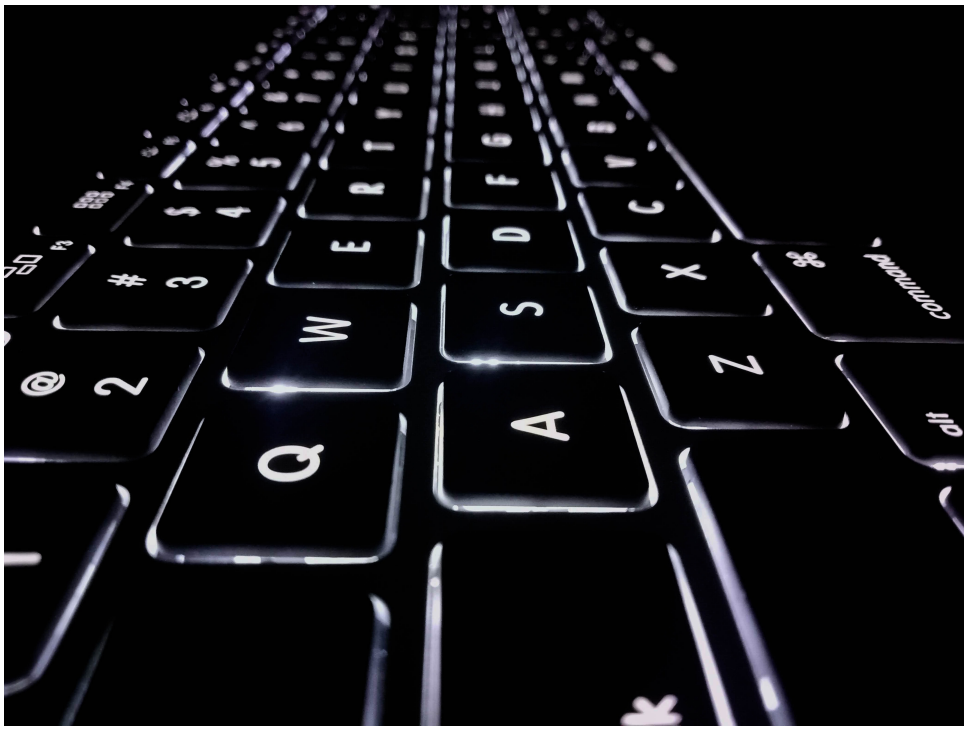

Assignment 4 ETC5513

Team name

XXXX

Contents

1 Introduction **3**

2 Data **4**
2.1 Data source 4
2.2 Data limitation 4

3 Methodology **5**

4 Section 1 **6**

5 Section 2 **7**

6 Section 3 **8**
6.1 The trend of occurence of the top five most frequent events changing by year 8

7 Section 4 **11**

8 Conclusion **12**

```r
knitr::opts_chunk$set(echo = TRUE, message=FALSE, warning= FALSE)
```

```r
library(tidyverse)
library(knitr)
library(readr)
library(lubridate)
library(bookdown)
library(kableExtra)
library(gridExtra)
theme_set(theme_minimal())
```

```r
# read data
storm17 <- read_csv("data/StormEvents_details-ftp_v1.0_d2017_c20200121.csv")
storm18 <- read_csv("data/StormEvents_details-ftp_v1.0_d2018_c20200317.csv")
storm19 <- read_csv("data/StormEvents_details-ftp_v1.0_d2019_c20200516.csv")

# change date format
clean <- function(x) {
x$BEGIN_DATE_TIME <- dmy_hms(x$BEGIN_DATE_TIME)
x$END_DATE_TIME <- dmy_hms(x$END_DATE_TIME)
x$BEGIN_DAY <- mday(x$BEGIN_DATE_TIME)
x$END_DAY <- mday(x$END_DATE_TIME)
x$BEGIN_TIME <- format(as.POSIXct(x$BEGIN_DATE_TIME) ,format = "%H:%M:%S")
x$END_TIME <- format(as.POSIXct(x$END_DATE_TIME) ,format = "%H:%M:%S")
x$MONTH_NAME <- month(x$BEGIN_DATE_TIME, label=TRUE)

#unselect column with NA value; unselect BEGIN MONTH and END MONTH because there are already MONTH colu
x <- x %>% select(-BEGIN_YEARMONTH, -END_YEARMONTH, -EPISODE_ID, -WFO, -CZ_TIMEZONE, -SOURCE, -MAGNITUDE
}

storm17 <- clean(storm17)
storm18 <- clean(storm18)
storm19 <- clean(storm19)
storm_all <- rbind(storm17,storm18,storm19)
```

1 Introduction

2 Data

2.1 Data source

We download the dataset from Storm Events Database data set which contains 31 variables and 187139 observations in total from NOAA https://www.ncdc.noaa.gov/stormevents/.

