

Code ▼

The Differences of the Causes of Happiness Between Female and Male

The dataset, HappyDB is a corpus of 100,000 crowd-sourced happy moments. The goal of the corpus is to advance the state of the art of understanding the causes of happiness that can be gleaned from text.

Environmental settings of this project

Hide

R.version

```
platform      _  
arch          aarch64-apple-darwin20  
arch          aarch64  
os            darwin20  
system        aarch64, darwin20  
status  
major         4  
minor         2.2  
year          2022  
month         10  
day           31  
svn rev       83211  
language      R  
version.string R version 4.2.2 (2022-10-31)  
nickname      Innocent and Trusting
```

Step 1: data processing and mining

Step 0 - Load all the required libraries

From the packages' descriptions:

- `tm` is a framework for text mining applications within R;
- `tidyverse` is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures;
- `tidytext` allows text mining using 'dplyr', 'ggplot2', and other tidy tools;

- DT provides an R interface to the JavaScript library DataTables.

[Hide](#)

```
library(tm)
library(tidytext)
library(tidyverse)
library(DT)
```

Step 1 - Load the data to be cleaned and processed

Step 2 - Preliminary cleaning of text

Clean the text by converting all the letters to the lower case, and removing punctuation, numbers, empty words and extra white space.

Step 3 - Stemming words and converting tm object to tidy object

Stemming reduces a word to its word *stem*. We stem the words here and then convert the “tm” object to a “tidy” object for much faster processing.

Step 4 - Creating tidy format of the dictionary to be used for completing stems

Need a dictionary to look up the words corresponding to the stems.

Step 5 - Removing stopwords that don't hold any significant information for our data set

Remove stopwords provided by the “tidytext” package and also add custom stopwords in context of our data.

Step 6 - Combining stems and dictionary into the same tibble

Combine the stems and the dictionary into the same “tidy” object.

Step 7 - Stem completion

Complete the stems by picking the corresponding word with the highest frequency.

Step 8 - Pasting stem completed individual words into their respective happy moments

We want our processed words to resemble the structure of the original happy moments. So we paste the words together to form happy moments.

Step 9 - Keeping a track of the happy moments with their own ID

Exporting the processed text data into a CSV file

[Hide](#)

```
write_csv(hm_data, "../output/processed_moments.csv")
```

The final processed data is ready to be used for analysis.

Step 2: Data Analysis and Visualization

Load all the required libraries

From the packages' descriptions:

- `tidyverse` is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures;
- `tidytext` allows text mining using `'dplyr'`, `'ggplot2'`, and other tidy tools;
- `DT` provides an R interface to the JavaScript library DataTables;
- `scales` map data to aesthetics, and provide methods for automatically determining breaks and labels for axes and legends;
- `wordcloud2` provides an HTML5 interface to wordcloud for data visualization;
- `gridExtra` contains miscellaneous functions for "grid" graphics;
- `ngram` is for constructing n-grams ("tokenizing"), as well as generating new text based on the n-gram structure of a given text input ("babbling");
- `Shiny` is an R package that makes it easy to build interactive web apps straight from R;

[Hide](#)

```
library(tidyverse)
library(tidytext)
library(DT)
library(scales)
library(wordcloud2)
library(gridExtra)
library(ngram)
library(shiny)
```

Load the processed text data along with demographic information on contributors

Use the processed data for our analysis and combine it with the demographic information available.

[Hide](#)

```
hm_data <- read_csv("../output/processed_moments.csv")

urlfile<-'https://raw.githubusercontent.com/rit-public/HappyDB/master/happydb/data/demographic.csv'
demo_data <- read_csv(urlfile)
```

Then combine both the data sets and keep the required columns for analysis. Select a subset of the data that satisfies specific row conditions.

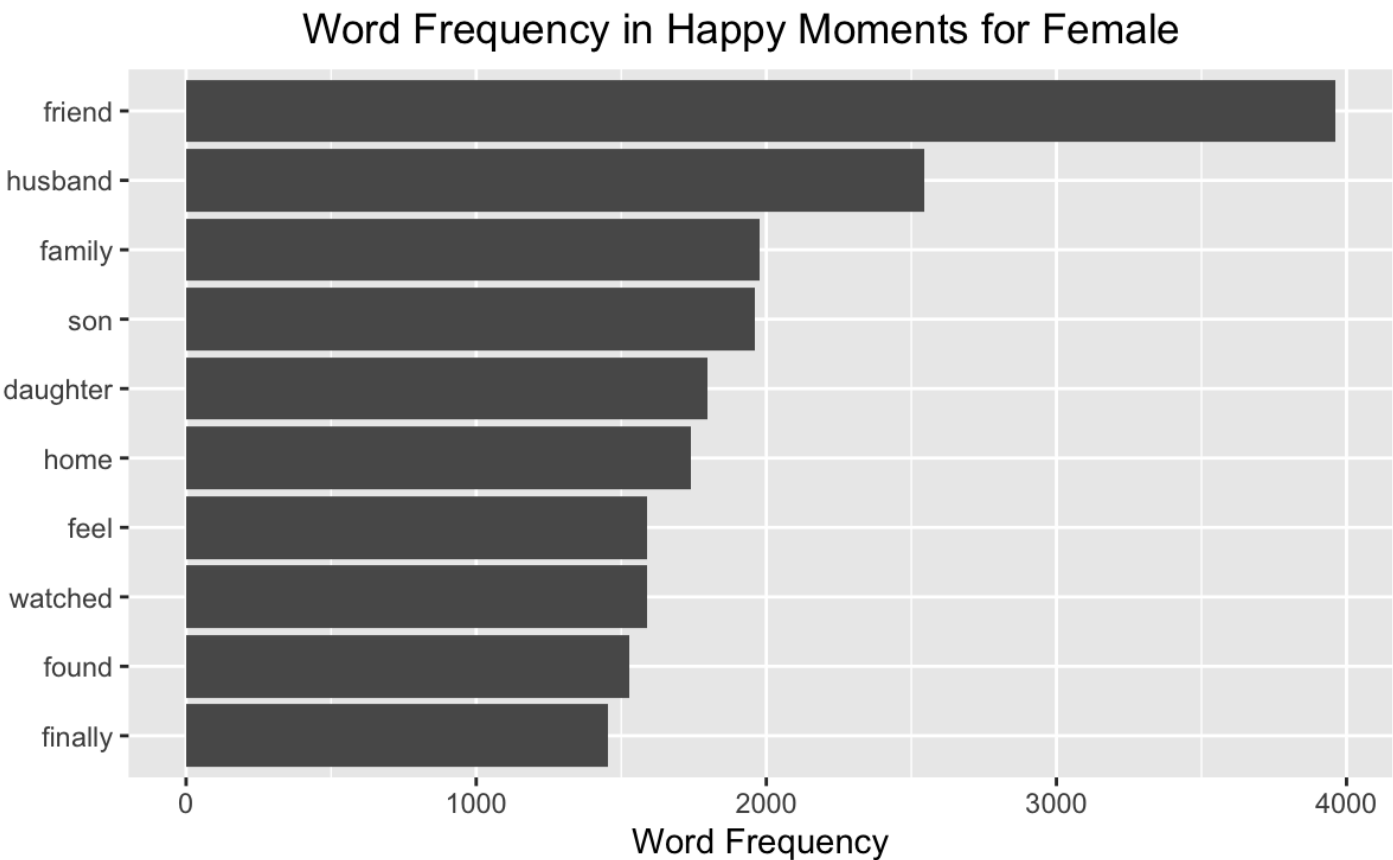
Question 1: What are the differences of the causes of happiness between female and male in general? Do female and male feel happy about the same things?

