plague_sent_anal

shimmy

8/3/2020

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------------

## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.2     v dplyr   1.0.0
## v tidyr   1.1.1     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0

## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(tidytext)
library(janeaustenr)
library(stringr)
library(readtext)
library(readr)
library(summarytools)
```

```
## Registered S3 method overwritten by 'pryr':
##   method      from
##   print.bytes Rcpp

## For best results, restart R session and update pander using devtools:: or remotes::install_github('ra

##
## Attaching package: 'summarytools'

## The following object is masked from 'package:tibble':
##
##     view
```

```
library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last

## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(wesanderson)
```

```
plague <- read.csv("plague_by_part_by_word.csv") %>% rename(no_punct = new_column) %>% drop_na(no_punct)
plague <- plague %>% subset(no_punct != "")
```

```
# p2 <- read.csv("df2.csv")
# p3 <- read.csv("df3.csv")
# p4 <- read.csv("df4.csv")
# p5 <- read.csv("df5.csv")
nrc <- get_sentiments("nrc") %>% rename(no_punct = word)
bing <- get_sentiments("bing") %>% rename(no_punct = word)
afin <- get_sentiments("afin") %>% rename(no_punct = word)
```

```
#plague %>% filter(Word == "plague" | Word == "Plague") %>% ggplot(aes(x = count)) + geom_histogram()
```

```
plague <-  plague%>% left_join(nrc, by = "no_punct")
plague <-  plague%>% left_join(bing, by = "no_punct")
plague <-  plague%>% left_join(afin, by = "no_punct")
plague <- plague %>% rename(nrc_sent = sentiment.x, afin_sent = sentiment.y)
# total5_sent$rows <- total5_sent%>% row.names()
# p1_sent <- p1 %>% left_join(sents, by = "Word")
# p2_sent <- p2 %>% left_join(sents, by = "Word")
# p3_sent <- p3 %>% left_join(sents, by = "Word")
# p4_sent <- p4 %>% left_join(sents, by = "Word")
# p5_sent <- p5 %>% left_join(sents, by = "Word")
# p1_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment != "negative") %>% ggplot((a
# p2_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment != "negative") %>% ggplot((a
# p3_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment != "negative") %>% ggplot((a
# p4_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment != "negative") %>% ggplot((a
# p5_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment != "negative") %>% ggplot((a
# total5_sent_filtered <- total5_sent %>% filter(!is.na(sentiment), sentiment != "positive", sentiment
# total5_sent_filtered$num <- total5_sent_filtered %>% row.names() %>% as.numeric()
# total5_sent_filtered %>% ggplot(aes(x = num))+ geom_histogram(bins = 50) + facet_wrap(~sentiment) +sc
```

```
think <- c(
"think",
"consider",
"determine",
"expect",
"feel",
"guess",
"judge",
"realize",
"see",
```

```
"take",
"understand",
"comprehend",
"conceive",
"conclude",
"credit",
"deem",
"envisage",
"envision",
"esteem",
"estimate",
"fancy",
"feature",
"foresee",
"gather",
"hold",
"image
imagine",
"presume",
"project",
"reckon",
"regard",
"sense",
"suppose",
"surmise",
"suspect",
"vision",
"visualize"
) %>% as.data.frame() %>% rename(no_punct = ".")
think$thought = rep("think", nrow(think))
plague <- plague%>% left_join(think,  by = "no_punct")
# total5_think %>%filter(!is.na(thought)) %>% ggplot(aes(x = rows)) + geom_histogram(bins = 20)
```

```
selfish <- c("egotistical",
"greedy",
"narcissistic",
"self-centered",
"egocentric",
"egoistic",
"egoistical",
"egomaniacal",
"egotistic",
"hoggish",
"mean",
"mercenary",
"miserly",
"narrow",
"parsimonious",
"prejudiced",
"self-indulgent",
"self-interested",
"self-seeking",
"stingy",
```

```
"ungenerous",
"think") %>% as.data.frame() %>% rename(no_punct = ".")
selfish$selfish <- rep("selfish", nrow(selfish))
plague <- plague %>% left_join(selfish, by = "no_punct")
# total5_selfish <- total5 %>% left_join(selfish, by = "Word")
# total5_selfish %>%filter(!is.na(selfish)) %>% ggplot(aes(x = rows)) + geom_histogram(bins = 20)
```

```
love <- c("love",
  "affection",
"appreciation",
"devotion",
"emotion",
"fondness",
"friendship",
"infatuation",
"lust",
"passion",
"respect",
"taste",
"tenderness",
"yearning",
"adulation",
"allegiance",
"amity",
"amorousness",
"amour",
"ardor",
"attachment",
"case",
"cherishing",
"crush",
"delight",
"devotedness",
"enchantment",
"enjoyment",
"fervor",
"fidelity",
"flame",
"hankering",
"idolatry",
"inclination",
"involvement",
"like",
"partiality",
"piety",
"rapture",
"regard",
"relish",
"sentiment",
"weakness",
"worship",
"zeal",
"ardency")%>% as.data.frame() %>% rename(no_punct = ".")
```

```r
love$love <- rep("love", nrow(love))
# total5_love <- total5 %>% left_join(love, by = "no_punct")
# total5_love %>%filter(!is.na(love)) %>% ggplot(aes(x = rows)) + geom_histogram()
plague <- plague %>% left_join(love, by = "no_punct")
```

```
{r} # # total5_sent_freq_part_1 <- total5_sent %>%filter(part
== "part_1") %>% freq(sentiment) %>% as.data.frame() %>% rename(percent_
= "% Total", percent_valid = "% Valid", percent_valid_cum = "%
Valid Cum.", percent_total_cum = "% Total Cum.") %>% setattr("row.names"
c("anger",        "anticipation", "disgust",      "fear",         "joy",
"negative",     "positive",     "sadness",      "surprise",
"trust",  "non_avail",      "Total")) # # total5_sent_freq_part_2
<- total5_sent %>%filter(part == "part_2") %>% freq(sentiment)
%>% as.data.frame() %>% rename(percent_total = "% Total", percent_valid
= "% Valid", percent_valid_cum = "% Valid Cum.", percent_total_cum
= "% Total Cum.") %>% setattr("row.names",  c("anger",        "anticipat
"disgust",      "fear",         "joy",          "negative",
"positive",     "sadness",      "surprise",     "trust",  "non_avail",
"Total")) # # total5_sent_freq_part_3<- total5_sent %>%filter(part
== "part_3") %>% freq(sentiment) %>% as.data.frame() %>% rename(percent_
= "% Total", percent_valid = "% Valid", percent_valid_cum = "%
Valid Cum.", percent_total_cum = "% Total Cum.") %>% setattr("row.names"
c("anger",        "anticipation", "disgust",      "fear",         "joy",
"negative",     "positive",     "sadness",      "surprise",
"trust",  "non_avail",      "Total")) # # total5_sent_freq_part_4
<- total5_sent %>%filter(part == "part_4") %>% freq(sentiment)
%>% as.data.frame() %>% rename(percent_total = "% Total", percent_valid
= "% Valid", percent_valid_cum = "% Valid Cum.", percent_total_cum
= "% Total Cum.") %>% setattr("row.names",  c("anger",        "anticipat
"disgust",      "fear",         "joy",          "negative",
"positive",     "sadness",      "surprise",     "trust",  "non_avail",
"Total")) # # total5_sent_freq_part_5 <- total5_sent %>%filter(part
== "part_5") %>% freq(sentiment) %>% as.data.frame() %>% rename(percent_
= "% Total", percent_valid = "% Valid", percent_valid_cum = "%
Valid Cum.", percent_total_cum = "% Total Cum.") %>% setattr("row.names"
c("anger",        "anticipation", "disgust",      "fear",         "joy",
"negative",     "positive",     "sadness",      "surprise",
"trust",  "non_avail",      "Total")) # #  # # total5_sent_freq
<- total5_sent %>% freq(sentiment) %>% as.data.frame() %>%
rename(percent_total = "% Total", percent_valid = "% Valid",
percent_valid_cum = "% Valid Cum.", percent_total_cum = "%
Total Cum.") %>% setattr("row.names",  c("anger",        "anticipation",
"disgust",      "fear",         "joy",          "negative",
"positive",     "sadness",      "surprise",     "trust",  "non_avail",
"Total")) #
```

