Programming assignment 2: reproducible research. Effects of major storms and weather events on human population and economy.

David García Sabaté

3/5/2019

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse)
library(lubridate)
```

Programming assignment 2: reproducible research. Effects of major storms and weather events on human population and economy.

Synopsis

By far, tornadoes are the type of natural catastrophe that has the greatest impact on human health, having added almost 6000 victims and 90000 injured in the period analyzed. However, in economic terms, floods are the most expensive natural phenomenon.

Data Processing

Downloading the "repdata_data_StormData.csv" dataset

```
if (!exists("dataframe")){
  zipUrl <-
"https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2
"
  zipFile <- "repdata_data_StormData.csv.bz2"
  download.file(zipUrl, zipFile, mode = "wb")
}
```

Now we have tib_observations tibble object, and see a sample of 10 observations

```
if (!exists("dataframe")){
  dataframe <- read.csv("repdata_data_StormData.csv.bz2")
}
tib_observations <- as_tibble(dataframe)
head(tib_observations, 10)

## # A tibble: 10 x 37
##    STATE__ BGN_DATE BGN_TIME TIME_ZONE COUNTY COUNTYNAME STATE EVTYPE
```

```
##      <dbl> <fct>    <fct>    <fct>      <dbl> <fct>      <fct> <fct>
##  1       1 4/18/19… 0130     CST           97 MOBILE     AL    TORNA…
##  2       1 4/18/19… 0145     CST            3 BALDWIN    AL    TORNA…
##  3       1 2/20/19… 1600     CST           57 FAYETTE    AL    TORNA…
##  4       1 6/8/195… 0900     CST           89 MADISON    AL    TORNA…
##  5       1 11/15/1… 1500     CST           43 