

Demographics on Messaging Platforms

Benliro, Duero, and Sadsad BSIT 2-B

2024-05-02

```
library(readr)
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
```

```
library(tidyr)
```

```
messRespon<-read.csv('C:/Users/User/Documents/Rstudio Files/midtermsurvey/FINAL EXCEL/RESPONDENTS.csv')
View(messRespon)
```

```
#GRAPHING DEMOGRAPHICS UTILIZED IN THE SURVEY
```

```
#ABOUT: This survey is about Messaging Applications Platforms in which it testifies the satisfaction, e
```

```
#DEMOGRAPHIC: Age
```

```
age_freq<-table(messRespon$Age)
```

```
#Viewing of the frequency of the ages
```

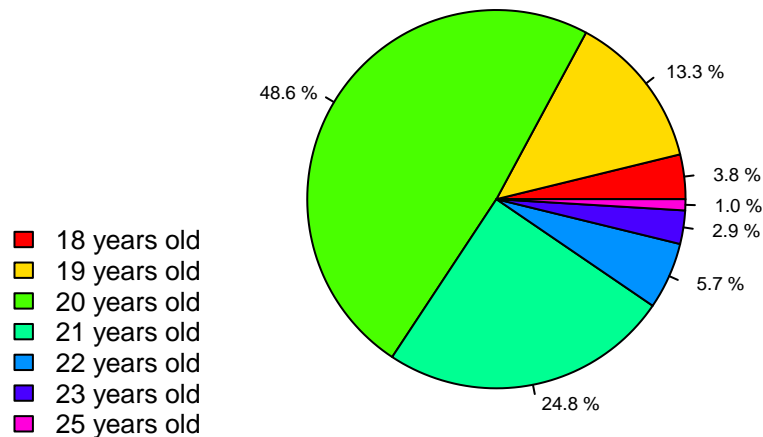
```
age_freq
```

```
##
## 18 19 20 21 22 23 25
##  4 14 51 26  6  3  1
```

```
percentages <- round((age_freq / sum(age_freq)) * 100, 1)
pie(age_freq,
    main = "Respondents by Age",
    col = rainbow(length(age_freq)),
    labels = paste("", format(percentages, nsmall = 1, digits = 2), "%"),
    cex = 0.6,
```

```
)
legend_labels <- paste(names(age_freq), "years old")
legend("bottomleft", legend = legend_labels, bty = "n", cex = 0.8, fill = rainbow(length(age_freq)))
```

Respondents by Age



#INSIGHTS: Based on the graph, it shows that mostly ages 19-21 are our primary respondents in our survey.

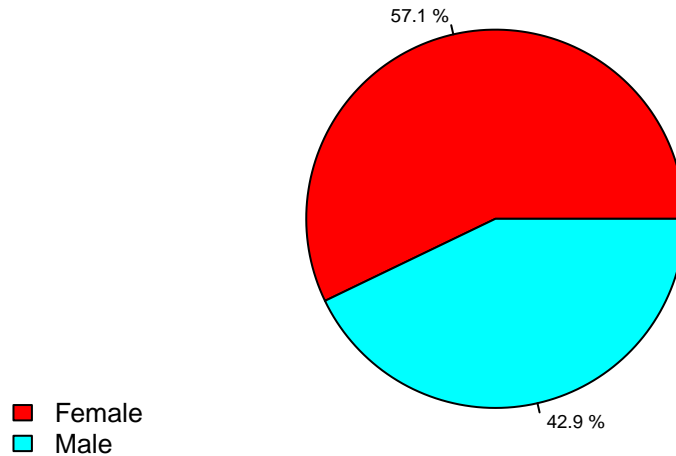
```
#DEMOGRAPHIC:Gender
gender_freq <- table(messRespon$Gender)
```

```
#Viewing of the frequency of the sex
gender_freq
```

```
##
## Female    Male
##      60      45
```

```
percentages <- round((gender_freq / sum(gender_freq)) * 100, 1)
pie(gender_freq,
    main = "Respondents by Sex",
    col = rainbow(length(gender_freq)),
    labels = paste("", format(percentages, nsmall = 1, digits = 2), "%"),
    cex = 0.6,
)
legend_labels <- paste(names(gender_freq))
legend("bottomleft", legend = legend_labels, bty = "n", cex = 0.8, fill = rainbow(length(gender_freq)))
```

Respondents by Sex



#INSIGHTS: Female respondents are our primary respondents for our survey.