```
anxiety <- c("angst",
"apprehension",
"concern",
"disquiet",
"doubt",
"dread",
"jitters",
"misery",
"misgiving",
"mistrust",
"nervousness",
"panic",
"restlessness",
"suspense",
"trouble",
"uncertainty",
"unease",
"uneasiness",
"botheration",
"butterflies",
"care",
"creeps",
"disquietude",
"distress",
"downer",
"drag",
"fidgets",
"flap",
"foreboding",
"fretfulness",
"fuss",
"heebie-jeebies",
"jumps",
"needles",
"shakes",
"shivers",
"solicitude",
"watchfulness",
"willies",
"worriment",
"all-overs",
"nail-biting",
"anxiety"
) %>% as.data.frame() %>% rename(no_punct = ".")
anxiety$anxiety<- rep("anxiety", nrow(anxiety))
plague <- plague %>% left_join(anxiety, by = "no_punct")
#plague %>% filter(!is.na(anxiety)) %>% ggplot(aes(x = count)) + geom_histogram(bins = 10)
```

```
#plague %>% filter(!is.na(value)) %>% ggplot(aes(x = count, y = value)) + geom_point()
```

```
we <-c("we") %>% as.data.frame() %>% rename(no_punct = ".")
we$we <- rep("we", nrow(we))
```

```
plague <- plague %>% left_join(we, by = "no_punct")
```

```
exile <- c("exile",
           "banishment",
            "diaspora",
            "dispersion",
            "displacement",
"exclusion",
"expatriation",
"expulsion",
"extradition",
"migration",
"ostracism",
"proscription",
"relegation",
"scattering",
"separation") %>% as.data.frame() %>% rename(no_punct = ".")
exile$exile<- rep("exile", nrow(exile))
plague <- plague %>% left_join(exile, by = "no_punct")
```

```
suffer <- c("suffer", "adversity",
"anguish",
"difficulty",
"discomfort",
"hardship",
"misery",
"misfortune",
"ordeal",
"torment",
"torture",
"affliction",
"distress",
"dolor",
"martyrdom",
"passion") %>% as.data.frame() %>% rename(no_punct = ".")
suffer$suffer <- rep("suffer", nrow(suffer))
plague <- plague %>% left_join(suffer, by = "no_punct")
```

```
# total5_sent_freq_part_1$row_names <- total5_sent_freq %>% rownames()
# total5_sent_freq_part_2$row_names <- total5_sent_freq %>% rownames()
# total5_sent_freq_part_3$row_names <- total5_sent_freq %>% rownames()
# total5_sent_freq_part_4$row_names <- total5_sent_freq %>% rownames()
# total5_sent_freq_part_5$row_names <- total5_sent_freq %>% rownames()
# total5_sent_freq$row_names <- total5_sent_freq %>% rownames()
```

```
#plague %>% group_by(Word) %>% View()
```

```
lonely <- c("lonely", "deserted",
"desolate",
"destitute",
"empty",
"homeless",
```

```
"isolated",
"lonesome",
"reclusive",
"solitary",
"abandoned",
"alone",
"apart",
"comfortless",
"companionless",
"disconsolate",
"down",
"estranged",
"forsaken",
"godforsaken",
"left",
"lone",
"outcast",
"rejected",
"renounced",
"secluded",
"single",
"troglodytic",
"unattended",
"unbefriended",
"uncherished",
"unsocial",
"withdrawn")%>% as.data.frame() %>% rename(no_punct = ".")
lonely$lonely<- rep("lonely", nrow(lonely))
plague<-plague %>% left_join(lonely, by = "no_punct")
```

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

```
# total5_sent_freq_part_1  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar
# total5_sent_freq_part_2  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar
# total5_sent_freq_part_3  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar
# total5_sent_freq_part_4  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar
# total5_sent_freq_part_5  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar
#
# total5_sent_freq  %>% slice(1:10) %>% ggplot(aes(x = row_names, y = percent_total)) + geom_bar(stat="

plague %>% write.csv("plague_edited_df.csv")
```

```
plague %>% filter(!is.na(nrc_sent)) %>% group_by(part) %>% freq(nrc_sent)
```

```
## Frequencies
## plague$nrc_sent
## Type: Character
## Group: part = part_1
##
```

```
##                     Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------------ ------ --------- -------------- --------- --------------
##              anger    173      5.77           5.77      5.77           5.77
##       anticipation    261      8.71          14.49      8.71          14.49
##            disgust    185      6.17          20.66      6.17          20.66
##               fear    301     10.05          30.71     10.05          30.71
##                joy    168      5.61          36.32      5.61          36.32
##           negative    495     16.52          52.84     16.52          52.84
##           positive    616     20.56          73.40     20.56          73.40
##            sadness    280      9.35          82.74      9.35          82.74
##           surprise    130      4.34          87.08      4.34          87.08
##              trust    387     12.92         100.00     12.92         100.00
##               <NA>      0                               0.00         100.00
##              Total   2996    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                     Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------------ ------ --------- -------------- --------- --------------
##              anger    282      5.37           5.37      5.37           5.37
##       anticipation    533     10.15          15.52     10.15          15.52
##            disgust    269      5.12          20.64      5.12          20.64
##               fear    498      9.48          30.12      9.48          30.12
##                joy    373      7.10          37.22      7.10          37.22
##           negative    834     15.88          53.10     15.88          53.10
##           positive   1041     19.82          72.92     19.82          72.92
##            sadness    519      9.88          82.81      9.88          82.81
##           surprise    271      5.16          87.97      5.16          87.97
##              trust    632     12.03         100.00     12.03         100.00
##               <NA>      0                               0.00         100.00
##              Total   5252    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                     Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------------ ------ --------- -------------- --------- --------------
##              anger    128      7.34           7.34      7.34           7.34
##       anticipation    139      7.97          15.30      7.97          15.30
##            disgust    122      6.99          22.29      6.99          22.29
##               fear    232     13.30          35.59     13.30          35.59
##                joy     83      4.76          40.34      4.76          40.34
##           negative    327     18.74          59.08     18.74          59.08
##           positive    247     14.15          73.24     14.15          73.24
##            sadness    229     13.12          86.36     13.12          86.36
##           surprise     64      3.67          90.03      3.67          90.03
##              trust    174      9.97         100.00      9.97         100.00
##               <NA>      0                               0.00         100.00
##              Total   1745    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                     Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------------ ------ --------- -------------- --------- --------------
##              anger    292      6.23           6.23      6.23           6.23
```

```
##       anticipation    441      9.40          15.63      9.40          15.63
##            disgust    281      5.99          21.62      5.99          21.62
##               fear    510     10.87          32.49     10.87          32.49
##                joy    306      6.52          39.02      6.52          39.02
##           negative    804     17.14          56.16     17.14          56.16
##           positive    824     17.57          73.73     17.57          73.73
##            sadness    473     10.09          83.82     10.09          83.82
##           surprise    224      4.78          88.59      4.78          88.59
##              trust    535     11.41         100.00     11.41         100.00
##               <NA>      0                                0.00         100.00
##              Total   4690    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                     Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------------ ------ --------- -------------- --------- --------------
##              anger    192      6.28           6.28      6.28           6.28
##       anticipation    325     10.64          16.92     10.64          16.92
##            disgust    160      5.24          22.16      5.24          22.16
##               fear    313     10.25          32.41     10.25          32.41
##                joy    233      7.63          40.03      7.63          40.03
##           negative    491     16.07          56.10     16.07          56.10
##           positive    526     17.22          73.32     17.22          73.32
##            sadness    315     10.31          83.63     10.31          83.63
##           surprise    154      5.04          88.67      5.04          88.67
##              trust    346     11.33         100.00     11.33         100.00
##               <NA>      0                                0.00         100.00
##              Total   3055    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(we)
```

