Individualized Learning Experience Framework: Teaching of Data Science to Non-Computer Science Students

Gholamreza Rafiee
Queen's University Belfast, The United Kingdom

Abstract

Data science is a difficult subject to learn because it requires a wide range of prerequisite skills. This paper describes a data science teaching practice that has been running for over three years for students in two different non-computer science pathways including CIT (computing and information technology) and BIT (business information technology) at the school of Electronics, Electrical Engineering and Computer Science (EEECS) in Queen's University Belfast (QUB). In this paper, a novel framework is proposed by which student learning experiences are personalized before the module begins in an academic year. By identifying each student's requirements and skills to focus on using a multilevel assessment approach, individualized content recommendations for each student will be provided in the first week of teaching. This enables each learner to concentrate on essential customized learning content materials and self-directed learning. Some preliminary evidence for the efficacy of this framework in improving student learning experience is discussed.

Keywords: content recommendation, data science, individualized learning experience framework, prerequisite skills, self-directed learning

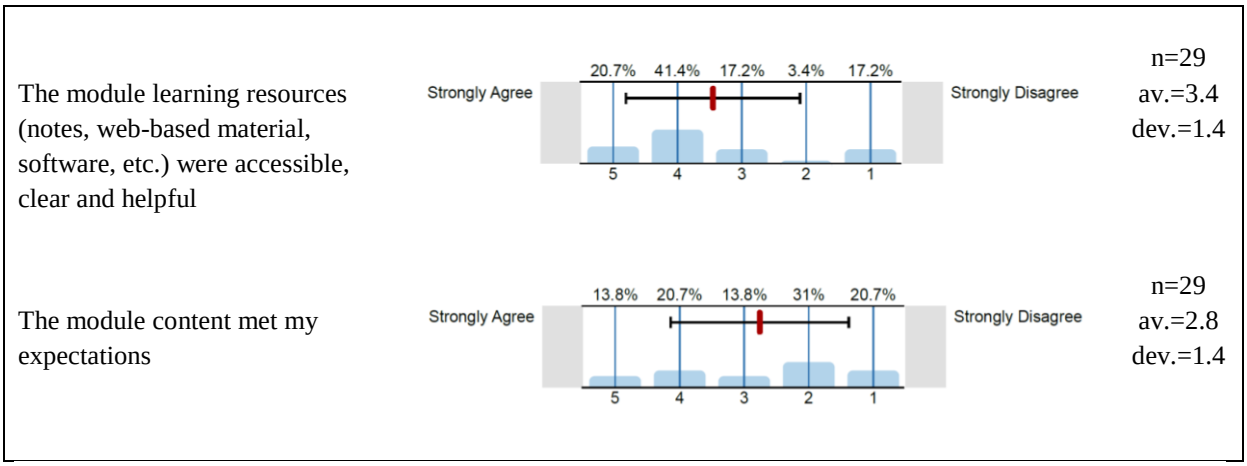
1. Introduction

Data science is a difficult subject to learn because it requires a wide range of prerequisite skills. As some of these skills are not directly as part of the learning outcomes of a data sceince based module, therefore, they could be easily neglected and resulted in a poor learning experience for students. This paper describes a teaching practice of a data science based module on campus to non-computer science students in two different undergraduate pathways, business information technology (BIT) and computing information technology (CIT). Figure 1 illustrates student feedback results in link with the *Module Learning Resources* and *Student Content Expectations* of this module in 2019 before starting the Pandemic. Despite having access to all module learning resources, students were not satisfied

with the module content materials. The average of fulfilling students‘ expectation was about 2.8 with the standard deviation of 1.4.

In this teaching and education research, the aim was to individually customise student learning experiences before the module starts in an academic year. We focused on identifying student weaknesses and skill gaps required in a data science based module by initially asking students to complete a multi-level assessment questionnaire before the week one of an academic year. Analysing this questionnaire helped us to suggest essential learning content materials for each learner resulted in an improved fulfilling expectations. Some preliminary evidence for the efficacy of this teaching practice in improving student learning experiences is discussed in the Results and Discussion section.

Figure 1: Student feedback results for the "Module Learning Resources and Content Expectations" before using the framework.



