NOAA Weather Analysis

Kim Kirk

January 16, 2018

U.S. National Oceanic and Atmospheric Administration's (NOAA) storm analysis

Storms and other severe weather events that have caused both public health and economic problems for communities and municipalities across the United States

Synopsis

The business questions, what is the highest ranking weather event for economic impact? and what is the highest ranking weather event for public health impact?, were answered. NOAA data was imported, cleaned, and analyzed for weather events that caused the most property and crop damage, as well as the most injuries and fatalities across the United States. Data was imported as csv file with variables not relevant to this analysis dropped to increase import speed. Weather event type (EVTYPE) was cleaned for obvious misspellings, see "Data Processing" for additional details. Injuries and fatalities (INJURIES, FATALITIES variables) were summed together to get a full scope for their impact on public health. Property and crops damage (PROPDMG, CROPDMG variables) were transformed into base 10 values and summed together to get a full scope for their impact on economic problems. The top ten highest ranking weather event types for public health impact and economic impact are presented as two separate visualizations. Top 2 weather events for public health impact include: Tornado and Thunderstorm Wind, for economic impact include: Flood and Hurricane/Typhoon.

Data Processing

The codebook provided by NOAA and their FAQ does not explicitly list variables and what they measure, so assumptions were made based on reading through the codebook and parsing together information gleaned there. The codebook designates that "only events permitted in Storm Data are listed in Table 1 of Section 2.1.1."; however, data entry personnel entered values not designated in the event types (EVTYPE). There may be values that essentially qualify as the same event type; however, they were entered as different event types and the variables for the data set were not clearly coded in the codebook, making it impossible to distinguish events types observed from designated event types per NOAA.

The variables PROPDMGEXP and CROPDMGEXP are assumed to be unit designators for the PROPDMG and CROPDMG variables. Exploratory analysis of those variables showed many missing values and values that don't match the codebook designators. Normally, one would go back to the data source (NOAA) to inquire as to what the column value represents, but there is no way to contact them directly; thus, these values were

dropped from the PROPDMG and CROPDMG variables. Breakdowns for percentage of values dropped include ~51% of total values PROPDMGEXP, 1% of total values PROPDMGEXP have values that do not make sense with the designators listed in the codebook (e.g. ? symbol, numbers, etc), ~68% of total CROPDMGEXP, 1% of total values CROPDMGEXP have values that do not make sense with the designators listed in the codebook (e.g. ? symbol, numbers, etc).

Percentage of values that can be analyzed include ~48% of total values PROPDMGEXP, ~31% of total values CROPDMGEXP. It is recognized more than 4% of values have to be dropped; thus, the analysis is not going to be as accurate. Normally, with missing data that is 4% or less, a valid solution would be to replace these values with the mean for the variable. However, upon additional research, the data is missing because there was bad reporting not due to errors in reporting; thus the data is sparse in nature.

NOAA data were downloaded from the specified URL and loaded into R as csv file with nonrelevant variables dropped to improve import speed and analysis.

```
##set file path to save download to
path <- file.path(paste(getwd(), 'repdata%2Fdata%2FStormData.csv.bz2', sep = "/"))

##set url for download
url <- "https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2"

##download file and save to working directory
download.file(url, path)

##read in dataset
stormDataset <- read.csv("repdata%2Fdata%2FStormData.csv.bz2", stringsAsFactors = FALSE, na.strings = c("NA", ""), colClasse
s = c("NULL", "NULL", "NULL","NULL","NULL","NULL", "NULL", "character", "NULL", "NULL", "NULL","NULL","NULL","NULL", "NULL",
 "NULL", "NULL", "NULL","NULL","NULL","NULL", "NULL", "numeric", "numeric", "numeric", "character", "numeric", "character",
"NULL", "NULL", "NULL","NULL","NULL","NULL", "NULL", "NULL", "NULL")
)
```

The packages DPLYR, GGPLOT2, and FORMATTABLE were loaded to be used as needed throughout the analysis.

```
##check environment for dplyr package and install or not as required
##credit Matthew on StackOverflow https://stackoverflow.com/users/4125693/matthew
using<-function(...) {
    libs<-unlist(list(...))
    req<-unlist(lapply(libs,require,character.only=TRUE))
    need<-libs[req==FALSE]
    n<-length(need)
    if(n>0){
        libsmsg<-if(n>2) paste(paste(need[1:(n-1)],collapse=", "),",",sep="") else need[1]
        print(libsmsg)
        if(n>1){
            libsmsg<-paste(libsmsg," and ", need[n],sep="")
        }
        libsmsg<-paste("The following packages could not be found: ",libsmsg,"\n\r\n\rInstall missing packages?",collapse=""
)
        if(winDialog(type = c("yesno"), libsmsg)=="YES"){
            install.packages(need)
            lapply(need,require,character.only=TRUE)
        }
    }
}

##install and load packages
using("dplyr")
```

```
## Loading required package: dplyr
```

```
## Warning: package 'dplyr' was built under R version 3.4.3
```

```
##
## 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
```

```
using("ggplot2")
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.4.3
```

```
using("formattable")
```

```
## Loading required package: formattable
```

```
## Warning: package 'formattable' was built under R version 3.4.3
```