```
## Frequencies
## plague$we
## Type: Character
## Group: part = part_1
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##          we      19    100.00         100.00      0.12           0.12
##        <NA>   16200                             99.88         100.00
##       Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##          we      32    100.00         100.00      0.12           0.12
##        <NA>   26012                             99.88         100.00
##       Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
```

```
##          we     10    100.00         100.00      0.14           0.14
##        <NA>   6951                             99.86         100.00
##       Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##          we      39    100.00         100.00      0.17           0.17
##        <NA>   22254                             99.83         100.00
##       Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##          we       9    100.00         100.00      0.07           0.07
##        <NA>   13433                             99.93         100.00
##       Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(love)
```

```
## Frequencies
## plague$love
## Type: Character
## Group: part = part_1
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##        love      84    100.00         100.00      0.52           0.52
##        <NA>   16135                             99.48         100.00
##       Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##        love     141    100.00         100.00      0.54           0.54
##        <NA>   25903                             99.46         100.00
##       Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##        love     48    100.00         100.00      0.69           0.69
##        <NA>   6913                             99.31         100.00
##       Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##        love     109    100.00         100.00      0.49           0.49
```

```
##        <NA>   22184                           99.51         100.00
##       Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##        love      85    100.00         100.00      0.63           0.63
##        <NA>   13357                           99.37         100.00
##       Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(anxiety)
```

```
## Frequencies
## plague$anxiety
## Type: Character
## Group: part = part_1
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       anxiety      55    100.00         100.00      0.34           0.34
##          <NA>   16164                           99.66         100.00
##         Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       anxiety      82    100.00         100.00      0.31           0.31
##          <NA>   25962                           99.69         100.00
##         Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------ --------- -------------- --------- --------------
##       anxiety     35    100.00         100.00      0.50           0.50
##          <NA>   6926                           99.50         100.00
##         Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       anxiety      51    100.00         100.00      0.23           0.23
##          <NA>   22242                           99.77         100.00
##         Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       anxiety      51    100.00         100.00      0.38           0.38
##          <NA>   13391                           99.62         100.00
```

```
##         Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(selfish)
```

```
## Frequencies
## plague$selfish
## Type: Character
## Group: part = part_1
##
##                  Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------- ------- --------- -------------- --------- --------------
##       selfish       7    100.00         100.00      0.04           0.04
##          <NA>   16212                             99.96         100.00
##         Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                  Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------- ------- --------- -------------- --------- --------------
##       selfish       8    100.00         100.00      0.03           0.03
##          <NA>   26036                             99.97         100.00
##         Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------- ------ --------- -------------- --------- --------------
##       selfish      2    100.00         100.00      0.03           0.03
##          <NA>   6959                             99.97         100.00
##         Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                  Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------- ------- --------- -------------- --------- --------------
##       selfish       5    100.00         100.00      0.02           0.02
##          <NA>   22288                             99.98         100.00
##         Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                  Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------- ------- --------- -------------- --------- --------------
##       selfish       2    100.00         100.00      0.01           0.01
##          <NA>   13440                             99.99         100.00
##         Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>% filter(!is.na(value)) %>% group_by(part) %>% freq(value)
```

```
## Frequencies
## plague$value
## Type: Numeric
## Group: part = part_1
```

```
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##          -4      1      0.07           0.07      0.07           0.07
##          -3    151     11.30          11.38     11.30          11.38
##          -2    377     28.22          39.60     28.22          39.60
##          -1    176     13.17          52.77     13.17          52.77
##           1    229     17.14          69.91     17.14          69.91
##           2    246     18.41          88.32     18.41          88.32
##           3    153     11.45          99.78     11.45          99.78
##           4      3      0.22         100.00      0.22         100.00
##        <NA>      0                                0.00         100.00
##       Total   1336    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##          -5      2      0.08           0.08      0.08           0.08
##          -4      5      0.21           0.30      0.21           0.30
##          -3    163      6.87           7.17      6.87           7.17
##          -2    617     26.01          33.18     26.01          33.18
##          -1    413     17.41          50.59     17.41          50.59
##           1    387     16.32          66.91     16.32          66.91
##           2    469     19.77          86.68     19.77          86.68
##           3    295     12.44          99.11     12.44          99.11
##           4     20      0.84          99.96      0.84          99.96
##           5      1      0.04         100.00      0.04         100.00
##        <NA>      0                                0.00         100.00
##       Total   2372    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##          -4      1      0.13           0.13      0.13           0.13
##          -3     92     11.96          12.09     11.96          12.09
##          -2    261     33.94          46.03     33.94          46.03
##          -1    128     16.64          62.68     16.64          62.68
##           1     94     12.22          74.90     12.22          74.90
##           2    112     14.56          89.47     14.56          89.47
##           3     74      9.62          99.09      9.62          99.09
##           4      7      0.91         100.00      0.91         100.00
##        <NA>      0                                0.00         100.00
##       Total    769    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##          -4     25      1.16           1.16      1.16           1.16
##          -3    226     10.52          11.69     10.52          11.69
##          -2    638     29.70          41.39     29.70          41.39
##          -1    341     15.88          57.26     15.88          57.26
```

```
##           1    310     14.43          71.69     14.43          71.69
##           2    377     17.55          89.25     17.55          89.25
##           3    223     10.38          99.63     10.38          99.63
##           4      8      0.37         100.00      0.37         100.00
##        <NA>      0                                0.00         100.00
##       Total   2148    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##          -3    113      7.72           7.72      7.72           7.72
##          -2    353     24.13          31.85     24.13          31.85
##          -1    237     16.20          48.05     16.20          48.05
##           1    174     11.89          59.95     11.89          59.95
##           2    341     23.31          83.25     23.31          83.25
##           3    202     13.81          97.06     13.81          97.06
##           4     43      2.94         100.00      2.94         100.00
##        <NA>      0                                0.00         100.00
##       Total   1463    100.00         100.00    100.00         100.00
```

```
plague %>%filter(!is.na(afin_sent)) %>% group_by(part) %>% freq(afin_sent)
```

```
## Frequencies
## plague$afin_sent
## Type: Character
## Group: part = part_1
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## -------------- ------ --------- -------------- --------- --------------
##       negative   1143     66.96          66.96     66.96          66.96
##       positive    564     33.04         100.00     33.04         100.00
##           <NA>      0                                0.00         100.00
##          Total   1707    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## -------------- ------ --------- -------------- --------- --------------
##       negative   1870     64.53          64.53     64.53          64.53
##       positive   1028     35.47         100.00     35.47         100.00
##           <NA>      0                                0.00         100.00
##          Total   2898    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## -------------- ------ --------- -------------- --------- --------------
##       negative    819     74.59          74.59     74.59          74.59
##       positive    279     25.41         100.00     25.41         100.00
##           <NA>      0                                0.00         100.00
##          Total   1098    100.00         100.00    100.00         100.00
##
## Group: part = part_4
```

```
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## -------------- ------ --------- -------------- --------- --------------
##       negative   1934     69.77          69.77     69.77          69.77
##       positive    838     30.23         100.00     30.23         100.00
##           <NA>      0                                0.00         100.00
##          Total   2772    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                 Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## -------------- ------ --------- -------------- --------- --------------
##       negative   1226     63.59          63.59     63.59          63.59
##       positive    702     36.41         100.00     36.41         100.00
##           <NA>      0                                0.00         100.00
##          Total   1928    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(suffer)
```

