R Notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the Run button within the chunk or by placing your cursor inside it and pressing Cmd+Shift+Enter.

The first part that I want to know whether the gender and marriage status (“single”&“married”) will influence the text,words that related to their happy moment. I will use the wordcloud and heatmap to visilize the relationship.

step 0 : Load the needed library

```
## [1] "R version 3.5.2 (2018-12-20)"
```

step 1 : Load data

prepared the data

Data cleaning

We are going to see what the description of the data and then clean the data by examining the features of country, gender, age, marital status, parenthood. We found the missing data in country, age and ground_truth_category.

```
## ground_truth_category              country                  age 
##                 81195                  147                   33 
##                   wid          original_hm               gender 
##                     0                    0                    0 
##               marital           parenthood     reflection_period 
##                     0                    0                    0 
##    predicted_category                 text            orig.count 
##                     0                    0                    0 
##                 count 
##                     0
```

Since country and age are the two topics we’re interested in, we omit the data points that are not having those information for our analysis.

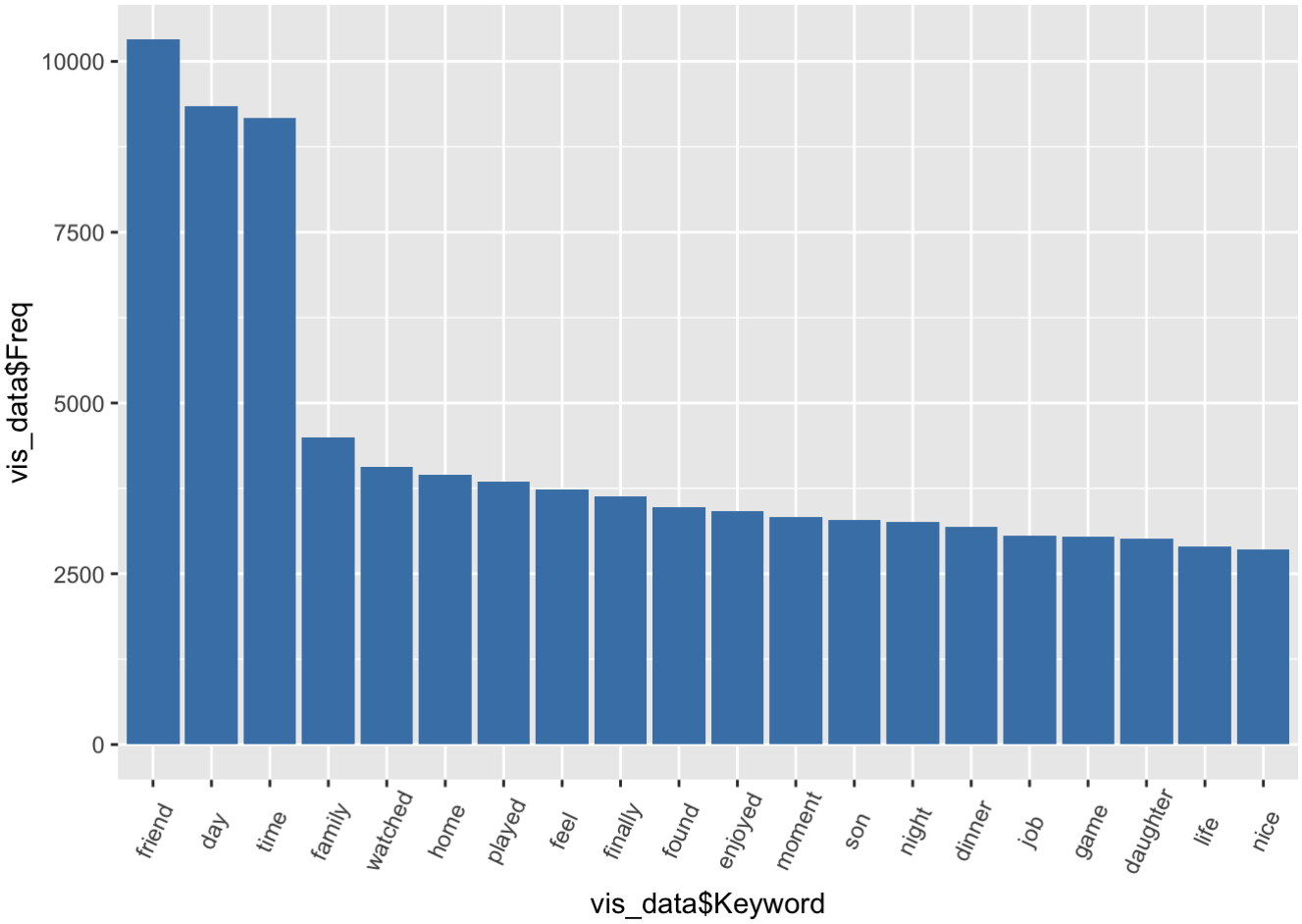
part 1

overiew wordcloud: Creating the word cloud and bar graph to acquire top 100 keywords

We would like to create a word cloud graph to explore what really makes staff happy.