The INJURIES and FATALITIES variables were combined into new column that is sum of fatalities and injuries; to make it easier to graph the total of fatalities and injuries.

```
##create new column that holds sum of INJURIES and FATALITIES columns
injuriesFatalitiesDataset <- mutate(stormDataset, Injuries_Fatalities = INJURIES + FATALITIES)
```

The associated EVTYPE was changed to all upper case to faciliate accurate grouping.

```
##convert all values for EVTYPE variable to upper case to faciliate spelling search
injuriesFatalitiesDataset$EVTYPE <- toupper(injuriesFatalitiesDataset$EVTYPE)
```

Any zero sum values were removed from the data set as they are not needed for analysis.

```
##retrieve injuries and fatalities that are greater than 0; subset the data to make it easier to clean
injuriesFatalitiesDataset <- injuriesFatalitiesDataset[injuriesFatalitiesDataset$Injuries_Fatalities != 0, ]
```

exploratory analysis of the misspelled words in the EVTYPE variable for injuries and fatalities analysis

```
##exploratory analysis of the misspelled words in the EVTYPE variable for injuries and fatalities analysis
table(injuriesFatalitiesDataset$EVTYPE)
```

```
##
##                     AVALANCE                     AVALANCHE
##                            1                           239
##                    BLACK ICE                     BLIZZARD
##                            1                            86
##                  BLOWING SNOW                   BRUSH FIRE
##                            2                             1
##                COASTAL FLOOD               COASTAL FLOODING
##                            4                             2
##       COASTAL FLOODING/EROSION                 COASTAL STORM
##                            1                             4
##                  COASTALSTORM                          COLD
##                            1                            33
##                COLD AND SNOW               COLD TEMPERATURE
##                            1                             2
##                    COLD WAVE                  COLD WEATHER
##                            1                             4
##               COLD/WIND CHILL                   COLD/WINDS
##                           75                             1
##                    DENSE FOG                       DROUGHT
##                           34                             1
##        DROUGHT/EXCESSIVE HEAT                     DROWNING
##                            1                             1
##                DRY MICROBURST           DRY MIRCOBURST WINDS
##                           14                             1
##                   DUST DEVIL                     DUST STORM
##                           15                            45
##                EXCESSIVE HEAT             EXCESSIVE RAINFALL
##                          678                             1
##                EXCESSIVE SNOW                 EXTENDED COLD
##                            2                             1
##                  EXTREME COLD        EXTREME COLD/WIND CHILL
##                          117                            91
##                  EXTREME HEAT              EXTREME WINDCHILL
##                           14                            15
##               FALLING SNOW/ICE                   FLASH FLOOD
##                            2                           931
##             FLASH FLOOD/FLOOD                 FLASH FLOODING
##                            6                            19
##           FLASH FLOODING/FLOOD                   FLASH FLOODS
```

```
##                            2                             2
##                        FLOOD              FLOOD & HEAVY RAIN
##                          410                             1
##               FLOOD/FLASH FLOOD               FLOOD/RIVER FLOOD
##                           16                             1
##                     FLOODING                           FOG
##                            5                            92
##      FOG AND COLD TEMPERATURES                       FREEZE
##                            1                             1
##                FREEZING DRIZZLE                 FREEZING RAIN
##                            4                             7
##              FREEZING RAIN/SNOW                FREEZING SPRAY
##                            1                             1
##                        FROST                  FUNNEL CLOUD
##                            1                             2
##                        GLAZE                GLAZE/ICE STORM
##                           16                             1
##                    GUSTY WIND                   GUSTY WINDS
##                            2                             8
##                         HAIL                HAZARDOUS SURF
##                          288                             1
##                         HEAT                     HEAT WAVE
##                          209                            31
##              HEAT WAVE DROUGHT                   HEAT WAVES
##                            1                             1
##                    HEAVY RAIN                   HEAVY RAINS
##                          121                             1
##                    HEAVY SEAS                   HEAVY SNOW
##                            2                           189
##      HEAVY SNOW AND HIGH WINDS             HEAVY SNOW SHOWER
##                            1                             1
##  HEAVY SNOW/BLIZZARD/AVALANCHE                HEAVY SNOW/ICE
##                            1                             1
##                    HEAVY SURF            HEAVY SURF AND WIND
##                           24                             1
##           HEAVY SURF/HIGH SURF                          HIGH
##                           39                             1
##                     HIGH SEAS                    HIGH SURF
##                            7                            91
##                   HIGH SWELLS                   HIGH WATER
##                            1                             2
```

```