```
## Frequencies
## plague$suffer
## Type: Character
## Group: part = part_1
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       suffer      16    100.00         100.00      0.10           0.10
##         <NA>   16203                               99.90         100.00
##        Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       suffer      78    100.00         100.00      0.30           0.30
##         <NA>   25966                               99.70         100.00
##        Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------ --------- -------------- --------- --------------
##       suffer     30    100.00         100.00      0.43           0.43
##         <NA>   6931                               99.57         100.00
##        Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       suffer      51    100.00         100.00      0.23           0.23
##         <NA>   22242                               99.77         100.00
##        Total   22293    100.00         100.00    100.00         100.00
##
```

```
## Group: part = part_5
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       suffer      15    100.00         100.00      0.11           0.11
##         <NA>   13427                             99.89         100.00
##        Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>% group_by(part) %>% freq(suffer) %>% saveRDS("suffer_freq.rds")
```

```
plague %>% group_by(part) %>% freq(exile) #%>% ggplot(aes(x = ))
```

```
## Frequencies
## plague$exile
## Type: Character
## Group: part = part_1
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##        <NA>   16219                             100.00         100.00
##       Total   16219      0.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##       exile      35    100.00         100.00      0.13           0.13
##        <NA>   26009                             99.87         100.00
##       Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##               Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------ --------- -------------- --------- --------------
##       exile     19    100.00         100.00      0.27           0.27
##        <NA>   6942                             99.73         100.00
##       Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##       exile       2    100.00         100.00      0.01           0.01
##        <NA>   22291                             99.99         100.00
##       Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ----------- ------- --------- -------------- --------- --------------
##       exile      38    100.00         100.00      0.28           0.28
##        <NA>   13404                             99.72         100.00
##       Total   13442    100.00         100.00    100.00         100.00
```

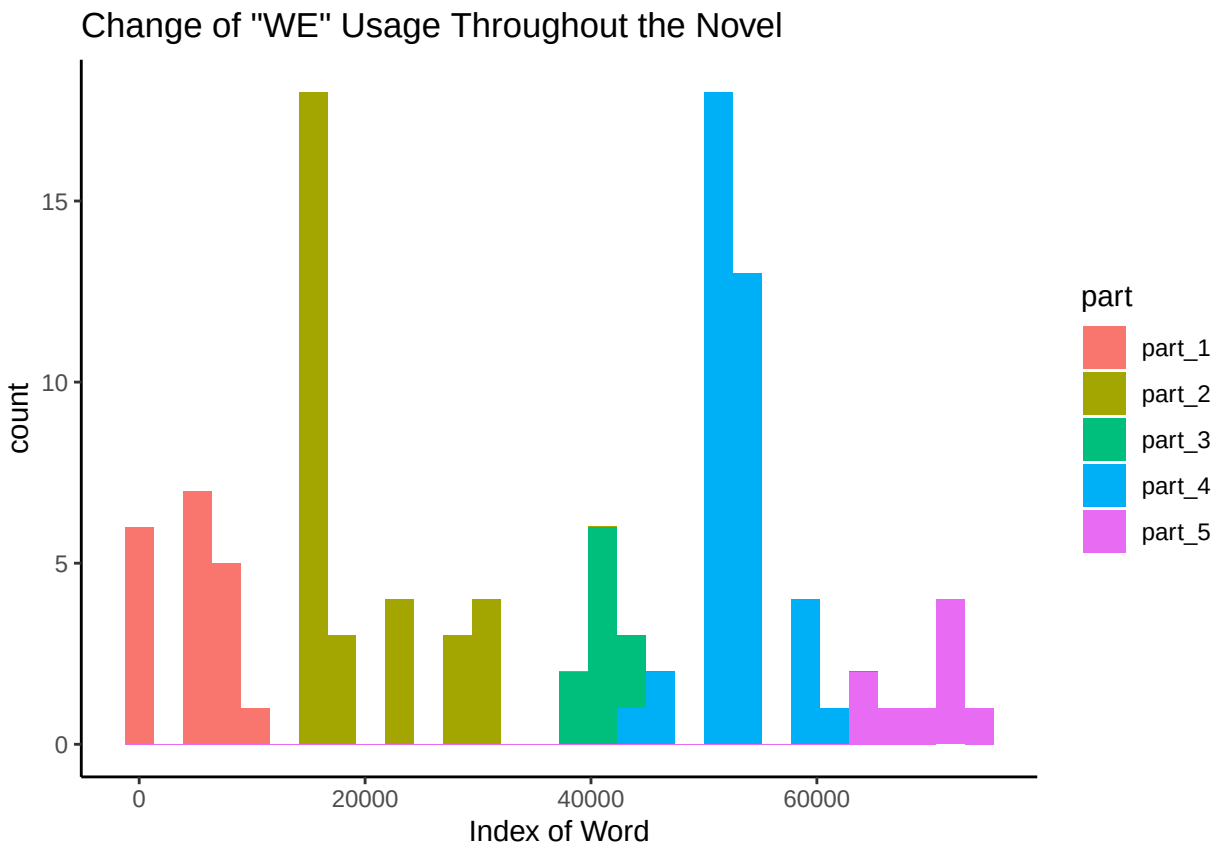
```
plague %>% group_by(part) %>% freq(lonely)
```

```
## Frequencies
## plague$lonely
## Type: Character
## Group: part = part_1
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       lonely      43    100.00         100.00      0.27           0.27
##         <NA>   16176                             99.73         100.00
##        Total   16219    100.00         100.00    100.00         100.00
##
## Group: part = part_2
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       lonely      92    100.00         100.00      0.35           0.35
##         <NA>   25952                             99.65         100.00
##        Total   26044    100.00         100.00    100.00         100.00
##
## Group: part = part_3
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------ --------- -------------- --------- --------------
##       lonely     14    100.00         100.00      0.20           0.20
##         <NA>   6947                             99.80         100.00
##        Total   6961    100.00         100.00    100.00         100.00
##
## Group: part = part_4
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       lonely      74    100.00         100.00      0.33           0.33
##         <NA>   22219                             99.67         100.00
##        Total   22293    100.00         100.00    100.00         100.00
##
## Group: part = part_5
##
##                Freq   % Valid   % Valid Cum.   % Total   % Total Cum.
## ------------ ------- --------- -------------- --------- --------------
##       lonely      47    100.00         100.00      0.35           0.35
##         <NA>   13395                             99.65         100.00
##        Total   13442    100.00         100.00    100.00         100.00
```

