

RepData_PeerAssessment2.Rmd

Me

2020/9/24

##Synopsis

This project involves exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. This database tracks characteristics of major storms and weather events in the United States, including when and where they occur, as well as estimates of any fatalities, injuries, and property damage. What I want to do is to identify which type of events, indicated in the EVTYPE variable, causes most harmful effects with respect to population health, and causes the greatest economic consequences.

Data Processing

At first, I try to set the environment in order to analyze the dataset properly. the "repdata_data_StormData.csv.bz2" can be downloaded at the coursera course project page. and then you can directly read the file by "read.csv()". There are 902297 rows and 37 columns in the csv file.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##   combine
```

```
Storm<-read.csv("repdata_data_StormData.csv.bz2")
dim(Storm)
```

```
## [1] 902297      37
```

And then, make a subset properly. In Jan. 1996, all EVTYPE appear, so the subset should contain after Jan 1996. And, we should omit some unnecessary variables.

```
dat<-as.Date(Storm$BGN_DATE, format = "%m/%d/%Y %H:%M:%S")
Storm$BGN_DATE<-sort(dat)
rp<-subset(Storm,Storm$BGN_DATE>="1996-1-1",select=c("BGN_DATE","EVTYPE","MAG","FATALITIES","INJURIES",
```

now, there are 653530 rows and 9 columns in the “rp” dataset. Secondly, we should correct “CROPDMG-EXP” and “PROPDMGEXP” in order to make the “CROPDMG” and “PROPDMG” measurable and comparable.

```
unique(rp$PROPDMGEXP)
```

```
## [1] "K" "" "M" "B" "0"
```

```
unique(rp$CROPDMGEXP)
```

```
## [1] "K" "" "M" "B"
```

There are five symbols “K” “” “M” “B” “0”. According to (https://rstudio-pubs-static.s3.amazonaws.com/58957_37b6723ee52b455990e149edde45e5b6.html), the symbols are modified like below, “” to 0, “0” to 10, “K” to 1000, “M” to 10^6 , “B” to 10^9 .

```
rp$PROPDMGEXP<-gsub("0",10,rp$PROPDMGEXP)
rp$PROPDMGEXP<-gsub("B",1000000000,rp$PROPDMGEXP)
rp$PROPDMGEXP<-gsub("M",1000000,rp$PROPDMGEXP)
rp$PROPDMGEXP<-gsub("K",1000,rp$PROPDMGEXP)
rp$PROPDMGEXP<-gsub("^$",0,rp$PROPDMGEXP)
rp$CROPDMGEXP<-gsub("B",1000000000,rp$CROPDMGEXP)
rp$CROPDMGEXP<-gsub("M",1000000,rp$CROPDMGEXP)
rp$CROPDMGEXP<-gsub("K",1000,rp$CROPDMGEXP)
rp$CROPDMGEXP<-gsub("^$",0,rp$CROPDMGEXP)
unique(rp$PROPDMGEXP)
```

```
## [1] "1000" "0" "1e+06" "1e+09" "10"
```

```
unique(rp$CROPDMGEXP)
```

```
## [1] "1000" "0" "1e+06" "1e+09"
```

And they are multiplied with “xxxxDMG”. So, I made a new variable by “mutate” function which multiply “xxxxDMG” by them. The new variables are named as “CROP” and “PROP”.

```
rp$PROPDMGEXP<-as.integer(rp$PROPDMGEXP)
rp$CROPDMGEXP<-as.integer(rp$CROPDMGEXP)
rp<-mutate(rp, CROP = CROPDMG * CROPDMGEXP)
head(rp$CROP,n=100)
```

```
## [1] 38000 0 0 0 0 0 0 0 0 0 0 0 0
## [13] 0 0 0 0 0 0 0 0 0 0 0 0 0
## [25] 0 0 0 0 0 0 0 0 0 0 0 0 0
## [37] 0 0 0 0 0 0 0 0 0 0 0 0 0
## [49] 0 0 0 0 0 0 0 0 0 0 1000 2000 10000
## [61] 0 2000 0 25000 0 0 0 0 0 0 25000 50000
## [73] 0 0 0 0 0 0 0 12000 10000 0 0 0
## [85] 0 0 0 0 0 0 0 0 0 0 0 0
## [97] 0 0 0 0
```

```
rp<-mutate(rp, PROP = PROPDMG * PROPDMGEXP)
head(rp$PROP,n=100)
```

```
## [1] 380000 100000 3000 5000 2000 0 400000 12000 8000
## [10] 12000 0 75000 2000 0 5000 2000 2000 2000
## [19] 100000 2000 2000 2000 2000 2000 2000 15000 150000
## [28] 0 0 595000 12000 0 10000 0 0 0
## [37] 195000 15000 0 2000 0 2500 15000 10000 0
## [46] 5000 0 25000 25000 60000 500000 0 20000 0
## [55] 8000 10000 2000 2000 2000 15000 8000 5000 75000
## [64] 200000 0 25000 10000 15000 45000 0 95000 8000000
## [73] 15000 1250000 700000 40000 1200000 800000 1500000 150000 80000
## [82] 65000 10000 12000 15000 10000 45000 10000 0 10000
## [91] 0 10000 0 0 12000 5000 0 10000 30000
## [100] 10000
```

```
names(rp)
```

```
## [1] "BGN_DATE" "EVTYPE" "MAG" "FATALITIES" "INJURIES"
## [6] "PROPDMG" "PROPDMGEXP" "CROPDMG" "CROPDMGEXP" "CROP"
## [11] "PROP"
```

