Data Analysis on Severe Weather events in the USA

Cédric HUMBERT

Friday, October 24, 2014

This report shows uses data from the NationaL Weather Service and National Climatic Data Center Storm Events in order to study the impacts of severe weather condition regarding the health of human population and economic consequences. In order to adress theses problems we are going to simplify the data set, acting only regarding from year by begining date and relevant data, we did not make any consideration regarding the location which might have been usefull in order to adress a specific city. To start this we imported the data :

```
## Data Processing :
RawData <- read.csv("repdata-data-StormData.csv")
RawData$EVTYPE = as.character(RawData$EVTYPE)
RawData$BGN_DATE <- as.Date(as.character(RawData$BGN_DATE), "%m/%d/%Y")
Data <- RawData[,c("BGN_DATE", "EVTYPE", "MAG", "FATALITIES","INJURIES",  "PROPDMGEXP", "CROPDMGEXP" )]
## Furthermore I have normalized the data in the following way, giving growing integer numbers to the facto
r :
Data$PROPDMGEXP <- as.numeric(Data$PROPDMGEXP)
Data$CROPDMGEXP <- as.numeric(Data$CROPDMGEXP)
```

Regarding the danger toward population health we started by looking for “human health” in the document “Storm Data Documentation” and extracted the corresponding daa from the Raw Data.

We are going to look at the event that have more that average of Injuries and Fatalities.

```
### Results : on Human Health
## Let us compute critera of our analysis :
Mean_Inj <- mean(Data$INJURIES)
Mean_Fat <- mean(Data$FATALITIES)
M_Mean_Pop_Risk <- which(Data$INJURIES > Mean_Inj & Data$FATALITIES > Mean_Fat)
## Risk Event Data Set is :
Risk_Event_Data <- Data[M_Mean_Pop_Risk,]

##The 15 most likely to happen which are also the more dangerous than average
Count_By_Type <- tapply(Risk_Event_Data$EVTYPE, Risk_Event_Data$EVTYPE, length)
Rank <- sort(Count_By_Type, decreasing = TRUE)
head(Rank, 15)
```

```
##          TORNADO         LIGHTNING         TSTM WIND       FLASH FLOOD
##             1378               263               148                85
##        HIGH WIND    EXCESSIVE HEAT      WINTER STORM       RIP CURRENT
##               69                64                52                47
##            FLOOD         AVALANCHE THUNDERSTORM WIND        HEAVY SNOW
##               44                42                40                32
##      RIP CURRENTS        ICE STORM               FOG
##               29                24                23
```

```
## Weight of differents components in Population health : \n 1) Injuries
Weight_Inj_By_Type <- tapply(Risk_Event_Data$INJURIES, Risk_Event_Data$EVTYPE, sum)
Weight_Inj_By_Type <- sort(Weight_Inj_By_Type, decreasing = TRUE)
head(Weight_Inj_By_Type, 15)
```

```
##           TORNADO    EXCESSIVE HEAT             FLOOD         ICE STORM
##             60187              4791              2679              1720
##              HEAT HURRICANE/TYPHOON          BLIZZARD         LIGHTNING
##              1420              1219               718               649
##         TSTM WIND       FLASH FLOOD      WINTER STORM               FOG
##               646               641               599               308
##         HIGH WIND    TROPICAL STORM         HEAT WAVE
##               308               274               269
```

