Makerspace Word Frequencies

```
library(tokenizers)
library(tm)
```

```
## Loading required package: NLP
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
data <- read.csv("MakerspaceSample.csv")
data$word_count <- sapply(strsplit(data$Description, "\\s+"), length)
dictionary <- paste(data$Description, collapse = ' ')
dictionary <- unlist(tokenize_words(tolower(dictionary)))
stop_words <- stopwords("en")
filtered_dictionary <- dictionary[!(dictionary %in% stop_words)]
```

```
# Calculate metrics
metrics <- data %>%
  summarise(
    average = mean(word_count),
    median = median(word_count),
    std_dev = sd(word_count),
    max = max(word_count),
    min = min(word_count)
  )

# View the calculated metrics
print(metrics)
```

```
##   average median  std_dev max min
## 1  217.74    201 114.9701 491  46
```

```r
#Top 30 words overall
word_frequencies <- table(filtered_dictionary)
word_frequencies <- sort(word_frequencies, decreasing = TRUE)
top_30 <- head(word_frequencies, 30)
print(top_30)
```

```
## filtered_dictionary
##  community      space      tools makerspace     makers     people   creative
##        124         80         65         63         57         55         51
##    members  equipment    classes        new        can     access       make
##         47         45         42         42         40         39         36
##   projects       shop      learn     create       work     making      build
##         36         35         33         32         32         30         28
##       open innovation   workshop      ideas      place creativity    artists
##         28         27         27         26         26         25         24
##   learning      maker
##         24         24
```

```r
#Top 30 words - nonprofit only
nonprofit <- data[data$Nonprofit_Commercial == "N",]
nonprofit <- paste(nonprofit$Description, collapse = ' ')
nonprofit <- unlist(tokenize_words(tolower(nonprofit)))
filtered_nonprofit <- nonprofit[!(nonprofit %in% stop_words)]
total_nonprofit <- length(filtered_nonprofit)
print(total_nonprofit)
```

```
## [1] 3189
```

```r
word_frequencies <- table(filtered_nonprofit)
word_frequencies <- sort(word_frequencies, decreasing = TRUE)
top_30 <- head(word_frequencies, 30)
print(top_30)
```

```
## filtered_nonprofit
##    community       people       makers        space        tools   makerspace
##           77           37           35           35           30           29
##     creative          new          can      classes       access       making
##           25           20           19           19           18           18
##      members            3      mission         open        build       create
##           18           17           17           17           16           16
##        learn     projects   innovation     together    equipment          non
##           16           16           15           15           14           14
##         work   creativity    knowledge     learning organization        place
##           14           13           13           13           13           13
```