##                    HIGH WAVES                     HIGH WIND
##                             1                           525
##                  HIGH WIND 48             HIGH WIND AND SEAS
##                             1                             1
##           HIGH WIND/HEAVY SNOW                 HIGH WIND/SEAS
##                             1                             1
##                    HIGH WINDS                HIGH WINDS/COLD
##                            98                             2
##                HIGH WINDS/SNOW                     HURRICANE
##                             1                            31
##     HURRICANE-GENERATED SWELLS              HURRICANE EDOUARD
##                             2                             1
##                HURRICANE EMILY                 HURRICANE ERIN
##                             1                             4
##                HURRICANE FELIX                 HURRICANE OPAL
##                             1                             2
##      HURRICANE OPAL/HIGH WINDS              HURRICANE/TYPHOON
##                             1                            26
##          HYPERTHERMIA/EXPOSURE                    HYPOTHERMIA
##                             1                             1
##           HYPOTHERMIA/EXPOSURE                           ICE
##                             6                            10
##                   ICE ON ROAD                     ICE ROADS
##                             1                             1
##                     ICE STORM          ICE STORM/FLASH FLOOD
##                            95                             1
##                     ICY ROADS                     LANDSLIDE
##                            15                            20
##                   LANDSLIDES                    LIGHT SNOW
##                             1                             1
##                     LIGHTNING LIGHTNING AND THUNDERSTORM WIN
##                          3305                             1
##               LIGHTNING INJURY                    LIGHTNING.
##                             1                             1
##                LOW TEMPERATURE                MARINE ACCIDENT
##                             7                             1
##               MARINE HIGH WIND                  MARINE MISHAP
##                             1                             2
##             MARINE STRONG WIND       MARINE THUNDERSTORM WIND
##                            16                            11
##               MARINE TSTM WIND                 MINOR FLOODING
```

```
##                          6                          1
##                MIXED PRECIP                   MUDSLIDE
##                          6                          2
##                  MUDSLIDES       NON-SEVERE WIND DAMAGE
##                          1                          1
##              NON TSTM WIND                      OTHER
##                          1                          1
##                  RAIN/SNOW                  RAIN/WIND
##                          2                          1
##         RAPIDLY RISING WATER                RECORD COLD
##                          1                          1
##                RECORD HEAT       RECORD/EXCESSIVE HEAT
##                          2                          1
##                RIP CURRENT                RIP CURRENTS
##                        399                        238
##      RIP CURRENTS/HEAVY SURF                RIVER FLOOD
##                          2                          4
##              RIVER FLOODING                 ROGUE WAVE
##                          3                          1
##                ROUGH SEAS                   ROUGH SURF
##                          3                          2
##                       SLEET                 SMALL HAIL
##                          1                          3
##                        SNOW                SNOW AND ICE
##                          12                         3
##                SNOW SQUALL                 SNOW SQUALLS
##                          3                          1
##           SNOW/ BITTER COLD              SNOW/HIGH WINDS
##                          1                          2
##                STORM SURGE              STORM SURGE/TIDE
##                         11                          2
##                STRONG WIND                 STRONG WINDS
##                        217                         18
##                THUNDERSNOW                 THUNDERSTORM
##                          1                          4
##          THUNDERSTORM  WINDS             THUNDERSTORM WIND
##                          1                         682
##      THUNDERSTORM WIND (G40)          THUNDERSTORM WIND G52
##                          1                          1
##           THUNDERSTORM WINDS          THUNDERSTORM WINDS 13
##                        359                          1
```

```
##        THUNDERSTORM WINDS/HAIL            THUNDERSTORM WINDSS
##                              1                              1
##            THUNDERSTORMS WINDS                 THUNDERSTORMW
##                              1                              1
##              THUNDERTORM WINDS                TIDAL FLOODING
##                              1                              1
##                        TORNADO                    TORNADO F2
##                           7928                              2
##                     TORNADO F3    TORNADOES, TSTM WIND, HAIL
##                              1                              1
##           TORRENTIAL RAINFALL                TROPICAL STORM
##                              1                             42
##         TROPICAL STORM GORDON                     TSTM WIND
##                              1                          2930
##               TSTM WIND (G35)               TSTM WIND (G40)
##                              1                              1
##               TSTM WIND (G45)                TSTM WIND/HAIL
##                              1                             28
##                        TSUNAMI                     TYPHOON
##                              2                              2
##            UNSEASONABLY COLD             UNSEASONABLY WARM
##                              1                              6
##     UNSEASONABLY WARM AND DRY URBAN AND SMALL STREAM FLOODIN
##                              1                              1
##          URBAN/SML STREAM FLD                  WARM WEATHER
##                             43                              1
##                    WATERSPOUT            WATERSPOUT TORNADO
##                              5                              1
##             WATERSPOUT/TORNADO                    WHIRLWIND
##                              2                              1
##                    WILD FIRES              WILD/FOREST FIRE
##                              1                            135
##                      WILDFIRE                          WIND
##                            196                             39
##                    WIND STORM                         WINDS
##                              1                              2
##                  WINTER STORM        WINTER STORM HIGH WINDS
##                            227                              1
##                 WINTER STORMS                WINTER WEATHER
##                              1                             49
##             WINTER WEATHER MIX            WINTER WEATHER/MIX
```

```
##                             2                             25
##                    WINTRY MIX
##                             3
```