```
plague %>%filter(!is.na(we))%>% ggplot(aes(x = count, fill = part)) + geom_histogram() + ggtitle("Chang
```

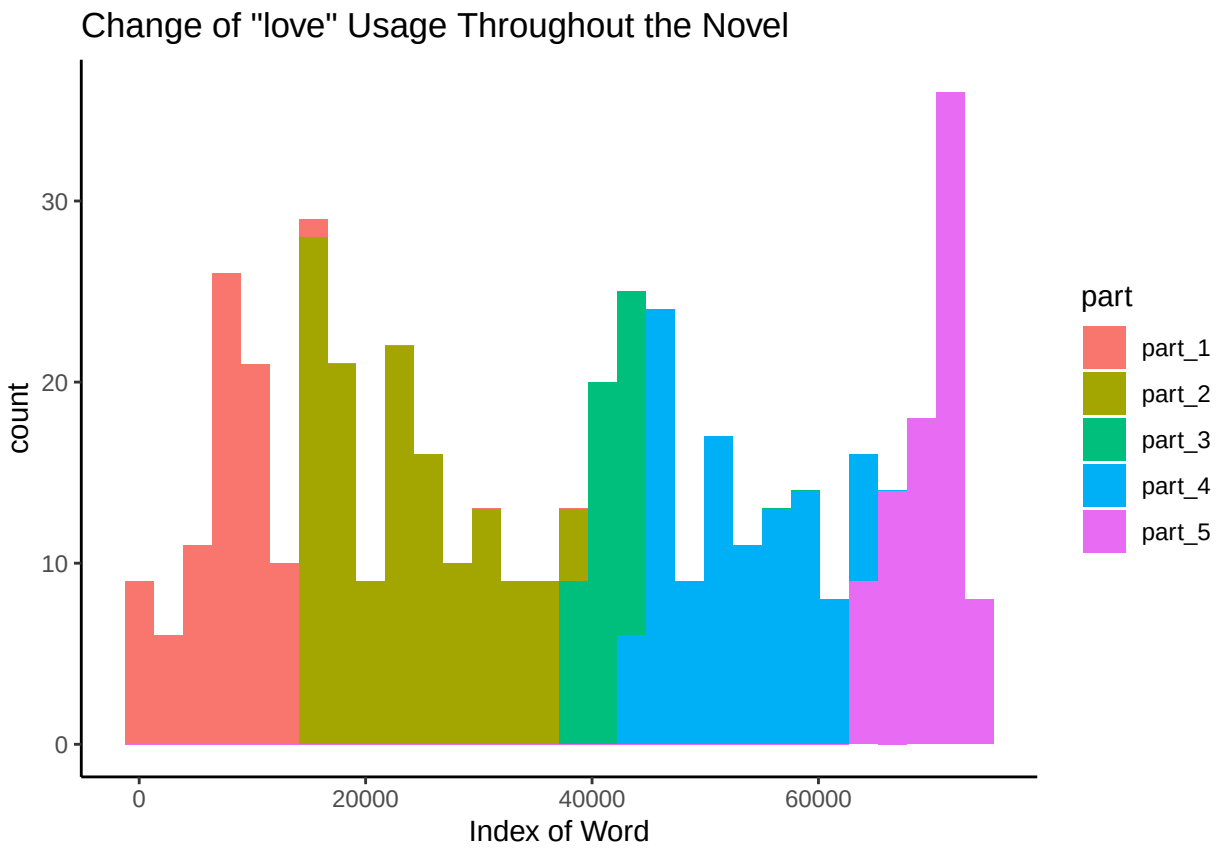
```
## Saving 6.5 x 4.5 in image
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



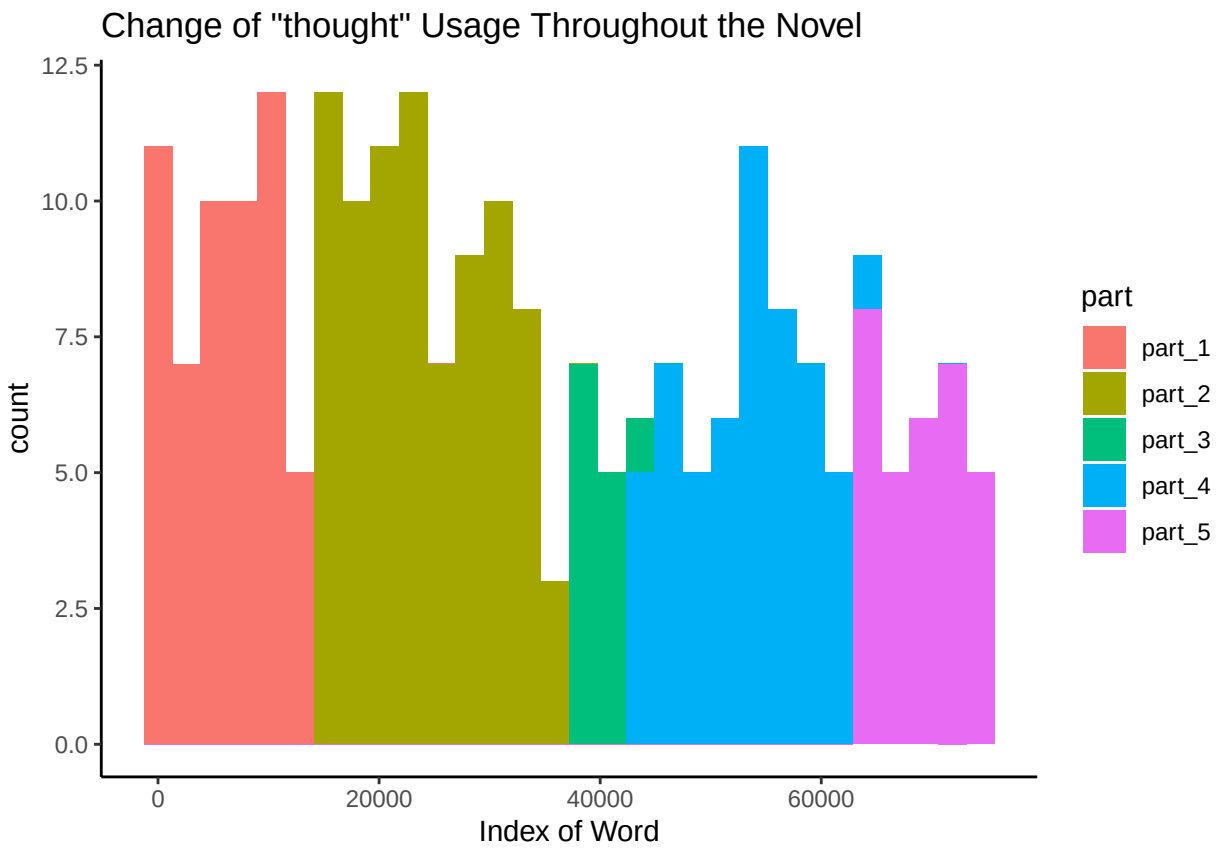

```
plague %>%filter(!is.na(love))%>% ggplot(aes(x = count, , fill = part)) + geom_histogram() + ggtitle("Cl
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



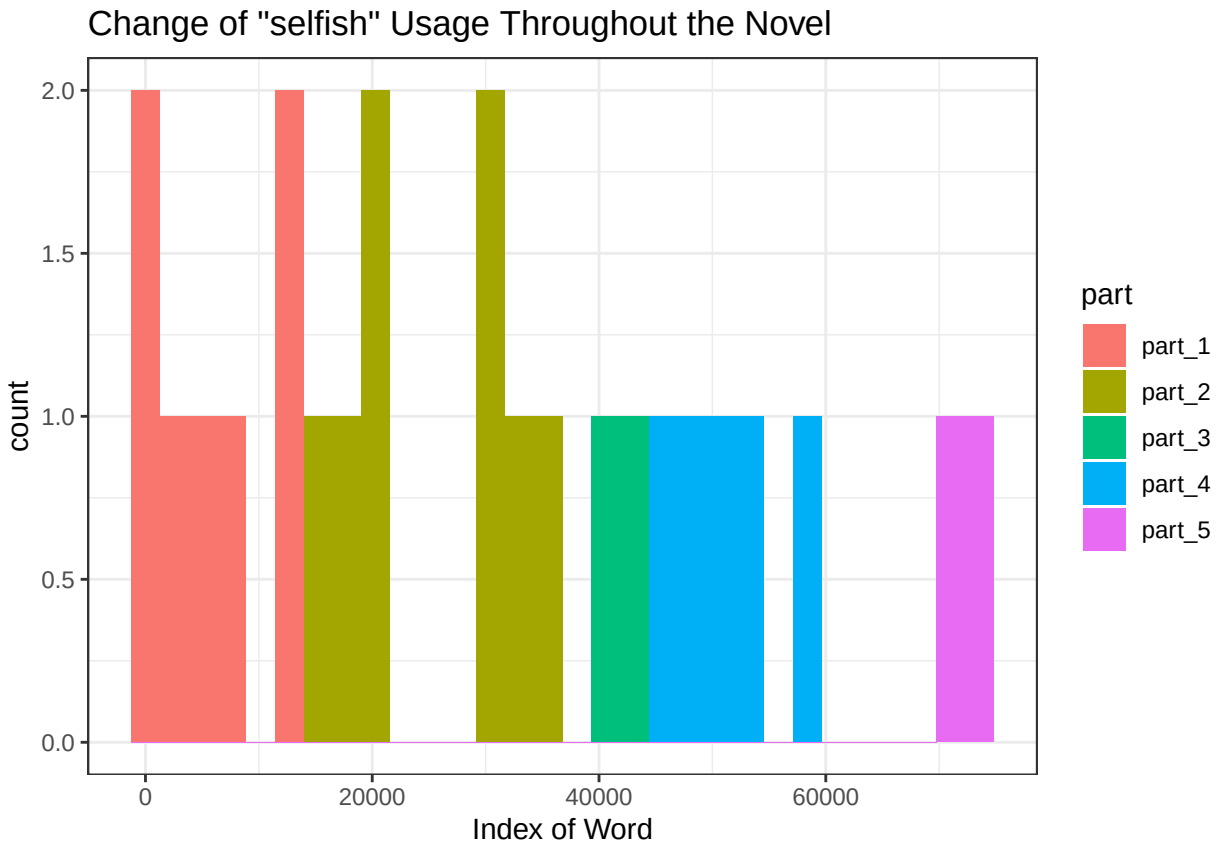

