

Game Of Thrones

Aaron Banlao

10/8/2021

```
suppressPackageStartupMessages(library("jsonlite"))
suppressPackageStartupMessages(library("tidyverse"))
library(jsonlite)
library(tidyverse)
library(ggplot2)
```

```
url <- fromJSON("https://raw.githubusercontent.com/jeffreylancaster/game-of-thrones/master/data/characters.json")
```

```
url <- as.data.frame(url)
```

```
url <- as.tibble(url)
```

```
## Warning: 'as.tibble()' was deprecated in tibble 2.0.0.
## i Please use 'as_tibble()' instead.
## i The signature and semantics have changed, see '?as_tibble'.
```

```
url <- url %>%
  rename(c(characterName = characters.characterName, houseName = characters.houseName, royal = characters.royal))
```

```
characters <- url %>%
  dplyr::select(-c("characters.characterLink", "characters.actorName", "characters.actorLink", "characters.characterLink"))
```

```
characters %>%
  head(n=5)
```

```
## # A tibble: 5 x 10
##   character~1 house~2 royal parents sibling~3 killed~4 nickn~5 killed paren~6 marri~7
##   <chr>         <list> <lgl> <list> <list> <list> <chr> <list> <list> <list>
## 1 Addam Ma~ <NULL> NA <NULL> <NULL> <NULL> <NA> <NULL> <NULL> <NULL>
## 2 Aegon Ta~ <chr> TRUE <chr> <chr> <chr> <NA> <NULL> <NULL> <NULL>
## 3 Aeron Gr~ <chr> NA <NULL> <chr> <NULL> Dampha~ <NULL> <NULL> <NULL>
## 4 Aerys II~ <chr> TRUE <NULL> <chr> <chr> The Ma~ <chr> <chr> <chr>
## 5 Akho <NULL> NA <NULL> <NULL> <chr> <NA> <NULL> <NULL> <NULL>
## # ... with abbreviated variable names 1: characterName, 2: houseName,
## # 3: siblings, 4: killedBy, 5: nickname, 6: parentOf, 7: marriedEngaged
```

```
#Filtering the dataset to only characters who have killed at least one other character
```

```
hasKilled <- characters %>%
  drop_na(killed)
```

```
head(hasKilled)
```

```
## # A tibble: 6 x 10
##   character~1 house~2 royal parents sibli~3 kille~4 nickn~5 killed paren~6 marri~7
##   <chr>         <list> <lgl> <list> <list> <list> <chr> <list> <list> <list>
## 1 Aerys II~ <chr> TRUE <NULL> <chr> <chr> The Ma~ <chr> <chr> <chr>
## 2 Alliser ~ <NULL> NA <NULL> <NULL> <chr> <NA> <chr> <NULL> <NULL>
## 3 Amory Lo~ <NULL> NA <NULL> <NULL> <chr> <NA> <chr> <NULL> <NULL>
## 4 Arya Sta~ <chr> NA <chr> <chr> <NULL> <NA> <chr> <NULL> <NULL>
## 5 Black Wa~ <chr> NA <chr> <NULL> <chr> <NA> <chr> <NULL> <NULL>
## 6 Bowen Ma~ <NULL> NA <NULL> <NULL> <chr> <NA> <chr> <NULL> <NULL>
## # ... with abbreviated variable names 1: characterName, 2: houseName,
## # 3: siblings, 4: killedBy, 5: nickname, 6: parentOf, 7: marriedEngaged
```

There are 80 characters who have killed at least 1 person.

```
nrow(hasKilled)
```

```
## [1] 80
```

The most prolific killers are Daenerys Targaryen & Sandor Clegane both with 16, followed by Arya Stark & Jon Snow with 14 kills, Gregor Clegane with 9 kills, etc.

```
characters <- hasKilled %>%
  unnest(killed) %>%
  dplyr::select(characterName, killed) %>%
  group_by(characterName) %>%
  summarize(numkills = n(), killed = paste(killed, collapse = ", ")) %>%
  arrange(desc(numkills))

print(characters)
```

```
## # A tibble: 80 x 3
##   characterName      numkills killed
##   <chr>             <int> <chr>
## 1 Daenerys Targaryen    16 Khal Drogo, Mirri Maz Duur, Doreah, Xaro Xhoan D~
## 2 Sandor Clegane       16 Mycah, Kings Landing Rioter #1, Kings Landing Ri~
## 3 Arya Stark           14 Red Keep Stableboy, Frey Soldier #1, Polliver, R~
## 4 Jon Snow             14 Othor, Qhorin Halfhand, Orell, Karl Tanner, Styr~
## 5 Gregor Clegane        9 Elia Martell, Rhaenys Targaryen, Aegon Targaryen~
## 6 Ramsay Snow           9 Master Torturer, Medger Cerwyn, Lady Cerwyn, Uni~
## 7 Cersei Lannister      7 Lancel Lannister, High Sparrow, Loras Tyrell, Ma~
## 8 Daario Naharis        6 Mero, Prendahl na Ghezn, Oznak zo Pahl, Mossador~
## 9 Jaime Lannister       6 Aerys II Targaryen, Jory Cassel, Alton Lannister~
## 10 Theon Greyjoy        5 Stiv, Rodrik Cassel, Myranda, Bolton Soldier, Ha~
## # ... with 70 more rows
```

```

characters %>%
  head(10) %>%
  ggplot(aes(x = numkills, y = reorder(characterName, numkills), fill = characterName)) +
  geom_bar(stat = "identity") +
  labs(title = "Top 10 Prolific Killers of Game of Thrones", x = "Character", y = "Number of Kills") +
  scale_x_continuous(breaks = seq(0, max(characters$numkills), by = 2)) +
  guides(fill = F)

```

Warning: The 'scale' argument of 'guides()' cannot be 'FALSE'. Use "none" instead as ## of ggplot2 3.3.4.