The weather event type variable (EVTYPE) was cleaned to correct obvious errors in spelling. Vectors of misspelled words were matched with correctly spelled words and reassigned to the found values. The decision to further clean the EVTYPE variable so that all values are recoded to codebook designated values was not carried out because some of the weather values reported may be observations that don't fit the codebook designated values; better to keep original values that show the reporter's observations than incorrectly recode values or mislabel values and miss crucial observations.

```
##vector of misspelled words
misspelled <- c("AVALANCE","COASTAL FLOODING", "COASTALSTORM", "FLASH FLOODING",
                    "FLOODING", "COLD TEMPERATURE",  "FLASH FLOODS", "FLASH FLOODING",
                    "FLASH FLOODS", "FLASH FLOODING/FLOOD", "FLOOD/FLASH FLOOD",
                    "FLOODING", "GUSTY WINDS", "HEAT WAVES", "HEAVY RAINS" , "HIGH WINDS",
                    "HIGH WIND AND SEAS", "HURRICANE OPAL/HIGH WINDS", "HYPOTHERMIA",
                    "ICE ON ROAD", "ICE ROADS", "ICE ROADS", "LANDSLIDES", "LIGHTNING.",
                    "MUDSLIDES", "MARINE ACCIDENT", "MARINE TSTM WIND", "RIP CURRENTS",
                    "RIVER FLOODING","SNOW SQUALLS", "STRONG WINDS", "THUNDERSNOW",
                    "THUNDERSTORM WINDS", "THUNDERSTORMW", "THUNDERSTORM WINDSS",
                    "THUNDERSTORMS WINDS", "THUNDERSTORM  WINDS", "THUNDERTORM WINDS", "TSTM WIND", "URBAN AND SMALL STREAM
 FLOODIN",
                    "URBAN/SML STREAM FLD", "WATERSPOUT TORNADO", "WILD FIRES", "WINDS",
                    "WINTER STORMS", "WINTER WEATHER MIX", "WINTRY MIX", "FLOODS", "HEAVY SURF/HIGH SURF",
                    "SEVERE THUNDERSTORM WINDS", "SEVERE THUNDERSTORMS", "THUDERSTORM WINDS",
                    "THUNDERSTORM WINDS HAIL", "THUNDERSTORMS", "THUNDERSTORMS WIND", "TSTM WIND/HAIL",
                    "WILD/FOREST FIRES", "WILDFIRES", "WIND DAMAGE", "HAIL/WINDS", "HEAT WAVE",
                    "URBAN FLOODING")

##vector of values used to correct misspelled words
correct <- c("AVALANCHE", "COASTAL FLOOD", "COASTAL STORM", "FLASH FLOOD",
                    "FLOOD", "COLD", "FLASH FLOOD", "FLASH FLOOD","FLASH FLOOD",
                    "FLASH FLOOD/FLOOD", "FLASH FLOOD/FLOOD", "FLOOD", "GUSTY WIND",
                    "HEAT WAVE", "HEAVY RAIN", "HIGH WIND", "HIGH WIND/SEAS", "HURRICANE OPAL",
                    "HYPOTHERMIA/EXPOSURE", "ICY ROADS", "ICY ROADS", "ICY ROADS",
                    "LANDSLIDE", "LIGHTNING", "MUDSLIDE", "MARINE MISHAP",
                    "MARINE THUNDERSTORM WIND", "RIP CURRENT", "RIVER FLOOD",
                    "SNOW SQUALL", "STRONG WIND", "THUNDERSTORM", "THUNDERSTORM WIND",
                    "THUNDERSTORM WIND", "THUNDERSTORM WIND", "THUNDERSTORM WIND",
                    "THUNDERSTORM WIND", "THUNDERSTORM WIND", "THUNDERSTORM WIND", "URBAN/SMALL STREAM FLOOD",
                    "URBAN/SMALL STREAM FLOOD", "WATERSPOUT/TORNADO", "WILDFIRE",
                    "WIND", "WINTER STORM", "WINTER WEATHER/MIX", "WINTER WEATHER/MIX",
                    "FLOOD", "HIGH SURF", "SEVERE THUNDERSTORM WIND", "SEVERE THUNDERSTORM WIND",
                    "THUNDERSTORM WIND","THUNDERSTORM WINDS/HAIL", "THUNDERSTORM", "THUNDERSTORM", "THUNDERSTORM WINDS/HAIL"
,
                    "WILD/FOREST FIRE", "WILDFIRE", "WIND", "HAIL/WIND", "HEAT", "URBAN FLOOD")


##correct the misspelled words; it takes more lines of code to create an object that can be passed by
for(i in 1:length(misspelled)) {
```

```
    injuriesFatalitiesDataset[["EVTYPE"]][injuriesFatalitiesDataset[["EVTYPE"]] %in% misspelled[i] ] <- correct[i]
}
```

