Women in Sport: Role Models

Introduction:

Many youth, especially athletes, look to a famous athlete as a role model. A role model is "a person whose behavior, example, or success is or can be emulated by others, especially by younger people" (Dictionary.com). Therefore, it is important to be mindful of the influence one has and the persona that he or she displays.

Since you, a professional female athlete in particular, are looking to advance in your career, and you aspire to be a role model for many young girls, I analyzed the sentiments and prevalence of what people have said about their role models. By identifying what characteristics are related to choosing female athlete role models, we can act intentionally in displaying these characteristics and having a positive effect on those to come after you.

Methods:

I started by asking people who their female athlete role models are. I gathered responses such as Alex Morgan, Abby Wombach, Serena Williams, Paige Alms, Allyson Felix, Simone Biles, and Megan Rapinoe. From there I found articles on these women which led me to find articles on other women and women who have won awards for their contributions, achievements, leadership, and character. I was able to perform text analysis to identify words that appear most frequently in these articles as well as in the Nike's Dream Crazier commercial. I then pulled tweets from Twitter with the hashtags '#athlete' and '#inspiration' to analyze data from a different audience and compare it to the previously analyzed data.

Analysis:

According to research, children report mixed results of take-aways from famous athletes' actions. Most kids believe that these athletes teach them that "being a good sport and playing fair

are as important as winning," but then they also report seeing athletes use illegal substances, trash talk, play dirty, or have angry outbursts (Miller and Skinner, 2000). The same source reports that women are seen as typically exemplifying better sportsmanship and teamwork. One common comment about Alex Morgan is that she is a great teammate (Aurora, 2019). Morgan and her teammates also take their leadership outside of the pitch which is often noted when choosing her, Abby Wombach, Carli Lloyd, or Megan Rapinoe as role models. They have been a major part of the #EqualPlayEqualPay campaign (Draguca, 2016). In fact, in every article used in my analysis role models were chosen based more on character than their physical ability or winning records.

Of course, as you can see by the word cloud (figure 1) formed from words that show up in the articles with positive and negative sentiments, represented in gray and black respectively, people do look at an athlete's accomplishments and records, but most of the words are related more to the character of the athlete. When citing why someone is a role model or why an athlete deserves an award people look at compassion and how one carries herself ("excited", "consistent", "dominated"). "Healthy" also shows up as a common word which is important for youth to see and emulate. According to Newport Academy, "80 percent of teen girls compare themselves to images they see of celebrities" and "almost half said that celebrity images make them feel dissatisfied with the way they look" (Newport Academy, 2018). Fortunately, role models such as Alex Morgan have campaigned for body positivity. Serena

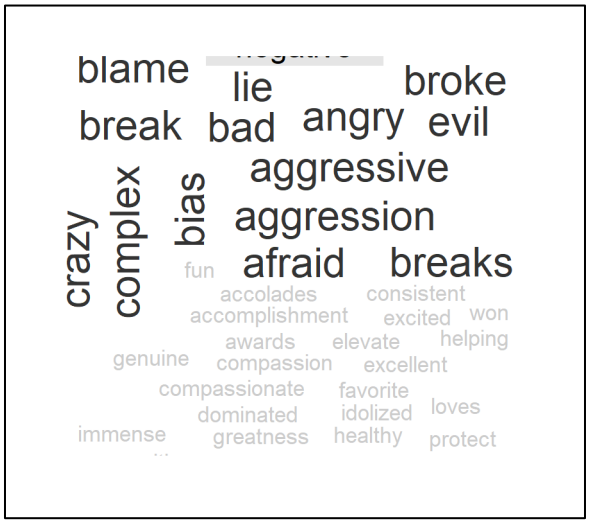


Figure 1: Article Word Cloud

Williams is constantly criticized for her body-build, yet she radiates confidence nonetheless. Allyson Felix spoke on her appreciation of her body and what it could do, teaching girls that they are amazing not for their looks, but by what their bodies are even capable of (Experience Life Team, 2004). This is further supported by the positive sentiments bar chart (figure 2) which includes “strong”, “confidence”, and “commitment.”

