**DISCUSSION GROUP ONE (JULY 16TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | From past experience: Grading portion, not really about the content which is okay. The difficulties are much more on examination or in the other words, the assessment. The teaching content was comfortable.  For the future expectation: from my current standpoint,I think I will face no difficulty. This perhaps also due to that I will not enroll in such courses in near future |
| After doing these four activities, can you briefly explain to me what data storytelling is? | Data storytelling: how using data to form the conclusion of a story (expressing some unconfident) |
| How do you feel about the data stories given? | Quite long. That’s all, I think I have another thing to say |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | The preference: the storytelling is easier to understand. In terms of differences, the conventional one is quite basic and “dry”, while the storytelling approach makes the concept easier to digest. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | Prior understanding of concepts, visualizations, and explanations: Roughly, I understand what the approach is trying to convey such as concepts, visualizations, and explanations.  For those who unfamiliar: this approach could help For someone who didn’t know the data science, since it provides the context (I can see the confidence) |
| Do you have any suggestions for the design or composition of DS? | No proper responses |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | The use of GenAI: Yes, I use Chatgpt for brainstorming some ideas.  The use of GenAI for Data Science courses: No, I did not use that in supporting any DS course that I enrolled in. While for the future, I also will not use it, since based on my own experience, it is not appropriate for conducting further analysis since it could only perform some easier task such as writing etc. |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | Importance of Data literacy Skill in this era compared before: the data literacy skill is still important as before. Even with the advancement of AI, since this skill could help us, for example like arranging the report.  If the courses related to DS Literacy are not provided in NUS, do you still consider its importance? Will you go to another institution that provides it: not to the extent that I would go to other institutions (seeking another alternative) if NUS is not providing any courses related to data skill literacy. even though I knew its importance before. I probably won't go to another uni.  The aspect that I want to learn in Data related courses that I have taken: Nope, so far it was sufficiently covered important aspects  The biggest takeaways: the final conclusion that provide recommendation how the data could be utilized |

**DISCUSSION GROUP TWO (JULY 17TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: The topic explanation is just too quickly, every covered topic just goes on and on. Yet, our prior class did help us, and for those who are just new to this course, they will be struggling.  B: Quite similar, since I do have some background with a specific course did help me, so then my other friend that was just new was struggling. For me, I found out that some courses related to DS that I enrolled in are pretty okay since it's comprehensive enough. Yet, there was a course that we found confusing, luckily, the TA did help provide a step by step approach. The TA explained the basics of python which was very helpful. While the class session only covered the theoretical parts. There is mismatch between class and tutorial  C: the content is condensed and focused much more on how to apply the code, without explaining what it was. Eventually, it led to copy and paste. So, the takeaway is how we understand the question rather than understand the data, which I believe is not the outcome that school is looking for. Also, I found it is difficult to link what I am doing to the real life things. But the other course that led me to have my own data project has provided me with knowledge and skill specifically when it comes to real life applicability. Hence, it was enjoyable.  The general consensus in terms of differences between some courses related to DS.  C: there is a course that its “bulk” was steep bcs of excel usage, but there is also a course that utilized R, so I would consider this as a fundamental difference, in which the tool involved. Which in my personal opinion, I would prefer excel, since I found learning R was not enjoyable, since R was quite difficult to learn |
| After doing these four activities, can you briefly explain to me what data storytelling is? | A: presenting the data in a more digestible approach  B: It presents the narrative which is much nicer since it provides the step by step and the difficulty found which hence led to some changes. These kind of information made the data easier to understand  C: is much like focusing on the interpretation aspect, since the data is number in nature, so it provides the narrative that makes it easier to be interpreted. |
| How do you feel about the data stories given? | A: Initially, I was not sure what it is, since it is quite lengthy. So, I think it would be much more effective if they could condense it by making it into a paragraph only, especially the first part. While for the second and third one, even though it is a full story, I still got the sense how it flows, yet it is still very long.  B: I share the same sentiment as A, since I found the full story in the first part was long and the dependent put a lot of effort. Also in the second and third part, I could see that some lines were crafted personable, yet I do appreciate that. Back again, the first part which provides the setting is still too long, about one page, and I was thinking of skimming it.  C: On the contrary, I found myself enjoying the sections, which from the content wise, it is how they link the students in university caught my interest, particularly how they feel. Including the data visualization, I still found no major difficulty to understand.  