Cambodia Academy of Digital Technology
Institute of Digital Research and Innovation
Institute of Digital Governance

Project Report

Digital Skill Essential (DSE) Analysis

Prepared by
Thear Sophal
Monirath Sokchovy

Given to
Institute of Digital Governance

Date
October 2023

Table of Contents

Abstract

Digital Skill Essential or known as DSE, is a set of subjects that provide the digital skills and knowledge necessary in the work and daily life of people, benefit from personal and organizational work, as well as be ready to step in the direction of digital evolution, increase efficiency and efficiency to keep up with the digital economy and the fourth generation of industry, especially to contribute to the implementation of digital government policies and economic policy framework, and digital society. In this report, we have a study to answer questions on satisfied passing score for examinees, the insight of question level of difficulties, and how course preparation affect examination.

1. Introduction

Digital Skill Essential or known as **DSE**, is a set of subjects that provide the digital skills and knowledge necessary in the work and daily life of people, benefit from personal and organizational work, as well as be ready to step in the direction of digital evolution, increase efficiency and efficiency to keep up with the digital economy and the fourth generation of industry, especially to contribute to the implementation of digital government policies and economic policy framework, and digital society.

The goal of this study is to answer these questions and improve the DSE implementation process so that it is fair and challenging for all trainees and examinees, as well as organizer. Here is a more detailed each raised question:

- Passing score: The passing score for the DSE is currently 60%. However, it is not clear if this is a satisfactory passing score. Some people believe that the passing score should be higher, while others believe that it should be lower. This study will help to determine the optimal passing score for the DSE.

- Preparation course: The DSE preparation course is a 88-hours course that covers 13 subjects. It is not clear if this course is sufficient for examinees to pass the exam. Some people believe that the course should be longer, while others believe that it should be shorter. This study will help to determine the optimal length of the DSE preparation course.

- Level of difficulty of question for exam: The DSE exam is divided into three levels of difficulty: easy (L1), medium (L2), and hard (L3). It is not clear at which level of difficulty examiners score the highest. Some people believe that examinees score the highest at the easy level, while others believe that they score the highest at the hard level. This study will help to determine the level of difficulty at which examinees score the highest.

2. Scope

So the scope of this project is to analyse and give solution to the problem that we have related to DSE. We want to achieve our goals and objectives throughout the analysis by using all data that we obtain within a timeline.

3. Methodology

3.1 Dataset

Currently, we have 5 types of data that we get from the Institute of Digital Governance. There are Test Results, Pre-test Results, Question design in XML format, Logging Progress, and Survey. Out of 5 types of datasets that we mentioned earlier, there are only three datasets that seem to be useful for analysis.

3.2 Tools and Technologies

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft with the Electron Framework, for Windows, Linux and macOS. Features include

support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git.

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured, object oriented and functional programming.

Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license.

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK. Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

4. Analysis

4.1 Passing Score Analysis

4.1.1 Examination Performance

From Test Result, there are 8 different set of data which come from:

- ព្រះវិហារ ឧត្តរមានជ័យ សៀមរាប
- ភូមិន្ទរដ្ឋបាល
- កែប កំពត ព្រះសីហនុ កោះកុង
- ព្រឹទ្ធសភា សាលាក្រុង ក្រសួងអប់រំ
- បន្ទាយមានជ័យ បាត់ដំបង ពោធិ៍សាត់ កំពង់ឆ្នាំង ប៉ៃលិន
- ក្រសួងប្រៃសណីយ៍ និងទូរគមនាគមន៍
- បណ្ឌិត្យសភាបច្ចេកវិទ្យាឌីជីថលកម្ពុជា
- ៦ខេត្តផ្សេង

Out of the 8 province, city, and other institutes, we can see that there are most examinees from ភូមិន្ទរដ្ឋបាល with a total number of 672 and the least examinees of the exam are from បណ្ឌិត្យសភា បច្ចេកវិទ្យាឌីជីថលកម្ពុជា with a number of 169. See Figure 1
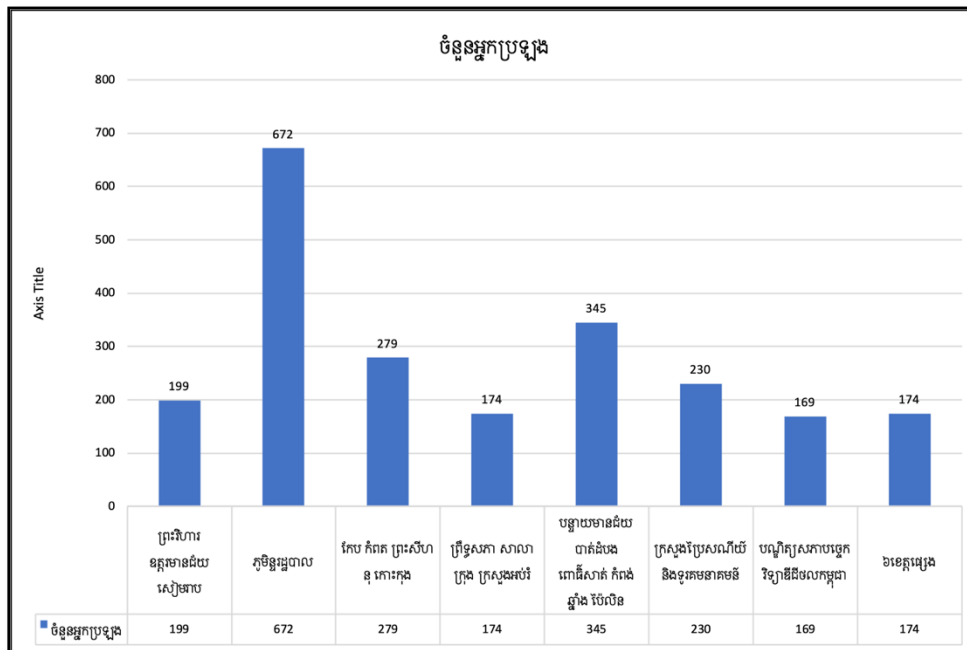
Figure 1: Number of Examinees from Province, City and Other Institutes

If we look this below pie charts, we can see that number of examinees from ភូមិន្ទរដ្ឋបាល is 30% of total examinees. See Figure 2
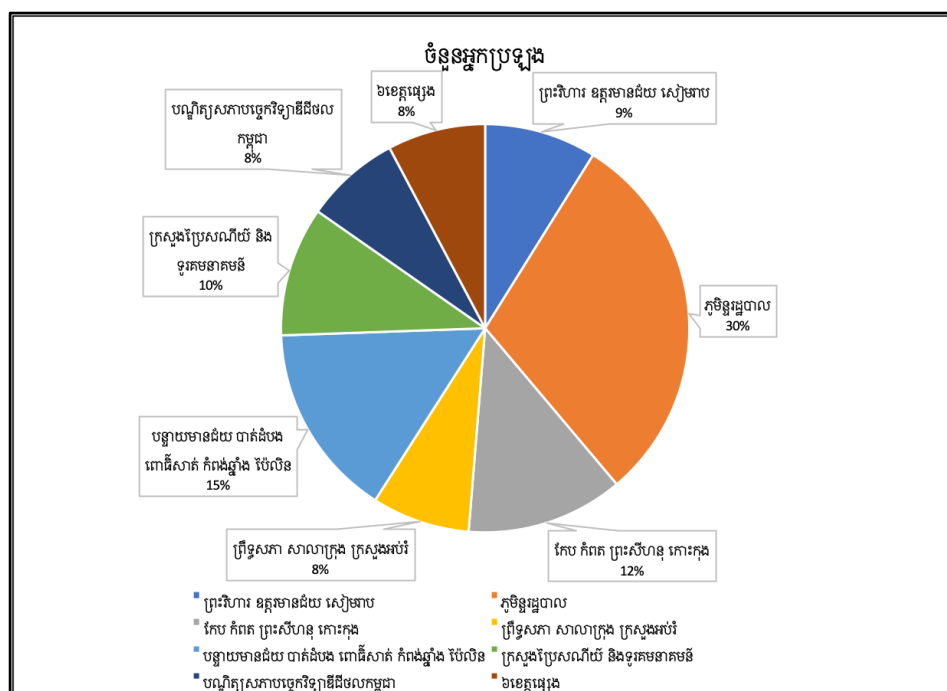