In daily life, it is not difficult to realize that women and men sometimes have different interests. For example, when taking a break, most women like shopping, while men like playing video games. So what factors best make women feel happy, and what factors best make men feel happy?

To answer these questions, I compare the “happy moments” between female and male in general first, regardless their marital status, age, and so on.

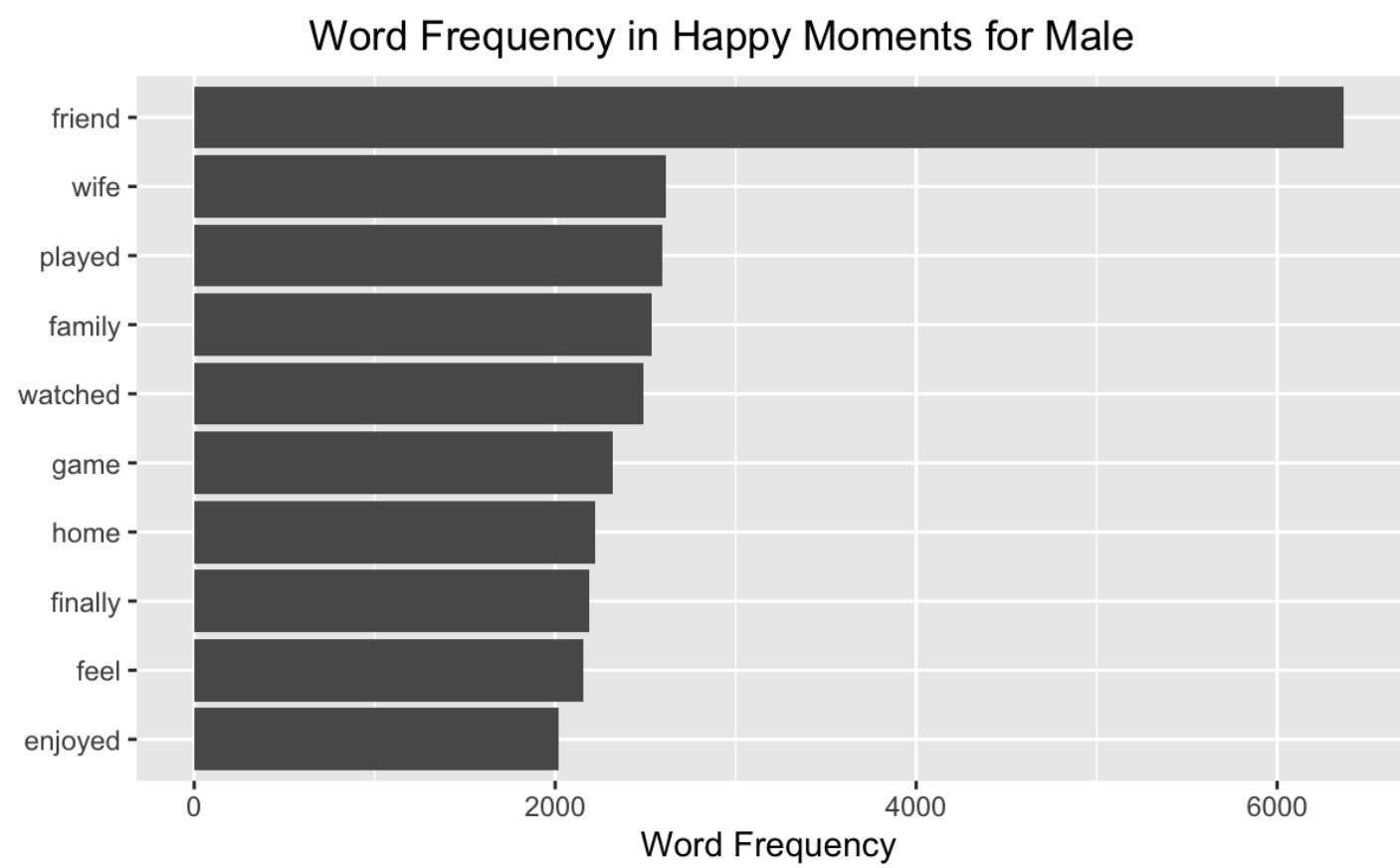
We need create a bag of words using the text data (female)

For female:



Create a bag of words using the text data (male)

For male:



To answer questions above, I first create two tables of hm_data for female and male separately. Then I use the text data to create bag of words to get two plots of word frequency in happy moments for female and male.

