Package ‘schrute’

March 17, 2020

Title  The Entire Transcript from the Office in Tidy Format
Version  0.2.0
Description  The complete scripts from the American version of the Office television show in tibble format. Use this package to analyze and have fun with text from the best series of all time.
License  MIT + file LICENSE
URL  https://github.com/bradlindblad/schrute
BugReports  https://github.com/bradlindblad/schrute/issues
Depends  R (>= 2.10)
Imports  dplyr, ggplot2, magrittr, stringi, stringr, tibble, tidytext
Suggests  covr, knitr, rmarkdown, testthat (>= 2.1.0)
VignetteBuilder  knitr
Encoding  UTF-8
LazyData  true
RoxygenNote  7.0.2
NeedsCompilation  no
Author  Brad Lindblad [aut, cre], Lucas Graybuck [ctb]
Maintainer  Brad Lindblad <bradley.lindblad@gmail.com>
Repository  CRAN
Date/Publication  2020-03-17 21:20:02 UTC

R topics documented:
schrute .................................................. 2
theoffice .................................................. 2
Index ..................................................... 3
Description

Analyze and have fun with the text from the best series of all time

The entire script transcriptions from The Office

Format

A tibble with 55130 observations of 12 variables:

- **index**: row index
- **season**: season number
- **episode**: episode number
- **episode_name**: name of the episode
- **director**: director of the episode
- **writer**: name(s) of writers of the episode
- **character**: name of the character saying the line
- **text**: words spoken by that actor
- **text_w_direction**: words spoken by that actor with stage direction included
- **imdb_rating**: rating from imdb
- **total_votes**: total votes for episode on imdb
- **air_date**: date the episode originally aired

Source

https://transcripts.foreverdreaming.org
Index

*Topic datasets
  theoffice, 2

schrute, 2

theoffice, 2