The possibility of fictional setting on the data stories into real life:  B: Yes, since I majoring life sciences and from what I heard from some alumnus who embark on some jobs, they say that often time they are assigned to analyze the data, so it is relevant |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | A: Actually, I don't mind the conventional one, since perhaps it is quite straightforward. While for the data storytelling one, I do believe that it requires other effort just to link it and translate into concrete and succinct note  B: I prefer a mixture between them. For example, in the beginning, it would be better if the presentation utilizes the data storytelling approach to better ground the understanding for students and be much more fun. And perhaps, this approach is also suitable to the juniors who are just onboard to the program so they may find themselves comfortable in learning this kind of course. And in the same time, the conventional approach such as using the summary note could be utilized for some important formulas  C: I also think the mixture one could be an option. The data storytelling is suitable for demonstrating the application and the conventional method is suitable for breaking down more on the theoretical aspect. Yet, I still prefer the data storytelling because I want to understand something more comprehensively before I apply it |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | For newest one:  A: I think yes, because the  B: I think for someone with zero background on it, a supportive sentence on a specific concept such as Mean and Median perhaps could help them more. For example the advance terms such as coefficient regression and coefficient correlation which some may think it somehow identical, so then, a more explanation on it could help.  C: Yes, but I think some parts need slight improvement. I meant, the story is good, yet for the technical part, I believe some extra explanations are required, for example adding the definitions for some technical terms |
| Do you have any suggestions for the design or composition of DS? | A: I think it would be better if the step-by-step approach for creating data visualization could be included as well through excel for example. Because, the mean, median, mode and standard deviation could be generated quickly by using excel or other tools  B: On the contrary, I think that the part of data visualization was not that necessary, since we were meant to understand and analyze the data rather than creating the visualization by ourselves. One thing that could be used is perhaps some quizzes at the end of each part and pop out the answer once we respond to it, so then we could evaluate how far we understand the given concept. So it will enhance our confidence as well.  C: I think if the team could provide a conjoint or chance for us to modify the scenario that fits our interest, it would also enhance engagement level. |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | Knowledge of GenAI  All: ChatGPT  A: using AI for summarizing the articles  B (Gemini): for generating question and resume writing  GenAI for DS:  A: I have my friend who utilized ChatGPT to finish up their assignment which it could provide some coding  B: I used chatgpt to check whether my coding was right, and I thought that chatgpt is quite good at providing coding. Yet, for some advance problem set, the chatgpt also could not solve it, led us to the conclusion that chatgpt has limitation as well  C: chatgpt is my best friend, since the course that I take on programming taught something more abstract, so then I asked chatgpt to pinpoint some easy questions. Yet, for long coding, oftentimes it led to confusion. But overall, it helpful |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | A: even more importance, since the information is more and more accessible these days, so then the skill in understanding the raw data and interpret on our own point is still important  B: yes because these days there are some instances where people manipulate the stats for their own benefit, so someone with low data literacy skill will highly likely fall into this trap. Also, understanding the data could avoid us from “cherry-picking” algorithm results. Since, it will enhance our awareness that everything that showed up from our social media for example, has been modified/set based on specific settings, hence it is not reflecting the whole truth.  C: Yes, it is more important, since it could help us distinguish which one is legitimate data and its analysis  The aspect that want to learn  A: I want to learn how the formula is given or derived (math parts) and how the stats/formula is applied  B:I prefer something on the surface, for example something related to our understanding such as time to use specific design/tool/approach and interpretation.  C: I want to see the involvement of more real life aspects to increase our applicability skills. |

**DISCUSSION GROUP THREE (JULY 18TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: The challenge was about navigating the correct direction; I meant in which part should I start (how do I start) once the data is given and what kind of approach should I take. Since the courses that I enrolled in mostly focusing on theoretical aspect, and merely emphasize the meaning or perhaps the applicability in real world, like (why I do need certain things)  B: in my side, I am thinking about the contenr’s coverage, yet, the professor did good job, not everything could be covered or comprehensively explained, though, I also do realize that since the course was kind of broad in the nature, it is almost imposibble to cover everything. Hence, we had to figure out by ourselves some specific aspects that attracted our interest yet weren’t covered in class. And, if you don't have or don't study with another peers, you will be overwhelmed. for example, in python, I did experience such a great support from my peer, since we helped each other  C (Not taking any DS related courses): I believe the maths or theoretical aspect is the things that I will face as challenges. |
| After doing these four activities, can you briefly explain to me what data storytelling is? | C: the approach that presents how the stats could be utilized in real life  B: the approach that allows us to be on the shoes of the one who work with the data itself, like how they approach that from different angles.  A: Use the story to bring up the data conceptualization, hence easier to understand |