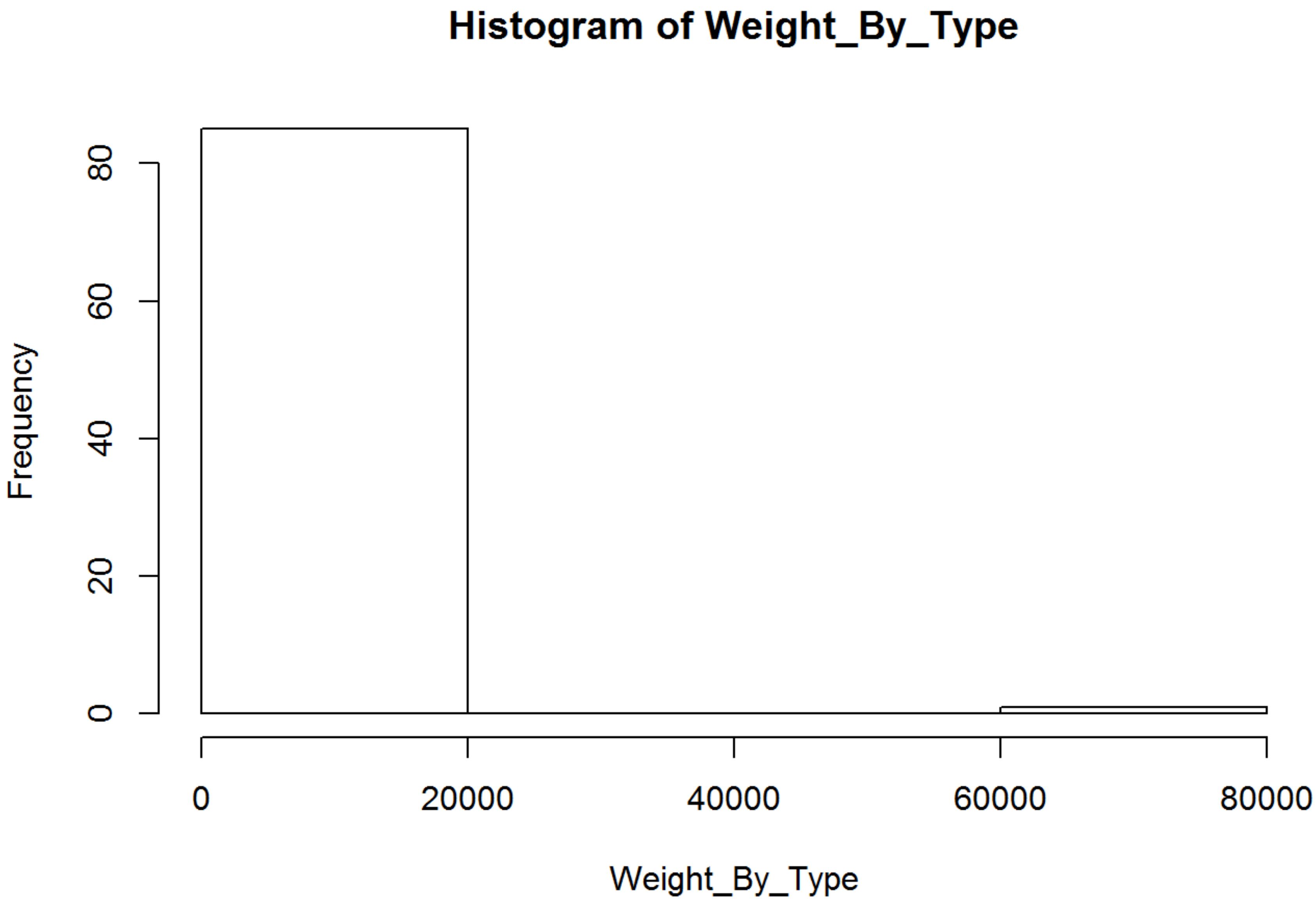
```
## 2) Fatalities
Weight_Fat_By_Type <- tapply(Risk_Event_Data$FATALITIES, Risk_Event_Data$EVTYPE, sum)
Weight_Fat_By_Type <- sort(Weight_Fat_By_Type, decreasing = TRUE)
head(Weight_Fat_By_Type, 15)
```

```
##           TORNADO    EXCESSIVE HEAT         LIGHTNING         TSTM WIND
##              5227               402               283               199
##       FLASH FLOOD             FLOOD         HIGH WIND      WINTER STORM
##               171               104               102                85
##              HEAT          WILDFIRE THUNDERSTORM WIND         AVALANCHE
##                73                55                54                52
##        HEAVY SNOW       RIP CURRENT          BLIZZARD
##                51                50                48
```

```
## 3) Sum of the Two :
Weight_By_Type <- Weight_Inj_By_Type + Weight_Fat_By_Type
head(Weight_By_Type, 15)
```

```
##          TORNADO    EXCESSIVE HEAT            FLOOD        ICE STORM
##            65414              5193             2962             1919
##             HEAT HURRICANE/TYPHOON         BLIZZARD        LIGHTNING
##             1591              1323              820              734
##        TSTM WIND       FLASH FLOOD     WINTER STORM              FOG
##              719               696              653              360
##        HIGH WIND    TROPICAL STORM        HEAT WAVE
##              359               324              317
```

```
## 4) Lets us revue the proportion of damages caused by thoses envents :
hist(Weight_By_Type, breaks = 3)
```



We see here that the imense majority of theses envents have a very small impact ont the Population healt side of the problem. Only a few at the far right are causing big problems.