According to these two plots, we can see that ‘friend’ is the most frequent word in the plot of Word Frequency in Happy Moments for Female, followed by ‘husband’. Other terms like ‘family’, ‘son’, and ‘daughter’ also feature prominently, followed by ‘home’, ‘feel’, ‘watched’, etc.

While for Word Frequency in Happy Moments for Male, ‘friend’ is also the most frequent word, and ‘wife’ afterward. The frequency of words ‘played’, ‘family’, ‘watched’, ‘game’ are also high.

Both female and male feel happy for ‘friend’, but for male, the proportion of ‘friend’ is much bigger than female. In addition, both female and male feel happy with their spouse. However, females exhibit a stronger inclination towards family and children-related joys. Males, in contrast, seem to derive more pleasure from solitary activities or hobbies.



The word cloud gives the same result.

Not surprisingly, women seem to get more happiness from their families. So what is the difference between happy moments for single men and women, and married men and women? Will it make a huge difference to the results above?

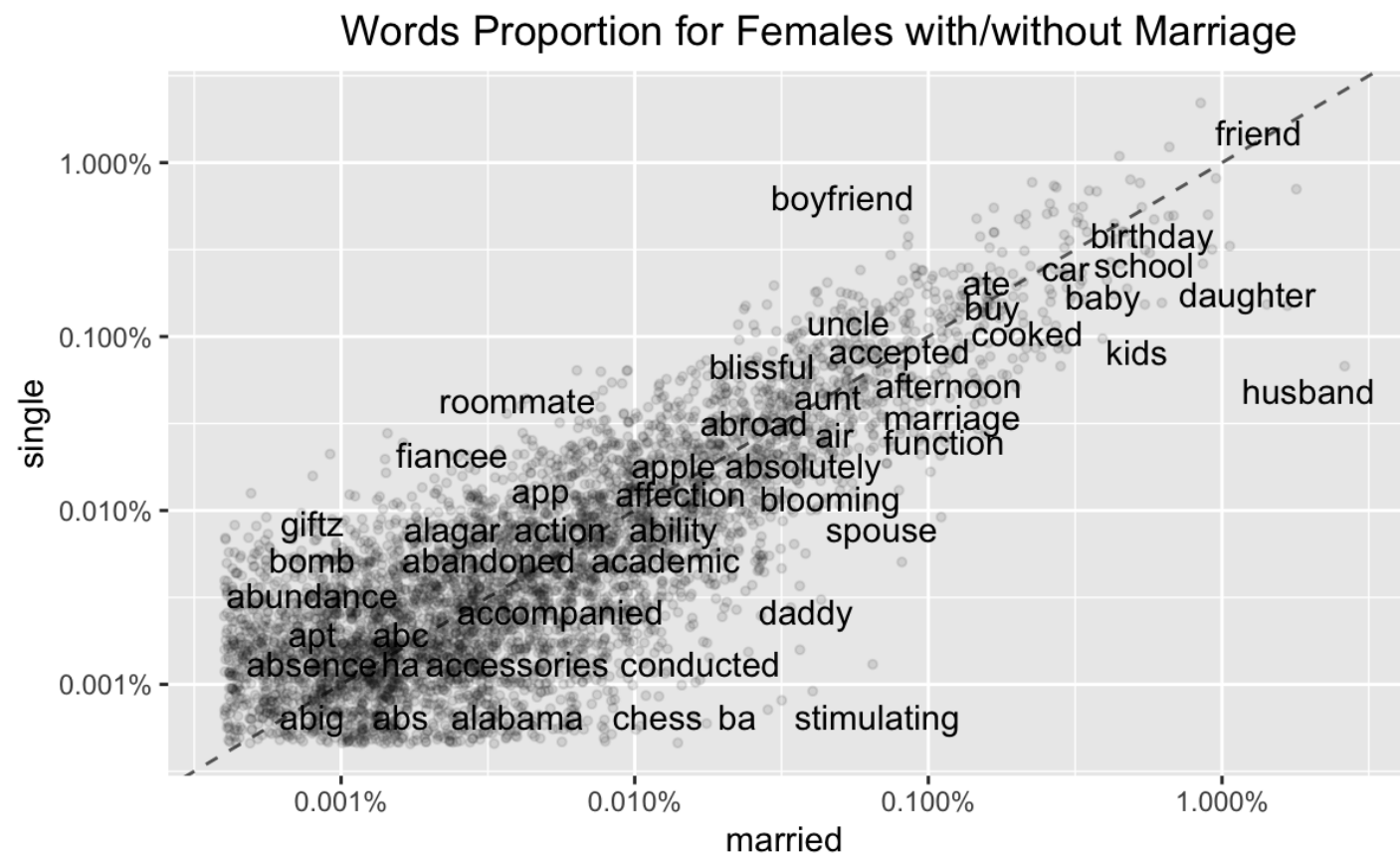
Question 2: What makes men and women happy when they are single? What makes men and women happy after marriage? Do 'happy moments' vary?

After observing the different causes of happiness in men and women in general, I find that they agree in respect of 'friend' and spouses; but women are more inclined to derive their happiness from their families, while men are more inclined to derive it from themselves. So does this result apply to single men and women and married men and women?

The first thing that is certain is the results for single women will be different from the results for women as a whole as described above.

Let's see...

