

SENTIMENT ANALYSIS ON AMAZON REVIEWS

B. E. Information Technology

By

Group No. - 20

Akshay Gunjal	34
Shaun Dsilva	35
Davon Carvalho	36

Ms. Vandana Patil

Assistant Professor



Department of Information Technology
St. Francis Institute of Technology
(Engineering College)

University of Mumbai
2020-2021

DECLARATION

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources.

We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in this submission.

We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

1. _____
(Signature)

(Akshay Gunjal, Roll Number 34)

2. _____
(Signature)

(Shaun Dsilva, Roll Number 35)

3. _____
(Signature)

(Davon Carvalho, Roll Number 36)

Date:

CERTIFICATE

This R programming Mini-project *Sentiment Analysis on Amazon Reviews* by *Akshay Gunjal (34), Shaun Dsilva (35), Davon Carvalho (36)* is complete in all respects and was successfully demonstrated on {Final external presentation date}.

Name : _____

Signature : _____

(Internal examiner)

Name : _____

Signature : _____

(External examiner)

Date:

Place: SFIT, Mumbai

INDEX

Sr. No.	Content	Page No.
1.	Project Overview	
2.	Data Extraction	
3.	Exploratory Data Analysis	
4.	Application of Mining Algorithm	
5.	Data Visualization and Interpretation	
6.	Conclusion	
7.	Acknowledgement	

Signature:

Ms. Vandana Patil
(Internal Guide)

Project Overview

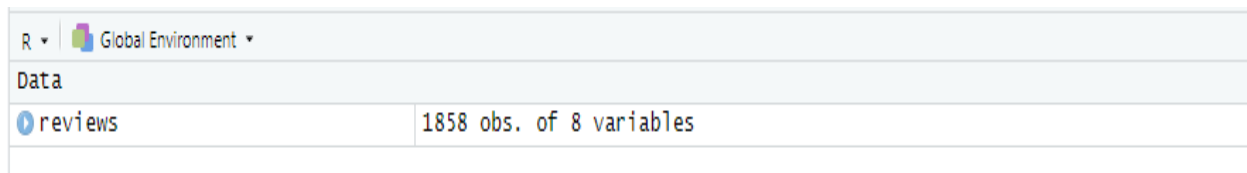
PROJECT OVERVIEW		
1.	Project Title -	Sentiment analysis on Amazon reviews
2.	Data set Name -	Amazon product reviews data set
3.	Introduction of Data set -	<p>The data set considered for analysis is the Amazon Product Review Data Set. The dataset consists of 8 columns for data such as:</p> <ol style="list-style-type: none">1. Id- It contains the registered number for the following product2. Profile name- It is the name of a person who has reviewed the product3. Text- These are the comments about the product which are written by the reviewer4. Date- Exact date when the review was given5. Title- Written in bold statement which the reviewer want to highlight6. Rating- The rating out of five which the customer gives about the product7. Images- The image of the product uploaded by the reviewer
4.	Length of Data set	12000
5.	Name of the source website -	Amazon review exporter
6.	URL -	https://mxnpro.ovh/ https://chrome.google.com/webstore/detail/amazon-reviews-exporter-c/njlpnciolcibljfdbobcefcngiampidm/

Data Extraction

Data Extraction

2.1 Import Data (.csv) ,

The read.csv() function is used to read the file in table format and create a data frame from it.
`reviews <- read.csv(file.choose(),header=T)`

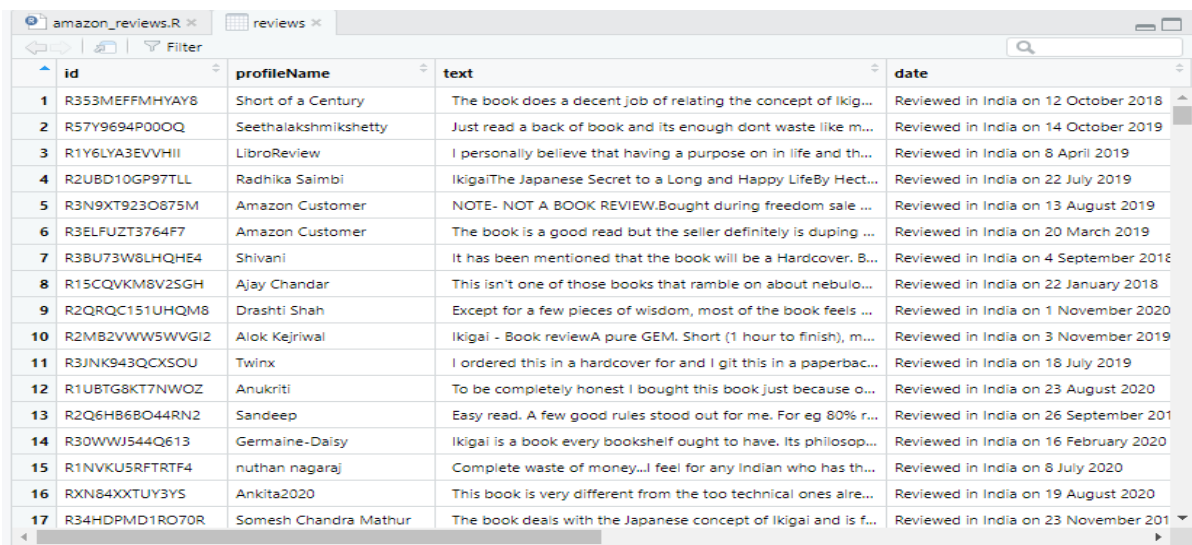


Global Environment	
Data	
reviews	1858 obs. of 8 variables

2.2 Viewing the data

2.2.1 View(reviews)

The View() function is used to invoke a spreadsheet-style data viewer on a matrixlike R object.



