FinalPresentation

Alice Wang, Jerry Haoming Jiang, Andrew Linxie

December 3, 2015

Getting the Raw Data

Set Working Directory and install packages

```
setwd("/users/alicewang/desktop/stat133/finalproject/")
#install.packages("dplyr")
library(dplyr)
```

Natural Disasters Data

```
##
## 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
```

```
library(readr)
rawnd <- read.csv("./rawdata/naturaldisasters.csv")
types <- c()
type <- as.character(rawnd[,2])
subtype <- as.character(rawnd[,3])
subsubtype <- as.character(rawnd[,11])
for (i in 1:155) {
  if (subsubtype[i] == "--"|| subsubtype[i] == "") {
    if (subtype[i] == "--") {
      types <- c(types, type[i])
    } else if (subtype[i] != "--") {
      types <- c(types, subtype[i])
    }
  } else if (subsubtype[i] != "--") {
    types <- c(types, subsubtype[i])
  }
}
nd <- data.frame("year" = rawnd[,1], "type" = types, "casualty" = rawnd[,5]+rawnd[,6]+rawnd[,7], "cost"
nd
```

```
##     year                          type casualty      cost
```

```
## 1   1995                      Cold wave      710         0
## 2   1995                      Heat wave   115319   3245000
## 3   1995                          Storm    84242  10645000
## 4   1995                 Riverine flood       68   3000000
## 5   1996                          Flood   243187   2720000
## 6   1996              Convective storm     4039   3400000
## 7   1996              Tropical cyclone       36         0
## 8   1996                Riverine flood      132         0
## 9   1996         Winter storm/Blizzard      211         0
## 10  1996              Convective storm       51      8500
## 11  1997              Tropical cyclone   191046   7345000
## 12  1997                   Forest fire      616    150000
## 13  1997 Land fire (Brush, Bush, Pastur      628    180000
## 14  1997                     Lightning       21    150000
## 15  1997                Riverine flood        0         0
## 16  1997                       Tornado     3743   1260000
## 17  1997              Convective storm       64    502000
## 18  1997              Tropical cyclone     1000      2500
## 19  1998                   Forest fire      130   4275000
## 20  1998                      Heat wave    32031   1062000
## 21  1998                          Storm   143646   6527450
## 22  1998                    Flash flood     5409   3263500
## 23  1998                Riverine flood     2200         0
## 24  1998                       Tornado      983    262500
## 25  1998              Convective storm        7         0
## 26  1998              Tropical cyclone    41158    281100
## 27  1999 Land fire (Brush, Bush, Pastur        0   1100000
## 28  1999                       Drought      257   1000000
## 29  1999                      Heat wave      100       300
## 30  1999                          Storm  3045095   7432500
## 31  1999                           Hail     1251    450000
## 32  1999              Convective storm      950    300500
## 33  1999              Tropical cyclone     4763   3417500
## 34  1999                   Forest fire       48   1000000
## 35  1999 Land fire (Brush, Bush, Pastur      801     92000
## 36  2000                       Drought        0   1100000
## 37  2000               Ground movement    25070     50000
## 38  2000                      Heat wave       35         0
## 39  2000                          Flood     1654    286000
## 40  2000                    Flash flood    14511   1069000
## 41  2000                Riverine flood        0         0
## 42  2000                       Tornado     5072    786600
## 43  2000              Convective storm        0         0
## 44  2000                   Forest fire    35416   2500000
## 45  2001               Ground movement      401   2000000
## 46  2001                      Heat wave       56         0
## 47  2001                          Flood    19312     33000
## 48  2001                Riverine flood   102276   6000000
## 49  2001                       Tornado     7350    328800
## 50  2001              Convective storm      111         0
## 51  2001              Tropical cyclone      141         0
## 52  2002                   Forest fire        0   3300000
## 53  2002                       Drought     3838         0
## 54  2002           Extreme temperature        14         0
```

```
## 55  2002                  Viral disease   145776   1013000
## 56  2002                      Heat wave    14050   2900500
## 57  2002                 Riverine flood     4749   2924000
## 58  2002          Winter storm/Blizzard       92    450500
## 59  2002               Convective storm     3511    232100
## 60  2003               Tropical cyclone       42    200000
## 61  2003                    Forest fire       29         0
## 62  2003 Land fire (Brush, Bush, Pastur     3224    123000
## 63  2003                 Ground movement       30         0
## 64  2003                   Viral disease   226292   5920000
## 65  2003                     Flash flood     2046   9000000
## 66  2003                   Riverine flood      180    233600
## 67  2003                       Landslide        0         0
## 68  2003                        Wildfire    27162   3500000
## 69  2004               Convective storm        3         0
## 70  2004               Tropical cyclone     5341    505500
## 71  2004                        Ash fall  5070178  53062500
## 72  2004                    