

Comic Characters 538

2019/04/12

สมาชิกกลุ่ม

สมาชิกคนที่ 1

ชื่อ: นายวิพุธ ภูทอง

รหัสนักศึกษา: 60070090

เวลาเรียน: อังคาร 13:30 - 16:30 น.

สมาชิกคนที่ 2

ชื่อ: นายธีรภัทร ไกรศรีศิริกุล

รหัสนักศึกษา: 60070183

เวลาเรียน: อังคาร 13:30 - 16:30 น.

โจทย์ปัญหา

1. Marvel หรือ DC เขียนตัวละครมาแล้วเสียชีวิตไปมากกว่ากัน?
2. เพศและพลังของตัวละครมีผลหรือแนวโน้มที่จะถูกเขียนให้เสียชีวิตมากน้อยเท่าไร?

ชุดข้อมูลที่ใช้จาก fivethirtyeight : comic_characters

ขั้นตอนการดำเนินการ

1.) โหลด Library

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.2.1 --
## v ggplot2 3.1.0      v purrr   0.3.1
## v tibble  2.0.1      v dplyr   0.8.0.1
## v tidyr   0.8.3      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(fivethirtyeight)
library(stringr)
library(ggthemes)
```

2.) ทดลองดูข้อมูล

comic_characters

```
## # A tibble: 23,272 x 16
##   publisher page_id name urlslug id align eye hair sex gsm
##   <chr> <int> <chr> <chr> <chr> <ord> <chr> <chr> <chr> <chr>
## 1 Marvel 1678 Spid~ "\\Sp~ Secr~ Good~ Haze~ Brow~ Male~ <NA>
## 2 Marvel 7139 Capt~ "\\Ca~ Publ~ Good~ Blue~ Whit~ Male~ <NA>
## 3 Marvel 64786 "Wol~ "\\Wo~ Publ~ <NA> Blue~ Blac~ Male~ <NA>
## 4 Marvel 1868 "Iro~ "\\Ir~ Publ~ Good~ Blue~ Blac~ Male~ <NA>
## 5 Marvel 2460 Thor~ "\\Th~ No D~ Good~ Blue~ Blon~ Male~ <NA>
## 6 Marvel 2458 Benj~ "\\Be~ Publ~ Good~ Blue~ No H~ Male~ <NA>
## 7 Marvel 2166 Reed~ "\\Re~ Publ~ Good~ Brow~ Brow~ Male~ <NA>
## 8 Marvel 1833 Hulk~ "\\Hu~ Publ~ Good~ Brow~ Brow~ Male~ <NA>
## 9 Marvel 29481 Scot~ "\\Sc~ Publ~ <NA> Brow~ Brow~ Male~ <NA>
## 10 Marvel 1837 Jona~ "\\Jo~ Publ~ Good~ Blue~ Blon~ Male~ <NA>
## # ... with 23,262 more rows, and 6 more variables: alive <chr>,
## # appearances <int>, first_appearance <chr>, month <chr>, year <int>,
## # date <date>
```

3.) คำตอบ

3.1.) ดูว่าตัวละครจากค่ายหนังสือการ์ตูนทั้ง Marvel และ DC นั้นมีอัตราการเสียชีวิตเท่าไร

```
by_publisher <- comic_characters %>%
  count(publisher, alive) %>%
  filter(!is.na(alive))
```

by_publisher

```
## # A tibble: 4 x 3
##   publisher alive      n
##   <chr> <chr> <int>
## 1 DC Deceased Characters 1693
## 2 DC Living Characters 5200
## 3 Marvel Deceased Characters 3765
## 4 Marvel Living Characters 12608
```

3.2.) ดูปัจจัยอื่นอย่างเพศและฝั่งของตัวละคร ว่ามีอัตราการเสียชีวิตเท่าไร

```
by_sex <- comic_characters %>%
  count(sex, alive) %>%
  filter(!is.na(alive)) %>%
  filter((sex == 'Male Characters') | (sex == 'Female Characters'))
```

by_sex

```
## # A tibble: 4 x 3
##   sex alive      n
##   <chr> <chr> <int>
## 1 Female Characters Deceased Characters 1155
## 2 Female Characters Living Characters 4648
```

```
## 3 Male Characters    Deceased Characters    4140
## 4 Male Characters    Living Characters      12280

by_align <- comic_characters %>%
  count(aligned, alive) %>%
  filter(!is.na(alive)) %>%
  filter(aligned == 'Bad Characters' | aligned == 'Good Characters' | is.na(aligned))
```

