INFO1111: Computing 1A Professionalism

2022 Semester 1

Practice: Team Project Report

Submission number: ??

Team Members:

|  |  |  |
| --- | --- | --- |
| Name | Student ID | Levels being attempted in this submission |
| Xiyan Huang | 510221401 | 2 |
| ?? | ?? | ?? |
| ?? | ?? | ?? |
| ?? | ?? | ?? |

Contents

[1. Level 1: Basic Skills 4](#_Toc6592)

[1.1. Developing industry skills 4](#_Toc6593)

[1.2. Skills: add student 1 name here : Computer Science 4](#_Toc6594)

[1.3. Skills: add student 2 name here : Data Science 4](#_Toc6595)

[1.4. Skills: add student 3 name here : Software Development 4](#_Toc6596)

[1.5. Skills: add student 4 name here : Cyber Security 4](#_Toc6597)

[2. Level 2: Basic Technology 5](#_Toc6598)

[2.1. Tech Stack: add student 1 name here 5](#_Toc6599)

[2.2. Tech Stack: add student 2 name here 5](#_Toc6600)

[2.3. Tech Stack: add student 3 name here 5](#_Toc6601)

[2.4. Tech Stack: add student 4 name here 5](#_Toc6602)

[3. Level 3: Advanced Skills 6](#_Toc6603)

[3.1. Advanced features: add student 1 name here 6](#_Toc6604)

[3.2. Advanced features: add student 2 name here 6](#_Toc6605)

[3.3. Advanced features: add student 3 name here 6](#_Toc6606)

[3.4. Advanced features: add student 4 name here 6](#_Toc6607)

[4. Level 4: Advanced Knowledge 7](#_Toc6608)

[4.1. Advanced Knowledge: add student 1 name here 7](#_Toc6609)

[4.2. Advanced Knowledge: add student 2 name here 7](#_Toc6610)

[4.3. Advanced Knowledge: add student 3 name here 7](#_Toc6611)

[4.4. Advanced Knowledge: add student 4 name here 7](#_Toc6612)

[References 8](#_Toc6613)

General Instructions

You should use this LATEX template to generate your team project report. Keep in mind the following key points:

When we assess your report, you are not given a mark. Instead we will indicate (separately, for each team member) whether each level is achieved .

In order to pass the unit, you must achieve at least level 1.

In order to achieve level 2, you must rst have achieved level 1, and so on for each level up to level 4. This means that we will not assess a higher level until a lower level has been achieved (though we will review one level higher and give you feedback to help you in re ning your work).

Some parts of the report are completed as a team and other parts require each student to complete a di erent section. This means that for each submission, some members of the team may have completed their work for a given section, but other members may not. It also is therefore possible that some members of the team may achieve a speci ed level and other members of the team may not yet have achieved that level.

Even if some members are completing their material for a given level, and others are not, your team members will still need to work together to edit and compile the report. The only exception to this is where a member of the team has already achieved the level they are targeting in a previous submission and has decided to not attempt higher levels, and so is not contributing any further (this should be obvious because no level is indicated for that student on the cover page).

When completing each section you should remove the explanation text and replace it with your material.

For each submission you will add new details to this report, and/or update previous sections (where previous work was not good enough to have achieved the relevant level). In particular:

General: For each submission, each student can attempt up to 2 levels. You must also successfully achieve each lower level before you can be assessed at a higher level. For example, in the rst submission you might attempt only level 1, but not be successful in achieving that level. You then reattempt level 1 and add in level 2 in the second submission and are successful in achieving level 1 but not level 2. For the third and nal submission you could then attempt level 2, or levels 2 and 3 - or even just choose to not submit anything further and remain at level 1).

Submission 1: You should complete at least the material for level 1 (since achieving level 1 is required to pass the unit). Each member of the team can also optionally choose to complete the material for level 2.