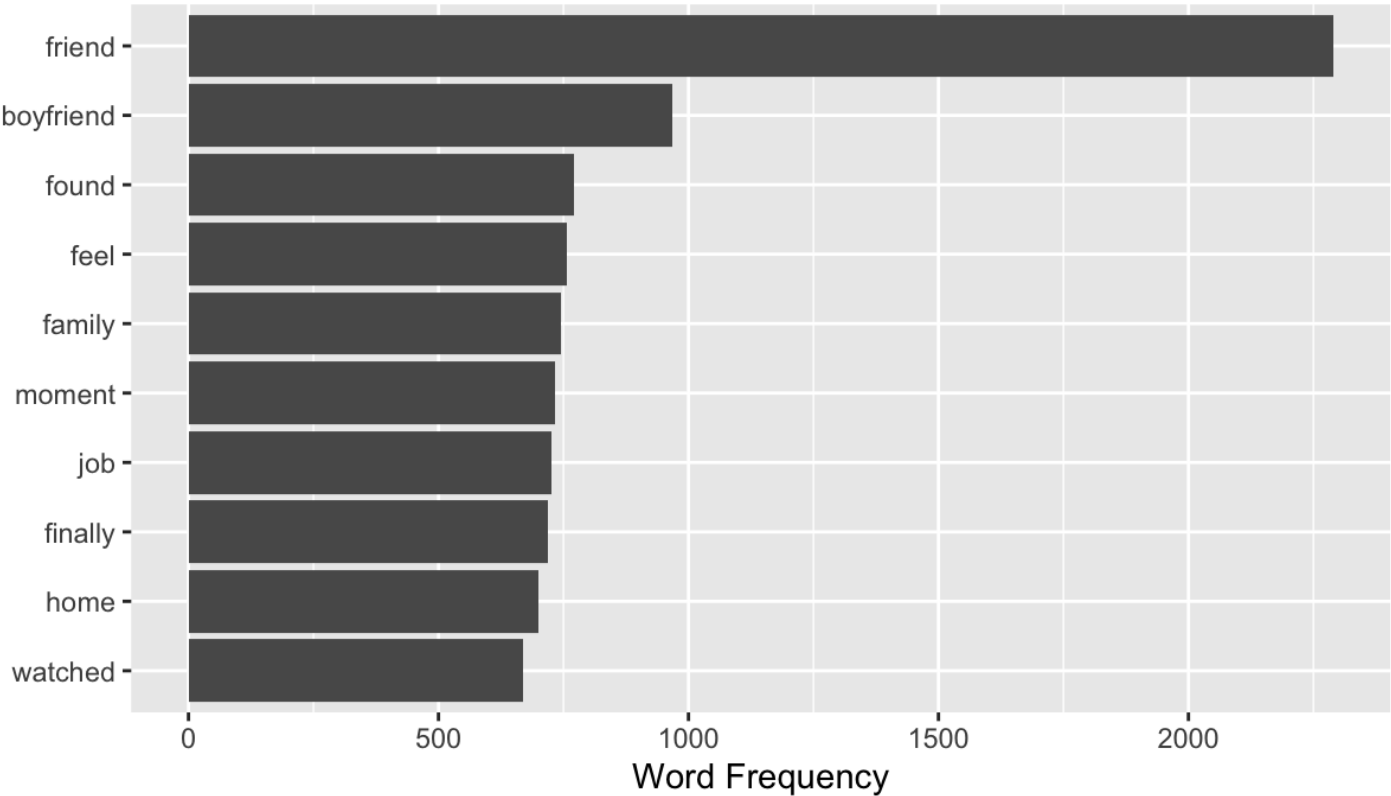
For single and married women



This graph clearly illustrates the differences in what brings happiness to single and married women. Single women’s happiness is frequently associated with terms such as ‘boyfriend’, ‘friend’, ‘school’, ‘roommate’, ‘ate’, and ‘buy’, indicating a focus on personal interests, social connections, and academic pursuits. On the other hand, married women’s happiness is connected with words like ‘husband’, ‘kids’, ‘friend’, ‘cooked’, and ‘marriage’, suggesting a greater emphasis on family life and domestic activities. This distinction underscores the shift in sources of joy from personal and educational achievements for single women to familial and relationship-centered happiness for married women.