Forest fire      364   1100000
## 73  2004                      Cold wave     7235    715000
## 74  2004                   Riverine flood     1679    309000
## 75  2004                        Wildfire    15205         0
## 76  2005               Convective storm       64         0
## 77  2005               Tropical cyclone    18311    730330
## 78  2005                    Forest fire   831852 157530000
## 79  2005                      Heat wave       53    350000
## 80  2005                    Flash flood      120         0
## 81  2005                   Riverine flood       30    350000
## 82  2005               Convective storm        5    100000
## 83  2006               Tropical cyclone     3529    150000
## 84  2006                    Forest fire      188         0
## 85  2006                 Ground movement    74100   1478500
## 86  2006                      Heat wave      167     32860
## 87  2006                   Riverine flood     4301   2200000
## 88  2006                       Lightning      601    450000
## 89  2006               Convective storm      684   1385000
## 90  2006               Tropical cyclone        3    660000
## 91  2006                        Wildfire       23     66000
## 92  2007 Land fire (Brush, Bush, Pastur        0    300000
## 93  2007                        Drought     8890    728000
## 94  2007                    Flash flood     1544   4600000
## 95  2007                   Riverine flood     6802   1080000
## 96  2007          Winter storm/Blizzard       16    140000
## 97  2007               Convective storm   650072   2815000
## 98  2008               Tropical cyclone     2103         0
## 99  2008                    Forest fire        0         0
## 100 2008 Land fire (Brush, Bush, Pastur 11032692  10002000
## 101 2008                 Ground movement  2300556  39540000
## 102 2008       Severe winter conditions      844   5860000
## 103 2008                   Riverine flood        5    360000
## 104 2008               Convective storm    55021   2000000
## 105 2009           Extra-tropical storm       58   1100000
## 106 2009               Tropical cyclone     9071    666000
## 107 2009                          Storm       18    600000
## 108 2009                    Forest fire       31   1500000
```

```
## 109 2009                      Cold wave        7   1240000
## 110 2009                  Riverine flood      907   7050000
## 111 2009                Convective storm       19         0
## 112 2009                Tropical cyclone       31    100000
## 113 2010 Land fire (Brush, Bush, Pastur      630     12500
## 114 2010                 Ground movement     1020         0
## 115 2010                  Riverine flood       20   1600000
## 116 2010                Convective storm      120         0
## 117 2010                Tropical cyclone     8050   2350000
## 118 2010                         Drought     2866   2700000
## 119 2010                       Heat wave        3   2500000
## 120 2011                  Riverine flood        0   8000000
## 121 2011                Convective storm       22         0
## 122 2011                Tropical cyclone    25117   6600000
## 123 2011                     Forest fire   370677  11550000
## 124 2011 Land fire (Brush, Bush, Pastur     4183  27000000
## 125 2011                         Drought       95   4900000
## 126 2011                       Heat wave       13   1383000
## 127 2012                  Riverine flood        0  20000000
## 128 2012                Convective storm      123         0
## 129 2012                Tropical cyclone      302    174000
## 130 2012                           Storm    77065  52210000
## 131 2012       Severe winter conditions     9012   7610000
## 132 2012                  Riverine flood      129   8500000
## 133 2012                Convective storm     3693   9025000
## 134 2012          Winter storm/Blizzard       28    150000
## 135 2012                     Forest fire     1812    800000
## 136 2013                         Drought       19         0
## 137 2013                 Ground movement    29356   2275000
## 138 2013                       Cold wave        4   1200000
## 139 2013                     Flash flood        9   4900000
## 140 2013                  Riverine flood   172068   7100000
## 141 2013                       Landslide     4933   1400000
## 142 2013                Convective storm      150    706400
## 143 2014                Tropical cyclone        0   2200000
## 144 2014                     Forest fire     3251    700000
## 145 2014                     Flash flood       21   2500000
## 146 2014                Convective storm   102002   1690000
## 147 2014                       Landslide       43     20000
## 148 2014                           Storm      835     66000
## 149 2014                       Lightning      340   5940000
## 150 2014                    Severe storm       55   2000000
## 151 2014          Winter storm/Blizzard       91   1560000
## 152 2014                        Wildfire      431    100000
## 153 2015                           Flood    24048   1000000
## 154 2015                         Tornado       82   2750000
## 155 2015          Winter storm/Blizzard       82    775000
```

```
if (!dir.exists("./data")) dir.create("./data")
file.create("./data/NaturalDisasters.csv")
```

```
## [1] TRUE
```

```
write.csv(nd,"./data/NaturalDisasters.csv")
```