	id	profileName	text	date
1	R353MEFFMHYAY8	Short of a Century	The book does a decent job of relating the concept of Ikigai...	Reviewed in India on 12 October 2018
2	R57Y9694P00OQ	Seethalakshmikshetty	Just read a back of book and its enough dont waste like m...	Reviewed in India on 14 October 2019
3	R1Y6LYA3EUVHII	LibroReview	I personally believe that having a purpose on in life and th...	Reviewed in India on 8 April 2019
4	R2UBD10GP97TLL	Radhika Saimbi	IkigaiThe Japanese Secret to a Long and Happy LifeBy Hect...	Reviewed in India on 22 July 2019
5	R3N9XT923O875M	Amazon Customer	NOTE- NOT A BOOK REVIEW.Bought during freedom sale ...	Reviewed in India on 13 August 2019
6	R3ELFUZT3764F7	Amazon Customer	The book is a good read but the seller definitely is duping ...	Reviewed in India on 20 March 2019
7	R3BU73W8LHQHE4	Shivani	It has been mentioned that the book will be a Hardcover. B...	Reviewed in India on 4 September 2018
8	R15CQVKM8V25GH	Ajay Chandar	This isn't one of those books that ramble on about nebulo...	Reviewed in India on 22 January 2018
9	R2QRQC151UHQM8	Drashti Shah	Except for a few pieces of wisdom, most of the book feels ...	Reviewed in India on 1 November 2020
10	R2MB2VWW5WVG12	Alok Kejriwal	Ikigai - Book reviewA pure GEM. Short (1 hour to finish), m...	Reviewed in India on 3 November 2019
11	R3JNK943QCXSOU	Twinx	I ordered this in a hardcover for and I git this in a paperbac...	Reviewed in India on 18 July 2019
12	R1UBTG8KT7NWOZ	Anukriti	To be completely honest I bought this book just because o...	Reviewed in India on 23 August 2020
13	R2Q6HB6BO44RN2	Sandeep	Easy read. A few good rules stood out for me. For eg 80% r...	Reviewed in India on 26 September 201
14	R30WWJ544Q613	Germaine-Daisy	Ikigai is a book every bookshelf ought to have. Its philosop...	Reviewed in India on 16 February 2020
15	R1NVKU5RFRTRF4	nuthan nagaraj	Complete waste of money...I feel for any Indian who has th...	Reviewed in India on 8 July 2020
16	RXN84XXTUY3YS	Ankita2020	This book is very different from the too technical ones alre...	Reviewed in India on 19 August 2020
17	R34HDPMD1RO70R	Somesh Chandra Mathur	The book deals with the Japanese concept of Ikigai and is f...	Reviewed in India on 23 November 201

2.2.2 head(reviews)

```
> head(reviews)
      id      profileName
1 R353MEFFMHYAY8 short of a Century
2 R57Y9694P00OQ seethalakshmikshetty
3 R1Y6LYA3EUVHII LibroReview
4 R2UBD10GP97TLL Radhika Saimbi
5 R3N9XT923O875M Amazon Customer
6 R3ELFUZT3764F7 Amazon Customer
```


2.3 Knowing the data

2.3.1 dim(reviews)

The dim() function is used to retrieve or set the dimension of an object.

```
> dim(reviews)
[1] 1858    8
> |
```

2.3.2 str(reviews)

The str() function is used to compactly display the internal structure of an R object

```
> str(reviews)
'data.frame': 1858 obs. of 8 variables:
 $ id      : chr  "R353MEFFMHYAY8" "R57Y9694P000Q" "R1Y6LYA3EUVHII" "R2UBD10GP97TLL" ...
 $ profileName: chr  "Short of a Century" "SeethalakshmiKshetty" "LibroReview" "Radhika Saimbi" ...
 $ text     : chr  "\n The book does a decent job of relating the concept of Ikigai to modern day
 psychology (with Frankl's Logoth)| __truncated__ "\n Just read a back of book and its enough dont
 waste like meðÿ\230t\n" "\n I personally believe that having a purpose on in life and then giving
 it your all is the most important to "| __truncated__ "\n IkigaiThe Japanese Secret to a Long and
 Happy LifeBy Hector Garcia and Francesc MirallesI will be confessin"| __truncated__ ...
 $ date     : chr  "Reviewed in India on 12 October 2018" "Reviewed in India on 14 October 2019"
 "Reviewed in India on 8 April 2019" "Reviewed in India on 22 July 2019" ...
 $ title    : chr  "Simple & Light Reading. May Disappoint - Dependent on Reader's Expectation."
 "Ikigai" "The best book to read during your break time." "Little hyped!!" ...
 $ rating   : int  3 2 4 3 4 1 1 5 2 4 ...
 $ images   : chr  "https://images-na.ssl-images-amazon.com/images/I/71+tJMyIj6L.jpg\nhttps://imag
 es-na.ssl-images-amazon.com/image"| __truncated__ "https://images-na.ssl-images-amazon.com/images/I/
 711BauE1RYL.jpg\nhttps://images-na.ssl-images-amazon.com/image"| __truncated__ "" "https://images-n
 a.ssl-images-amazon.com/images/I/71tzysoukOL.jpg" ...
 $ helpful  : int  656 425 308 43 23 28 29 33 10 10 ...
```

Exploratory Data Analysis

Exploratory Data Analysis

- Text Mining using tm

- 3.1 tm installation
install.packages('tm')
- 3.2 library(tm)
- 3.3 corpus <- iconv(reviews\$text)

Values	
corpus	chr [1:943] "\n I saw my father die right infront of me in the ambulanc...
review_score	Named num [1:11] 470 894 433 633 562 ...
sentiment_data	chr [1:943] "\n I saw my father die right infront of me in the ambulanc...
w	Named num [1:3351] 197 150 114 103 100 99 96 93 91 70 ...

