Annex 1

Computing Worcester case data with R programming language

This annex is for replication purposes

```
worcester_case <- read.csv("~/Documents/RKC/Dissertation/Case study/worcester_case.csv")

install.packages("dplyr")
library("dplyr")

res <- count(worcester_case,date,media)

install.packages("xlsx")
library("xlsx")
```

Check the trend of published articles

```
res <- worcester_case[worcester_case$type.of.media=='Newspaper web page',]

res2 <- count(res,date,media)
```

Date field from files comes as text, try read ods file

```
install.packages("readODS")
library("readODS")
```

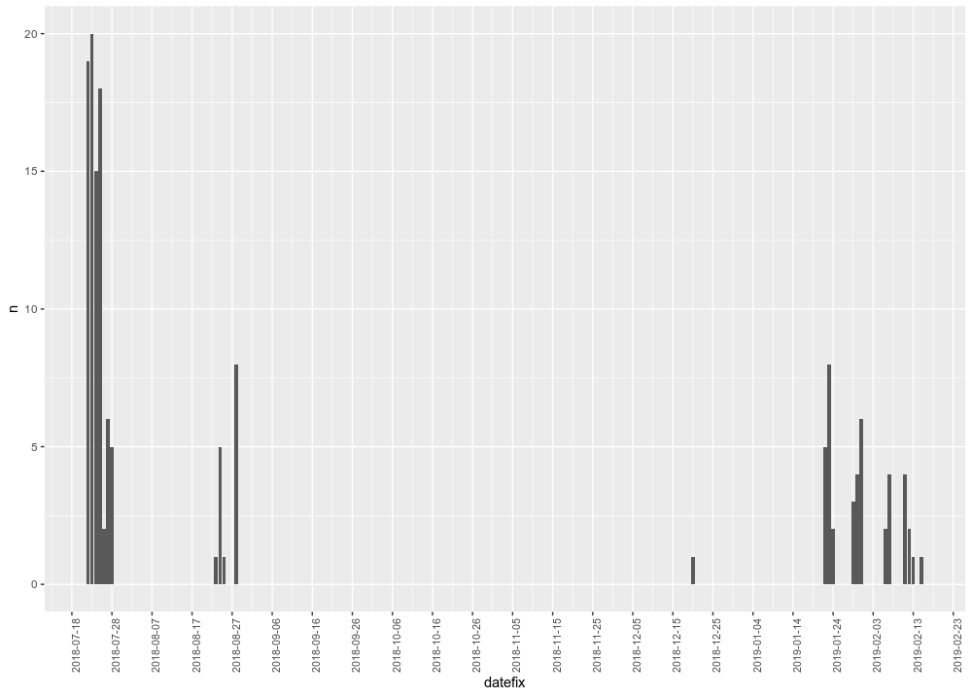
Warning, cell comments are loaded too, need to erase them from source

```
worcester_case <- read_ods(path = "/home/oreste/Documents/RKC/Dissertation/Case
study/accidents.ods", sheet = 2)
```

Convert dates in R environment, there is no way to convert them while loading the file.

```
worcester_case$datefix <- as.Date(worcester_case$date,format="%d/%m/%y")

res <- count(worcester_case[worcester_case$'type of media'=='Newspaper web
page',],datefix)

write_ods(res, "/home/oreste/Downloads/date_media_freq.ods")

p <- ggplot(data=res, aes(x=datefix, y=n)) + geom_bar(stat="identity") +
theme(axis.text.x = element_text(size=8, angle=90)) + scale_x_date(date_breaks = "10
days")
```

timeline articles



Keep date and datefix close

```
refcols <- c("date", "datefix")
worcester_case2 <- worcester_case[, c(refcols, setdiff(names(worcester_case), refcols))]
rm(worcester_case,res,p,refcols)
```

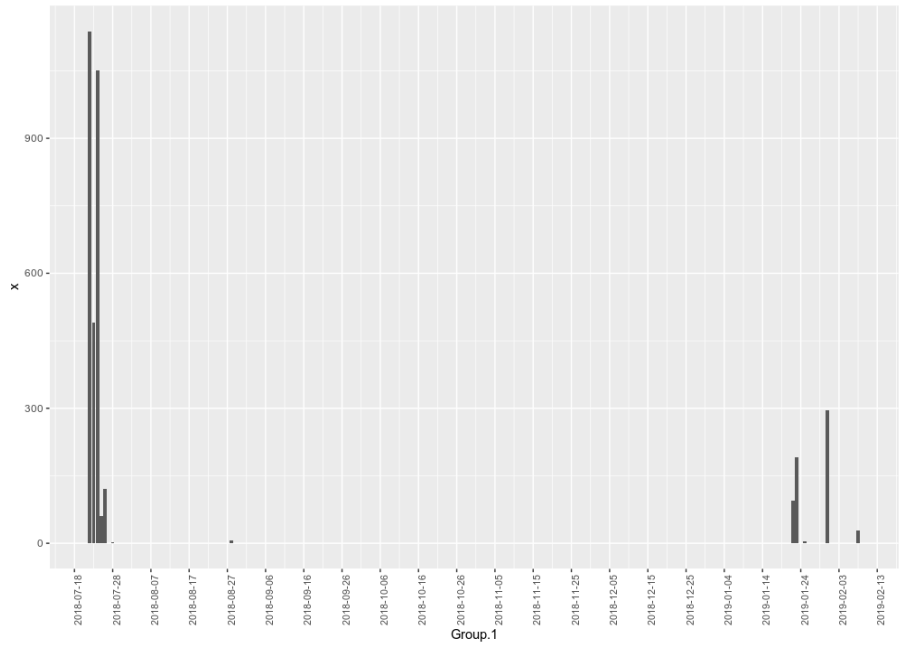
Check data type

```
> str(worcester_case2$'n. of comments')
 num [1:226] 174 NA NA 3 452 NA NA NA 17 36 ...
```

