USA Storm socio-economic report

Fideborios

Saturday, February 21, 2015

Storms and other severe weather events can cause both public health and economic problems for communities and municipalities. Many severe events can result in fatalities, injuries, and property damage, and preventing such outcomes to the extent possible is a key concern. 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. The data for this assignment are saved in Storm Data data-set.

```
StormData=read.csv("repdata-data-StormData.csv")

dim(StormData)
```

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

As we can see there are 902.297 observations and 37 variables. For the first question "Across the United States, which types of events are most harmful with respect to population health?" we need to find the sum of Injuries and Fatalities indicating the total impact of the natural disasters.

```
Injuries=aggregate(StormData$INJURIES, by=list(StormData$EVTYPE), FUN=sum)
Fatalities=aggregate(StormData$FATALITIES, by=list(StormData$EVTYPE), FUN=sum)
```

Next we will order it with the maximum value first and pick the top 6 natural disasters according public health problems.

```
Order.Injuries=Injuries[order(Injuries[,2],decreasing=TRUE),]
Order.Fatalities=Fatalities[order(Fatalities[,2],decreasing=TRUE),]

HOrder.Injuries= head(Order.Injuries)
HOrder.Fatalities= head(Order.Fatalities)
```

```
HOrder.Injuries
```

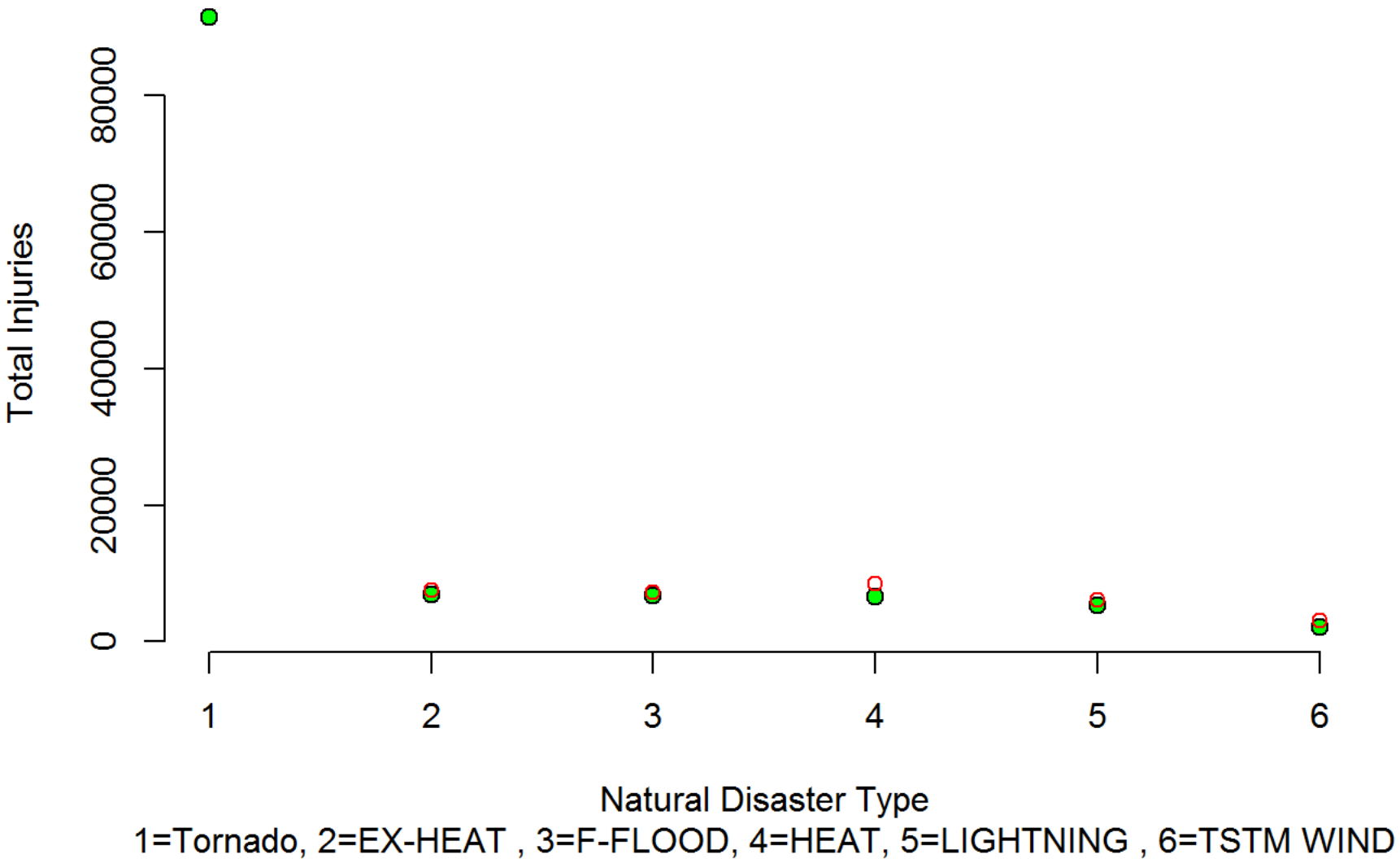
```
##            Group.1     x
## 834        TORNADO 91346
## 856      TSTM WIND  6957
## 170          FLOOD  6789
## 130 EXCESSIVE HEAT  6525
## 464      LIGHTNING  5230
## 275           HEAT  2100
```

```
HOrder.Fatalities
```

```
##            Group.1    x
## 834        TORNADO 5633
## 130 EXCESSIVE HEAT 1903
## 153    FLASH FLOOD  978
## 275           HEAT  937
## 464      LIGHTNING  816
## 856      TSTM WIND  504
```

