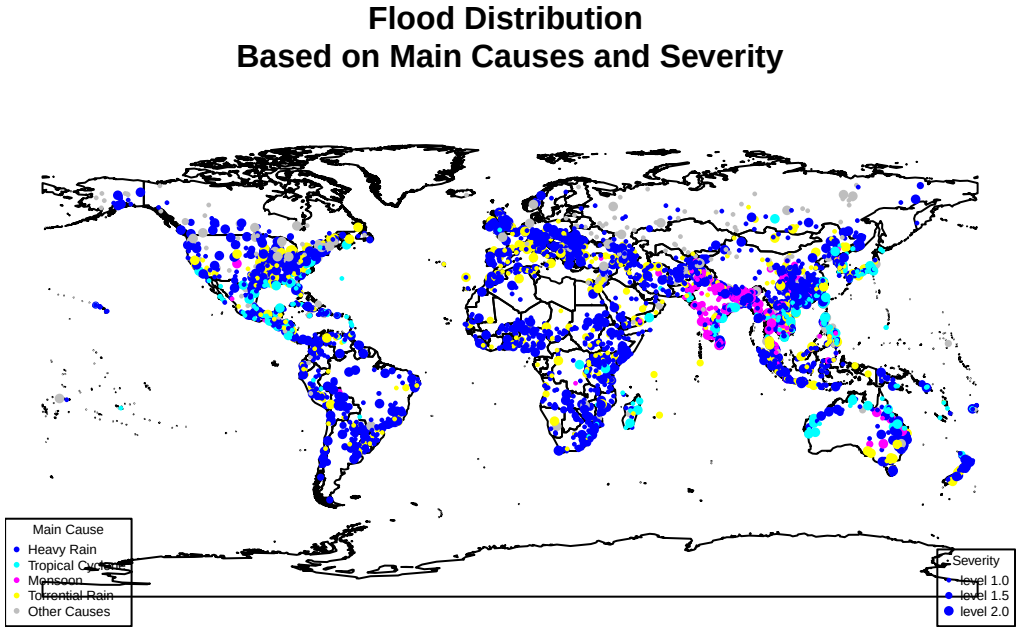
proj2_tt_v1

Tian Tan

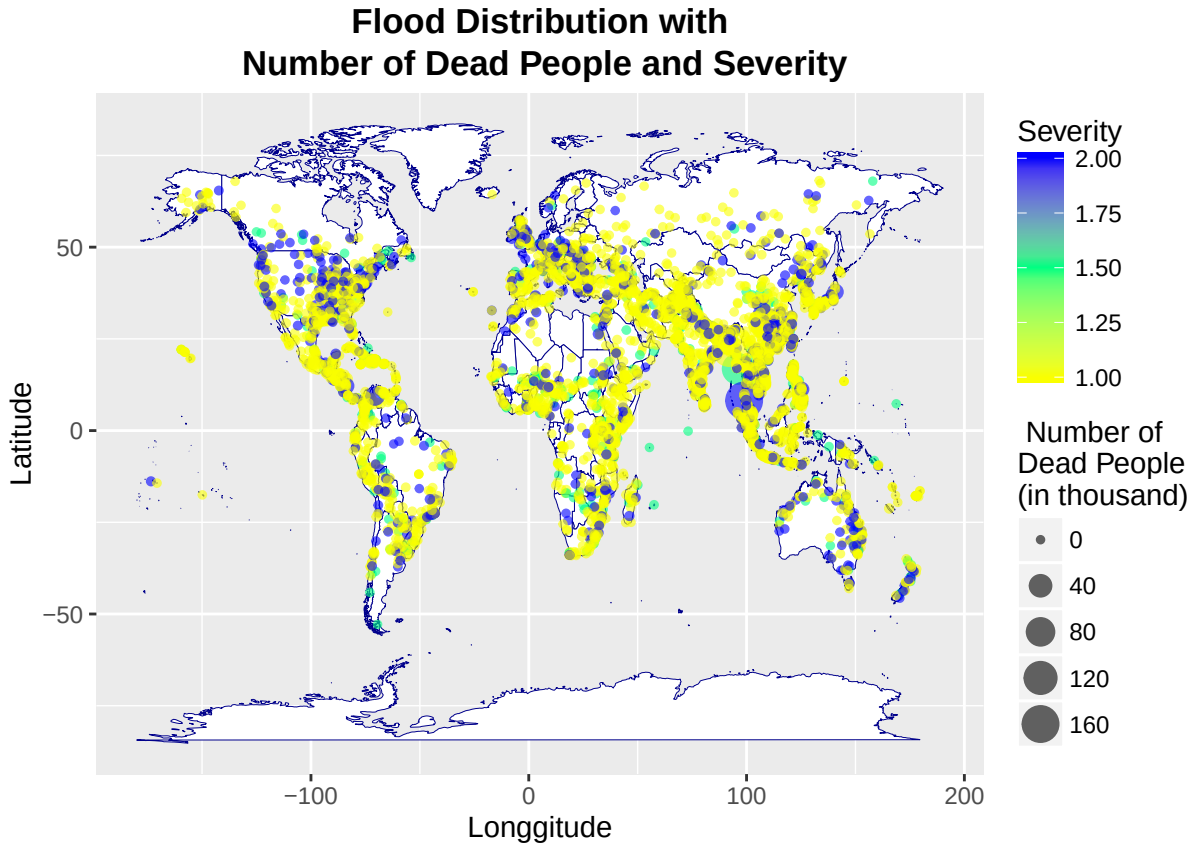
February 25, 2016

```
setwd("/Users/MaxTan/Documents/CU_16spring/EDAV/proj2")
library(fields)
library(maptools)
library(ggplot2)
library(ggmap)
library(maps)
library(plyr)
library(lattice)
library(Rmisc)
library(mapproj)
library(rgbif)

#######################################
#Preprocess data:
filename <- "GlobalFloodsRecordMaster.csv"
df <- read.csv(filename, as.is = TRUE)
df$Centroid.X <- as.numeric(df$Centroid.X)
df$Centroid.Y <- as.numeric(df$Centroid.Y)
df$Severity..<- as.numeric(df$Severity..)
df <- df[-which(is.na(df$Centroid.X)),]
XLon <- as.numeric(df$Centroid.X)
YLat <- as.numeric(df$Centroid.Y)
Z <- as.numeric(df$Severity..)
Cause <- df$Main.cause
#rev(sort(table(Cause)))[1:6]
n <- length(Cause)
for (i in 1:n){
  if (grepl('eavy',Cause[i])){Cause[i] <- replace(Cause[i], grepl('eavy',Cause[i]),1) }
  #1 stands for 'Heavy Rain'
  else if(grepl('clone',Cause[i])){Cause[i] <- replace(Cause[i], grepl('clone',Cause[i]),2)}
  #2 stands for'Tropical Cyclone'
  else if(grepl('onsoon',Cause[i])){Cause[i] <- replace(Cause[i], grepl('onsoon',Cause[i]),3)}
  #3 stands for 'Monsoon'
  else if(grepl('orrential',Cause[i])){Cause[i] <- replace(Cause[i], grepl('orrential',Cause[i]),4)}
  #4 stands for 'Torrential Rain'
  else {Cause[i] <- replace(Cause[i],TRUE,5)}