#DEMOGRAPHIC: Type of Community

```
community_freq <- table(messRespon$Type.of.community)
```

#Viewing of the frequency of the community types

```
community_freq
```

```
##
```

```
## Rural Urban
```

```
##    50    55
```

```
percentages <- round((community_freq / sum(community_freq)) * 100, 1)
```

```
pie(community_freq,
```

```
  main = "Respondents by Type of Community",
```

```
  col = rainbow(length(community_freq)),
```

```
  labels = paste("", format(percentages, nsmall = 1, digits = 2), "%"),
```

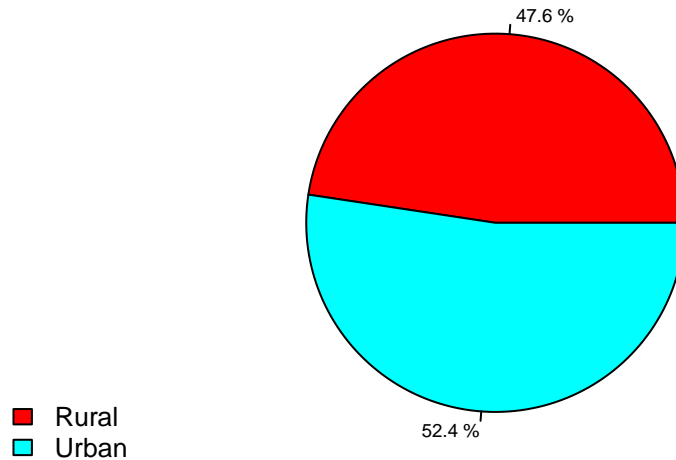
```
  cex = 0.6,
```

```
)
```

```
legend_labels <- paste(names(community_freq))
```

```
legend("bottomleft", legend = legend_labels, bty = "n", cex = 0.8, fill = rainbow(length(community_freq)))
```

Respondents by Type of Community



#INSIGHTS: Based on the graph, it shows that there are more than half of our respondents live in towns

```
#DEMOGRAPHIC:College
college_freq <- table(messRespon$College)

#Viewing of the frequency of the colleges (before transforming)
college_freq
```

```
##
##      Central Philippine University (CPU)
##                                     3
##      Iloilo City Community College
##                                     1
##      Iloilo Doctor's College
##                                     2
## Iloilo Science and Technology University
##                                     49
##      John B. Lacson University
##                                     1
##      St.Vicent College
##                                     1
## Univeristy of the Philippines - Visayas
##                                     1
##      University of Iloilo (UI)
##                                     18
```

```
##          University of San Agusin (USA)
##                                7
##      West Visayas State University (WVSU)
##                                18
##      Western Institute of Technology (WIT)
##                                4
```

```
single_respondent_colleges <- names(college_freq[college_freq == 1])
single_respondent_colleges
```

```
## [1] "Iloilo City Community College"
## [2] "John B. Lacson University"
## [3] "St.Vicent College"
## [4] "Univeristy of the Philippines - Visayas"
```

```
messRespon$College[messRespon$College %in% single_respondent_colleges] <- "Others"
```

```
college_freq <- table(messRespon$College)
```

```
#Viewing of the frequency of the colleges (after transforming)
college_freq
```

```
##
##      Central Philippine University (CPU)
##                                3
##      Iloilo Doctor's College
##                                2
## Iloilo Science and Technology University
##                                49
##                                Others
##                                4
##      University of Iloilo (UI)
##                                18
##      University of San Agusin (USA)
##                                7
##      West Visayas State University (WVSU)
##                                18
##      Western Institute of Technology (WIT)
##                                4
```

```
percentages <- round((college_freq / sum(college_freq)) * 100, 1)
```

```
pie_data <- college_freq
others_index <- "Others"
if (others_index %in% names(pie_data)) {
  pie_data[others_index] <- sum(pie_data[pie_data < 2])
  pie_data <- pie_data[pie_data >= 2]
}
```

