# **Textual Data**

⚠ Changes before 22/23

- $\Box$  Add Learning Outcomes for each week
- $\Box$  Link to regex/working with strings in Week 2 CASA0005

### **Overview**

! Important

This week's Learning Outcomes are:

- 2.
- 3.

## **Preparation**

#### Lectures

COme to class prepared to present:

Session	Video	Presentation	Notes
Notebooks as	Video	Slides	Notes
Documents			
Patterns in Text	Video	Slides	Notes
Cleaning Text	Video	Slides	Notes
Analysing Text	Video	Slides	Notes

#### Readings

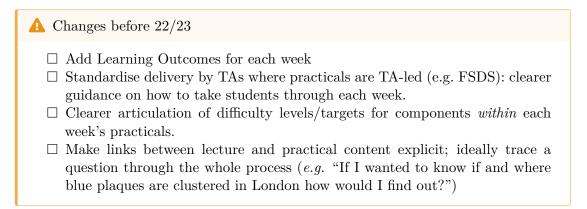
Come to class prepared to present:

- (Ladd 2020) URL
- (Lavin 2019) URL

#### **Activities**

- Padlet: [Collaborative Agenda]
- Complete the short Moodle quiz associated with this week's activities.

#### **Practical**



The practical can be downloaded from GitHub.

#### References

Ladd, John R. 2020. "Understanding and Using Common Similarity Measures for Text Analysis." *The Programming Historian*, no. 9. https://doi.org/10.46430/phen008 9.

Lavin, Matthew J. 2019. "Analyzing Documents with TF-IDF." The Programming Historian, no. 8. https://doi.org/10.46430/phen0082.