iemisctext: Example Analysis using the tm package

Irucka Embry
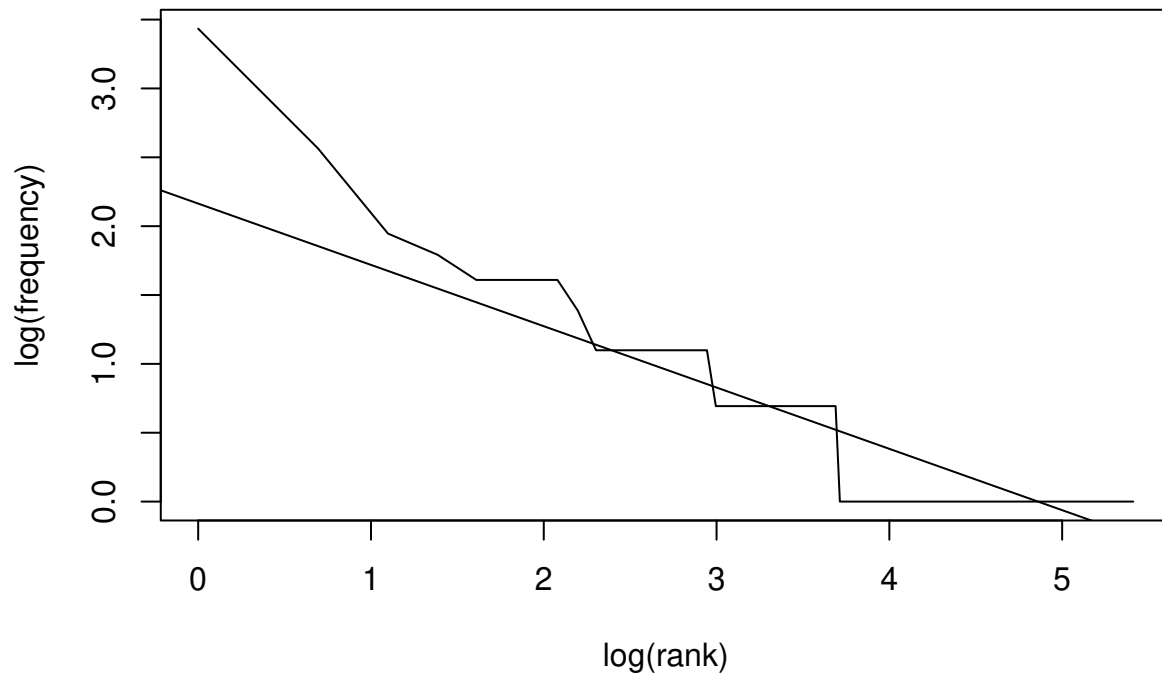
2024-05-26

Contents

Examples

Example 1

```r
install.load::load_package("iemisctext", "tm")
# load needed packages using the load_package function from the install.load
# package (it is assumed that you have already installed these packages)


# appeal_womanhood_world
data(appeal_womanhood_world)

# create a document term matrix of appeal_womanhood_world
a <- DocumentTermMatrix(appeal_womanhood_world)

# find terms used > 5 times
findFreqTerms(a, 5)
```

```
## [1] "and"   "great" "have"  "not"   "that"  "the"   "with"  "women"
```

```r
# Zipf's law
Zipf_plot(a)
```

```
## (Intercept)             x
##   2.1639648  -0.4452613
```

```
# anarchy
data(anarchy)
```

```
# create a document term matrix of anarchy
a <- DocumentTermMatrix(anarchy)
```
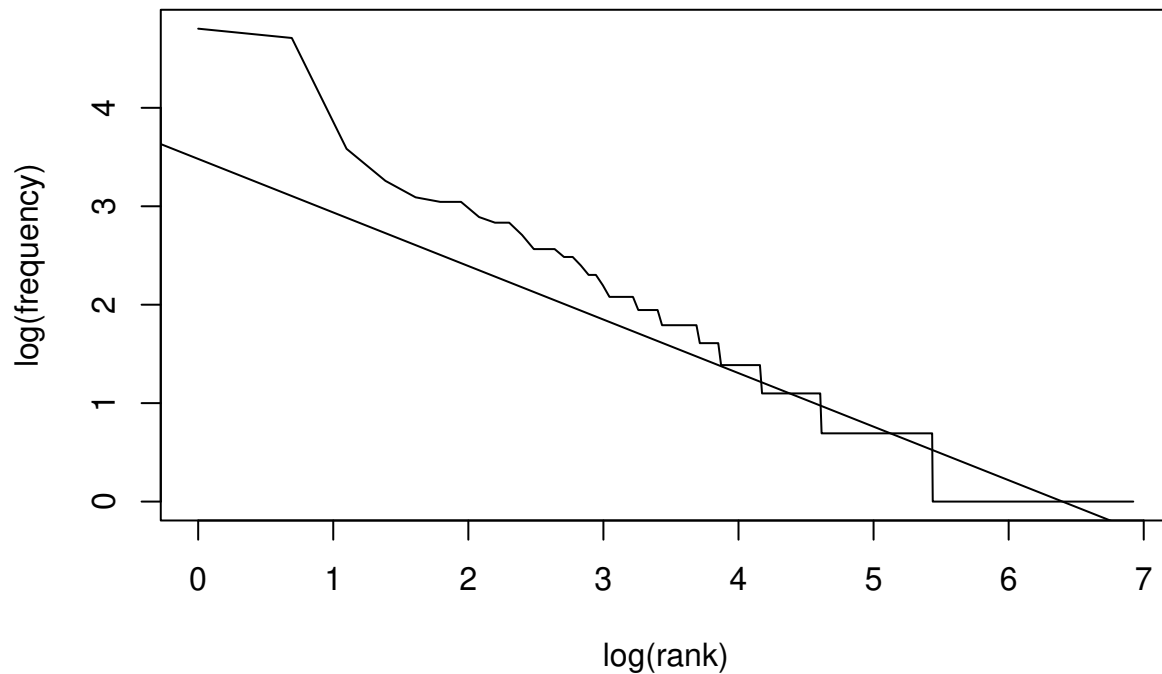
```
# find terms used > 5 times
findFreqTerms(a, 5)
```

```
##  [1] "`'tis"    "`and"     "`let"     "`thou"    "all"      "and"      "are"
##  [8] "art"      "blood"    "but"      "earth"    "england"  "english"  "every"
## [15] "for"      "from"     "had"      "have"     "her"      "his"      "its"
## [22] "let"      "like"     "make"     "not"      "one"      "over"     "own"
## [29] "such"     "that"     "the"      "their"    "them"     "they"     "those"
## [36] "thou"     "through"  "thy"      "upon"     "was"      "when"     "where"
## [43] "which"    "who"      "will"     "with"     "your"
```