```
unique(rp$EVTYPE)
```

```
## [1] "WINTER STORM" "TORNADO"
## [3] "TSTM WIND" "HAIL"
## [5] "HIGH WIND" "HEAVY RAIN"
## [7] "FLASH FLOOD" "FREEZING RAIN"
## [9] "EXTREME COLD" "EXCESSIVE HEAT"
## [11] "LIGHTNING" "FUNNEL CLOUD"
## [13] "EXTREME WINDCHILL" "BLIZZARD"
## [15] "URBAN/SML STREAM FLD" "FLOOD"
## [17] "TSTM WIND/HAIL" "WATERSPOUT"
## [19] "RIP CURRENTS" "HEAVY SNOW"
## [21] "Other" "Record dry month"
## [23] "Temperature record" "WILD/FOREST FIRE"
```

## [25] "Minor Flooding"	"ICE STORM"
## [27] "STORM SURGE"	"Ice jam flood (minor)"
## [29] "High Wind"	"DUST STORM"
## [31] "STRONG WIND"	"DUST DEVIL"
## [33] "Tstm Wind"	"DROUGHT"
## [35] "DRY MICROBURST"	"FOG"
## [37] "ROUGH SURF"	"Wind"
## [39] "THUNDERSTORMS"	"Heavy Surf"
## [41] "HEAVY SURF"	"Dust Devil"
## [43] "Wind Damage"	"Marine Accident"
## [45] "Snow"	"AVALANCHE"
## [47] "Freeze"	"TROPICAL STORM"
## [49] "Snow Squalls"	"Coastal Flooding"
## [51] "Heavy Rain"	"Strong Wind"
## [53] "WINDS"	"WIND"
## [55] "COASTAL FLOOD"	"COASTAL STORM"
## [57] "COASTALFLOOD"	"Erosion/Cstl Flood"
## [59] "Heavy Rain and Wind"	"Light Snow/Flurries"
## [61] "Wet Month"	"Wet Year"
## [63] "Tidal Flooding"	"River Flooding"
## [65] "SNOW"	"DAMAGING FREEZE"
## [67] "Damaging Freeze"	"HURRICANE"
## [69] "Beach Erosion"	"Hot and Dry"
## [71] "Flood/Flash Flood"	"Icy Roads"
## [73] "High Surf"	"Heavy Rain/High Surf"
## [75] "HIGH SURF"	"Thunderstorm Wind"
## [77] "Rain Damage"	"ICE JAM"
## [79] "Unseasonable Cold"	"Early Frost"
## [81] "Wintry Mix"	"blowing snow"
## [83] "STREET FLOODING"	"Record Cold"
## [85] "Extreme Cold"	"Ice Fog"
## [87] "Excessive Cold"	"Torrential Rainfall"
## [89] "Freezing Rain"	"Landslump"
## [91] "Late-season Snowfall"	"Hurricane Edouard"
## [93] "Coastal Storm"	"Flood"
## [95] "HEAVY RAIN/WIND"	"TIDAL FLOODING"
## [97] "Winter Weather"	"Snow squalls"
## [99] "Strong Winds"	"Strong winds"
## [101] "RECORD WARM TEMPS."	"Ice/Snow"
## [103] "Mudslide"	"Glaze"
## [105] "Extended Cold"	"Snow Accumulation"
## [107] "Freezing Fog"	"Drifting Snow"
## [109] "Whirlwind"	"Heavy snow shower"
## [111] "Heavy rain"	"COASTAL FLOODING"
## [113] "LATE SNOW"	"Record May Snow"
## [115] "Record Winter Snow"	"Heavy Precipitation"
## [117] " COASTAL FLOOD"	"Record temperature"
## [119] "Light snow"	"Late Season Snowfall"
## [121] "Gusty Wind"	"small hail"
## [123] "Light Snow"	"MIXED PRECIP"
## [125] "Black Ice"	"Mudslides"
## [127] "Gradient wind"	"Snow and Ice"
## [129] "COLD"	"Freezing Spray"
## [131] "DOWNBURST"	"Summary Jan 17"

## [133]	"Summary of March 14"	"Summary of March 23"
## [135]	"Summary of March 24"	"Summary of April 3rd"
## [137]	"Summary of April 12"	"Summary of April 13"
## [139]	"Summary of April 21"	"Summary August 11"
## [141]	"Summary of April 27"	"Summary of May 9-10"
## [143]	"Summary of May 10"	"Summary of May 13"
## [145]	"Summary of May 14"	"Summary of May 22 am"
## [147]	"Summary of May 22 pm"	"Heatburst"
## [149]	"Summary of May 26 am"	"Summary of May 26 pm"
## [151]	"Metro Storm, May 26"	"Summary of May 31 am"
## [153]	"Summary of May 31 pm"	"Summary of June 3"
## [155]	"Summary of June 4"	"Summary June 5-6"
## [157]	"Summary June 6"	"Summary of June 11"
## [159]	"Summary of June 12"	"Summary of June 13"
## [161]	"Summary of June 15"	"Summary of June 16"
## [163]	"Summary June 18-19"	"Summary of June 23"
## [165]	"Summary of June 24"	"Summary of June 30"
## [167]	"Summary of July 2"	"Summary of July 3"
## [169]	"Summary of July 11"	"Summary of July 22"
## [171]	"Summary July 23-24"	"Summary of July 26"
## [173]	"Summary of July 29"	"Summary of August 1"
## [175]	"Summary August 2-3"	"Summary August 7"
## [177]	"Summary August 9"	"Summary August 10"
## [179]	"Summary August 17"	"Summary August 21"
## [181]	"Summary August 28"	"Summary September 4"
## [183]	"Summary September 20"	"Summary September 23"
## [185]	"Summary Sept. 25-26"	"Summary: Oct. 20-21"
## [187]	"Summary: October 31"	"Summary: Nov. 6-7"
## [189]	"Summary: Nov. 16"	"Microburst"
## [191]	"wet micoburst"	"HAIL/WIND"
## [193]	"Hail(0.75)"	"Funnel Cloud"
## [195]	"Urban Flooding"	"No Severe Weather"
## [197]	"Urban flood"	"Urban Flood"
## [199]	"Cold"	"WINTER WEATHER"
## [201]	"Summary of May 22"	"Summary of June 6"
## [203]	"Summary August 4"	"Summary of June 10"
## [205]	"Summary of June 18"	"Summary September 3"
## [207]	"Summary: Sept. 18"	"Coastal Flood"
## [209]	"coastal flooding"	"Small Hail"
## [211]	"Record Temperatures"	"Light Snowfall"
## [213]	"Freezing Drizzle"	"Gusty wind/rain"
## [215]	"GUSTY WIND/HVY RAIN"	"Blowing Snow"
## [217]	"Early snowfall"	"Monthly Snowfall"
## [219]	"Record Heat"	"Seasonal Snowfall"
## [221]	"Monthly Rainfall"	"Cold Temperature"
## [223]	"Sml Stream Fld"	"Heat Wave"
## [225]	"MUDSLIDE/LANDSLIDE"	"Saharan Dust"
## [227]	"Volcanic Ash"	"Volcanic Ash Plume"
## [229]	"Thundersnow shower"	"NONE"
## [231]	"COLD AND SNOW"	"DAM BREAK"
## [233]	"RAIN"	"RAIN/SNOW"
## [235]	"OTHER"	"FREEZE"
## [237]	"TSTM WIND (G45)"	"RECORD WARMTH"
## [239]	"STRONG WINDS"	"FREEZING DRIZZLE"

## [241] "UNSEASONABLY WARM"	"SLEET/FREEZING RAIN"
## [243] "BLACK ICE"	"WINTRY MIX"
## [245] "BLOW-OUT TIDES"	"UNSEASONABLY COLD"
## [247] "UNSEASONABLY COOL"	"TSTM HEAVY RAIN"
## [249] "UNSEASONABLY DRY"	"Gusty Winds"
## [251] "GUSTY WIND"	"TSTM WIND 40"
## [253] "TSTM WIND 45"	"HARD FREEZE"
## [255] "TSTM WIND (41)"	"HEAT"
## [257] "RIVER FLOOD"	"TSTM WIND (G40)"
## [259] "RIP CURRENT"	"TSTM WND"
## [261] "DENSE FOG"	"Wintry mix"
## [263] " TSTM WIND"	"MUD SLIDE"
## [265] "MUDSLIDES"	"MUDSLIDE"
## [267] "Frost"	"Frost/Freeze"
## [269] "SNOW AND ICE"	"WIND DAMAGE"
## [271] "RAIN (HEAVY)"	"Record Warmth"
## [273] "Prolong Cold"	"Cold and Frost"
## [275] "RECORD COLD"	"PROLONG COLD"
## [277] "AGRICULTURAL FREEZE"	"URBAN/SML STREAM FLDG"
## [279] "SNOW SQUALL"	"HEAVY SNOW SQUALLS"
## [281] "SNOW/ICE"	"GUSTY WINDS"
## [283] "SMALL HAIL"	"SNOW SQUALLS"
## [285] "LAKE EFFECT SNOW"	"STRONG WIND GUST"
## [287] "LATE FREEZE"	"RECORD TEMPERATURES"
## [289] "ICY ROADS"	"RECORD SNOWFALL"
## [291] "BLOW-OUT TIDE"	"THUNDERSTORM"
## [293] "Hypothermia/Exposure"	"HYPOTHERMIA/EXPOSURE"
## [295] "Lake Effect Snow"	"Mixed Precipitation"
## [297] "Record High"	"COASTALSTORM"
## [299] "LIGHT SNOW"	"Snow and sleet"
## [301] "Freezing rain"	"Gusty winds"
## [303] "FUNNEL CLOUDS"	"WATERSPOUTS"
## [305] "Blizzard Summary"	"FROST"
## [307] "ICE"	"SUMMARY OF MARCH 24-25"
## [309] "SUMMARY OF MARCH 27"	"SUMMARY OF MARCH 29"
## [311] "GRADIENT WIND"	"Icestorm/Blizzard"
## [313] "Flood/Strong Wind"	"TSTM WIND AND LIGHTNING"
## [315] "gradient wind"	"SEVERE THUNDERSTORMS"
## [317] "EXCESSIVE RAIN"	"Freezing drizzle"
## [319] "Mountain Snows"	"URBAN/SMALL STRM FLDG"
## [321] "WET MICROBURST"	"Heavy surf and wind"
## [323] "Mild and Dry Pattern"	"COLD AND FROST"
## [325] "RECORD HEAT"	"TYPHOON"
## [327] "LANDSLIDES"	"HIGH SWELLS"
## [329] "HIGH SWELLS"	"VOLCANIC ASH"
## [331] "HIGH WINDS"	"DRY SPELL"
## [333] " LIGHTNING"	"BEACH EROSION"
## [335] "UNSEASONAL RAIN"	"EARLY RAIN"
## [337] "PROLONGED RAIN"	"WINTER MIX"
## [339] "COASTAL FLOODING/EROSION"	"UNSEASONABLY WET"
## [341] "HOT SPELL"	"HEAT WAVE"
## [343] "UNSEASONABLY HOT"	"UNSEASONABLY WARM AND DRY"
## [345] " TSTM WIND (G45)"	"TSTM WIND (G45)"
## [347] "HIGH WIND (G40)"	"TSTM WIND (G35)"

## [349] "DRY WEATHER"	"TSTM WINDS"
## [351] "FREEZING RAIN/SLEET"	"ABNORMAL WARMTH"
## [353] "UNUSUAL WARMTH"	"GLAZE"
## [355] "WAKE LOW WIND"	"MONTHLY RAINFALL"
## [357] "COLD TEMPERATURES"	"COLD WIND CHILL TEMPERATURES"
## [359] "MODERATE SNOW"	"MODERATE SNOWFALL"
## [361] "URBAN/STREET FLOODING"	"COASTAL EROSION"
## [363] "UNUSUAL/RECORD WARMTH"	"BITTER WIND CHILL"
## [365] "BITTER WIND CHILL TEMPERATURES"	"SEICHE"
## [367] "TSTM"	"COASTAL FLOODING/EROSION"
## [369] "SNOW DROUGHT"	"UNSEASONABLY WARM YEAR"
## [371] "HYPERTHERMIA/EXPOSURE"	"SNOW/SLEET"
## [373] "ROCK SLIDE"	"ICE PELLETS"
## [375] "URBAN FLOOD"	"PATCHY DENSE FOG"
## [377] "RECORD COOL"	"RECORD WARM"
## [379] "HOT WEATHER"	"RIVER FLOODING"
## [381] "RECORD TEMPERATURE"	"SAHARAN DUST"
## [383] "TROPICAL DEPRESSION"	"VOLCANIC ERUPTION"
## [385] "COOL SPELL"	"WIND ADVISORY"
## [387] "GUSTY WIND/HAIL"	"RED FLAG FIRE WX"
## [389] "FIRST FROST"	"EXCESSIVELY DRY"
## [391] "HEAVY SEAS"	"FLASH FLOOD/FLOOD"
## [393] "SNOW AND SLEET"	"LIGHT SNOW/FREEZING PRECIP"
## [395] "VOG"	"EXCESSIVE RAINFALL"
## [397] "FLASH FLOODING"	"MONTHLY PRECIPITATION"
## [399] "MONTHLY TEMPERATURE"	"RECORD DRYNESS"
## [401] "EXTREME WINDCHILL TEMPERATURES"	"MIXED PRECIPITATION"
## [403] "EXTREME WIND CHILL"	"DRY CONDITIONS"
## [405] "HEAVY RAINFALL"	"REMNANTS OF FLOYD"
## [407] "EARLY SNOWFALL"	"FREEZING FOG"
## [409] "LANDSPOUT"	"DRIEST MONTH"
## [411] "RECORD COLD"	"LATE SEASON HAIL"
## [413] "EXCESSIVE SNOW"	"WINTER MIX"
## [415] "DRYNESS"	"FLOOD/FLASH/FLOOD"
## [417] "WIND AND WAVE"	"SEVERE THUNDERSTORM"
## [419] "LIGHT FREEZING RAIN"	" WIND"
## [421] "MONTHLY SNOWFALL"	"DRY"
## [423] "RECORD RAINFALL"	"RECORD PRECIPITATION"
## [425] "ICE ROADS"	"HIGH SEAS"
## [427] "SLEET"	"ROUGH SEAS"
## [429] "UNSEASONABLY WARM/WET"	"UNSEASONABLY COOL & WET"
## [431] "UNUSUALLY WARM"	"TSTM WIND G45"
## [433] "NON SEVERE HAIL"	"RECORD SNOW"
## [435] "SNOW/FREEZING RAIN"	"SNOW/BLOWING SNOW"
## [437] "NON-SEVERE WIND DAMAGE"	"UNUSUALLY COLD"
## [439] "WARM WEATHER"	"LANDSLUMP"
## [441] "THUNDERSTORM WIND (G40)"	"LANDSLIDE"
## [443] "WALL CLOUD"	"HIGH WATER"
## [445] "UNSEASONABLY WARM & WET"	" FLASH FLOOD"
## [447] "LOCALLY HEAVY RAIN"	"WIND GUSTS"
## [449] "UNSEASONAL LOW TEMP"	"HIGH SURF ADVISORY"
## [451] "LATE SEASON SNOW"	"GUSTY LAKE WIND"
## [453] "ABNORMALLY DRY"	"WINTER WEATHER MIX"
## [455] "RED FLAG CRITERIA"	"WND"