We note that the greatest overall factor of destruction toward human healt are Tornados. Given what we have said before, an aproach to the problem could be to focus emergency ressources on the top 4 or 5 most harmfull risks.

Let us study the impacts on Econmical critera, we will repeat the same analysis :

```
### Results : on Economic Consequences :
Mean_Cro <- mean(Data$CROPDMGEXP)
Mean_Pro <- mean(Data$PROPDMGEXP)
M_Mean_Ec_Risk <- which(Data$CROPDMGEXP > Mean_Cro & Data$PROPDMGEXP > Mean_Pro)
## Risk Event Data Set is :
Risk_Event_Data <- Data[M_Mean_Ec_Risk,]

##The 15 most likely to happen which are also the more dangerous than average
Count_By_Type <- tapply(Risk_Event_Data$EVTYPE, Risk_Event_Data$EVTYPE, length)
Rank <- sort(Count_By_Type, decreasing = TRUE)
head(Rank, 15)
```

```
##        THUNDERSTORM WIND                     HAIL              FLASH FLOOD
##                    81417                    79967                    21623
##                    FLOOD                HIGH WIND                  TORNADO
##                    13548                    11494                     9382
##             WINTER STORM           WINTER WEATHER               HEAVY SNOW
##                     6714                     6659                     6014
## MARINE THUNDERSTORM WIND                TSTM WIND               HEAVY RAIN
##                     5812                     5613                     5258
##                LIGHTNING              STRONG WIND             FUNNEL CLOUD
##                     4168                     2572                     2382
```

```
## Weight of differents components regarding Economic consequence: \n 1) Crop :
Weight_Cro_By_Type <- tapply(Risk_Event_Data$CROPDMGEXP, Risk_Event_Data$EVTYPE, sum)
```

```r
Weight_Cro_By_Type <- sort(Weight_Cro_By_Type, decreasing = TRUE)
head(Weight_Cro_By_Type, 15)
```

```
##        THUNDERSTORM WIND                     HAIL              FLASH FLOOD 
##                   570059                   560566                   151689 
##                    FLOOD                HIGH WIND                  TORNADO 
##                    95502                    80548                    65824 
##             WINTER STORM           WINTER WEATHER               HEAVY SNOW 
##                    47010                    46615                    42114 
## MARINE THUNDERSTORM WIND                TSTM WIND               HEAVY RAIN 
##                    40684                    39555                    36854 
##                LIGHTNING              STRONG WIND             FUNNEL CLOUD 
##                    29180                    18008                    16674
```

```r
## 2) Property :
Weight_Pro_By_Type <- tapply(Risk_Event_Data$PROPDMGEXP, Risk_Event_Data$EVTYPE, sum)
Weight_Pro_By_Type <- sort(Weight_Pro_By_Type, decreasing = TRUE)
head(Weight_Pro_By_Type, 15)
```