Note 1: If you do not complete the level 2 information then you obviously cannot achieve level 2 at this stage. This does not stop you from attempting level 2 in Deliverable 2 or 3, but it will make it more di cult to achieve the higher levels later in the semester. Note 2: To be able to achieve Level 1 in submission one your team has to achieve level 1 in the group component (Section 1.1) and you have to achieve Level 1 in the individual component (i.e. your assigned section 1.2, 1.3, 1.4 or 1.5) Submission 2: Each member of your team will complete additional sections, but because you are submitting a single document, you need to work together to compile your results together and generate the nal submission.

If you did not achieve level 1 in your rst submission, then you should revise the material for level 1 based on the feedback, and optionally you can also complete level

2.

If you achieved level 1 in your rst submission, then each team member can optionally complete the material for levels 2 and 3. Note: If you do not achieve level 1 with this submission then the highest level you will be able to achieve in the nal submission will be level 2. If you achieve level 1, but not level 2, with this submission then the highest level you will be able to achieve with the nal submission is level 3.

Submission 3: Again, you can correct sections where you did not achieve the speci ed level in the previous submission, and you complete additional sections. If you still have not achieved level 1, then you should revise the material for level 1 based on the feedback, and again optionally you can also complete level 2.

For those at level 1, you can choose to complete the material for levels 2 and 3.

For those at level 2, you can choose to complete the material for levels 3 and 4. For those at level 3, you can choose to complete the material for level 4.

Whilst the team project is just that a team project it has been designed to also allow di erent members of the team to achieve di erent outcomes. We do expect you to work together as a team. If you do come across problems working together then the rst step should be to discuss this with your tutor. Note: If you are having problems you should approach your tutor as soon as you can to make them aware of the di culties you are having with your team.

Finally, you should also ensure that any resources you use are suitably referenced, and references are included into the reference list at the end of this document. You should use APA 6th reference style (The University of Sydney, 2022).

## 1. Level 1: Basic Skills

Level 1 focuses on basic technical skills (related to LATEX and Git) and the types of skills used in di erent computing jobs.

### 1.1. Developing industry skills

This section is completed as a team.

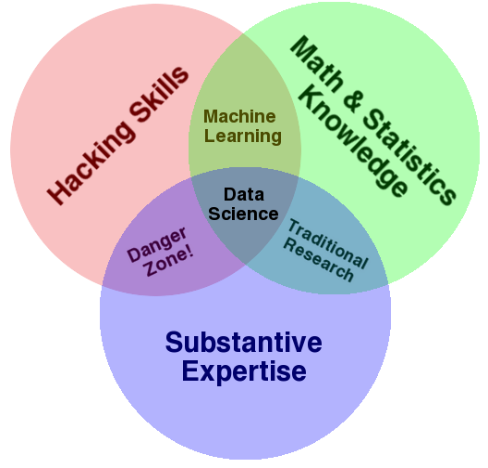
Throughout your Computing degree we will help you learn a range of new skills. Once you graduate however you will need to continue to learn new languages, new tools, new applications, etc. For this section you need to identify 5 approaches you can take to this continual learning. You should then put these in order from most e ective to least e ective, and then explain the circumstances in which each approach might be appropriate. (Target = ∼100 words per skill = ∼500 words total).

This section is completed individually. Each member of the team should independently complete a separate copy of this section.

You should begin by allocating to each team member a di erent major to focus on (i.e. one of: Computer Science; Data Science; Software Development; Cyber Security). If you have a fth member, then your tutor will suggest a fth topic to cover. You should then undertake research into the typical practical skills that you believe would be most important to someone who graduates with this major and is then working in industry. You should list the 8 skills that you believe are most important and for each one give a short explanation as to why you feel it is important. (Target = ∼100 words per skill ∼800 words total per student).

### 1.2. Skills: Xiangyi Chen: Computer Science

### 1.3. Skills: Xiyan Huang: Data Science

(Conway, 2010)

While some of the junction labels are a little sarcastic, I believe this graphic captures the heart of what I believe people mean when they say "data science": it is inherently an interdisciplinary field. The skills of a statistician who knows how to model and summarize datasets; the skills of a computer scientist who can design and use algorithms to efficiently store, process, and visualize this data; and the domain expertise—what we might think of as "traditional" training in a subject—necessary both to formulate the right questions and to put their answers in context. (Vanderplas, 2017) There are some kinds of skills below that are considered as the most useful ones in our future learning.

