SINGAPORE UNIVERSITY OF SOCIAL SCIENCES

ANL501

End-of-Course Assessment - July Semester 2023

Data Visualisation and Storytelling

INSTRUCTIONS TO STUDENTS:

1. This End-of-Course Assessment paper comprises **7** pages (including the cover page).

2. You are to include the following particulars in your submission: Course Code, Title of the ECA, SUSS PI No., Your Name, and Submission Date.

3. Late submission will be subjected to the marks deduction scheme. Please refer to the Student Handbook for details.

IMPORTANT NOTE

ECA Submission Deadline: Tuesday, 26 September 2023 12:00 pm

ECA Submission Guidelines

Please follow the submission instructions stated below:

A - What Must Be Submitted

You are required to submit the following ONE (1) item for marking and grading:

- *A Report.*

Please verify your submissions after you have submitted the above ONE (1) item.

B - Submission Deadline

- *The ONE (1) item of Report is to be submitted **by 12 noon** on the submission deadline.*
- *You are allowed multiple submissions till the cut-off date for the ONE (1) item.*
- *Late submission of the ONE (1) item **will be subjected to mark-deduction scheme** by the University. Please refer to Section 5.2 Para 2.4 of the Student Handbook.*

C - How the ONE (1) Item Should Be Submitted

- *The Report: submit online to Canvas via TurnItIn (for plagiarism detection)*
- *Avoid using a public WiFi connection for submitting large files. If you are using public wireless (WiFi) connection (e.g. SG Wireless at public areas), you might encounter a break in the connection when sending large files.*

D - Additional guidelines on file formatting are given as follows:

1. Report	• *Please ensure that your Microsoft Word document is generated by Microsoft Word 2016 or higher.* • *The report must be saved in **.docx** format.*

E - Please be Aware of the Following:

Submission in hardcopy or any other means not given in the above guidelines will not be accepted. You do not need to submit any other forms or cover sheets (e.g. form ET3) with your ECA.

*You are reminded that electronic transmission is not immediate. The network traffic may be particularly heavy on the date of submission deadline and connections to the system cannot be guaranteed. Hence, you are advised to submit your work early. **Canvas will allow you to submit your work late but your work will be subjected to the mark-deduction scheme**. You should therefore not jeopardise your course result by submitting your ECA at the last minute.*

It is your responsibility to check and ensure that your files are successfully submitted to Canvas.

F - Plagiarism and Collusion

Plagiarism and collusion are forms of cheating and are not acceptable in any form in a student's work, including this ECA. Plagiarism and collusion are taking work done by others or work done together with others respectively and passing it off as your own. You can avoid plagiarism by giving appropriate references when you use other people's ideas, words or pictures (including diagrams). Refer to the APA Manual if you need reminding about quoting and referencing. You can avoid collusion by ensuring that your submission is based on your own individual effort.

The electronic submission of your ECA will be screened by plagiarism detection software. For more information about plagiarism and collusion, you should refer to the Student Handbook (Section 5.2.1.3). You are reminded that SUSS takes a tough stance against plagiarism or collusion. Serious cases will normally result in the student being referred to SUSS's Student Disciplinary Group. For other cases, significant mark penalties or expulsion from the course will be imposed.

G - Use of Generative AI Tools (Allowed)

The use of generative AI tools is allowed for this assignment.

- *You are expected to provide proper attribution if you use generative AI tools while completing the assignment, including appropriate and discipline-specific citation, a table detailing the name of the AI tool used, the approach to using the tool (e.g. what prompts were used), the full output provided by the tool, and which part of the output was adapted for the assignment;*

- *To take note of section 3, paragraph 3.2 and section 5.2, paragraph 2A.1 (Viva Voce) of the Student Handbook;*

- *The University has the right to exercise the viva voce option to determine the authorship of a student's submission should there be reasonable grounds to suspect that the submission may not be fully the student's own work.*

- *For more details on academic integrity and guidance on responsible use of generative AI tools in assignments, please refer to the TLC website for more details;*

- *The University will continue to review the use of generative AI tools based on feedback and in light of developments in AI and related technologies.*

Please read the following instructions carefully.

Please submit your responses to all questions as a single MS Word document (in .docx extension) and include your RMarkdown script at the end of the document. You may embed your RMarkdown script as an object in your MS Word document. This is the preferred way for you to include your source code. Alternatively, you may copy-and-paste your RMarkdown codes using the "Courier New" font at the end of your document as an appendix.

Do check that your RMarkdown script can be opened as an embedded object and that your report is reproducible from your codes.

Please ensure that your responses to the questions are contained in the Word document you submit. Marks will **NOT** be awarded if your responses are omitted from your Word document, even if they are contained in your RMarkdown script.

Therefore, in your response, **do** describe what you have done (say, explain the general data cleaning and data wrangling strategy), and if appropriate, include code snippets as well as intermediate and final outputs as support.

Marks will be deducted if your Word document is not generated from RMarkdown, for instance, if your submission is typed entirely using MS Word. However, formatting works done outside of RMarkdown are acceptable.

When developing the data visualisations, please pay attention to 1) their design and visual appeal, i.e. the visualisations should look clean and not cluttered, 2) the insights that can be uncovered by these visualisations, and 3) the quality of the data story being told, i.e. the message conveyed by the visualisations should be clear, catchy, and easily understood.

For all questions below, data preprocessing and data cleaning could be required. Please explore the data, clean them or generate new variables if necessary, before plotting.

The R code chunks in your RMarkdown script should be clearly documented to ensure that the logic and purpose of these codes are clear. Clear documentation is especially important for explaining the steps taken for data preprocessing and data cleaning.

Finally, please remember to cite your sources and include a reference or bibliography at the end of your report using the APA citation format.