We observe that the worst Natural Disaster is the Tornado in both Fatalities and Injuries.

```
y= c(5633 , 504 , 470 , 1903, 816, 937)
plot(head(Order.Injuries[,2]),
     xlab = "Natural Disaster Type",
     ylab = "Total Injuries",
     bg = "green",
     cex = 1.1, # size of dots#
     pch = 21,  # type of dots#
     frame = FALSE, # frame or not#
     sub= "1=Tornado, 2=EX-HEAT , 3=F-FLOOD, 4=HEAT, 5=LIGHTNING , 6=TSTM WIND"
     )
    points(head(Order.Injuries[,2]) + y, col="red")
```



We check that graphically that EXCESSIVE HEAT comes second in the Fatalities and Injuries combined.

Now we want to see which Natural Disaster has the most total expenses for USA properties and crops. The values are coded as “K”=1000,“k”=1000, “H”=100 , “h”=100 ,“m”=1000000 ,“M”=1000000, “B”=1000000000 so we are transforming

them to numeric values.

```
library(plyr)
library(knitr)

StormData$PROPDMGEXP=revalue(StormData$PROPDMGEXP, c("K"=1000,"k"=1000, "H"=100 , "h"=100 ,"m"=1000000 ,"M"
=1000000, "B"=1000000000))

StormData$CROPDMGEXP=revalue(StormData$CROPDMGEXP, c("K"=1000,"k"=1000, "H"=100 , "h"=100 ,"m"=1000000 ,"M"
=1000000, "B"=1000000000))
```

Next step is to sum the total expenses by Disaster type, and as before take the top 6 natural disasters.

```
PROP.DMG.EXP=aggregate(as.numeric(StormData$PROPDMGEXP), by=list(StormData$EVTYPE), FUN=sum)
CROP.DMG.EXP=aggregate(as.numeric(StormData$CROPDMGEXP), by=list(StormData$EVTYPE), FUN=sum)

Head.Order.PROP.DMG.EXP=head(PROP.DMG.EXP[order(PROP.DMG.EXP[,2],decreasing=TRUE),])
Head.Order.CROP.DMG.EXP=head(CROP.DMG.EXP[order(PROP.DMG.EXP[,2],decreasing=TRUE),])

Head.Order.PROP.DMG.EXP
```

