Step by Step Guide

Text Mining and Sentiment Analysis with R

by R-Tutorials.com

Preparation Steps

Setup Twitter developer account and collect the keys

Download and activate required R packages

Connect R to Twitter

Download and unzip sentiment lexicons, and put them into the working directory

Preparation Steps

Lecture: *"Twitter Developer Account"*

Lecture: *"Required Packages"*

Lecture: *"Connection: R-Twitter"*

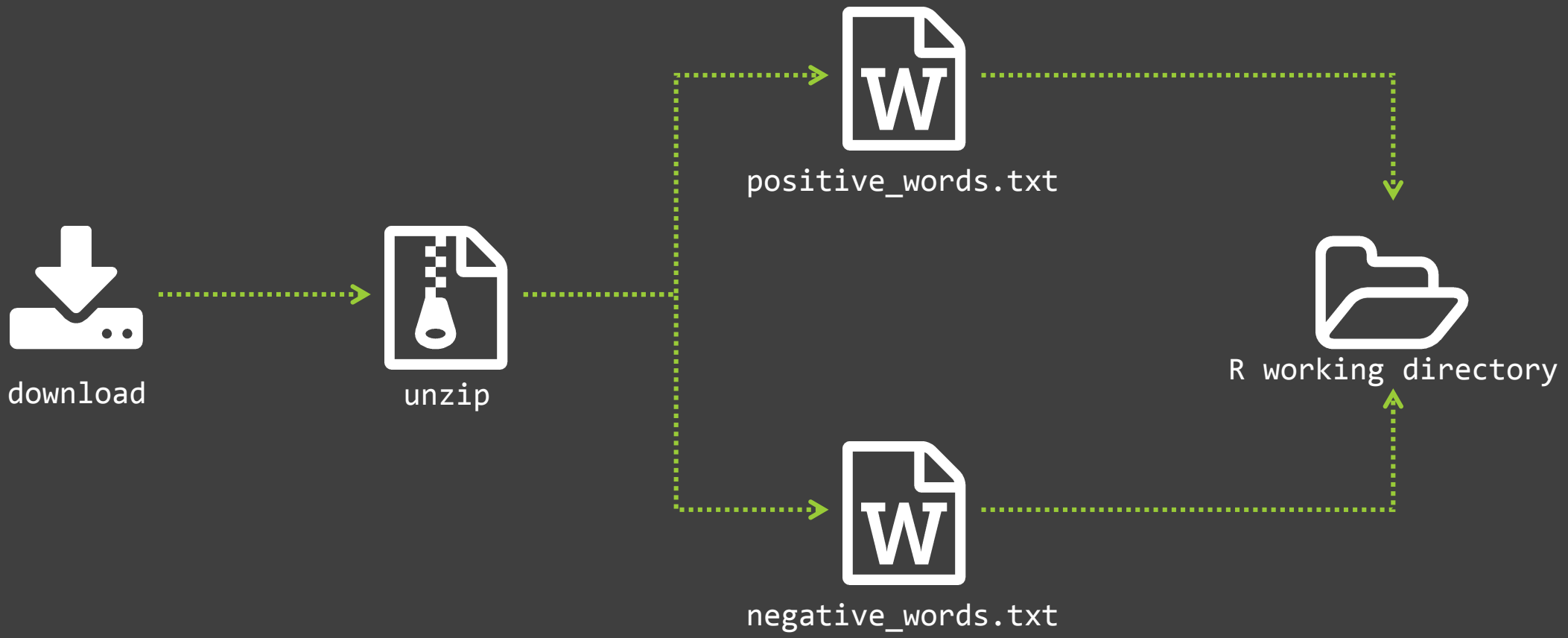
Lecture: *"Sentiment Lexicon"*

Required Packages

- ROAuth
- RCurl
- httr
- twitteR
- tm
- stringr
- plyr
- dplyr


R
Tokens, secrets, keys
API
Connection verification

Sentiment Lexicons

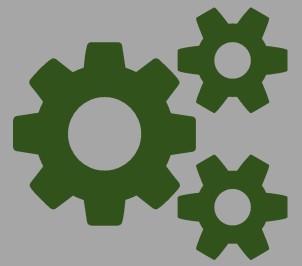


Text Mining and Analysing Steps

Scraping Twitter using key words

Text cleaning phase I

Writing the sentiment score function



Applying the sentiment score function

Text Mining and Analysing Steps

Lecture: *"Twitter Scraping"*

Lecture: *"Cleaning Phase I"*

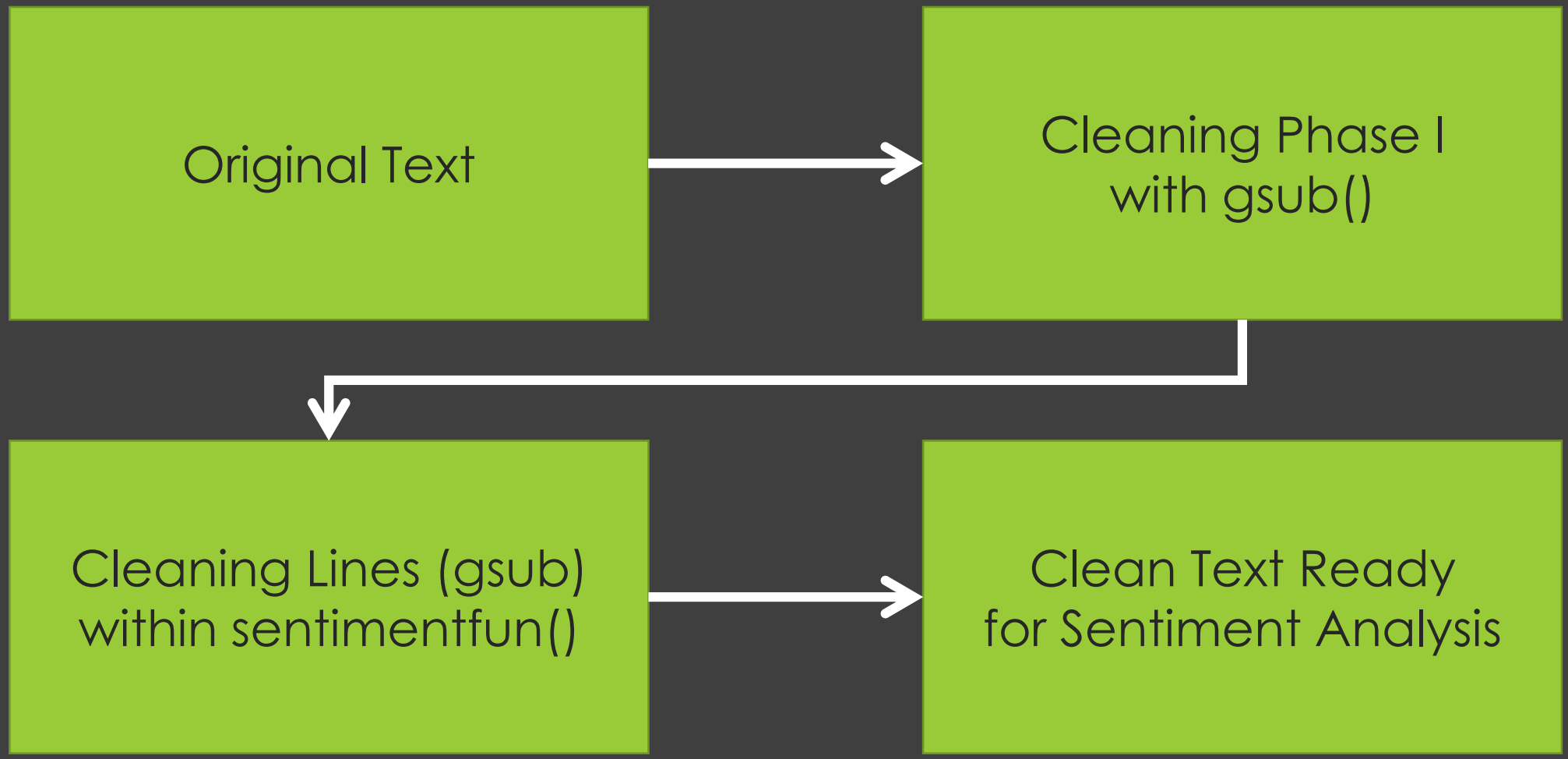
Lecture: *"Sentiment Score Function"*

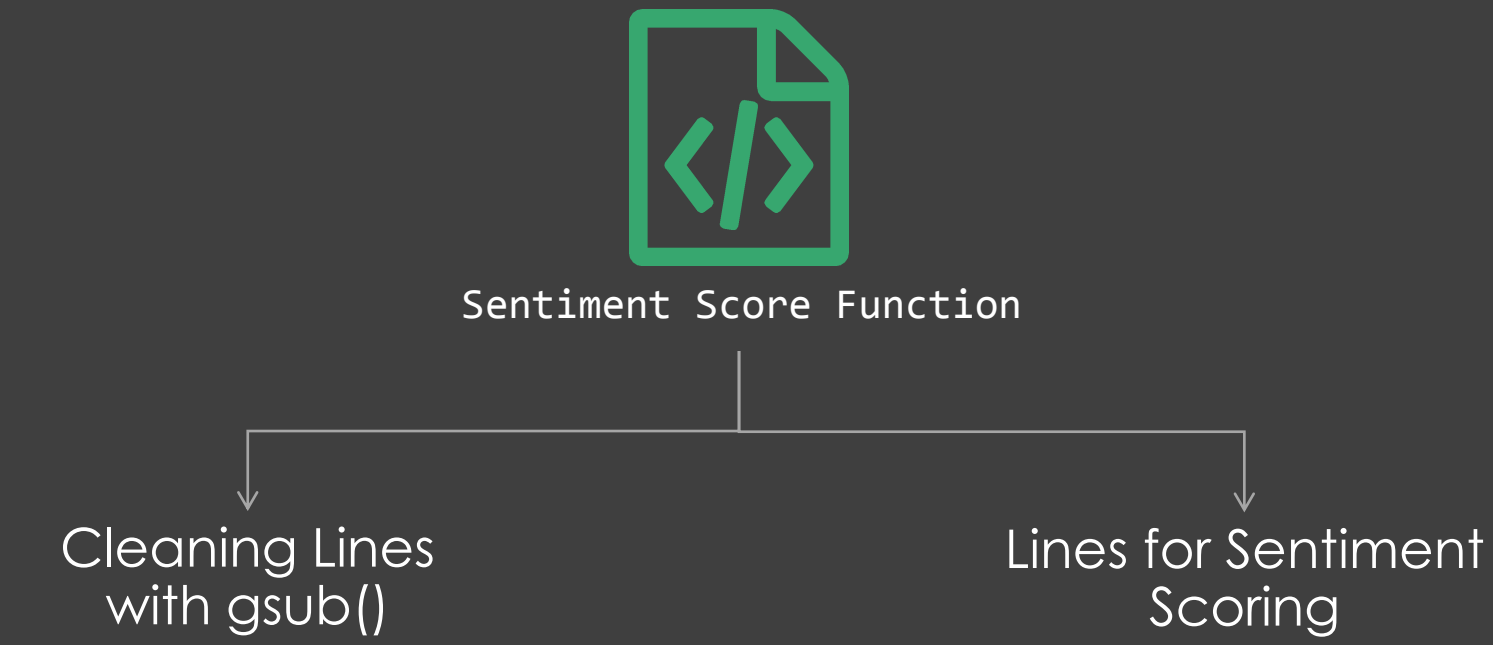
