SRV_Analysis

Justin A. Gould

July 15, 2018

Required Packages

```
#install.packages("tm")
#install.packages("topicmodels")
#install.packages("sentimentr")
#install.packages("wordcloud")
#install.packages("sqldf")
#install.packages("SnowballC")
#install.packages("RColorBrewer")
#install.packages("ggplot2")
library(tm)
```

```
## Warning: package 'tm' was built under R version 3.4.4
```

```
## Loading required package: NLP
```

```
library(topicmodels)
```

```
## Warning: package 'topicmodels' was built under R version 3.4.4
```

```
library(sentimentr)
```

```
## Warning: package 'sentimentr' was built under R version 3.4.4
```

```
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 3.4.4
```

```
## Loading required package: RColorBrewer
```

```
library(sqldf)
```

```
## Warning: package 'sqldf' was built under R version 3.4.4
```

```
## Loading required package: gsubfn
```

```
## Warning: package 'gsubfn' was built under R version 3.4.4
```

```
## Loading required package: proto
```

```
## Warning: package 'proto' was built under R version 3.4.4
```

```
## Loading required package: RSQLite
```

```
## Warning: package 'RSQLite' was built under R version 3.4.4
```

```
library(SnowballC)
library(RColorBrewer)
library(ggplot2)
```

```
##
## Attaching package: 'ggplot2'
```

```
## The following object is masked from 'package:NLP':
##
##     annotate
```

Topic Modeling

```
#Loading the individual song txt files into corpus
setwd("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R")
filenames = list.files(getwd(),pattern="*.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Creating the document-term matrix
#Remove punctuation
documents = tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Text to lower case
documents = tm_map(documents,content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Eliminate digits
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove stopwords from standard stopword list
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Create document-term matrix
dtm = DocumentTermMatrix(documents)
rownames(dtm) = filenames
dtm
```

```
## <<DocumentTermMatrix (documents: 5, terms: 178)>>
## Non-/sparse entries: 206/684
## Sparsity           : 77%
## Maximal term length: 12
## Weighting          : term frequency (tf)
```

```
#Run LDA to find topics in text (since I am using top 5 songs, I will use 5 topics)
burnin = 1000
iter = 2000
thin = 500
nstart = 5
seed = list(365783, 210, 233998, 256730148, 3)
best = TRUE
k = 5
ldaOut = LDA(dtm, k, method = "Gibbs", control = list(nstart = nstart, seed = seed, best = best,
 burnin = burnin, iter = iter, thin = thin))
```

Reviewing the Topics from the LDA

```
#Review the top 4 terms for each of the 5 topics from LDA
terms(ldaOut,4)
```

```
##      Topic 1    Topic 2  Topic 3    Topic 4  Topic 5
## [1,] "hey"      "baby"   "yeah"     "little" "well"
## [2,] "ive"      "back"   "day"      "shes"   "ever"
## [3,] "roughest" "clouds" "stay"     "sweet"  "lord"
## [4,] "wine"     "man"    "anything" "love"   "cwell"
```

Assigning Topics to Documents

```
topics(ldaOut)
```

```
##         littleWing.txt maryHadALittleLamb.txt        prideAndJoy.txt
##                      3                      3                      4
##         texasFlood.txt        tinPanAlley.txt
##                      2                      1
```

From a high level, the topics reflect:

Topic #1: With the use of the exclamation “hey,” as well as a reference to alcohol and “roughest,” the topic seems to be negative in nature.

Topic #2: SRV uses “baby” quite often, in numerous contexts. For example, in live versions of “Lenny,” despite having no lyrics, he explains that he wrote the song for his wife (his “baby”), Lenny. However, SRV also uses the word similar to “hey” and “yeah” – as an exclamation. In fact, 50% of the words in topic 2 are exclamation, suggesting that this topic is neither positive, nor negative, and that these songs are more bluesy, due to these exclamations.