Figure 2: Proportion of Examinees

Here is the statistic information from the combined data of all province, city and other institution, we have total examinees of 2242 and completed examinees of 2240. See Figure 3
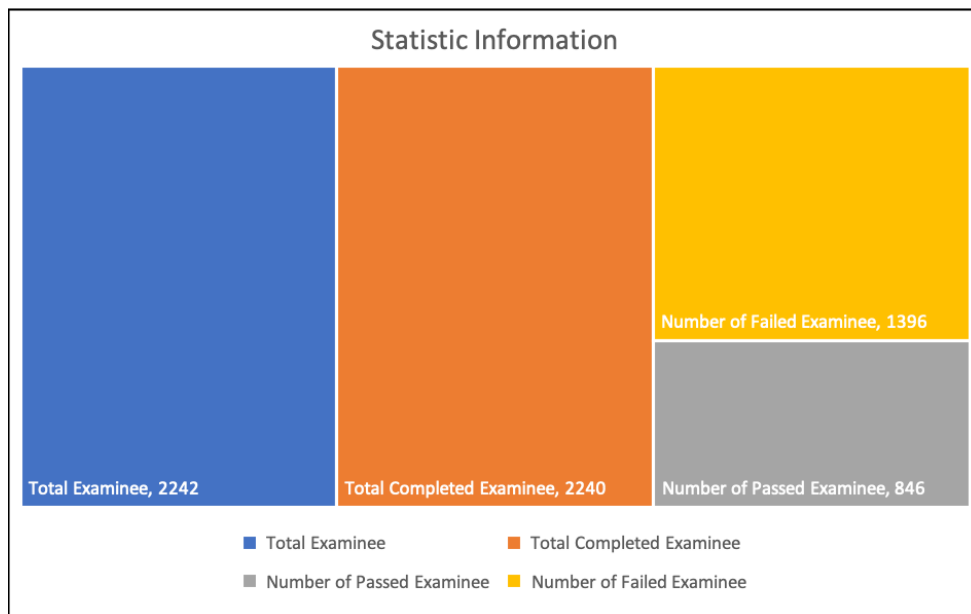
Figure 3: Statistical Information

Amongst the total number of 2240, there are only 846 candidates who passed and the amount who failed is almost twice more than passed candidate, 1396. See Figure 4
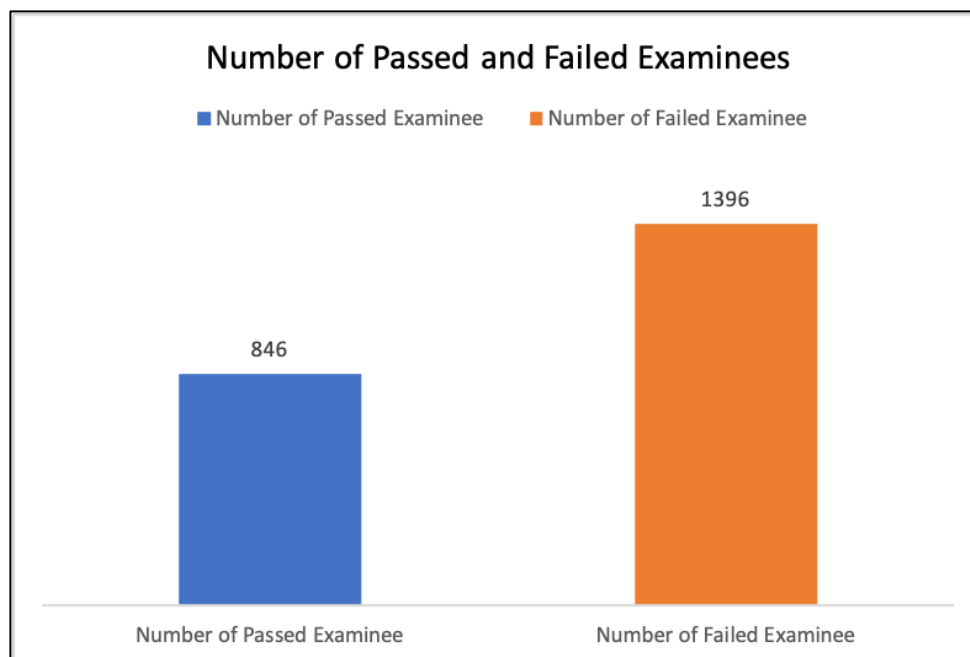

Figure 4: Number of Passed and Failed Examinees

In this study, the passing score was set to 60 and as a result, the passing rate is almost twice as low as the failed rate. The passing rate of the examination is only 38% while the failed rate is 62%. Figure 5
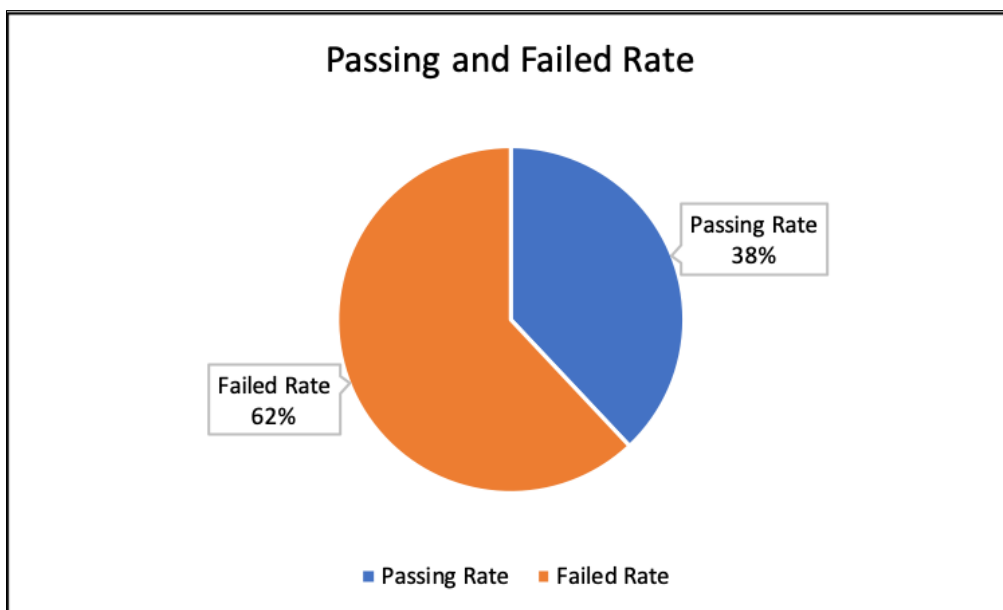
Figure 5: Passing and Failed Rate

Here is also statistical information from examination result data that we combined. The maximum score across all examinees is 94.18 while the minimum score is only 3. What is more, the average score is 55.45 which is around 5 points below the passing score. See Figure 6

Description	Data
Maximum Score	94.18
Minimum Score	3
Average Score	55.45
Standard Deviation	14.65
Median	55.14