```r
#Top 30 words - commercial only
commercial <- data[data$Nonprofit_Commercial == "C",]
commercial <- paste(commercial$Description, collapse = ' ')
commercial <- unlist(tokenize_words(tolower(commercial)))
filtered_commercial <- commercial[!(commercial %in% stop_words)]
total_commercial <- length(filtered_commercial)
print(total_commercial)
```

```
## [1] 3461
```

```
word_frequencies <- table(filtered_commercial)
word_frequencies <- sort(word_frequencies, decreasing = TRUE)
top_30 <- head(word_frequencies, 30)
print(top_30)
```

```
## filtered_commercial
##  community      space      tools makerspace  equipment    members   creative
##         47         45         35         34         31         29         26
##       shop       make    classes     makers        new     access        can
##         25         24         23         22         22         21         21
##   projects     people       work      learn        art   business     create
##         20         18         18         17         16         16         16
##        get industrial      ideas     studio   workshop    artists       like
##         15         15         14         14         14         13         13
##      maker       need
##         13         13
```

```
#Socioemotional coding
community <- c("collaboration", "community", "team", "unity", "collective", "networking", "support", "t
dei <- c("diversity", "equity", "inclusion", "everyone", "anyone", "all kinds", "for all", "supportive"
personal_growth <- c("empower", "empowerment", "autonomy", "agency", "self-efficacy", "confidence", "co
```

```
#Instrumental coding
creativity <- c("creativity", "creative", "creator", "creators", "imagine", "imagination", "new")
skills <- c("skill", "skills", "learn", "learning", "problem solving", "employment", "job", "jobs")
entrepreneurship <- c("entrepreneur", "entrepreneurs", "entrepreneurship", "innovation", "innovator", "
prototype <- c("prototype", "prototypes", "prototyping")
economy <- c("economic opportunity", "innovation economy", "economic growth", "opportunity", "economy",
```

```
#Socioemotional - nonprofit
community_freq_n <- sum(sapply(community, function(word) grepl(word, nonprofit)))
dei_freq_n <- sum(sapply(dei, function(word) grepl(word, nonprofit)))
personal_growth_freq_n <- sum(sapply(personal_growth, function(word) grepl(word, nonprofit)))

cat("Community Frequency:", community_freq_n, "\n")
```

```
## Community Frequency: 210
```

```
cat("Community Ratio:", community_freq_n/3189, "\n")
```

```
## Community Ratio: 0.06585136
```

```
cat("DEI Frequency:", dei_freq_n, "\n")
