

GEOG5129E – Introduction to Data Science with Python for Engineers and Researchers

Final Assessment 1 (70% final grade)

Deadline: Wednesday the 6th July, 2022 4 p.m. (submission via Moodle).

Markers: J.F. Einsle and GTA team Moderator: Josh Einsle

Business Brief/ Case for Support:

Imagine writing a business brief to your current manger (or a future) manager arguing why you need to change an established data analysis procedure in your place of work. Keep in mind processes are hard to change so quantitative metrics demonstrating new insights are key. Leverage the power of visualisations to make your case.

In this proposal you will briefly describe the current problems with how things are done, and how your more data centric approach using Python and the data science techniques from this course will produce a new a better outcome as demonstrated on this pilot data set.

1. Apply the **data science techniques covered in this course** to address a data science problem of your choice. Refer to the course outline and includes (but not limited to) ANOVA / Tukey testing, Linear Regression, Nearest Neighbor Analysis, PCA, Clustering)
2. You need to clearly state the problem, describe the data used, and the methodology selected for analysis.
3. As part of this you need to describe the recommended actions that would be taken as a result or insight gained through this analysis and how it will result in new actions.
4. Sections should include:
 1. Introduction and problem statement
 2. Description of data
 3. Methodology (techniques used)
 4. Results
 5. Discussion and action plan
 6. References (Suggested between 3 and 10)
 - Use IEEE or Nature citation and bibliography style
 - References used to advance argument. (Do not cite just to cite)
5. 1000 –1500 words
6. Report should be submitted as a Word file or PDF. PDF can be generated from the notebook using markdown.
7. The jupyter notebook used for analysis also needs to be submitted

The marking proforma used to grade your Business Brief is reported below (not all categories will be weighted equally):

	A	B	C	D	Poor
Does the structure reflect the outline above?					
Is the language used accessible to a non-specialized audience?					
Is the data problem well formulated?					

Does the technique selected seem appropriate for the task?					
Are the conclusions and actions supported by the results of the analysis?					
Is the literature used to advance the argument?					
How relevant and up to date is the used literature?					

Final Assessment 2 (30% Final Grade)

Deadline: Wednesday the 6th July, 2022 4 p.m. (submission via Moodle).

Markers: J.F. Einsle and GTA team Moderator: Josh Einsle

Voice over slides presentation:

This compliments and extends the written report. Aside from having to document your project and its outcomes you will be tasked with presenting to your manger a short report demonstrating why your data analysis should be adopted.

1. Produce a presentation summarising your data project and the outcome/s of your analysis.
2. The presentation needs to serve as an oral defence of the data, methods and conclusions reached in the written report.
3. Equivalent of a 500-word essay
4. 5-minute / ~5 slides

The marking proforma used to grade your Business Brief is reported below:

	A	B	C	D	Poor
Does the presentation clearly state the problem being addressed?					
Is the language used accessible to a non-specialized audience?					
Is the data problem well formulated?					
Does the technique selected seem appropriate for the task?					
Are the conclusions and actions supported by the results of the analysis?					
Are figures and slide captions used to advance the argument?					