```
plague %>%filter(!is.na(thought))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() +  ggtitle
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



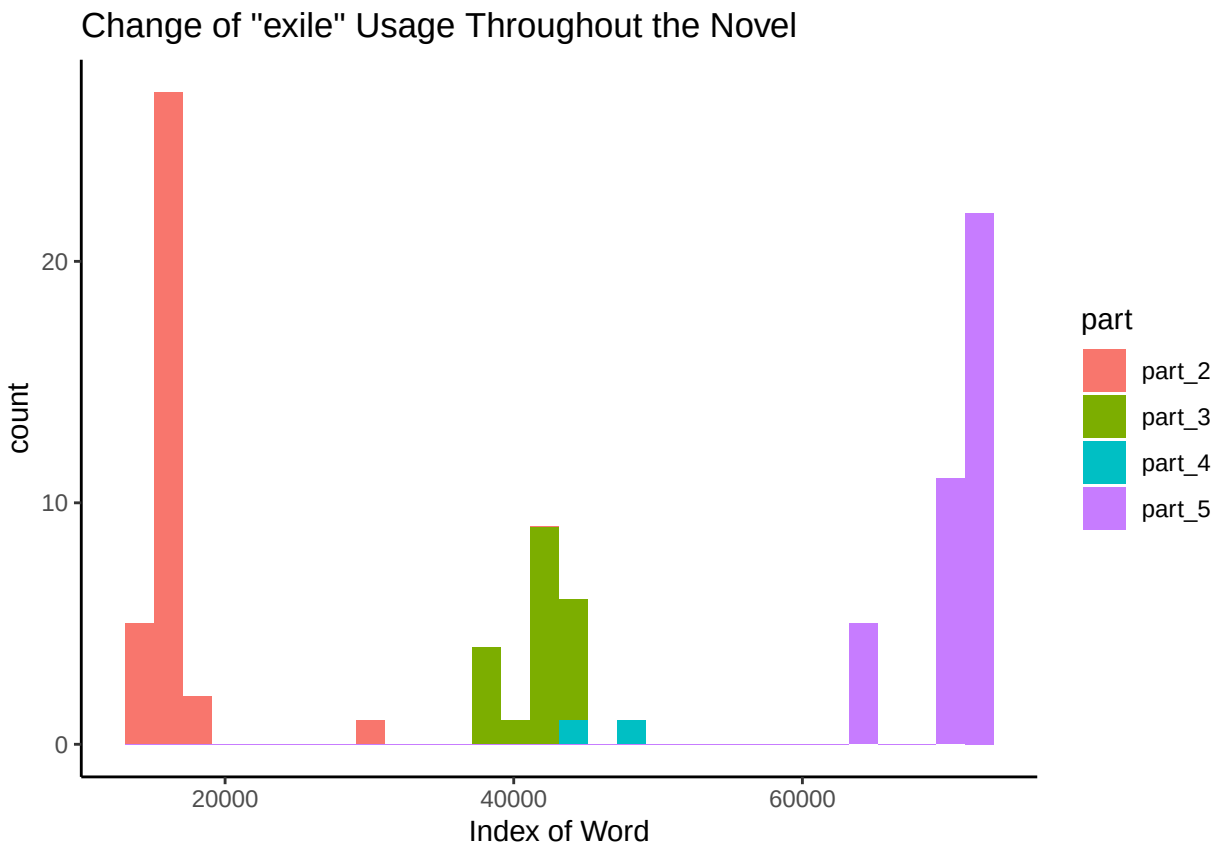

```
plague %>%filter(!is.na(selfish))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() + ggtitle(
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



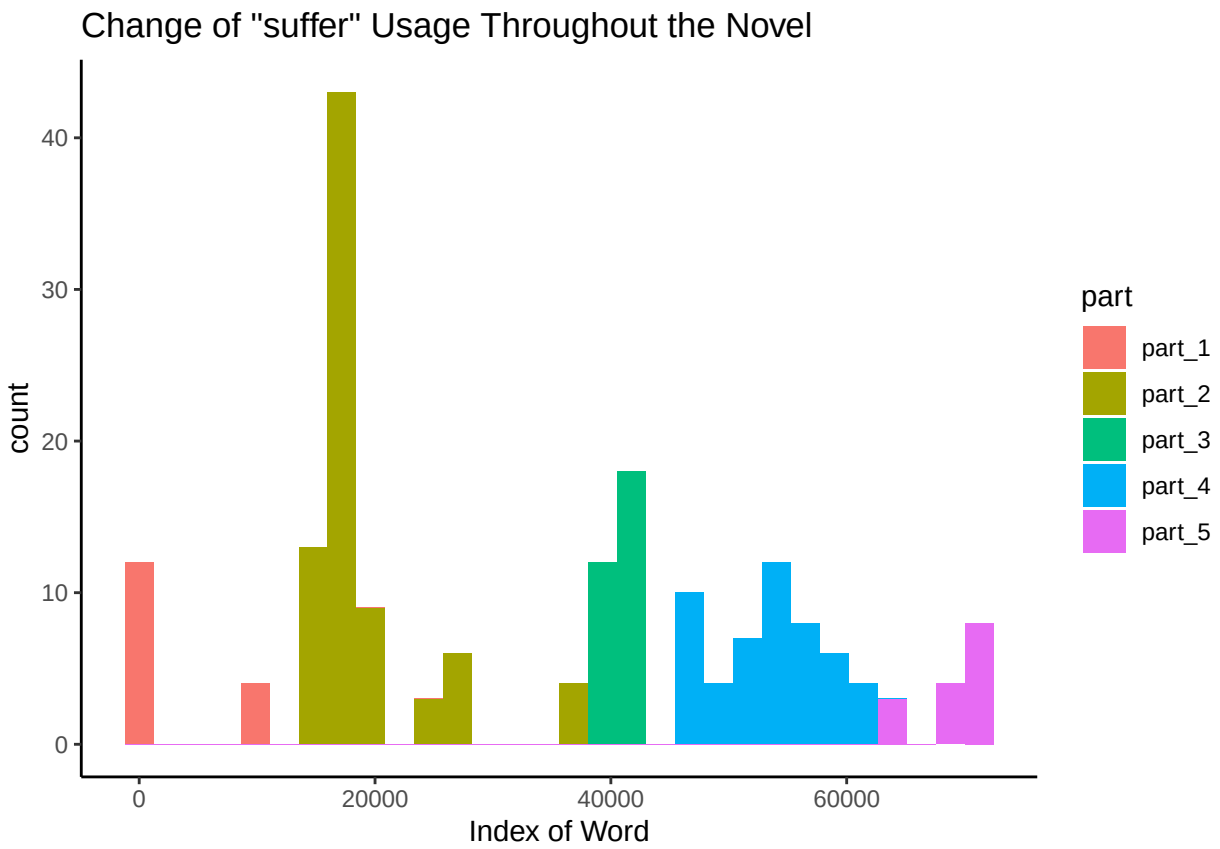