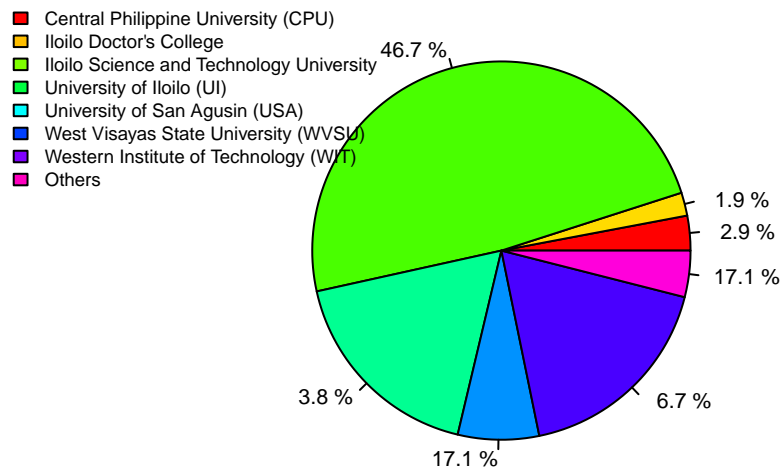
```
pie(pie_data,
    main = "Respondents by College",
    col = rainbow(length(pie_data)),
```

```
labels = paste("", format(percentages[pie_data >= 2], nsmall = 1, digits = 2), "%"),
cex = 0.7)
```

```
legend_labels <- names(pie_data)
```

```
legend("topleft", legend = c(legend_labels, "Others"), bty = "n", cex = 0.6, fill = rainbow(length(pie_data)))
```

Respondents by College



#Explanation: The data shows that the majority of the respondents are from Iloilo Science and Technology University.

#DEMOGRAPHIC: Courses

#As our data for courses are mixed, we first transform it by lowercasing the texts.

```
corrected_courses <- tolower(messRespon$Course)
```

```
valid_courses <- c(
  "bachelor of science in information technology",
  "bachelor of science in architecture",
  "bachelor of science in civil engineering",
  "bachelor of elementary education major in physical science",
  "bachelor of science in information systems",
  "bachelor of science in industrial technology (bit) - level iii",
  "bachelor in human services",
  "bachelor of secondary education major in science",
  "bachelor of science in computer science",
  "bachelor in fashion design and merchandising (bsfdm) - level ii",
```

```

"bachelor of science in nursing",
"bachelor of science in business administration major in financial management",
"bachelor of science in medical laboratory science",
"bachelor of science in hospitality management",
"bachelor of science in criminology",
"bachelor of science in economics",
"bachelor of science in biology with specialization in microbiology",
"bachelor of science in forestry",
"bachelor of science in marine engineering",
"bachelor of science in business administration",
"bachelor of art in english language studies",
"bachelor of science in accountancy",
"bachelor of library and information science",
"bachelor of science in development communication",
"bachelor of secondary major in english",
"bachelor of science in office administration",
"bachelor of science in pharmacy"
)

```

```

find_closest_match <- function(course) {
  distances <- sapply(valid_courses, function(valid_course) {
    sum(tolower(course) != tolower(valid_course))
  })
  closest_index <- which.min(distances)
  if (distances[closest_index] > 2) {
    return(course)
  } else {
    return(valid_courses[closest_index])
  }
}

```

```

corrected_courses <- sapply(corrected_courses, find_closest_match)

```

```

course_freq <- table(corrected_courses)
single_respondent_courses <- names(course_freq[course_freq == 1])
corrected_courses[corrected_courses %in% single_respondent_courses] <- "Others"
course_freq <- table(corrected_courses)

```

```

course_freq <- course_freq[order(-course_freq)]

```

```

#Viewing of the frequency of the courses
course_freq

```

```

## corrected_courses
##          bachelor of science in information technology
##                                     59
##                                     Others
##                                     11
##          bachelor of science in civil engineering
##                                     7
##          bachelor of science in nursing

```

```
##
##          bachelor of science in accountancy          6
##
##          bachelor of science in architecture          5
##
##          bachelor of science in hospitality management 3
##
##          bachelor of secondary education major in science 3
##
##          bachelor of library and information science    2
##
## bachelor of science in industrial technology (bit) - level iii 2
##
##          bachelor of science in information systems    2
##
##          bachelor of science in medical laboratory science 2
##
```

```
percentages <- round((course_freq / sum(course_freq)) * 100, 1)
```