CULLMAN    AL    TORNA…
##  6       1 11/15/1… 2000     CST           77 LAUDERDALE AL    TORNA…
##  7       1 11/16/1… 0100     CST            9 BLOUNT     AL    TORNA…
##  8       1 1/22/19… 0900     CST          123 TALLAPOOSA AL    TORNA…
##  9       1 2/13/19… 2000     CST          125 TUSCALOOSA AL    TORNA…
## 10       1 2/13/19… 2000     CST           57 FAYETTE    AL    TORNA…
## # … with 29 more variables: BGN_RANGE <dbl>, BGN_AZI <fct>,
## #   BGN_LOCATI <fct>, END_DATE <fct>, END_TIME <fct>, COUNTY_END
<dbl>,
## #   COUNTYENDN <lgl>, END_RANGE <dbl>, END_AZI <fct>, END_LOCATI
<fct>,
## #   LENGTH <dbl>, WIDTH <dbl>, F <int>, MAG <dbl>, FATALITIES <dbl>,
## #   INJURIES <dbl>, PROPDMG <dbl>, PROPDMGEXP <fct>, CROPDMG <dbl>,
## #   CROPDMGEXP <fct>, WFO <fct>, STATEOFFIC <fct>, ZONENAMES <fct>,
## #   LATITUDE <dbl>, LONGITUDE <dbl>, LATITUDE_E <dbl>, LONGITUDE_
<dbl>,
## #   REMARKS <fct>, REFNUM <dbl>
```

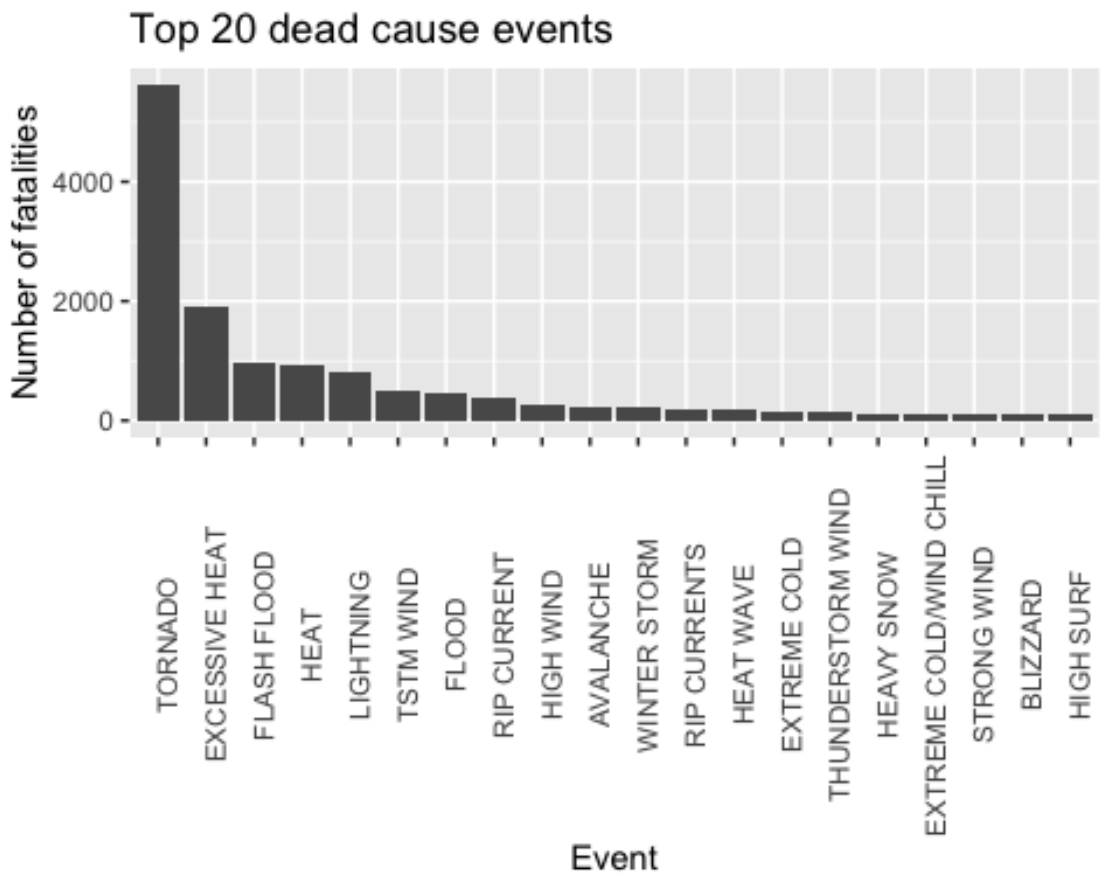
RESULTS

Which types of events are most harmful to population health?

To answer this question we will measure Fatalities and Injuries. We will group the sum of fatalities and injuries by type of event.

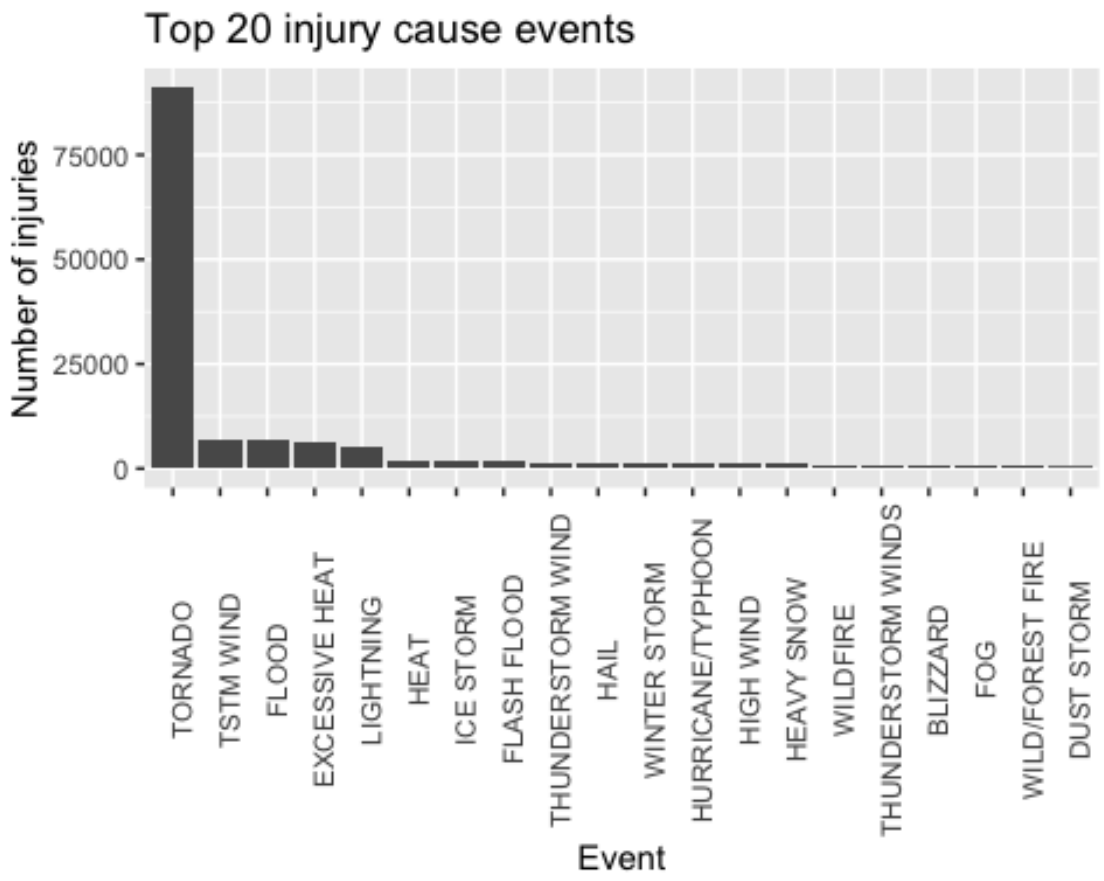
```
tib_observations_fatal <- tib_observations %>%
  group_by(EVTYPE) %>%
  summarize(fatalities = sum(FATALITIES)) %>%
  arrange(desc(fatalities)) %>%
  head(20)

ggplot(tib_observations_fatal, mapping = aes(reorder(EVTYPE, -
fatalities), fatalities)) +
  geom_bar(stat = "identity") +
  theme(axis.text.x = element_text(angle = 90)) +
   labs(
    title = "Top 20 dead cause events",
    x = "Event",
    y = "Number of fatalities"
  )
```