1. Mathematical skills: From my perspective, mathematic especially linear algebra, calculus, probability and statistics are quite important for the future learning. Math helps us better understand the principle behind the formula and the code we use.

Linear algebra is an area of mathematics that comes in handy when it comes to data science and machine learning. The majority of machine learning models may be written as a matrix, which is a common representation of a dataset. Data preparation, data transformation, and model assessment all involve linear algebra. (“Essential Linear Algebra for Data Science and Machine Learning,” 2021)

In machine learning models, calculus and optimization are used to arrive at the final answer. When we create a suitable algorithm model, the problem's final solution frequently includes an optimization problem. Without the help of differential theory and calculation methods, no elegant model can be produced in the process of exploring the extremes of data space. As a result, the only approach to the final answer is to grasp the basic idea of multivariate differentiation and the optimization realization technique. (Zhang, 2020)

Probability statistics is a technique for deducing rules from data and speculating on the unknown. (Zhang, 2020) For hypothesis testing and distributions like the Gaussian distribution and probability density function, probability crucial. (“Math and Data Science: What Do You Need to Know?,” 2020)

Statistics is at the heart of advanced machine learning algorithms in data science, capturing and translating data patterns into actionable evidence. Statistics are used by data scientists to collect, review, analyze, and draw conclusions from data, as well as to apply quantified mathematical models to appropriate variables. Data scientists operate in a variety of capacities, including as programmers, academics, and corporate executives. However, there is one thing that all of these fields have in common: a statistical foundation. (“How Much Do Data Scientists Need to Know about Statistics?,” 2021)

1. Python skills: Python is a high-level, open-source, interpreted programming language that offers an excellent approach to object-oriented programming. It is one of the most popular languages used by data scientists for a variety of projects and applications. Python has a lot of features for dealing with arithmetic, statistics, and scientific functions. (Chan07, 2018) Python has established itself as a first-class tool for scientific computing activities, such as the analysis and visualization of huge datasets, during the previous two decades. This may have surprised early Python proponents, given the language was not built with data analysis or scientific computing in mind. NumPy for manipulation of homogeneous array-based data, Pandas for manipulation of heterogeneous and labeled data, SciPy for common scientific computing tasks, Matplotlib for publication-quality visualizations, IPython for interactive execution and sharing of code, Scikit-Learn for machine learning, and many more tools will be discussed in the following sections. (Vanderplas, 2017)

R skills: R is a suitable tool for applying different statistical operations on Data Science since it is heavily reliant on statistics. R is appealing for a variety of data science applications because it has aesthetically pleasing visualization tools such as ggplot2, scatterplot3D, lattice, and highcharter. (Pandurang Choudhari says, 2019) R is a statistical computing and graphics programming language that you may use to clean, analyze, and graph your data. It is frequently used to estimate and present results by researchers from many fields, as well as by lecturers of statistics and research methodologies. (Yee, 2017)

### 1.4. Skills: Felicia Chen: Software Development

fs Your text goes here

### 1.5. Skills: Chenrui Li: Cyber Security

Your text goes here

## 2. Level 2: Basic Technology

Level 2 focuses on initial evaluation of the tech stack that is used by a selected company. All companies make use of a range of technologies, and these technologies need to work together. A tech stack is basically just this collection of technologies that collectively enable a company’s systems. As an example, one of the most common technology stacks for supporting web servers is LAMP: Linux as the underlying operating system; Apache as a web server; MySQL as the supporting database; and Perl (or more recently PHP or Python) as the programming language.

Each student should choose a di erent tech stack and explain the role of each of the di erent technologies in that stack. Note that prior to researching your proposed tech stack and spending time writing about it, it might be a good idea to check with your tutor as to whether your chosen stack is suitable. (Target = ∼200-400 words per student).