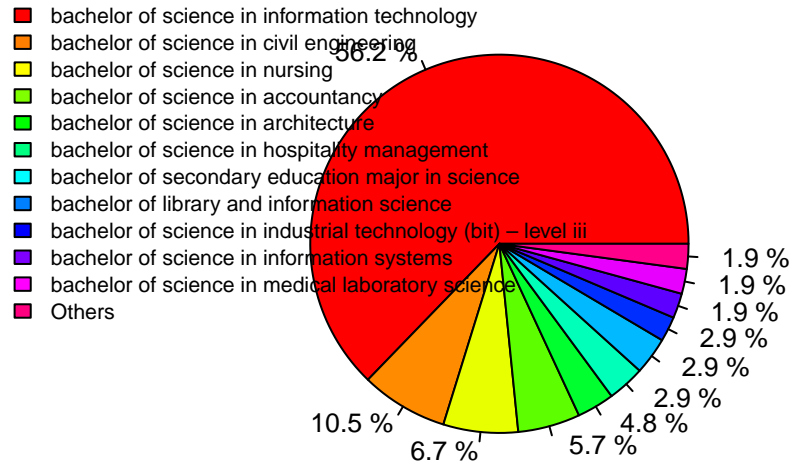
```
pie_data <- course_freq
others_index <- "Others"
if (others_index %in% names(pie_data)) {
  pie_data[others_index] <- sum(pie_data[pie_data < 2])
  pie_data <- pie_data[pie_data >= 2]
}
```

```
pie(pie_data,
    main = "Respondents by Courses",
    col = rainbow(length(pie_data)),
    labels = paste("", format(percentages[pie_data >= 2], nsmall = 1, digits = 2), "%"),
    cex = 0.8)
```

```
legend_labels <- names(pie_data)
```

```
legend("topleft", legend = c(legend_labels, "Others"), bty = "n", cex = 0.7, fill = rainbow(length(pie_
```


Respondents by Courses



#Explanation: The data shows that the majority of the respondents are from Bachelor of Science in Inform

#DEMOGRAPHIC:By Year Level

```
yearlvl_freq <- table(messRespon$Year.Level)
yearlvl_freq
```

```
##
## 1st Year 1st Year 2nd Year 3rd Year 4th Year
##      5      5      65      25      5
```

```
yearlvl_freq["1st Year"] <- yearlvl_freq["1st Year"] + yearlvl_freq["1st Year"]
yearlvl_freq <- yearlvl_freq[!names(yearlvl_freq) %in% "1st Year"]
```

```
percentages <- round((yearlvl_freq / sum(yearlvl_freq)) * 100, 1)
```

#Viewing of the frequency of the number of the year level.

```
yearlvl_freq
```

```
## 1st Year 2nd Year 3rd Year 4th Year
##     10     65     25     5
```