Figure 6: Aggregation and Statistical Information

Below bar chart below shows the average passing and failed scores from all 8 provinces, cities, and other institutes. If we look at failed scores, they are more or less the same, which range between 45 to 47 and we look at passing average score, we can see that examinees from ក្រសួងប្រៃសណីយ៍ និងទូរគមនាគមន៍ and បណ្ឌិត្យសភាបច្ចេកវិទ្យាឌីជីថលកម្ពុជា is the highest, 73. See Figure 7
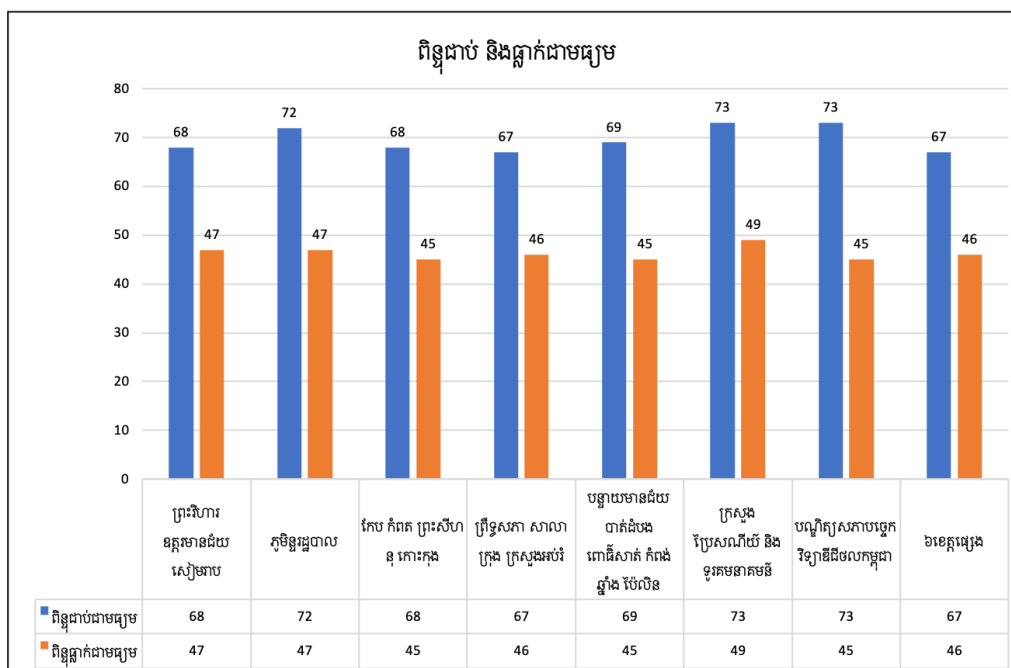
Figure 7: Passing and Failed Score of Each Province, City and Other Institutions

Below stacked bar chart below shows the passing and failed rates from each province, city, and other institutions. Looking at result from ក្រសួងប្រៃសណីយ៍ និងទូរគមនាគមន៍ and បណ្ឌិត្យសភាបច្ចេកវិទ្យាឌីជីថលកម្ពុជា, their passing rate is above 50%, which means that more examinees passed the DSE test. In contrast, the passing rate in others is below 50%. See Figure 8
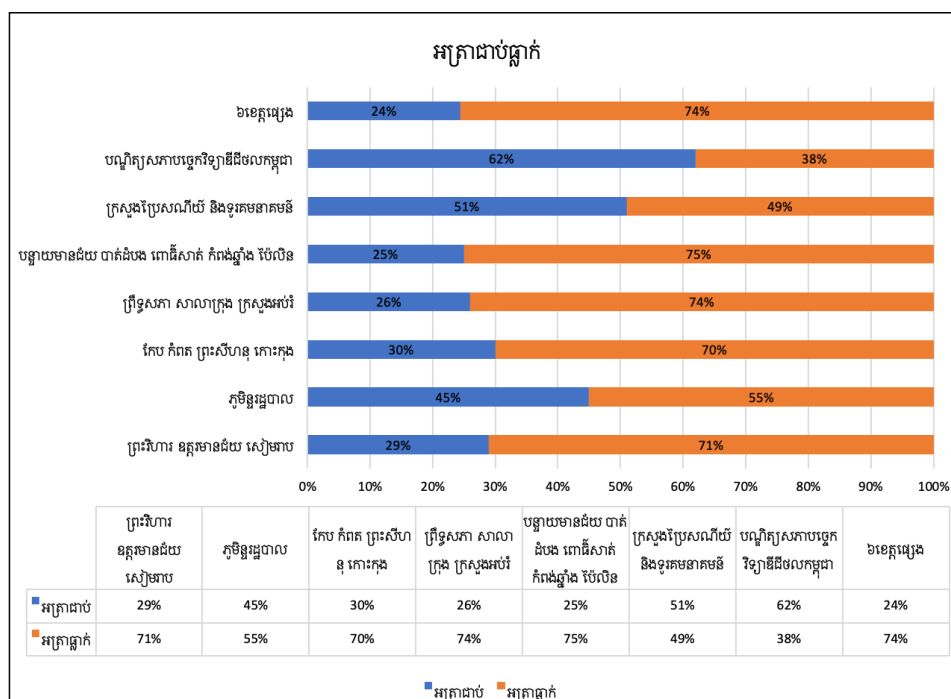

Figure 8: Passing and Failed Rate of Each Province, City and Other Institutions

We have been analyzing data and information when the passing is set to 60, and this pie chart shows the passing rate and failed in case the score is set to 50. As a result of setting the passing score to 50, we can see that we increase passing from 38% to 64%. See Figure 9
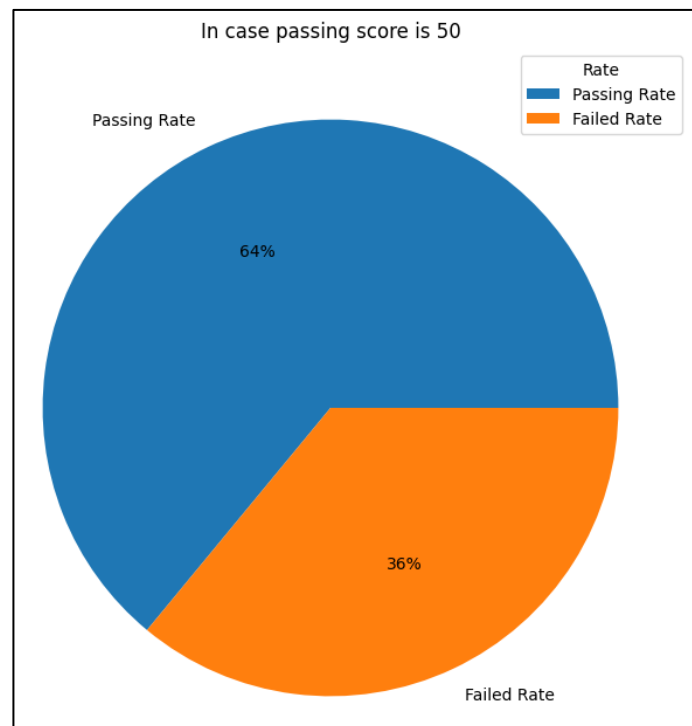
Figure 9: Passing and Failed Rate If Score is Set to 50

4.1.2 Student's Behaviour on Examination

In this bar chart, it shows the average score that examinees score during the exam. We scaled from 1-100 to 0-1. We can see that the average score in the 10th Subject, "កម្មវិធីដែលមានលក្ខណៈភាពលើទូរសព្ទ ឆ្លាតវៃ **Productivity Apps on Smartphone**" is the highest, 0.64. It means that examinees perform well in this subject.

Note:

១ កុំព្យូទ័រចាំបាច់ Computer Essentials

២ អនឡាញចាំបាច់ Online Essentials

៣ អក្ខរកម្មព័ត៌មាន Information Literacy

៤ ឯកជនភាពទិន្នន័យ និងការការពារ Data Privacy and Protection

៥ បង្កើតឯកសាររដ្ឋបាល Word Processing

៦ កម្មវិធីបង្កើតបញ្ជី និងវិភាគទិន្នន័យ Spreadsheets

៧ បង្កើតមាតិកាឌីជីថល Digital Content Creation

៨ អក្ខរកម្មទិន្នន័យ Data Literacy

៩ កម្មវិធីសហការតាមអនឡាញ Online Collaboration Tools

១០ កម្មវិធីដែលមានលក្ខណៈភាពលើទូរសព្ទឆ្លាតវៃ Productivity Apps on Smartphone

១១ បរិវត្តកម្ម និងសេដ្ឋកិច្ចឌីជីថល Digital Economy and Transformation

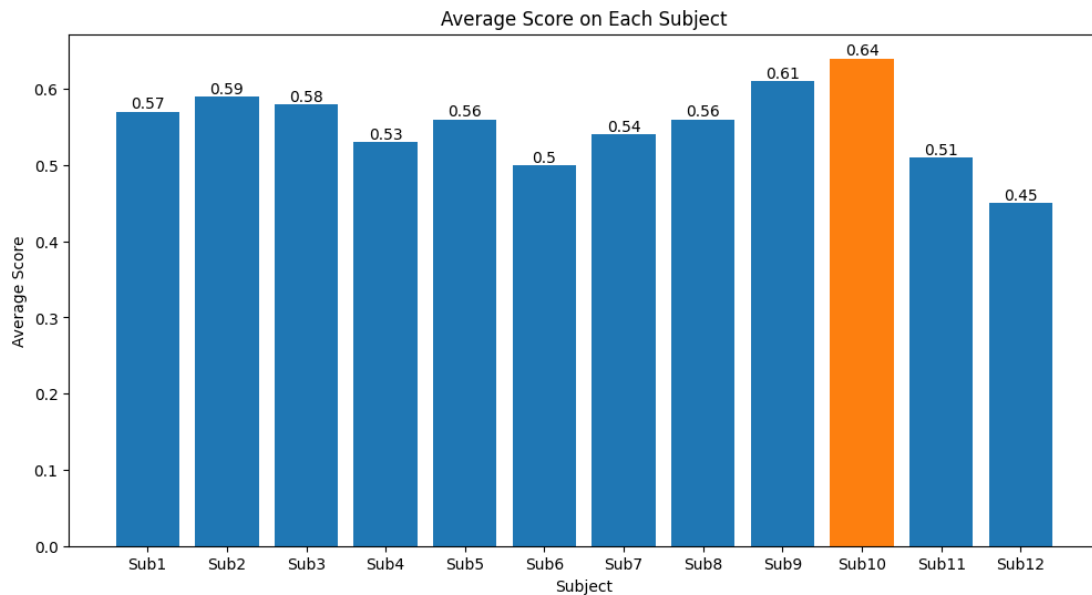១២ អភិបាលកិច្ចឌីជីថល Digital Governance

Figure 10: Average Score on Each Subject

For the students who pass the exam, we can find the average time they spent on their exam by sorting it out by labeling it as Passed and Failed.

After analyzing based on the time they spent on the test, we can see that the Figure 11 represents the average passing and failing time of all examinees. For examinees who passed the exam, the average time to complete the exam was 1 hour 50 minutes converted into minutes, which is 110 minutes, while examinees who failed the exam took 1 hour 40 minutes converted into minutes 109 minutes.
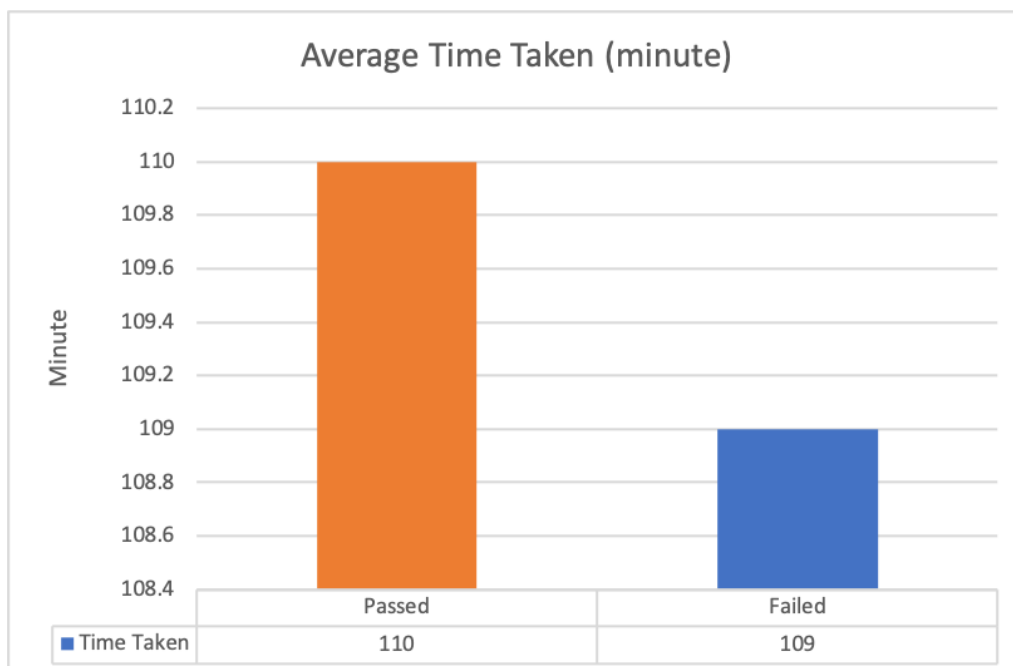

Figure 11: Average Time Taken of Exam (minute)

In addition, the maximum time for examinees who pass the exam is 2 hours 22 minutes converted into minutes, which is 142 minutes, and for examinees who fail the exam, it is 2 hours 2 minutes, converted into minutes, which is 122 minutes, as shown on Figure 12.
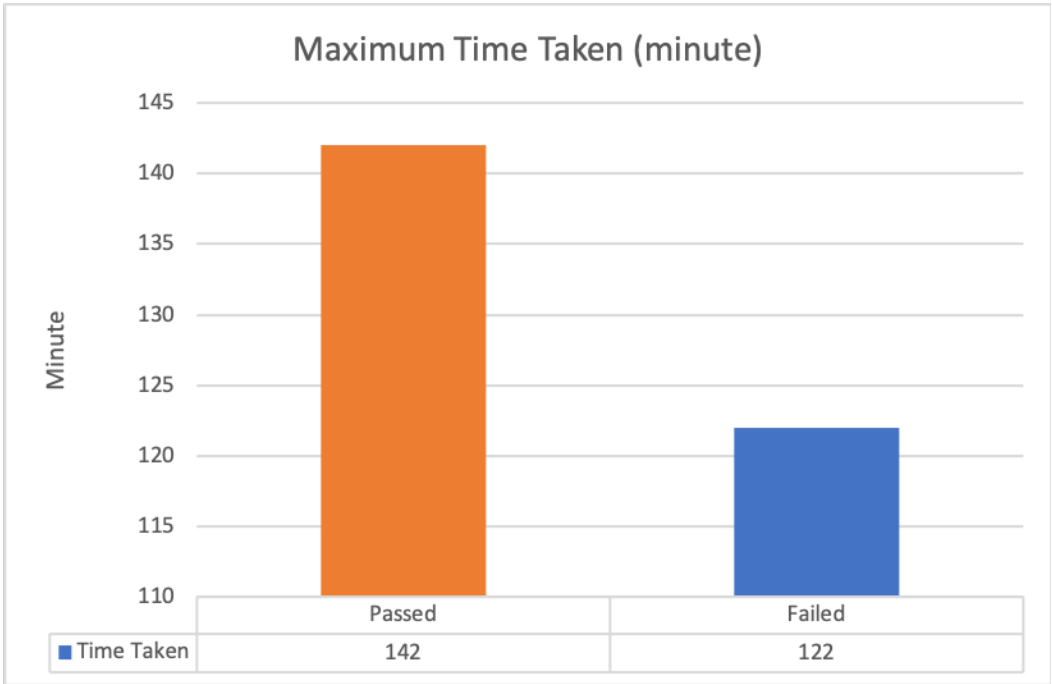

Figure 12: Maximum Time Taken of Exam (minute)