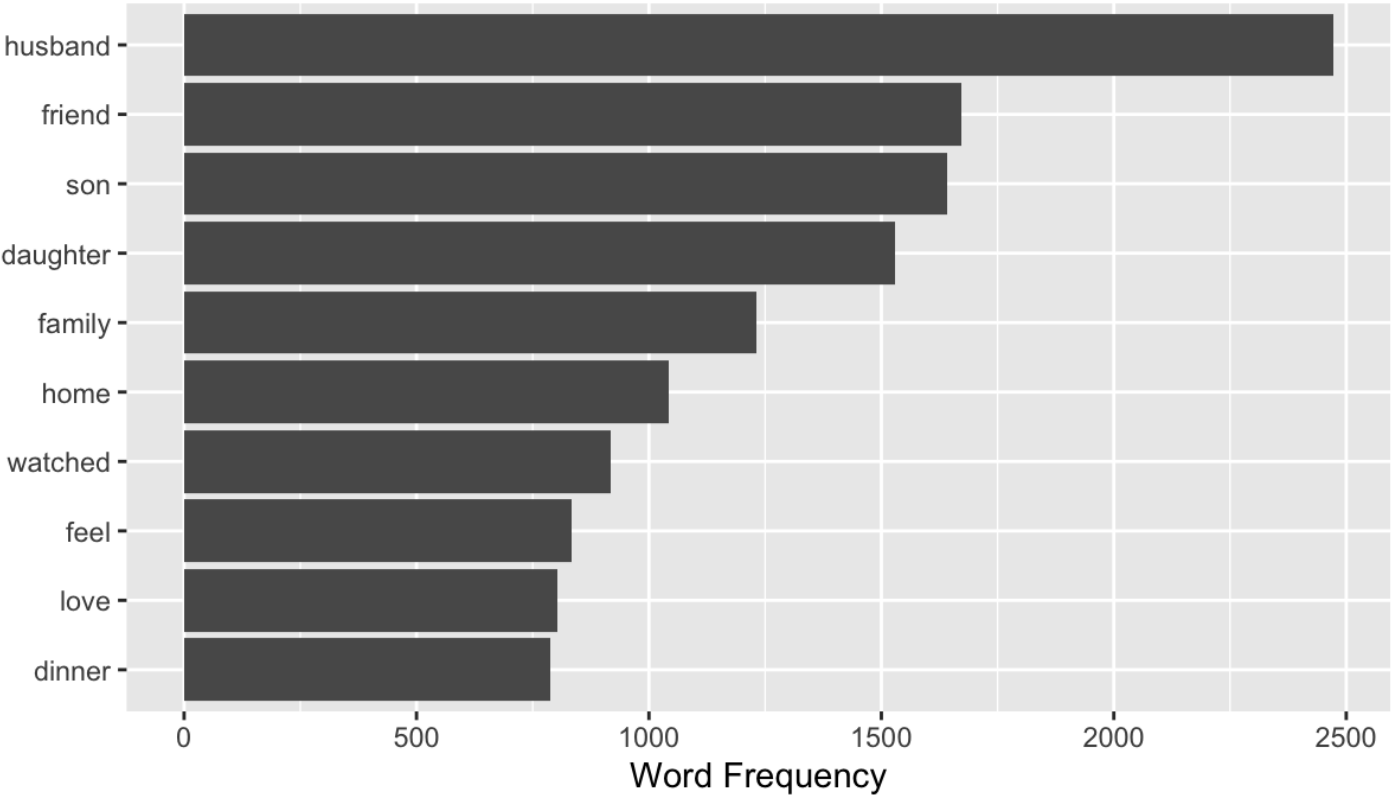
For single women:

Word Frequency in Happy Moments for Single Women



For married women:

Word Frequency in Happy Moments for Married Women



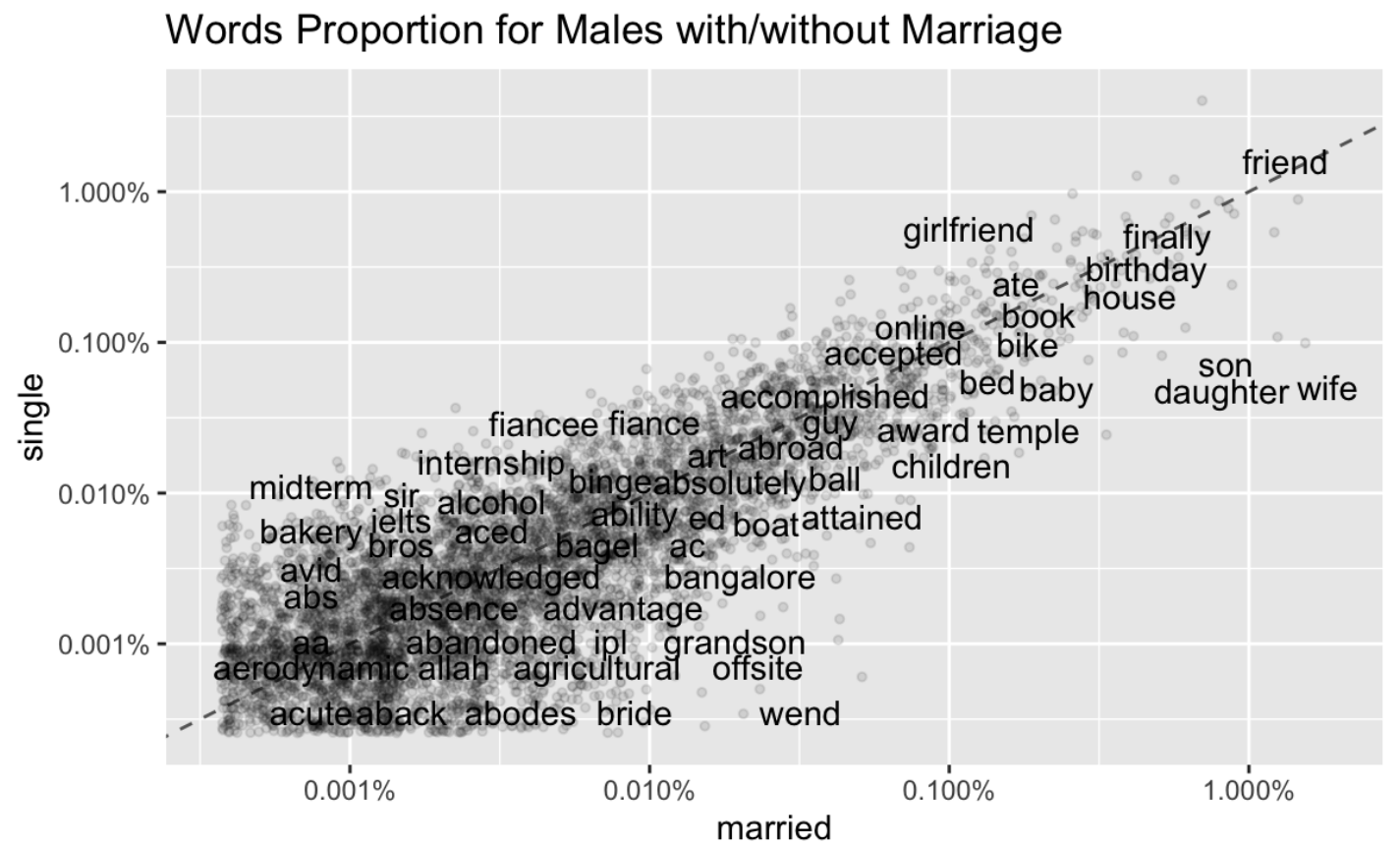
Analyzing the word frequency plot for happy moments among single women, it is evident that the top five words are ‘friend’, ‘boyfriend’, ‘found’, ‘feel’, and ‘family’.

