

DIGI405 Corpus exploratory analysis assignment

Key information:

Assignment length: Between 1150-1300 words, excluding table text
Due date: Due date for Assignment 1 on Learn
Submit to: 'Assignment 1 – Corpus exploratory analysis' link on Learn

Learning goals:

- Demonstrate understanding of concepts and techniques for obtaining and explaining frequency and co-occurrence evidence about texts
- Explain how text data are connected to social, cultural, and linguistic structures and contexts
- Appropriately link concepts from relevant research domains to interpret text analysis outputs in written form

Instructions:

In this assignment you will use the RNZ news corpus to investigate the representation of climate change in NZ news media. You should start by undertaking some exploratory analysis before deciding how you will answer the questions below.

Part 1

For each of 1 & 2 below, **identify a pattern of language use** that offers some insight into news reporting on climate change by Radio New Zealand. For each of these patterns, **create a table** to present evidence from the corpus, and **write between 350-400 words explaining what the evidence shows us** about that pattern of language use.

1. Frequency-based evidence (ie frequency, keyness)
2. Co-occurrence evidence (ie concordances, collocations/ngrams)

In total, Part 1 will present two tables and associated paragraphs. Each table should have a suitable title and caption that incorporates key information about the methods and/or quantitative measures used. You may consider elements of *content* (the topic of climate change) and *discourse* (the characteristics of news and different social / cultural groups or perspectives reflected in the news).

Part 2

Choose **one** of your answers from Part 1 and **write further paragraphs (450-500 words) to develop your analysis.**

- You should use other types of evidence than you covered for this pattern of language use in Part 1, ie you should supplement frequency-based evidence with co-occurrence evidence and vice-versa.
- You should relate your answer to one piece of relevant prior research. You should compare or contrast your findings, or provide connecting analysis with this research. You should **not** merely summarise other research without relating it to your own work.
- You may include other approaches covered in the course to date, eg parts of speech, dispersion or diversity, if they support or extend your analysis.
- You should draw on analysis strategies discussed in Topic 2.10 "Approaches to corpus analysis" in the online learning materials.

Notes:

- All corpora needed for this assignment are provided on Learn.
- For Part 2, you should only use academic journal articles or books sourced through the UC Library.
- You should use APA referencing style. The details of the referencing style are not a focus here, but you must ensure you provide a clear citation, including page number, whenever you draw on ideas or quotations from an academic article or book.
- You are expected to use methods taught in this course. You can assume your marker is familiar with them, and you are not expected to define or explain them in detail. You should, however, be clear about how you used the methods in your table captions.
- Tables should be numbered and presented using [these guidelines](#). Table captions should state how you have used analysis methods (eg for collocation evidence giving specific details such as node words examined, window size, min. document frequency threshold applied). Be clear and concise.
- Tables, captions and quotations from secondary sources are not included in the word count.
- We expect you to write every part of this assignment yourself. Writing develops your ability to think and analyse text data as well as communicating your ideas. For that reason, use of generative AI tools is not permitted. Your assignment will be checked by Turnitin, which detects machine-generated text.
- Please consider making use of the Academic Skills Centre's [resources on writing](#).