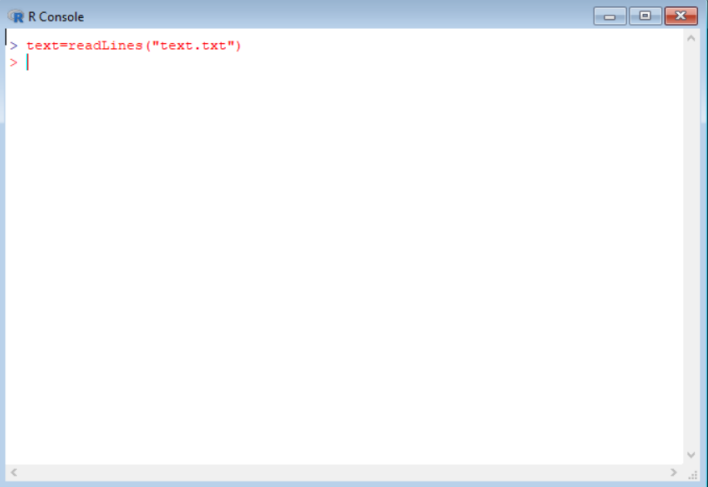
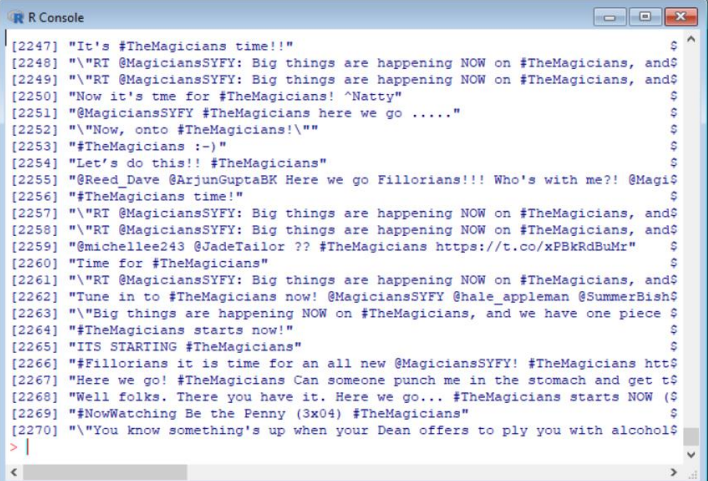


R-Programming Fundamentals for Business Students — Sentiment Analysis

Nick V. Flor, University of New Mexico (nickflor@unm.edu)

Assumptions. This tutorial assumes (1) that you had an Excel worksheet with a single column A, where each row represented a social media post (e.g., a tweet), and (2) that you at least removed carriage returns and/or line feeds from each post; and (3) that you saved that file as *text.txt* and, finally, (4) that you ran R and did a *File > Change dir...* to the folder containing *text.txt*

ACTION	REACTION
IMPORTING A TEXT FILE INTO R	
<ul style="list-style-type: none"> Type <code>text=readLines("text.txt")</code> then press Enter <p><u>Explanation:</u> This command reads every post into a separate row in the variable <code>text</code>. If you do not see an error message, you have done this correctly! If you do get an error, retype the command (don't copy and paste)</p>	
<ul style="list-style-type: none"> Type <code>text</code> (or whatever you named your variable) then press Enter to see your data <p><u>Explanation:</u> Typing any variable allows you to see its content.</p> <p><u>Note:</u> The number of rows (e.g., 2270) should agree with the number of rows in your Excel spreadsheet. If not, then you did not remove all Enters (carriage returns / line feeds) and you will have to go back and start over.</p>	

(continued on next page)