Topic #3: Topic 3 has an exclamation; however, with the use of “stay” and “anything,” considering the rather melancholic mood of many of the songs from SRV’s first album, which makes up $\frac{3}{5}$ of the songs used, the implication is someone or something may be leaving, and he is willing to do “anything” to get them/it to “stay.”

Topic #4: Topic 4, with the numerous references to romance, is positive. Although these words primarily come from the song “Pride and Joy,” which is interesting, as this is the only song associated with that topic, topic 4 shows the romantic, amorous side of SRV and his music.

Topic #5: These words really do not provide much meaning and insight; since none of the 5 songs have this topic, it will be ignored.

Sentiment Analysis

```
#Loading the csv containing all lyrics, aggregated at sentence-level, by song title
lyrics = read.csv("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\ly
rics.csv", header = T, sep = ",")
lyrics$title = as.character(lyrics$title)
lyrics$text = as.character(lyrics$text)
lyrics$rowNum = 1:nrow(lyrics)

#Running the sentiment analysis
textSen = data.frame(sentiment(lyrics$text))

#JOINing the sentiment scores ON the lyrics data.frame
lyrics = sqldf("SELECT lyrics.*, textSen.word_count AS word_count, textSen.sentiment AS sentimen
t FROM lyrics, textSen WHERE lyrics.rowNum = textSen.element_id")
lyrics$rowNum = NULL

#Average sentiment score aggregated by song
avgScore = sqldf("SELECT title, avg(sentiment) AS avgSentiment, sum(word_count) AS total_words F
ROM lyrics GROUP BY title")
avgScore
```

```
##                    title avgSentiment total_words
## 1            Little_Wing   0.03186577          71
## 2 Mary_Had_A_Little_Lamb   0.05993735          64
## 3          Pride_And_Joy   0.30540653         185
## 4            Texas_Flood   0.15878487         123
## 5          Tin_Pan_Alley  -0.04181070         185
```

```
lyrics
```

```
##                    title
## 1             Little_Wing
## 2             Little_Wing
## 3             Little_Wing
## 4             Little_Wing
## 5             Little_Wing
## 6             Little_Wing
## 7             Little_Wing
## 8             Little_Wing
## 9             Little_Wing
## 10            Little_Wing
## 11            Little_Wing
## 12            Little_Wing
## 13            Little_Wing
## 14 Mary_Had_A_Little_Lamb
## 15 Mary_Had_A_Little_Lamb
## 16 Mary_Had_A_Little_Lamb
## 17 Mary_Had_A_Little_Lamb
## 18 Mary_Had_A_Little_Lamb
## 19 Mary_Had_A_Little_Lamb
## 20 Mary_Had_A_Little_Lamb
## 21 Mary_Had_A_Little_Lamb
## 22 Mary_Had_A_Little_Lamb
## 23 Mary_Had_A_Little_Lamb
## 24 Mary_Had_A_Little_Lamb
## 25 Mary_Had_A_Little_Lamb
## 26          Pride_And_Joy
## 27          Pride_And_Joy
## 28          Pride_And_Joy
## 29          Pride_And_Joy
## 30          Pride_And_Joy
## 31          Pride_And_Joy
## 32          Pride_And_Joy
## 33          Pride_And_Joy
## 34          Pride_And_Joy
## 35          Pride_And_Joy
## 36          Pride_And_Joy
## 37          Pride_And_Joy
## 38          Pride_And_Joy
## 39          Pride_And_Joy
## 40          Pride_And_Joy
## 41          Pride_And_Joy
## 42          Pride_And_Joy
## 43          Pride_And_Joy
## 44          Pride_And_Joy
## 45          Pride_And_Joy
## 46            Texas_Flood
## 47            Texas_Flood
## 48            Texas_Flood
## 49            Texas_Flood
## 50            Texas_Flood
## 51            Texas_Flood
## 52            Texas_Flood
```

```
## 53            Texas_Flood
## 54            Texas_Flood
## 55            Texas_Flood
## 56            Texas_Flood
## 57            Texas_Flood
## 58            Texas_Flood
## 59            Texas_Flood
## 60            Texas_Flood
## 61            Texas_Flood
## 62            Texas_Flood
## 63            Texas_Flood
## 64          Tin_Pan_Alley
## 65          Tin_Pan_Alley
## 66          Tin_Pan_Alley
## 67          Tin_Pan_Alley
## 68          Tin_Pan_Alley
## 69          Tin_Pan_Alley
## 70          Tin_Pan_Alley
## 71          Tin_Pan_Alley
## 72          Tin_Pan_Alley
## 73          Tin_Pan_Alley
## 74          Tin_Pan_Alley
## 75          Tin_Pan_Alley
## 76          Tin_Pan_Alley
## 77          Tin_Pan_Alley
## 78          Tin_Pan_Alley
## 79          Tin_Pan_Alley
## 80          Tin_Pan_Alley
## 81          Tin_Pan_Alley
## 82          Tin_Pan_Alley
## 83          Tin_Pan_Alley
## 84          Tin_Pan_Alley
## 85          Tin_Pan_Alley
## 86          Tin_Pan_Alley
## 87          Tin_Pan_Alley
##                                                          text
## 1                                Well she's walking through the clouds.