### 2.1. Tech Stack: add student 1 name here

Your text goes here

### 2.2. Tech Stack: add student 2 name here

Your text goes here

### 2.3. Tech Stack: add student 3 name here

Your text goes here

### 2.4. Tech Stack: add student 4 name here

Your text goes here

## 3. Level 3: Advanced Skills

Level 3 focuses on more advanced technical skills (LATEX and Git) and analysis of linkages and relationships between the items in the company tech stack.

The following is a list of advanced Git and LATEX skills/features. Each student should select one pair of items from each list and demonstrate actual use of each item (either through activity in Git, or through including items in this report). (Target = ∼100 words per student for each feature).

Git

Rebasing and Ignoring les

Forking and Special les

Resetting and Tags

Reverting and Automated merges Hooks and Tags

LATEX

Cross-referencing and Custom commands

Footnotes/margin notes and creating new environments

Floating gures and editing style sheets

Graphics and advanced mathematical equations Macros and hyperlinks

### 3.1. Advanced features: add student 1 name here

Explain your use of the advanced Git and LATEX features.

### 3.2. Advanced features: add student 2 name here

Explain your use of the advanced Git and LATEX features.

### 3.3. Advanced features: add student 3 name here

Explain your use of the advanced Git and LATEX features.

### 3.4. Advanced features: add student 4 name here

Explain your use of the advanced Git and LATEX features.

## 4. Level 4: Advanced Knowledge

Level 4 focuses on analysing your particular tech stack and considering alternatives. Each student should consider the tech stack they described for Level 2, and then discuss each of the following points:

What are the strengths and limitations of this stack? (Target = ∼200 words).

What alternatives exist, and under what situations might these alternatives be a better choice? (Target = ∼200 words).

### 4.1. Advanced Knowledge: add student 1 name here

Your text goes here

### 4.2. Advanced Knowledge: add student 2 name here

Your text goes here

### 4.3. Advanced Knowledge: add student 3 name here

Your text goes here

### 4.4. Advanced Knowledge: add student 4 name here

Your text goes here

# References

* The University of Sydney. (2022). APA referencing style (6th edition). (See <https://libguides.library.usyd.edu.au/ld.php?content_id=47913440>)
* Essential Linear Algebra for Data Science and Machine Learning. (2021, May 21). Retrieved from KDnuggets website: https://www.kdnuggets.com/2021/05/essential-linear-algebra-data-science-machine-learning.html
* Math and Data Science: What Do You Need to Know? (2020, September 7). Retrieved March 8, 2022, from Simplilearn.com website: https://www.simplilearn.com/math-and-data-science-article
* Zhang, Y. (2020, February 20). What roles do calculus, probability, and linear algebra play in machine learning? - Cloud + Community - Tencent Cloud. Retrieved from cloud.tencent.com website: https://cloud.tencent.com/developer/article/1590800
* How Much Do Data Scientists Need to Know about Statistics? (2021, September 24). Retrieved from School of Data Science website: https://datascience.virginia.edu/news/how-much-do-data-scientists-need-know-about-statistics
* Chan07, D. (2018, October 23). Python for Data Science. Retrieved from GeeksforGeeks website: https://www.geeksforgeeks.org/python-for-data-science/
* Conway, D. (2010, September 30). Drew Conway. Retrieved from Drew Conway website: http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram
* Vanderplas, J. T. (2017). Python data science handbook : essential tools for working with data. Beijing Etc.: O’reilly, Cop.
* Pandurang Choudhari says. (2019, April 22). Why Choose R for Data Science - Discover Top Features and Companies. Retrieved from DataFlair website: <https://data-flair.training/blogs/using-r-for-data-science/>
* Yee, S. J. W. and D. (2017). Why You Should Become a UseR: A Brief Introduction to R. APS Observer, 30(3). Retrieved from https://www.psychologicalscience.org/observer/why-you-should-become-a-user-a-brief-introduction-to-r