

# [PDF] Text Mining With R: A Tidy Approach

Julia Silge, David Robinson - pdf download free book

O'REILLY Copyrighted Material

## Text Mining with R

Much of the data available today is unstructured and text-heavy, making it challenging for analysts to apply their usual data wrangling and visualization tools. With this practical book, you'll explore text-mining techniques with tidytext, a package that authors Julia Silge and David Robinson developed using the tidy principles behind R packages like ggplot2 and dplyr. You'll learn how tidytext and other tidy tools in R can make text analysis easier and more effective.

The authors demonstrate how treating text as data frames enables you to manipulate, summarize, and visualize characteristics of text. You'll also learn how to integrate natural language processing (NLP) into effective workflows. Practical code examples and data explorations will help you generate real insights from literature, news, and social media.

- Learn how to apply the tidytext format to NLP
- Use sentiment analysis to mine the emotional content of text
- Identify a document's most important terms with frequency measurements
- Explore relationships and connections between words with the ggraph and wdyr packages
- Convert back and forth between R's tidy and non-tidy text formats
- Use topic modeling to classify document collections into natural groups
- Examine case studies that compare Twitter archives, dig into NASA metadata, and analyze thousands of Usenet messages

**Julia Silge** is a data scientist at Stack Overflow, analyzes complex datasets and communicates technical topics to diverse audiences. She has a PhD in astrophysics and loves line art and making beautiful charts.

**David Robinson** is a data scientist at Stack Overflow with a PhD in Quantitative and Computational Biology from Princeton University. He enjoys developing open-source R packages, including broom, gganimate, fuzzytext, and wdyr.

"David and Julia have transformed text analysis in R. With tidytext, it's no longer necessary to learn a complex set of tools; you can now apply everything you already know about manipulating and visualizing data frames."

—Hadley Wickham  
Chief Scientist, RStudio  
Adjunct Professor of Statistics at the  
University of Auckland  
Stanford University  
and UC Berkeley



Twitter: @oreiljmsa  
facebook.com/oreilly

## Books Details:

Title: Text Mining with R: A Tidy Ap

Author: Julia Silge, David Robinson

Released:

Language:

Pages: 194

ISBN: 1491981652

ISBN13: 9781491981658

ASIN: 1491981652

[\*\*CLICK HERE FOR DOWNLOAD\*\*](#)

pdf, mobi, epub, azw, kindle

## Description:

Much of the data available today is unstructured and text-heavy, making it challenging for analysts to apply their usual data wrangling and visualization tools. With this practical book, you'll explore text-mining techniques with tidytext, a package that authors Julia Silge and David Robinson developed using the tidy principles behind R packages like ggplot2 and dplyr. You'll learn how tidytext and other tidy tools in R can make text analysis easier and more effective.

The authors demonstrate how treating text as data frames enables you to manipulate, summarize, and visualize characteristics of text. You'll also learn how to integrate

natural language processing (NLP) into effective workflows. Practical code examples and data explorations will help you generate real insights from literature, news, and social media.

- Learn how to apply the tidy text format to NLP
  - Use sentiment analysis to mine the emotional content of text
  - Identify a document's most important terms with frequency measurements
  - Explore relationships and connections between words with the *ggraph* and *widyr* packages
  - Convert back and forth between R's tidy and non-tidy text formats
  - Use topic modeling to classify document collections into natural groups
  - Examine case studies that compare Twitter archives, dig into NASA metadata, and analyze thousands of Usenet messages
- 

- Title: Text Mining with R: A Tidy Approach
  - Author: Julia Silge, David Robinson
  - Released:
  - Language:
  - Pages: 194
  - ISBN: 1491981652
  - ISBN13: 9781491981658
  - ASIN: 1491981652
-