Powerpoint Presentation Project Brief

AI UX & Data Visualisation Design Principles (CA6002)

1. Introduction

The use of data visualisation is an integral part of the process of developing AI solutions. This often begin with the visual exploration of the dataset characteristics. Then the use of appropriate visualisation techniques to design and evaluate the AI algorithms and performance characteristics of the resulting model. Finally, to complete the entire workflow, there is often the need to make a presentation of the insights and results derived from the dataset used in the exercise.

The process of effective data visualisation design is a confluence of many different disciplines in both the sciences and the arts. It is technical because it requires knowledge to interpret, manipulate and transform digital data into appropriate visual forms. At the same time, it has elements of the visual arts and psychology as it is to the human mind that we need to visually and effectively communicate the meaning behind this data. Evidence of mastery in data visualisation will therefore require the demonstration of one's ability to bring all these different knowledge and skills together to discover and then tell a compelling visual story behind the dataset used in the intelligent inference process. This group assignment will provide you a creative opportunity to put into practice the skills and knowledge you have acquired in the data visualisation (Part 1) aspects of CA6002.

2. The Domain

Demonstrating your data visualisation skills requires a context or domain in which you can first find a suitable dataset. With it, you can then investigate, explore, design and develop appropriate data visualisation and AI algorithms to discover insights embedded within the data. This will then allow you to tell an effective visual story of the interesting insights you have discovered. **You are free to choose your own dataset** for this purpose. It is recommended that you pick a domain that is of the collective interest of the group and within the expertise of some members, as the interpretation of insights within the dataset requires some degree of domain knowledge (e.g. if you have little interest or understanding of songs, you may not want to look at datasets related to songs). There are many online sources and repository where you can freely download datasets. (See **Appendix B** for some suggestions).

3. Powerpoint Presentation

Presentation Coverage – Your slide deck should at least include discussion of the following stages of your investigative process (as shown in **Figure 1**):

1. **Introduction** - Project objectives, motivation, process overview, etc.
2. **Exploration of Dataset** – Brief description of dataset, characteristic affecting implementation (i.e. its strengths and weaknesses of dataset characteristics), feature engineering (selection), conditioning done to dataset (if any), etc
3. **Design of AI Algorithm** - Selection of learning algorithm based on dataset and objectives. The design of the AI algorithm, including its parameter selection, parameter fine tuning, learning curve, explainability visualisation (if applicable), etc
4. **Model Evaluation and Performance** - Present visualisation that highlights the model evaluation and performance. Highlight what were the factors that had significant influence on its performance and the reasons for these observations.
5. **Visual Storytelling** - Present the findings and insights from what you have gleaned from the analysis of your dataset. Treat this segment as if you were making a presentation to your bosses on what the data that you have evaluated is telling the company.

6. **Conclusions** - Quick summary of the contributions of the group in the assignment, especially highlighting what was novel and original about the contributions. Avoid just repeating what have already been covered earlier.

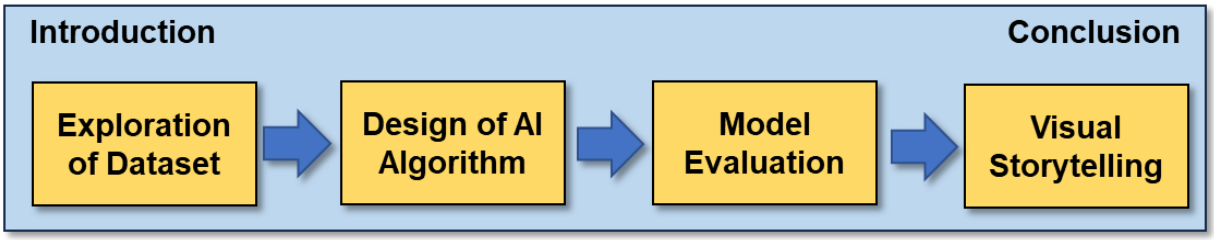

Figure 1 – The main topics to be covered in the Powerpoint presentation slides

Slide Notes – It is important that you embed in each of your slides appropriate and clear notes describing the visual content of your presentation and its design rationale. These notes will be marked for their correctness and clarity. You can consider designing these notes as if you were using it to make your presentation of the slides. However, you do need to keep these notes succinct and avoid exceeding one page for each slide (if possible). The font size of the notes should not go below 10 points as it becomes difficult to read beyond that. Your oral presentation (if assignment is selected) is likely to be a more brief and selective version of what you will write in your slide notes as there is a strict 15-minute group presentation time limit.

Slide count limit and responsibility – You must limit your slide count to no more than **20 slides**. You are free to decide among yourselves who gets allocated how many slides to work on. The bottom of each slide must carry the name of the person(s) responsible for its content and design (use the attached *CA6002 Assignment_Group no – Template.pptx* file). Please be fair to all group members.

Content focus - The contents of this assessment are mostly about your data visualisation techniques. For this reason, slides that are excessively wordy will not score well. Your goal in this assignment is to use the development of an AI solution to demonstrate your competency in exploratory and explanatory data visualisation using a succinct form of Powerpoint presentation.

Originality and novelty - The dataset analysis, AI algorithm selection & implementation, data visualisation design & implementation and presentation design must be done by you. **Plagiarism** or submission of content not done by you is considered **a serious violation** and can severely impact the group's assessment. You may glean ideas from the work of others but do note that originality and novelty is a large component of the assessment criteria. Do come up with some new ideas and approaches to the problem you have selected. You may seek the aid of Generative AI tools in generating the Python codes but such used should be appropriately acknowledged in your slide notes.

