Guidelines for submitting the report

These pages should NOT be submitted with your report.

Upload software and design documents

You should make two file submissions as follows:

- A completed report answering the given questions as a single file in either .docx or .pdf format. This should **NOT** be included in the zipped file and should not exceed any specified word/page counts.
- A single zipped file containing your program or programs. If a database been used you should produce a file dump of the table structure to include here. This should NOT contain the original data set.

NOTE: Failure to submit the programme/s will result in a zero grade.

The report must adhere to the word count limits stated for EACH question and the page limit for the appendices. All code and images should be presented in the appendices.

Report contents

The report consists of three sections, each containing specific questions targeting what needs to be address. It should be clear in your report which section and question you are answering. In each section is a maximum word count and pages in the appendices (where applicable) and this should be adhered to. All images and diagrams should be of a size that enables them to be clearly viewed. You should refer to the assessment outline further formatting and submission guidance.

The sections contribute to the module learning outcomes as follows:

Section 1: Theory supported by code samples (50%, 1400 words plus code samples)

Evidence for learning outcome: *Demonstrate critical understanding of the theory and application of advanced programming techniques*; *Design and implement programs for real world problems*.

Section 2: Design decisions supported by code samples (40% 1200 words plus code samples)

Evidence for learning outcome: *Communicate design decisions for the selection, storage and manipulation of data; Design and implement programs for real world problems*.

Section 3: Reflection on the ethics, moral and legal aspects (10% 400 words)

Evidence for learning outcome: *Critically evaluate the legal and ethical impact of software developments within real world contexts.*

Report references

Provide a correctly structured list of references to all the resources used for this development and report. Your responses should be appropriately supported by references to the literature and relevant resources using the computer science department's referencing standard.

Report Outline: Data Analysis Tool

DO NOT PUT ANY PERSONALLY IDENTIFY INFORMATION ON YOUR DOCUMENT

Section 1: Theory supported by code samples (50%, 1400 words plus code samples)

a) *[20 marks]* Identify **one** part of your program design that could be redesigned to use **Python Threads**. Discuss what specific issues need to be considered given the overall design of your program, clearly demonstrate how data and/or communications will be passed between threads, and which Python constructs would support this. You should provide an overview diagram of your existing program flow and clearly identify where concurrency could be of potential benefit.
(500 words maximum, 1 page maximum of supporting code examples, and 2 page diagrams)

b) *[20 marks]* With specific refence to GUI interface constructs and best practice regarding interface layouts, discuss how your GUI supports the required user interactions, given the client's requirements. You should clearly identify the different interactions required by the user, the GUI and code construct/s that support it, and your reasoning for selecting these over other viable options.
(500 words maximum, 1 page maximum of supporting code examples, and 2 pages of wire frame/state designs)

c) *[10 marks]* Java and Python are both high-level programming languages, each providing similar functionality but via different mechanisms and syntax. Having experienced both languages, discuss which you believe to be the most effective in terms of 'speed of development' and give specific Python examples (only) from your program that support your argument.
(400 words maximum)

Section 2: Design decisions supported by code samples (40% 1200 words plus code samples)

a) *[10 marks]* With specific reference to the client's data manipulation requirements, discuss your reasoning for your selected data format (JSON, XML, entity relationship structure), and what advantages/disadvantages it has demonstrated in this context?
(300 words maximum, and 1 page maximum of supporting code examples, 1 page data format diagram)

b) *[10 marks]* With support of appropriate code sample/s discuss how you implemented both parts client's 3rd requirement as stated below. You should use appropriate terminology and clearly identify the code constructs/functions/APIs you have selected and why you selected them over alternatives. You should make reference to any data cleaning needed to get to this stage in the application.

> 3. Produce the mean, mode and median for the **inspection score per year**:
>   a. For each type of vender's seating
>   b. For each 'Zip Code'

(300 words maximum, 1 page maximum of supporting code examples, and 1 page maximum of screen shots)

c) [10 marks] With support of appropriate code sample/s discuss how you implemented appropriate visualisations to meet the clients 4th requirement as stated below. You should clearly identify the APIs you selected and why you choose them over other viable options, ensuring you make clear

reference to the advantages/disadvantages in this context. You should provide screenshots to demonstrate the visualisation/s you selected. You should make reference to any data cleaning need to get to this stage in the application.

4. Produce a suitable graph that displays the number of establishments that have committed each type of violation, you may need to consider how you group this data to make visualisation/s feasible

(300 words maximum, 1 page maximum of supporting code examples, and 1 page maximum of screen shots)

d) *[10 marks]* Using code and screen shots of your output demonstrate how you have determined if there is a significant correlation between the number of violations committed per vendor and their zip code. Clearly state If you consider what you have identified this to be a significant correlation and why you think this is the case. You should make reference to any data cleaning need to get to this stage in the application.
(300 words maximum, 1 page maximum of supporting code examples, and 1 page maximum of screen shots)

5. Determine if there is any significant correlation between the number of violations committed per vendor and their zip code, 'Is there a tendency for facilities in specific locations to have more violations?'. You will need to select an appropriate visualisation to demonstrate this.

Section 3: Reflection on the ethics, moral and legal aspects (10% 400 words)

a) [10 marks] Reflect on the ethics, moral and legal aspects of computing as discussed in week 5 of the module. Drawing on discussion boards and the topics investigated as a source of examples, respond to the following statement (400 words maximum):

"Digital technologies are increasingly penetrating all aspects of the social, civil, and economic domains. These are often based upon algorithms whose designs are protected behind proprietary licences and organisational secrecy (Google's search algorithm, for instance). To protect the security of both the state and the individual, legislation is required to make these algorithms available for public scrutiny."

Report references

Provide a correctly structured list of references to all the resources used for this development and report here.