Package 'bardr'

March 24, 2021

Type Package

Title Complete Works of William Shakespeare in Tidy Format

Version 0.0.9

Maintainer Zane Billings <wz.billings@gmail.com>

Description Provides R data structures for Shakespeare's complete works,
as provided by Project Gutenberg <https:www.gutenberg.org/ebooks/100>.

License GPL-3

Encoding UTF-8

LazyData true

Depends R (>= 3.6.0)

RoxygenNote 7.1.1

NeedsCompilation no

Author Zane Billings [aut, cre]

Repository CRAN

Date/Publication 2021-03-24 09:20:02 UTC

R topics documented:

---

all_works_df   *Contents of Complete Works of William Shakespeare (dataframe)*

Description

A dataframe containing the full text of all of the complete works of William Shakespeare, as provided by Project Gutenberg.

Usage

```
all_works_df
```

Format

A data frame with 166340 rows and 4 variables:

name short (or common) name of the work

content the full contents of the work. Each line is ~70 characters

full_name the complete name of the work, as listed

genre whether the work is poetry, history, comedy, or tragedy

Source

<http://www.gutenberg.org/files/100/100-0.txt>

Examples

```
works <- bardr::all_works_df
subset(works, works$genre == "History")
```

---

all_works_list   *Contents of Complete Works of William Shakespeare (list)*

Description

A list containing the full text of all of the complete works of William Shakespeare, as provided by Project Gutenberg.

Usage

```
all_works_list
```

Format

A list with 44 elements, each one containing a character vector containing the full text of a work, given in the element name.

Source

http://www.gutenberg.org/files/100/100-0.txt

---

bardr   *bardr: providing the complete works of the Bard in tidy format.*

Description

The bardr package provides R data structures for all of William Shakespeare's works available in the Project Gutenberg ebook. The provided data are designed to seamlessly work in R without the hassle of data wrangling and cleaning, which has already been performed.

Details

Inspired by the janeaustenr package by Julia Silge: see https://github.com/juliasilge/janeaustenr .

Complete collections

The complete works are available all at one time in two separate formats.

One is a named list, where each entry is a named character vector. The name of the vector is the name of the work, and the contents of the vector are lines of the associated text file (all lines are <= 70 characters).

The other is a data frame with a column for the name of the work (repeated as many times as there are lines of content) and a column for the content of the work, where each cell in the content column is one line of text.

Index