```
tib_observations_injuries <- tib_observations %>%
  group_by(EVTYPE) %>%
  summarize(injuries = sum(INJURIES)) %>%
  arrange(desc(injuries)) %>%
  head(20)

ggplot(tib_observations_injuries, mapping = aes(reorder(EVTYPE, -
injuries), injuries)) +
  geom_bar(stat = "identity") +
  theme(axis.text.x = element_text(angle = 90)) +
   labs(
    title = "Top 20 injury cause events",
    x = "Event",
    y = "Number of injuries"
  )
```

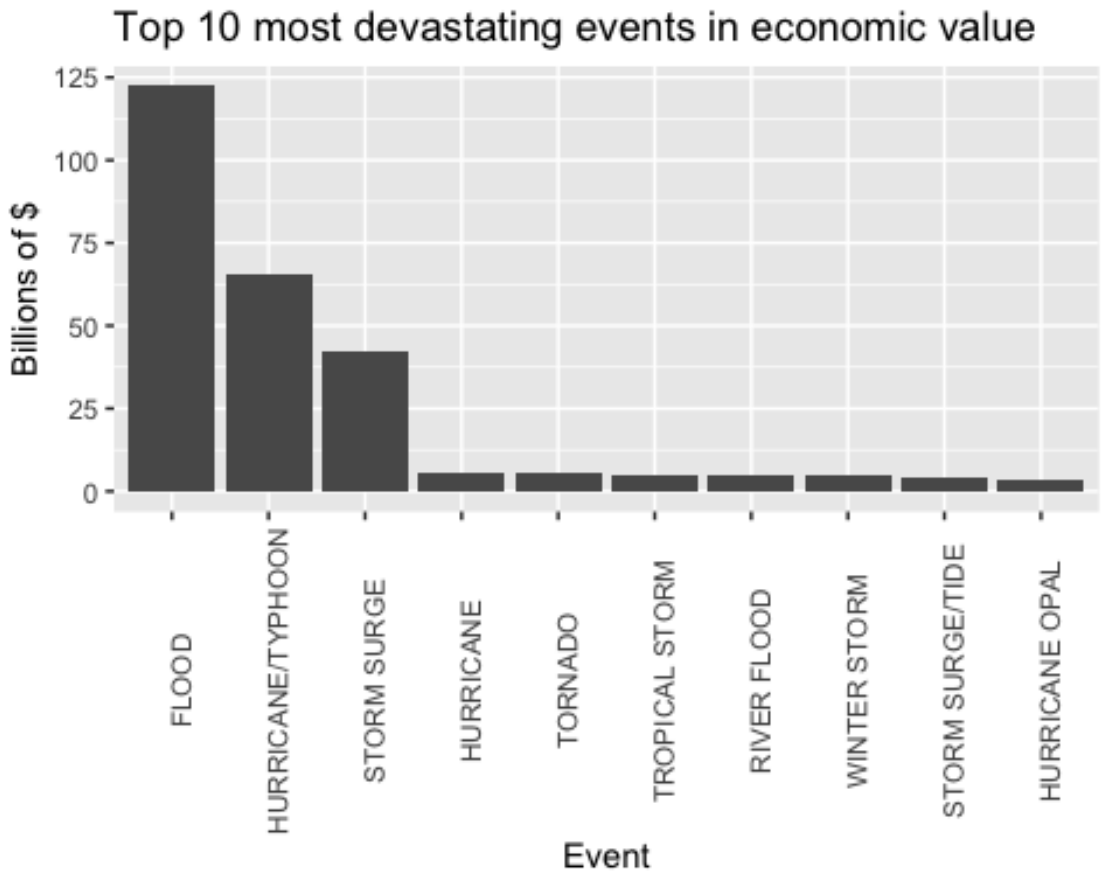


Which types of events have the greatest economic consequences?

We are only taking into account figures whose PROPDMGEXP variable value is B (Billions). These are the most expensive.

```
tib_observations_economic <- tib_observations %>%
  filter(PROPDMGEXP == "B") %>%
  group_by(EVTYPE) %>%
  summarize(total_amount = sum(PROPDMG)) %>%
  arrange(desc(total_amount)) %>%
  head(10)

ggplot(tib_observations_economic, mapping = aes(reorder(EVTYPE, -
total_amount), total_amount)) +
  geom_bar(stat = "identity") +
  theme(axis.text.x = element_text(angle = 90)) +
   labs(
    title = "Top 10 most devastating events in economic value",
    x = "Event",
    y = "Billions of $"
  )
```


Top 10 most devastating events in economic value
Billions of $
125
100
75
50
25
0
FLOOD
HURRICANE/TYPHOON
STORM SURGE
HURRICANE
TORNADO
TROPICAL STORM
RIVER FLOOD
WINTER STORM
STORM SURGE/TIDE
HURRICANE OPAL
Event