Lecture: *"Sentiment Score Function"*

Keywords and Locations Used

- *Each combination requires a separate scraping session*
- *Results are to be stored in different objects*

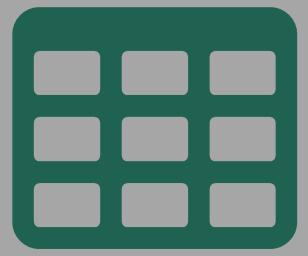
- "apple+iphone" – Los Angeles
- "apple+iphone" – New York
- "apple+iphone" – Toronto
- "apple+iphone" – London
- "samsung+galaxy" – Los Angeles
- "samsung+galaxy" – New York
- "samsung+galaxy" – Toronto
- "samsung+galaxy" – London


Original Text
Cleaning Phase I with gsub()
Cleaning Lines (gsub) within sentimentfun()
Clean Text Ready for Sentiment Analysis


Sentiment Score Function
Cleaning Lines with gsub()
Lines for Sentiment Scoring

- Apply the function '`sentimentfun()`' separately on each object containing tweet text (*8 sessions*)
  - Tweets about Apple iPhone from Los Angeles
  - Tweets about Samsung Galaxy from Los Angeles
  - and so on…
- Store the results (sentiment score) in separate objects (*8 objects*)
  - Sentiment score of Tweets about Apple iPhone from Los Angeles
  - Sentiment score of Tweets about Samsung Galaxy from Los Angeles