```
# Zipf's law
Zipf_plot(a)
```

```
## (Intercept)            x
##   3.4810675  -0.5440241
```

```r
# war_prayer
data(war_prayer)

# create a document term matrix of war_prayer
wp <- DocumentTermMatrix(war_prayer)

# find terms used > 5 times
findFreqTerms(wp, 5)
```
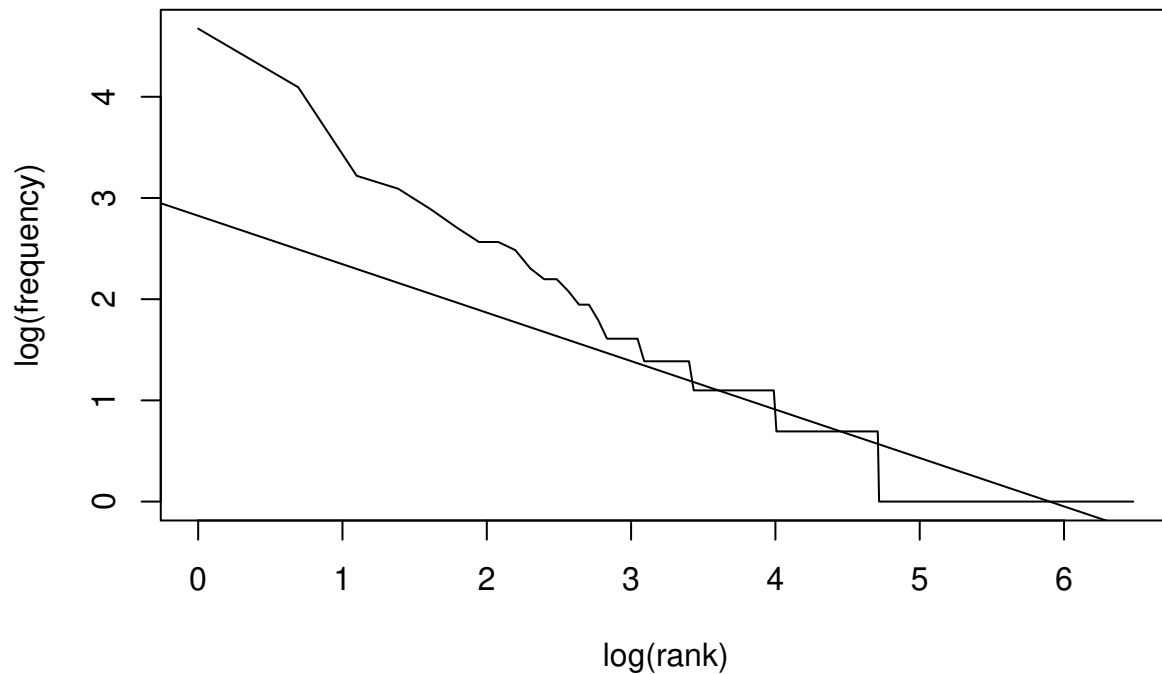
```
##  [1] "and"    "for"    "from"   "god"    "have"   "help"   "his"    "our"
##  [9] "prayed" "that"   "the"    "their"  "them"   "upon"   "war"    "was"
## [17] "which"  "who"    "with"   "you"    "your"
```

```r
# Zipf's law
Zipf_plot(wp)
```

```
## (Intercept)              x
##    2.824209   -0.478659
```

```r
# war_racket
data(war_racket)
```

```r
# create a document term matrix of war_racket
wr <- DocumentTermMatrix(war_racket)
```
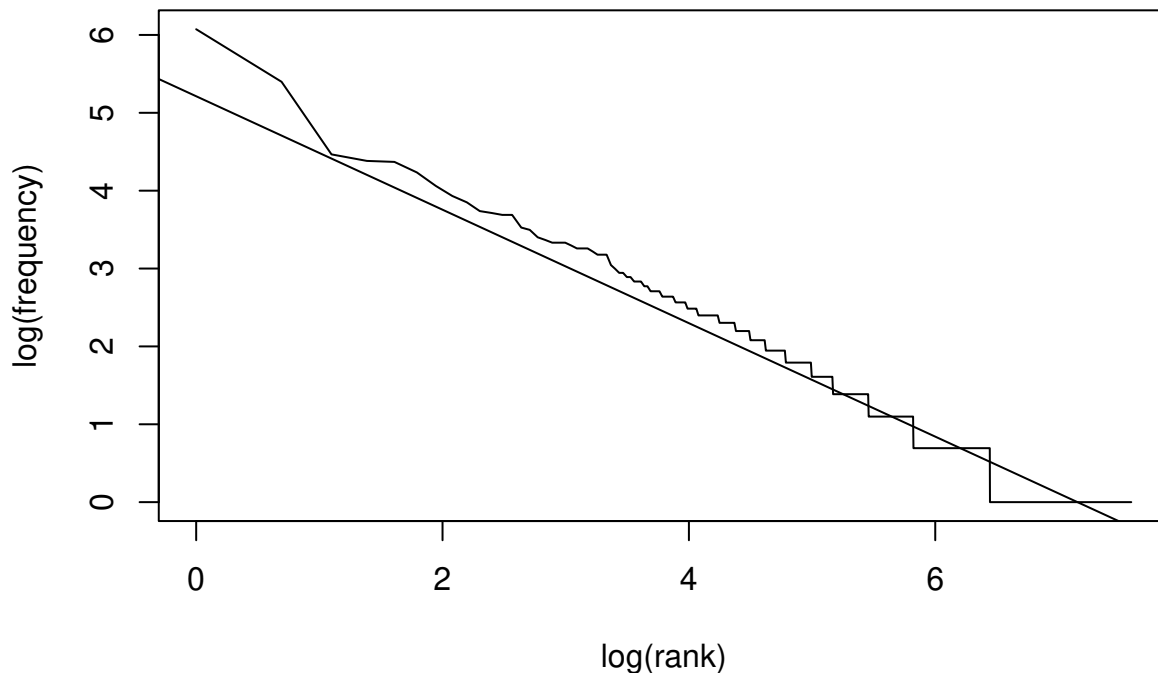
```r
# find terms used > 5 times
findFreqTerms(wr, 5)
```

```
##    [1] "about"         "all"           "american"      "and"
##    [5] "another"       "any"           "are"           "arms"
##    [9] "average"       "back"          "bankers"       "been"
##   [13] "before"        "being"         "bonds"         "bought"
##   [17] "boys"          "business"      "but"           "called"
##   [21] "came"          "can"           "can't"         "cannot"
##   [25] "capital"       "cent"          "cent."         "chapter"
##   [29] "cost"          "could"         "did"           "does"
##   [33] "dollars"       "don't"         "during"        "each"
##   [37] "earnings"      "eighteen"      "end"           "even"
##   [41] "every"         "few"           "for"           "france"
##   [45] "from"          "general"       "germany"       "get"
##   [49] "got"           "government"    "great"         "had"
##   [53] "has"           "have"          "his"           "home"
##   [57] "how"           "huge"          "hundred"       "increase"
##   [61] "industry"      "international" "into"          "it,"
##   [65] "it."           "its"           "just"          "kill"
##   [69] "less"          "let"           "let's"         "liberty"
##   [73] "like"          "limit"         "little"        "lot"
##   [77] "made"          "make"          "manufacturers" "many"
##   [81] "men"           "might"         "more"          "mosquito"
```

```
##  [85] "munitions"     "must"          "nation"        "never"
##  [89] "no."           "normal"        "not"           "now"
##  [93] "off"           "one"           "only"          "other"
##  [97] "our"           "out"           "over"          "own"
## [101] "paid"          "part"          "pay"           "pays"
## [105] "people"        "per"           "period"        "profit"
## [109] "profits"       "profits."      "put"           "racket"
## [113] "sam"           "same"          "ships"         "should"
## [117] "smash"         "sold"          "soldier"       "soldiers"
## [121] "some"          "states"        "steel"         "still"
## [125] "take"          "than"          "that"          "the"
## [129] "their"         "them"          "then"          "there"
## [133] "these"         "they"          "this"          "those"
## [137] "thousands"     "three"         "time"          "told"
## [141] "too"           "too."          "total"         "two"
## [145] "uncle"         "united"        "until"         "upon"
## [149] "use"           "very"          "war"           "war,"
## [153] "war."          "was"           "well"          "well,"
## [157] "went"          "were"          "what"          "when"
## [161] "where"         "whether"       "who"           "will"
## [165] "with"          "without"       "world"         "would"
## [169] "year"          "year."         "yearly"        "years"
## [173] "yes,"          "you"           "your"
```

```r
# Zipf's law
Zipf_plot(wr)
```



```
## (Intercept)            x
##    5.212469    -0.728547
```

```r
# connect_dots
data(connect_dots)
```

```r