2.2 Data limitation

One of the defects of this dataset is that there is no further division of the scale of disasters, because the losses, casualties and even distribution areas caused by disasters of different scales are different, for example, it is quite difficult to distinguish between small versus large hail through this dataset. So, if there is a further division of the scale of disasters, it will be more conducive to analysis.

3 Methodology

4 Section 1

5 Section 2

6 Section 3

6.1 The trend of occurence of the top five most frequent events changing by year

In this section, the main purpose is to explore the trend of occurence of the top five most frequent events changing by year both across the whole USA and in each state. First, create a table and group the data by "STATE" and "YEAR". Then count the times of occurence for all events and rename the name of this column as "FREQ" to present the frequency. And next, create a new table to arrange the data by descending order with displaying the top 5 events that happen most frequently. Then, use "right_join" to connect these two tables. Finally, create two figures to display the trend of occurence of the top 5 event changing by year.

Table 1: The frequency of the events

STATE	YEAR	EVENT_TYPE	FREQ
ALABAMA	2017	Coastal Flood	4
ALABAMA	2017	Cold/Wind Chill	1
ALABAMA	2017	Drought	92
ALABAMA	2017	Extreme Cold/Wind Chill	1
ALABAMA	2017	Flash Flood	132
ALABAMA	2017	Flood	24
ALABAMA	2017	Frost/Freeze	11
ALABAMA	2017	Funnel Cloud	2
ALABAMA	2017	Hail	201
ALABAMA	2017	Heavy Rain	6

Table 2: Top five most frequent events

EVENT_TYPE
Thunderstorm Wind
Hail
Flood
Flash Flood
Winter Weather

Table 3: Occurence of top five most frequent events

STATE	YEAR	EVENT_TYPE	FREQ
ALABAMA	2017	Thunderstorm Wind	630
ALABAMA	2018	Thunderstorm Wind	495
ALABAMA	2019	Thunderstorm Wind	591
ALASKA	2019	Thunderstorm Wind	2
ARIZONA	2017	Thunderstorm Wind	116
ARIZONA	2018	Thunderstorm Wind	185
ARIZONA	2019	Thunderstorm Wind	97
ARKANSAS	2017	Thunderstorm Wind	423
ARKANSAS	2018	Thunderstorm Wind	411
ARKANSAS	2019	Thunderstorm Wind	359