```

```
## DEI Frequency: 57
```

```r
cat("DEI Ratio:", dei_freq_n/3189, "\n")
```

```
## DEI Ratio: 0.01787394
```

```r
cat("Personal Growth Frequency:", personal_growth_freq_n, "\n")
```

```
## Personal Growth Frequency: 61
```

```r
cat("Personal Growth Ratio:", personal_growth_freq_n/3189, "\n")
```

```
## Personal Growth Ratio: 0.01912825
```

```r
#Socioemotional - commercial
community_freq_c <- sum(sapply(community, function(word) grepl(word, commercial)))
dei_freq_c <- sum(sapply(dei, function(word) grepl(word, commercial)))
personal_growth_freq_c <- sum(sapply(personal_growth, function(word) grepl(word, commercial)))

cat("Community Frequency:", community_freq_c, "\n")
```

```
## Community Frequency: 145
```

```r
cat("Community Ratio:", community_freq_c/3461, "\n")
```

```
## Community Ratio: 0.04189541
```

```r
cat("DEI Frequency:", dei_freq_c, "\n")
```

```
## DEI Frequency: 44
```

```r
cat("DEI Ratio:", dei_freq_c/3461, "\n")
```

```
## DEI Ratio: 0.01271309
```

```r
cat("Personal Growth Frequency:", personal_growth_freq_c, "\n")
```

```
## Personal Growth Frequency: 40
```

```r
cat("Personal Growth Ratio:", personal_growth_freq_c/3461, "\n")
```

```
## Personal Growth Ratio: 0.01155735
```

```r
#Instrumental - nonprofit
creativity_freq_n <- sum(sapply(creativity, function(word) grepl(word, nonprofit)))
skills_freq_n <- sum(sapply(skills, function(word) grepl(word, nonprofit)))
entrepreneurship_freq_n <- sum(sapply(entrepreneurship, function(word) grepl(word, nonprofit)))
prototype_freq_n <- sum(sapply(prototype, function(word) grepl(word, nonprofit)))
economy_freq_n <- sum(sapply(economy, function(word) grepl(word, nonprofit)))

cat("Creativity Frequency:", creativity_freq_n, "\n")
```

```
## Creativity Frequency: 62
```

```
cat("Creativity Ratio:", creativity_freq_n/3189, "\n")
```

```
## Creativity Ratio: 0.01944183
```

```
cat("Skills Frequency:", skills_freq_n, "\n")
```

```
## Skills Frequency: 78
```

```
cat("Skills Ratio:", skills_freq_n/3189, "\n")
```

```
## Skills Ratio: 0.02445908
```

```
cat("Entrepreneurship Frequency:", entrepreneurship_freq_n, "\n")
```

```
## Entrepreneurship Frequency: 58
```

```
cat("Entrepreneurship Ratio:", entrepreneurship_freq_n/3189, "\n")
```

```
## Entrepreneurship Ratio: 0.01818752
```

```
cat("Prototype Frequency:", prototype_freq_n, "\n")
```

```
## Prototype Frequency: 4
```

```
cat("Prototype Ratio:", prototype_freq_n/3189, "\n")
```

```
## Prototype Ratio: 0.001254312
```

```
cat("Economy Frequency:", economy_freq_n, "\n")
```

```
## Economy Frequency: 29
```

```
cat("Economy Ratio:", economy_freq_n/3189, "\n")
```

```
## Economy Ratio: 0.00909376
```

```
#Instrumental - commercial
creativity_freq_c <- sum(sapply(creativity, function(word) grepl(word, commercial)))
skills_freq_c <- sum(sapply(skills, function(word) grepl(word, commercial)))