```
## [457] "CSTL FLOODING/EROSION"      "SMOKE"
## [459] " WATERSPOUT"                  "SNOW ADVISORY"
## [461] "EXTREMELY WET"                 "UNUSUALLY LATE SNOW"
## [463] "VERY DRY"                      "RECORD LOW RAINFALL"
## [465] "ROGUE WAVE"                    "SNOWMELT FLOODING"
## [467] "PROLONG WARMTH"                "ACCUMULATED SNOWFALL"
## [469] "FALLING SNOW/ICE"              "DUST DEVEL"
## [471] "NON-TSTM WIND"                 "NON TSTM WIND"
## [473] "BRUSH FIRE"                    "GUSTY THUNDERSTORM WINDS"
## [475] "PATCHY ICE"                   "SNOW SHOWERS"
## [477] "HEAVY RAIN EFFECTS"            "BLOWING DUST"
## [479] "EXCESSIVE HEAT/DROUGHT"        "NORTHERN LIGHTS"
## [481] "MARINE TSTM WIND"              "  HIGH SURF ADVISORY"
## [483] "WIND CHILL"                    "HAZARDOUS SURF"
## [485] "WILDFIRE"                      "FROST/FREEZE"
## [487] "WINTER WEATHER/MIX"            "ASTRONOMICAL HIGH TIDE"
## [489] "COLD WEATHER"                  "WHIRLWIND"
## [491] "VERY WARM"                     "ABNORMALLY WET"
## [493] "TORNADO DEBRIS"                "EXTREME COLD/WIND CHILL"
## [495] "ICE ON ROAD"                   "FIRST SNOW"
## [497] "ICE/SNOW"                       "DROWNING"
## [499] "GUSTY THUNDERSTORM WIND"        "MARINE HAIL"
## [501] "HIGH SURF ADVISORIES"          "HURRICANE/TYPHOON"
## [503] "HEAVY SURF/HIGH SURF"          "SLEET STORM"
## [505] "STORM SURGE/TIDE"              "COLD/WIND CHILL"
## [507] "LAKE-EFFECT SNOW"              "MARINE HIGH WIND"
## [509] "THUNDERSTORM WIND"             "TSUNAMI"
## [511] "DENSE SMOKE"                   "LAKESHORE FLOOD"
## [513] "MARINE THUNDERSTORM WIND"       "MARINE STRONG WIND"
## [515] "ASTRONOMICAL LOW TIDE"         "VOLCANIC ASHFALL"
```

Finally, EVTYPE has 516 variables but, they are integrated into 48 proper variables and the others. The 48 variables are listed at page 6 of https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2_doc%2Fpd01016005curr.pdf