- 3.4 inspect(corpus[1:5])

```
> inspect(corpus[1:5])
<<simpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 5

[1] c("\n The book does a decent job of relating the concept of Ikigai to modern day psychology (with Frankl's Logotherapy from Man's Search for Meaning among others) and a few scientific references in a simple manner. It talks about how purpose play s an important role in a man's life and the different ways in which it manifests itself. It also tackles some ways to 'find yo ur flow' and ensure that what you do receives 100% of your attention and that you enjoy whatever you are creating.The book als o discusses certain other Japanese concepts like takumi (specialized workers) and moai (connections with community or friend-c ircle). The brief discussions have the benefit of being to the point and simple but also pose the risk of trivializing them in to regular self-help advice. The book also delves into Japanese perspectives on living life and persevering without getting ca ught up in artificially-created urgency. But again, maybe the authors wished for the readers to research more or meditate more on the content given the concise treatment of the same.The chapters on diet and exercises have more details and thus, may be more useful. Certain foods are dealt with in greater detail as is the concept of 'hara hachi bu' wherein one eats only 80% of what would actually assuage his hunger. The chapter on exercises includes illustrations and steps. while they may suffice for some of the purposes mentioned in the book - the philosophy behind them, progressive increments and other essential details a re missing or insufficient.I was interested in the concept of Ikigai and wanted to read more about it. Despite being well-writ ten and presenting modern applications, the book did not fulfill my requirements at all. Some of the condensed content made me think that the extended research, including on-site interviews, done by the authors for writing this book was clearly lost in translation or presentation in certain parts.The hardcover is pretty and soothing with its matte texture. Inner pages are smo oth and heavv with a cream tinge. The font size is good. The spine as well as pages hold up well. Overall. the book is quite l
```

- 3.5 Cleaning the corpus to lower case
corpus <- tm_map(corpus,tolower)

```
> inspect(corpus[1])
<<simpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 1

[1] \n the book does a decent job of relating the concept of ikigai to modern day psychology (with frankl's logotherapy from ma n's search for meaning among others) and a few scientific references in a simple manner. it talks about how purpose plays an impo rtant role in a man's life and the different ways in which it manifests itself. it also tackles some ways to 'find your flow' and ensure that what you do receives 100% of your attention and that you enjoy whatever you are creating.the book also discusses cer tain other japanese concepts like takumi (specialized workers) and moai (connections with community or friend-circle). the brief discussions have the benefit of being to the point and simple but also pose the risk of trivializing them into regular self-help advice. the book also delves into japanese perspectives on living life and persevering without getting caught up in artificially -created urgency. but again, maybe the authors wished for the readers to research more or meditate more on the content given the concise treatment of the same.the chapters on diet and exercises have more details and thus, may be more useful. certain foods a re dealt with in greater detail as is the concept of 'hara hachi bu' wherein one eats only 80% of what would actually assuage his hunger. the chapter on exercises includes illustrations and steps. while they may suffice for some of the purposes mentioned in the book - the philosophy behind them, progressive increments and other essential details are missing or insufficient.i was inte rested in the concept of ikigai and wanted to read more about it. despite being well-written and presenting modern applications, the book did not fulfill my requirements at all. some of the condensed content made me think that the extended research, includi ng on-site interviews, done by the authors for writing this book was clearly lost in translation or presentation in certain part s.the hardcover is pretty and soothing with its matte texture. inner pages are smooth and heavy with a cream tinge. the font size of inr 499.\n
```

3.6 Cleaning the corpus to remove punctuations

```
corpus <- tm_map(corpus, removePunctuation)
```

```
> inspect(corpus[1])  
<<SimpleCorpus>>  
Metadata: corpus specific: 1, document level (indexed): 0  
Content: documents: 1  
  
[1] \n the book does a decent job of relating the concept of ikigai to modern day psychology with frankls logotherapy from mans search for meanin  
g among others and a few scientific references in a simple manner it talks about how purpose plays an important role in a mans life and the differ  
ent ways in which it manifests itself it also tackles some ways to find your flow and ensure that what you do receives 100 of your attention and t  
hat you enjoy whatever you are creatingthe book also discusses certain other japanese concepts like takumi specialized workers and moai connection  
s with community or friendcircle the brief discussions have the benefit of being to the point and simple but also pose the risk of trivializing th  
em into regular selfhelp advice the book also delves into japanese perspectives on living life and persevering without getting caught up in artifi  
ciallycreated urgency but again maybe the authors wished for the readers to research more or meditate more on the content given the concise treatm  
ent of the samethe chapters on diet and exercises have more details and thus may be more useful certain foods are dealt with in greater detail as  
is the concept of hara hachi bu wherein one eats only 80 of what would actually assuage his hunger the chapter on exercises includes illustration  
s and steps while they may suffice for some of the purposes mentioned in the book the philosophy behind them progressive increments and other ess  
ential details are missing or insufficienti was interested in the concept of ikigai and wanted to read more about it despite being wellwritten and  
presenting modern applications the book did not fulfill my requirements at all some of the condensed content made me think that the extended rese  
arch including onsite interviews done by the authors for writing this book was clearly lost in translation or presentation in certain partsthe har  
dcover is pretty and soothing with its matte texture inner pages are smooth and heavy with a cream tinge the font size is good the spine as well a  
s pages hold up well overall the book is quite light and sturdy bought it for inr 460 against mrrp of inr 499\n  
> |
```

3.7 Cleaning the corpus to remove the numbers

```
corpus <- tm_map(corpus, removeNumbers)
```

```
> inspect(corpus[1])  
<<SimpleCorpus>>  
Metadata: corpus specific: 1, document level (indexed): 0  
Content: documents: 1  
  
[1] \n the book does a decent job of relating the concept of ikigai to modern day psychology with frankls logother  
apy from mans search for meaning among others and a few scientific references in a simple manner it talks about how  
purpose plays an important role in a mans life and the different ways in which it manifests itself it also tackles  
some ways to find your flow and ensure that what you do receives of your attention and that you enjoy whatever yo  
u are creatingthe book also discusses certain other japanese concepts like takumi specialized workers and moai conn  
ections with community or friendcircle the brief discussions have the benefit of being to the point and simple but  
also pose the risk of trivializing them into regular selfhelp advice the book also delves into japanese perspectiv  
es on living life and persevering without getting caught up in artificiallycreated urgency but again maybe the auth  
ors wished for the readers to research more or meditate more on the content given the concise treatment of the same  
the chapters on diet and exercises have more details and thus may be more useful certain foods are dealt with in gr  
eater detail as is the concept of hara hachi bu wherein one eats only of what would actually assuage his hunger th  
e chapter on exercises includes illustrations and steps while they may suffice for some of the purposes mentioned i  
n the book the philosophy behind them progressive increments and other essential details are missing or insufficie  
nti was interested in the concept of ikigai and wanted to read more about it despite being wellwritten and presenti  
ng modern applications the book did not fulfill my requirements at all some of the condensed content made me think  
that the extended research including onsite interviews done by the authors for writing this book was clearly lost  
in translation or presentation in certain partsthe hardcover is pretty and soothing with its matte texture inner p  
ages are smooth and heavy with a cream tinge the font size is good the spine as well as pages hold up well overall  
the book is quite light and sturdy bought it for inr 460 against mrrp of inr 499\n  
> |
```

