Data Science Foundation Assessment (MSCDSA01)

1. General Assessment Guidance

The summative assessment for this module consists of two parts:

(1) a written report with word limit of 3000 accounting for 70% of the marks,

(2) a recorded presentation (max 15 min) with word limit of 1500 accounting for 30% of the marks.

The submission deadline is **7 January 2021, 13:00 pm**.

2. Dataset

You are required to use only the provided dataset available under the Assessment section.

3. Tasks

1. Conduct exploratory data analysis using descriptive statistical tools with Python language.

2. Use data visualisation tools/techniques to present and communicate findings.

3. Prepare a set of slides, record a presentation (max 15 min) presenting analysis process and findings.

4. Deliverables

1. Report named with your student number included with the link to your presentation (1234567. pdf)

2. Python source code named with your student number (1234567.py)

Note 1: *Your code should run without issues on the university machines and produces the reported output. You will not be assessed on the quality/performance of your code.*

Note 2: You are required to submit (1) report with a link to your recorded presentation and (2) Python code for a valid submission.

5. Report Structure

The report should be structured according to the followings:

Table of Contents

1. Introduction

2. Body Sections

3. Findings and Conclusions

Appendices

A: Copy of source code as text (not screen shots)

B: Presentation slides

C: Link to your presentation

References

Note: *All the sections/subsections must be numbered except, Appendices and References. All the figures/tables in the report must have captions (e.g. Figure 1, Table 1), properly referenced and explained in the report.*

6. Assessment Criteria

6.1 Report

	Criterion 1: Analysis & Report (weight 40%)	**Criterion 2: Visualisation (weight 20%)**	**Criterion 3: Presentation (weight 10%)**
80% and above	Analysis is complete and free from errors with application of methods that may be insightful or original. Report: Extremely robust and consistent argument making a convincing argument with evidence of originality. Impressive ability in the use of generated information to support arguments. Python code runs without errors and produces all the reported results.	Excellent choice of data visualisation tools/techniques to communicate to specialist and non-specialist audiences.	1. Excellent structure 2. Excellent presentation 3. Excellent writing skills 4. Precise, full and appropriate references.
70 to 79	Analysis is relevant and free from errors with appropriate application of methods. Report: Extremely robust and consistent argument that convincingly addresses issues including uncertainties and conflicts. Excellent use of generated information support arguments. Python code runs without errors and produces all the reported results.	Very good choice of data visualisation tools/techniques to communicate to specialist and non-specialist audiences.	1. Very good structure 2. Very good presentation 3. Very good writing skills 4. Very good and appropriate references.
60 - 69	Analysis is relevant and mostly free from errors with relevant and effective application of methods. Report: Evidence of an argument that is generally convincing with a good internal consistency and addresses most issues. Very good use of generated information to support arguments. Python code runs without errors and produces most the reported results.	Good choice of data visualisation tools/techniques to communicate to specialist and non-specialist audiences.	All (4): 1) Presentation is good but basic 2) no typos 3) good writing skills 4) referencing complies with the Harvard style

50 - 59	Analysis is relevant, free from significant or critical errors with appropriate application of methods. Report: Evidence of an overall convincing argument but may have weaknesses, gaps or inconsistencies. Clear use of generated information but may have some weaknesses in the integration into the argument. Python code runs without errors and produces some the reported results.	Satisfactory choice of data visualisation tools/techniques to communicate to specialist and non-specialist audiences.	3 out of 4
40 - 49	Analysis is relevant, but contains errors with significant effect, or methods that are applied inappropriately. Report: Evidence of a consistent argument but may have weaknesses, significant gaps or be unconvincing. Clear use of generated information but may not be sufficient to sustain the argument. Python code runs without errors and produces some the reported results.	Poor choice of data visualisation tools/techniques, poor communication to specialist and non-specialist audiences.	2 out of 4
30 - 39	Analysis is irrelevant, incomplete or contains errors which have critical effect, or methods that are applied inappropriately. Report: Lack of consistency or structure in the argument. Serious weaknesses in the integration of generated information. Python code runs with errors and the reported results cannot be produced.	No clear choice of data visualisation tools/techniques, limited communication to specialist and non-specialist audiences.	1 out 4

Below 30	Analysis is irrelevant or it is almost non-existent and also incorrect. Report: Total lack of consistency or structure in the argument. Nil or limited integration of generated information. Python code runs with errors and the reported results cannot be produced.	Lack of data visualisation.	None

6.1 Recorded Presentation

	Criterion 1: **Structure and Coherence (weight 15%)**	**Criterion 2:** **Independence from notes** **(weight 15%)**
80% and above	Extremely coherent; good steady pace in presenting material.	Excellent engagement with audience maintained throughout. Excellent flow of presentation.
70 to 79	Very coherent with very good link with report and findings; easy to follow when presenting material.	Very good engagement with audience maintained throughout without interruption in the flow of presentation.
60 - 69	Very good structure and good link with report and findings; reasonably easy to follow when presenting material.	Good engagement with audience with some interruption in the flow of presentation.
50 - 59	Good structure, but structure not always clear, good link with report and findings; pace not always appropriate when presenting material.	Good engagement with audience with interruptions in the flow of presentation. Some dependence on slides with some sections are read out.
40 - 49	Fair structure with fair link with report and findings; fair pace when presenting material.	Fair engagement with audience with many interruptions in the flow of presentation. Heavy dependence on slides with most sections are read out.

30 - 39	Poor structure and poor link with report and findings; poor pace when presenting material.	Poor engagement with audience, poor flow of presentation. Very Heavy dependence on slides with all sections are read out.
Below 39	Lack of structure and link with report and findings; no apparent order; delivery of material is much too fast or too slow.	Very poor engagement with audience, lack of flow of presentation. Constant dependence on slides; the presentation is mostly read.