Admittedly, some of the words classified as negative are not truly negative in a sports context. Aggression and aggressiveness were found mostly in the articles where soccer (futbol) players were cited as being the role model. When referring to how a soccer player plays, aggression is a positive attribute. Crazy is another word that was used in a positive manner in Nike’s Dream Crazier campaign for women empowerment in sports. Break, breaks, and broke appear in context of breaking records as well as women returning to sport after a major injury or starting a family. In the bar chart (figure 2), you will see fierce as a negatively classified word, but this is actually positive. Words that remain negative are lie, scandal, inequality, and doubt. These are truly negative terms and they remind us that to be seen as a role model, it is important to remain moral, abide by the rules, and stand for something meaningful such as fighting inequality to empower more women.

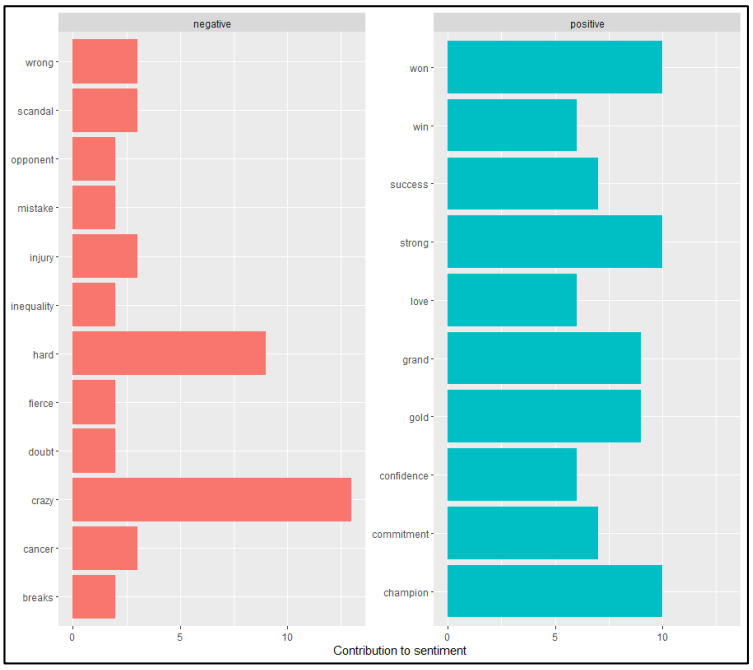


Figure 2: Article Bar Charts

The word cloud based on the Twitter data (figure 3) supports these findings. It includes "trust", "healthy", "confidence", "honor", and "work" in its positive words. In its negative words it has "cheat", "animosity", and "guilt". These again confirm that an athlete's character off the court/pitch/arena are also crucially important. The sentiment bar chart (figure 5) from the twitter data also adds "passion" as an attribute with positive sentiments. The frequency counts (figure 4) from the twitter data show that nutrition, fitness, and motivation are commonly associated with athlete role models.