Exploit the medium - You should exploit the features available to you on Powerpoint, which include colour, animation, sequencing, transition, text and figure annotation, etc to demonstrate your ability to create, interact, sequence and present visual information derived from your analysis and visualisation of the dataset and AI algorithm. You are graded on the effectiveness of your visual storytelling.

The audience - In your presentation, you should be mindful of your audience and provide sufficient background information to help them understand the message you are attempting to convey.

Assessment criteria – The assessment criteria for the Powerpoint presentation slides are given in **Appendix A**. Read through this list of criteria carefully so you know exactly what to address when creating your visualisation for the Powerpoint slides.

Uploading your assignment – On behalf of the group, the assigned group leader can submit a copy of the team's assignment on the NTULearn Blackboard UCV portal under the **Assignment** folder or go directly to the **Gradebook**. The title of the assignment is "**Powerpoint Group Assignment**". This

Youtube video (https://www.youtube.com/watch?v=tDOkc_6c5h0) gives more details on the submission process. You can make multiple submissions (up to 3 times) but only the last submission just after the deadline will be the one assessed. However, I strongly encourage you to upload your assignment only when you are satisfied that it is your final version. Avoid making multiple submission as sometimes the earlier version may be downloaded for assessment.

4. Peer Evaluation

Peer Evaluation Rubrics – At the end of the assignment, it is necessary for all students to assess the contributions of each team member based on:

1. **Teamwork**: demonstrating proactiveness in collaborating with team members and respect for each other.
2. **Quantity of work**: demonstrating fair share in the overall workload throughout the team project.
3. **Quality of work**: contributing ideas and research efforts that enhance the overall quality of the team's output.

Scoring – Peer evaluation exercise is **confidential** and will be carried out after the assignment submission deadline is over. All students must complete their peer evaluation before stipulated deadline. Failing which, they will receive **zero marks** for their peer evaluation component, regardless of what scores they received from their peers. Each student will give a rating scale between 1 to 5 on the above three assessment criteria for each of his or her group members. Please carefully note the rating rubric below. Do note that a score of "1" is not the best rating, but the worst.

Rating Scale	
5	Strongly agree (*best rating*)
4	Agree
3	Neutral
2	Disagree
1	Strongly disagree (*lowest rating*)

Managing Your Team – Setup regular meeting (face to face or online) to monitor assignment progress. If team members are not contributing to their allocated assigned task and responsibilities, do provide timely feedback to them. Encourage them to contribute and remind them that their individual marks in the course can be impacted by a poor peer evaluation rating from the teammates.

5. Deliverables and Deadlines

The table below outlines the various deliverables for Powerpoint Group Assignment and their respective weightage.

Deliverables	Submission Mode	Weightage	Deadline
Powerpoint Presentation – Completed Powerpoint presentation (≤ 20 slides) will figures & well-narrated notes. Note1: Append group number to filename of your pptx file. Note2: Slides beyond the 20-slide limit will not be assessed	Upload to NTULearn Assignment portal	95%	Before Monday, **Week #5** (2 Feb@12 noon)
Peer Evaluation – Completed peer assessment for each of your teammates using the peer review system on NTULearn, which should appear in the **Assignment** folder.	NTULean peer review system	5%	Before Saturday, **Week #6** (14 Feb@12 noon)

Appendix A – Assessment Criteria for Powerpoint Presentation

Presentation - Assessment Criteria	Weightage
Appropriateness – Uses of most appropriate data visualisation techniques and plots at each stage of the process	25%
Correctness & Clarity - Work done at each stage of the process are technically correct. Including analysis of dataset, selection of learning algorithms, evaluation of model and results, insights derived from investigation and the way these insights were communicated.	25%
Apply Human Visual Perception Principles – Use appropriate human visual perception and psychological principles to make data visualisation more accessible and effective.	25%
Novelty and Originality – Originality and novelty of problem addressed, dataset used, visualisation techniques, use of AI techniques and visual storytelling design.	25%
Total	100%

Appendix B – Some possible online sources for datasets

[1] Kaggle - huge repository of community published data & code - https://www.kaggle.com/datasets

[2] Data.world open datasets - https://data.world/datasets/open-data

[3] Singapore Statistics - https://www.singstat.gov.sg/

[4] Singapore Government Published Dataset - https://data.gov.sg/

[5] Singapore Geo Data dataset - https://data.gov.sg/dataset/national-map-line

[6] US COVID-19 Datasets - https://data.cdc.gov/browse?limitTo=datasets

[7] COVID-19 data - https://github.com/owid/covid-19-data/tree/master/public/data

[8] Our World in Data - Coronavirus Source Data - https://ourworldindata.org/coronavirus-source-data

[9] UCI Machine Learning Repository - https://archive.ics.uci.edu/

[10] Asian Development Bank (ADB) dataset - https://data.adb.org/search/content/type/dataset

[11] DataHub.io Stock Market Data - https://datahub.io/collections/stock-market-data

[12] Tableau - Free Public Data Sets for Analysis - https://www.tableau.com/learn/articles/free-public-data-sets

[13] Dataquest - 21 Places to Find Free Datasets for Data Science Projects - https://www.dataquest.io/blog/free-datasets-for-projects/

[14] Atlassian - Where to Find Free Datasets & How to Know if They're Good Quality - https://www.atlassian.com/data/business-intelligence/free-datasets

[15] Google's Dataset Search - https://datasetsearch.research.google.com/

Note: Some sites will require you to sign up as a member to gain access.