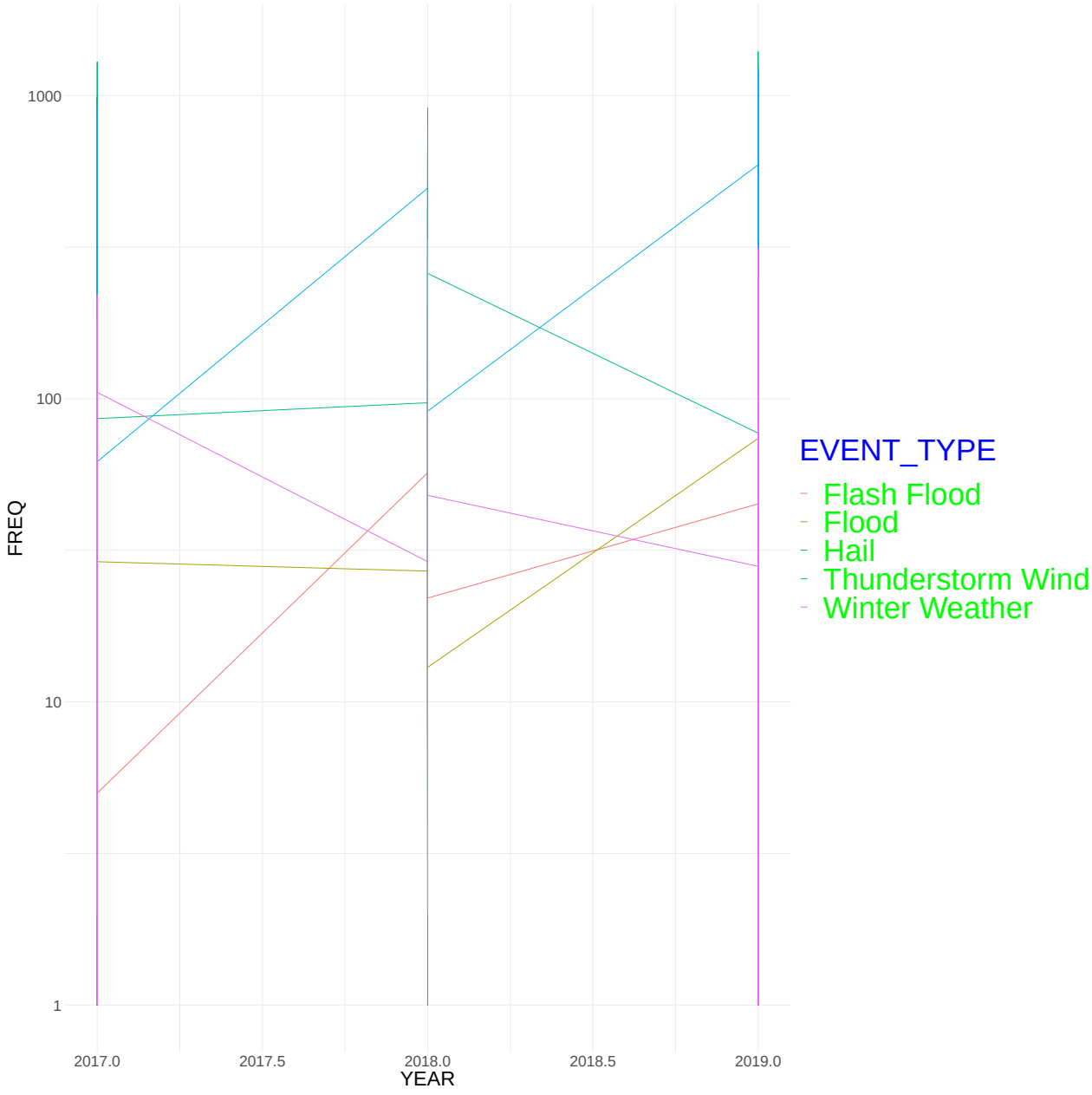

Figure 1: Occurence of top five most frequent events changing by year

Table 1 displays the first ten lines of the frequency of all events. And Table 2 displays the top five most frequent events. And Table 3 displays the first ten lines of the final results of the occurence of top five most frequent events. It can be seen that top 5 events have occurred in some regions in three years for example, thunderstorm wind has occurred in Alabama in three yearsfrom 495 times to 630 times , while it has only occurred twice in Alaska in 2019. Figure 1 presents the occurence of top five most frequent events across the whole USA, we could see that thunderstorm wind showed an increasing trend during three years. And hail showed a slow growth in 2017-2018, followed by a sharp decline in 2018-2019. As for the winter weather, it has been on a downward trend for three years, but the slowdown from 2017 to 2018 was larger than that from 2018 to 2019. What's more, flash flood kept growing, but slowed down between 2018 and 2019. At last, flood kept a slow decline at first, then turned into a significant rise. We could see that most events have a similar trend, but hail has decreased in 2018 to 2019 years, which is possibly due to increased melting level heights and greater atmospheric instability.(Dessens u. a., 2015) Correspondingly, the number of floods increased from 2018 to 2019, which is increasingly common due to years of relative sea level increases and El Nino.(floodingdaysin2018)

7 Section 4

8 Conclusion

References

[floodingdaysin2018] : *U.S. ties record for number of high tide flooding days in 2018.* – URL https://www.noaa.gov/media-release/us-ties-record-for-number-of-high-tide-flooding-days-in-2018

[Dessens u. a. 2015] Dessens, J. ; Berthet, C. ; Sanchez, J.l.: Change in hailstone size distributions with an increase in the melting level height. In: *Atmospheric Research* 158-159 (2015), S. 245–253