

# MATH513

## Big Data and Social Network Visualization

### Practical 2023–24

#### Practical Information

The Practical for MATH513 will take the form of a **presentation**.

**Please read the following notes before working on the presentation:**

- The deadline for this assignment is **Thursday, 7th December, 2023 at 10:00am**. You should submit your work through the MATH513 Big Data and Social Network Visualization DLE site. Your submission will be marked anonymously.
- **This is a group presentation. Please work in self-assigned groups of four people. The groups must be the same as those of the Coursework task.** Details of the Coursework task will be given to you soon.
- Along with this assignment, you need to submit the **minutes of 4 of the meetings you have held to produce the presentation**, where each time a different member of the group is chair and minute-taker. Note that 5 marks are allocated to the submissions of the minutes.
- You should keep notes of all your meetings. Each member of the group will receive the same mark, unless any member chooses to make use of the Peer Assessment option. If you wish to make use of the Peer Assessment option, you will need to contact the Module Leader **Dr Luciana Dalla Valle** by **Wednesday, 6th December, 2023** to make an appointment.
- This presentation counts for 40% of your final mark on this module. Marks will be assigned according to the marking grid on page 5.
- Please note that 5 marks are allocated to the submission of the Ground Rules Contract for your group by the 2nd November, 2023, as discussed in class and as explained on the DLE.
- Your presentation slides should be prepared using R/RStudio and associated packages, and reported using R Markdown.
- You are required to submit a video of your presentation, recorded via Zoom, for example. Zoom allows you to share your screen with your slides, and video-record your presentation. During the presentation, each member of the group should turn their cameras off and use student IDs as display names. For more information about recording videos on Zoom, please click [here](#).
- The presentation must not last more than **10 minutes**. Presentations exceeding the time limit will be penalized.
- The video presentation should be shared equally between the group members.
- Marked assessments will be returned within 20 working days of the submission date. In particular, you will get full feedback on your work by January 9th, 2024.

The relevant MATH513 Assessed Learning Outcomes (ALOs) for this assignment are:

**ALO1** Critically select and use a broad range of techniques to perform Big Data manipulation and visualization;

**ALO2** Perform exploratory analyses to extract information, insight and innovation from data.

**ALO3** Collaborate with others to produce and document R code and to present its professional use for Big Data or Social Network Visualization.

You should keep these ALOs in mind when working on this presentation.

You are reminded of the University's Academic Regulations:

Academic offences occur when activity is undertaken which could confer an unfair advantage to any candidate(s) in assessment. The University recognises the following (including any attempt to carry out the actions described) as academic offences, regardless of intent:

- a. Plagiarism, which is copying or paraphrasing of other people's work or ideas into a submitted assessment without full acknowledgement. More information on plagiarism is available here:  
<https://www.plymouth.ac.uk/student-life/your-studies/essential-information/regulations/plagiarism>
- b. Collusion, which is unauthorised collaboration of students (or others) in producing a submitted assessment. The offence of collusion occurs if a student copies any part of another student's work, or allows their own work to be copied. Collusion also occurs if other people contribute significantly to work that a student submits as their own.

The complete list of regulations can be found here:

<https://www.plymouth.ac.uk/student-life/your-studies/essential-information/regulations>

By submitting this practical assessment, all the group members confirm that they have understood the University's policy on plagiarism and collusion.

## Your Task

Your group has been assigned a dataset, available on the MATH513 DLE, containing data extracted from Twitter (now called “X”).

You need to perform a full analysis of the twitter dataset, drawing useful and informative conclusions from your outputs.

Your analyses should include a range of suitable data visualization techniques and appropriate sentiment analysis.

In your presentation, you also need to discuss the  $t$ -test, explaining whether it would be appropriate for your Twitter dataset, motivating your answer (your answer will also depend on your specific dataset!). If a  $t$ -test is appropriate, you should apply it to your data and present the results clearly, together with brief justification of its use.

Your presentation should include a short critical discussion about what stakeholders can learn from your analyses, together with some strategic suggestions. As an example, your discussion could include whether reactions to new products indicate that they have been well received.

You should prepare slides for your presentation using R Markdown. Please remember to build up your R Markdown presentation step by step, adding a small amount of code at a time. If you include in your R Markdown file a large amount of code, it probably will not work.

Your presentation should report your findings in a well structured and coherent way that is suitable for a general audience.

Your presentation should not exceed **fifteen** slides, including references. Presentations exceeding the slides limit will be penalized. This slide limit should not be considered as a target! Considerable credit will be given to concise, well designed and constructed presentations.

**Note 1:** You are welcome to use/adapt the code of Practical 6 that we studied in class. However, please, note that your dataset might not contain the same/all the variables of the datasets that we analysed in class. Therefore, **you might not be able to implement the same/all the analyses** that we have seen.

**Note 2:** The slides of your presentation should show the main output and major findings of your analysis. Please, **do not include the code in your slides**. You might include a limited part of your code, if this is essential for the presentation of specific results.

## Submission

One member of your group needs to submit the following **seven files** electronically using the online submission facility on the DLE by **Thursday, 7th December, 2023 at 10:00am**:

- An mp4 file containing the video of your presentation, named  
SocialMedia\_First.Second.Third.Fourth.StudentID.mp4  
where you substitute in the Student Identification Numbers of all group members.  
For example,  
SocialMedia\_11034023.12504043.12830176.13643987.mp4
- The pdf file containing your presentation slides, named in a similar way. For example,  
SocialMedia\_11034023.12045043.12830176.13643987.pdf
- The R Markdown file that produces your presentation, named in a similar way. For example,  
SocialMedia\_11034023.12045043.12830176.13643987.Rmd
- The minutes of four of your meetings, named similarly as explained above. For example,
  - Minutes1\_11034023.12045043.12830176.13643987.docx
  - Minutes2\_11034023.12045043.12830176.13643987.docx
  - Minutes3\_11034023.12045043.12830176.13643987.docx
  - Minutes4\_11034023.12045043.12830176.13643987.docx

Please ask if you do not clearly understand these submission instructions.

**Note 1:** You should **not submit any other files** on the DLE (including any images/photos you might have included in your presentation).

**Note 2:** Please **do not submit your files in any zipped folder or sub-folder**, otherwise you will be penalized: 5 marks are allocated to the correct submission of the required files.

## MATH513 Big Data and Social Network Visualization: Practical Marking Grid

Assessment Area	Maximum Mark	Awarded Mark	Feedback
<i>Ground Rules Contract:</i> submission by 2/11/22 at 10:00	5		
Submission of the <i>minutes of four of your meetings</i>	5		
Submission of the correct file formats (no zipped folders)	5		
<i>Structure:</i> clear introduction and conclusions, logical and coherent structure.	20		
<i>Presentation:</i> presentation style conveys interest, high quality images / text, even pace throughout, slides appearance.	20		
<i>Understanding:</i> deep understanding of content, including statistical test, well-explained analysis.	25		
<i>Accuracy:</i> technical correctness, including any analytic or statistical notation and R / R Markdown code.	20		
<b>Total</b>	100		