Reproducible Research Course-2

Maulid Hussein Bwabo

2022-06-12

# Weather Events on Personal and Property Damage Using NOAA Storm Database

## Synopsis

1. Downloading data

download.file("https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2",destfile = "repdata\_data\_StormData.csv.bz2")  
StormData = read.csv(bzfile("repdata\_data\_StormData.csv.bz2"))  
summary(StormData)  
dim(StormData)  
tail(StormData)  
head(StormData)

## Data Proccesing

1. Cleaning the data The key constructs that are used for this analysis are *EVTYPE : Type of the event* FATALITIES : Number of fatalities out of the event *INJURIES : It try to figure out injuries from the event* PROPDMG : It measured the number of property damage measured *CROPDMG : It measured the number crop damage* PROPDMGEXP : It measured the Property damage exponent (Mns, Bns etc) \*CROPDMGEXP : It measured the crop damage exponent (Mns, Bns etc) The cleaning data has been extended by assigning the numerical values to exponential power

StormData$PROPDMGEXP<- toupper(StormData$PROPDMGEXP)  
StormData$PROPDMGEXP[StormData$PROPDMGEXP %in% c("", "+", "-", "?")] <- "0"  
StormData$PROPDMGEXP[StormData$PROPDMGEXP %in% c("B")] <- "9"  
StormData$PROPDMGEXP[StormData$PROPDMGEXP %in% c("M")] <- "6"  
StormData$PROPDMGEXP[StormData$PROPDMGEXP %in% c("K")] <- "3"  
StormData$PROPDMGEXP[StormData$PROPDMGEXP %in% c("H")] <- "2"  
  
StormData$CROPDMGEXP <- toupper(StormData$CROPDMGEXP)  
StormData$CROPDMGEXP[StormData$CROPDMGEXP %in% c("", "+", "-", "?")] <- "0"  
StormData$CROPDMGEXP[StormData$CROPDMGEXP %in% c("B")] <- "9"  
StormData$CROPDMGEXP[StormData$CROPDMGEXP %in% c("M")] <- "6"  
StormData$CROPDMGEXP[StormData$CROPDMGEXP %in% c("K")] <- "3"  
StormData$CROPDMGEXP[StormData$CROPDMGEXP %in% c("H")] <- "2"

1. Developing new Variable

StormData$PROPDMGTOTAL <- StormData$PROPDMG \* (10 ^ as.numeric(StormData$PROPDMGEXP))  
StormData$CROPDMGTOTAL <- StormData$CROPDMG \* (10 ^ as.numeric(StormData$CROPDMGEXP))  
StormData$DMGTOTAL <- StormData$PROPDMGTOTAL + StormData$CROPDMGTOTAL  
##Analyzing the critical events   
SummStormData<- StormData %>%  
 group\_by(EVTYPE) %>%  
 summarize(SUMFATALITIES = sum(FATALITIES),  
 SUMINJURIES = sum(INJURIES),  
 SUMPROPDMG = sum(PROPDMGTOTAL),  
 SUMCROPDMG = sum(CROPDMGTOTAL),  
 TOTALDMG = sum(DMGTOTAL))

head(SummStormData)  
tail(SummStormData)

##Analysis for critical events Total fatalities, injuries and economic damage are summarized for each weather event 1. Events with the Most Fatalities

SummStormDataFatality <- arrange(SummStormData, desc(SUMFATALITIES))  
FatalityData <- head(SummStormDataFatality)  
summary(SummStormData)

1. Events caused more injuries

SummStormDataInjury <- arrange(SummStormData, desc(SUMINJURIES))  
InjuryData <- head(SummStormDataInjury)

1. Events that caused damaged

SummStormDataDamage <- arrange(SummStormData, desc(TOTALDMG))  
DamageData <- head(SummStormDataDamage)

##Results 1. Fatalities

FatalityData$EVTYPE <- with(FatalityData, reorder(EVTYPE, -SUMFATALITIES))  
ggplot(FatalityData, aes(EVTYPE, SUMFATALITIES, label = SUMFATALITIES)) +  
 geom\_bar(stat = "identity") +  
 geom\_text(nudge\_y = 200) +  
 xlab("Event Type") +  
 ylab("Total Fatalities") +  
 ggtitle("Most Fatal Events") +  
 theme(plot.title = element\_text(hjust = 0.5))

1. Injuries

InjuryData$EVTYPE <- with(InjuryData, reorder(EVTYPE, -SUMINJURIES))  
ggplot(InjuryData, aes(EVTYPE, SUMINJURIES, label = SUMINJURIES)) +  
 geom\_bar(stat = "identity") +  
 geom\_text(nudge\_y = 3000) +  
 xlab("Event Type") +  
 ylab("Total Injuries") +  
 ggtitle("Most Injury Events") +  
 theme(plot.title = element\_text(hjust = 0.5))

1. Damaged

DamageData$EVTYPE <- with(DamageData, reorder(EVTYPE, -TOTALDMG))  
DamageDataLong <- DamageData %>%  
 gather(key = "Type", value = "TOTALDAMAGE", c("SUMPROPDMG", "SUMCROPDMG")) %>%  
 select(EVTYPE, Type, TOTALDAMAGE)  
DamageDataLong$Type[DamageDataLong$Type %in% c("SUMPROPDMG")] <- "Property damage"  
DamageDataLong$Type[DamageDataLong$Type %in% c("SUMCROPDMG")] <- "Crop damage"  
  
ggplot(DamageDataLong, aes(x = EVTYPE, y = TOTALDAMAGE, fill = Type)) +  
 geom\_bar(stat = "identity", position = "stack") +  
 xlab("Event Type") +  
 ylab("Total Damage") +  
 ggtitle("Events with Most Damage") +  
 theme(plot.title = element\_text(hjust = 0.5), legend.position = "bottom")