```
plague %>%filter(!is.na(exile))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() +  ggtitle("
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



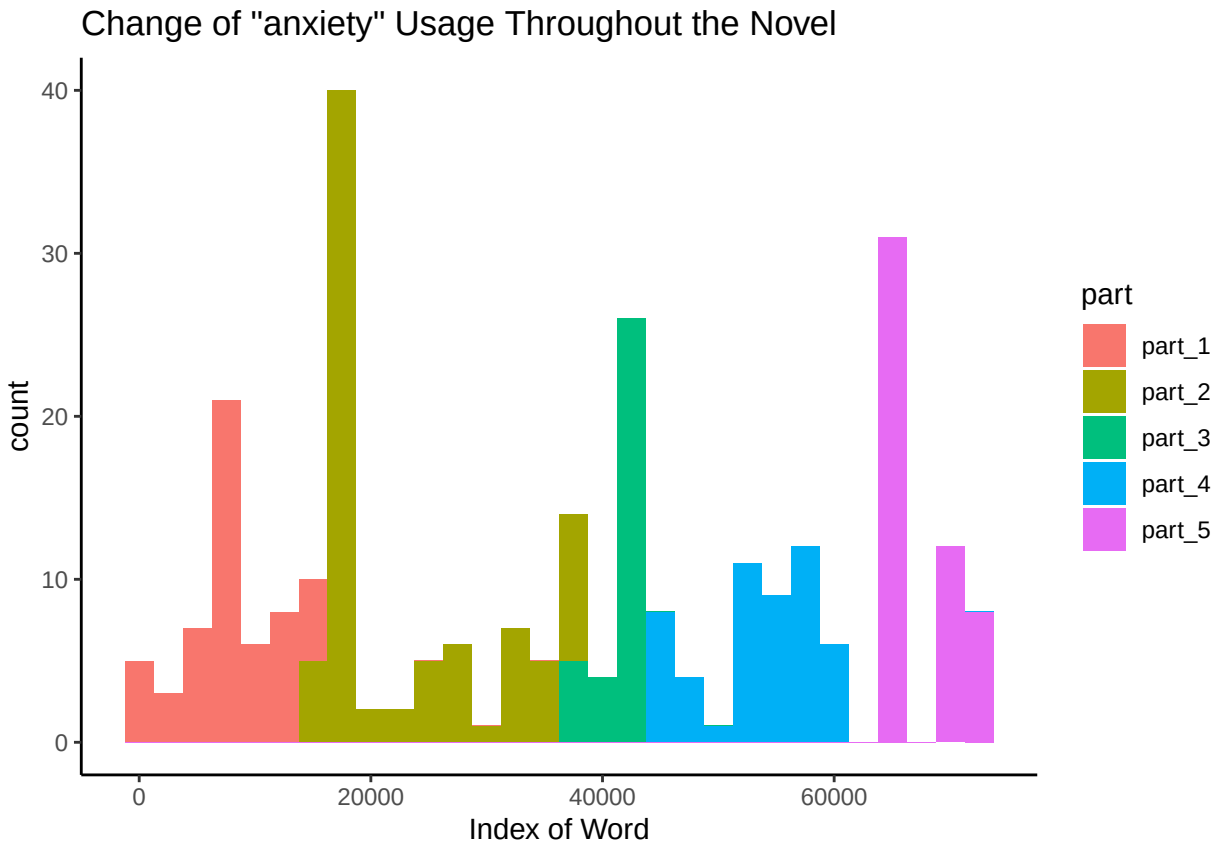

```
plague %>%filter(!is.na(suffer))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() +  ggtitle(
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



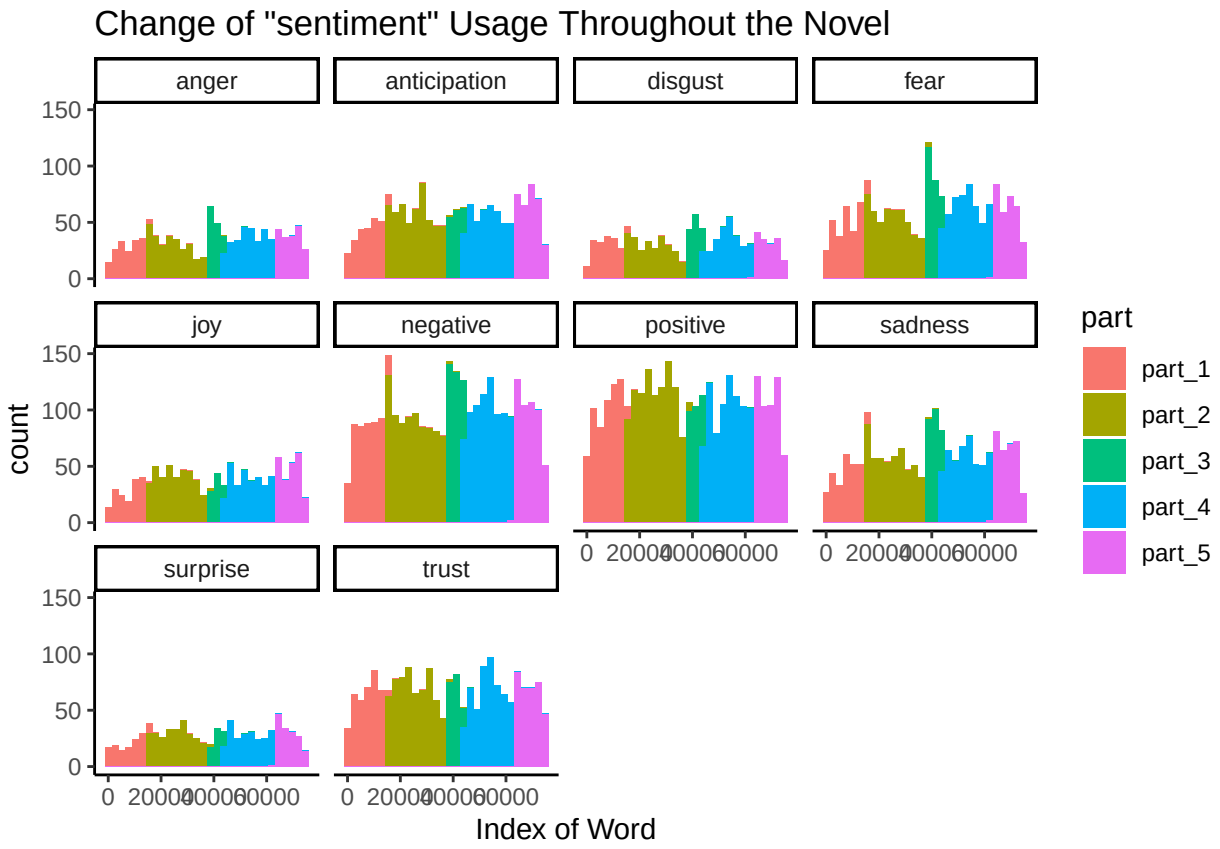

```
plague %>%filter(!is.na(anxiety))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() +  ggtitle
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



