



## Contents

Create the document-term matrix  
5m 29s video

Weighting the document-term matrix  
3m 56s video

Focus the document-term matrix  
4m 13s video

Chapter Quiz  
4 questions

## 6. Apply Statistics to Text

Word and document frequency  
3m 29s video

Hierarchical clustering  
3m 20s video

Associated terms  
3m 55s video

Chapter Quiz  
3 questions

## 7. Sentiment Analysis

What is sentiment analysis?  
2m 40s video

Real-world example of sentiment

## INSTRUCTOR



Mark Niemann-Ross  
Technologist experienced  
in hardware, software,  
and science fiction

+ Follow on  
LinkedIn

## RELATED TO THIS COURSE

Exercise Files · [Show all](#)

Certificates · [Show all](#)

## Course details

2h 32m · Advanced · Released: 5/19/2023

4.9 ★★★★★ (16)

Natural language processing (NLP) is one of the most important components of artificial intelligence. It allows you to process, analyze, and understand large amounts of data in the form of natural language. In this course, instructor Mark Niemann-Ross shows you how to get started implementing NLP algorithms using R, the popular programming language for statistical computing and graphics.

Explore the basics of manipulating matrices and producing statistics, both of which are core to successful NLP. Learn how to use tools and text-mining frameworks such as tm, quanteda, and tidytext, as well as work with corpora, sources, and other types of NLP document metadata. Mark covers the best practices for preprocessing text in preparation for NLP, creating structured data, applying statistics to text, performing sentiment analysis, visualizing datasets, and more.

## Skills covered