| How do you feel about the data stories given? | A: they presented data is new to me, but it is good idea, and it addresses the problem that I found in first question, which is I dunno where and how to apply the data analysis approach  B: quite similar, which is when and how to apply the concept, for me all of them are quite neutral, I meant all four parts are equally important. I did enjoy the flow overall.  C: after a while, I found myself a little bit boring, that’s because the concept is quite intuitive so somehow not that interesting |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | A: I prefer the mixture of them. The conventional lecture notes provide the essential explanation, for example what is this and what is that. And the data storytelling approach is like its supplementary which explains how to apply it.  B: Similarly, I prefer both. The lecture notes are still very important to me. I do prefer to see the courses that ground the theoretical framework first while also involving a kind of interactive learning, for example asking us to work on regression coefficient or equation earlier, after that reveal the answer and approach to see its correctness. Which I believe the data storytelling may do have similar feature as this interactive approach  C: Both theory from lecture notes and the practicality from data storytelling. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | For someone who didn’t know these before:  A: I found a story a little bit complicated then without any initial sufficient explanation, the beginners may find some terms and approach difficult to understand and confusing  B: I think the beginners need to read whole parts then perhaps get to sense the meaning of it. Yet, the extra explanation either at introduction or middle parts, may also help them  C: I think more definitions will be much helpful, for example I had some difficulties in distinguishing the regression coefficient and correlation. |
| Do you have any suggestions for the design or composition of DS? | C: I think it could be more condensed, rather the lengthy one  A: The reading is too long, so sometimes I get lost of what I have read initially. Hence, I suggest making it shorter and easier to be remembered |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | B: Chatgpt: I am using of open book exam, like MC Qs, and I played with the prompt to try which one provide more accurate answer  C: Chatgpt, gemini; I am using in summarize the condensed paragraph but not in the position that asking it to answer specific question, since I know this machine is something that we can’t rely.  A chatgpt: I only use for personal purpose  For DS  A: No, but I am thinking perhaps there is a possibility where AI could be utilized in DS learning process, it is to ask the insight of specific data visualization or pattern, yet I am not sure whether this is feasible right now. I also thinking that the quick response that AI could provide as a way how it could support the DS learning process  C: I think AI could assist us in programming and explain certain concepts  B: I used it on several occasions for debunking and writing the code and ask its insight about the code that I made, since there is a part that covers how to use chatgpt for coding. |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | B: I think the skill itself still importance, yet perhaps the change is in which sub-skills that turned to be more crucial given the raise of AI, for example hard coding skill is not that important instead understanding and having the skill in the design large language model is more and more important these days.  C: The fundamental one like understanding the nature of data is still relevant. Yet the ground understanding like on the beneath process is less important than the skill and understanding on the meaning of it, and how it could be applied and how to improve it.  A: I am thinking about the timeless skill that we should have, I may take an analogy as calculator, no matter how advanced it is, we should have prior knowledge on it. This means that we are not completely rely on it, since it could be misleading. Similarly, I may use chatgpt to provide any coding, yet I will check and verify the output before utilizing it. this emphasizes how the skill and understanding fundamental concepts remain important |

**DISCUSSION GROUP FOUR (JULY 19TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: I think the challenging part was deciding the appropriate design at the very beginning. I mean, the decision process (including the purpose of analysis, and what kind of audience) itself requires us to have such comprehensive knowledge so we could provide and produce meaningful analysis.  