Pokemon Cleanup

Andrew Linxie

December 3, 2015

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)
```

```
##
## 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)
```

Getting the Tables we need

```
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

```
if (!dir.exists("./data")) dir.create("./data")
write.csv(x = result, file = "./data/types.csv", row.names = FALSE)
```