For the Figure 13, the minimum time for those who pass the exam is 1 hour 16 minutes converted into minutes is 76 minutes, and for examinees who fail the examinee spend their time in exam is 1 hour 15 minutes converted into minutes is 75 minutes.
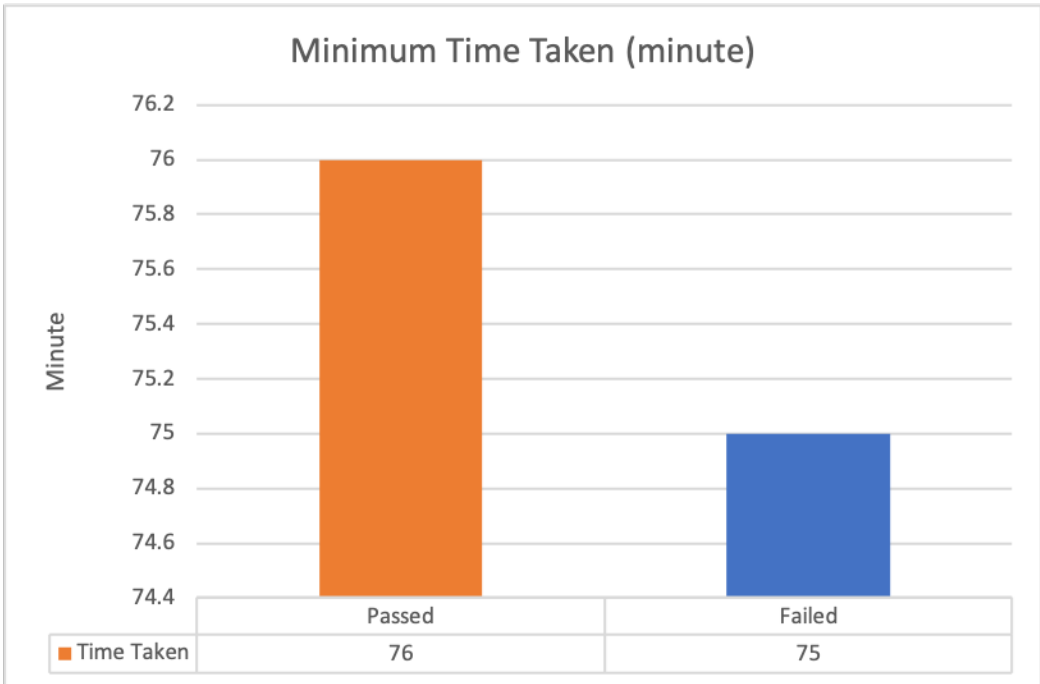

*Figure 13: Minimum Time Taken of Exam (minute)*

4.2 Question Difficulty Level Analysis

4.2.1 Difficult Level from Question Pool Sets and Question Statistical Result

To analyze the difficulty level of the question, we first need to take a look at the question level status in each of the subjects to be randomly chosen to appear in the exam.

In the pie charts below, we show the amount of question difficulty levels in each of the subjects and their proportion in each question pool. Actually, there are 3 levels of