1.1 Background

Applying new teaching innovations and strategies in higher education and teaching practice is an ongoing pedagogy (Struyven, 2003). Most of these strategies inspired by learning theories considering two factors including how a student learns and how to provide an effective educational content (Hughes, 2010; Coderch, 2021; QIA, 2008). In this paper, by focusing on both factors (i.e., identifying student study skills and previous learning experiences), a new framework is proposed by which learning experiences are customized. This framework will be providing an effective starting point for each learner in a data science module based on a diagnostic assessment.

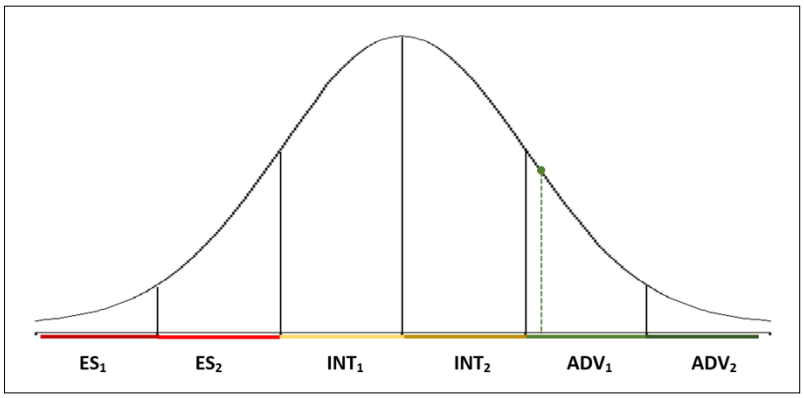
At the school of Electronics, Electrical Engineering and Computer Science (EEECS) in Queen's University Belfast (QUB), 'Data Analysis and Visualisation' module was initially taught in R programming for all pathways including computer science (CS), software engineering (SE), CIT and BIT till 2020. Since 2021, this module has been taught using Python (Jupyter Notebook platform) for both CIT students as a compulsory and BIT students as an optional module. This paper focuses on teaching practice of this module to both CIT and BIT pathways with the total number of 72 students while utilising the diagnostic

assessment (also called baseline assessment) to identify students' learning strengths and needs.

After compeleting the baseline assessment by almost all students, student skills are evaluated and categorised in six various levels including essential 1 (ES_1), essential 2 (ES_2), intermediate 1 (INT_1), intermediate 2 (INT_2), advanced 1 (ADV_1) and advanced 2 (ADV_2). Figure 2 demonstrates the various levels of student skills and the performance (i.e., ability score) of a BIT student after completing this assessment. It is worth mentioning that about 65% of CIT students' ability scores are categorised in ES_1 and ES_2.

We aimed to establish this framework by whcih students could receive some suggestions for content materials required to learn in this module as self-directed learning (Knowles, 1970, 1975; Loeng, 2020).

Figure 2: The performance of a student categorsied in one the six levels. Each student is given an ability score by completing the baseline assessment. 79th percentile is the ability score of a BIT student illustrated as a green dashed line.



2. Baseline Assessment

The module, Data Analysis and Visualisation, runs throughout the third year of undergraduate in two pathways, BIT and CIT with 27 and 45 students, respectively starting in September 2022. This module has two main coursework projects each 45% alongside 10% for engagement and self-directed learning. The learning outcomes include pre-processing (such as visualisation, dimension reduction and addressing missingness), model development (such as unsupervised and supervised learning) and interpretation with discovery.

This 12-week module delivery includes two hours in-person teaching, one-hour online session with live Python scripting and two hours practical sessions per week. A week before starting the semester (from September to December), all students are asked to complete a baseline assessment (i.e., online questionnaire). To better understand the ability of each learner and customise their prerequisite learning contents, potential skills or abilities and performance of every student is assessed.

A questionnaire designed based on assessing different skills was presented to students via CANVAS, the University's virtual learning environment (VLE). Three main assessments in

this questionnaire include *study skills*, *background in Python programming* and *learning style*.