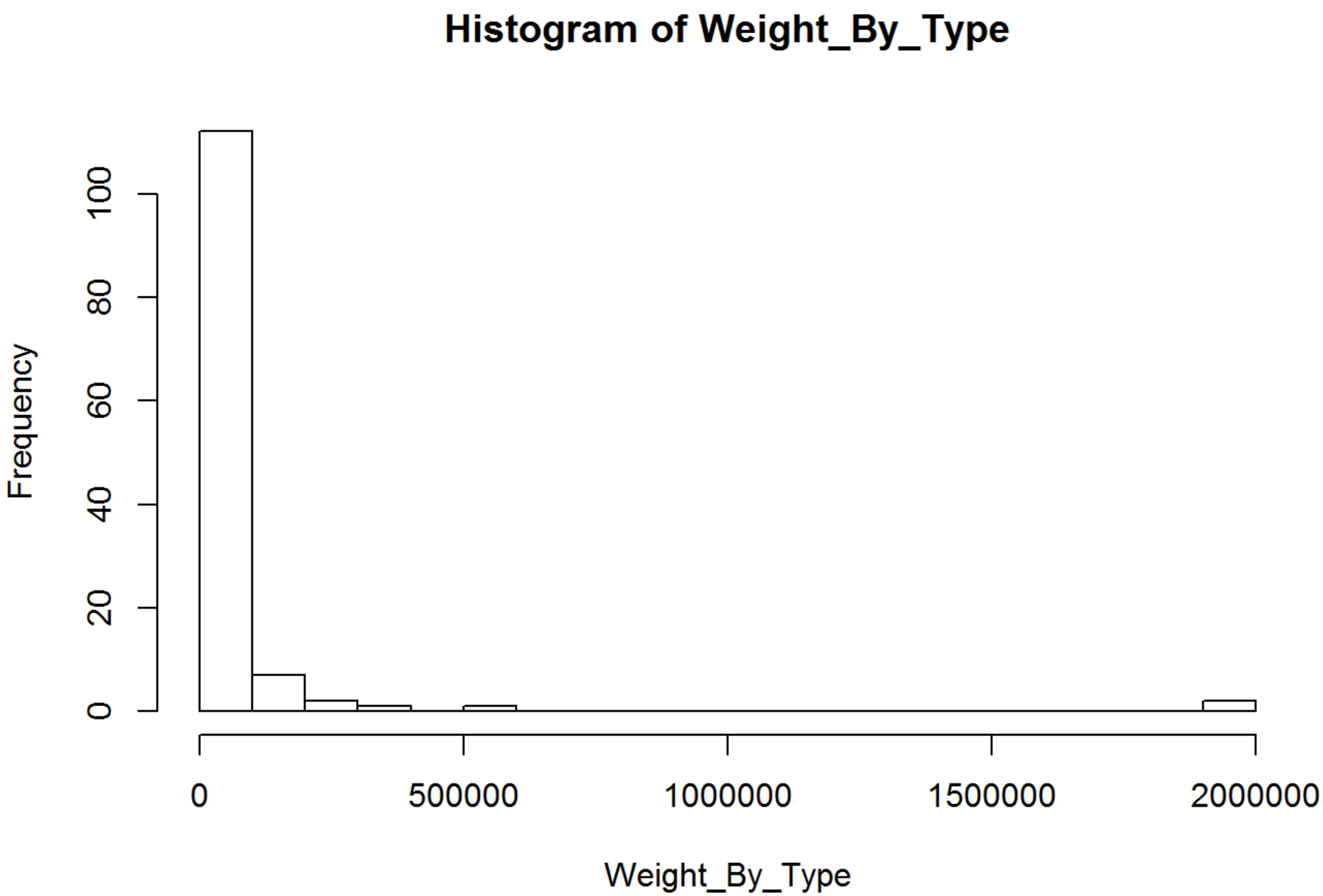
```
##        THUNDERSTORM WIND                     HAIL              FLASH FLOOD 
##                  1384613                  1360169                   368953 
##                    FLOOD                HIGH WIND                  TORNADO 
##                   232122                   195716                   161065 
##             WINTER STORM           WINTER WEATHER               HEAVY SNOW 
##                   114238                   113215                   102284 
## MARINE THUNDERSTORM WIND                TSTM WIND               HEAVY RAIN 
##                    98804                    95599                    89442
```

```
##            LIGHTNING          STRONG WIND         FUNNEL CLOUD
##                70962                43732                40494
```

```r
## 3) Sum of the Two :
Weight_By_Type <- Weight_Cro_By_Type + Weight_Pro_By_Type
head(Weight_By_Type, 15)
```

```
##    THUNDERSTORM WIND                 HAIL          FLASH FLOOD
##              1954672              1920735               520642
##                FLOOD            HIGH WIND              TORNADO
##               327624               276264               226889
##         WINTER STORM       WINTER WEATHER           HEAVY SNOW
##               161248               159830               144398
## MARINE THUNDERSTORM WIND            TSTM WIND           HEAVY RAIN
##               139488               135154               126296
##            LIGHTNING          STRONG WIND         FUNNEL CLOUD
##               100142                61740                57168
```

```r
## 4) Lets us revue the proportion of damages caused by thoses envents :
hist(Weight_By_Type, breaks = 25)
```



We note a very strong correlation bettwen the 2 Economic sets.

However the relation bettwen Econimic

risks and Human health is very thin, there are not many top risks identified here for human health and for the Economie.