## 2                              With a circus mind that's running wild.
## 3                                          Butterflies and zebras.
## 4                                  And moonbeams and a fairy tale.
## 5                                That's all she ever thinks about.
## 6                                              Riding the wind.
## 7                                   When I'm sad she comes to me.
## 8                          With a thousand smiles she gives to me free.
## 9                                  It's alright she says, it's alright.
## 10                             Take anything you want from me Anything.
## 11                                         Fly on little wing.
## 12                                           Yeah, yeah, yeah.
## 13                                         Fly on little wing.
## 14                                      Mary had a little lamb.
## 15                                   His fleece was white as snow.
## 16                                     Everywhere that child went.
## 17                                 Your little lamb was sure to go.
## 18                               He followed her to school one day.
```

```
## 19                                        And broke the teacher's rule.
## 20                                           What a time did they have.
## 21                                                  That day at school.
## 22                                                  Tisket tasket baby.
## 23                                           A green and yellow basket.
## 24                                            Sent a letter to my baby.
## 25                                               On my way I passed it.
## 26                Well you've heard about love givin' sight to the blind.
## 27                                 My baby's lovin' cause the sun to shine.
## 28                   She's my sweet little thang, she's my pride and joy.
## 29                 She's my sweet little baby, I'm her little lover boy.
## 30                                  Yeah I love my baby, heart and soul.
## 31                                  Love like ours won't never grow old.
## 32                   She's my sweet little thang, she's my pride and joy.
## 33                 She's my sweet little baby, I'm her little lover boy.
## 34                              Yeah I love my lady, she's long and lean.
## 35                      You mess with her, you'll see a man get mean.
## 36                   She's my sweet little thang, she's my pride and joy.
## 37                 She's my sweet little baby, I'm her little lover boy.
## 38                            Well I love my baby, like the finest wine.
## 39                                 Stick with her until the end of time.
## 40                   She's my sweet little thang, she's my pride and joy.
## 41                 She's my sweet little baby, I'm her little lover boy.
## 42                               Yeah I love my baby, my heart and soul.
## 43                                  Love like ours won't never grow old.
## 44                   She's my sweet little thang, she's my pride and joy.
## 45                 She's my sweet little baby, I'm her little lover boy.
## 46                                  Well there's floodin' down in Texas.
## 47                                  All of the telephone lines are down.
## 48                                  Well there's floodin' down in Texas.
## 49                                  All of the telephone lines are down.
## 50                                 And I've been tryin' to call my baby.
## 51                                  Lord and I can't get a single sound.
## 52                                      Well dark clouds are rollin' in.
## 53                                    Man I'm standin' out in the rain.
## 54                                      Well dark clouds are rollin' in.
## 55                                    Man I'm standin' out in the rain.
## 56                                      Yeah flood water keep a rollin'.
## 57                                Man it's about to drive poor me insane.
## 58                                           Well I'm leavin' you baby.
## 59                                 Lord and I'm goin' back home to stay.
## 60                                           Well I'm leavin' you baby.
## 61                                 Lord and I'm goin' back home to stay.
## 62                               Well back home are no floods or tornados.