```
grep("^[ ]",unique(rp$EVTYPE),value = T)
```

```
## [1] " COASTAL FLOOD"      " TSTM WIND"          " LIGHTNING"
## [4] " TSTM WIND (G45)"    " WIND"               " FLASH FLOOD"
## [7] " WATERSPOUT"        "  HIGH SURF ADVISORY"
```

```
s2<-gsub("^[ ]+", "", rp$EVTYPE)
s1<-unique(s2)
grep("[a]s|[a]v",s1,ignore.case = T, value = T)
```

```
## [1] "AVALANCHE"           "ASTRONOMICAL HIGH TIDE" "ASTRONOMICAL LOW TIDE"
```

```
grep("[b]li",s1,ignore.case = T,value = T)
```

```
## [1] "BLIZZARD"           "Blizzard Summary"
```



```
s2<-gsub("Blizzard Summary","BLIZZARD",s2)
grep("[c]oa",s1,ignore.case = T,value = T)
```

```
## [1] "Coastal Flooding"      "COASTAL FLOOD"
## [3] "COASTAL STORM"         "COASTALFLOOD"
## [5] "Coastal Storm"         "COASTAL FLOODING"
## [7] "Coastal Flood"         "coastal flooding"
## [9] "COASTALSTORM"          "COASTAL FLOODING/EROSION"
## [11] "COASTAL EROSION"       "COASTAL FLOODING/EROSION"
```

```
s2<-gsub("[c]oastal [f]+","COASTAL FLOOD",s2,ignore.case = T)
grep("c.*s.*t.*l",s1,ignore.case = T,value = T)
```

```
## [1] "Coastal Flooding"      "COASTAL FLOOD"
## [3] "COASTAL STORM"         "COASTALFLOOD"
## [5] "Erosion/Cstl Flood"    "Coastal Storm"
## [7] "COASTAL FLOODING"      "Coastal Flood"
## [9] "coastal flooding"      "COASTALSTORM"
## [11] "Icestorm/Blizzard"     "COASTAL FLOODING/EROSION"
## [13] "COASTAL EROSION"       "COASTAL FLOODING/EROSION"
## [15] "CSTL FLOODING/EROSION"
```

```
s2<-gsub(".*s.*t.*l.*f.*","Coastal Flood",s2,ignore.case = T)
s2<-gsub("floo[d].+$","FLOOD",s2,ignore.case = T)
grep("cold|wind chill",s1,ignore.case = T,value = T)
```

```
## [1] "EXTREME COLD"          "Unseasonable Cold"
## [3] "Record Cold"           "Extreme Cold"
## [5] "Excessive Cold"        "Extended Cold"
## [7] "COLD"                  "Cold"
## [9] "Cold Temperature"      "COLD AND SNOW"
## [11] "UNSEASONABLY COLD"     "Prolong Cold"
## [13] "Cold and Frost"        "RECORD COLD"
## [15] "PROLONG COLD"          "COLD AND FROST"
## [17] "COLD TEMPERATURES"     "COLD WIND CHILL TEMPERATURES"
## [19] "BITTER WIND CHILL"      "BITTER WIND CHILL TEMPERATURES"
## [21] "EXTREME WIND CHILL"     "RECORD COLD"
## [23] "UNUSUALLY COLD"        "WIND CHILL"
## [25] "COLD WEATHER"          "EXTREME COLD/WIND CHILL"
## [27] "COLD/WIND CHILL"
```

There are a lot of cold / wind chill types. So, I classify “BITTER WIND CHILL”, “BITTER WIND CHILL TEMPERATURES”, “Cold”, “COLD”, “Cold and Frost”, “COLD AND FROST”, “COLD AND SNOW”, “Cold Temperature”, “COLD TEMPERATURES”, “COLD WEATHER”, “COLD WIND CHILL TEMPERATURES”, “COLD/WIND CHILL”, “WIND CHILL” into “Cold/Wind Chill” and “Excessive Cold”, “Extended Cold”, “Extreme Cold”, “EXTREME COLD”, “EXTREME COLD/WIND CHILL”, “EXTREME WIND CHILL”, “RECORD COLD” are classified into “Extreme Cold/Wind Chill”. Then, if a word integrated by “/” which consists both two words listed in the 48 variables, basically left word(left of /) is adopted.

```
s2<-gsub("^[c]old.*$|^[w]ind chill.*$|^[b]itter.+$","Cold/Wind Chill",s2,ignore.case = T)
grep("deb",s1,ignore.case = T, value = T)
```

```
## [1] "TORNADO DEBRIS"
```

There is no “Debris Flow” in the rp dataset, so I dismiss the variable.

```
grep("den",s1,ignore.case = T,value = T)
```

```
## [1] "Marine Accident" "DENSE FOG" "PATCHY DENSE FOG" "DENSE SMOKE"
```

```
s2<-gsub(".*DENSE FOG","Dense Fog",s2,ignore.case = T)
grep("dro",s1,ignore.case = T, value = T)
```

```
## [1] "DROUGHT" "SNOW DROUGHT" "EXCESSIVE HEAT/DROUGHT"
## [4] "DROWNING"
```

```
s2<-gsub("drought|SNOW DROUGHT","Drought",s2, ignore.case=T)
grep("dust",s1,ignore.case = T, value = T)
```

```
## [1] "DUST STORM" "DUST DEVIL" "Dust Devil" "Saharan Dust" "SAHARAN DUST"
## [6] "DUST DEVEL" "BLOWING DUST"
```

```
s2<-gsub("^[d]ust devil|^[d]ust devel","Dust Devil",s2,ignore.case = T)
grep("ex",s1,ignore.case = T, value = T)
```

```
## [1] "EXTREME COLD" "EXCESSIVE HEAT"
## [3] "EXTREME WINDCHILL" "Extreme Cold"
## [5] "Excessive Cold" "Extended Cold"
## [7] "Hypothermia/Exposure" "HYPOTHERMIA/EXPOSURE"
## [9] "EXCESSIVE RAIN" "HYPERTHERMIA/EXPOSURE"
## [11] "EXCESSIVELY DRY" "EXCESSIVE RAINFALL"
## [13] "EXTREME WINDCHILL TEMPERATURES" "EXTREME WIND CHILL"
## [15] "EXCESSIVE SNOW" "EXTREMELY WET"
## [17] "EXCESSIVE HEAT/DROUGHT" "EXTREME COLD/WIND CHILL"
```

```
grep("heat",s1,ignore.case = T, value = T)
```

```
## [1] "EXCESSIVE HEAT" "Heatburst" "Record Heat"
## [4] "Heat Wave" "HEAT" "RECORD HEAT"
## [7] "HEAT WAVE" "EXCESSIVE HEAT/DROUGHT"
```

There are 10 “Heat” types. So, “HEAT”, “Heat Wave”, “HEAT WAVE” are classified into “HEAT” and “EXCESSIVE HEAT”, “EXCESSIVE HEAT/DROUGHT”, “Record Heat”, “RECORD HEAT” are classified into “Excessive Heat”.