B: Even though I enjoyed both lecture and tutorial sessions which have provided and explained step by step approach, but I think there was a lack practice, for example like the additional one so then we could practice, apply and play on. Also, like the targeted audience also wasn't comprehensively elaborated  C: As the one with no math background, I found that course was challenging in essence, and it was hard to understand. Also, as the background of class was varied so then it added another complexity to the lecture since he can’t accommodate all needs based on our level of prior understanding. The time spent in class also limited, so we didn’t get the chance to deep dive to the concepts and its practical application  D: Mostly my concern like C which for the basic concept like descriptive statistics almost all students could get it, but when moving forward to inferential statistics like p-value concept, I saw some students struggling to it, including myself. |
| After doing these four activities, can you briefly explain to me what data storytelling is? | A: It’s like how to understand the data from its beginning, I mean the design, which the first stage we should understand before proceeding to more technical steps. I also thinking that DS is kind of whole process including the interpretation and report for audience  B: DS kind of explain the process or step by step approach before the result is reported, in shorter word, it’s like a logical linking among the whole process  C: DS is like a composition that include the hypothesis, RQs, analysis approach. And it is descriptive  D: DS is like ground the context of the things that we would like to find out, most importantly, it provides the flow of process, including how they approach the data. |
| How do you feel about the data stories given? | B: Overall, I feel so interested to this story, since it showcases what happen in the middle of data process, instead of just present the result  D: I do kind of appreciate the DS since it provides the context and the flow, since it presents the rationale of certain concepts so then easier for us to apply in real life. Hence overall, it is more useful  A: The format itself is like the typical data analysis process, so I would say it is informative so it could be more effective for students to know better the data related courses.  C: It does help me to connect all the process, and I could imagine myself being in that kind of specific situation that specifically provided by the story. So, it helps me to understand more comprehensively |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | A: I think the DS is much more practical and useful for the students since the final output or expected outcome for students is when they could apply the data analysis concept to real context. The DS offers hands-on for students to get more engaged with practical things.  B: Personally, I still prefer both. The conventional approach will ground something more like theoretical approaches, for example like the variation of methods. And the DS could serve as a complementary which will cover the practical aspects. Essentially, this is also important considering the students themselves have different approaches to the learning process.  C: Even though I do recognize the balance between these approaches, I still prefer to see the learning approach that puts much more emphasis on the DS side.  D: I prefer both. The lecture notes could provide and teach the technical aspects in shaping our coding skill for example. While the DS could be utilized to apply it, for example like when we should use it. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | C: Somehow, I think this DS could be difficult to understand given that there is no explanation why they chose specific tool and method. But I also think that DS also could provide easier explanation since they are breaking down the process that they have taken, including the step-by-step approach.  B: I do think that some features featured at this DS could be challenging to be understood by those with no prior knowledge, for example like the scatter plot concept. Yet I do think that DS could be considered as an alternative to introduce Data related concept  D: I also think for those with no prior knowledge, certain term could be challenging for them, hence, provide a supplementary explanation of each term could be beneficial  A: I do think that the application of DS such as flip class approach which we initially focus much on the hands-on and implementation could be beneficial for all students |
| Do you have any suggestions for the design or composition of DS? | C: I think the provision of additional explanation of any statistical tools featured in the study will be helpful. So, then it will enhance the understanding specifically for those with zero/different background. Also, the rationale or justification why the included tools/methods/analyses are chosen.  B: I think that adding the additional explanation could be beneficial, but it also could distract the flow of story. So, I think instead of included on main passage, it could be better if they could put on the box beside the formula or visualization, such note for example  D: quite like B and C and nothing to add  A: Apart from DS materials, I think the lecture also could ask students to solve the problems by themselves. Also, the systematic tutoring of the concept of data analysis could be featured as well. |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | C: Chatgpt, I use to summarize the long articles and sometimes asking nonacademic option  B: Chatpgt, I use to summarize the long articles as well, and use for some courses like coding, which I hesitate to ask my friends. While sometimes I do understand the output produced by chatgpt, but often I find peer explanation is much more understandable.  D: I use for spanish which was quite accurate for trranslation and coding as well, for example like writing the long coding  A: Frankly speaking, I use chatgpt heavily and consider it as main assistance. For data analysis, even though I know how to code, I still prefer to utilize chatgpt to help me for long coding and indeed it is time saving and efficient. But I acknowledge that the overreliance could be the con of it so then diminishing our skill |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | AI and data science learning process: D: I think Chatgpt still could be beneficial in supporting the data related courses for example whether it could provide the definition of certain concepts and providing coding as well  B: AI still could help us clarify the concept but when it comes to large and longer coding, it could confuse as well. Yet, chatgpt could save our time  C: I think chatgpt helps our learning process more efficiently, specifically when you find none that could help you. Based on my own experience, it helped me to identify the error/mistakes I made.  