The *study skills* questionnaire giving a general idea of students' abilities is divided into 8 sections (UHCL, 2022). Each section has an assessment including 8 questions with the posssible responses between 1 to 4 for each question (1=Never, 2=Sometimes, 3=Usually, 4=Always). This quetionnarie usually takes less than 10 minutes to be completed by students as follows:

Time Management and Procrastination

- *I arrived at classes and other meetings on time.*
- *I devoted sufficient study time to each of my courses.*
- *I schedule definite times and outline specifies goals of my study time.*
- *I prepare a "to do" list daily.*
- *I avoid activities which tend to interfere with my planned schedule.*
- *I use prime time when I am most alert for study.*
- *At the beginning of the term, I make up daily activity and study schedules.*
- *I begin major course assignments well in advance.*

Concentration and Memory

- *I have the "study-place habit", that is, merely being at a certain place at a certain time means time to study.*
- *I study in a place free from auditory and visual distraction.*
- *I find that I am able to concentrate – that is, give undivided attention to the task for at least 20 minutes.*
- *I am confident with the level of concentration I am able to maintain.*
- *I have an accurate understanding of the material I wish to remember.*
- *I learn with the intention of remembering.*
- *I practice the materials I am learning by reciting out load.*
- *I recall readily those things which I have studied.*

Study Aids and Note-Taking

- *While I am taking notes I think about how I will use them later.*
- *I understand the lecture and classroom discussion while I am taking notes.*
- *I organize my notes in some meaningful manner (such as outline format).*
- *I review and edit my notes systematically.*
- *I take notes on supplementary reading materials.*
- *I have a system for marking textbooks.*
- *When reading, I mark or underline parts I think are important.*
- *I write notes in the book while I read.*

Test Strategies and Test Anxiety

- *I try to find out what the exam will cover and how the exam is to be graded.*
- *I feel confident that am prepared for the exam.*
- *I try to imagine possible test questions during my preparation for an exam.*
- *I take time to understand the exam questions before starting to answer.*
- *I follow directions carefully when taking an exam.*
- *I usually get a good night's rest prior to a scheduled exam.*
- *I am calmly able to recall what I know during an exam.*
- *I understand the structure of different types of tests, and am able to prepare for each type.*

Organizing and Processing Information

- *When reading, I can distinguish readily between important and unimportant points.*
- *I break assignments into manageable parts.*
- *I maintain a critical attitude during my study - thinking before accepting or rejecting.*
- *I relate material learned in one course to materials of other courses.*
- *I try to organize facts in a systematic way.*
- *I use questions to better organize and understand the material I am studying.*
- *I try to find the best method to do a given job.*
- *I solve a problem by focusing on its main point.*

Motivation and Attitude

- *I sit near the front of the class if possible.*
- *I am alert in classes.*
- *I ask the instructor questions when clarification is needed.*
- *I volunteer answers to questions posed by instructors in the class.*
- *I participate in meaningful class discussions.*
- *I attend class regularly.*
- *I take the initiative in group activities.*
- *I use a study method which helps me develop an interest in the material to be studied.*

Reading and Selecting the Main Idea

- *I survey each chapter before I begin reading.*
- *I follow the writer's organization to increase meaning.*
- *I review reading material several times during a semester.*

- *When learning a unit of material, I summarize it in my own words.*
- *I am comfortable with my reading rate.*
- *I look up parts I don't understand.*
- *I am satisfied with my reading ability.*
- *I focus on the main point while reading.*