## 63                                     Baby and the sun shines every day.
## 64                                            Went down to Tin Pan Alley.
## 65                                                 See what was going on.
## 66                                       Things was too hot down there.
## 67                                             Couldn't stay very long.
## 68            Hey hey hey hey, Alley's the roughest place I've ever been.
## 69     All the people down there Livin' for their whiskey, wine, and gin.
## 70                                              I heard a woman scream.
## 71                                   Yeah and I peeped through the door.
## 72           Some cat was workin' on Annie with a Lord with a two-by-four.
```

```
## 73             Hey hey hey hey, Alley's the roughest place I've ever been.
## 74                                                 All the people down there.
## 75                                Livin' for their whiskey, wine, and gin.
## 76                                                    I heard a pistol shoot.
## 77                                                     Yeah and it was a .44.
## 78 Somebody killed a crap shooter cause he didn't Shake, rattle, and roll.
## 79             Hey hey hey hey, Alley's the roughest place I've ever been.
## 80                                                 All the people down there.
## 81                               Killin' for their whiskye, wine, and gin.
## 82                                                     I saw a cop standing.
## 83                                                   With his hand on his gun.
## 84                               He said "this is a raid, boy Nobody run".
## 85             Hey hey hey hey, Alley's the roughest place I've ever been.
## 86                                       Yeah they took me away from Alley.
## 87                                Lord they took me right back to the pen.
##    word_count   sentiment
## 1           6  0.32659863
## 2           7 -0.05669467
## 3           3  0.00000000
## 4           6  0.00000000
## 5           6  0.00000000
## 6           3  0.00000000
## 7           7 -0.18898224
## 8           9  0.33333333
## 9           6  0.00000000
## 10          7  0.00000000
## 11          4  0.00000000
## 12          3  0.00000000
## 13          4  0.00000000
## 14          5  0.00000000
## 15          6  0.16329932
## 16          4  0.30000000
## 17          7  0.00000000
## 18          7  0.00000000
## 19          5 -0.33541020
## 20          6  0.00000000
## 21          4  0.00000000
## 22          3  0.34641016
## 23          5  0.00000000
## 24          6  0.24494897
## 25          6  0.00000000
## 26         10  0.25298221
## 27          8  0.38890873
## 28         10  0.36366193
## 29         10  0.13281566
## 30          8  0.47729708
## 31          7  0.39686270
## 32         10  0.36366193
## 33         10  0.13281566
## 34          9  0.45000000
## 35         10 -0.23717082
## 36         10  0.36366193
## 37         10  0.13281566
## 38          9  1.05000000
```

```
## 39          8  0.00000000
## 40         10  0.36366193
## 41         10  0.13281566
## 42          9  0.45000000
## 43          7  0.39686270
## 44         10  0.36366193
## 45         10  0.13281566
## 46          6  0.32659863
## 47          7  0.00000000
## 48          6  0.32659863
## 49          7  0.00000000
## 50          8  0.21213203
## 51          8  0.00000000
## 52          6  0.08164966
## 53          7  0.00000000
## 54          6  0.08164966
## 55          7  0.00000000
## 56          6  0.00000000
## 57          8 -0.44194174
## 58          5  0.62609903
## 59          8  0.00000000
## 60          5  0.62609903
## 61          8  0.00000000
## 62          8  0.56568542
## 63          7  0.45355737
## 64          6 -0.40824829
## 65          5  0.00000000
## 66          6 -0.10206207
## 67          4  0.00000000
## 68         11  0.00000000
## 69         12  0.00000000
## 70          5 -0.44721360
## 71          7  0.00000000
## 72         14  0.00000000
## 73         11  0.00000000
## 74          5  0.00000000
## 75          7  0.00000000
## 76          5 -0.40249224
## 77          5  0.00000000
## 78         12  0.57735027
## 79         11  0.00000000
## 80          5  0.00000000
## 81          7  0.00000000
## 82          5  0.00000000
## 83          6 -0.20412415
## 84          9 -0.28333333
## 85         11  0.00000000
## 86          7  0.00000000
## 87          9  0.26666667
```

```
write.csv(lyrics, "C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\lyricsSENTIMENT.csv")
```