Conversely, the analysis for married women highlights ‘husband’, ‘friend’, ‘son’, ‘daughter’, and ‘family’ as the five most frequent terms.

This comparison reveals a shift in the sources of happiness: while ‘friend’ ranks as the primary source of joy for single women, ‘husband’ assumes the top position for married women, with ‘friend’ moving to second place.

These observations suggest that married women are more likely to derive happiness from their immediate family. In contrast, single women find greater joy in their social life and personal emotions.

For single and married men

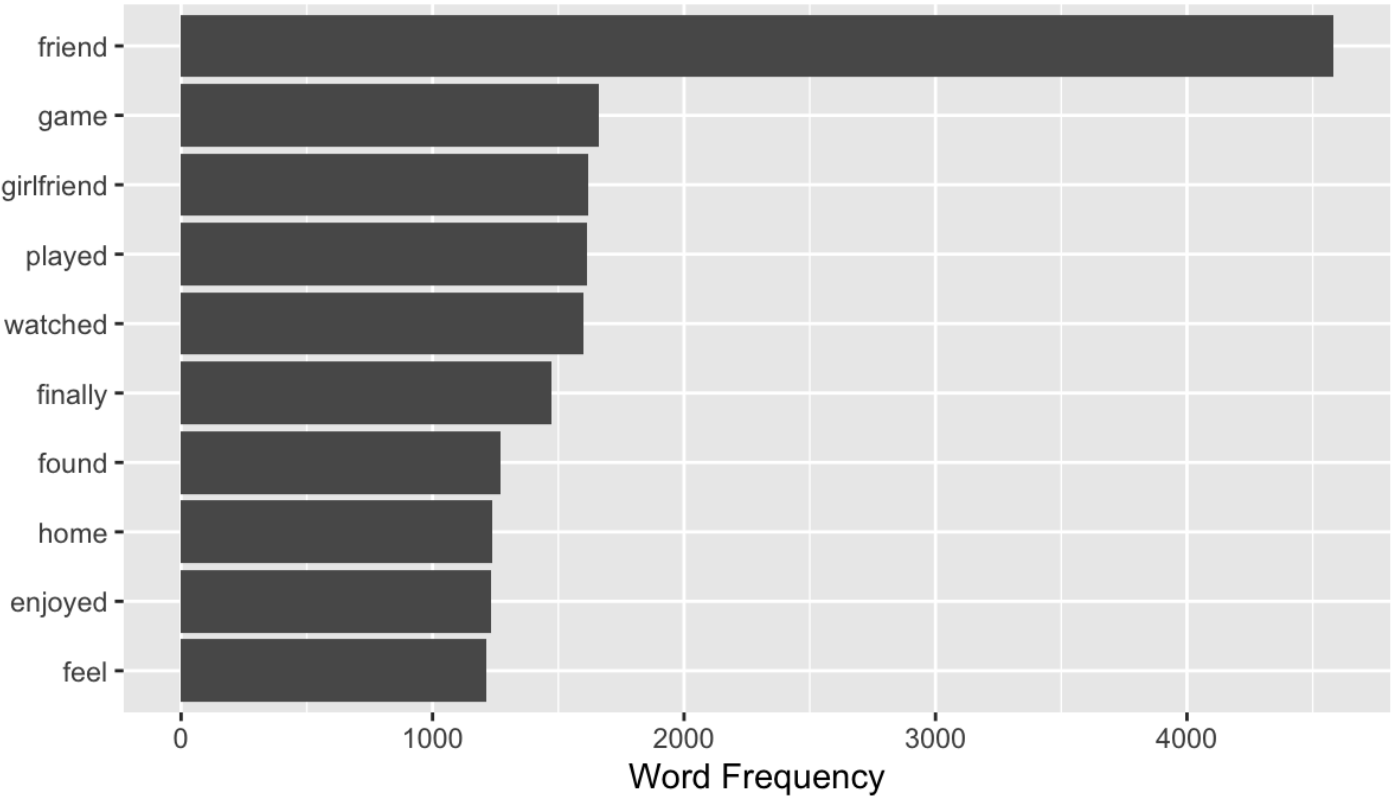


This graph clearly demonstrates that single and married men derive happiness from distinct sources. For single men, prevalent terms such as ‘girlfriend’, ‘friend’, ‘internship’, ‘midterm’, ‘ate’, ‘book’, and similar words suggest a strong emphasis on personal achievements, social connections, and academic pursuits. Meanwhile, married men’s happiness is more frequently associated with words like ‘wife’, ‘children’, ‘friend’, ‘house’, ‘award’, indicating a deeper focus on family life and achievements.

The analysis indicates that single men’s happiness is significantly tied to personal successes, social interactions, and educational accomplishments, whereas married men prioritize family-related joys above all.