```
s2<-gsub(".+heat.*","Excessive Heat",s2,ignore.case = T)
grep("extreme|cold|wind chill",s1,ignore.case = T,value = T)
```

```
## [1] "EXTREME COLD" "EXTREME WINDCHILL"
## [3] "Unseasonable Cold" "Record Cold"
## [5] "Extreme Cold" "Excessive Cold"
## [7] "Extended Cold" "COLD"
## [9] "Cold" "Cold Temperature"
## [11] "COLD AND SNOW" "UNSEASONABLY COLD"
## [13] "Prolong Cold" "Cold and Frost"
## [15] "RECORD COLD" "PROLONG COLD"
## [17] "COLD AND FROST" "COLD TEMPERATURES"
## [19] "COLD WIND CHILL TEMPERATURES" "BITTER WIND CHILL"
## [21] "BITTER WIND CHILL TEMPERATURES" "EXTREME WINDCHILL TEMPERATURES"
## [23] "EXTREME WIND CHILL" "RECORD COLD"
## [25] "UNUSUALLY COLD" "EXTREMELY WET"
## [27] "WIND CHILL" "COLD WEATHER"
## [29] "EXTREME COLD/WIND CHILL" "COLD/WIND CHILL"
```

```
s2<-gsub("(^ex|^record)(.+cold|.+Wind.*Chill.*)","Extreme Cold/Wind Chill",s2,ignore.case = T)
grep("fla|flood",s1,ignore.case = T, value = T)
```

```
## [1] "FLASH FLOOD" "FLOOD"
## [3] "Minor Flooding" "Ice jam flood (minor"
## [5] "Coastal Flooding" "COASTAL FLOOD"
## [7] "COASTALFLOOD" "Erosion/Cstl Flood"
## [9] "Tidal Flooding" "River Flooding"
## [11] "Flood/Flash Flood" "STREET FLOODING"
## [13] "Flood" "TIDAL FLOODING"
## [15] "COASTAL FLOODING" "Urban Flooding"
## [17] "Urban flood" "Urban Flood"
## [19] "Coastal Flood" "coastal flooding"
## [21] "RIVER FLOOD" "Flood/Strong Wind"
## [23] "COASTAL FLOODING/EROSION" "URBAN/STREET FLOODING"
## [25] "COASTAL FLOODING/EROSION" "URBAN FLOOD"
## [27] "RIVER FLOODING" "RED FLAG FIRE WX"
## [29] "FLASH FLOOD/FLOOD" "FLASH FLOODING"
## [31] "FLOOD/FLASH/FLOOD" "RED FLAG CRITERIA"
## [33] "CSTL FLOODING/EROSION" "SNOWMELT FLOODING"
## [35] "LAKESHORE FLOOD"
```

There are a lot of “Flood” types. So, “FLASH FLOOD”, “FLASH FLOOD”, “FLASH FLOOD/FLOOD”, “FLASH FLOODING”, “Flood/Flash Flood”, “FLOOD/FLASH/FLOOD” are classified into “Flash Flood”. “LAKESHORE FLOOD” is classified into “Lakeshore Flood”. “Flood” and “FLOOD” are classified into “Flood”. “Flood/Strong Wind” is classified into “Strong Wind”.

```
s2<-gsub(".*flash.*","Flash Flood",s2,ignore.case = T)
s2<-gsub("^flood","Flood",s2,ignore.case = T)
grep("frost|freez",s1,ignore.case = T, value = T)
```

```
## [1] "FREEZING RAIN" "Freeze"
```

```
## [3] "DAMAGING FREEZE"          "Damaging Freeze"
## [5] "Early Frost"              "Freezing Rain"
## [7] "Freezing Fog"             "Freezing Spray"
## [9] "Freezing Drizzle"         "FREEZE"
## [11] "FREEZING DRIZZLE"         "SLEET/FREEZING RAIN"
## [13] "HARD FREEZE"              "Frost"
## [15] "Frost/Freeze"             "Cold and Frost"
## [17] "AGRICULTURAL FREEZE"      "LATE FREEZE"
## [19] "Freezing rain"            "FROST"
## [21] "Freezing drizzle"         "COLD AND FROST"
## [23] "FREEZING RAIN/SLEET"      "FIRST FROST"
## [25] "LIGHT SNOW/FREEZING PRECIP" "FREEZING FOG"
## [27] "LIGHT FREEZING RAIN"      "SNOW/FREEZING RAIN"
## [29] "FROST/FREEZE"
```

“Frost” and “Freeze” types also confuse you, so “AGRICULTURAL FREEZE”, “Cold and Frost”, “COLD AND FROST”, “Damaging Freeze”, “DAMAGING FREEZE”, “Early Frost”, “FIRST FROST”, “Freeze”, “FREEZE”, “Frost”, “FROST”, “Frost/Freeze”, “FROST/FREEZE”, “HARD FREEZE”, “LATE FREEZE” are classified into “Frost/Freeze”. “Freezing Fog” and “FREEZING FOG” are classified into “Freezing Fog”.

```
s2<-gsub(".*freeze.*|.*frost.*","Frost/Freeze",s2, ignore.case=T)
grep("Fun|Clo",s1,ignore.case = T, value = T)
```

```
## [1] "FUNNEL CLOUD" "Funnel Cloud" "FUNNEL CLOUDS" "WALL CLOUD"
```

```
s2<-gsub("Funnel Cloud.*","Funnel Cloud",s2, ignore.case=T)
s2<-gsub("Freezing Fog","Freezing Fog",s2, ignore.case=T)
grep("hail",s1,ignore.case = T, value = T)
```

```
## [1] "HAIL"          "TSTM WIND/HAIL" "small hail"      "HAIL/WIND"
## [5] "Hail(0.75)"    "Small Hail"     "SMALL HAIL"      "GUSTY WIND/HAIL"
## [9] "LATE SEASON HAIL" "NON SEVERE HAIL" "MARINE HAIL"
```

There are a lot of “Hail” types. So, “GUSTY WIND/HAIL”, “HAIL”, “Hail(0.75)”, “HAIL/WIND”, “LATE SEASON HAIL”, “NON SEVERE HAIL”, “small hail”, “Small Hail”, “SMALL HAIL”, “TSTM WIND/HAIL” are classified into “Hail”. “MARINE HAIL” is classified into “Marine HAIL”.