A: I think the tutor could encourage the students to use chatgpt for real business cases and they should control its usage as well. I do admit that chatgpt is quite efficient and may increase the student’s performance  Data literacy importance:  C: of course, it still important, since chatgpt not always giving correct answer and it still requires our own knowledge to assess the chatgpt’s output accuracy  D: Yes, I do agree with the argument before which chatgpt could give you wrong answer so then we should have solid and fundamental knowledge. Also, I do realize that one thing that chatgpt can’t do yet is data cleaning, which we still need human recognizition  A: I think this question could be answered by time. I believe that data literacy skills are still important, but I believe that chatgpt could bring something different, for example, back years ago if we wanted to conduct advance analysis such as random forest, we should resort to some professionals. But now chatgpt could do that. So, what the school could do is to craft the courses that focusing on AI skill as response to this kind of situation, where changes have taken place  B: so far, I do agree with all 3 answers and nothing to add |

**DISCUSSION GROUP FIVE (JULY 23RD)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: Due to the lack of explanation of some concepts, often I just copy and paste. Also, there was a chance for us to come up with our own data storytelling project, but because of lack of examples, I still found it was difficult, even though the dataset was provided.  B: since I am not that good in math, so I was little bit struggling to the principal of data analysis,  C: The concepts that have been covered and explained in class, I found there was a lack of time for us to apply it. Also, there was no project that allowed me to apply for it as well. Hence, I may not be able to use it in the future. |
| After doing these four activities, can you briefly explain to me what data storytelling is? | C: DS is making use of story to teach the concept relevant to data analytics  B: pretty much the same  A: it’s like putting kind of messy data into a more organized one |
| How do you feel about the data stories given? | C: Personally, it’s interesting to see the data is presented in such manner, in the real-world scenario, then it could help the organization and interpreration  B: I like the step-by-step approach, I felt that it is easier to understand  C: I thought the flow was cool, particularly the data filter’s section in which the ground context is presented and how the anomaly is found. |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | C: for me, I prefer the DS approach one. I believe, this kind of approach makes me remember the concept better  B: for me, I still prefer the conventional one, even though I also could say that I understand the DS approach. Since the lecture notes help me to review the basic concept  A: I think the DS is good, but still having the lecture note is important as well. The DS could accompany the lecture note. Also, if we need to go back to review, the DS may be fluffy and troublesome. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | The importance:  C: I understood the concept  B: I understood as well, but unsure whether it is correct.  Someone with no prior knowledge:  C: I think depends, for example the usage of certain terms such as coefficient correlation that states in weak degree, may lead to confusion or partial knowledge for someone with zero prior knowledge. Since the content did not include the whole spectrum/category of which one is strong or weak correlation.  A: I found that some parts weren’t in-depth explained so I was little bit confuse, and I think those with no prior knowledge may experience the same thing  B: I think for beginners it is important to ground the basic concepts and definitions, hence to make everything clearer. |
| Do you have any suggestions for the design or composition of DS? | B: I think it would be better if the approach is also including us in a more interactive manner so we can also follow what the steps included  C: I somehow agree to that argument before. The data was given, but there was lack of explanation in how the data was processed and just showed the result such as Mean, Median and SD. Despite I am not sure either whether this kind of aspect is important for student  A: I think it depends on the purpose of this approach, whether it wants to teach how students get the result of data analysis or just the concept itself |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | General usage C: Chatgpt, I use that for checking grammar errors. I also use for some statistical analysis  B: Personally, I am not really using any AI tools, I just use it for coming up with ideas and summarizing some proposals. I also use that for Excel formula  A: I use it all the time, including modifying the coding and coming up with the designing. Since some parts of my work is repetitive, so I ask chatgpt to do that and I will modify that on my end  Usage of AI for DS:  A: I used chatgpt to provide some explanations of certain codings that not covering during class  AI for DS purpose suggestion:  C: I think it depends on the student themselves, since each student is different  B: I think the school could suggest the students resort to chatgpt once they aren’t understanding some concepts before approaching the lectures. And provide some kind of question/prompt templates to student that may be better for student  A: I think the AI could be beneficial to breaking down the complicated coding. The student also should specify the question that they’d like to ask to AI so then it can provide the relevant responses. |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | The importance of data literacy skills  A: I think it is still important considering the data itself may be messy (and a lot of its variation) and need for cleaning and modification and it requires the skills and knowledge to do so. And having the skill in understanding the data can help better understand the formulation.  B: In my opinion, data literacy still important since the design of analysis itself requires certain knowledge and skill  C: I think it’s still important since we can’t fully rely on any AI including Chatgpt without they are coming up with certain conclusion |