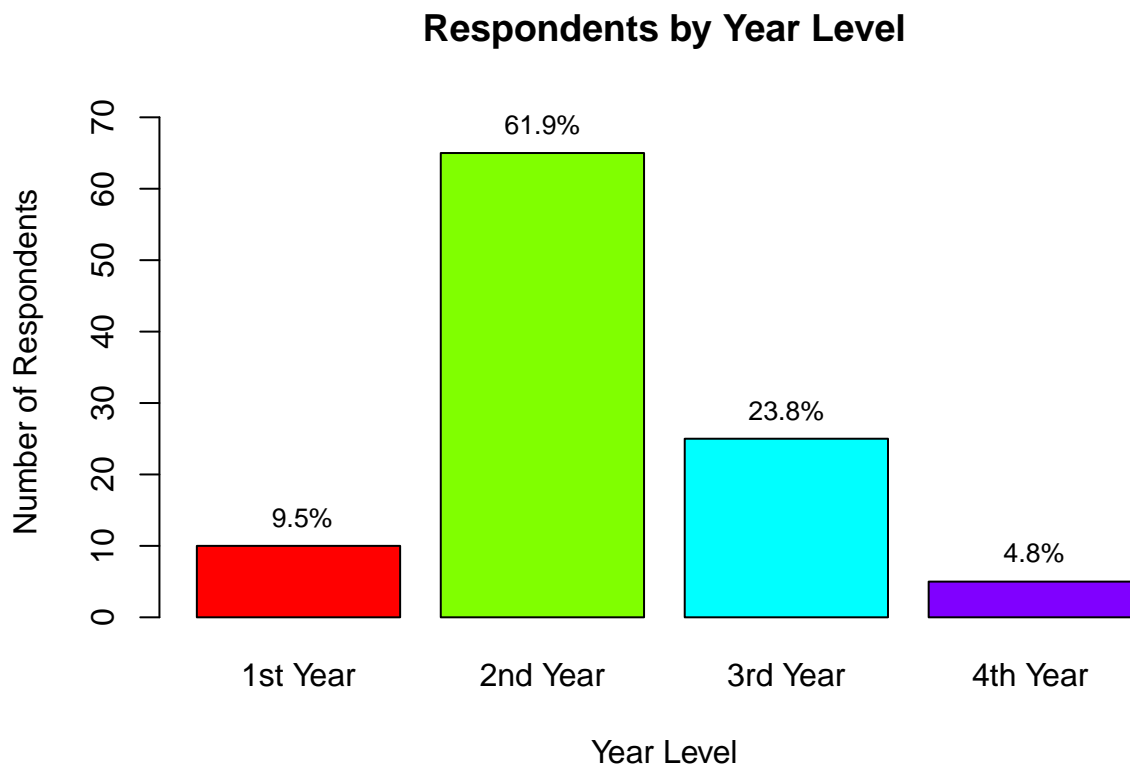
```
barplot(as.vector(yearlvl_freq),
        main = "Respondents by Year Level",
        col = rainbow(length(yearlvl_freq)),
```

```

xlab = "Year Level",
ylab = "Number of Respondents",
ylim = c(0, max(yearlvl_freq) * 1.1),
names.arg = names(yearlvl_freq))

text(x = barplot(as.vector(yearlvl_freq), plot = FALSE),
     y = yearlvl_freq,
     label = paste0(format(percentages, nsmall = 1, digits = 2), "%"),
     pos = 3, cex = 0.8)

```



#INSIGHTS: Mostly from 2nd year level answered our survey with a percentage of 61.9%, approximately 64

```

#DEMOGRAPHIC:Application used in messaging
messagingapp_freq <- table(messRespon$Which.messaging.app.do.you.use.most.frequently.)

#Viewing of the frequency of the applications used in messaging
messagingapp_freq

```

```

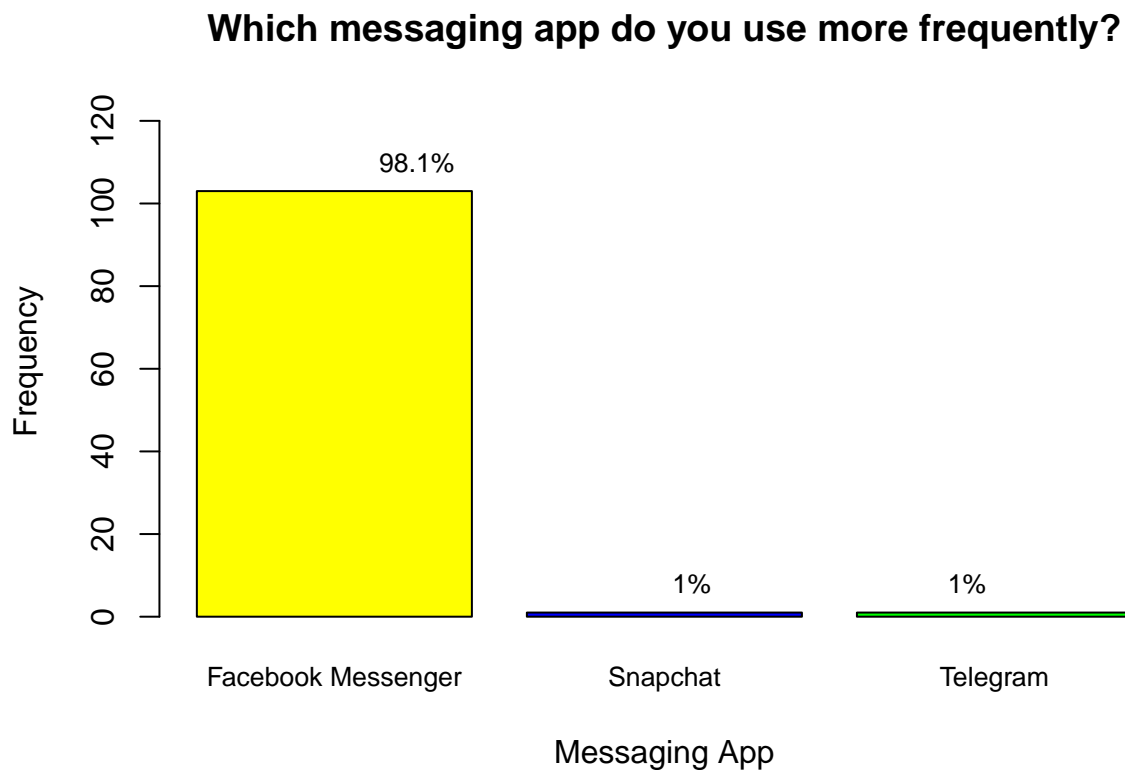
##
## Facebook Messenger      Snapchat      Telegram
##           103                1            1