```
s2<-gsub("^hail.*","Hail",s2, ignore.case=T)
s2<-gsub("[^m].+hail.*","Hail",s2, ignore.case=T)
s2<-gsub("^heat wave|^heat","Heat",s2,ignore.case = T)
grep("rain|snow",s1,ignore.case = T,value=T)
```

```
## [1] "HEAVY RAIN"          "FREEZING RAIN"
## [3] "HEAVY SNOW"          "Snow"
## [5] "Snow Squalls"        "Heavy Rain"
## [7] "Heavy Rain and Wind" "Light Snow/Flurries"
## [9] "SNOW"                "Heavy Rain/High Surf"
## [11] "Rain Damage"         "blowing snow"
## [13] "Torrential Rainfall" "Freezing Rain"
## [15] "Late-season Snowfall" "HEAVY RAIN/WIND"
## [17] "Snow squalls"        "Ice/Snow"
```

## [19] "Snow Accumulation"	"Drifting Snow"
## [21] "Heavy snow shower"	"Heavy rain"
## [23] "LATE SNOW"	"Record May Snow"
## [25] "Record Winter Snow"	"Light snow"
## [27] "Late Season Snowfall"	"Light Snow"
## [29] "Snow and Ice"	"Light Snowfall"
## [31] "Gusty wind/rain"	"GUSTY WIND/HVY RAIN"
## [33] "Blowing Snow"	"Early snowfall"
## [35] "Monthly Snowfall"	"Seasonal Snowfall"
## [37] "Monthly Rainfall"	"Thundersnow shower"
## [39] "COLD AND SNOW"	"RAIN"
## [41] "RAIN/SNOW"	"SLEET/FREEZING RAIN"
## [43] "TSTM HEAVY RAIN"	"SNOW AND ICE"
## [45] "RAIN (HEAVY)"	"SNOW SQUALL"
## [47] "HEAVY SNOW SQUALLS"	"SNOW/ICE"
## [49] "SNOW SQUALLS"	"LAKE EFFECT SNOW"
## [51] "RECORD SNOWFALL"	"Lake Effect Snow"
## [53] "LIGHT SNOW"	"Snow and sleet"
## [55] "Freezing rain"	"EXCESSIVE RAIN"
## [57] "Mountain Snows"	"UNSEASONAL RAIN"
## [59] "EARLY RAIN"	"PROLONGED RAIN"
## [61] "FREEZING RAIN/SLEET"	"MONTHLY RAINFALL"
## [63] "MODERATE SNOW"	"MODERATE SNOWFALL"
## [65] "SNOW DROUGHT"	"SNOW/SLEET"
## [67] "SNOW AND SLEET"	"LIGHT SNOW/FREEZING PRECIP"
## [69] "EXCESSIVE RAINFALL"	"HEAVY RAINFALL"
## [71] "EARLY SNOWFALL"	"EXCESSIVE SNOW"
## [73] "LIGHT FREEZING RAIN"	"MONTHLY SNOWFALL"
## [75] "RECORD RAINFALL"	"RECORD SNOW"
## [77] "SNOW/FREEZING RAIN"	"SNOW/BLOWING SNOW"
## [79] "LOCALLY HEAVY RAIN"	"LATE SEASON SNOW"
## [81] "SNOW ADVISORY"	"UNUSUALLY LATE SNOW"
## [83] "RECORD LOW RAINFALL"	"SNOWMELT FLOODING"
## [85] "ACCUMULATED SNOWFALL"	"FALLING SNOW/ICE"
## [87] "SNOW SHOWERS"	"HEAVY RAIN EFFECTS"
## [89] "FIRST SNOW"	"ICE/SNOW"
## [91] "LAKE-EFFECT SNOW"	

There are a lot of rain and snow terms. So, "EXCESSIVE RAIN", "EXCESSIVE RAINFALL", "GUSTY WIND/HVY RAIN", "Heavy rain", "Heavy Rain", "HEAVY RAIN", "Heavy Rain and Wind", "HEAVY RAIN EFFECTS", "Heavy Rain/High Surf", "HEAVY RAIN/WIND", "HEAVY RAINFALL", "LOCALLY HEAVY RAIN", "RAIN (HEAVY)", "RECORD RAINFALL", "Torrential Rainfall", "TSTM HEAVY RAIN" are translated into "Heavy Rain". And, "EXCESSIVE SNOW", "HEAVY SNOW", "Heavy snow shower", "HEAVY SNOW SQUALLS", "RECORD SNOW", "RECORD SNOWFALL", "Record Winter Snow" are translated into "Heavy Snow".

```
grep("rainfall",s1,ignore.case = T,value = T)
```

```
## [1] "Torrential Rainfall" "Monthly Rainfall"      "MONTHLY RAINFALL"
## [4] "EXCESSIVE RAINFALL"  "HEAVY RAINFALL"        "RECORD RAINFALL"
## [7] "RECORD LOW RAINFALL"
```

```
s2<-gsub(".*heavy rain.*|.hvy rain.*|rain.+heavy.|^[^m].+rainfall|EXCESSIVE RAIN","Heavy Rain",s2,
        ignore.case=T)
grep("snowfall",s1,ignore.case = T,value = T)
```

```
## [1] "Late-season Snowfall" "Late Season Snowfall" "Light Snowfall"
## [4] "Early snowfall"      "Monthly Snowfall"      "Seasonal Snowfall"
## [7] "RECORD SNOWFALL"      "MODERATE SNOWFALL"      "EARLY SNOWFALL"
## [10] "MONTHLY SNOWFALL"     "ACCUMULATED SNOWFALL"
```

```
grep("snow[^f]|snow$",s1,ignore.case = T,value = T)
```

```
## [1] "HEAVY SNOW"          "Snow"
## [3] "Snow Squalls"       "Light Snow/Flurries"
## [5] "SNOW"               "blowing snow"
## [7] "Snow squalls"       "Ice/Snow"
## [9] "Snow Accumulation"  "Drifting Snow"
## [11] "Heavy snow shower"  "LATE SNOW"
## [13] "Record May Snow"    "Record Winter Snow"
## [15] "Light snow"         "Light Snow"
## [17] "Snow and Ice"       "Blowing Snow"
## [19] "Thundersnow shower" "COLD AND SNOW"
## [21] "RAIN/SNOW"         "SNOW AND ICE"
## [23] "SNOW SQUALL"       "HEAVY SNOW SQUALLS"
## [25] "SNOW/ICE"          "SNOW SQUALLS"
## [27] "LAKE EFFECT SNOW"   "Lake Effect Snow"
## [29] "LIGHT SNOW"         "Snow and sleet"
## [31] "Mountain Snows"     "MODERATE SNOW"
## [33] "SNOW DROUGHT"       "SNOW/SLEET"
## [35] "SNOW AND SLEET"     "LIGHT SNOW/FREEZING PRECIP"
## [37] "EXCESSIVE SNOW"     "RECORD SNOW"
## [39] "SNOW/FREEZING RAIN" "SNOW/BLOWING SNOW"
## [41] "LATE SEASON SNOW"   "SNOW ADVISORY"
## [43] "UNUSUALLY LATE SNOW" "SNOWMELT FLOODING"
## [45] "FALLING SNOW/ICE"   "SNOW SHOWERS"
## [47] "FIRST SNOW"         "ICE/SNOW"
## [49] "LAKE-EFFECT SNOW"
```

```
s2<-gsub("RECORD SNOWFALL|HEAVY SNOW.*|EXCESSIVE SNOW|RECORD[^m]+SNOW","Heavy Snow",s2,ignore.case = T)
grep("High|Surf",s1,ignore.case = T,value = T)
```

```
## [1] "HIGH WIND"          "High Wind"          "ROUGH SURF"
## [4] "Heavy Surf"         "HEAVY SURF"         "High Surf"
## [7] "Heavy Rain/High Surf" "HIGH SURF"          "Record High"
## [10] "Heavy surf and wind" "HIGH SWELLS"        "HIGH SWELLS"
## [13] "HIGH WINDS"         "HIGH WIND (G40)"    "HIGH SEAS"
## [16] "HIGH WATER"         "HIGH SURF ADVISORY" "HAZARDOUS SURF"
## [19] "ASTRONOMICAL HIGH TIDE" "HIGH SURF ADVISORIES" "HEAVY SURF/HIGH SURF"
## [22] "MARINE HIGH WIND"
```

```
s2<-gsub(".*surf.*","High Surf",s2,ignore.case = T)
s2<-gsub("^high wind.*","High Wind",s2,ignore.case = T)
```

```
s2<-gsub("^HURRICANE.*|. *TYPHOON$", "Hurricane (Typhoon)", s2, ignore.case = T)
s2<-gsub("^Ice ?Storm.*", "Ice Storm", s2, ignore.case = T)
s2<-gsub("lake.e.+ ", "Lake-Effect Snow", s2, ignore.case = T)
grep("Lightn", s1, ignore.case = T, value = T)
```

```
## [1] "LIGHTNING" "TSTM WIND AND LIGHTNING"
```

“LIGHTNING” and “LIGHTNING” are names as “Lightning” and “TSTM WIND AND LIGHTNING” belongs to “Thunderstorm Wind”, because TSTM means Thunderstorm.