The data set was grouped by EVTYPE and plotted (see “Results” section for visualization).

```
##group the data by event type and order by highest number of injuries and fatalities, retrieve the top 10 entries
finalInjuriesFatalities<- injuriesFatalitiesDataset %>% group_by(EVTYPE) %>% summarise(InjuriesAndFatalities = sum(Injuries_
Fatalities)) %>% arrange(desc(InjuriesAndFatalities)) %>% filter(InjuriesAndFatalities >=1554 )
```

Unit value designators from the PROPDMGEXP and CROPDMGEXP variables were used to index the data set.

```
##finds unit value designators for damage cost
findDesignators <- function(x) {
    return(stormDataset[[x]] %in% c("K", "m", "M", "B", "k"))
}

##get index positions of CROPDMGEXP values that are complete unit designators
indexPositions <- findDesignators("CROPDMGEXP")

##get index positions of PROPDMGEXP values that are complete unit designators
indexPositions2 <- findDesignators("PROPDMGEXP")

##get CROPDMGEXP values that are to be included in data analysis
##get PROPDMGEXP values that are to be included in data analysis
damageDataset <- stormDataset[indexPositions & indexPositions2, ]
```

The associated EVTYPE was changed to all upper case to faciliate accurate grouping.

```
##convert all values for EVTYPE variable to upper case to faciliate spelling search
damageDataset$EVTYPE <- toupper(damageDataset$EVTYPE)
```

Exploratory analysis of the misspelled words in the EVTYPE variable for property and crops analysis

```
##exploratory analysis of the misspelled words in the EVTYPE variable for property and crops analysis
table(damageDataset$EVTYPE)
```

```
##
##        ASTRONOMICAL HIGH TIDE         ASTRONOMICAL LOW TIDE
##                             1                           174
##                     AVALANCHE                     BLIZZARD
##                           154                          1742
##                 COASTAL FLOOD              COASTAL FLOODING
##                           522                             2
##               COLD AIR TORNADO               COLD/WIND CHILL
##                             1                           520
##                      DENSE FOG                  DENSE SMOKE
##                           920                             9
##                       DROUGHT                DRY MICROBURST
##                          1400                             4
##                    DUST DEVIL                    DUST STORM
##                            52                           201
##         DUST STORM/HIGH WINDS                EXCESSIVE HEAT
##                             1                           712
##                  EXTREME COLD       EXTREME COLD/WIND CHILL
##                            89                           718
##                   FLASH FLOOD             FLASH FLOOD/FLOOD
##                         21622                            11
##                FLASH FLOODING          FLASH FLOODING/FLOOD
##                            60                             8
##                         FLOOD             FLOOD/FLASH FLOOD
##                         13548                            14
##                      FLOODING                       FLOODS
##                            25                             1
##                           FOG                  FOREST FIRES
##                             1                             1
##                        FREEZE                  FREEZING FOG
##                             1                            39
##                  FROST/FREEZE                  FUNNEL CLOUD
##                          1023                          2382
##                     GLAZE ICE                      GUSTNADO
##                             2                             3
##                   GUSTY WINDS                          HAIL
##                             1                         79965
##                      HAIL 100                     HAIL/WIND
##                             1                             1
##                     HAIL/WINDS                          HEAT
```

```
##                          1                          626
##                  HEAT WAVE             HEAT WAVE DROUGHT
##                          1                            1
##                 HEAVY RAIN          HEAVY RAIN/HIGH SURF
##                       5258                            1
##                HEAVY RAINS          HEAVY RAINS/FLOODING
##                          4                            5
##                HEAVY SNOW HEAVY SNOW/HIGH WINDS & FLOOD
##                       6014                            1
##         HEAVY SURF/HIGH SURF                    HIGH SURF
##                          4                          401
##                  HIGH WIND                   HIGH WINDS
##                      11494                           11
##        HIGH WINDS HEAVY RAINS               HIGH WINDS/COLD
##                          1                            5
##                  HURRICANE               HURRICANE ERIN
##                         68                            3
##              HURRICANE FELIX               HURRICANE OPAL
##                          1                            3
##     HURRICANE OPAL/HIGH WINDS             HURRICANE/TYPHOON
##                          1                           33
##             ICE JAM FLOODING                    ICE STORM
##                          1                          960
##                  ICY ROADS             LAKE-EFFECT SNOW
##                          4                          612
##              LAKESHORE FLOOD                   LANDSLIDE
##                         23                          320
##                  LIGHTNING                  MARINE HAIL
##                       4168                          208
##             MARINE HIGH WIND           MARINE STRONG WIND
##                        123                           46
##      MARINE THUNDERSTORM WIND                  RIP CURRENT
##                       5812                          277
##                 RIVER FLOOD               RIVER FLOODING
##                         16                            4
##                      SEICHE     SEVERE THUNDERSTORM WINDS
##                          7                            2
##          SEVERE THUNDERSTORMS                        SLEET
##                          1                           54
##                   SMALL HAIL                         SNOW
##                          1                            1
```

```
##                   STORM SURGE              STORM SURGE/TIDE
##                             7                           136
##                   STRONG WIND            THUDERSTORM WINDS
##                          2572                             1
##             THUNDERSTORM HAIL            THUNDERSTORM WIND
##                             1                         81417
##           THUNDERSTORM WINDS       THUNDERSTORM WINDS HAIL
##                           816                             3
##  THUNDERSTORM WINDS LIGHTNING     THUNDERSTORM WINDS/ FLOOD
##                             1                             1
##       THUNDERSTORM WINDS/HAIL          THUNDERSTORM WINDSS
##                             2                            11
##                THUNDERSTORMS           THUNDERSTORMS WIND
##                             1                             1
##           THUNDERSTORMS WINDS                      TORNADO
##                             1                          9378
##                   TORNADO F0    TORNADOES, TSTM WIND, HAIL
##                             2                             1
##           TROPICAL DEPRESSION               TROPICAL STORM
##                            27                           417
##           TROPICAL STORM DEAN        TROPICAL STORM GORDON
##                             1                             1
##          TROPICAL STORM JERRY                    TSTM WIND
##                             2                          5613
##                TSTM WIND/HAIL                    TSUNAMI
##                            70                            19
##                       TYPHOON                 URBAN FLOOD
##                             5                            58
##                URBAN FLOODING         URBAN/SML STREAM FLD
##                            49                           161
##             VOLCANIC ASHFALL                   WATERSPOUT
##                             3                          1060
##             WILD/FOREST FIRE            WILD/FOREST FIRES
##                            29                             1
##                     WILDFIRE                   WILDFIRES
##                          1783                             1
##                 WIND DAMAGE                        WINDS
##                             1                             1
##                WINTER STORM       WINTER STORM HIGH WINDS
##                          6714                             1
```

```
##                WINTER STORMS               WINTER WEATHER
##                            1                         6659
```