3.8 Cleaning the corpus to remove stop words.

```
corpus <- tm_map(corpus, removeWords, stopwords("english"))
```

```
> inspect(corpus[1])  
<<SimpleCorpus>>  
Metadata: corpus specific: 1, document level (indexed): 0  
Content: documents: 1  
  
[1] \n book decent job relating concept ikigai modern day psychology frankls logotherapy mans search mea  
ning among others scientific references simple manner talks purpose plays important role mans life dif  
ferent ways manifests also tackles ways find flow ensure receives attention enjoy whatever cre  
atingthe book also discusses certain japanese concepts like takumi specialized workers moai connections communit  
y friendcircle brief discussions benefit point simple also pose risk trivializing regular selfhelp ad  
vice book also delves japanese perspectives living life persevering without getting caught artificiallycreate  
d urgency maybe authors wished readers research meditate content given concise treatment samethe chap  
ters diet exercises details thus may useful certain foods dealt greater detail concept hara hachi bu  
wherein one eats actually assuage hunger chapter exercises includes illustrations steps may suffice  
purposes mentioned book philosophy behind progressive increments essential details missing insufficienti  
interested concept ikigai wanted read despite wellwritten presenting modern applications book fulfill  
requirements condensed content made think extended research including onsite interviews done authors w  
riting book clearly lost translation presentation certain partsthe hardcover pretty soothing matte texture  
inner pages smooth heavy cream tinge font size good spine well pages hold pretty well overall book quite ligh  
t sturdy bought inr mrrp inr \n  
> |
```

3.9 Cleaning the corpus to remove white spaces.

```
corpus <- tm_map(corpus, stripWhitespace)
```

```
> corpus <- tm_map(corpus, stripWhitespace)
warning message:
In tm_map.SimpleCorpus(corpus, stripWhitespace) :
  transformation drops documents
> inspect(corpus[1])
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 1

[1] book decent job relating concept ikigai modern day psychology frankls logotherapy mans search meaning among ot
hers scientific references simple manner talks purpose plays important role mans life different ways manifests also
tackles ways find flow ensure receives attention enjoy whatever creatingthe book also discusses certain japanese c
oncepts like takumi specialized workers moai connections community friendcircle brief discussions benefit point sim
ple also pose risk trivializing regular selfhelp advice book also delves japanese perspectives living life persever
ing without getting caught artificiallycreated urgency maybe authors wished readers research meditate content given
concise treatment samethe chapters diet exercises details thus may useful certain foods dealt greater detail conce
pt hara hachi bu wherein one eats actually assuage hunger chapter exercises includes illustrations steps may suffic
e purposes mentioned book philosophy behind progressive increments essential details missing insufficienti interest
ed concept ikigai wanted read despite wellwritten presenting modern applications book fulfill requirements condense
d content made think extended research including onsite interviews done authors writing book clearly lost translati
on presentation certain partsthe hardcover pretty soothing matte texture inner pages smooth heavy cream tinge font
size good spine well pages hold well overall book quite light sturdy bought inr mrp inr
> |
```

3.10 Storing the cleaned corpus in review_final variable.

```
reviews_final <- corpus
```

3.11 Creating a term document

```
tdm <- TermDocumentMatrix(reviews_final)
tdm <- as.matrix(tdm)
tdm[1:10,1:5]
```

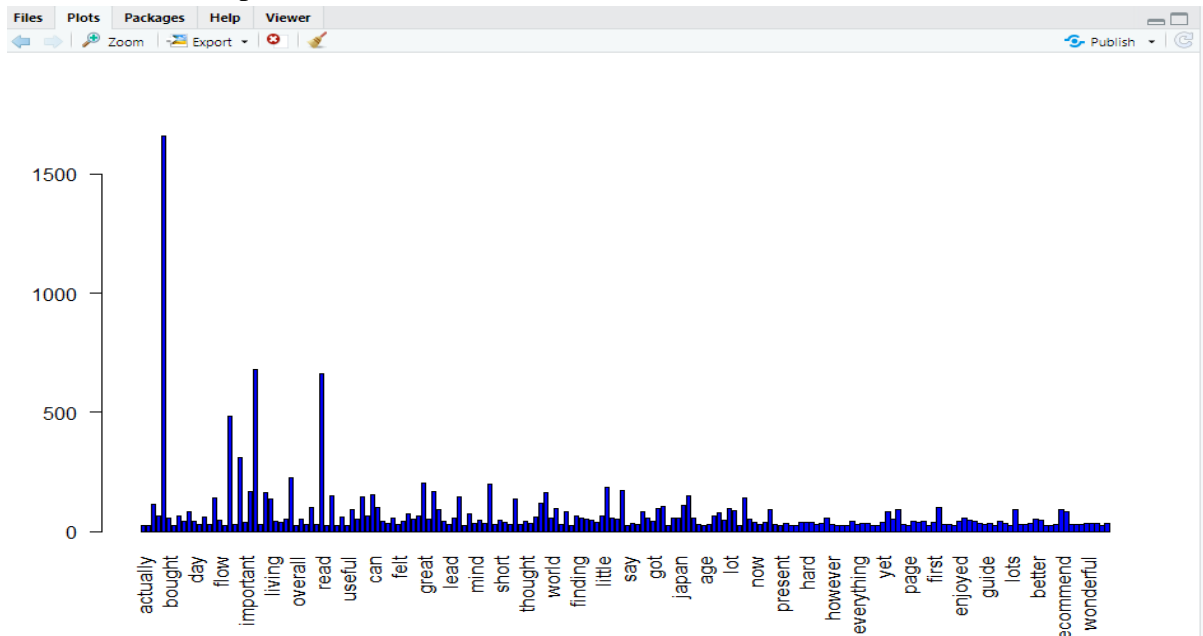
```
> tdm <- TermDocumentMatrix(reviews_final)
> tdm <- as.matrix(tdm)
> tdm[1:10,1:5]
```