Economic Cost Data

Process to Cleaning the Economics Data:

```
setwd("/users/alicewang/desktop/stat133/finalproject/")
Raw_GDP <- read.csv("./rawdata/GDP_Data.csv", stringsAsFactors = FALSE,col.names = c("Quarter","GDP"))
```

1) Reading in the File

2) Clean Data:

```
Quarterly_GDP_Data <- Raw_GDP[-1:-10,]
Quarterly_GDP_Data <- data.frame(
                          'Quarter' = Quarterly_GDP_Data[,1],
                          'GDP' = Quarterly_GDP_Data[,2], stringsAsFactors = FALSE)
Quarterly_GDP_Data[,"GDP"] <- as.numeric(Quarterly_GDP_Data[,"GDP"])
```

I wanted to remove the introductory piece, specifically, where this Data was obtained from (sorry), so that only the numbers remain.

```
Annual_GDP_Data <- data.frame(
                       'Years' = 1995:2015,
                       'GDP' = Quarterly_GDP_Data[c(seq(0,nrow(Quarterly_GDP_Data),by = 4),nrow(Quarterly
```

3) Create Yearly GDP

```
setwd("/users/alicewang/desktop/stat133/finalproject/")
if (!dir.exists("./data")) dir.create("./data")
file.create("./data/Econ.csv")
```

4) Export Data

```
## [1] TRUE
```

```
write.csv(x = Annual_GDP_Data, file = "./data/Econ.csv", row.names = FALSE)
```

Pokemon Data

Goal: Return a clean data table, types.csv, containing pokemon types, the number of pokemon per type, a power metric for each type, and an average power metric per pokemon for each type

Packages and Working Directory

```
#install.packages("dplyr")
library(dplyr)
library(readr)
```

Getting the Tables we need

```
setwd("/users/alicewang/desktop/stat133/finalproject/")
type_reference <- read_csv("./rawdata/types.csv")
pokemon_stats <- read_csv("./rawdata/pokemon_stats.csv")
pokemon_types <- read_csv("./rawdata/pokemon_types.csv")
stat_names <- read_csv("./rawdata/stat_names.csv")
```

What Types we want

```
types <-  c('water', 'fire', 'flying', 'ground', 'poison', 'dragon', 'dark', 'ice', 'electric', 'rock'
```

Function to get a type given a type ID. Use the data frame, type_reference

```
get_type_by_id <- function(id) {
  return(type_reference[type_reference$id == id,]$identifier[1])
}
```

Group pokemon_types by first type, get actual types from type ID. Get the number of pokemon per type to get values for the count of each type in our types table.

```
get_first <- function(arr) {
  return(arr[1])
}
pokemon_id_types <- group_by(pokemon_types, pokemon_id) %>%
  summarise(type_id = get_first(type_id)) %>%
  mutate(type = sapply(type_id, get_type_by_id)) %>%
  select(-type_id)

type_counts <- group_by(pokemon_id_types, type) %>%
  summarise(pokemon_count = length(pokemon_id)) %>%
  filter(type %in% types)

type_counts
```

```
## Source: local data frame [10 x 2]
##
##        type pokemon_count
##       (chr)         (int)
## 1      dark            30
## 2    dragon            29
```

```
## 3  electric            44
## 4      fire            52
## 5    flying             4
## 6    ground            31
## 7       ice            24
## 8    poison            28
## 9      rock            43
## 10    water           110
```

Link pokemon_id to types filtered to attack and special attack stats. Add these stats for a power metric. Group by type and add powers to get a total power metric per type

```
stats_and_types <- left_join(pokemon_id_types,pokemon_stats) %>% filter(stat_id %in% c(2,4)) %>% group_
```

```
## Joining by: "pokemon_id"
```

```
type_power <- group_by(stats_and_types, type) %>% summarise(total_power = sum(power))
result <- left_join(type_counts, type_power) %>% mutate(avg_power = as.integer(floor(total_power / poke
```

```
## Joining by: "type"
```

```
result
```

```
## Source: local data frame [10 x 4]
##
##        type pokemon_count total_power avg_power
##       (chr)         (int)       (int)     (int)
## 1      dark            30        4884       162
## 2    dragon            29        5842       201
## 3  electric            44        7001       159
## 4      fire            52        8910       171
## 5    flying             4         692       173
## 6    ground            31        4541       146
## 7       ice            24        3507       146
## 8    poison            28        3783       135
## 9      rock            43        6553       152
## 10    water           110       15966       145
```

Store clean file as types_clean.csv

```
setwd("/users/alicewang/desktop/stat133/finalproject/")
if (!dir.exists("./data")) dir.create("./data")
file.create("./data/types.csv")
```

```
## [1] TRUE
```

```
write.csv(x = result, file = "./data/types.csv", row.names = FALSE)
```

Exploratory Data Analysis

Modelling