difficulty. They are L1, L2, and L3. L1 is an "**easy**" level. L2 is the "**medium**" level, and L3 is the "**hard**" level. If we look at those pie charts, we can see that most of the questions are marked with either easy or medium levels, whereas the hard level is the least among all questions in each pool. We also see that there are some questions in some question pools that are not set level yet. They are still in the status of either "**pending**" or "**reviewing**".
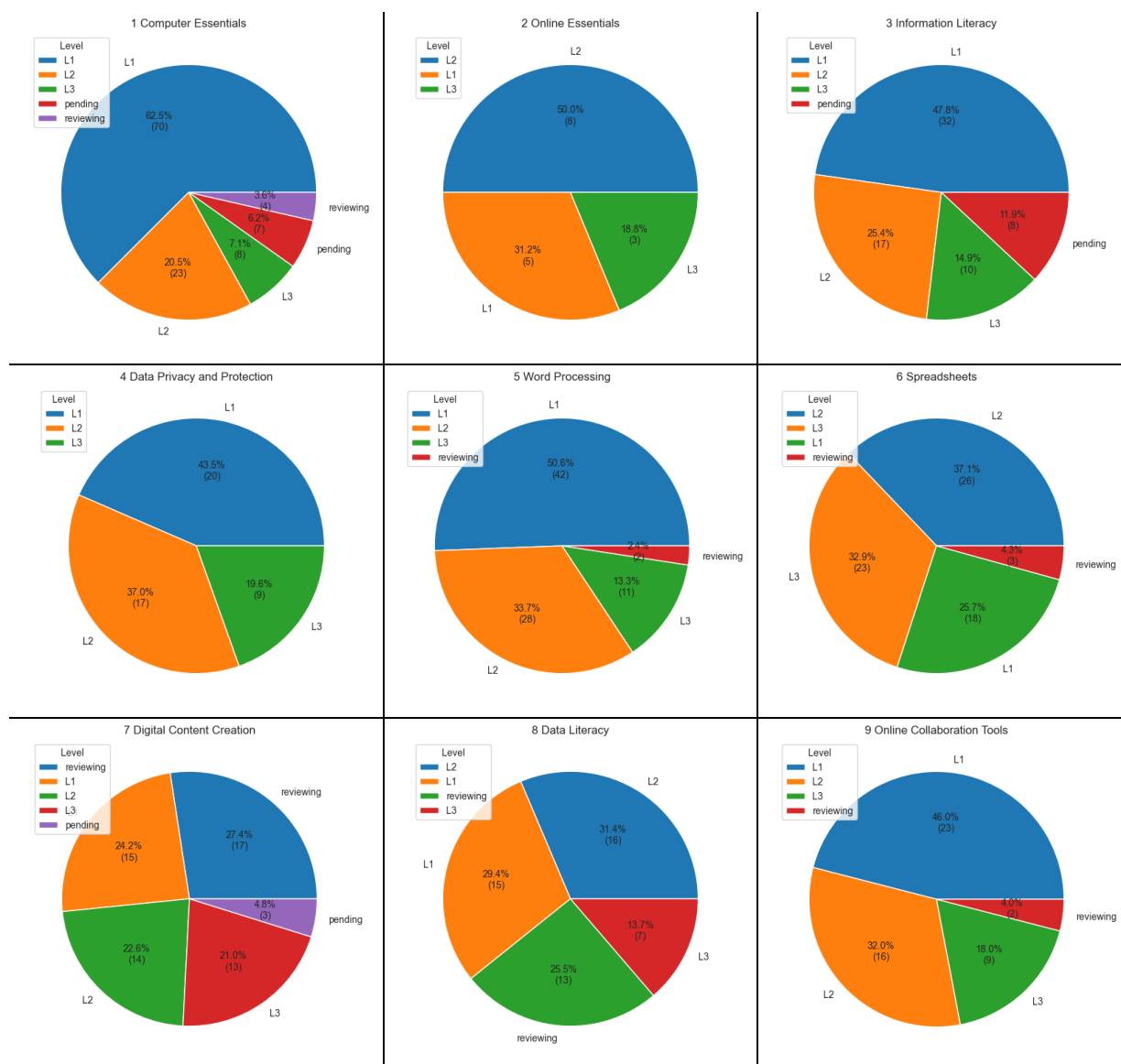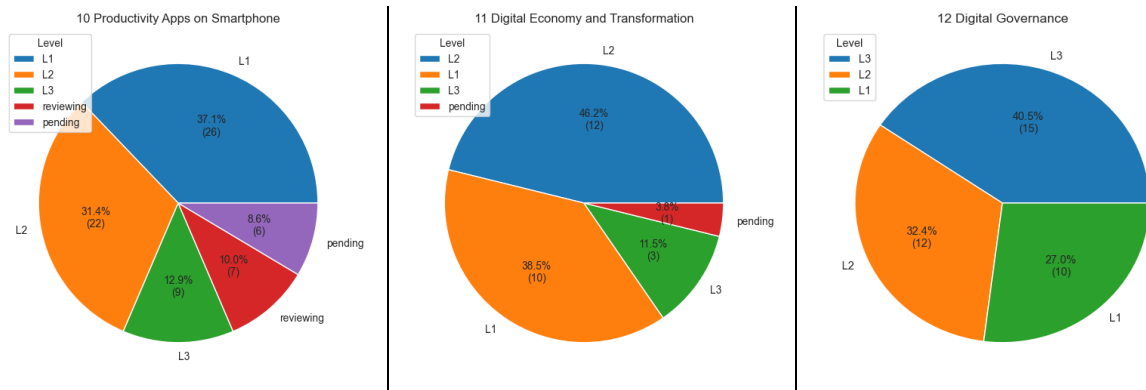
Here is the total amount of questions at all level except questions with status "pending" and "reviewing". There are 286 L1 questions, 211 L2 questions, and 120 L3 questions.
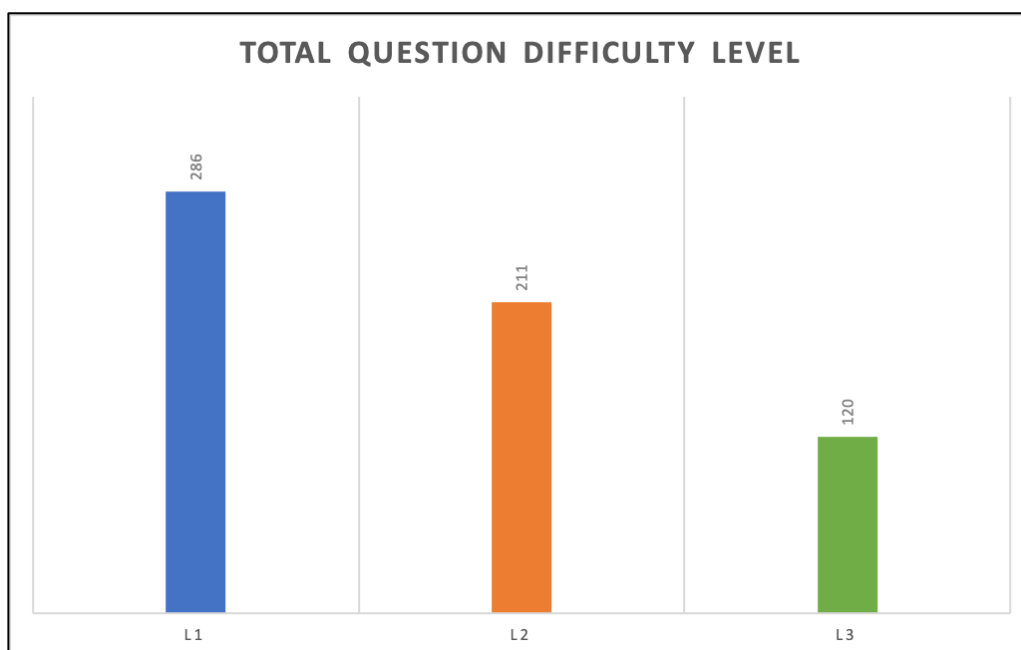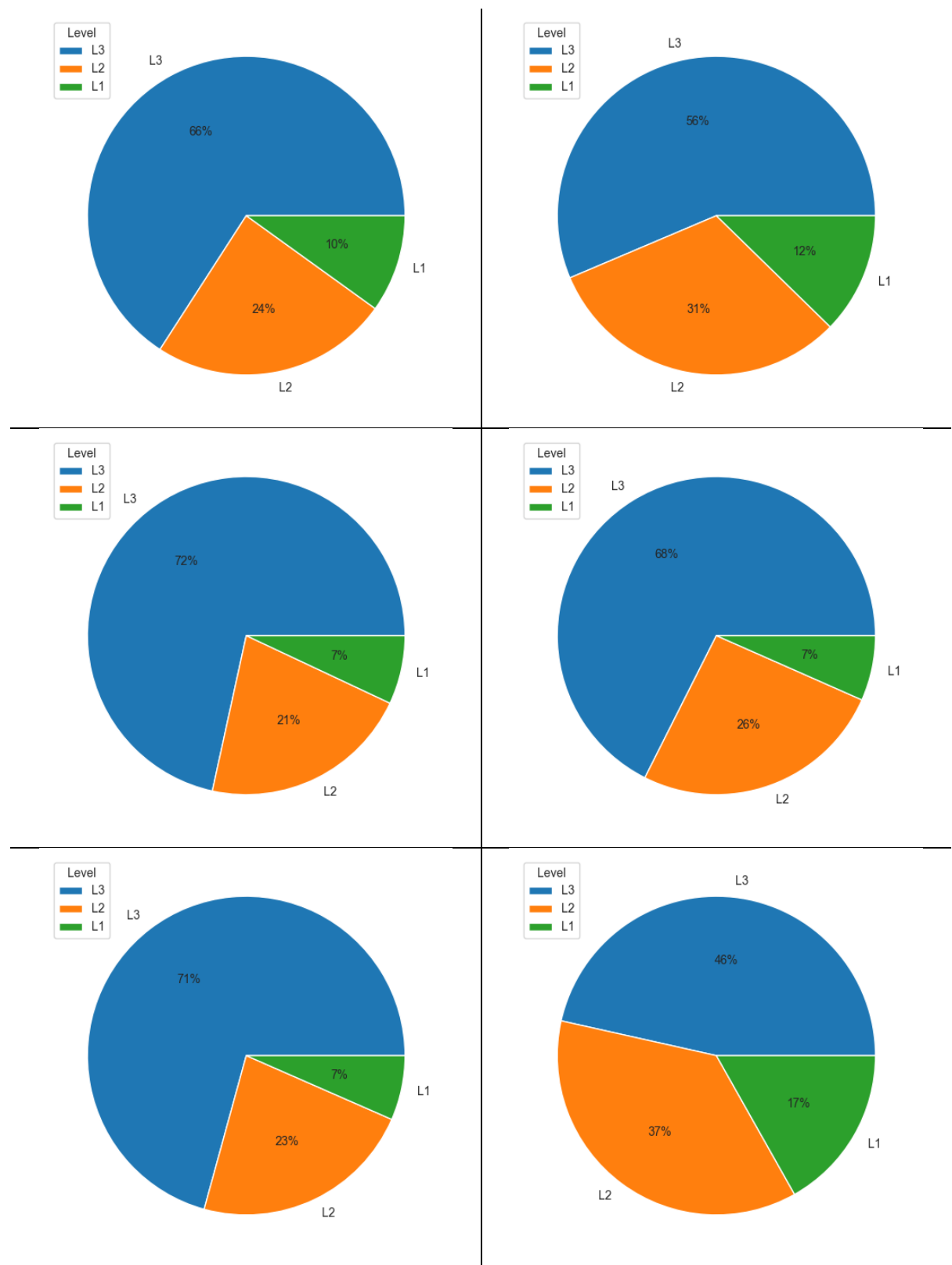