```

```
percentages <- round((messagingapp_freq / sum(messagingapp_freq)) * 100, 1)

barplot(messagingapp_freq,
  main = "Which messaging app do you use more frequently?",
  col = c("yellow", "blue", "green"),
  names.arg = names(messagingapp_freq),
  ylim = c(0, max(messagingapp_freq) * 1.2),
  xlab = "Messaging App",
  ylab = "Frequency",
  cex.names = 0.8
)

text(x = 1:length(messagingapp_freq), y = messagingapp_freq, labels = paste0(percentages, "%"), pos = 3)
```



#INSIGHTS: Only one respondent used Telegram and also for the Snapchat. The rest used Facebook Messenger.

```
#DEMOGRAPHIC: Frequency of usage using messaging apps
usage_length <- table(messRespon$How.often.do.you.use.messaging.apps.)

#Viewing of the frequency of the usage length.
usage_length
```

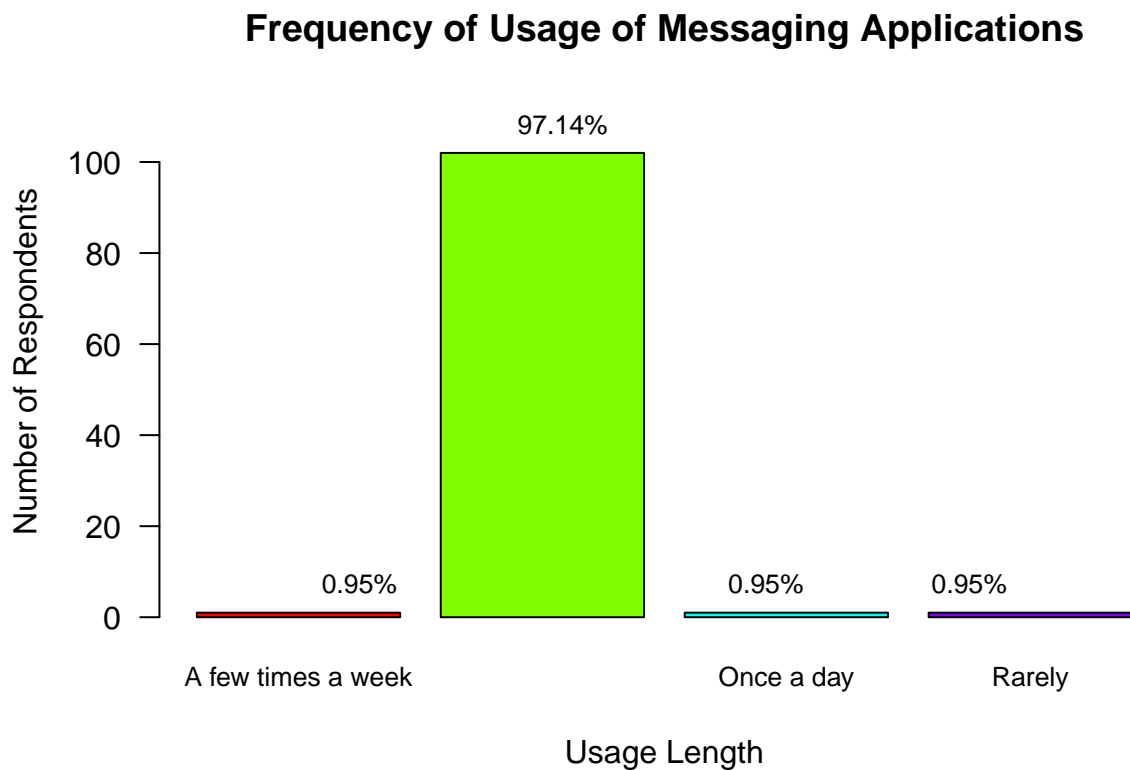
```
##
```

```
##      A few times a week Multiple times a day      Once a day
##              1              102              1
##              Rarely
##              1
```

```
percentages <- round(prop.table(usage_length) * 100, 2)
```

```
barplot(usage_length,
        main = "Frequency of Usage of Messaging Applications",
        col = rainbow(length(usage_length)),
        names.arg = names(usage_length),
        xlab = "Usage Length",
        ylab = "Number of Respondents",
        cex.names = 0.8,
        las = 1,
        ylim = c(0, max(usage_length) * 1.1)
)
```

```
text(x = 1:length(usage_length), y = usage_length, labels = paste0(percentages, "%"), pos = 3, cex = 0.8)
```