Figure 3: Twitter Word Cloud

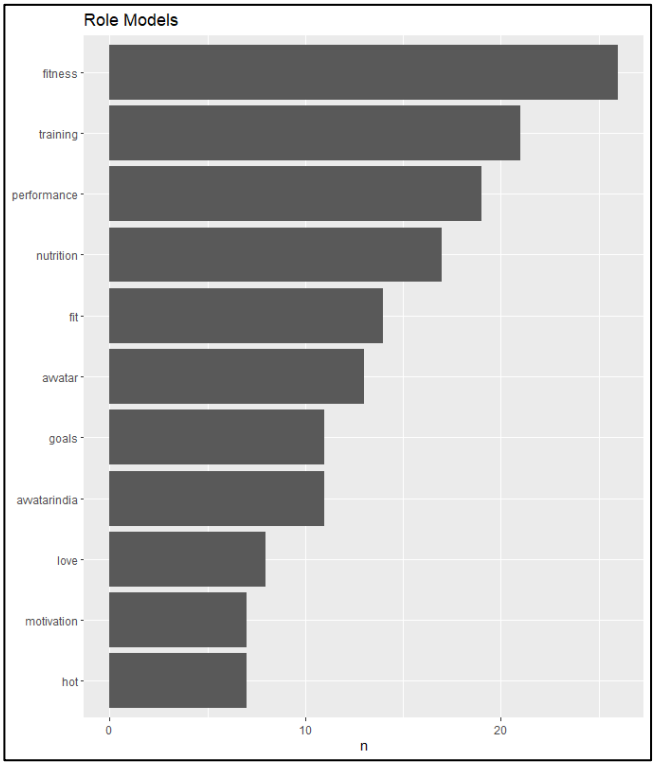


Figure 2: Twitter Frequency Counts

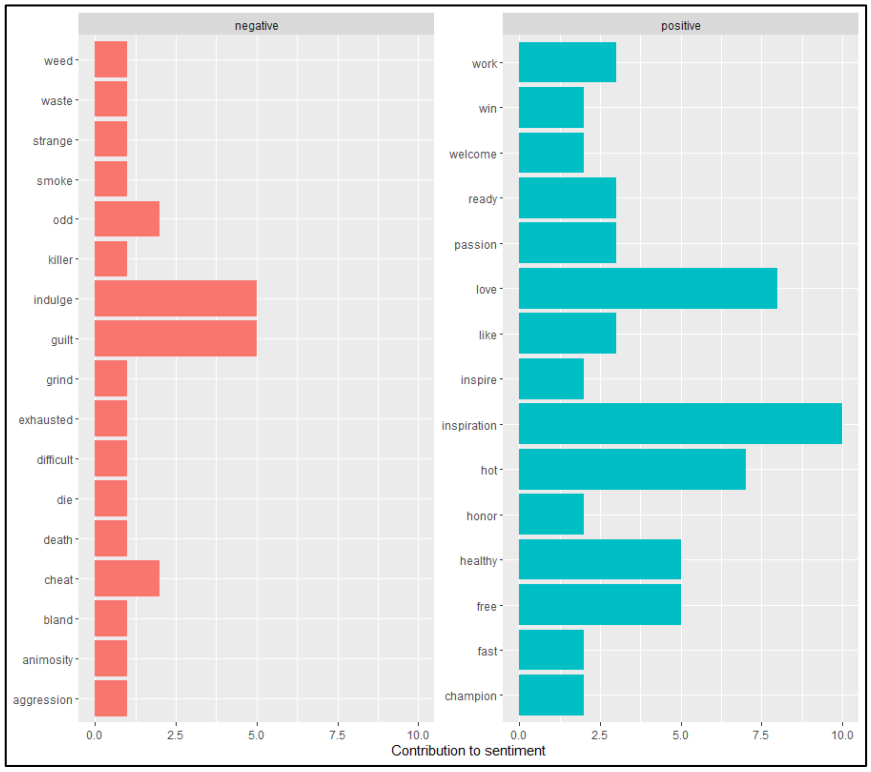


Figure 1: Twitter Sentimen Bar Chart

Recommendation:

Based on these findings, sentiments are largely linked to controllable attributes. If you truly aspire to become young girls' role model, it is essential to obey the law. Do not break the rules. Play your game and continue to work to be the best athlete you can be, because that will be your platform and your way to become more visible. However, to remain in the spotlight with positive sentiments and to be deemed worthy of being a role model, stay vigilant in everything you do. When Abby Wombach got a DUI it was heartbreaking to Americans who felt conflicted about whether they viewed her the same afterwards (Mape, 2016). It is important to encourage body positivity since your audience is very sensitive to that topic. Showing confidence and passion also go a long way.