The higher a sentiment score, the more positive a song; the lower, the more negative.

The result of the average sentiment score by song title reveals:

“Little Wing”: slightly positive This result makes sense. The song makes numerous references to a guardian angel-like figure, but through metaphorical passages, so the algorithm had a tough time picking up on the nuances of the lyrics, but successfully detected positivity associated with the lyrics.

“Mary Had a Little Lamb”: slightly positive This result makes sense. “Mary Had a Little Lamb” is SRV’s take on the nursery rhyme, and successfully detected the positivity in the lyrics.

“Pride and Joy”: quite positive This is the most accurate result of the 5. “Pride and Joy” is essentially a declaration of love, in which SRV states his emphtic love for a woman who is his “Pride and Joy”!

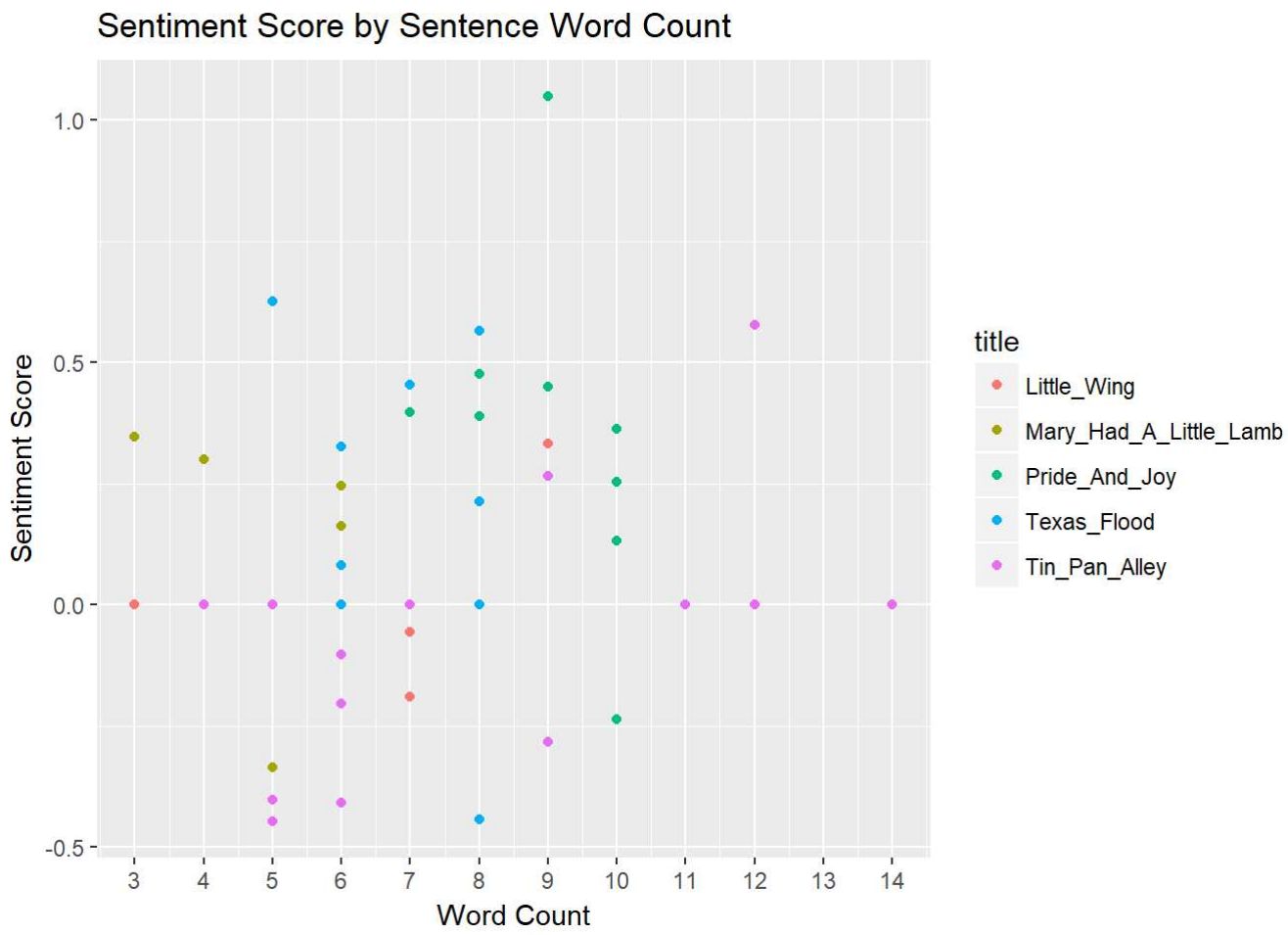
“Texas Flood”: positive Unfortunately, the sentimentr package, and all sentiment analyses come with drawbacks. SRV’s “Texas Flood” is neither a negative nor a positive song. With its numerous references to deadly storms in Texas, there is a great deal of negativity in the song, which the algorithm typically successfully detected and classified; however, SRV states how he wants to leave his “baby,” Texas is personified, here, for a safer land without fear of “tornadoes.” SRV’s leaving could be interpreted negatively (leaving his personified home), or as positive (embarking on a new journey to better his safety.) With the ambiguity in this song, I expected the sentiment score $= 0 \pm 0.025$.

“Tin Pan Alley”: slightly negative This song is definitely negative, with numerous references to violence, police brutality, and other issues in “Tin Pan Alley,” which was a community of New York City music publishers who monopolized popular US music of the early 20th century. Essentially, the song is metaphorical, in the sense that it is SRV’s anthem of the difficulty associated with the music industry – from getting his break through continuously rotating new music into his set list during performances.