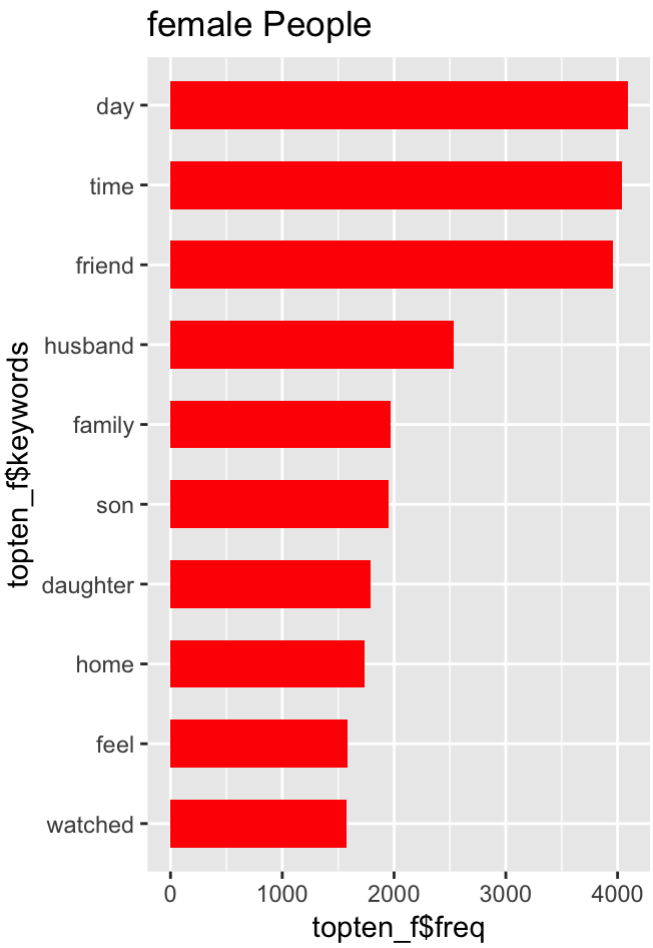
bar plot

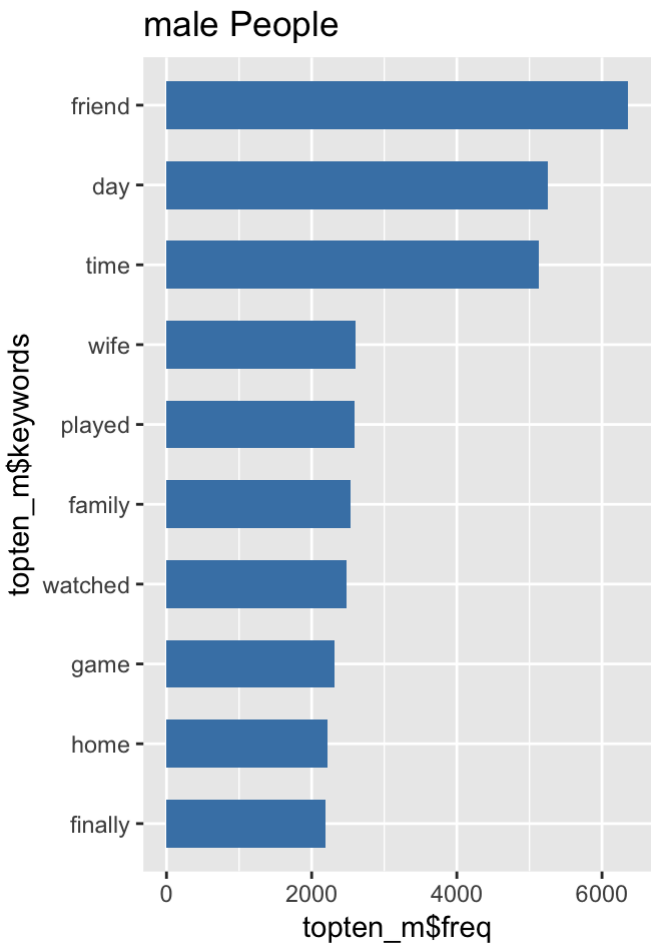
We select the top-20 highest frequence words to explore the significant factors of happiness.



PART 2: deference of words frequence between gender

bar plot





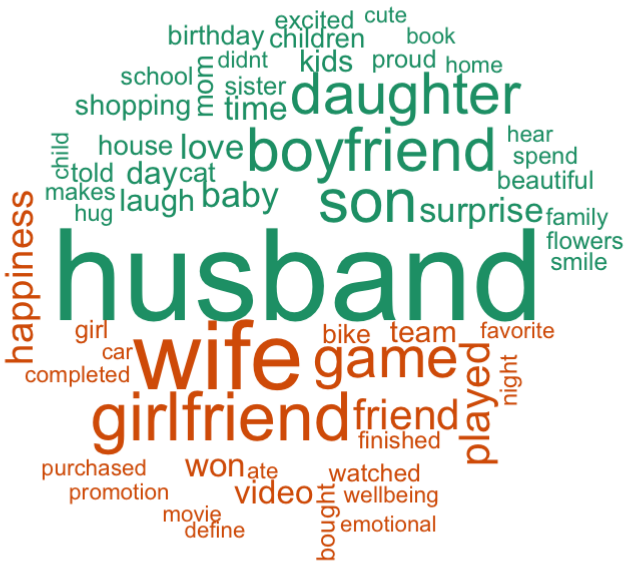
female/male individual wordcloud

walk
birthday
sister
boyfriend
watched
feel
nice
home
talked
friend
time
favorite
hours
surprise
enjoyed
visit
school
day
im
played
dinner
found
job
family
received
life
love
moment
son
night
event
morning
daughter
dog
finally
bought
house
started
mother
husband

started
birthday
event
received
nice
morning
home
favorite
won
finished
dog
talked
day
wife
love
found
played
feel
hours
dinner
movie
girlfriend
enjoyed
friend
night
time
game
son
life
family
car
finally
bought
school
watched
job
visit
daughter
moment
happiness