Ultimately following those guidelines will help to keep you from being cast from young girls' hearts. However, to really become an athlete that a young girl looks up to, we should formulate a platform. Noncontroversial topics are safer such as equal pay: something that many Americans can admit sounds right. If there is something more controversial that you are more passionate about such as LGBTQ rights, we should formulate an intentional platform and plan of action to be one that is respected without being seen as rebellious or ungrateful. Megan Rapinoe and Colin Kaepernik come to mind of having controversial standpoints and various methods of showing their beliefs.

All in all, to be one of the greatest and most memorable, you have to make a memorable impact on and off the court. However, all actions should be intentional and well thought out to have minimal backlash.

Let's meet after you've had a chance to think about what issues you care about and what you would like to have an impact in.

References

"Role Model". N.d. Dictionary.com. Retrieved from https://www.dictionary.com/browse/role-model?s=t

Arenas. (2019, June 28). Women in Sports: Role Models for Life. Retrieved from https://www.employdiversitynetwork.com/blog/2019/6/18/women-in-sports-role-models-for-life-1

Aurora. (2016, August 30). Why Alex Morgan Is The Best Role Model For Students And Student Athletes. Retrieved from https://www.theodysseyonline.com/alex-morgan-role-model-students-student-athletes

Draguca. (2016, August 10). 8 Reasons Alex Morgan Is a Role Model of Olympic Proportions. Retrieved from https://www.instyle.com/awards-events/sports/olympics/reasons-alex-morgan-role-model

Experience Life Team. (2004, April). Full Speed Ahead: Allyson Felix. Retrieved from https://experiencelife.com/article/full-speed-ahead-allyson-felix/

Henneman, K. (2020, February 5). Team USA's Top Female Fencers Discuss Their Athletic Role Models on National Girls & Women in Sports Day. Retrieved from https://www.usafencing.org/news_article/show/1083848-team-usa-s-top-female-fencers-discuss-their-athletic-role-models-on-national-girls-and-women-in-sports-day

Maghersim. (2018, September 12). 7 Reasons Why Serena Williams Still One Of My Role Models. Retrieved from https://www.theodysseyonline.com/serena-williams-role-model

Mape. (2016, April 21). Dear Abby Wambach, Can You Still Be My Child's Role Model After A DUI? Retrieved from https://ilovetowatchyouplay.com/2016/04/21/why-im-okay-if-abby-wambach-is-still-my-childs-role-model-despite-a-dui/

McCarty, P. (2018, April 10). 5 Strong Female Role Models to Inspire the Next Generation. Retrieved from https://breakingmuscle.com/fitness/5-strong-female-role-models-to-inspire-the-next-generation

Miller and Skinner. (2000, October 12). Children Get Mixed Messages from Famous Athletes, Both On and Off the Field. Retrieved from https://www.kff.org/wp-content/uploads/2000/09/3070-nr-national-survey-of-kids-and-their-parents-about-famous-athletes-as-role-models.pdf

Mott, S. (2015, August 9). The 50 most influential women in sport: The full list. Retrieved from https://www.independent.co.uk/sport/the-50-most-influential-women-in-sport-the-full-list-10446935.html

Newport Academy. (2018, August 28). How Celebrities Influence Teens and Why it Matters. Retrieved from https://www.newportacademy.com/resources/empowering-teens/celebrities-influence/

Nike. (2019, February 24). *Dream Crazier* | *Nike* [Video]. Youtube. https://www.youtube.com/watch?v=whpJ19RJ4JY

Smith. (2016, December 19). Serena Williams as a role model. Retrieved from https://theundefeated.com/features/serena-williams-as-role-model/

Team All Good. (2019, February 27). Welcome to the All Good family, Paige Alms! Retrieved from https://allgoodproducts.com/all-good-stories/welcome-to-the-all-good-family-paige-alms/