Writing

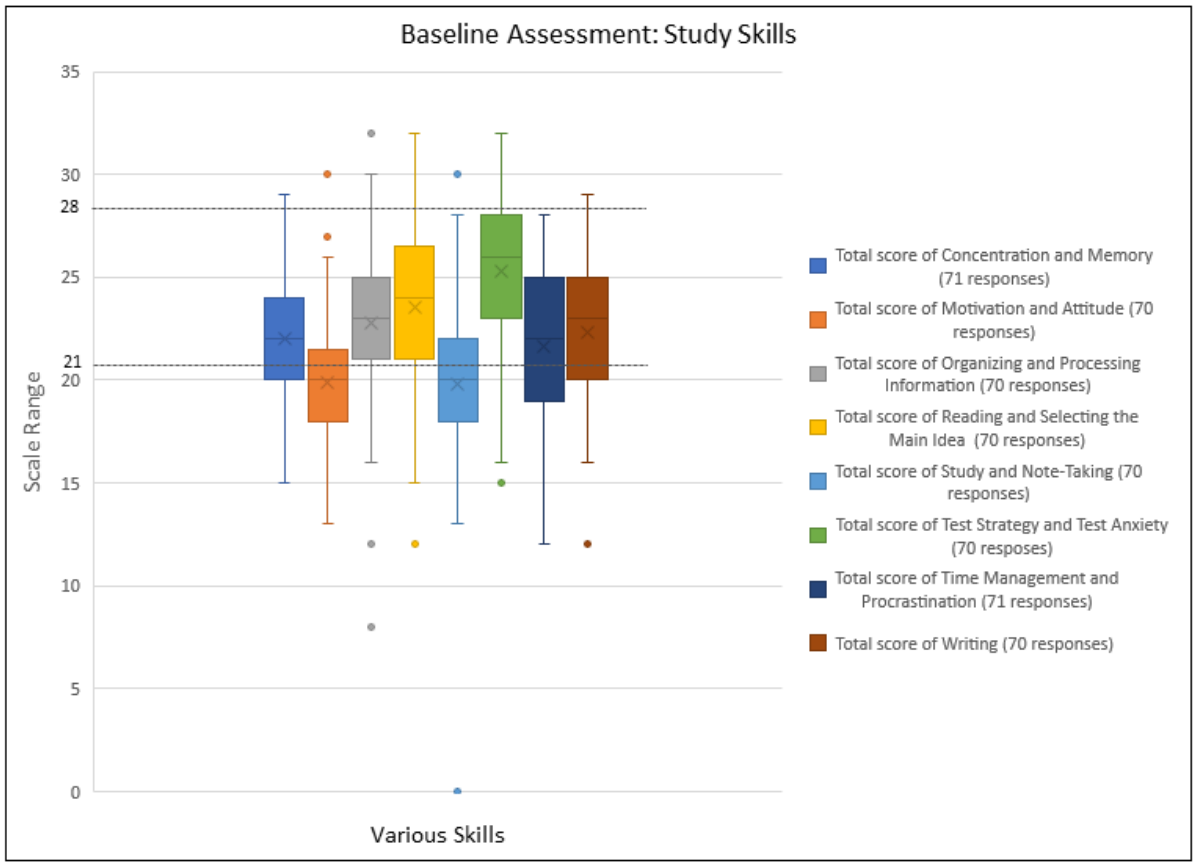
- *I find that I am able to express my thoughts well in writing.*
- *I write rough drafts quickly and spontaneously from notes.*
- *I put aside a written assignment for a day or so, then rewrite it.*
- *I review my writing for grammatical errors.*
- *I have someone else read my written work and consider their suggestions for improved writing.*
- *I am comfortable using library resources for research.*
- *I am able to narrow a topic for an essay, research paper, etc.*
- *I allow sufficient time to collect information, organize material, and write the assignment.*

The *Python background programming* questionnaire includes 6 questions. The *learning style* questionnaire including 20 questions was adapted from the Pennsylvania Higher Education Assistance Agency (PHEAA), one of the nation's leading student aid organizations (PHEAA, 2022).

2.1 Results and Discussion

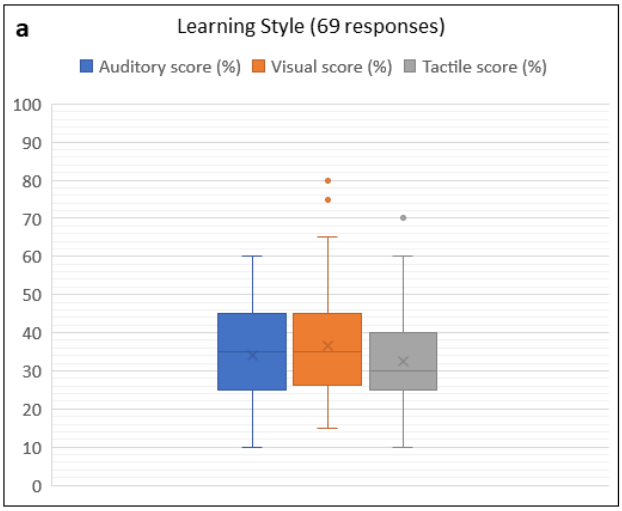
The baseline assessment was completed by 71 students and all answers were analyzed accordingly. The Figure 3 illustrates a summary of overall responses in a boxplot format representing the distributional characteristics of each assessment in study skills.