Terms	1	2	3	4	5
actually	1	0	0	0	0
advice	1	0	0	0	0
also	4	0	2	0	0
among	1	0	0	0	0
applications	1	0	0	0	0
artificiallycreated	1	0	0	0	0
assuage	1	0	0	0	0
attention	1	0	0	0	0
authors	2	0	0	0	0
behind	1	0	0	1	0

```
> |
```

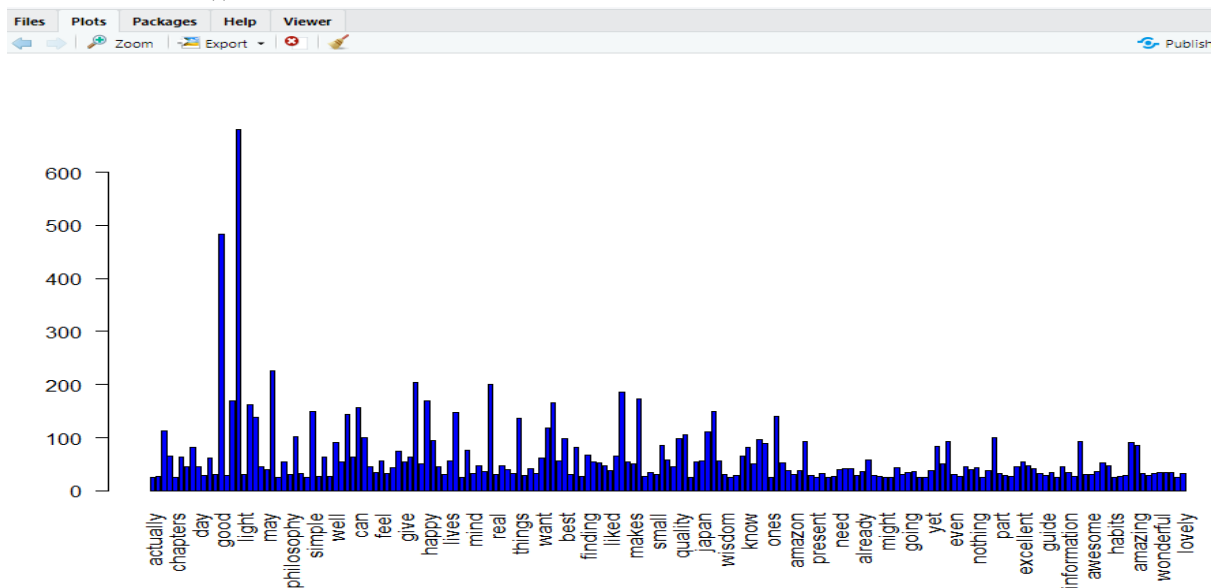
3.12 Creating a bar plot for sum of tdm

```
w <- rowSums(tdm)
w <- subset(w, w>=25)
barplot(w, las = 2, col = "blue")
```



3.13 Cleaning the data again for specific non-sentimental words

```
corpus <- tm_map(corpus, removeWords,
c("given", "bought", "read", "book", "important", "flow", "ikigai", "find", "meaning"))
```



Application of Mining Algorithm

Application of Mining/Analytics Algorithm

4.1 Sentiment Analysis

There are a variety of methods and dictionaries that exist for evaluating the opinion or emotion in text. The tidy text package provides access to several sentiment lexicons. Three general-purpose lexicons are AFINN from Finn Arup Nielsen, Bing from Bing Liu and collaborators, and NRC from Saif Mohammad and Peter Turney. All three of these lexicons are based on unigrams, i.e., single words. These lexicons contain many English words, and the words are assigned scores for positive/negative sentiment, and also possibly emotions like joy, anger, sadness, and so forth. The NRC lexicon categorizes words in a binary fashion (“yes” / “no”) into categories of positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The Bing lexicon categorizes words in a binary fashion into positive and negative categories. The AFINN lexicon assigns words with a score that runs between -5 and 5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment.

NRC Sentiments

The NRC Emotion Lexicon is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive). The annotations were manually done by crowdsourcing.

4.2 Implementation of NRC Sentiments

```
sentiment_data <- iconv(reviews$text)
s <- get_nrc_sentiment(sentiment_data)
s[1:10,] #viewing sentiments for 10 records
```

```
> s <- get_nrc_sentiment(sentiment_data)
> s[1:10,]
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     1             7      1   3   9         3       3    11        6        24
2     0             0      1   0   0         0       0     0        1         0
3     3            14      1   2   9         3       1    13        6        22
4     2            10      2   1   9         4       3    12        4        14
5     0             0      0   0   1         0       0     2        0         2
6     0             1      0   0   2         0       1     2        2         2
7     0             0      0   0   0         0       0     1        0         1
8     0             2      0   0   3         0       0     5        0         3
9     1             0      0   0   1         1       1     3        2         6
10    1             3      1   0   2         2       0     4        2        11
> |
```


4.3 Creating a new variable in the sentiment table.

A new variable “score” is created in the sentiment table which is the difference of negative and positive sentiment score.

```
s$score <- s$positive - s$negative
s[1:10,]
```

```
> s$score <- s$positive - s$negative
> s[1:10,]
  anger anticipation disgust fear joy sadness surprise trust negative positive score
1     1           7       1    3    9         3         3    11         6        24     18
2     0           0       1    0    0         0         0     0         1         0     -1
3     3          14       1    2    9         3         1    13         6        22     16
4     2          10       2    1    9         4         3    12         4        14     10
5     0           0       0    0    1         0         0     2         0         2      2
6     0           1       0    0    2         0         1     2         2         2      0
7     0           0       0    0    0         0         0     1         0         1      1
8     0           2       0    0    3         0         0     5         0         3      3
9     1           0       0    0    1         1         1     3         2         6      4
10    1           3       1    0    2         2         0     4         2        11      9
> |
```