  #5 stands for 'Other Causes'
}

#######################################

#Try simple plot of "Main Causes" and "Severity":
data(wrld_simpl)
plot(wrld_simpl)
points(XLon, YLat, pch = 16, cex = Z/3, col = as.numeric(Cause)+3)
title(main = "Flood Distribution \nBased on Main Causes and Severity", cex.main =1)
legend("bottomleft",legend = c("Heavy Rain","Tropical Cyclone","Monsoon",
```

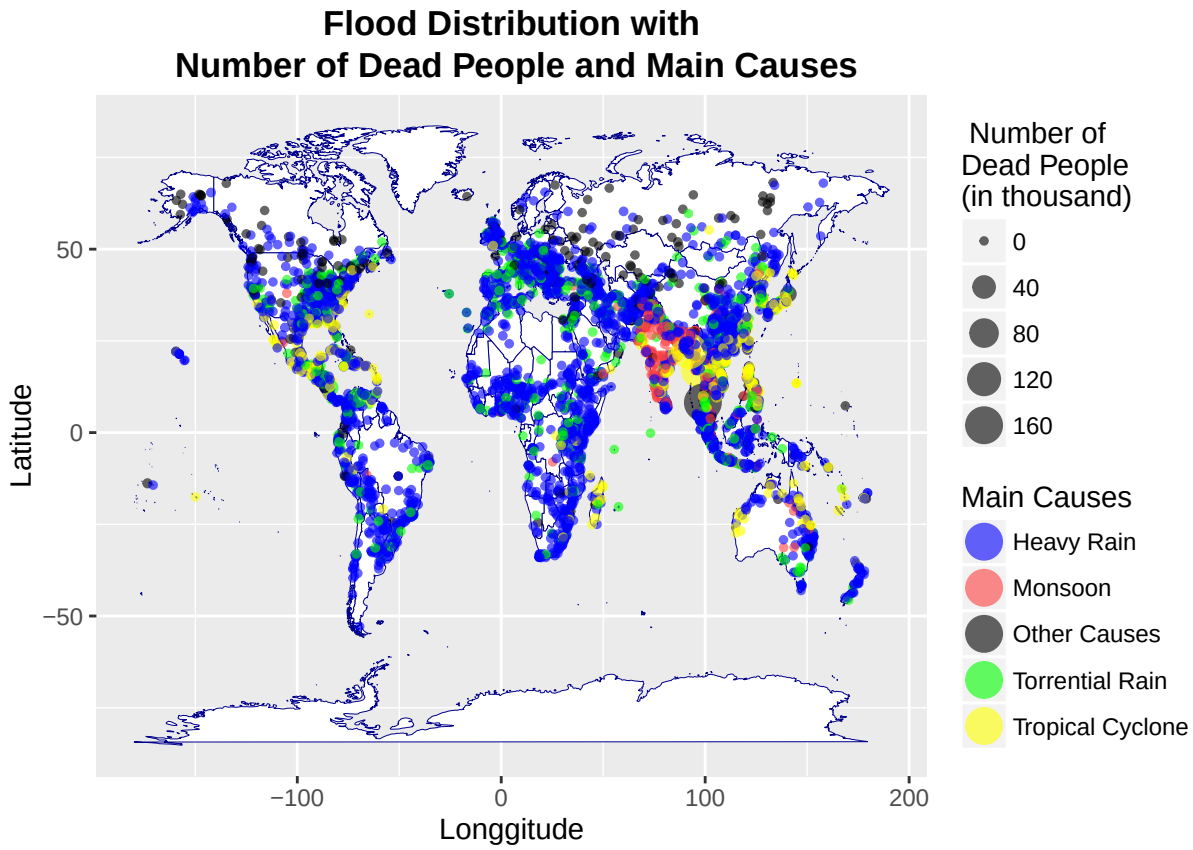
```
                                "Torrential Rain","Other Causes"),
       cex = 0.4, pch = 16, col = c(4:8), title ="Main Cause",title.adj = .5)
legend("bottomright",legend = c("level 1.0","level 1.5","level 2.0"),
       pt.cex = c(1,1.5,2)/3, cex = .4, pch = 16,col =4, title = "Severity")
```




```
#######################################

#Try ggplot of "Number of Dead People" and "Severity"
Dead <- as.numeric(df$Dead)
df_new <- data.frame(XLon,YLat,Z,Dead)

world <- map_data("world")
ggplot(world, aes(long, lat)) +
  geom_polygon(aes(group=group), fill = "White", color ="Dark Blue", size = 0.05) +
  geom_jitter(data=df_new, aes(XLon, YLat, color = Z, size = Dead/1000), alpha = 0.6) +
  scale_colour_gradientn(colours = rainbow(3, start = 0.17, alpha = 0.2)) +
  labs(title = "Flood Distribution with\n Number of Dead People and Severity", x = "Longgitude",
       y = "Latitude", size = " Number of\nDead People\n(in thousand)", color = "Severity")+