#INSIGHTS: Out of 105 respondents, the 102 respondents used messaging applications multiple times in a

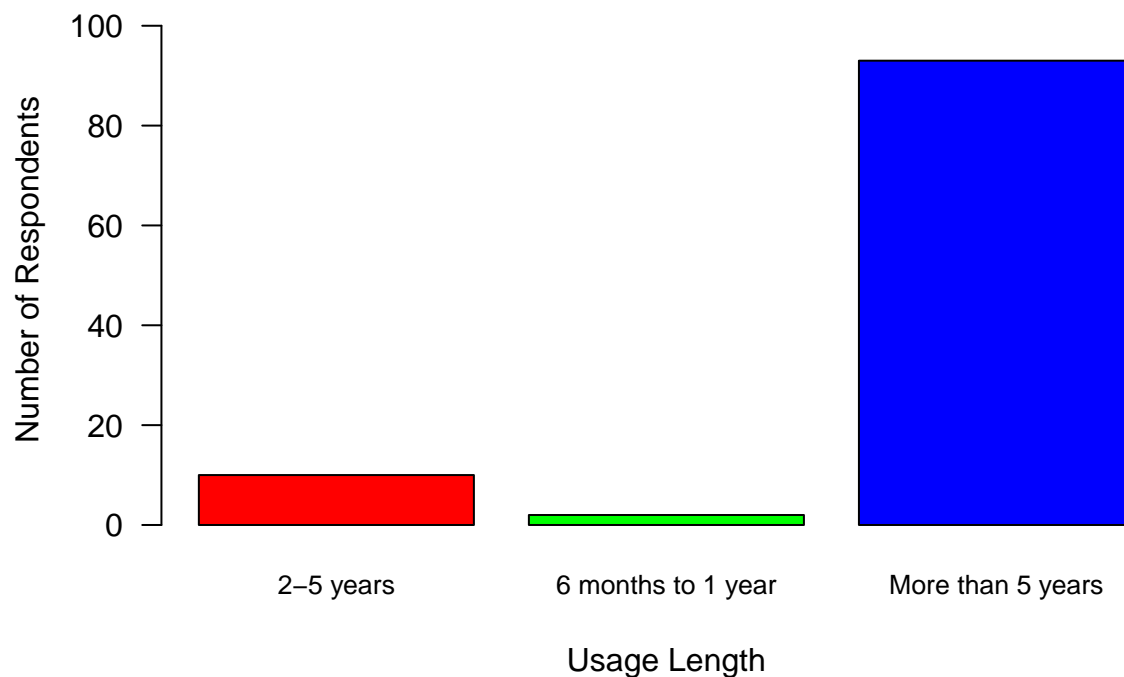
```
#DEMOGRAPHIC:How long have you been using the messaging app?
usage_length <- table(messRespon$How.long.have.you.been.using.messaging.apps.)
```

```
#Viewing of the frequency of the usage of the messaging app
usage_length
```

```
##
##          2-5 years 6 months to 1 year  More than 5 years
##                10                   2                93
```

```
percentages <- round(prop.table(usage_length) * 100, 2)
barplot(usage_length,
        main = "How long have you been using the messaging app?",
        col = rainbow(length(usage_length)),
        names.arg = names(usage_length),
        xlab = "Usage Length",
        ylab = "Number of Respondents",
        cex.names = 0.8,
        las = 1,
        ylim = c(0, max(usage_length) * 1.1)
)
```