The Definitive Guide to Female Athlete Role Models in 2020. (2019, May 29). Retrieved from https://www.uplifterinc.com/female-athlete-role-models

Women Leaders in College Sports Awards. Retrieved from

https://www.womenleadersincollegesports.org/WL/For_Members/awards.aspx

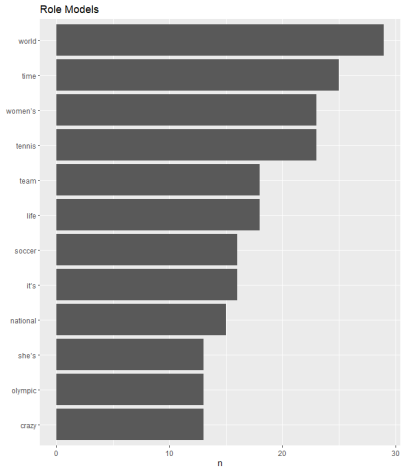
R Code and Output:

```
#######################################################
# Read in libraries
#######################################################

library(dplyr)
library(stringr)
library(tidytext)
library(tidyr)
library(textreadr)
library(ggplot2)
library(tm)
library(twitteR)
library(RColorBrewer)
library(wordcloud)
library(reshape2)

setwd("C:/Users/egome/Documents/Hult/Text Analytics")

#######################################################
# For sentiment analysis
#######################################################

afinn <- get_sentiments("afinn")
nrc <- get_sentiments("nrc")
bing <- get_sentiments("bing")
data("stop_words")

sentiments <- bind_rows(mutate(afinn, lexicon="afinn"),
                        mutate(nrc, lexicon= "nrc"),
                        mutate(bing, lexicon="bing"))

sport_stop <- tibble(
  word = c("women", "woman", "women\'s", "female", "females", "girls", "girl", "sports", "sport", "role",  "model", "models", "athelete",
"athletes",
          "serena", "williams", "alex", "morgan", "abby", "wombach", "athletics", "it\'s"),
  lexicon = rep("SPORTS" , each = 22)
)

#######################################################
# Importing DataFrame
#######################################################

women_sports <- read_document(file="C:/Users/egome/Documents/Hult/Text Analytics/women_in_sports.docx")
women_df <- as.data.frame(women_sports)
women_df <- women_df %>%
                rename(text = women_sports)
women_df$text <- as.character(women_df$text)

#######################################################
# Tokenizing and frequency counts
#######################################################

women_tokens <- women_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(sport_stop) %>%
  count(word, sort=T)

women_tokens
```

```
word        n
  <chr>    <int>
1 world      29
2 time       25
3 tennis     23
4 women’s    23
5 life       18
6 team       18
```

```
 7 it’s       16
 8 soccer     16
 9 national   15
10 crazy      13
# ... with 1,600 more rows
```

```
women_freq <- women_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(sport_stop) %>%
  count(word, sort=TRUE) %>%
  top_n(10) %>%
  mutate(word=reorder(word,n)) %>%
  ggplot(aes(word, n))+
  geom_col()+
  ggtitle("Role Models")+
  xlab(NULL)+
  coord_flip()
print(women_freq)
```




```
#######################################################
# Tokens with sentiment
#######################################################

women_tokens_bing <- women_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(sport_stop) %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=T)

women_tokens_nrc <- women_df %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(sport_stop) %>%
  inner_join(get_sentiments("nrc")) %>%
  count(word, sentiment, sort=T)

women_tokens_bing
```

