---
output:
  pdf_document: default
  html_document: default

```{r}
library(rwhatsapp)
library(ggplot2)
library(ggimage)
library(tidyverse)
library(tidytext)
library(lubridate)
library(stopwords)
library(dplyr)
library(tidyr)
#load Data
#Import and Check Structure of Data
chat <- rwa_read("C:/Users/Talha/Downloads/_chat.txt") %>%
  filter(!is.na(author)) # remove messages without author
chat

#Messages Per day
chat%>%
    mutate(day = date(time)) %>%
  count(day) %>%
  ggplot(aes(x = day, y = n)) +
  geom_bar(stat = "identity") +
  ylab("y axis") + xlab("x axis") +
```

```
  ggtitle("Messages in 1 day day")

#Number Of Messages
chat %>%
  mutate(day = date(time)) %>%
  count(author) %>%
  ggplot(aes(x = reorder(author, n), y = n)) +
  geom_bar(stat = "identity") +
  ylab("") + xlab("") +
  coord_flip() +
  ggtitle("Number of messages")

#Most Often Used Emojis
chat %>%
  unnest(emoji) %>%
  count(author, emoji, sort = TRUE) %>%
  group_by(author) %>%
  top_n(n = 6, n) %>%
  ggplot(aes(x = reorder(emoji, n), y = n, fill = author)) +
  geom_col(show.legend = FALSE) +
  ylab("") +
  xlab("") +
  coord_flip() +
  facet_wrap(~author, ncol = 2, scales = "free_y")  +
  ggtitle("Most often used emojis")

#Most Often Used Words
chat %>%
  unnest_tokens(input = text,
```

```
            output = word) %>%
  count(author, word, sort = TRUE) %>%
  group_by(author) %>%
  top_n(n = 6, n) %>%
  ggplot(aes(x = reorder_within(word, n, author), y = n, fill = author)) +
  geom_col(show.legend = FALSE) +
  ylab("") +
  xlab("") +
  coord_flip() +
  facet_wrap(~author, ncol = 2, scales = "free_y") +
  scale_x_reordered() +
  ggtitle("Most often used words")
```