MAST90106 Data Science Project Pt1: Assessment

Task 1: 10 minute Zoom presentation	5% of assessment in MAST90106 and MAST90107 combined

In this task, your group will present a 10 minute Zoom/online overview of your project. The presentation should describe and motivate your problem, provide a preliminary analysis of your data, plus a summary of the findings to date. You should also describe your plans for the coming semester, identifying key challenges that you are likely to face.

It is expected that ALL groups members will participate in the presentation. You should consider how best to use illustrations, charts, animations or other visual media to help convey your ideas.

Your project supervisor and a subject coordinator will allocate marks to the presentation.

The group presentation is a **hurdle** task. That is, your group must achieve a satisfactory mark on the group presentation if you are to continue with the project in semester 2.

The schedule for the presentations will be published on the LMS in Week 10. At this stage, our plan is for the presentations to take place in Week 11 and/or Week 12.

Task 2: Written report	5% of assessment in MAST90106 and MAST90107 combined

Your group will produce a written report (approximately 2500 words), which describes the problem domain, discusses the related literature and presents a concrete plan for semester 2.

The report will be submitted as a PDF document using the assignment link on the LMS. You should submit the written report after your presentation. The deadline for submission is **Friday June 5, 11:59pm**.

We recommend the use of LaTeX to create the written report. LaTex will allow you to create a professional standard document very easily. It also has a particularly good automatic bibliography processing tools (bibtex). Please use an 11 pt font and 1.5 line spacing. We also recommend that you use Overleaf, a collaborative tool for writing in LaTeX.

The report should contain the following sections (suggested lengths in parentheses):

1. **Introduction (1 page)**: What is your problem? Why is it important? How is the problem difficult from a Data Science perspective? For example, difficulty might stem from sheer size, sparsity, degree of noise, structure, etc.

2. **Related work (2 pages)**: What research from the literature tackles the same or similar problems?

   We recommend that you cite at least 10 different sources. It is important that you identify the key resources/research papers describing models/methods that will be most useful when developing a solution for your problem. You may need to consider if there are any commercial systems that are relevant to your problem, and include these in your review.

   In your description of the related work, you should paint a coherent picture of the background rather than simply saying . . . *'paper X said blah, web site Y said blah'*.

3. **Data (8-10 pages)**: What data do you have? What are its properties? Show some basic analysis of the data, from the size and scale, to ranges of values for response variables and input features, degrees of correlations etc. Outline your data cleaning and processing steps. Feel free to include charts and tables.

4. **Method (2 pages)**: How do you plan to use your data, in terms of models, algorithms, and other techniques for analysis and prediction?

   What techniques from the **Related work** section will you use? How might you adapt them? How will you evaluate success? What hurdles are you likely to face?

   If you have preliminary results, you should include the results in this section.

5. **Timeline (0.5 pages)**: Include a clear project plan (a Gantt chart perhaps). Key tasks and milestones should be listed. The role of each group member should be clearly documented.

The report should be written such that it is accessible to others in your student cohort, as well as to your client. You should include illustrations, tables and graphs.

Given the fact that this is a group task, it is imperative that you proof read your work carefully to eliminate presentation errors.

The group written report is a **hurdle** task. That is, your group must achieve a satisfactory mark on the group written report if you are to continue with the project in semester 2.

Task 3: Peer assessment	0 marks

A peer assessment task – an online quiz/survey – must also be completed as part of the assessment in this subject.

In this task, you will also be asked to rate your team members' contributions to the project as a whole and to the development of the presentation and written report. Outcomes from the quiz/survey will be used when determining your final mark in MAST90106-MAST90107.

Your answers to the peer assessment task will also provide us with important information related to group dynamics and insight into any problems that your team has faced.