Remove NA

```
res2 <- worcester_case2[!is.na(worcester_case2$'n. of comments'),c('datefix','n. of
comments')]

res2 <- res2[res2$'n. of comments'>0,]

res <- aggregate(x=res2$'n. of comments', by=list(res2$datefix), FUN=sum)

write_ods(res, "/home/oreste/Downloads/date_comments_freq.ods")

p <- ggplot(data=res, aes(x=Group.1, y=x)) + geom_bar(stat="identity") +
theme(axis.text.x = element_text(size=8, angle=90)) + scale_x_date(date_breaks = "10
days")
```

timeline comments



Type of comments

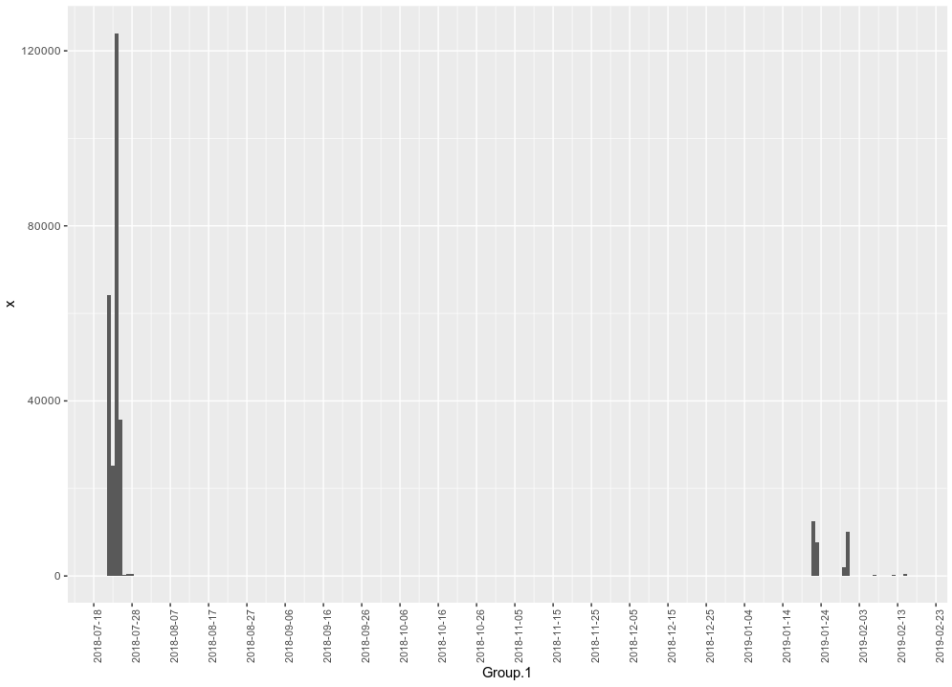
```
colsnum = c("n. of comments","political","religious","anti islamic","anti
foreigner","death penalty")

res3 <- worcester_case2[,colsnum]

colSums (res3, na.rm = TRUE)
```

```
n. of comments      political      religious   anti islamic anti foreigner  death penalty
          3484             47              4             17             75            181
```

```
n. of comments   political    religious     anti islamic   anti foreigner  death penalty
          3484   1.349024 %   0.1148106 %   0.4879449 %    2.152698 %      5.195178 %
```

