# STAT 451 - Visualizing Data - Autumn 2025

# Ariane Ducellier

# 11/20/2025

```
## corrplot 0.95 loaded
library(RColorBrewer)
library(tm)
## Loading required package: NLP
library(wordcloud)
Make a word cloud
# Set number of colors and palette
pal = brewer.pal(6, "RdGy")
# Choose minimum frequency and the range of the size of the words
wordcloud("The objective of this course is to provide students with a comprehensive understanding of da
## Warning in tm_map.SimpleCorpus(corpus, tm::removePunctuation): transformation
## drops documents
## Warning in tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,
## tm::stopwords())): transformation drops documents
         fundamentals effectivelybetter perception through visualizations
                 popular design Will cognition
        handson of creation creation crovide of the communicate practices based interactivity studen
                                                                             interactivityStudents
coordinates programming packages packages
                                                                         insights types decisions
                                                                        enabling geographic
             clear shiny exercises
                                                                                            learning
            projects informed C
                                                                      ⊦focus ഗ്ര
                                                                  8 also
                 choosing three
                         make USE advanced
                                                                                             topics right
                           atic colors word the second the s
                                                                                              course
             creating compelling tell main main main
                                    interactive practical
                                         zoomingpanning principles
```

To use a list of words and their frequencies.

techniques

Load R packages.

library(corrplot)



To read a text file and preprocess it, before doing the word cloud.

```
file = readLines("../../data/syllabus.txt")
doc = Corpus(VectorSource(file))
doc = tm_map(doc, tolower)
## Warning in tm_map.SimpleCorpus(doc, tolower): transformation drops documents
doc = tm_map(doc, removePunctuation)
## Warning in tm_map.SimpleCorpus(doc, removePunctuation): transformation drops
## documents
doc = tm_map(doc, removeNumbers)
## Warning in tm_map.SimpleCorpus(doc, removeNumbers): transformation drops
## documents
doc = tm_map(doc, removeWords, stopwords("english"))
## Warning in tm_map.SimpleCorpus(doc, removeWords, stopwords("english")):
## transformation drops documents
wordcloud(as.character(doc), scale=c(2, 0.5))
## Warning in tm_map.SimpleCorpus(corpus, tm::removePunctuation): transformation
## drops documents
## Warning in tm_map.SimpleCorpus(corpus, function(x) tm::removeWords(x,
## tm::stopwords())): transformation drops documents
     interactive interactive interactive
         visualizations
            cover
```

# Make two word clouds

```
files = DirSource("../../data/debate/")
data = Corpus(DirSource("../../data/debate/"))
data = tm_map(data, tolower)
data = tm_map(data, removePunctuation)
data = tm_map(data, removeNumbers)
data = tm_map(data, removeWords, c(stopwords("english"), "biden", "trump"))
data = TermDocumentMatrix(data)
data = as.matrix(data)
colnames(data) = c("biden", "trump")
comparison.cloud(data, max.words=100, title.size=2, colors=c("blue", "red"))
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : closed could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : obamacare could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : judges could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : november could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : radical could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : never could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : counted could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : together could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : states could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : thing could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : shut could not be fit on page. It will not be plotted.
## Warning in comparison.cloud(data, max.words = 100, title.size = 2, colors =
## c("blue", : management could not be fit on page. It will not be plotted.
```

```
home everybody police money biden president simply take people american create discredited doesn't united number totally america america able talk able talk can son covid by covid by also chris call lem much want look also chris call lem much want look and covid by also chris call lem much want look and look
```

# Plot correlations between texts

```
data(crude)
data = tm_map(crude, content_transformer(tolower))
data = tm_map(data, removePunctuation)
data = tm_map(data, removeNumbers)
data = tm_map(data, removeWords, stopwords("english"))
data = TermDocumentMatrix(data)
data = as.matrix(data)
crf = cor(data)
corrplot(crf, method = c("ellipse"))
```