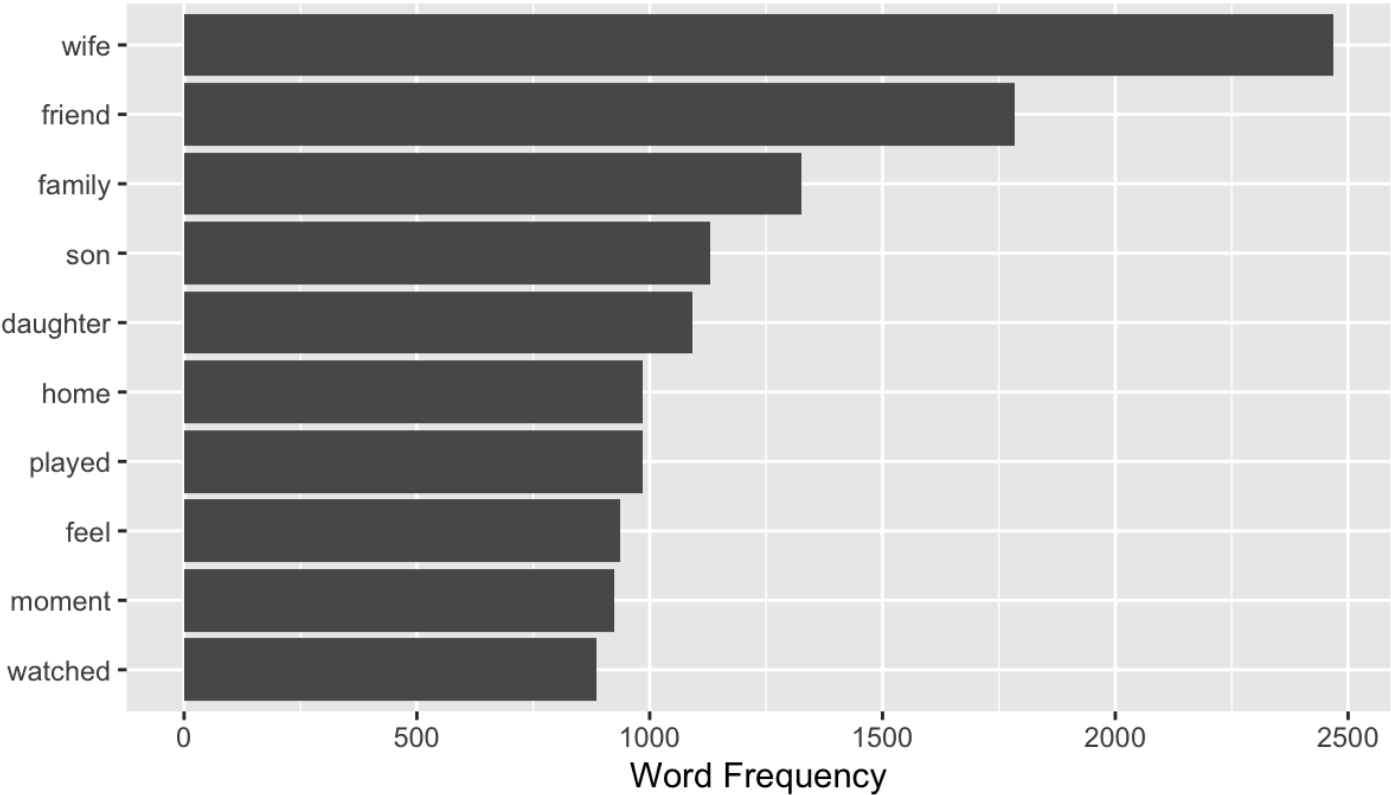
For single men:

Word Frequency in Happy Moments for Single Men



For married men:

Word Frequency in Happy Moments for Married Men



The analysis of the word frequency plot for happy moments among single men reveals that the top five keywords are 'friend', 'game', 'girlfriend', 'played', and 'watched'.

In contrast, for married men, the most frequently occurring words are 'wife', 'friend', 'family', 'son', and 'daughter'.

This comparison clearly indicates a shift in the sources of happiness from personal interests and social activities among single men to a more family-centric joy among married men. The emphasis moves from games and personal entertainment for the single to the joys of family life and relationships for the married.

Based on the analysis of happy moments for single and married individuals, both women and men, I find some differences and similarities in the cause of happiness across these groups:

The Differences in the Cause of Happiness:

1. Gender-specific Interests and Social Connections:

- Single men's happiness is strongly tied to games, personal entertainment, and romantic relationships ('girlfriend'), reflecting a blend of personal hobbies and social connections.
- Single women emphasize broader social connections and achievements ('boyfriend', 'friend', 'school') but also show a strong inclination towards personal feelings and experiences.

The Similarities in the Cause of Happiness:

1. Universal Value of Friendships:

- 'Friend' is a significant source of happiness across all groups, underscoring the universal value of friendships irrespective of marital status or gender. This indicates that social connections remain a constant source of joy.

2. Shift Towards Domestic and Family Life in Marriage:

- Both genders, upon marriage, show a marked shift towards family as a primary source of happiness. This shift reflects a common cultural or personal inclination to focus on familial bonds and responsibilities post-marriage.

3. Continued Importance of Personal Interests:

- Even though the emphasis shifts, personal interests (like 'game' for single men and 'book' for single women) do not disappear entirely upon marriage but may be expressed differently or become integrated into family life.

Conclusion:

These findings highlight the dynamic nature of happiness sources across different life stages and relationship statuses. They reveal a universal trend of moving from individualistic pursuits and social connections to more family-oriented joys upon entering marriage, with friendships remaining a constant source of happiness. Additionally, they underscore the nuanced differences in how men and women, whether single or married, articulate and experience happiness.

The outcomes for single and married men and women diverge significantly from those observed for men and women as a whole above. For instance, it appears that married men derive a greater sense of joy from their familial connections, while single women find more happiness in pursuing personal interests. This observation challenges the initial opinion that women generally prioritize family and child-related happiness, and men prefer solitary pursuits or hobbies. The detailed analysis reveals a more complex picture, indicating that marital status plays a crucial role in shaping the sources of happiness, rather than gender alone.

Question 3: What are the top predicted categories of ‘Happy Moments’ for both women and men, and how do they differ between the two genders? What insights do these differences provide?