The weather event type variable (EVTYPE) was cleaned to correct obvious errors in spelling. Vectors of misspelled words were matched with correctly spelled words and reassigned to the found values. The decision to further clean the EVTYPE variable so that all values are recoded to codebook designated values was not carried out because some of the weather values reported may be observations that don't fit the codebook designated values; better to keep original values that show the reporter's observations than incorrectly recode values or mislabel values and miss crucial observations.

```
##use correctTheWords to correct EVTYPES misspellings
for(i in 1:length(misspelled)) {
    damageDataset[["EVTYPE"]][damageDataset[["EVTYPE"]] %in% misspelled[i] ] <- correct[i]
}
```

A list was created to hold the unit value designators to search on and their associated base 10 value to multiply by for the found PROPDMG and CROPDMG values; these values were computed and assigned to the data set.

```
##holds variable values for converting *DMG variables to base 10 using corresponding designator from *DMGEXP variable
my_list <-
    list(
        list(
            c("K", "k"),
            1000,
            list("CROPDMGEXP", "PROPDMGEXP"),
            list("CROPDMG", "PROPDMG")
        ),
        list(
            c("B", "b"),
            1000000000,
            list("CROPDMGEXP", "PROPDMGEXP"),
            list("CROPDMG", "PROPDMG")
        ),

        list(
            c("M", "m"),
            1000000,
            list("CROPDMGEXP", "PROPDMGEXP"),
            list("CROPDMG", "PROPDMG")
        )

    )


##convert *DMG variables to base 10 using corresponding designator from *DMGEXP variable
for(i in 1:length(my_list)) {
    for (j in 1:length(my_list[[1]][[3]])) {
        damageDataset[damageDataset[[my_list[[i]][[3]][[j]]]] %in% c(my_list[[i]][[1]]), my_list[[i]][[4]][[j]]] <-
            damageDataset[damageDataset[[my_list[[i]][[3]][[j]]]] %in% c(my_list[[i]][[1]]), my_list[[i]][[4]][[j]]] * my_li
st[[i]][[2]]
    }
}
```

Additionally PROPDMG and CROPDMG were combined into a new column that is sum of damage for each this was done to make it easier to graph total of property and crops damages.

```
##create new column to add the values together
damageDataset <- mutate(damageDataset, Crops_Property_Damage = PROPDMG + CROPDMG)
```