Comparison Cloud

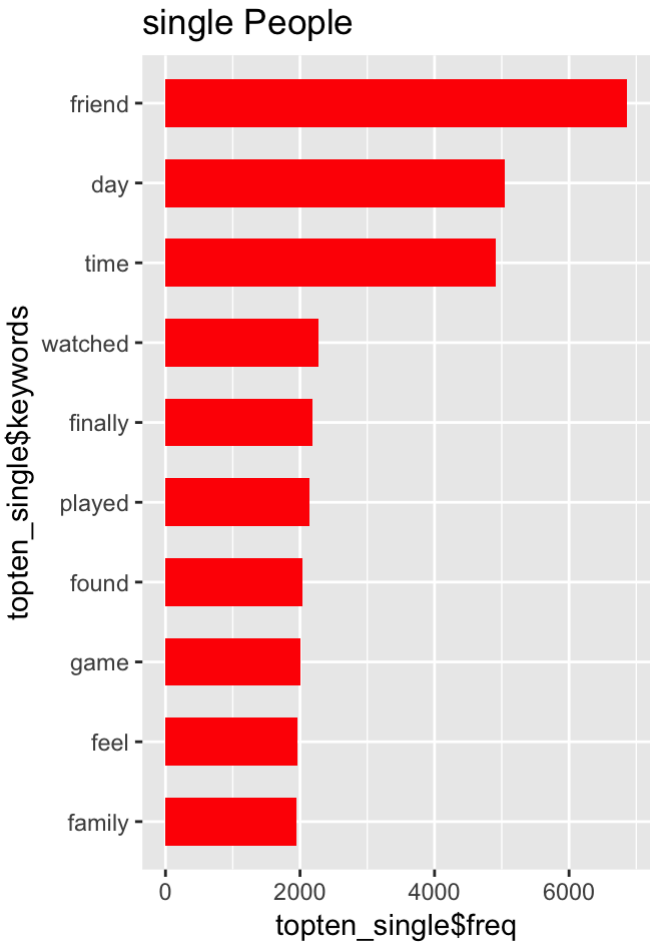
f

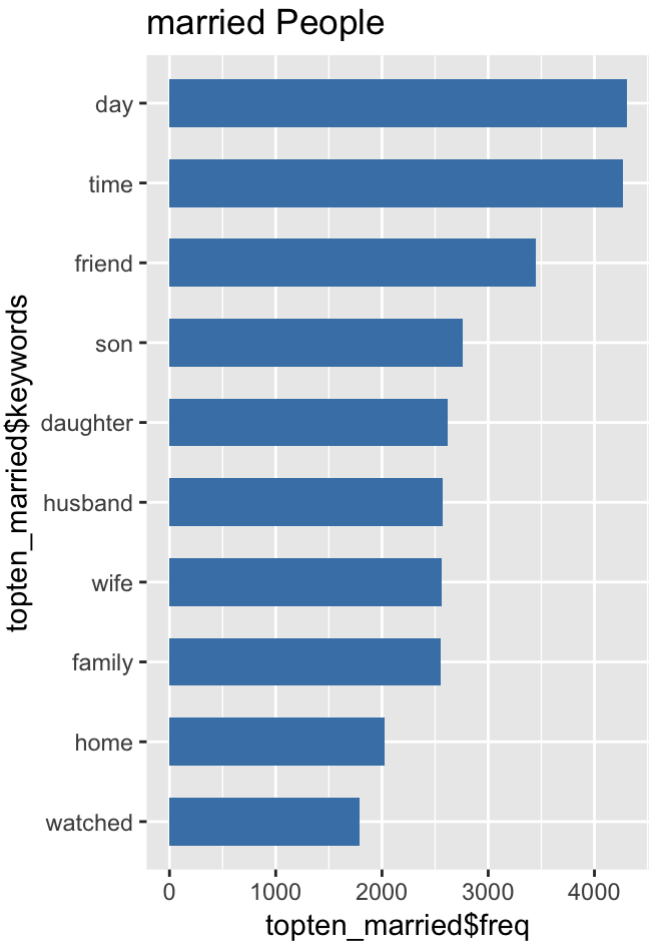


m

part3:deference of word frequence between martial status

bar plot





word clouds of different marital status

movie
visit
favorite
moment
job
received
sister
home
dinner
talked
car
found
finally
ate
bought
morning
enjoyed
feel
time
boyfriend
started
friend
game
hours
im
money
life
birthday
day
played
love
night
school
dog
watched
family
event
nice
finished
girlfriend

```
## NULL
```