Figure 14: Total Question Difficulty Level

After we studied the difficulty level of question pools from each subject, we studied question statistical information from each of our datasets, each province, city, and institution. In this section, we study the difficulty level based on the passing score of 60 that was set. Because passing is 60, we decided to set the L3 level range from 1 to 59 score. L2 level if the score ranges from 60 to 79, and the L3 level base score range from 80 to 100.

With this study, we also use pie charts to analyze (See Figure 15). Based on the level that we set according to the score, we can see that most of the questions are in the L3 level (Hard). Except question level from ក្រសួងប្រៃសណីយ៍ និងទូរគមនាគមន៍ and បណ្ឌិត្យសភាបច្ចេកវិទ្យាឌីជីថលកម្ពុជា because these examinees have background in Information Communication Technology (ICT). What is more, we also see a big proportion of L2 level (Medium) in the charts and less proportion of L1 (Easy). If we look at Figure 16, we can see that 60% of the questions are Hard, then 29% are Medium level and only 11% are Easy for them.

It can be inferred that most of the questions that randomly appeared in the exam are difficult for examinees.
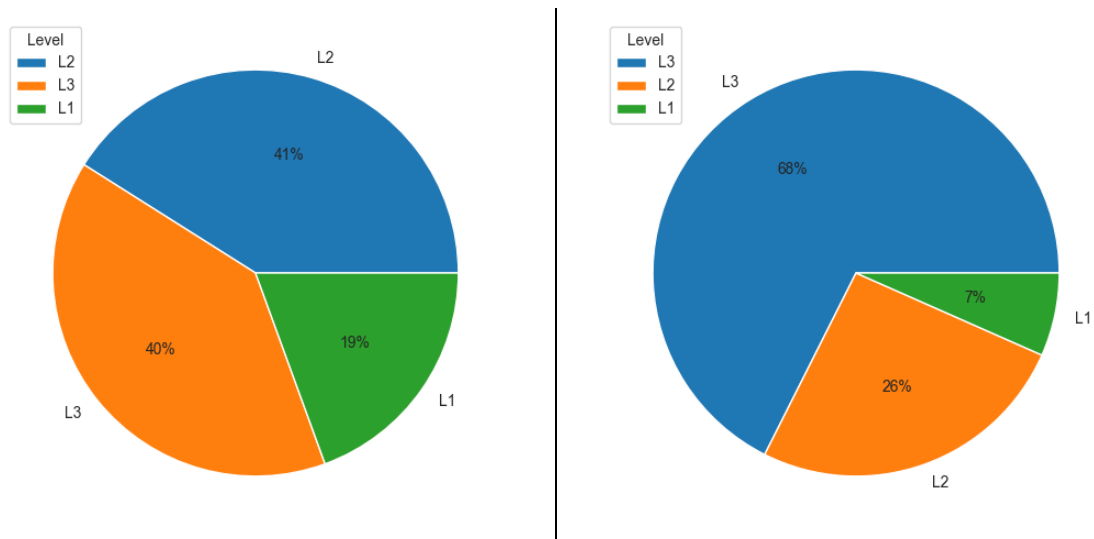
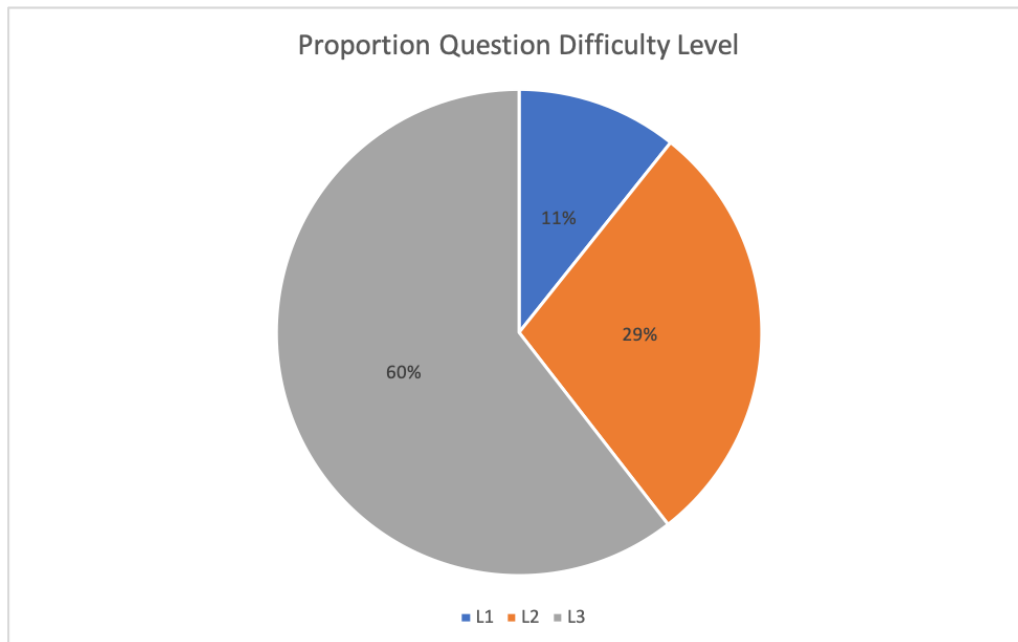*Figure 15: Proportion of Question Level of each Province, City and Institution*

Figure 16: Proportion of Question Difficulty Level from QuestionStats

4.3 Course Preparation Analysis

4.3.1 Student Engagement Before Exam

For the study of the preparation course, we analyze the data from CADT and B018 as the sample represents the whole dataset.

Looking at Figure 17, on the left pie chart, the average completion rate of each examinee is only 16% while the lessons that are not completed are 84%. Moreover, on the right pie chart shows similar information that the average completed rate of the lessons is 2% less than CADT. So, with online course preparation, examinees spend less time with the provided materials and lessons and this could be one factor that affects their performance in real tests.
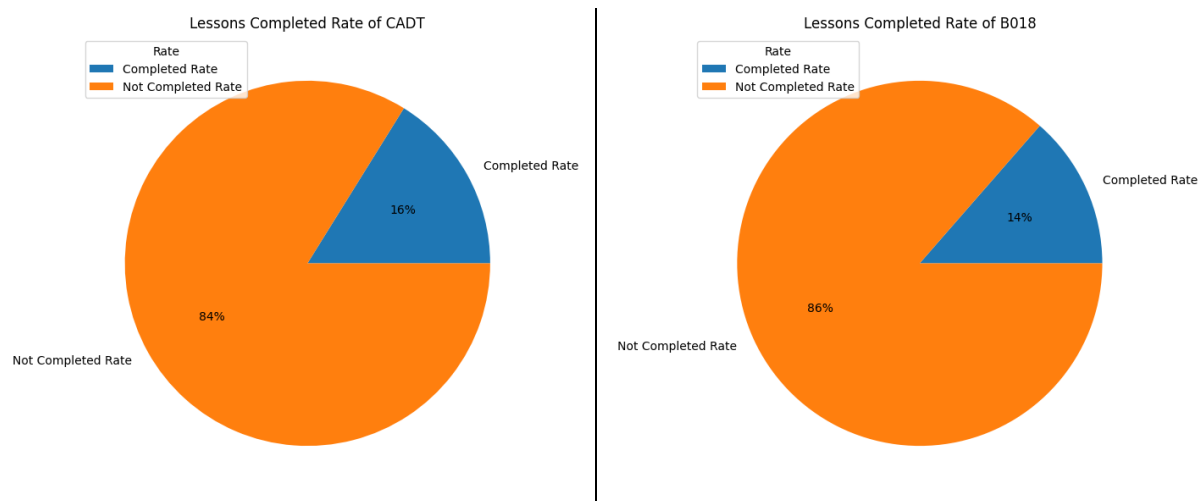
Figure 17: Completed Rate of Online Lessons

In this line graph (Figure 18), show the trend that examinees spend some time completing the online slide and video lessons, as we can see in the chart, it shows that most of the examinees spend most of the time on the first 50 lessons then less time on other lessons.

We can explain that there seem to be many lessons to complete and examinees tend not to do so. The number of slides and videos should be reduced to attract learner to spend time on the provided lessons.
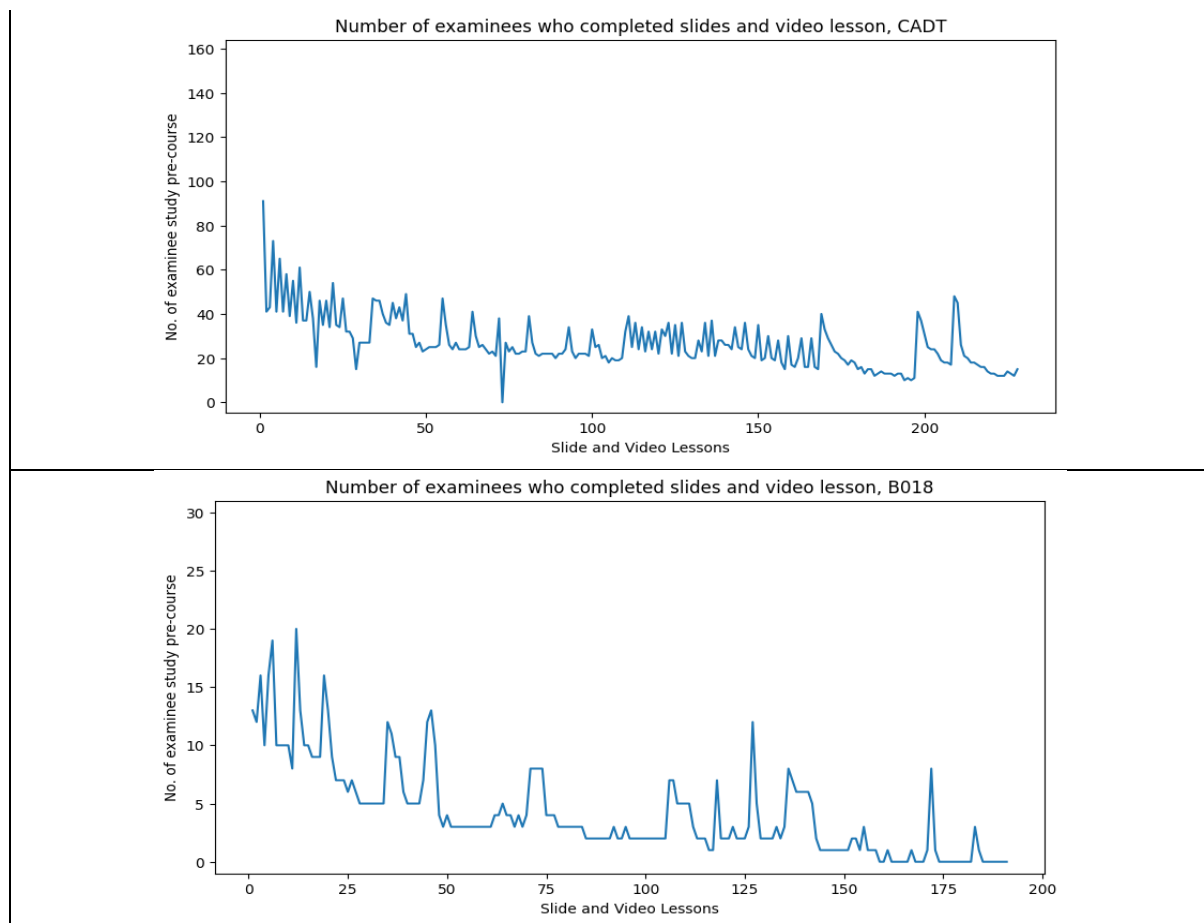