INSTALL PACKAGES

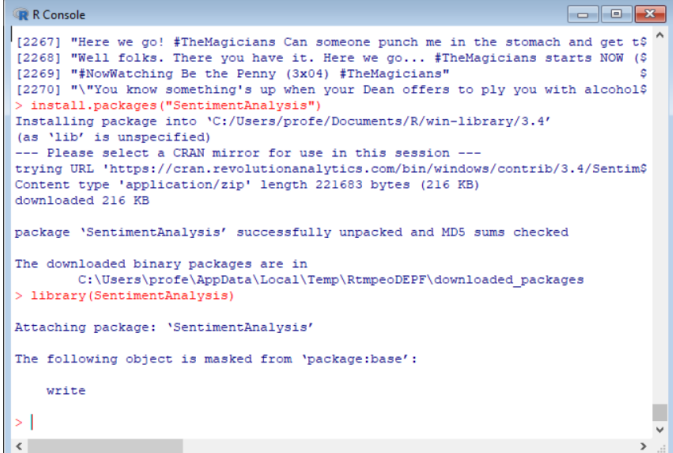
- Type: `install.packages("SentimentAnalysis")`
- Type: `library(SentimentAnalysis)`

Note: This can take several minutes and freeze your keyboard. Just wait it out.

Explanation:

"SentimentAnalysis" is R's sentiment analysis package

`install.packages` merely loads the packages onto your computer. The `library` command allows you to use the packages in your current R session.



```
R Console
[2267] "Here we go! #TheMagicians Can someone punch me in the stomach and get t
[2268] "Well folks. There you have it. Here we go... #TheMagicians starts NOW ($
[2269] "#NowWatching Be the Penny (3x04) #TheMagicians"
[2270] "\nYou know something's up when your Dean offers to ply you with alcohol
> install.packages("SentimentAnalysis")
Installing package into 'C:/Users/profe/Documents/R/win-library/3.4'
(as 'lib' is unspecified)
--- Please select a CRAN mirror for use in this session ---
trying URL 'https://cran.revolutionanalytics.com/bin/windows/contrib/3.4/Sentim
Content type 'application/zip' length 221683 bytes (216 KB)
downloaded 216 KB

package 'SentimentAnalysis' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/profe/AppData/Local/Temp/RtmpoDEPF/downloaded_packages
> library(SentimentAnalysis)

Attaching package: 'SentimentAnalysis'

The following object is masked from 'package:base':

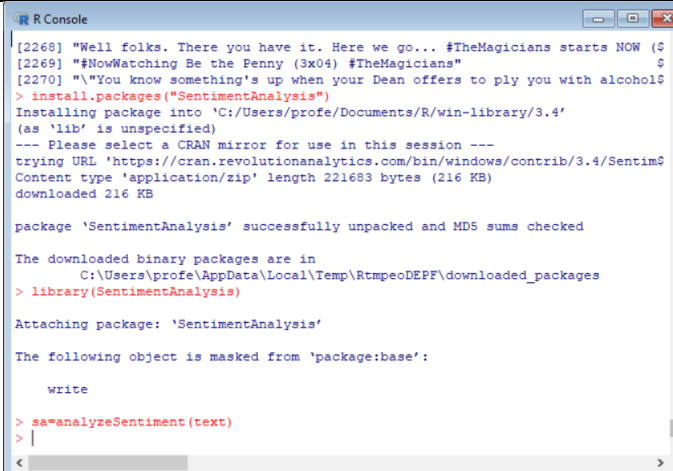
  write

> |
<
```

