Package 'speechanalysis'

January 14, 2021

Title A Package to Provide Graphical Summaries of Speech Data

Version 0.0.0.9000

Date 2021-01-14

Description Provides a number of functions which produce graphical summaries of speech data. An example dataset is provided, which consists of 10 speeches made by Donald Trump throughout September 2020.

License GPL-2 | GPL-3

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RogygenNote 7.1.1

Depends R (>= 2.10)

Imports tidyverse, tidytext, lubridate, scales

RoxygenNote 7.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

`plot_tf_idf` *Most Frequent Words per Speech, by tf-idf Index*

Description

Returns a plot of the most frequent words used (in descending order) during a collection of speeches. The plot is faceted by the location of the speech, and uses the tf-idf index to determine word frequencies.

Usage

```
plot_tf_idf(df, n_words = 10)
```

Arguments

df	A dataframe with the following columns: 'speech' <chr>, 'location' <chr>, 'date' <date>
n_words	<int> The number of 'top' values to plot

Value

<ggplot> A plot of the most frequent words used

Author(s)

10570155 <10570155>

Examples

```
plot_tf_idf(trump_speeches, 10)
```

---

trump_speeches *Trump Speech Data*

Description

A dataset consisting of 10 speeches made by Donald Trump throughout September 2020.

Usage

```
trump_speeches
```

Format

A data.frame with 10 rows and 3 columns:

speech Speech transcript

location Location of speech

date Date of speech

Source

University of Plymouth, MATH513 Assignment 2020/21

Examples

```
plot_tf_idf(trump_speeches, 10)
word_frequency(trump_speeches, c("america", "biden", "China"))
zipfs_law(trump_speeches, 1)
```

word_frequency *Graph of word usage of a data set*

Description

Returns a plot of the frequencies of the words provided over time in the dataset.

Usage

```
word_frequency(df, words_to_graph)
```

Arguments

df	A dataframe with the following columns: 'speech' <chr>, 'location' <chr>, 'date' <date>
words_to_graph	A collection of words for which the user wants graphs plotted

Value

<ggplot> A plot of the word frequency over time

Author(s)

10701983 <10701983>

Examples

```
word_frequency(trump_speeches, c("america", "biden", "China"))
```

zipfs_law *Graph of word frequencies against linear regression and against the theoretical Zipf's Law.*

Description

This function provides a graph using ggplt2 of a data set containing speeches.

Usage

```
zipfs_law(df, plot_type)
```

Arguments

df	A dataframe with the following columns: 'speech' <chr>, 'location' <chr>, 'date' <date>
plot_type	A number 1, 2, or 3 which indicates of the type of plot to be displayed. 1 stands for 'frequency plot', 2 for 'frequency + linear', 3 for 'theoretical zipf's law and practical frequencies'

Value

<ggplot> Graph of word frequencies against linear regression and against the theoretical Zipf's Law.

Author(s)

10696253 <10696253>

Examples

```
zipfs_law(trump_speeches, 1)
```

Index