How long have you been using the messaging app?



```
#INSIGHTS: Most of the respondents used the messaging for a long time, approximately more than 5 years.
```

```
#DEMOGRAPHIC: Device used for messaging
device_type <- table(messRespon$What.type.of.device.do.you.primarily.use.for.messaging.)

#Viewing of the frequency of the devices used.
device_type
```

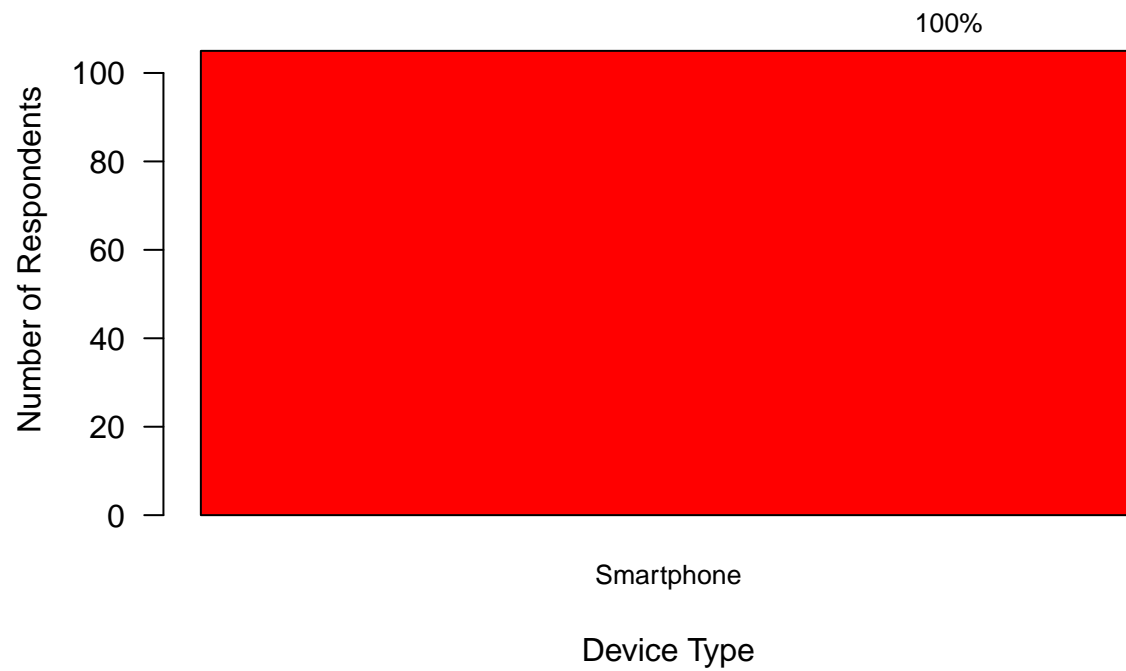
```
##
## Smartphone
##      105
```

```
percentages <- round(prop.table(device_type) * 100, 2)
```

```
barplot(device_type,
        main = "Type of device primarily use for messaging",
        col = rainbow(length(device_type)),
        names.arg = names(device_type),
        xlab = "Device Type",
        ylab = "Number of Respondents",
        cex.names = 0.8,
        las = 1,
        ylim = c(0, max(device_type) * 1.1)
)
```

```
text(x = 1:length(device_type), y = device_type, labels = paste0(percentages, "%"), pos = 3, cex = 0.8,
```

Type of device primarily use for messaging



#INSIGHTS: All of the respondents primarily used smartphone device as it its easy to access and navigat