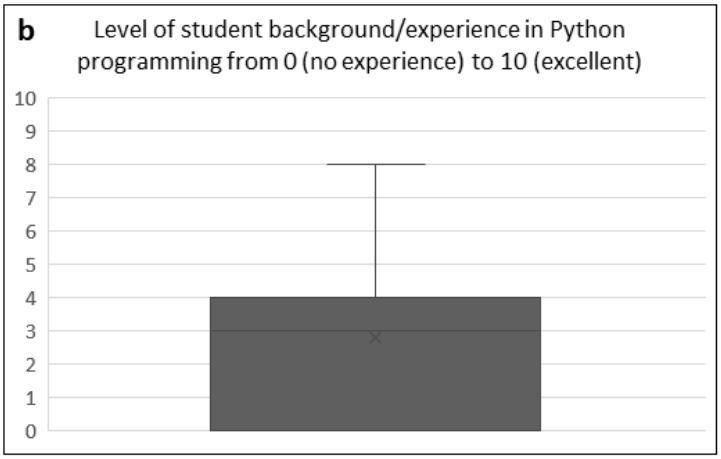
Figure 3. The Study Skills as the first part of the baseline assessment.

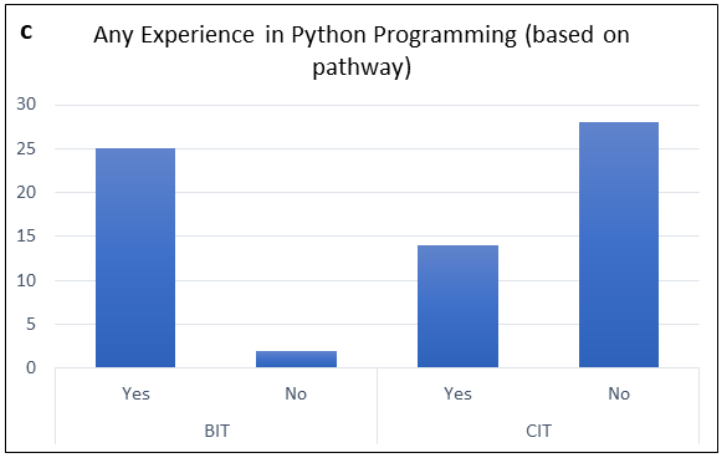


In this assessment, two thresholds (i.e., 21 & 28) were considered as illustrated in black dashed line in Figure 3. A learner with the score of 20 or less in a particular skill was suggested to make use of some resources in that skill. The resources were provided via CANVAS VLE. A learner with a score between 21 to 28 (i.e., moderately good) was provided few optional resources for improving some abilities in that skill. Any score above 28 is considered as an excellent skill strategy by which a learner does not require to focus on any essential resource. In this questionnaire, the medians of 'Motivation and Attitude' (illustrated by orange color) and 'Study and Note-Taking' (illustrated by light blue color) scores are both at the same level (i.e., around 20) and have rather similar distributions of scores for most of the students in this module. The second and third parts of the baseline assessment including *learning style* and *Python programming background* are shown in Figure 4 (a, b & c).

Figure 4. The Learning Style (a) and Background in Python Programming (b & c) as the second and third part of the baseline assessment.