**DISCUSSION GROUP SIX (JULY 25TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: I found the learning curve to be quite steep. We had to learn how to use R and other tools that I didn’t have much experience with, especially since in my secondary school, we didn’t really engage with these kinds of software. Adapting to this new content was challenging.  B: I have never enrolled in such a class, yet I do believe that some of the content felt like what I had learned before, but it was presented in a more advanced way, which made it a bit of a struggle to keep up. I felt that it wasn’t easy to understand and required more time to grasp compared to other subjects I’ve studied. Concepts in data science seemed more challenging to me, and I needed more effort to fully understand them. |
| After doing these four activities, can you briefly explain to me what data storytelling is? | A: I think data storytelling involves using real-life examples to make data science concepts more relatable and easier to understand. By applying data in a practical context, it becomes more accessible for students, helping them to see how it can be used in real-world situations.  B: For me, data storytelling makes it easier to folldigesow complex concepts. When data is presented in a way that connects to real-world scenarios, it becomes simpler to understand and more engaging, which enhances the learning experience. |
| How do you feel about the data stories given? | A: I found the storyline quite engaging and easy to follow, likely because the content was relevant to us as students. The topics were relatable, which made me interested in seeing what the data revealed. I also appreciated the explanations provided after each story, as they helped clarify the information.  B: For the first story, I found it a bit tricky to grasp initially. I prefer content that is concise and straightforward, so I struggled with it at first. However, the subsequent stories (2, 3, and 4) were much clearer and more enjoyable. I was able to follow along better with those. |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | A: Personally, I prefer content that is straightforward. I like to first get familiar with the material through traditional notes, and if I still don't understand something, I can then refer to stories or contextual examples to help clarify the concepts. I think stories can be useful as supplementary material, especially when there are concepts that are difficult to grasp, but I still value direct and clear explanations.  B: I believe that while visualizations and explanations in stories can help deepen my understanding, I wouldn’t say they make me completely confident in mastering the topic. I think the traditional format has its strengths, especially when it includes strong explanations for why certain methods were chosen. While stories can offer valuable insights, I still find conventional lecture notes important, particularly when they provide detailed justifications behind the concepts being taught. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | A: Possibly, but some parts of the stories were tricky, even for me, and I was already familiar with the content. For example, when explaining why certain elements, like background color or correlation coefficients, are used, I found it challenging to grasp the reasoning behind them. It might be helpful to include more examples or comparisons to make these concepts clearer, especially for someone unfamiliar with them.  B: I think it could be beneficial to include more situations or examples that illustrate the concepts, which would make the explanations easier to understand. While I understood some parts, I felt that certain visualizations and concepts could be better explained, especially for those who aren’t already familiar with the material. Adding more context and accessible explanations could improve |
| Do you have any suggestions for the design or composition of DS? | A: I found the design of the data stories to be quite easy to understand, and the stories were relatable. However, I think it might not be the best approach to have big chunks of text. It would be helpful to emphasize key points with paragraphs or other formatting techniques, making the content more digestible.  B: I think the explanations within the data stories were integrated smoothly into the storyline, which helped maintain a clear flow. However, it’s important to ensure that the explanations are detailed enough to help readers keep up with the content. For me, the way the explanations were presented was very well done, but it’s crucial to maintain that clarity throughout. |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | A: Yes, I've used Generative AI tools, mainly for tasks like summarizing content and helping me make my writing more formal, especially using ChatGPT. However, most of my usage has been for simple tasks like sentence generation. When it comes to using these tools for learning data science, I think they could be useful for interpreting data or helping students understand concepts by summarizing complex information. However, I believe the current tools are still somewhat limited and might not be fully reliable for more advanced or precise tasks.  B: terms of applying these tools to data science education, they could be helpful in breaking down complex data or summarizing difficult concepts to make them more accessible to students. However, I wouldn’t say these tools are completely accurate or reliable yet; there's still a lot of room for improvement, especially for more advanced data science applications |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | A: Yes, these skills are still as important because there isn’t an alternative yet. Since AI isn’t advanced enough to replace these skills, learning them yourself remains crucial. The relevance of these skills hasn’t diminished.  B: I believe these skills are just as important as before, especially in the context of data storytelling. There are many different approaches, and we still need people with diverse mindsets and unique ways of thinking to tackle these challenges. Even with AI, human input remains vital. |