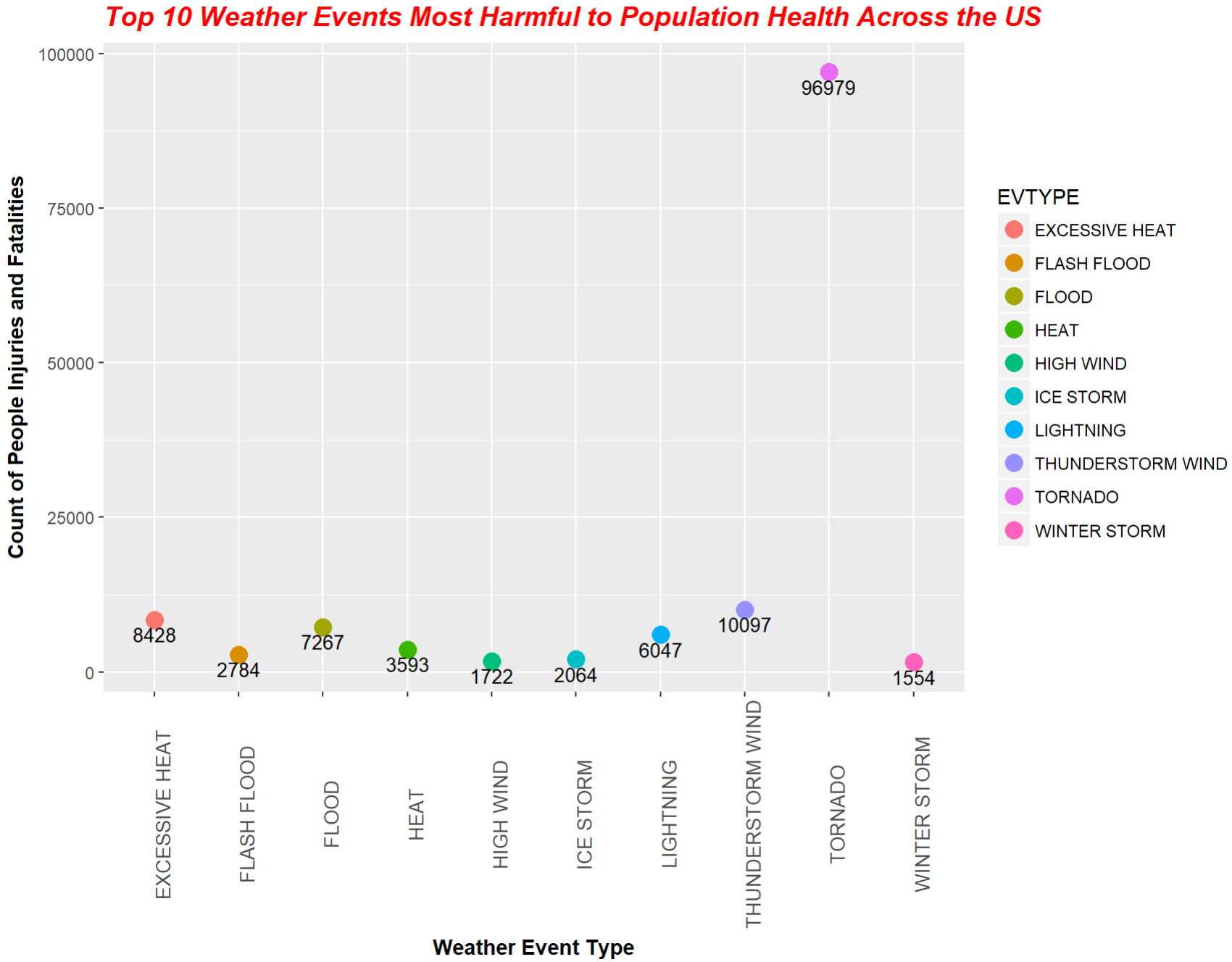
The data set was grouped by EVTYPE, formatted for easier display in the visualization, and plotted (see “Results” section for visualization).

```
##group/filter as did with previous question
damageDataset<- damageDataset %>% group_by(EVTYPE) %>% summarise(Damage_Total = sum(Crops_Property_Damage)) %>% arrange(desc
(Damage_Total)) %>% filter(Damage_Total >=4641493000 )


##format currency for more legibility
damageDataset["Damage_Total"] <- damageDataset["Damage_Total"] / 1000000000
damageDataset$Damage_Total <- currency(damageDataset$Damage_Total)
```