ANALYZE SENTIMENT

- Type: `sa=analyzeSentiment(text)`

Explanation: `sa` is just a variable, `analyzeSentiment(text)` runs the sentiment analyzer on the entire text.



```
R Console
[2268] "Well folks. There you have it. Here we go... #TheMagicians starts NOW ($
[2269] "#NowWatching Be the Penny (3x04) #TheMagicians"
[2270] "\nYou know something's up when your Dean offers to ply you with alcohol
> install.packages("SentimentAnalysis")
Installing package into 'C:/Users/profe/Documents/R/win-library/3.4'
(as 'lib' is unspecified)
--- Please select a CRAN mirror for use in this session ---
trying URL 'https://cran.revolutionanalytics.com/bin/windows/contrib/3.4/Sentim
Content type 'application/zip' length 221683 bytes (216 KB)
downloaded 216 KB

package 'SentimentAnalysis' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/profe/AppData/Local/Temp/RtmpoDEPF/downloaded_packages
> library(SentimentAnalysis)

Attaching package: 'SentimentAnalysis'

The following object is masked from 'package:base':

  write

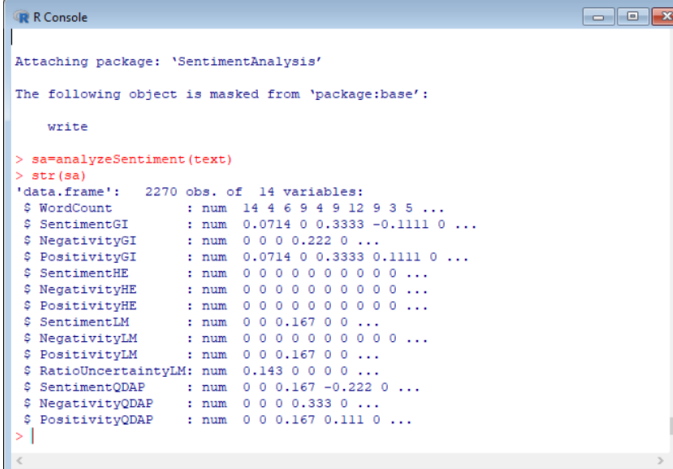
> sa=analyzeSentiment(text)
> |
<
```

EXAMINE, EXPORT, & IMPORT THE RESULTS INTO EXCEL

- Type: `str(sa)`

Explanation: `str` shows the structure of the `sa` variable, which in this example is a data.frame with 2270 points and 14 "columns" including: WordCount, SentimentGI, NegativityGI, PositivityGI, etc. In all there are 4 dictionaries: GI, HE, LM, and QDAP.

You can stay in R to do further processing, or import the results into Excel



```
R Console
Attaching package: 'SentimentAnalysis'

The following object is masked from 'package:base':

  write

> sa=analyzeSentiment(text)
> str(sa)
'data.frame': 2270 obs. of 14 variables:
 $ WordCount      : num 14 4 6 9 4 9 12 9 3 5 ...
 $ SentimentGI    : num 0.0714 0 0.3333 -0.1111 0 ...
 $ NegativityGI   : num 0 0 0 0.222 0 ...
 $ PositivityGI   : num 0.0714 0 0.3333 0.1111 0 ...
 $ SentimentHE    : num 0 0 0 0 0 0 0 0 0 ...
 $ NegativityHE   : num 0 0 0 0 0 0 0 0 0 ...
 $ PositivityHE   : num 0 0 0 0 0 0 0 0 0 ...
 $ SentimentLM    : num 0 0 0.167 0 0 ...
 $ NegativityLM   : num 0 0 0 0 0 0 0 0 0 ...
 $ PositivityLM   : num 0 0 0.167 0 0 ...
 $ RatioUncertaintyLM: num 0.143 0 0 0 0 ...
 $ SentimentQDAP  : num 0 0 0.167 -0.222 0 ...
 $ NegativityQDAP : num 0 0 0 0.333 0 ...
 $ PositivityQDAP : num 0 0 0.167 0.111 0 ...

> |
<
```