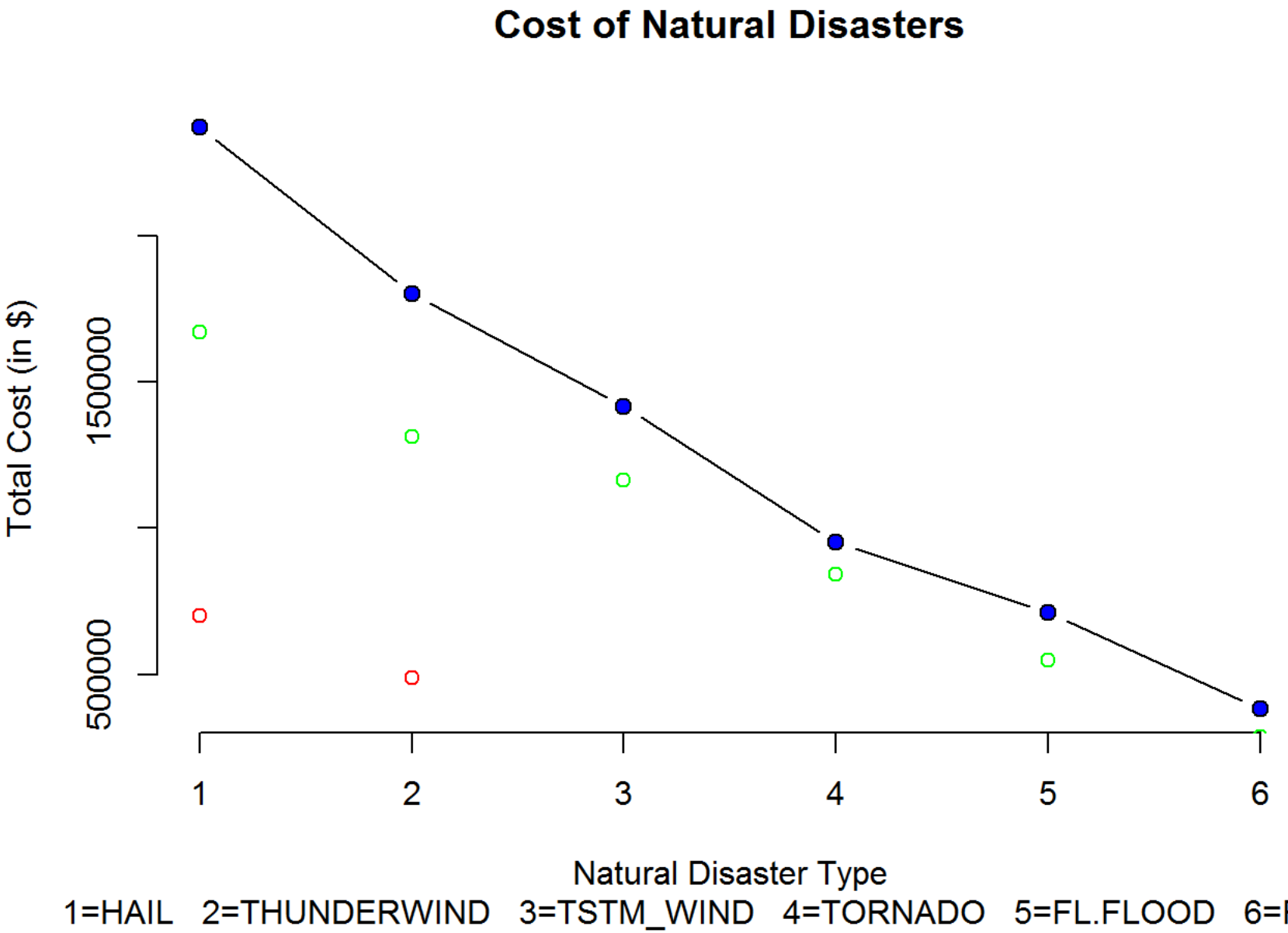
```
##               Group.1       x
## 244              HAIL 1669218
## 760 THUNDERSTORM WIND 1311962
## 856         TSTM WIND 1163141
## 834           TORNADO  842638
## 153       FLASH FLOOD  549998
## 170             FLOOD  287756
```

```
Head.Order.CROP.DMG.EXP
```

```
##               Group.1      x
## 244              HAIL 700757
## 760 THUNDERSTORM WIND 489759
## 856         TSTM WIND 253082
## 834           TORNADO 108711
## 153       FLASH FLOOD 162853
## 170             FLOOD  93786
```

As we can see in both property and corps Hail is the most expensive followed by THUNDERSTORM WIND, TSTM WIND , TORNADO , FLASH FLOOD & FLOOD

```
plot(Head.Order.PROP.DMG.EXP[,2] + Head.Order.CROP.DMG.EXP[,2],
     type="b",
     main = "Cost of Natural Disasters",
     xlab = "Natural Disaster Type",
     ylab = "Total Cost (in $)",
     sub = "1=HAIL   2=THUNDERWIND   3=TSTM_WIND   4=TORNADO   5=FL.FLOOD   6=FLOOD",
     bg = "blue",
     cex = 1.1, # size of dots#
     pch = 21,  # type of dots#
     frame = FALSE, # frame or not#
     )
points(Head.Order.CROP.DMG.EXP[,2], col="red")
points(Head.Order.PROP.DMG.EXP[,2], col="green")
```



We check it

graphically Red = Only Crop Green = Only Property & Blue = Combined Cost and we end up with the same conclusions.