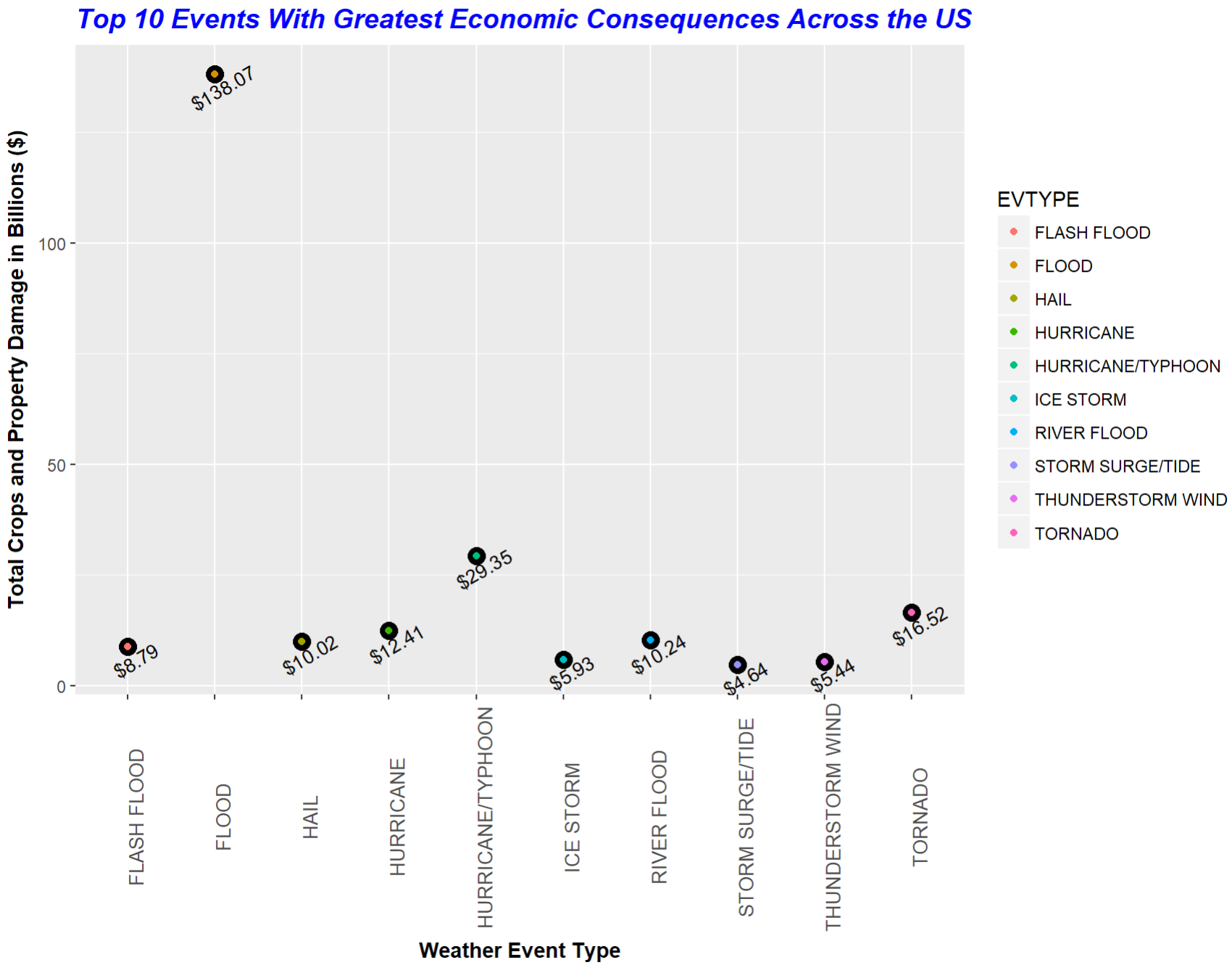
Results

```
##graph the top 10 entries
ggplot(finalInjuriesFatalities, aes(x = EVTYPE, y = InjuriesAndFatalities)) + geom_point(aes(color = EVTYPE), size = 4) +
    labs(x = "Weather Event Type", y = "Count of People Injuries and Fatalities", subtitle = "1950 - 2011")  +
    geom_text(aes(label = InjuriesAndFatalities, vjust = 1.45),  size = 3.5) +
    ggtitle("Top 10 Weather Events Most Harmful to Population Health Across the US" ) +
    theme(plot.title = element_text(color="red", size=14, face="bold.italic"), axis.title.x = element_text(face = "bold"),ax
is.title.y = element_text(face = "bold"), axis.text.x = element_text(size = 10, angle = 90))
```



Top 10 weather event types and the associated count of people who sustained injuries or fatalities across the United States.

```
##graph the top 10 entries
ggplot(damageDataset, aes(x = EVTYPE, y = Damage_Total)) +
labs(x = "Weather Event Type", y = "Total Crops and Property Damage in Billions ($)", subtitle = "1950 - 2011")  +
geom_text(aes(label = Damage_Total, vjust = 1.46), size = 3.5, angle = 30) + ggtitle("Top 10 Events With Greatest Economic C
onsequences Across the US") +
theme(plot.title = element_text(color="blue", size=14, face="bold.italic"), axis.title.x = element_text(face = "bold"), axi
s.title.y = element_text(face = "bold"), axis.text.x = element_text(size = 10, angle = 90)) + geom_point(size = 4) +
geom_point(aes(color = EVTYPE))
```



Top 10 weather event types and the associated economic damage via property and crops damage across the United States.

Considerations

Some considerations to improve analysis. Moving data reporting to centralized cloud computing database would save client money, allow NOAA access to data in realtime, and apply constraints on database fields to facilitate cleaner data which will improve data analysis performance on the client end.