# create a document term matrix of connect_dots
```

```
cd <- DocumentTermMatrix(connect_dots)

# find terms used > 5 times
findFreqTerms(cd, 5)

## [1] "and"       "for"       "has"       "have"      "must"      "our"
## [7] "petroleum" "that"      "the"

# Zipf's law
Zipf_plot(cd)
```
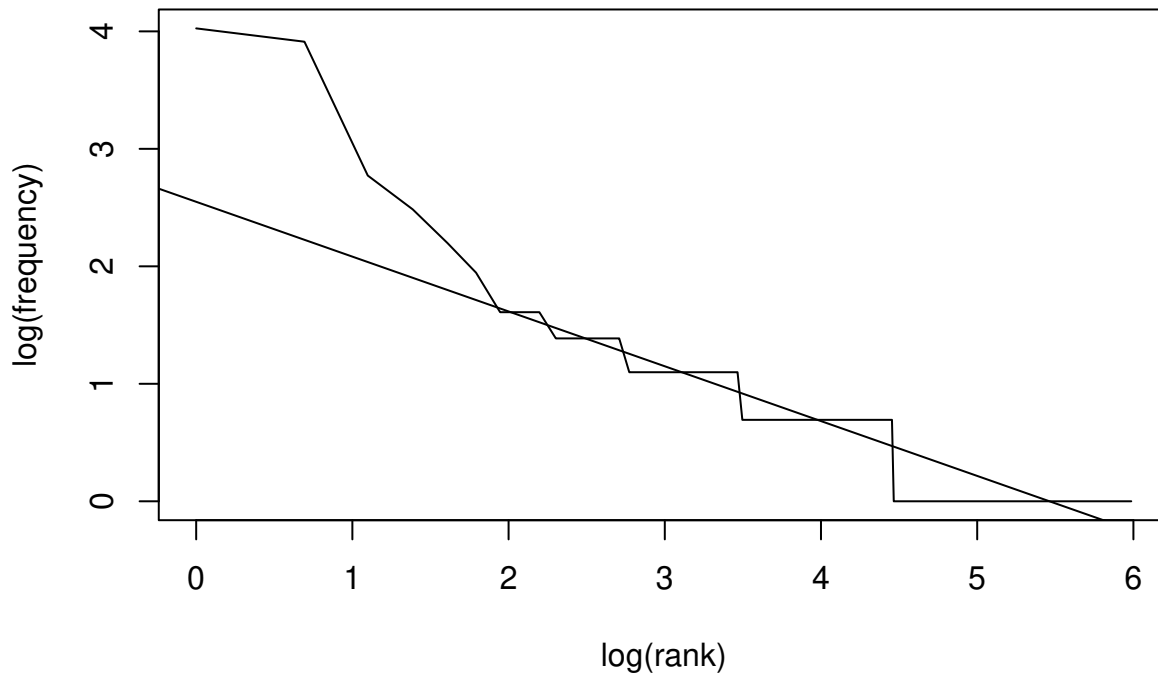


```
## (Intercept)          x
##   2.5493059  -0.4666954

# us_them
data(us_them)

# create a document term matrix of us_them
ut <- DocumentTermMatrix(us_them)

# find terms used > 5 times
findFreqTerms(ut, 5)

##  [1] "all"      "and"      "for"      "freedom"  "gave"      "know"     "not"
##  [8] "our"      "tha"      "the"      "their"    "them"      "they"      "will"
## [15] "won't"

# Zipf's law
Zipf_plot(ut)
```
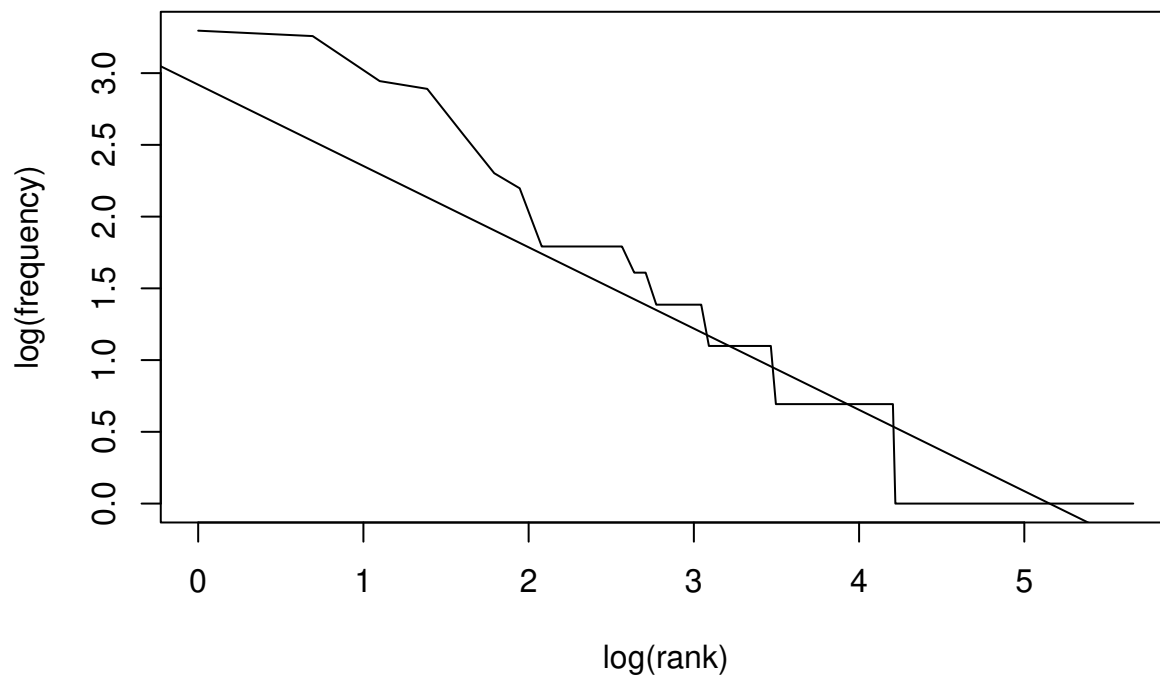
```
## (Intercept)             x
##   2.9193372  -0.5665568
```
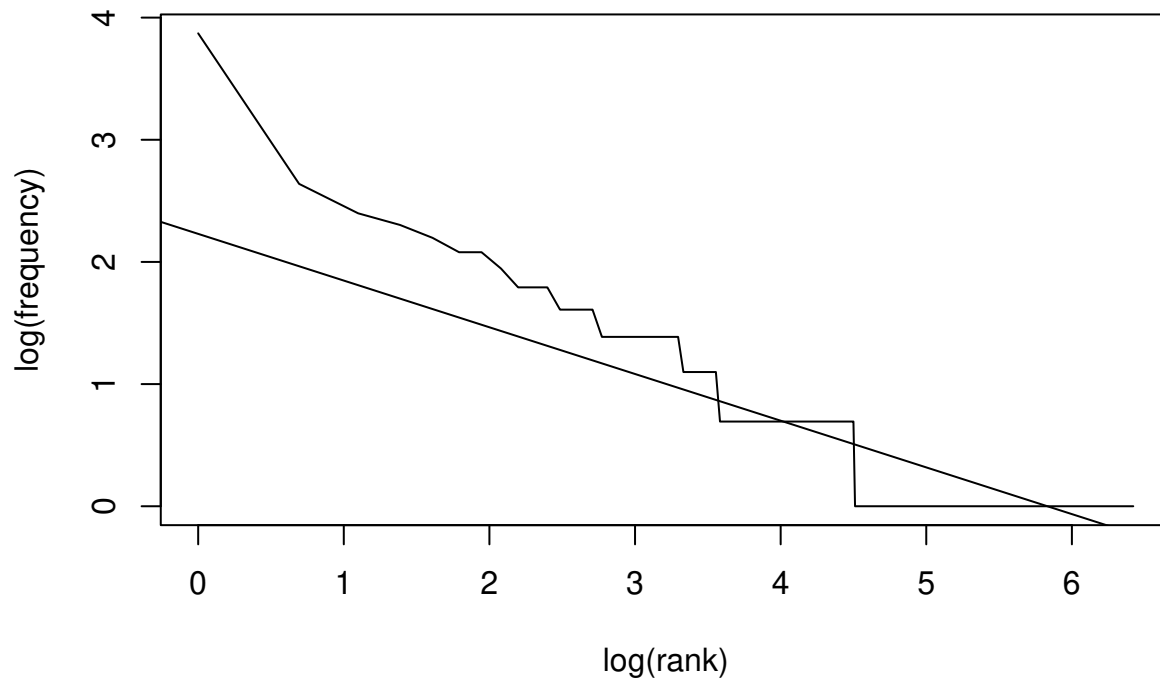
```r
# climate_strange
data(climate_strange)

# create a document term matrix of climate_strange
cs <- DocumentTermMatrix(climate_strange)

# find terms used > 5 times
findFreqTerms(cs, 5)
```

```
##  [1] "all"     "are"     "can"     "earth"   "for"     "have"    "life"
##  [8] "must"    "not"     "our"     "science" "that"    "the"     "this"
## [15] "your"
```

```r
# Zipf's law
Zipf_plot(cs)
```

```
## (Intercept)              x
##   2.2297149  -0.3823361
```