entrepreneurship_freq_c <- sum(sapply(entrepreneurship, function(word) grepl(word, commercial)))
prototype_freq_c <- sum(sapply(prototype, function(word) grepl(word, commercial)))
economy_freq_c <- sum(sapply(economy, function(word) grepl(word, commercial)))

cat("Creativity Frequency:", creativity_freq_c, "\n")
```

```
## Creativity Frequency: 69
```

```r
cat("Creativity Ratio:", creativity_freq_c/3461, "\n")
```

```
## Creativity Ratio: 0.01993643
```

```r
cat("Skills Frequency:", skills_freq_c, "\n")
```

```
## Skills Frequency: 78
```

```r
cat("Skills Ratio:", skills_freq_c/3461, "\n")
```

```
## Skills Ratio: 0.02253684
```

```r
cat("Entrepreneurship Frequency:", entrepreneurship_freq_c, "\n")
```

```
## Entrepreneurship Frequency: 66
```

```r
cat("Entrepreneurship Ratio:", entrepreneurship_freq_c/3461, "\n")
```

```
## Entrepreneurship Ratio: 0.01906963
```

```r
cat("Prototype Frequency:", prototype_freq_c, "\n")
```

```
## Prototype Frequency: 11
```

```r
cat("Prototype Ratio:", prototype_freq_c/3461, "\n")
```

```
## Prototype Ratio: 0.003178272
```

```r
cat("Economy Frequency:", economy_freq_c, "\n")
```

```
## Economy Frequency: 18
```

```r
cat("Economy Ratio:", economy_freq_c/3461, "\n")
```

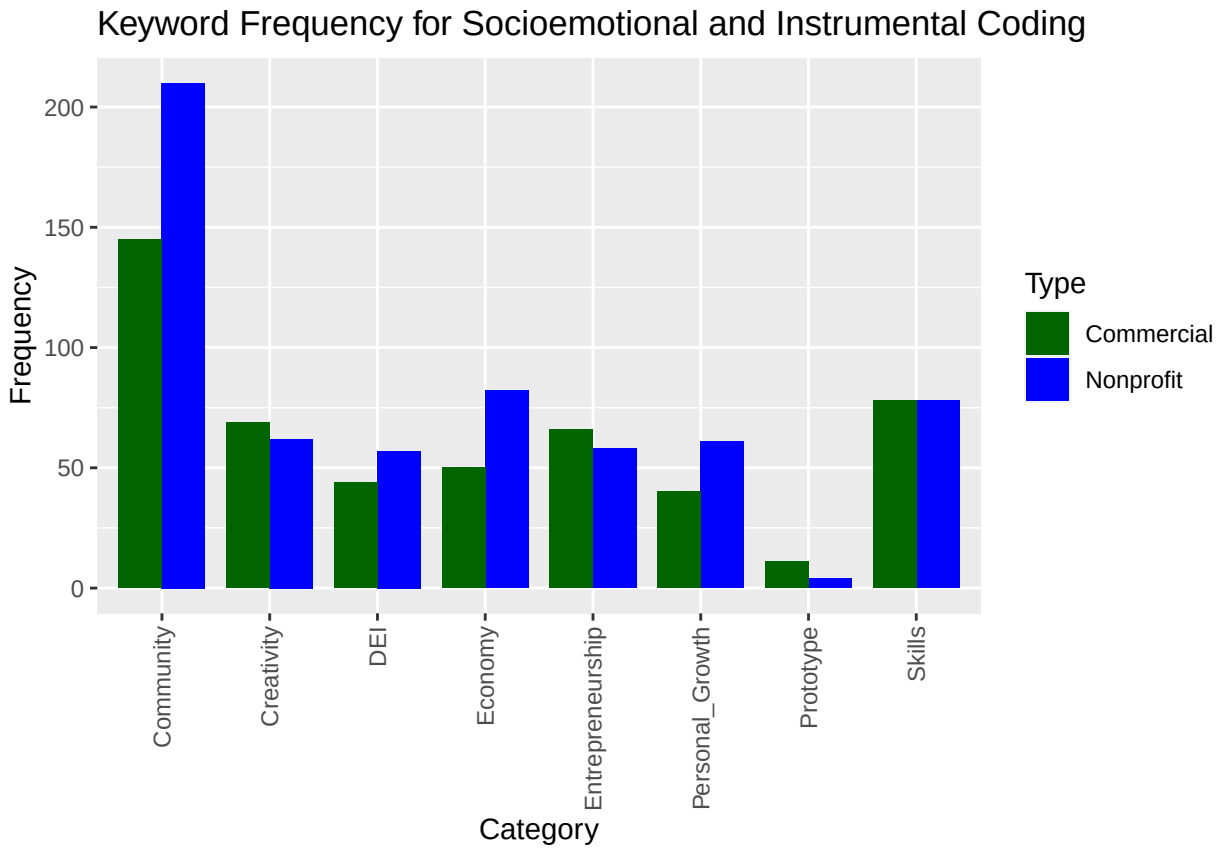
```
## Economy Ratio: 0.005200809
```

```r
library(ggplot2)
```

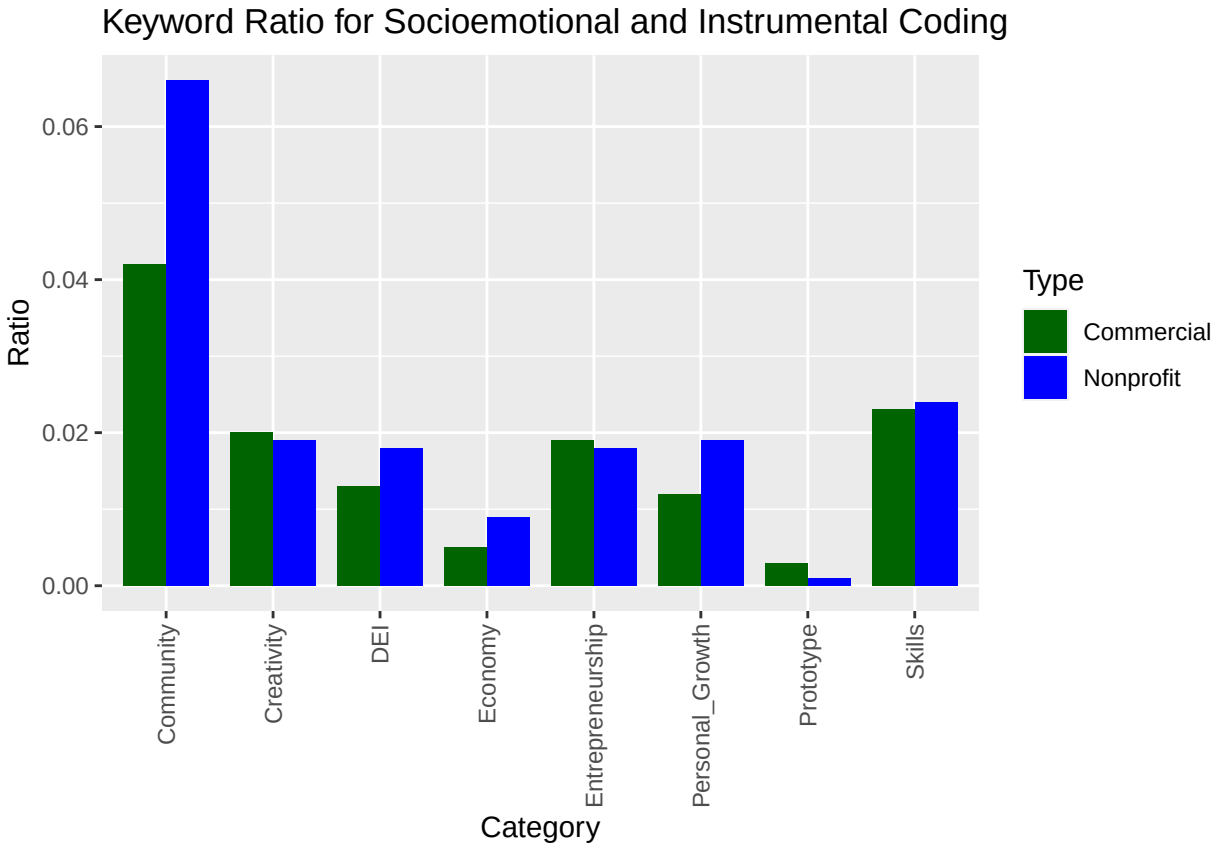
```
##
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:NLP':
##
##     annotate
```

```r
library(tidyr)

plot_df <- data.frame(Type=c("Nonprofit", "Commercial"), Community=c(210, 145), DEI=c(57, 44), Personal_

plot_df_long <- gather(plot_df, key = "Category", value = "Frequency", -Type)

ggplot(plot_df_long, aes(x = Category, y = Frequency, fill = Type)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.8) +
  labs(title = "Keyword Frequency for Socioemotional and Instrumental Coding",
       x = "Category",
       y = "Frequency") +
  scale_fill_manual(values = c("darkgreen", "blue"), name = "Type") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1))
```




```r
plot_df <- data.frame(Type=c("Nonprofit", "Commercial"), Community=c(0.066, 0.042), DEI=c(0.018, 0.013)

plot_df_long <- gather(plot_df, key = "Category", value = "Ratio", -Type)

ggplot(plot_df_long, aes(x = Category, y = Ratio, fill = Type)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.8) +
  labs(title = "Keyword Ratio for Socioemotional and Instrumental Coding",
       x = "Category",
       y = "Ratio") +
  scale_fill_manual(values = c("darkgreen", "blue"), name = "Type") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1))
```


Keyword Ratio for Socioemotional and Instrumental Coding
Ratio
0.06
0.04
0.02
0.00
Community
Creativity
DEI
Economy
Entrepreneurship
Personal_Growth
Prototype
Skills
Category
Type
Commercial
Nonprofit