  - and so on…

Data Gathering and Exporting Steps



Gather relevant data into a data frame

Export data frames as .CSV files

Data Gathering and Exporting Steps

Lecture: "*Data Gathering and Export as .CSV*"

Data Gathering

Create a data frame for each search term (*Apple iPhone, Samsung Galaxy*) and location (*New York, Los Angeles, Toronto, London*) combination containing:

- Tweet Text
- Tweet Date
- Is Retweet
- Retweet Count
- Favourite Count
- Score
- Product
- City
- Country

Result: 8 data frames

Export Data as .CSV files

Export the data frames as .CSV files:

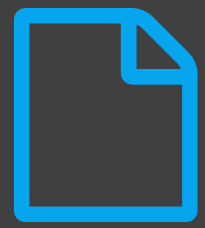
- Apple iPhone + Los Angeles
- Apple iPhone + Now York
- Apple iPhone + Toronto
- Apple iPhone + London
- Samsung Galaxy + Los Angeles
- Samsung Galaxy + New York
- Samsung Galaxy + Toronto
- Samsung Galaxy + London

Result: 8 .CSV files

apple_langeles.CSV

apple_newyork.CSV

apple_toronto.CSV



apple_london.CSV

samsung_langeles.CSV

samsung_newyork.CSV

samsung_toronto.CSV

samsung_london.CSV
Apparently omitted due to the low number of Tweets