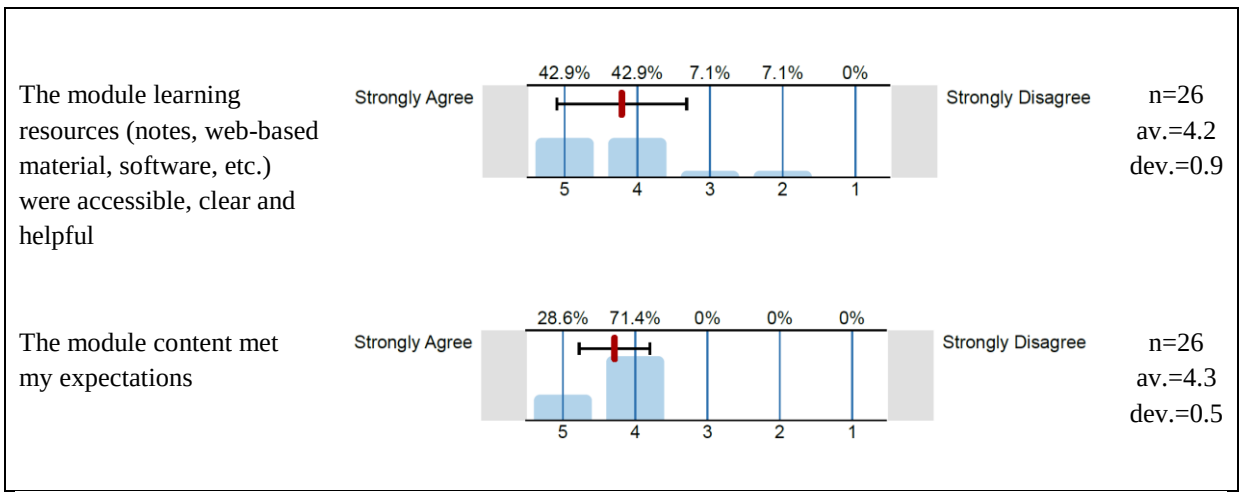
According to the results of the baseline assessment, students were recommended different content materials to complete as self-directed study and learning within week 1 and week 2 of the semester. The recommended educational resources included educational videos and

practices in link with Python programming at different levels such as 'Getting started with Python', 'Loading data using Panadas', 'Data manipulation with Pandas' and 'Data types for data science in Python'. The recommended educational resources including videos and practices were provided by *DataCamp* platform (Datacamp, 2022).

Students are also given several tips and guidance according to their *learning styles* to learn better during lectures. For example, a learner who has been considered as a visual learner is advised for a regular eyesight checking, using flashcards to learn new content, visualising things that student hears, writing down keywords, ideas and instruction and avoiding distractions during study times (PHEAA, 2022). Completion of these tasks by the learners in this module were also considered as part of the engagement and self-directed learning assessment.

Figure 5 illustrates student feedback results when using this framework for the *Module Learning Resources* and *Student Content Expectations* of this module obtained in December 2022. As it has been demonstrated in this figure, learners were found the module learning resources helpful and content expectations were successfully met. The average of fulfilling student expectations was about 4.3 with the standard deviation of 0.5.

Figure 5: Student feedback results for the "Module Learning Resources and Content Expectations" when using the framework.



3. Conclusion and Future Work

Applying this framework has several interesting points.

- Students from two different pathways can sit in a same lecture while focusing and working on their own customized learning materials.

- Student learning experience can be improved. The overall satisfaction achieved with the module learning resources and content expectations given by using this framework is about 4.3 out of 5.
- This framework can be applied to any other data science related modules including students from different pathways.

As the required educational content of a learner is identified according to a baseline assessment, customizing and then offering this content for each learner is a time-consuming task for an instructor. To resolve this issue and speed up the process of assigning self-directed learning content materials, an automated content recommendation system is suggested by which each learner's content is automatically provided.

Acknowledgment

The support of the PHEAA and the University of Houston in providing the questionnaires are gratefully acknowledged.

References

Coderch, M. (2021), "An initial assessment of the academic and professional profile of modern foreign languages' teachers in UK higher education" Language Learning in Higher Education, vol. 11, no. 2, pp. 527-535. https://doi.org/10.1515/cercles-2021-2033

Datacamp online coding platform (2022): https://www.datacamp.com/

Hughes, N., Schwab, I. (2010), Teaching Adult Literacy: Principles and Practice, Open University Press.

Knowles, M. S. (1970), The Modern Practice of Adult Education: Andragogy versus Pedagogy, Association Press, New York, NY, USA

Knowles, M. S. (1975), Self-directed Learning. Cambridge, MA: Cambridge Adult Education

S. Loeng, Self-directed learning: A core concept in adult education. Education Research International, 2020

Struyven, K., Dochy, F., Janssens, S. (2003), Students' Perceptions about New Modes of Assessment in Higher Education: A Review

The Pennsylvania Higher Education Assistance Agency (PHEAA) (2022), http://www.educationplanner.org/

University of Houston (UHCL) Counseling Services (2021), SSCB Suite 3103, www.uhcl.edu/counselingservices

QIA (2008), Initial and Diagnostic Assessment: a learner centred process, www.excellencegateway.org.uk