4.4 Saving the scores in a CSV file.

```
write.csv(x = s, file = "C:/Users/Shawn Dsilva/Desktop/SEM 8/R/Amazon
Reviews/Final_score_ikigai.csv")
```

	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	negative	positive	score
1	1	7	1	3	9	3	3	11	6	24	18
2	0	0	1	0	0	0	0	0	1	0	-1
3	3	14	1	2	9	3	1	13	6	22	16
4	2	10	2	1	9	4	3	12	4	14	10
5	0	0	0	0	1	0	0	2	0	2	2
6	0	1	0	0	2	0	1	2	2	2	0
7	0	0	0	0	0	0	0	1	0	1	1
8	0	2	0	0	3	0	0	5	0	3	3
9	1	0	0	0	1	1	1	3	2	6	4
10	1	3	1	0	2	2	0	4	2	11	9
11	0	2	0	0	3	0	2	4	0	4	4
12	1	14	1	4	15	5	2	13	4	27	23
13	1	4	0	2	3	2	2	4	2	6	4
14	0	3	1	1	3	0	1	2	2	5	3
15	1	2	1	0	1	0	1	1	2	3	1
16	1	3	1	2	3	2	2	4	6	6	0
17	1	3	0	0	3	0	1	4	2	5	3
18	2	10	2	1	7	3	2	13	7	18	11
19	1	5	0	1	7	0	1	6	1	20	19
20	1	1	0	1	1	2	0	0	3	2	-1
21	3	14	1	3	14	8	8	15	8	26	18
22	1	2	2	1	4	2	2	2	2	8	6

4.5 Checking the overall sentiments.

```
review_score <- colSums(s[,])  
print(review_score)
```

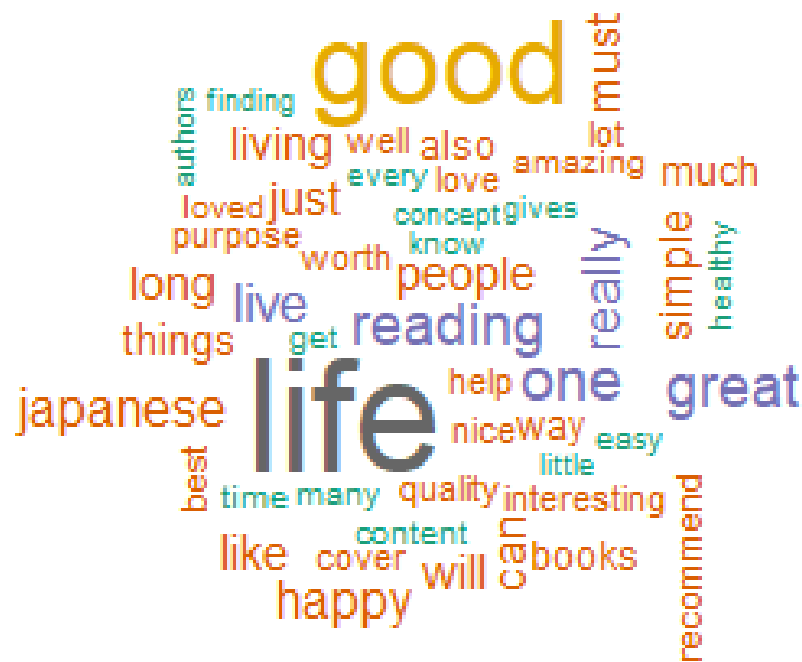
```
> review_score <- colSums(s[,])  
> print(review_score)  
      anger anticipation      disgust      fear      joy      sadness      surprise      trust  
      241      1772         201      286      1882         361         868      2252  
negative      positive      score  
      672      3989      3317  
> |
```

Data Visualization and Interpretation

Data Visualization and Interpretation

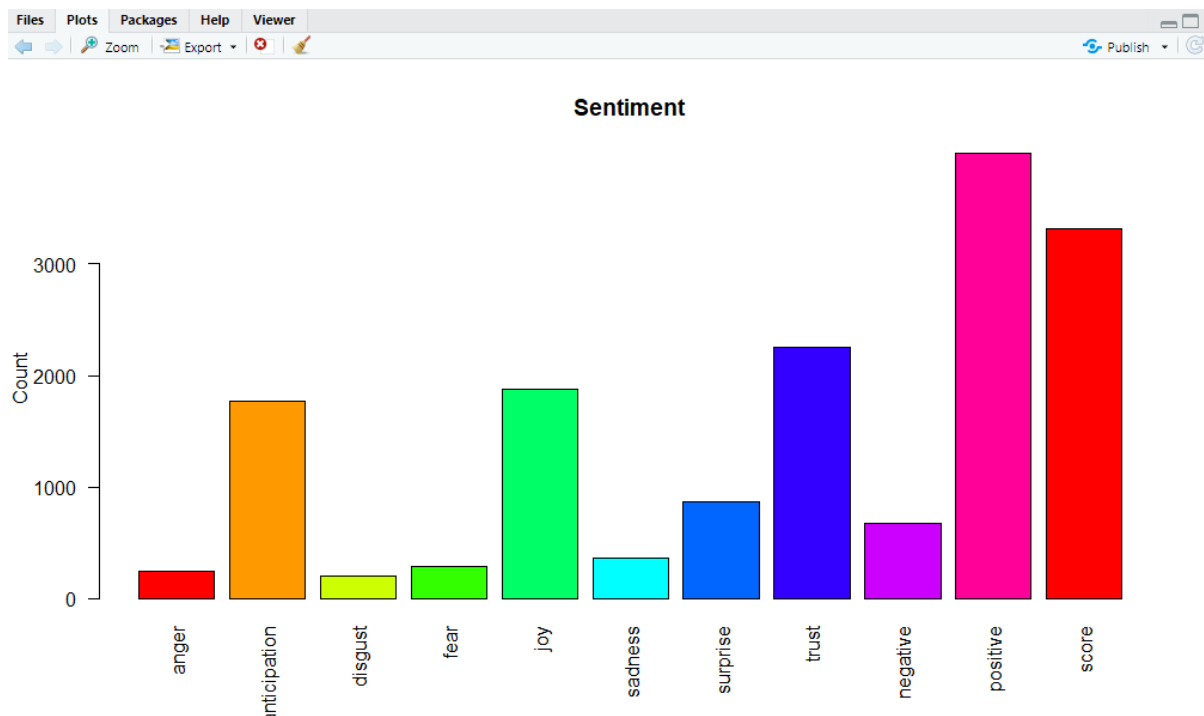
5.1 Creating a Word Cloud of TDM

```
w <- sort(rowSums(tdm), decreasing = T)
set.seed(2000)
wordcloud(words = names(w),
freq = w,
max.words = 50,
random.order = T,
min.freq = 5,
colors = brewer.pal(25,"Dark2"),
scale = c(3,0.3))
```