  theme(plot.title = element_text(lineheight=1, face="bold"))
```




```
#####################################

#Try ggplot of "Number of Dead People" and "Main Causes"
for (i in 1:n){
  Cause[i] <- replace(Cause[i], Cause[i]=='1','Heavy Rain')
  Cause[i] <- replace(Cause[i], Cause[i]=='2','Tropical Cyclone')
  Cause[i] <- replace(Cause[i], Cause[i]=='3','Monsoon')
  Cause[i] <- replace(Cause[i], Cause[i]=='4','Torrential Rain')
  Cause[i] <- replace(Cause[i], Cause[i]=='5','Other Causes')
}
df_new2 <- data.frame(XLon,YLat,Cause,Dead)
ggplot(world, aes(long, lat)) +
  geom_polygon(aes(group=group), fill = "White", color ="Dark Blue", size = 0.05) +
  geom_jitter(data=df_new2, aes(XLon, YLat, color = Cause, size = Dead/1000), alpha = 0.6) +
  scale_colour_manual(values = c("blue","brown1","black","green","yellow"))+
  labs(title = "Flood Distribution with\n Number of Dead People and Main Causes", x = "Longgitude",
       y = "Latitude", size = " Number of\nDead People\n(in thousand)", color = "Main Causes")+
  guides(colour = guide_legend(override.aes = list(size=6)))+
  theme(plot.title = element_text(lineheight=1, face="bold"))
```


Flood Distribution with
Number of Dead People and Main Causes
Latitude
50
0
−50
−100
0
100
200
Longgitude
Number of
Dead People
(in thousand)
0
40
80
120
160
Main Causes
Heavy Rain
Monsoon
Other Causes
Torrential Rain
Tropical Cyclone