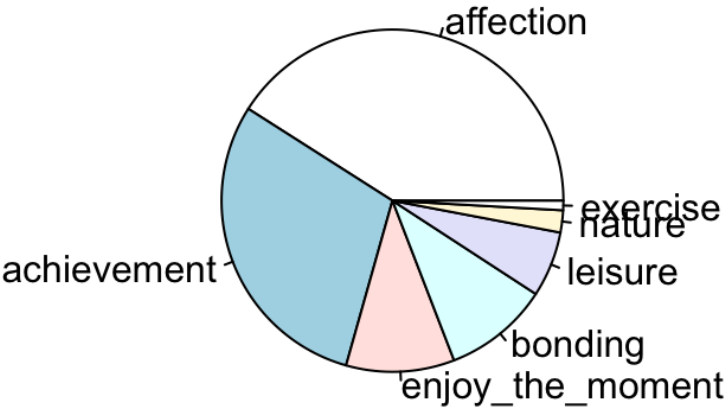
Having examined the variations in sources of happiness across genders as well as between single and married individuals, I am curious to explore whether men and women share similar predicted categories of ‘Happy Moments’.

For female:

hm_data_female_category				
affection	achievement	enjoy_the_moment	bonding	
15864	11476	3966	3854	
leisure	nature	exercise		
2398	809	346		

hm_data_female_category				
affection	achievement	enjoy_the_moment	bonding	
41.0	29.6	10.2	10.0	
leisure	nature	exercise		
6.2	2.1	0.9		

Predicted Categories in Happy Moments for Female

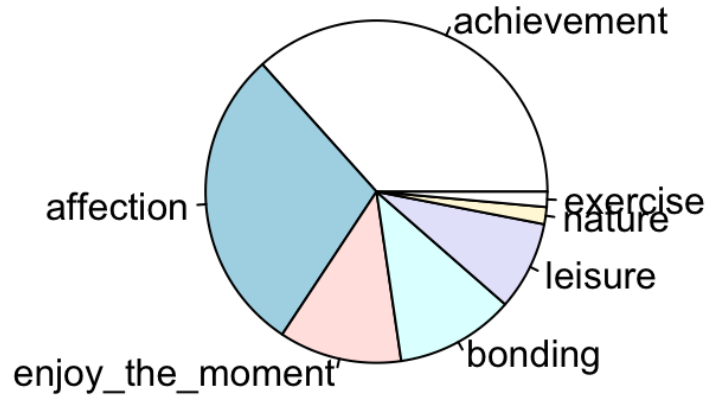


For male:

hm_data_male_category						
achievement	affection	enjoy_the_moment	bonding			
20442	16221	6480	6277			
leisure	nature	exercise				
4680	899	801				

hm_data_male_category						
achievement	affection	enjoy_the_moment	bonding			
36.6	29.1	11.6	11.2			
leisure	nature	exercise				
8.4	1.6	1.4				

Predicted Categories in Happy Moments for Male



Upon analyzing the tables and pie charts detailing the predicted categories of 'Happy Moments' for women, it emerges that 'affection' is the predominant category, followed by 'achievement' and 'enjoy_the_moment', with 'affection' accounting for 41%, 'achievement' approximately 30%, and 'enjoy_the_moment' around 10%. Together, 'affection' and 'achievement' comprise over half of the total reported happy moments.

Similarly, when examining the data for men, 'achievement' stands out as the most common category of 'Happy Moments', succeeded by 'affection' and 'enjoy_the_moment', where 'achievement' represents roughly 37%, 'affection' about 30%, and 'enjoy_the_moment' around 10%. Like with women, 'achievement' and 'affection' together make up more than half of the total.

While both genders predominantly share the same categories—'achievement', 'affection', and 'enjoy_the_moment' — the distribution differs. Women report a higher prevalence of 'affection', at 40%, indicating a significant portion of their happiness is derived from relationships with family and friends. Conversely, men attribute more of their happy moments to 'achievement', also around 40%, highlighting a focus on personal successes.

In conclusion, the data suggests women tend to find more joy in moments of 'affection', such as those spent with family and friends, whereas men place a greater emphasis on 'achievement', reflecting a preference for personal accomplishments.

Step 3: Conclusion

The analysis of 'Happy Moments' across different demographics and relationship statuses within this project reveals insightful patterns about the diverse sources of happiness for men and women, both when single and married. The findings show the complexity and variability in what brings joy to individuals, highlighting several key themes:

- 1. Gender and Relationship Status Influence Happiness Sources:** While both men and women value friendships, indicating the universal appeal of social connections, the primary sources of happiness shift notably between genders and according to relationship status. Single individuals tend to find more joy in personal achievements and social activities, whereas married individuals prioritize family life and relationships. This shift suggests that life stages and relationship commitments significantly shape individuals' happiness.
- 2. Divergent Priorities between Genders:** Women show a stronger inclination towards affection and relational joys, particularly valuing time spent with family and friends. Men, while also valuing affection, place a higher emphasis on achievement, suggesting a preference for personal success and accomplishments as significant sources of happiness. This distinction offers a view of gender-specific happiness factors, with women leaning more towards emotional bonds and men towards personal and external achievements.
- 3. Marriage Transforms Happiness Sources:** The transition from single to married life marks a clear pivot in happiness sources for both men and women. Marriage introduces a broader emphasis on family, underscoring the importance of spousal and parental relationships in contributing to individuals' sense of joy. This change reflects the deep-seated cultural and personal values placed on family as a cornerstone of happiness once in a committed relationship.
- 4. Shared and Unique Categories of Happiness:** The analysis of predicted 'Happy Moments' categories reveals that 'achievement', 'affection', and 'enjoy_the_moment' are common to both men and women, albeit with different rankings in prevalence. This similarity suggests shared values across genders, while the differences in rankings highlight the unique ways in which men and women experience and prioritize happiness.

Summary:

This project illuminates the intricate relationship between gender, marital status, and happiness, revealing that while there are universal themes of joy, the specific sources of happiness are deeply influenced by individual life circumstances and roles. It shows the evolving nature of happiness as individuals transition through life stages, with a notable shift towards family and relational joys in marriage. By acknowledging these patterns, we gain a deeper understanding of the multifaceted concept of happiness and the factors that contribute to it across different segments of society. In general, these insights can inform approaches to improving well-being and fostering happiness in various contexts, recognizing the diversity of experiences and values that shape our lives.