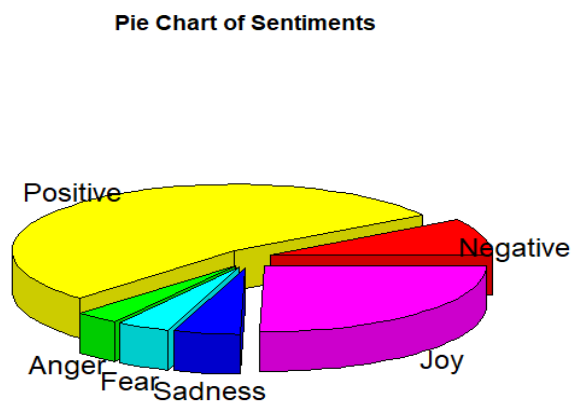
5.2 Creating a Bar plot of the sentiments

```
barplot(colSums(s),
las = 2,
col = rainbow(10),
ylab = 'Count',
main = 'Sentiment')
```



5.3 Creating a 3D Pie chart of selected few sentiments

```
neg_score <- colSums(s[9])  
pos_score <- colSums(s[10])  
anger_score <- colSums(s[1])  
fear_score <- colSums(s[4])  
sadness_score <- colSums(s[6])  
joy_score <- colSums(s[5])  
  
x <- c(neg_score, pos_score, anger_score, fear_score, sadness_score, joy_score)  
labels <- c("Negative", "Positive", "Anger", "Fear", "Sadness", "Joy")  
  
pie3D(x, labels = labels, explode = 0.1, main = "Pie Chart of Sentiments ")
```



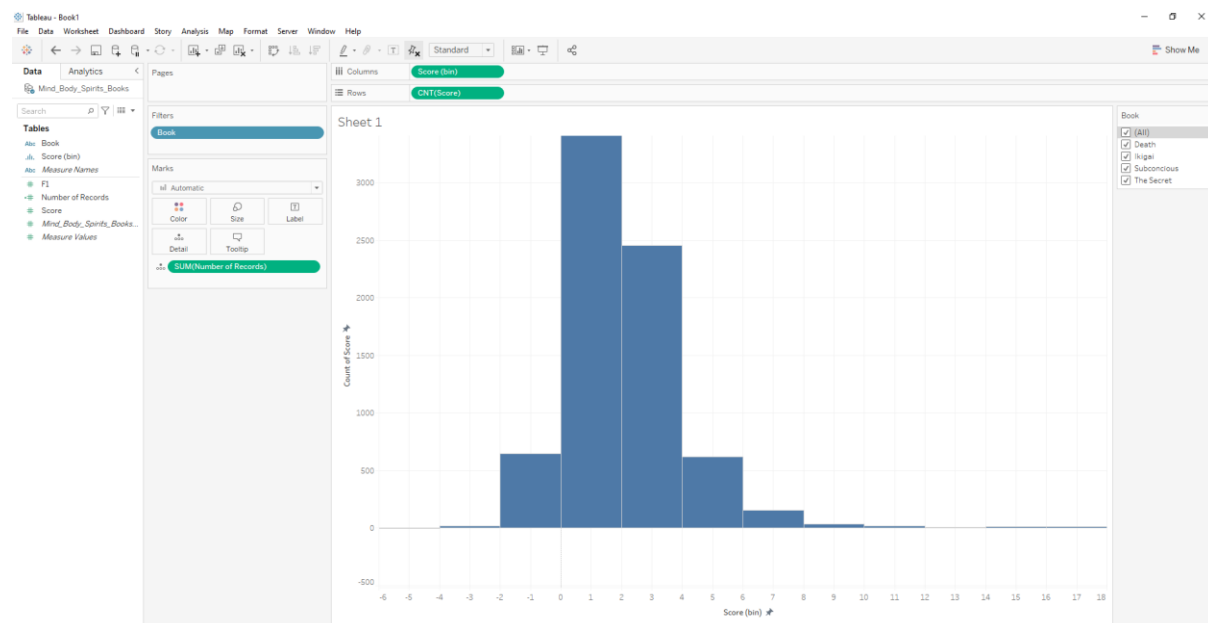
5.3 More to the list.

Adding more books / items from amazon reviews extractor and following the process all over again for these items. Then we merge all the CSV files of the sentiment scores together to get a single CSV file.