Example 2

```r
install.load::load_package("iemisctext", "tm")
# load needed packages using the load_package function from the install.load package (it is assumed tha

# appeal_womanhood_world
data(appeal_womanhood_world)

# View text representation of appeal_womanhood_world
inspect(appeal_womanhood_world[[1]])



# anarchy
data(anarchy)

# View text representation of anarchy
inspect(anarchy[[1]])



# war_prayer
data(war_prayer)
```

```
# View text representation of war_prayer
inspect(war_prayer[[1]])




# war_racket
# Example 2

data(war_racket)

# View text representation of war_racket
inspect(war_racket[[1]])




# connect_dots
data(connect_dots)

# View text representation of connect_dots
inspect(connect_dots[[1]])




# us_them
data(us_them)

# View text representation of us_them
inspect(us_them[[1]])




# climate_strange
data(climate_strange)

# View text representation of climate_strange
inspect(climate_strange[[1]])
```

EcoC²S Links

EcoC²S Home – https://www.ecoccs.com/
About EcoC²S – https://www.ecoccs.com/about-ecoc2s.html
Services – https://www.ecoccs.com/services.html
1 Stop Shop – https://www.ecoccs.com/other-biz.html
Products – https://www.questionuniverse.com/products.html
Media – https://www.ecoccs.com/media.html

Resources – https://www.ecoccs.com/resources.html
R Trainings and Resources provided by EcoC²S (Irucka Embry, E.I.T.) – https://www.ecoccs.com/rtraining.html

Copyright and License

All R code written by Irucka Embry is distributed under the GPL-3 (or later) license, see the GNU General Public License {GPL} page.

All written content originally created by Irucka Embry is copyrighted under the Creative Commons Attribution-ShareAlike 4.0 International license. All other written content retains the copyright of the original author(s).

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International license.