Relationship between word_count and sentiment

I hypothesize that there will be no relationship between a sentence’s word count, and its assigned sentiment score.

```
lyrics$title = as.factor(lyrics$title)
myColors = brewer.pal(5,"Set1")
names(myColors) = levels(lyrics$title)
colScale <- scale_colour_manual(name = "title",values = myColors)
j = ggplot(lyrics, aes(word_count, sentiment, colour = title))
wordCountSen = j + geom_point() + labs(x = "Word Count", y = "Sentiment Score", title = "Sentime
nt Score by Sentence Word Count") + scale_x_continuous(breaks = seq(1,15,1))
wordCountSen
```



No relationship; my hypothesis was correct!

Word Clouds

Pride and Joy

```
#Import the lyrics
filenames = list("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\pri
deAndJoy.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Preprocessing/preparing the data
#Convert the text to lower case
documents = tm_map(documents, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Remove numbers
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove english common stopwords
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Remove punctuation
documents <- tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Document term matrix creation
dtm = TermDocumentMatrix(documents)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

#Build the word cloud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 200, random.order = FALSE, ro
t.per = 0.35, colors = brewer.pal(8, "Dark2"))
```

As previously stated, “Pride and Joy” is essentially a declaration of love, and the word cloud reflects that, too. The song’s most frequently-used words, within the context of the song, are multiple references to romance, his love, etc.

