

Package ‘TidytextPlot’

December 3, 2020

Title A Package to plot change in proportion of given words over time.

Version 1.0.0.11

Description The package transform a dataframe into a tidy text data structure and produces a plot showing how the proportion of given words changes over time.

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.1.1

Suggests knitr,
rmarkdown

VignetteBuilder knitr

Imports dplyr,
ggplot2,
scales,
tidytext

Depends R (>= 2.10)

R topics documented:

createTidyData	2
debateData	3
plotDebateWords	3
Index	5

createTidyData	Create tidytext data structure
----------------	--------------------------------

Description

This function transforms a data.frame having 3-columns of "person", "dialogue" and "turn" into tidytext data structure. It tokenizes the dialogue, splitting each sentence in separate words and remove stopwords. It also calculates count, and proportion of word(s) within each turn.

Usage

```
createTidyData(input_DF)
```

Arguments

data_df A data.frame containing the data

Value

A data.frame consisting turn, the tokenized dialogue, count and proportion of words within each turn.

turn Number indicating the turn of talk.

word The tokenized dialogue.

n The count of words per turn.

p The proportion of words per turn.

Author(s)

Sewedo Abolade Todowede <todowede@gmail.com>

Examples

```
fict_df<-data.frame(person=c("John","Todd","Texas","Rama"),  
  dialogue=c("We gather here", "meet basic needs","people of this nation", "focus on implementation"),  
  turn=c(1,2,3,4), stringsAsFactors = FALSE)  
createTidyData(fict_df)
```

debateData	<i>2012 US Presidential Debate Data Set</i>
------------	---

Description

The 2012 US Presidential Debate Data Set contains a cleaned version of dialogues for the three main presidential debates

Usage

```
debateData
```

Format

A data.frame with 2912 rows and 3 columns having columns header:

person The speaker. Obama and Romney were the candidates, Cowlwy, Lehrer, Schieffer were moderators and “question” indicates questions raised by the public.

dialogue The words spoken by each person.

turn Progressive number indicating the turn of talk.

Source

Data provided by Dr. Luciana Dalla Valle.

Examples

```
with(debateData, plotDebateWords("iran"))
with(debateData, plotDebateWords(c("iran", "people", "china")))
```

plotDebateWords	<i>Plot of change in proportion of given word(s) over time</i>
-----------------	--

Description

This function creates a ggplot of change in proportion of given word(s) over time, where time is indicated by turn It calls the createTidyData function which prepares the data for plotting.

Usage

```
plotDebateWords(word = NULL, data = debateData)
```

Arguments

word	Single or combination of characters (words) for plotting
data	A data.frame containing the data - Not required if supplied data is used.

Value

A ggplot graphic showing change in proportion of given word(s) over time.

Author(s)

Sewedo Abolade Todowede <todowede@gmail.com>

Examples

```
fict_df<-data.frame(person=c("John","Todd","Texas","Rama"),  
  dialogue=c("We gather here", "meet basic needs","people of this nation", "focus on implementation"),  
  turn=c(1,2,3,4), stringsAsFactors = FALSE)  
plotDebateWords(c("people","focus"),fict_df)
```

Index

*Topic **datasets**

debateData, [3](#)

createTidyData, [2](#)

debateData, [3](#)

plotDebateWords, [3](#)