(Full marks: 100)

Answer all questions (100 marks)
Answer all questions in this section.

Question 1

Emergency food aid serves as a lifeline for vulnerable populations, providing crucial support during times of crisis to alleviate hunger and malnutrition. However, there is a pressing concern about the unintended consequences of food aid, specifically its potential to be diverted and used to fuel insurgency and conflict. This has sparked significant debate among researchers, policymakers, and humanitarian organisations regarding the relationship between emergency food aid and civil conflict.

On one hand, proponents argue that food aid can contribute to conflict prevention and peacebuilding efforts. By meeting the basic food needs of affected populations, it is believed that food aid can help reduce grievances and tensions, promote stability, and foster social cohesion. Moreover, food aid programs often incorporate elements of community engagement, livelihood support, and capacity building, which can contribute to long-term resilience and reduce the risk of conflicts.

On the other hand, skeptics highlight the potential negative effects of food aid on conflict. They argue that food aid may, unintendedly, fuel conflicts by creating competition, unequal distribution, and dependency among communities. In some cases, food aid has been reported to be diverted by armed groups, leading to the militarisation of aid and the perpetuation of conflicts.

You are a consultant for the Global Peace Research Institute (GPRI), a think-tank focusing on economic and social issues in developing countries. The GPRI is interested in studying civil conflict in Sub-Saharan Africa, which is one of the most conflict-affected regions in the world. SSA countries have received significant amounts of emergency food aid. As a consultant, your goal to is conduct an extensive exploratory data analysis (EDA) using data visualisation to try to understand the overall background of civil conflict and food aid in the region.

You are given the dataset ANL501_FoodAid.xlsx containing two workbooks. The first workbook, named “data”, contains the dataset. The second workbook, named “description”, contains the variable description and the data sources. The dataset contains data on civil conflict, natural disasters, aid (humanitarian aid, non-humanitarian aid, and other aid), and other data on the SSA countries like inflation, GDP per capita, and a measure of the type of regime known as the Polity2 score. Aid, in this context, refers to the amount of aid received by an SSA country from 30 Development Assistant Committee (DAC) Donor Countries.

Using the dataset, you are to compose data rich visualisations to shed light on conflict and food aid in Sub-Saharan Africa. You may construct cross-sectional plots, time series plots, spatial plots, or use visualisations not covered in this course wherever

appropriate or interesting. Your visualisations should ultimately demonstrate whether there is an association between food aid and conflict. You are welcome to draw from other data sources based on your own research on this issue (e.g. use a shape file to construct choropleth maps).

Here are the tasks for this assignment.

(i) Before we begin any data analysis, the first thing we should do is to check the data. For instance, some variables are derived from other variables. Are these values consistent? Are there potential concerns about the quality of certain variables in the dataset?

(ii) Use an API to extract a data series from the World Development Indicators (see https://datahub.io/collections/world-bank#api) that you feel is relevant to the topic. The data series obtained from the API is usually in a JSON format, so be sure to import it into R correctly. Merge this data series with the data provided to you for this assignment.

(iii) Conduct a comprehensive exploratory data analysis to gain a good understanding on conflict and food aid in the SSA region. You should also explore the associations between conflict, aid provision, disasters, economic indicators, institutions, and the variable you extracted in part (ii) by designing the appropriate data visualisations to illustrate these associations. Note that you should pick the appropriate associations to report, and not report all pairwise associations. Therefore, you should explore the dataset carefully first before choosing the key visualisations to show in your report.

You should also pay attention to effective data storytelling with the help of appropriate data visualisation elements (e.g. the use of aesthetics, settings, facets, etc.). As a guide, you may consider answering the following issues:

- What are the patterns or trends of civil conflict in the region?
- What is the distribution of the impact of natural disasters?
- Do natural disasters affect the amount of food aid received?
- Is there a relationship between food aid and conflict?
- Does political institution matter in whether food aid "leads" to more conflict, if such a relationship exist?

(iv) Create your report using an RMarkdown document. All the codes for your data visualisations must be contained in your RMarkdown document. Using your RMarkdown document, generate a report in MS Word not exceeding 3000 words (excluding tables and references), to document your findings. This report will be your end-of-course assessment submission.

Your report should follow the following format:

- Executive Summary

- Introduction
  - Provide an introduction on the topic and objective.
  - Summarize the work you have done and your key findings.
  - Write a brief conclusion.
- Data
  - Discuss the key features of your dataset. For example, how many countries are there? What is the sample period? How many observations are there? Which variables have many missing values?
  - Discuss the work you have done to check for the quality of the dataset.
  - Explain why you have selected a certain data series from the World Development Indicators for your analysis.
  - Explain how you extracted this data series and merged it with the dataset.
- Exploratory Analysis
  - Discuss your insights from the data visualisations you constructed.
- Conclusion
  - Discuss your key findings, and whether it is possible to conclude that emergency food aid causes more conflict in the SSA region.

Note that your data visualisations and the insights you obtained from them will form the main part of your report.

The distribution of the scores is as follows:

- Quality of the report – 45 marks
- Quality of visualisations – 45 marks
- Quality of codes – 10 marks

For the quality of the report, you will be assessed on the quality of your executive summary, whether your report is clearly written with a good logical flow, the insights you obtained, and your conclusions. For the quality of visualisations, you will be assessed on your choice of visualisations for various data types, the clarity of the visualisations in terms of data storytelling, and their overall appearance (e.g. good use aesthetics and pre-attentive attributes). For the quality of codes, you will be assessed on whether your code chunks in RMarkdown are intelligible and cleanly presented. It would be a good idea to use comments to explain your codes and increase their readability.

(100 marks)

----- END OF ECA PAPER -----