**DISCUSSION GROUP SEVEN (JULY 30TH)**

|  |  |
| --- | --- |
| Guided questions | Responses |
| When you were learning data science related courses at NUS/when you will be learning data science related courses, what difficulties did you face/do you think that you will  face? | A: I didn't have any coding background, so when we first started using R Studio, I was very confused. Because of this, I mainly prefer to using Excel for my coursework.  B: In the course 8101, one major difficulty we all faced was answering pause questions. The questions required deep thinking, and I often found it hard to understand the thought process behind them. We ended up focusing more on just finishing the questions rather than truly grasping the concepts. Additionally, when working with code, many of us struggled because it was challenging to memorize everything, so we often resorted to copying and pasting the code provided during class.  C: In G1000, the main challenge I faced was with the campus quizzes. The multiple-choice questions seemed straightforward at first, but the statements often sounded very similar, making them tricky to differentiate. This made it hard to pinpoint the correct answers, especially early on.  D: For G1000, I found the content relatively easy to understand, but applying it was much harder. This was particularly evident during standardized exams, where the preparation wasn’t ideal. We often had to rely on tutors and peers, which wasn’t the best for exam preparation. In 1704, a business school course with basic coding, the challenge was that we weren't really involved in the coding itself. We just observed different functions, and the professor’s approach didn’t fully prepare us for the exams, leading to difficulties in handling the exam questions. |
| After doing these four activities, can you briefly explain to me what data storytelling is? | A: Data storytelling is like a narrative that explains and teaches new concepts. It reminds me of the CHS common curriculum in humanities, where scenarios with different characters are used to help us understand and answer questions. Essentially, it uses historical or contextual stories to convey new information.  B: It's about using narrative tracks to present data and, as mentioned earlier, using a story to demonstrate or prove a point. The story helps to convey what you want to communicate through the data.  C: Another aspect of data storytelling is making the content more relatable, which helps the audience understand the concepts better.  D: Data storytelling involves crafting a narrative with characters and, importantly, connecting the story to your research objectives. Instead of just explaining a concept like regression, you're applying it to a specific situation, making it more intuitive for the reader to see how the concept can be applied in real life. |
| How do you feel about the data stories given? | A: The data stories provided are quite engaging and directly relevant to subjects like climate science and management, which are of great interest to us. The content feels pertinent and resonates with our current studies, making it more relatable. However, the format of the presentation, especially the use of surveys, poses some challenges. Since we aren't always sure whether the choices we make are correct or not, it can hinder further learning. For example, if I misinterpret something in the first story, I may carry that misunderstanding throughout, potentially affecting my confidence and comprehension of subsequent material.  B: I found the story to be somewhat confusing. The inclusion of terms like regression, coefficient, and correlation required me to spend more time reading and trying to understand the content. It felt a bit overwhelming, and I needed to go back multiple times to grasp the concepts fully.  C: The content was interesting, but I found it a bit challenging to understand. Some concepts felt like they assumed prior knowledge, which made it harder for me to follow. For example, terms like "average score" were easier to grasp, but other concepts felt more complex. It would be helpful if the information was presented more directly, without unnecessary details or overly long sentences. Simplifying the data structure and making the narrative more straightforward would improve understanding and make the story more accessible.  D: I found the data storytelling somewhat difficult to follow due to the varying formats and complex details, especially in the sections related to computer-related topics. At times, it felt overwhelming and confusing, making it hard to confirm the accuracy of the data and background information. I think a simpler, more straightforward approach would make it easier to focus on the key points without getting lost in unnecessary complexity. |
| Can you tell us which one you prefer in your data science learning, conventional data science lecture notes or data stories such as the ones you just read? | A: I prefer a mix of both. I appreciate lecture-style notes for their straightforwardness, but incorporating case studies or stories can enhance understanding by illustrating concepts in context. While stories can be useful, they shouldn’t replace lecture notes entirely because they sometimes require you to infer information rather than providing clear instructions.  B: I prefer conventional lecture notes. They are more straightforward and easier to use for exam preparation, providing clear and concise information that is beneficial for studying and understanding key concepts.  C: I also prefer a mix. Lecture notes are essential for effective exam preparation and understanding basic concepts. Data stories are useful for illustrating complex ideas and answering long-form questions in research projects. A combination of both helps in answering and explaining questions more effectively.  D: I lean towards a mix as well. Data stories can be easier to visualize and less time-consuming when learning, as they provide practical examples. However, I believe they should supplement rather than replace conventional notes, which are still crucial for comprehensive learning. Combining both approaches can enhance understanding and learning effectiveness. |