Figure 18: Number of Examinee who completed slide and video lessons

4.3.2 Pre-test

Looking at the result from the pre-test, it shows that amongst 3064 examinees, there are only 429 passed the test and 2635 failed. Statistically, the number of passed examinees in the pre-test is twice less than the number of passed examinees in the real test. With the amount of passing and failing of examinees in the pre-test, the passing rate is only 14% and the failing rate is around 6 times more than the passing rate which is 86%. See Figure 19

There is a correlation between the result from the pre-test and the real test that the less the passing rate in the pre-test it makes less passing rate in the real test.
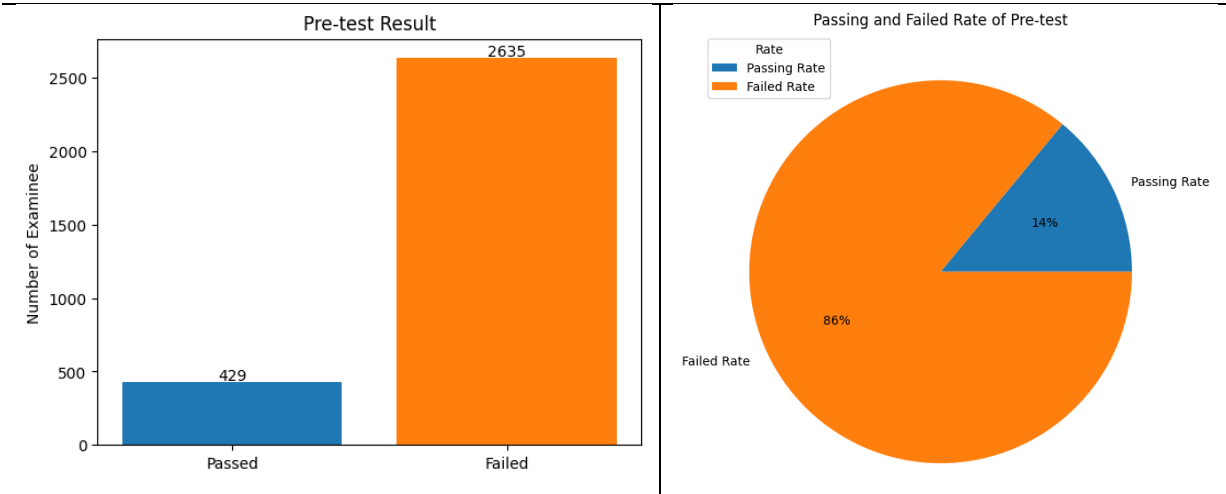

Figure 19: Pre-test Result and Passing, Failed Rate

Figure 20 shows the distribution score of the pre-test result with a vertical red line that indicates the passing score. We can see that it is a normal distribution like the distribution score of the real examination but the difference is that this distribution score is getting close to the score range between 20 to 60 while the real test's is getting close to the score range between 40 to 80.
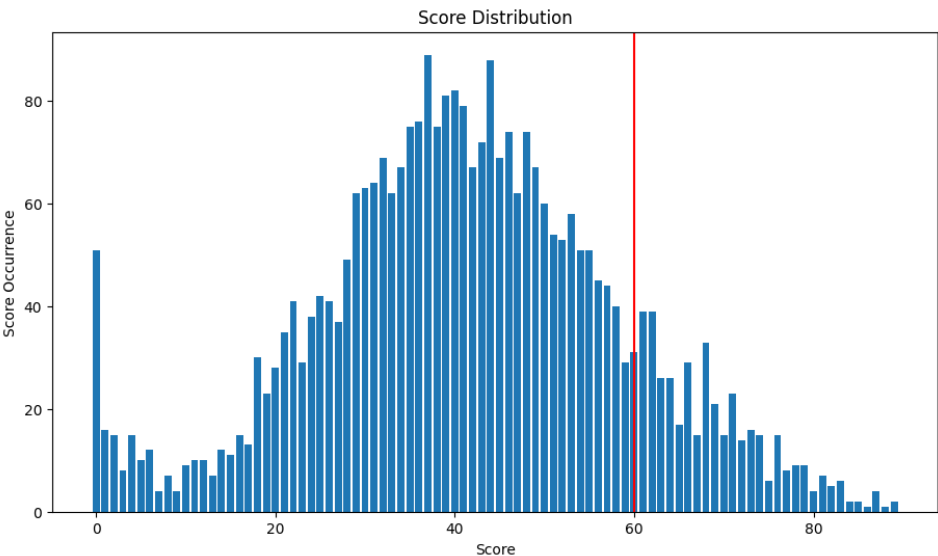

Figure 20: Pre-test Distribution Score

5. Result and Recommendation

As a result of the analysis study, we found some insights that can be useful for making consideration on questions that were raised.

First of all, we found that the passing score of 60 is a bit high for examinees from different backgrounds especially those who don't have an ICT background. Unlike ក្រសួងប្រៃសណីយ៍ និង ទូរគមនាគមន៍ and បណ្ឌិត្យសភាបច្ចេកវិទ្យាឌីជីថលកម្ពុជា, that they have passing rate above 50%, other examinees have passing rate in average less than 30%. In contrast, with a passing score of 50, we increase the amount of examinees who can pass the exam from 38% to 64%. In comparison, because of the different gap between those who have a background in ICT and those who don't, it can be inferred that the passing score of 50 should be re-considered since it can be in between the gap.

Secondly, with the analysis of course preparation, we can see that most of the examinees did not complete lessons or watch videos that were prepared by the committee and it made a poor performance in the real test. Furthermore, with pre-test data analysis, the passing rate is much less than the failed rate, with only 14%. It means that course preparation is not effective enough for examinees to take pre-tests and real tests. In our opinion, examinees seem to pay less attention to pre-tests and it makes the real test in bad result neither. For the suggestion of this section, there should be a possibility that we take the pre-test as an add-on score for the real test in order to encourage examinees to take the exam more seriously.

Finally, on the last point of analysis on question difficulty level, we gained big-picture insight that there are more L1 (easy) levels in the question pool but in contrast when we get the result from the exam. We could see that most of the questions are in L3 (hard) level for examinees and that level L3 is responsible for 60% of all questions that appeared in the real test. It can be inferred questions that were set to an easy level are not quite easy for examinees and the question difficulty level should be changed.

6. Conclusion

In conclusion, we managed to analyze and answer the question raised related to the Digital Skill Essential examination. There are interesting insights that we studied and made suggestions and there is some interesting information that we need to look at and make some changes accordingly.

The passing score of 60 should be considered to decrease to 50 to make no big gap between those who have an ICT background and those who do not. For course preparation analysis, there should be a reduction of time of self-study because it's not effective and students tend not to spend time on slides and video provided. Lastly, for question difficulty level, there should be a change or a restructure so that the question meets the performance of examinees who take the examination.