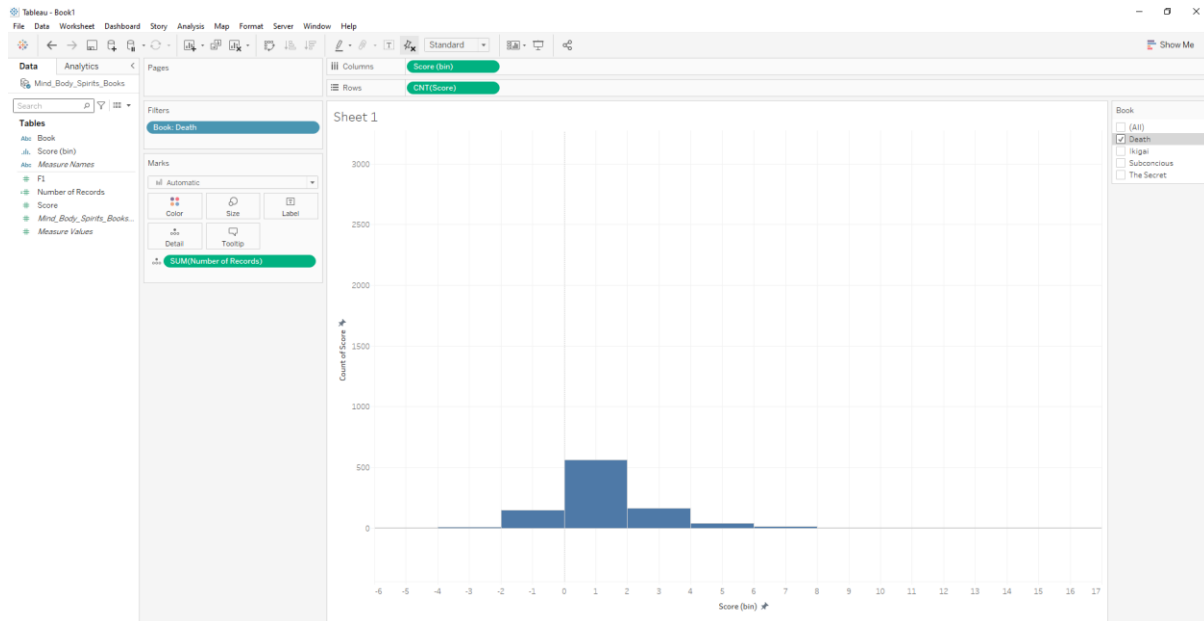
1		score	Book	
2	1	-1	Death	
3	2	6	Death	
4	3	8	Death	
5	4	6	Death	
6	5	5	Death	
7	6	-2	Death	
8	7	3	Death	
9	8	17	Death	
10	9	3	Death	
11	10	2	Death	
12	11	0	Death	
13	12	9	Death	
14	13	4	Death	
15	14	1	Death	
16	15	7	Death	
17	16	9	Death	
18	17	-1	Death	

5.4 Visualizing data in Tableau

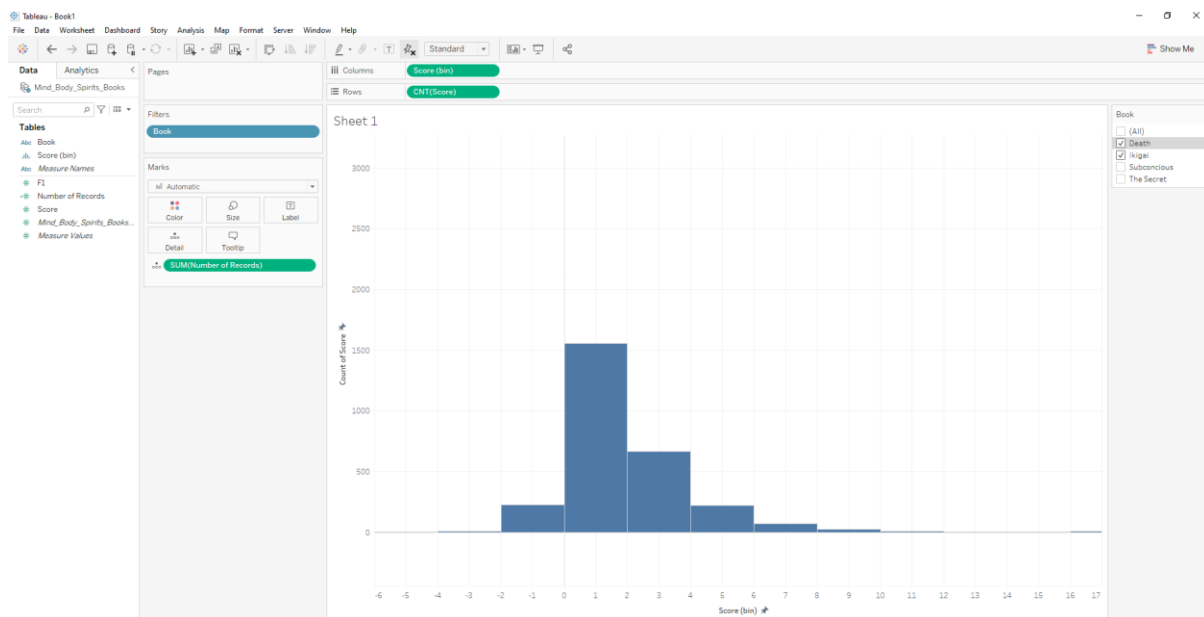
5.4.1 Filter for all books



5.4.2 Filter for book 1 (Death)



5.4.3 Filter for combination of Multiple books



Conclusion

Conclusion:

Sentiment analysis or opinion mining is one of the major tasks of NLP (Natural Language Processing). Sentiment analysis has gained much attention in recent years. In this paper, we aim to tackle the problem of sentiment polarity categorization, which is one of the fundamental problems of sentiment analysis. A general process for sentiment polarity categorization is proposed with detailed process descriptions. Data used in this study are online product reviews collected from Amazon.com. Experiments for both sentence-level categorization and review-level categorization are performed with promising outcomes. At last, we also give insight into our future work on sentiment analysis. Sentiment analysis or opinion mining is a field of study that analyses people's sentiments, attitudes, or emotions towards certain entities. This paper tackles a fundamental problem of sentiment analysis, sentiment polarity categorization. Online product reviews from Amazon.com are selected as data used for this study. A sentiment polarity categorization process has been proposed along with detailed descriptions of each step. Experiments for both sentence-level categorization and review-level categorization have been performed.

ACKNOWLEDGEMENT

We would like to express our gratitude to our guide Mrs. Vandana Patil who gave us the opportunity to do a project on, “Sentiment analysis on Amazon Reviews using R”. Her constant guidance and prompt suggestions have helped us make progress in our project. We would also like to extend our gratitude to the Head of Department (INFT) Dr. Joanne Gomes and our Principal and Director for providing us with the necessary facilities required for the project.