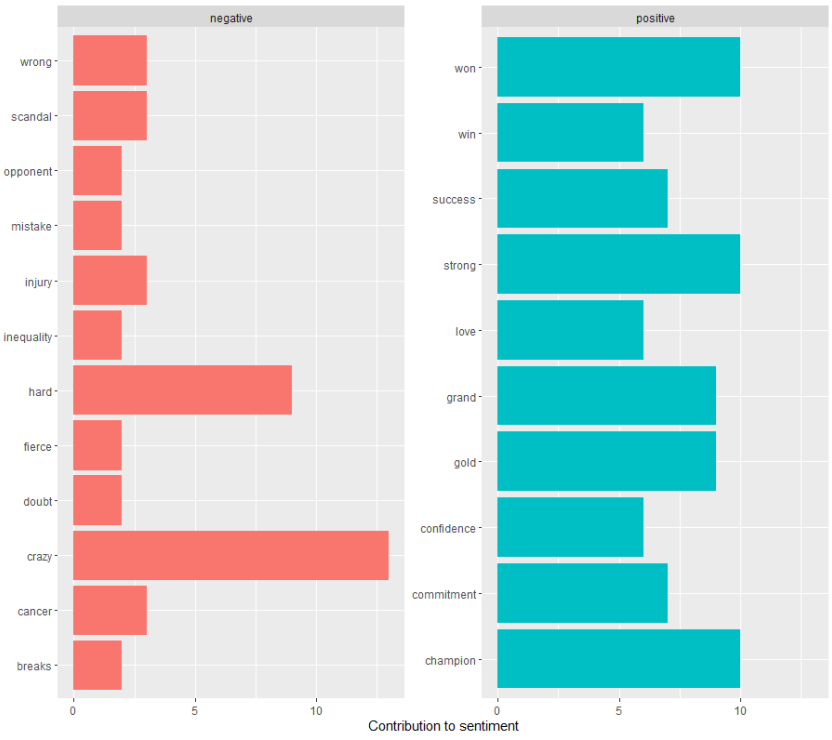
```
# A tibble: 224 x 3
   word      sentiment     n
   <chr>     <chr>     <int>
 1 crazy     negative     13
 2 champion  positive     10
 3 strong    positive     10
 4 won       positive     10
 5 gold      positive      9
 6 grand     positive      9
 7 hard      negative      9
```

```
 8 commitment positive     7
 9 success    positive     7
10 confidence positive     6
# ... with 214 more rows
```

```
women_tokens_nrc
```

```
# A tibble: 896 x 3
   word     sentiment       n
   <chr>    <chr>       <int>
 1 time     anticipation   25
 2 team     trust          18
 3 crazy    anger          13
 4 crazy    fear           13
 5 crazy    negative       13
 6 crazy    sadness        13
 7 athlete  positive       11
 8 champion anticipation   10
 9 champion joy            10
10 champion positive       10
# ... with 886 more rows
```

```
#########################################################
# Bar chart of tokens with sentiment
#########################################################

women_bar_sent <- women_tokens_bing %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  mutate(word=reorder(word, n)) %>%
  ggplot(aes(word, n, fill=sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y="Contribution to sentiment", x=NULL) +
  coord_flip()

women_bar_sent
```

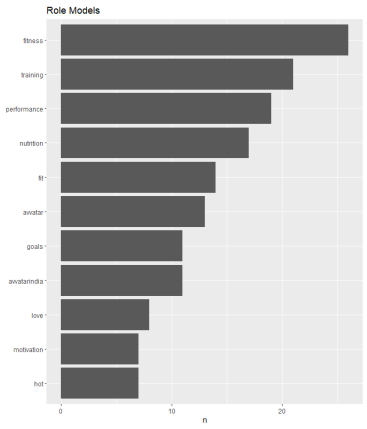



```
#########################################################
# Sentiment Cloud
#########################################################
```

```
women_cloud <- women_tokens %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=TRUE) %>%
  acast(word ~sentiment, value.var="n", fill=0) %>%
  comparison.cloud(colors = c("grey20", "gray80"), max.words=100)
```


```
#########################################################
# Bigram tokenization
#########################################################

women_bigrams <- women_df %>%
  unnest_tokens(bigram, text, token = "ngrams", n=2)

women_bigrams %>%
  count(bigram, sort = TRUE) #this has many stop words, need to remove them

# A tibble: 6,120 x 2
   bigram        n
   <chr>     <int>
 1 of the       38
 2 in the       37
 3 she is       25
 4 is a         24
 5 role model    24
 6 one of       20
 7 on the       18
 8 to be        18
 9 for the      17
10 as a         16
# ... with 6,110 more rows

bigrams_separated <- women_bigrams %>%
  separate(bigram, c("word1", "word2"), sep = " ")

bigrams_filtered <- bigrams_separated %>%
  filter(!word1 %in% stop_words$word) %>%
  filter(!word2 %in% stop_words$word) # filtering out sport_stop loses too much information

#creating the new bigram, "no-stop-words":
bigram_counts <- bigrams_filtered %>%
  count(word1, word2, sort = TRUE)
#want to see the new bigrams
bigram_counts

# A tibble: 1,104 x 3
   word1    word2       n
   <chr>    <chr>    <int>
```