- Type `write.csv(sa, "sa.csv")`

Explanation: This command writes out the variable `sa` into the comma-separated file `sa.csv`

```

R Console
Attaching package: 'SentimentAnalysis'

The following object is masked from 'package:base':

write

> sa=analyzeSentiment(text)
> str(sa)
'data.frame': 2270 obs. of 14 variables:
 $ WordCount      : num  14 4 6 9 4 9 12 9 3 5 ...
 $ SentimentGI    : num  0.0714 0 0.3333 -0.1111 0 ...
 $ NegativityGI   : num  0 0 0 0.222 0 ...
 $ PositivityGI   : num  0.0714 0 0.3333 0.1111 0 ...
 $ SentimentHE    : num  0 0 0 0 0 0 0 0 0 ...
 $ NegativityHE   : num  0 0 0 0 0 0 0 0 0 ...
 $ PositivityHE   : num  0 0 0 0 0 0 0 0 0 ...
 $ SentimentLM    : num  0 0 0 0.167 0 ...
 $ NegativityLM   : num  0 0 0 0 0 0 0 0 ...
 $ PositivityLM   : num  0 0 0 0.167 0 ...
 $ RatioUncertaintyLM: num  0.143 0 0 0 ...
 $ SentimentQDAP  : num  0 0 0.167 -0.222 0 ...
 $ NegativityQDAP : num  0 0 0 0.333 0 ...
 $ PositivityQDAP : num  0 0 0.167 0.111 0 ...

> write.csv(sa, "sa.csv")
>
  
```

- Open `sa.csv` in Excel

	WordCount	Sentiment	Negativity	Positivity	Sentiment	Negativity	Positivity	RatioUncertainty	Sentiment	Negativity	Positivity	QDAP
1	14	0.071429	0	0.071429	0	0	0	0.142857	0	0	0	0
2	4	0	0	0	0	0	0	0	0	0	0	0
3	6	0.333333	0	0.333333	0	0	0.166667	0	0.166667	0	0.166667	0
4	9	-0.111111	0.222222	0.111111	0	0	0	0	-0.222222	0.333333	0.111111	0
5	4	0	0	0	0	0	0	0	0	0	0	0
6	9	0.333333	0	0.333333	0	0	0	0	0	0.333333	0.333333	0
7	12	0	0.083333	0.083333	0	0	0	0	0	0.083333	0.083333	0.166667
8	9	-0.222222	0	0.222222	0	0	0	0	-0.222222	0	0.222222	0
9	9	0.333333	0.333333	0	0	0	0	0	0	0	0	0
10	9	0	0	0	0	0	0	0	0	0	0	0
11	30	5	0	0.2	0.2	0	0	0	0	0	0	0
12	11	4	0	0	0	0	0	0	0	0	0	0
13	12	0	0	0	0	0	0	0	0	0	0	0
14	13	14	0.071429	0.071429	0.142857	0	0	0	0	0.142857	0.142857	0
15	14	7	0	0	0	0	0	0	0	0	0	0
16	15	7	-0.142857	0.285714	0.142857	0	0	0	0	0.142857	0.142857	0
17	16	7	0	0	0	0	0	0	0	0	0	0
18	17	7	0	0	0	0	0	0	0	0	0	0
19	18	10	0	0	0	0	0	0	0	0	0	0
20	19	10	0.2	0.1	0.1	0	0	0	0	0.2	0	0.2
21	20	4	0.25	0	0.25	0	0	0	0	0.25	0	0.25
22	21	0	-0.25	0.25	0	0	0	0	-0.25	0.25	0	0
23	22	7	0.285714	0	0.285714	0.142857	0.142857	0.142857	0.142857	0.285714	0.285714	0
24	23	30	0.1	0.1	0.2	0.1	0	0.1	0.1	0.1	0.2	0.2
25	24	9	0.111111	0.111111	0	0	0	-0.111111	0.111111	0	0.111111	0.111111
26	25	20	0	0.040816	0.040816	0	0	0	0	0	0	0.040816
27	26	9	0	0	0	0	0	0	0	0	0	0
28	27	0	-0.111111	0.111111	0	0	0	0	-0.111111	0.111111	0	0
29	28	0	-0.25	0.25	0	0	0	-0.25	0.25	0	0	0
30	29	0	0	0	0	0	0	0	0	0	0	0

You're done. You can now copy the sentiment columns next to your data rows. Choose the directionary that gives you the best results.