Little Wing

```
#Import the lyrics
filenames = list("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\lit
tleWing.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Preprocessing/preparing the data
#Convert the text to lower case
documents = tm_map(documents, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Remove numbers
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove english common stopwords
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Remove punctuation
documents <- tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Document term matrix creation
dtm = TermDocumentMatrix(documents)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

#Build the word cloud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 200, random.order = FALSE, ro
t.per = 0.35, colors = brewer.pal(8, "Dark2"))
```

At the center of this word cloud is “yeah,” an exclamation – especially in the world of Blues. This provides the implication that the song is extra-bluesy, which makes sense, as it is the SRV (Blues) take on Jimi Hendrix’s “Little Wing,” which was a blues-inspired ballad, instead of a more psychedelic-focused mood. SRV brings that out in his take on “Little Wing,” and the word cloud reflects that.

Mary Had a Little Lamb

```
#Import the lyrics
filenames = list("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\mar
yHadALittleLamb.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Preprocessing/preparing the data
#Convert the text to lower case
documents = tm_map(documents, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Remove numbers
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove english common stopwords
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Remove punctuation
documents <- tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Document term matrix creation
dtm = TermDocumentMatrix(documents)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

#Build the word cloud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 200, random.order = FALSE, ro
t.per = 0.35, colors = brewer.pal(8, "Dark2"))
```

Texas Flood

```
#Import the lyrics
filenames = list("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\texasFlood.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Preprocessing/preparing the data
#Convert the text to lower case
documents = tm_map(documents, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Remove numbers
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove english common stopwords
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Remove punctuation
documents <- tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Document term matrix creation
dtm = TermDocumentMatrix(documents)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

#Build the word cloud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 200, random.order = FALSE, ro
t.per = 0.35, colors = brewer.pal(8, "Dark2"))
```

Tin Pan Alley (AKA Roughest Place in Town)

```
#Import the lyrics
filenames = list("C:\\Users\\gould\\Documents\\Developer\\SRV-Topic-Sentiment-Analysis-in-R\\tin
PanAlley.txt")
files = lapply(filenames,readLines)
documents = Corpus(VectorSource(files))

#Preprocessing/preparing the data
#Convert the text to lower case
documents = tm_map(documents, content_transformer(tolower))
```

```
## Warning in tm_map.SimpleCorpus(documents, content_transformer(tolower)):
## transformation drops documents
```

```
#Remove numbers
documents = tm_map(documents, removeNumbers)
```

```
## Warning in tm_map.SimpleCorpus(documents, removeNumbers): transformation
## drops documents
```

```
#Remove english common stopwords
documents = tm_map(documents, removeWords, stopwords("english"))
```

```
## Warning in tm_map.SimpleCorpus(documents, removeWords,
## stopwords("english")): transformation drops documents
```

```
#Remove punctuation
documents <- tm_map(documents, removePunctuation)
```

```
## Warning in tm_map.SimpleCorpus(documents, removePunctuation):
## transformation drops documents
```

```
#Document term matrix creation
dtm = TermDocumentMatrix(documents)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

#Build the word cloud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1, max.words = 200, random.order = FALSE, ro
t.per = 0.35, colors = brewer.pal(8, "Dark2"))
```


hey
alleys
roughest
place
people
yeah
livin
alley
whiskey
ever
gin
wine
took
heard
lord
annie
stay
standing
whiskye
pan
nobody
woman
pen
long
see
killed
scream
hand
saw
run
roll
twofour
tin
gun
going
right
things
somebody
shake
pistol
rattle
cop
hot
doorcat
killin
shoot
raid
workin
shooter
peeped
away
said
cause
back
crap
boy
cwent