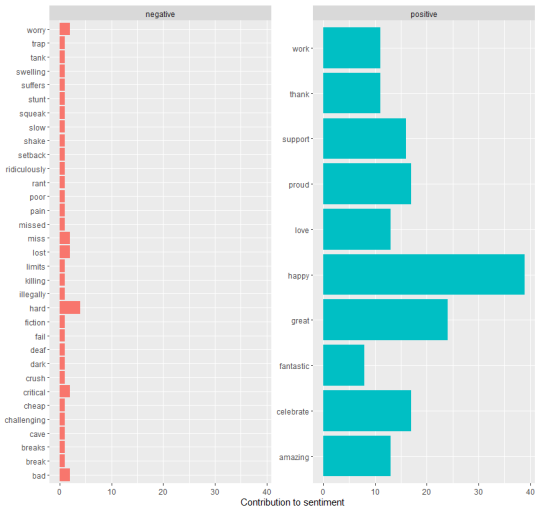
```
 1 role     model      24
 2 serena   williams   10
 3 female   athletes    9
 4 world    cup         9
 5 grand    slam        6
 6 women's  sports      6
 7 alex     morgan      5
 8 gold     medals      5
 9 national team        5
10 olympic  gold        5
# ... with 1,094 more rows

#################################################################################################
# Validating with Twitter
#################################################################################################


twit_stop <- tibble(
  word = c("t.co", "https", "rt", "http", "nextlevel", "athlete", "awatar", "boxing",
           "jarod_dag", "awatarindia", "swimbikerun", "triathlon", "inspiration", "athlete"),
  lexicon = rep("TWITTER" , each = 14)
)

consumer_key <- '2ogiEdY3HQ5LFEYi8X6hcweBZ'
consumer_secret <- 'rCT3WPSW46qPgb6dybVmMIJ1vb387oCuER26pkWzPTSAYVN4pe'
access_token <- '938787020993892353-7XERuUx1ZURkdXTFVYiQBd4VGgHeoIE'
access_secret <- 'edVxvbg1G2o9gYDAq3JC3WF2fb7ApTreoNUEQXPEHLH2s'

setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
Wins <- twitteR::searchTwitter('#womeninsport', n = 1000, since = '2011-06-01', retryOnRateLimit = 1e3)
w_twit = twitteR::twListToDF(Wins)

w_twit_filt <- w_twit %>% # removing retweeted data
  filter(w_twit$isRetweet == 'FALSE')

######## second try ##########

w_role <- twitteR::searchTwitter('#athlete + #inspiration', n = 1000, since = '2011-06-01', retryOnRateLimit = 1e3)
twit_role = twitteR::twListToDF(w_role) # keeping retweets, not limited to femaleathletes

#######################################################
# Tiwtter tokenization and frequency counts
#######################################################

twitter_freq <- twit_role %>%
  unnest_tokens(word, text) %>%
  anti_join(stop_words) %>%
  anti_join(twit_stop) %>%
  count(word, sort=TRUE) %>%
  top_n(10) %>%
  mutate(word=reorder(word,n)) %>%
  ggplot(aes(word, n))+
  geom_col()+
  ggtitle("Role Models")+
  xlab(NULL)+
  coord_flip()
print(twitter_freq)
```




```
#########################################################
# Tiwtter sentiment analysis
#########################################################

women_twit_token <- w_twit %>%
  unnest_tokens(word, text) %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=T) %>%
  ungroup()

women_twit_token # not insightful

# A tibble: 175 x 3
   word      sentiment     n
   <chr>     <chr>     <int>
 1 celebrate positive     77
 2 happy     positive     75
 3 positive  positive     66
 4 support   positive     61
 5 great     positive     48
 6 proud     positive     43
 7 thank     positive     40
 8 love      positive     33
 9 better    positive     32
10 amazing   positive     28
# ... with 165 more rows

women_twit_bar <- women_twit_token %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  mutate(word=reorder(word, n)) %>%
  ggplot(aes(word, n, fill=sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y="Contribution to sentiment", x=NULL) +
  coord_flip()

women_twit_bar # not insightful


# Filtering out retweets

women_twit_filt_token <- w_twit_filt %>%
  unnest_tokens(word, text) %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=T) %>%
  ungroup()

women_twit_filt_token # not insightful

# A tibble: 165 x 3
   word      sentiment     n
```

```
  <chr>     <chr>     <int>
 1 happy     positive     39
 2 great     positive     24
 3 celebrate positive     17
 4 proud     positive     17
 5 support   positive     16
 6 amazing   positive     13
 7 love      positive     13
 8 thank     positive     11
 9 work      positive     11
10 fantastic positive      8
# ... with 155 more rows
```