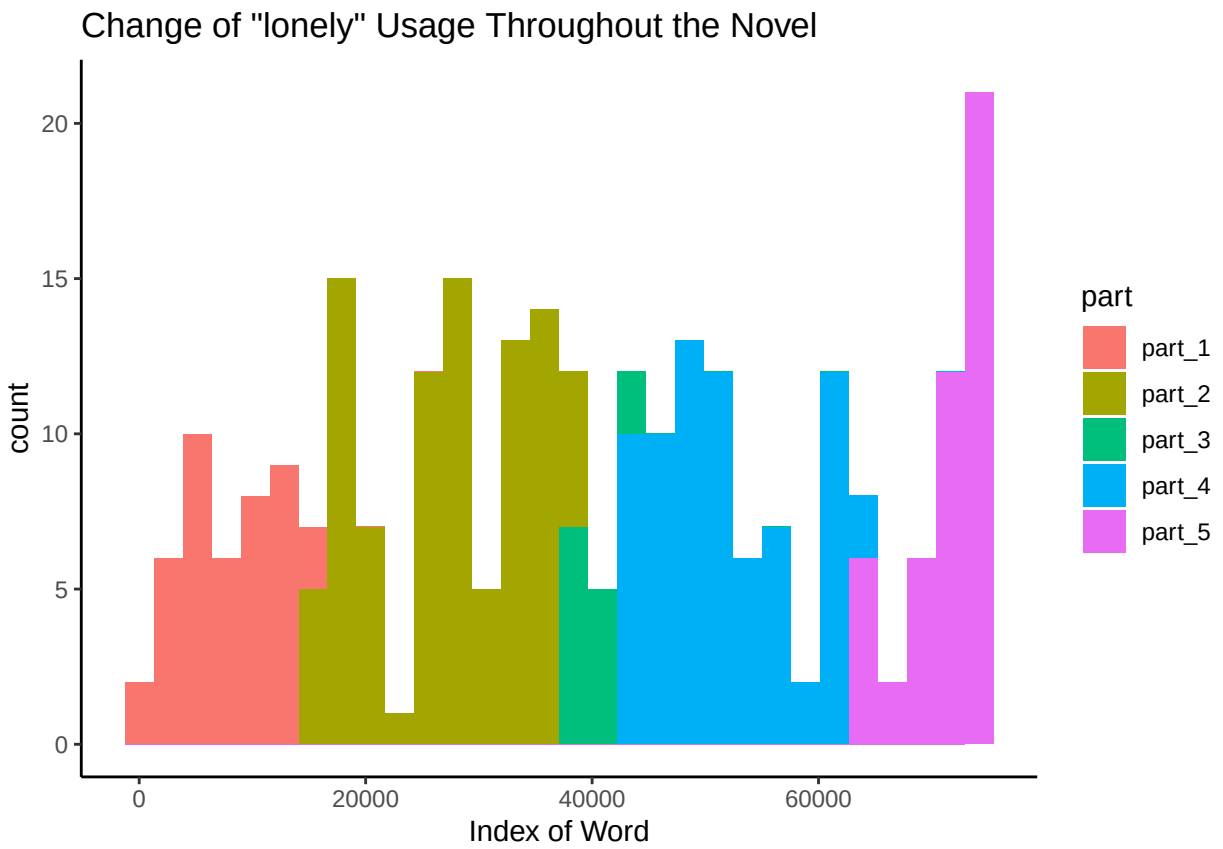

```
plague %>%filter(!is.na(nrc_sent))%>% ggplot(aes(x = count , fill = part)) + geom_histogram() + facet_wr
```

```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
plague %>%filter(!is.na(lonely))%>% ggplot(aes(x = count , fill = part)) + geom_histogram()  + ggtitle(
```

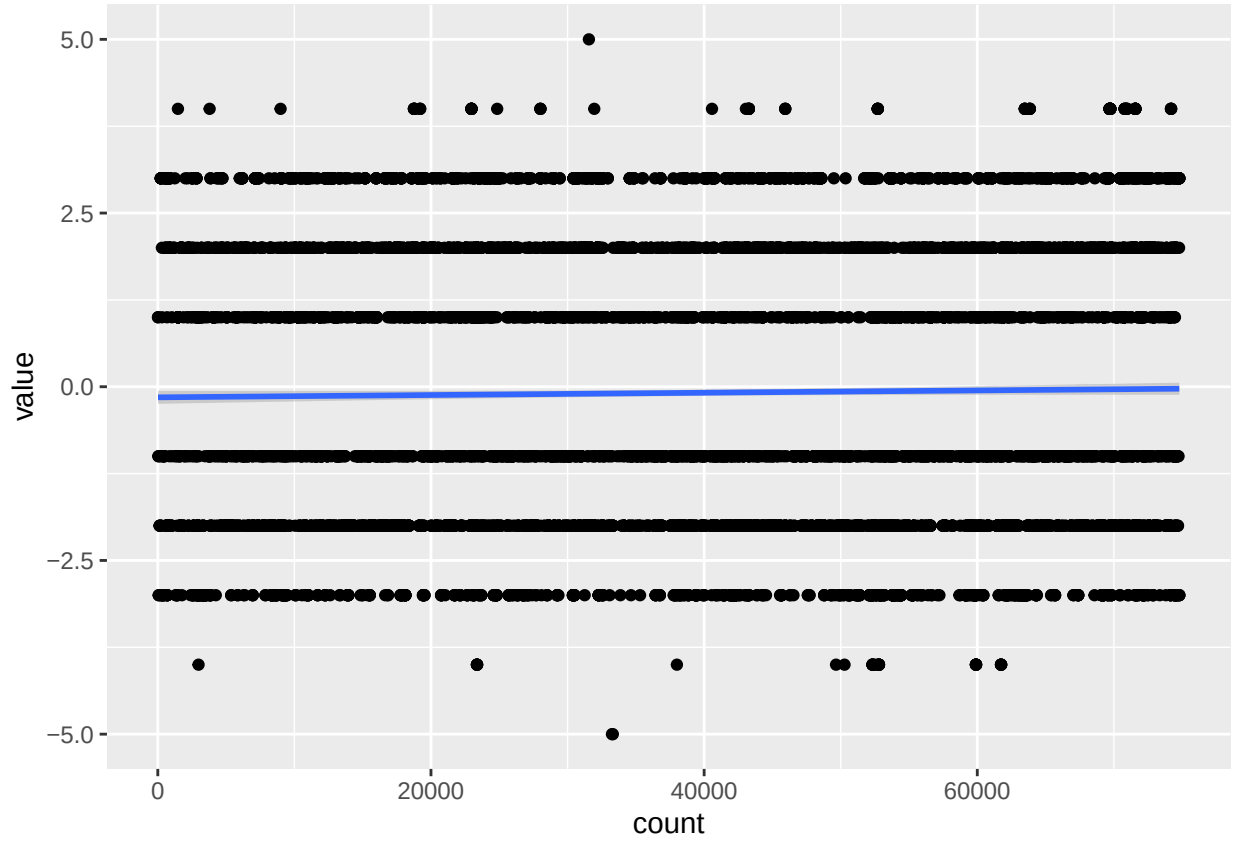
```
## Saving 6.5 x 4.5 in image
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
plague %>% filter(!is.na(value)) %>%ggplot(aes(x = count, y = value)) + geom_point() + geom_smooth(meth
```

```
## Saving 6.5 x 4.5 in image
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```




```
relation <-  lm(value ~count, data = plague)
summary(relation)
```

```
##
## Call:
## lm(formula = value ~ count, data = plague)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.9029 -1.9104 -0.8805  2.0670  5.1000
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.523e-01  4.778e-02  -3.187  0.00144 **
## count        1.657e-06  1.067e-06   1.553  0.12047
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.071 on 8086 degrees of freedom
##   (76871 observations deleted due to missingness)
## Multiple R-squared:  0.0002982,  Adjusted R-squared:  0.0001745
## F-statistic: 2.412 on 1 and 8086 DF,  p-value: 0.1205
```