```
s2<-gsub("LIGHTNING", "Lightning", s2, ignore.case = T)
grep("marine", s1, ignore.case = T, value = T)
```

```
## [1] "Marine Accident" "MARINE TSTM WIND"
## [3] "MARINE HAIL" "MARINE HIGH WIND"
## [5] "MARINE THUNDERSTORM WIND" "MARINE STRONG WIND"
```

```
s2<-gsub("Marine T.+ ", "Marine Thunderstorm Wind", s2, ignore.case = T)
s2<-gsub("Rip Current.*", "Rip Current", s2, ignore.case = T)
grep("Seiche", s1, ignore.case = T, value = T)
```

```
## [1] "SEICHE"
```

```
s2<-gsub(".*sleet.*", "Sleet", s2, ignore.case = T)
s2<-gsub("Storm Surge.*", "Storm Surge/Tide", s2, ignore.case = T)
s2<-gsub("^[^m]*Strong Wind.*", "Strong Wind", s2, ignore.case = T)
grep("Thunderstorm.*Wind|TSTM.*Wind", s1, ignore.case = T, value = T)
```

```
## [1] "TSTM WIND" "TSTM WIND/HAIL"
## [3] "Tstm Wind" "Thunderstorm Wind"
## [5] "TSTM WIND (G45)" "TSTM WIND 40"
## [7] "TSTM WIND 45" "TSTM WIND (41)"
## [9] "TSTM WIND (G40)" "TSTM WIND AND LIGHTNING"
## [11] "TSTM WIND (G45)" "TSTM WIND (G35)"
## [13] "TSTM WINDS" "TSTM WIND G45"
## [15] "THUNDERSTORM WIND (G40)" "NON-TSTM WIND"
## [17] "NON TSTM WIND" "GUSTY THUNDERSTORM WINDS"
## [19] "MARINE TSTM WIND" "GUSTY THUNDERSTORM WIND"
## [21] "THUNDERSTORM WIND" "MARINE THUNDERSTORM WIND"
```

There are a lot of “Thunderstorm Wind” types. So, “GUSTY THUNDERSTORM WIND”, “GUSTY THUNDERSTORM WINDS”, “Thunderstorm Wind”, “THUNDERSTORM WIND”, “THUNDERSTORM WIND (G40)”, “Tstm Wind”, “TSTM WIND”, “TSTM WIND”, “TSTM WIND”, “TSTM WIND (G45)”, “TSTM WIND (41)”, “TSTM WIND (G35)”, “TSTM WIND (G40)”, “TSTM WIND (G45)”, “TSTM WIND (G45)”, “TSTM WIND 40”, “TSTM WIND 45”, “TSTM WIND AND LIGHTNING”, “TSTM WIND G45”, “TSTM WINDS” are renamed as “Thunderstorm Wind”.

```
s2<-gsub("^[^m]*Thunderstorm Wind.*|^TSTM W.*ND$|TSTM WiND[/]+", "Thunderstorm Wind", s2, ignore.case = T)
grep("Tornado", s1, ignore.case = T, value = T)
```

```
## [1] "TORNADO" "TORNADO DEBRIS"
```

```
s2<-gsub("^Tornado.*","Tornado",s2,ignore.case = T)
grep("Tropical|Depression",s1,ignore.case = T,value = T)
```

```
## [1] "TROPICAL STORM"      "TROPICAL DEPRESSION"
```

```
grep("Tsunami",s1,ignore.case = T,value = T)
```

```
## [1] "TSUNAMI"
```

```
grep("Volcanic|Ash",s1,ignore.case = T,value = T)
```

```
## [1] "FLASH FLOOD"      "Flood/Flash Flood"  "Volcanic Ash"
## [4] "Volcanic Ash Plume" "VOLCANIC ASH"       "VOLCANIC ERUPTION"
## [7] "FLASH FLOOD/FLOOD" "FLASH FLOODING"     "FLOOD/FLASH/FLOOD"
## [10] "VOLCANIC ASHFALL"
```

```
s2<-gsub("^Volcanic Ash.*","Volcanic Ash",s2,ignore.case = T)
grep("Water|spout",s1,ignore.case = T,value = T)
```

```
## [1] "WATERSPOUT"  "WATERSPOUTS" "LANDSPOUT"   "HIGH WATER"
```

```
s2<-gsub("^Waterspout.*","Waterspout",s2,ignore.case = T)
grep("Wild|fire",s1,ignore.case = T,value = T)
```

```
## [1] "WILD/FOREST FIRE" "RED FLAG FIRE WX" "BRUSH FIRE"      "WILDFIRE"
```

```
s2<-gsub("Wild.*fire","Wildfire",s2,ignore.case = T)
grep("Winter.*Storm",s1,ignore.case = T,value = T)
```

```
## [1] "WINTER STORM"
```

```
grep("Winter|Weather",s1,ignore.case = T,value = T)
```

```
## [1] "WINTER STORM"      "Winter Weather"     "Record Winter Snow"
## [4] "No Severe Weather" "WINTER WEATHER"     "WINTER MIX"
## [7] "DRY WEATHER"       "HOT WEATHER"        "WINTER MIX"
## [10] "WARM WEATHER"      "WINTER WEATHER MIX" "WINTER WEATHER/MIX"
## [13] "COLD WEATHER"
```

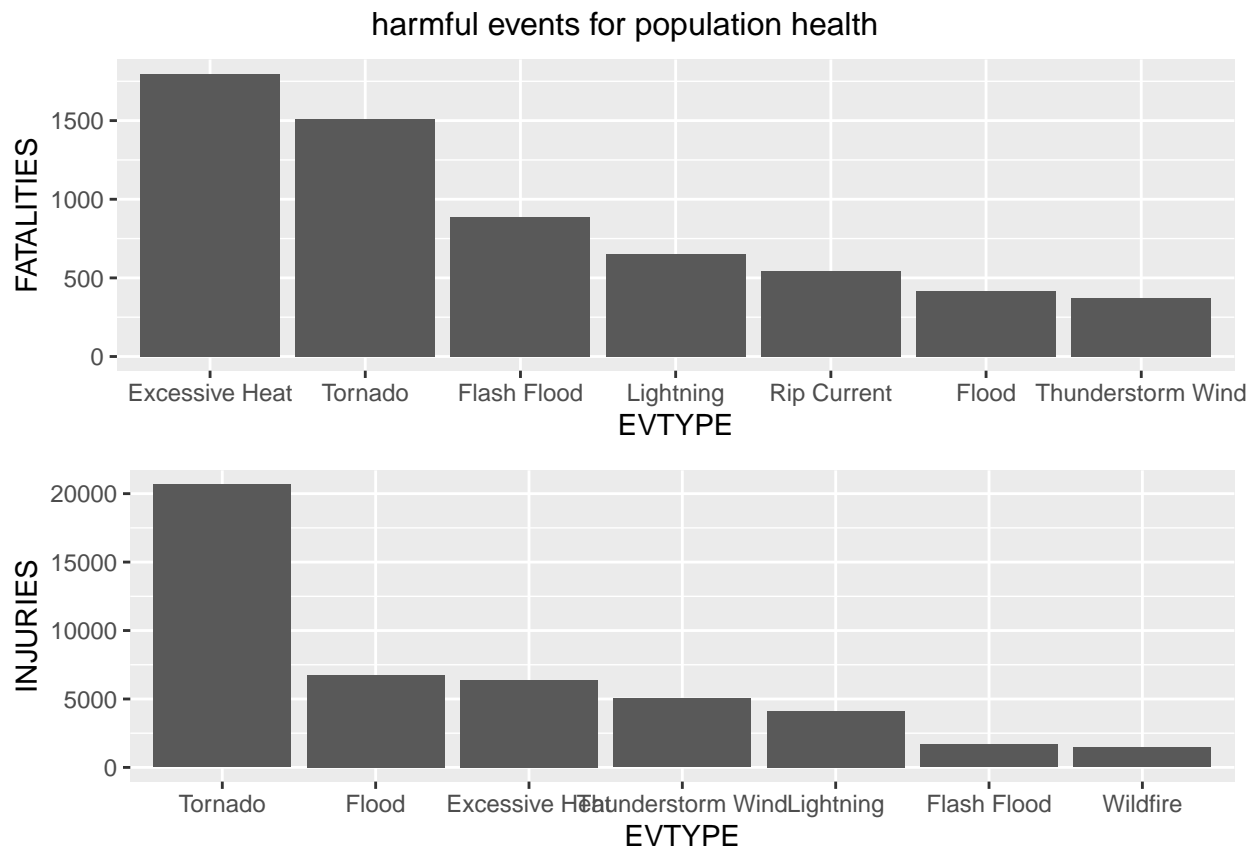
```
s2<-gsub(".*Winter Weather.*","Winter Weather",s2,ignore.case = T)
```

Rename EVTYPE names.

Result

So, make a population health plot. The top 7 events with the highest total fatalities and injuries are classified into “harmful events for population health”.


