

Package ‘UkraineWordTrends’

January 5, 2024

Type Package

Title Analysis of Ukraine Word Trends

Version 0.1.0

Author 10892938 10764009 10887206 10892109

Maintainer 10892938 <10892938@plymouth.ac.uk>

Description

Analyze and visualize the change in word frequency over time for Ukraine Conflict articles.

License GPL-2 | GPL-1

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Depends R (>= 2.10)

Suggests knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

Combined_Guardian_NYT	1
plot_word_trends	2

Index	4
--------------	----------

Combined_Guardian_NYT	<i>Combined Guardian and New York Times Dataset</i>
-----------------------	---

Description

A combined dataset of articles related to the Ukraine conflict from The Guardian and The New York Times.

Usage

```
data("Combined_Guardian_NYT")
```

Format

A data frame with 407 observations on the following 4 variables.

`published` a POSIXct object representing the publication date and time.

`headlines` a character vector containing the headlines of the articles.

`articles` a character vector containing the full text of the articles.

`journal` a character vector indicating the journal (source) of each article.

Details

This dataset includes information about the publication date, headlines, full articles, and the corresponding journal for each entry.

Source

Original data obtained from The Guardian and The New York Times.

References

The articles were obtained from The Guardian and The New York Times.

Examples

```
data(Combined_Guardian_NYT)
## Explore the structure of the dataset
str(Combined_Guardian_NYT)

## Summary statistics
summary(Combined_Guardian_NYT)

## Plot trends of selected words
plot_word_trends(Combined_Guardian_NYT, c("Ukraine", "conflict", "Russia"))

## Display the first few rows of the dataset
head(Combined_Guardian_NYT)
```

`plot_word_trends`*Plot Change of Word Frequency Over Time*

Description

This function takes a dataset and a vector of words as input and produces a plot showing how the proportion of given words in the articles changes over time.

Usage

```
plot_word_trends(data, words, na.rm = TRUE)
```

Arguments

data	A data frame containing articles with a 'published' column representing the publication date and 'articles' column containing the text of the articles.
selected_words	A character vector specifying the selected words of interest for which the word frequency change over time will be plotted.
na.rm	Should NA be removed? Default value TRUE.

Details

The function tokenizes the articles, removes stopwords, calculate word frequencies overtime, and plots the change in word frequency for the specified words. It also includes a regression line for each word.

Value

A plot illustrating the change of word frequency over time for the specified words in the provided dataset.

Author(s)

10892938 10764009 10887206 10892109

See Also

[ggplot](#), [tidytext](#), [dplyr](#)

Examples

```
data(Combined_Guardian_NYT)
combined_data <- Combined_Guardian_NYT

selected_words <- c("civilians")
plot_word_trends(combined_data, selected_words)

selected_words <- c("country", "putin", "zelensky")
plot_word_trends(combined_data, selected_words)

selected_words <- c("biden", "nato", "russia", "ukraine", "war", "weapons")
plot_word_trends(combined_data, selected_words)
```

Index

* **datasets**

Combined_Guardian_NYT, [1](#)

Combined_Guardian_NYT, [1](#)

dplyr, [3](#)

ggplot, [3](#)

plot_word_trends, [2](#)

tidytext, [3](#)