

# 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)
```

### Data Processing

#### 2. Cleaning the data

```
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"
```

#### 3. 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 eventts
SummStormData <- StormData %>%
  group_by(EVTYPE) %>%
  summarize(SUMFATALITIES = sum(FATALITIES),
```

```
SUMINJURIES = sum(INJURIES),
SUMPROPDMG = sum(PROPDMGTOTAL),
SUMCROPDMG = sum(CROPDMGTOTAL),
TOTALDMG = sum(DMGTOTAL))
```

```
head(SummStormData)
```

##Analysis for critical events 1. Events with the Most Fatalities

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

2. Events caused more injuries

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

3. 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))
```

2. 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))
```

3. 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")

```