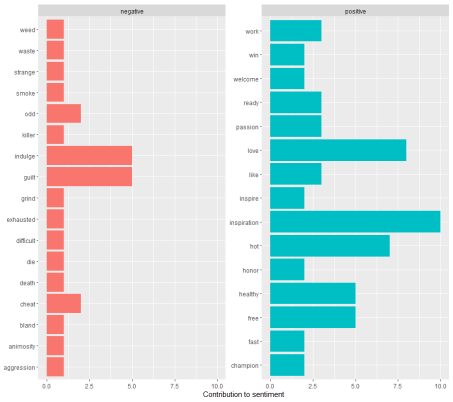
```
women_twit_filt_bar <- women_twit_filt_token %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  mutate(word=reorder(word, n)) %>%
  ggplot(aes(word, n, fill=sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y="Contribution to sentiment", x=NULL) +
  coord_flip()

women_twit_filt_bar # not insightful
```




```
##### second try ##############

twit_role_token <- twit_role %>%
  unnest_tokens(word, text) %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=T) %>%
  ungroup()

twit_role_token
```

```
# A tibble: 43 x 3
   word       sentiment     n
   <chr>      <chr>     <int>
 1 inspiration positive     10
 2 love       positive      8
 3 hot        positive      7
 4 free       positive      5
 5 guilt      negative      5
 6 healthy    positive      5
 7 indulge    negative      5
 8 like       positive      3
 9 passion    positive      3
10 ready      positive      3
```

```
# ... with 33 more rows

twit_role_bar <- twit_role_token %>%
  group_by(sentiment) %>%
  top_n(10) %>%
  mutate(word=reorder(word, n)) %>%
  ggplot(aes(word, n, fill=sentiment)) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~sentiment, scales = "free_y") +
  labs(y="Contribution to sentiment", x=NULL) +
  coord_flip()

twit_role_bar # much more insight
```




```
##########################################################
# Twitter word cloud
##########################################################

twit_cloud <- twit_role_token %>%
  inner_join(get_sentiments("bing")) %>%
  count(word, sentiment, sort=TRUE) %>%
  acast(word ~sentiment, value.var="n", fill=0) %>%
  comparison.cloud(colors = c("grey20", "gray80"), max.words=100)
```