| In the data stories given, there are some concepts, visualizations, and explanations. Did you understand them all? Did you know them prior to this research? For someone who didn’t know these before, do you think reading stories such as these can help them understand these knowledge better? | A: I mostly understood the concepts, although I had some trouble with the charts in one of the stories and needed to read the description several times. While I got a general idea of the concepts, I still lack a concrete understanding of the theories or definitions. For someone with no prior background, the stories might provide some help, but they might not fully grasp the underlying theories.  B: The stories explained the concepts, but they didn't show how to apply them in tools like R Studio, which is crucial. For those unfamiliar with the concepts like correlation coefficients or associations between variables, the stories may provide some understanding but won't help in practical application or coding within R Studio. Thus, it might not be sufficient for learning how to apply these concepts in practice.  C: It would be helpful to include explanations for all technical terms, similar to how 'swatting' was explained. Providing definitions or brackets beside terms would aid in understanding and applying the concepts better.  D: The stories give a general understanding of the basics, but additional materials or appendices are necessary for a deeper grasp of the concepts and methods used. Understanding how to achieve results with different methods is as important as memorizing the results themselves. Thus, supplementary materials are needed for a more comprehensive understanding. |
| Do you have any suggestions for the design or composition of DS? | A: The stories might benefit from being shorter and more concise.  B: Consider organizing the content into a more accessible format, perhaps by using articles or summaries from previous years. Providing a condensed background would make the material easier to understand.  Cl Using slides could be effective, with each slide focusing on a main point, research objective, or challenge. This format allows for easier navigation and understanding of specific parts of the content.  D: Some data stories are too lengthy. Using slides to visualize the key points and breaking down the content into bullet points rather than full sentences could make the material more digestible. |
| Have you used any Generative AI tools? How do you think these tools can be used in the teaching or learning of data science? | Use of GenAI  A: Generative AI is useful for composing emails and generating initial ideas. It helps in obtaining broad, generic input before refining content, which aids in starting tasks and brainstorming.  B: Generative AI supports brainstorming and creating names for new products. It efficiently generates images and avoids the need for manual creation, saving time while producing original designs.  C: Generative AI is beneficial for brainstorming and generating multiple solutions in group settings. It enhances the creative process by producing a variety of ideas and solutions quickly.  D: Generative AI is used for generating diverse solutions and ideas, especially in group activities. It helps in conceptualizing and developing new approaches efficiently.  GenAI for Data science courses:  A: Generative AI can be helpful in generating research ideas and study types for data science. However, it is less effective for interpreting data.  B: Generative AI is valuable for creating new ideas and developing FAQ documents for teaching purposes. It helps ensure that essential concepts are covered and facilitates better understanding.  C: Generative AI does not fully meet expectations for advanced coding tasks in data science. It is better suited for brainstorming rather than solving complex coding problems.  D: Generative AI provides extensive answers but is not yet effective for detailed coding assistance. It works better for generating initial ideas rather than advanced technical solutions. |
| In face of the new challenges such as Generative AI tools? Compared to the past, do you think these skills are still as important as before? | A: Data Skills are still important, though their nature may have evolved. While past methods such as computing standard deviations manually are now facilitated by tools like Excel, interpreting and applying recommendations remains essential. The focus has shifted, but fundamental skills are still relevant and necessary.  B: Data Skills have become increasingly significant due to higher expectations in the current environment. With abundant resources and free tools available, future employers might expect more advanced skills in data analysis. The emphasis is on adapting to modern practices and integrating new technologies into traditional skillsets.  C: Skills are becoming more crucial as data analysis and interpretation are integral to business operations. With the growing volume of data, the ability to analyze and derive insights is increasingly valuable. This trend suggests that data science skills will continue to be essential and may become more prominent in the future.  D: Skills are more important now than in the past. The increasing amount of data and its significance in decision-making underscore the necessity of strong analytical skills. These skills have gained prominence and are expected to remain vital as the field of data science evolves. |