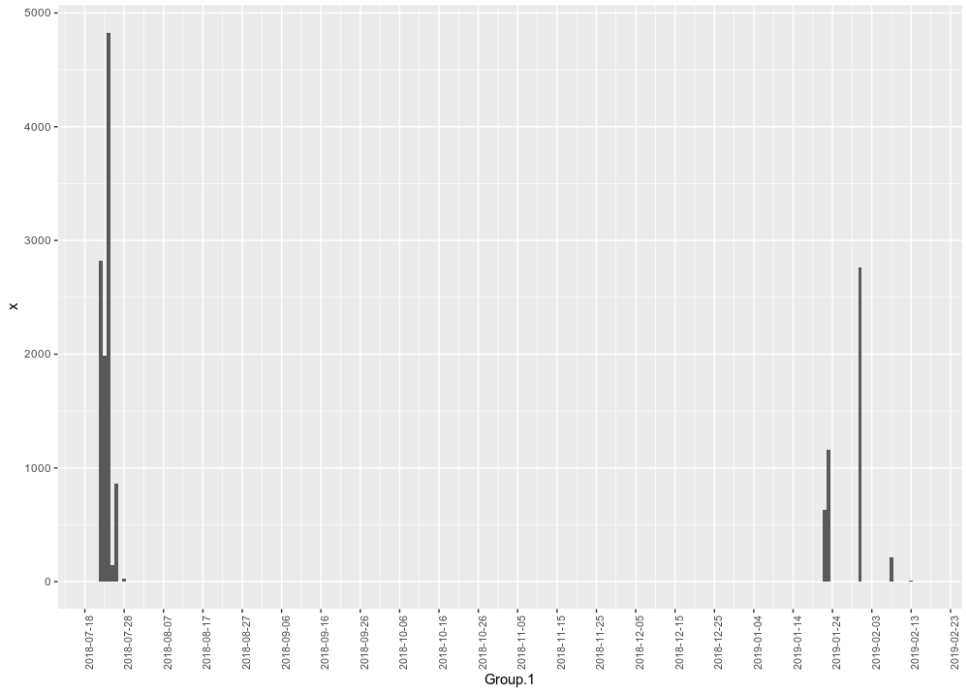
Share timeline

```
res <- worcester_case2[!is.na(worcester_case2$'n. of shares') & worcester_case2$'n. of
shares' > 0,c('datefix','n. of shares')]

res$'n. of shares' <- as.numeric(as.character(res$'n. of shares'))

res2 <- aggregate(x=res$'n. of shares', by=list(res$datefix), FUN=sum)

write_ods(res2, "/home/oreste/Downloads/date_shares_freq.ods")

p <- ggplot(data=res2, aes(x=Group.1, y=x)) + geom_bar(stat="identity") +
theme(axis.text.x = element_text(size=8, angle=90)) + scale_x_date(date_breaks = "10
days")
```



Emotional clicks timeline

```
res <- worcester_case[!is.na(worcester_case$'emotional clicks') &
worcester_case$'emotional clicks' > 0,c('datefix','emotional clicks')]

> str(worcester_case$'emotional clicks')
 num [1:226] 777 NA NA 4 NA NA NA NA 127 150 ...

res2 <- aggregate(x=res$'emotional clicks', by=list(res$datefix), FUN=sum)

p <- ggplot(data=res2, aes(x=Group.1, y=x)) + geom_bar(stat="identity") +
theme(axis.text.x = element_text(size=8, angle=90)) + scale_x_date(date_breaks = "10
days")
```



Comparision table

```
> names(worcester_case)
 [1] "date" "datefix" "id" "media" "type of media" "comments allowed"
 [7] "comments open" "n. of comments" "share avail" "n. of shares" "emotional clicks"
"link"
[13] "geo ref" "adv" "URL" "date of search" "text searched" "place searched"
[19] "political" "religious" "anti islamic" "anti foreigner" "death penalty" "anger"
[25] "notes"

> str(worcester_case$'n. of comments')
 num [1:226] 174 NA NA 3 452 NA NA NA 17 36 ...
> str(worcester_case$'n. of shares')
 chr [1:226] "522" "?" "?" "11" "9400" "20000" "6943" "54" "17" "76" "52" "?" "?" "?" "5"
"0" "88" "944" "25" "30" "572" "484" ...
> str(worcester_case$'emotional clicks')
 num [1:226] 777 NA NA 4 NA NA NA NA 127 150 ...

colclicks = c('datefix', 'n. of comments', 'n. of shares', 'emotional clicks')

res <- worcester_case[,colclicks]
```

Convert chr to num

```
> res$'n. of shares' <- as.numeric(as.character(res$'n. of shares'))
Warning message:
NAs introduced by coercion
```

Convert NA to 0 for later sum aggregation

```
res[c('n. of comments', 'n. of shares', 'emotional clicks')][is.na(res[c('n. of
comments', 'n. of shares', 'emotional clicks')])] <- 0
```

sum aggregation

```
res2 <- aggregate(list(comments=res$'n. of comments', shares=res$'n. of shares',
emotional=res$'emotional clicks'), by=list(res$datefix), FUN=sum)

write_ods(res2, "/home/oreste/Downloads/date_click_comparison.ods")
```

Anger analisys

```
colanger = c('datefix', 'n. of comments', 'type of media', 'geo ref', 'anger')

res <- worcester_case[,colanger]
```

Convert NA to 0

```
res['n. of comments'][is.na(res['n. of comments'])] <- 0
```

Delete rows with zero values and withouth anger value

```
res2 <- res[!(res$'n. of comments'==0) & ! is.na(res$'anger') ,]
```

Keep a copy

```
write_ods(res2, "/home/oreste/Downloads/anger.ods")
```

Anger intense aggregated by n. of comments

```
res2_intense_temp <- res2[res2$anger=='intense',]

res2_intense_temp$anger <- NULL

res2_intense <- aggregate(x=res2_intense_temp$'n. of comments',
by=list(res2_intense_temp$datefix), FUN=sum)

write_ods(res2_intense_temp, "/home/oreste/Downloads/res2_intense_temp.ods")

write_ods(res2_intense, "/home/oreste/Downloads/res2_intense.ods")
```