```
rp$EVTYPE<-s2
arpf<-aggregate(list(FATALITIES=rp$FATALITIES),by=list(EVTYPE=rp$EVTYPE),sum)
fatalities<-head(arrange(arpf,desc(FATALITIES)),n=7)
p1<-ggplot(fatalities,aes(x=reorder(EVTYPE,-FATALITIES),y=FATALITIES))+geom_bar(stat = "identity") +labs(x="EVTYPE",y="FATALITIES")
arpi<-aggregate(list(INJURIES=rp$INJURIES),by=list(EVTYPE=rp$EVTYPE),sum)
injuries<-head(arrange(arpi,desc(INJURIES)),n=7)
p2<-ggplot(injuries,aes(x=reorder(EVTYPE,-INJURIES),y=INJURIES))+geom_bar(stat = "identity") +labs(x="EVTYPE",y="INJURIES")
grid.arrange(p1,p2,ncol=1,top="harmful events for population health")
```



```
dev.copy(png,file="population health.png")
```

```
## png
## 3
```

```
dev.off()
```

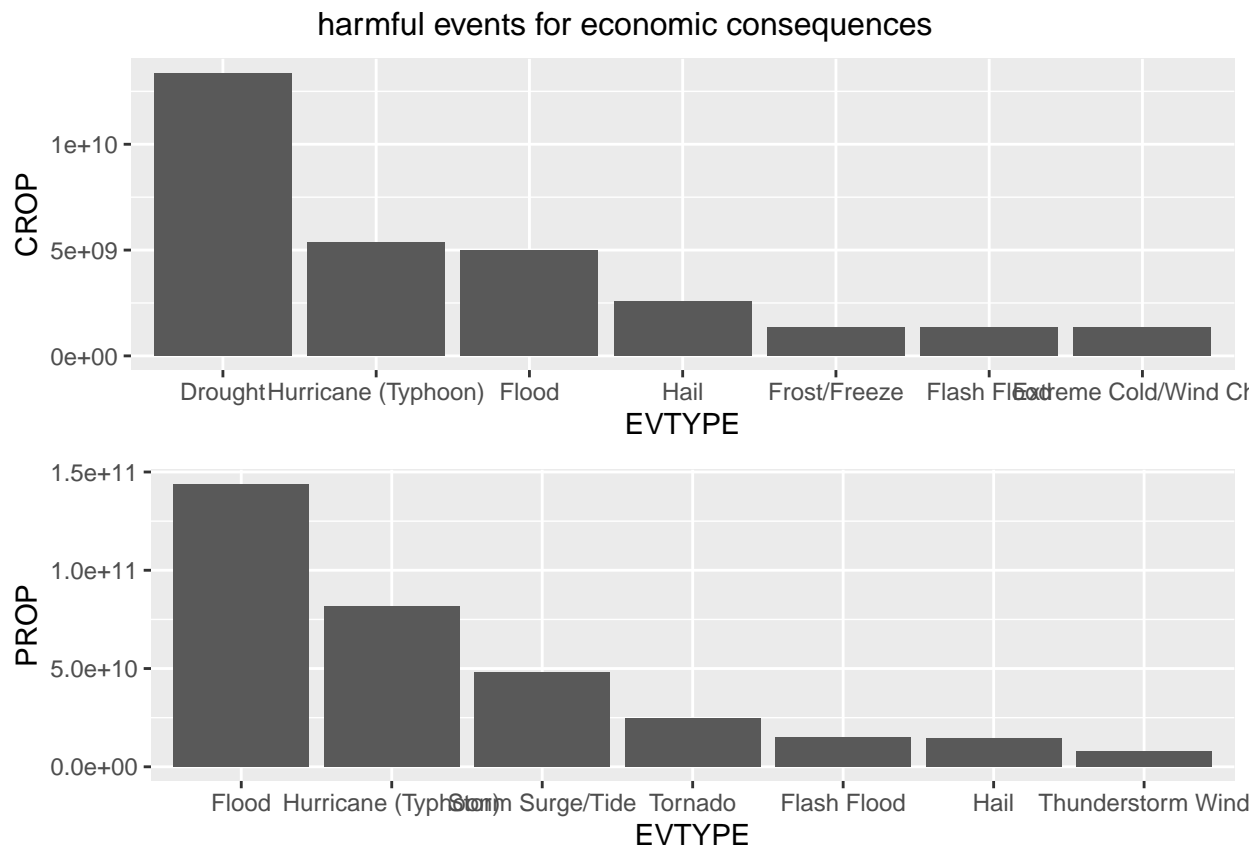
```
## pdf
## 2
```

Then, make a economic consequences plot. The top 7 events with the highest total “CROP” and “PROP” are classified into “harmful events for economic consequences”.

```

arpc<-aggregate(list(CROP=rp$CROP),by=list(EVTYPE=rp$EVTYPE),sum)
fatalities<-head(arrange(arpc,desc(CROP)),n=7)
p1<-ggplot(fatalities,aes(x=reorder(EVTYPE,-CROP),y=CROP))+geom_bar(stat = "identity") +labs(x="EVTYPE")
arpp<-aggregate(list(PROP=rp$PROP),by=list(EVTYPE=rp$EVTYPE),sum)
injuries<-head(arrange(arpp,desc(PROP)),n=7)
p2<-ggplot(injuries,aes(x=reorder(EVTYPE,-PROP),y=PROP))+geom_bar(stat = "identity") +labs(x="EVTYPE")
grid.arrange(p1,p2,ncol=1,top="harmful events for economic consequences")

```



```

dev.copy(png,file="economic consequences.png")

```

```

## png
## 3

```

```

dev.off()

```

```

## pdf
## 2

```