Package ‘janeaustenr’
June 10, 2017

Title  Jane Austen's Complete Novels
Version  0.1.5
Description  Full texts for Jane Austen's 6 completed novels, ready for text analysis. These novels are "Sense and Sensibility", "Pride and Prejudice", "Mansfield Park", "Emma", "Northanger Abbey", and "Persuasion".

URL  https://github.com/juliasilge/janeaustenr
BugReports  https://github.com/juliasilge/janeaustenr/issues
Depends  R (>= 3.1.2)
Suggests  dplyr, testthat
License  MIT + file LICENSE
LazyData  true
RoxygenNote  6.0.1
Encoding  UTF-8
NeedsCompilation  no
Author  Julia Silge [aut, cre]
Maintainer  Julia Silge <julia.silge@gmail.com>
Repository  CRAN
Date/Publication  2017-06-10 06:29:05 UTC

R topics documented:

austen_books ......................................................... 2
emma ................................................................. 2
janeaustenr .......................................................... 3
mansfieldpark ......................................................... 3
northangerabby ....................................................... 4
persuasion ............................................................. 4
prideprejudice ......................................................... 5
sensesensibility ...................................................... 5

Index 6
austen_books  

*Tidy data frame of Jane Austen’s 6 completed, published novels*

**Description**

Returns a tidy data frame of Jane Austen’s 6 completed, published novels with two columns: `text`, which contains the text of the novels divided into elements of up to about 70 characters each, and `book`, which contains the titles of the novels as a factor in order of publication.

**Usage**

```
austen_books()
```

**Details**

Users should be aware that there are some differences in usage between the novels as made available by Project Gutenberg. For example, "anything" vs. "any thing", "Mr" vs. "Mr.", and using underscores vs. all caps to indicate italics/emphasis.

**Value**

A data frame with two columns: `text` and `book`

**Examples**

```
library(dplyr)

austen_books() %>% group_by(book) %>% 
  summarise(total_lines = n())
```

---

emma  

*The text of Jane Austen’s novel “Emma”*

**Description**

A dataset containing the text of Jane Austen’s 1815 novel “Emma”. The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

**Usage**

```
emma
```
Format
A character vector with 15297 elements

Source
http://www.gutenberg.org/ebooks/158

---

janeaustenr  Jane Austen’s Complete Novels

Description
This package contains the complete text of Jane Austen’s 6 completed, published novels, formatted to be convenient for text analysis.

---

mansfieldpark  The text of Jane Austen’s novel "Mansfield Park"

Description
A dataset containing the text of Jane Austen’s 1814 novel "Mansfield Park". The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

Usage
mansfieldpark

Format
A character vector with 14768 elements

Source
http://www.gutenberg.org/ebooks/141
persuasion

northangerabbey  The text of Jane Austen’s novel “Northanger Abbey”

Description
A dataset containing the text of Jane Austen’s novel "Northanger Abbey", published posthumously in 1818. The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

Usage
northangerabbey

Format
A character vector with 7840 elements

Source
http://www.gutenberg.org/ebooks/121

persuasion  The text of Jane Austen’s novel “Persuasion”

Description
A dataset containing the text of Jane Austen’s novel "Persuasion", published posthumously in 1818. The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

Usage
persuasion

Format
A character vector with 8328 elements

Source
http://www.gutenberg.org/ebooks/105
prideprejudice

The text of Jane Austen’s novel “Pride and Prejudice”

Description
A dataset containing the text of Jane Austen’s 1813 novel “Pride and Prejudice”. The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

Usage
prideprejudice

Format
A character vector with 12447 elements

Source
http://www.gutenberg.org/ebooks/1342

sensesensibility

The text of Jane Austen’s novel “Sense and Sensibility”

Description
A dataset containing the text of Jane Austen’s 1811 novel “Sense and Sensibility”. The UTF-8 plain text was sourced from Project Gutenberg and is divided into elements of up to about 70 characters each. (Some elements are blank.)

Usage
sensesensibility

Format
A character vector with 12262 elements

Source
http://www.gutenberg.org/ebooks/161
Index

*Topic datasets
  emma, 2
  mansfieldpark, 3
  northangerabbey, 4
  persuasion, 4
  prideprejudice, 5
  sensesensibility, 5

austen_books, 2
emmma, 2
janeaustenr, 3
janeaustenr-package (janeaustenr), 3
mansfieldpark, 3
northangerabbey, 4
persuasion, 4
prideprejudice, 5
sensesensibility, 5