
2	www.cit.iit.bas.bg Internet Source	2%
3	monashdatafluency.github.io Internet Source	1%
4	dados.gov.pt Internet Source	1%
5	fict.utar.edu.my Internet Source	1%
6	www.mdpi.com Internet Source	1%
7	Ahmad, Kamsuriah, Nazlena Mohamad Ali, and Suhaila Zainudin. "An Improved Course Assessment Measurement for Analyzing Learning Outcomes Performance Using Rasch Model", Procedia - Social and Behavioral Sciences, 2011. Publication	<1%

Universiti Tunku Abdul Rahman			
Form Title : Supervisor's Comments on Originality Report Generated by Turnitin for Submission of Final Year Project Report (for Undergraduate Programmes)			
Form Number: FM-IAD-005	Rev No.: 0	Effective Date: 01/10/2013	Page No.: 1of 1

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY

Full Name(s) of Candidate(s)	CHEAH CHENG QIAN
ID Number(s)	17ACB04662
Programme / Course	BACHELOR OF INFORMATION SYSTEMS (HONOURS) BUSINESS INFORMATION SYSTEMS
Title of Final Year Project	A VISUALIZED ANALYSIS ON STUDENT PERFORMANCE BY USING CLOUD BUSINESS INTELLIGENCE

Similarity	**Supervisor's Comments (Compulsory if parameters of originality exceeds the limits approved by UTAR)**
Overall similarity index: 10 % **Similarity by source** Internet Sources: 9 % Publications: 4 % Student Papers: 0 %	
Number of individual sources listed of more than 3% similarity: 0%	
Parameters of originality required and limits approved by UTAR are as Follows: **(i) Overall similarity index is 20% and below, and** **(ii) Matching of individual sources listed must be less than 3% each, and** **(iii) Matching texts in continuous block must not exceed 8 words** ***Note: Parameters (i) – (ii) shall exclude quotes, bibliography and text matches which are less than 8 words.***	

Note Supervisor/Candidate(s) is/are required to provide softcopy of full set of the originality report to Faculty/Institute

Based on the above results, I hereby declare that I am satisfied with the originality of the Final Year Project Report submitted by my student(s) as named above.

Signature of Supervisor

Name: Ts Dr Chan Lee Kwun

Date: 02 September 2022

Signature of Co-Supervisor

Name: _____

Date: _____

UNIVERSITI TUNKU ABDUL RAHMAN

FACULTY OF INFORMATION & COMMUNICATION TECHNOLOGY
(KAMPAR CAMPUS)

CHECKLIST FOR FYP2 THESIS SUBMISSION

Student Id	17ACB04662
Student Name	CHEAH CHENG QIAN
Supervisor Name	TS DR CHAN LEE KWUN

TICK (√)	DOCUMENT ITEMS Your report must include all the items below. Put a tick on the left column after you have checked your report with respect to the corresponding item.
	Front Plastic Cover (for hardcopy)
√	Title Page
√	Signed Report Status Declaration Form
√	Signed FYP Thesis Submission Form
√	Signed form of the Declaration of Originality
√	Acknowledgement
√	Abstract
√	Table of Contents
√	List of Figures (if applicable)
√	List of Tables (if applicable)
	List of Symbols (if applicable)
√	List of Abbreviations (if applicable)
√	Chapters / Content
√	Bibliography (or References)
√	All references in bibliography are cited in the thesis, especially in the chapter of literature review
√	Appendices (if applicable)
√	Weekly Log
√	Poster
√	Signed Turnitin Report (Plagiarism Check Result - Form Number: FM-IAD-005)
√	I agree 5 marks will be deducted due to incorrect format, declare wrongly the ticked of these items, and/or any dispute happening for these items in this report.

*Include this form (checklist) in the thesis (Bind together as the last page)

I, the author, have checked and confirmed all the items listed in the table are included in my report.

_____
(Signature of Student)
Date: 02 September 2022