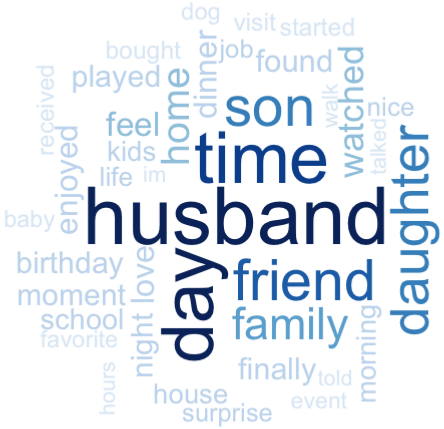
```
## NULL
```

Comparison Cloud

part4: defference of word frequence between gender and marital status

Wordcloud for single female

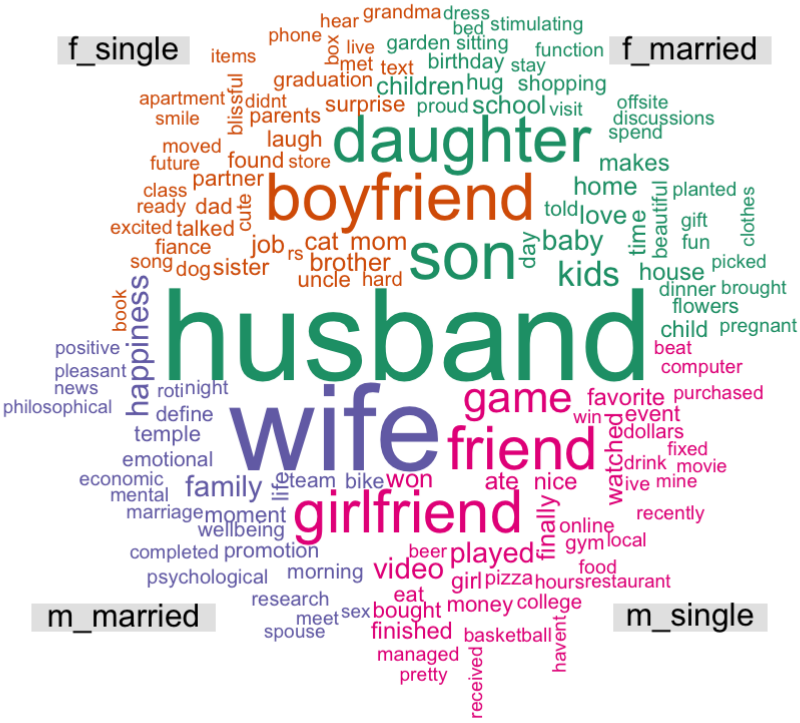
Wordcloud for married female



Wordcloud for single male

Wordcloud for married male

comparison wordcloud


f_single
f_married
m_married
m_single
husband
wife
daughter
boyfriend
son
friend
girlfriend
game
happiness
family
video
played
finally
watched
hear
grandma
dress
bed
stimulating
phone
box
live
garden
sitting
function
items
met
text
birthday
stay
graduation
children
hug
shopping
apartment
blissful
didnt
surprise
proud
school
visit
offsite
smile
parents
discussions
spend
moved
laugh
future
found
store
makes
class
partner
home
planted
ready
dad
cute
told
love
beautiful
clothes
excited
talked
time
gift
fiance
job
rs
cat
mom
day
baby
fun
song
dog
sister
brother
kids
house
picked
uncle
hard
dinner
brought
book
flowers
child
pregnant
positive
beat
pleasant
computer
news
roti
night
favorite
purchased
philosophical
define
win
event
temple
dollars
fixed
emotional
drink
movie
economic
life
team
bike
won
ate
nice
mental
ive
mine
marriage
moment
recently
wellbeing
online
completed
promotion
gym
local
beer
food
psychological
morning
girl
pizza
hours
restaurant
research
eat
money
college
meet
sex
bought
spouse
finished
basketball
havent
managed
pretty
received