repdata project2

morgan

2/15/2020

##Synopsis 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.The basic goal of this assignment is to explore the NOAA Storm Database and answer some basic questions about severe weather events. ##1.Data loading

##2.Data processing

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

data <- read.csv("repdata\_data\_StormData.csv.bz2")  
data1 <- subset(data, select = c("EVTYPE","FATALITIES","INJURIES","PROPDMG","PROPDMGEXP","CROPDMG","CROPDMGEXP"))##Subset columns relevant to the analysis  
##we define a new variable "harm", equals to total counts of fatalities and injuries  
data1$harm <- data1$FATALITIES+data1$FATALITIES

##3.Answer the question ##Across the United States, which types of events (as indicated in the EVTYPE variable) are most harmful with respect to population health?

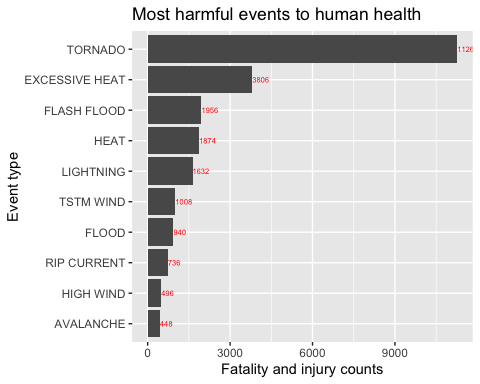
harm\_by\_evtype <- aggregate(data1$harm ~ data1$EVTYPE, data1,sum)  
harm\_by\_evtype <- harm\_by\_evtype[order(harm\_by\_evtype[,2], decreasing = TRUE),]  
##The answer is tornado

##Across the United States, which types of events have the greatest economic consequences?

##we define a new variable "eco", equals to total counts of crop damage and property damage.  
data1$eco <- data1$PROPDMG+data1$CROPDMG  
eco\_by\_evtype <- aggregate(data1$eco ~ data1$EVTYPE, data1,sum)  
eco\_by\_evtype <- eco\_by\_evtype[order(eco\_by\_evtype[,2], decreasing = TRUE),]  
##The answer is tornado

##4.Results ##The most harmful weather events to human health.

library(ggplot2)  
res <- harm\_by\_evtype[1:10,]  
name <- res[1:10,1]  
val <- res[1:10,2]  
res$name <- factor(res$`data1$EVTYPE`, levels = res$`data1$EVTYPE`[order(res$`data1$harm`)])  
  
pic<-ggplot(res, aes(x = name, y = val)) +  
 geom\_bar(stat="identity") +   
 geom\_text(aes(label = val), hjust=0, color="red", size=2) +   
 coord\_flip()+  
 labs(title = "Most harmful events to human health") +  
 labs(y = "Fatality and injury counts", x = "Event type")  
pic

 ##The most harmful weather events to economy.

res2 <- eco\_by\_evtype[1:10,]  
name <- res2[1:10,1]  
val <- res2[1:10,2]  
res2$name <- factor(res2$`data1$EVTYPE`, levels = res2$`data1$EVTYPE`[order(res2$`data1$eco`)])  
pic2<-ggplot(res2, aes(x = name, y = val)) +  
 geom\_bar(stat="identity") +   
 geom\_text(aes(label = val), hjust=0, color="red", size=2)+   
 coord\_flip()+  
 labs(title = "Most harmful events to economy") +  
 labs(y = "Crop and property damage", x = "Event type")  
pic2