```
## Warning: Factor `aligned` contains implicit NA, consider using
## `forcats::fct_explicit_na`
```

```
by_align
```

```
## # A tibble: 6 x 3
##   aligned      alive      n
##   <ord>      <chr>    <int>
## 1 Bad Characters Deceased Characters 2732
## 2 Bad Characters Living Characters 6882
## 3 Good Characters Deceased Characters 1513
## 4 Good Characters Living Characters 5954
## 5 <NA>          Deceased Characters 1212
## 6 <NA>          Living Characters 4970
```

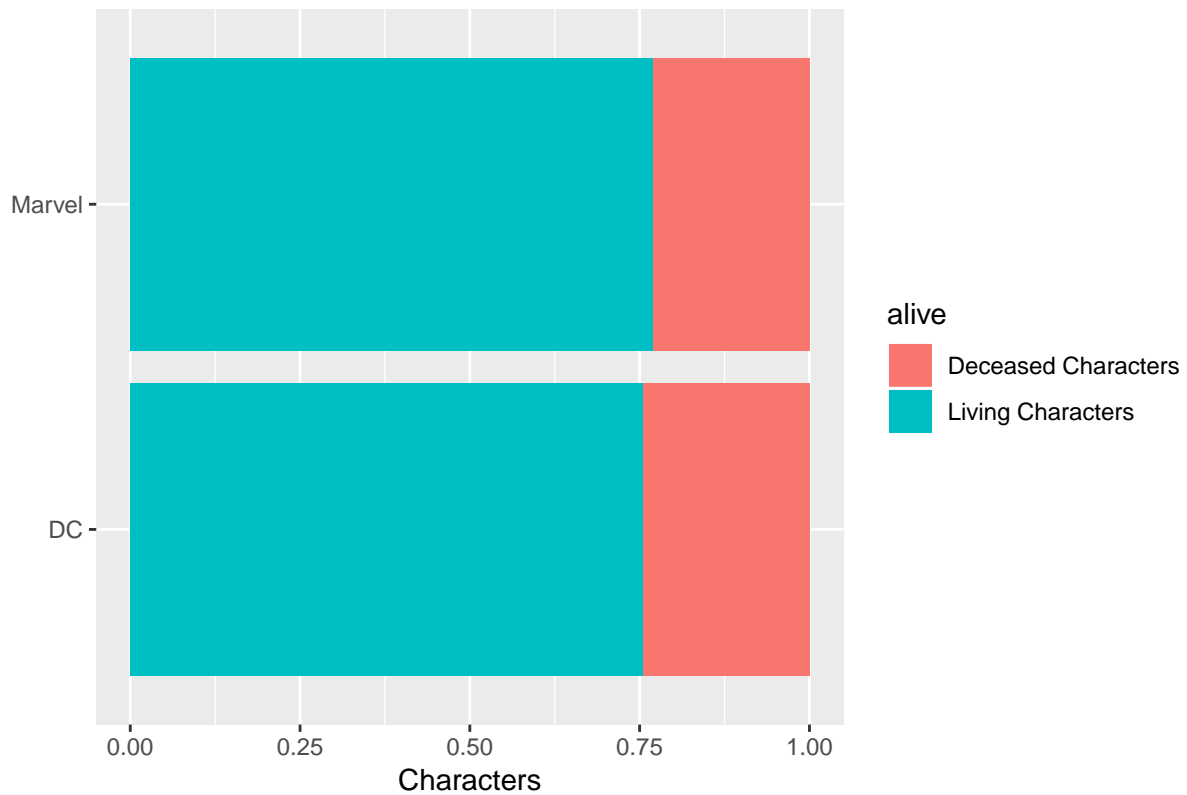
4.) นำข้อมูลมานำเสนอ (Data Visualization)

4.1.) ใช้ Bar Chart

```
p1 <- ggplot(data=by_publisher, aes(x = publisher, y = n)) +
  # map "alive" of beverage to aes "fill" of the bar
  geom_col(aes(fill = alive), position = position_fill()) +
  # flip coordinate
  coord_flip() +
  # add labels
  labs(x="", y="Characters", title="Comic characters status from each publisher (In percentage)")

p1
```

Comic characters status from each publisher (In percentage)

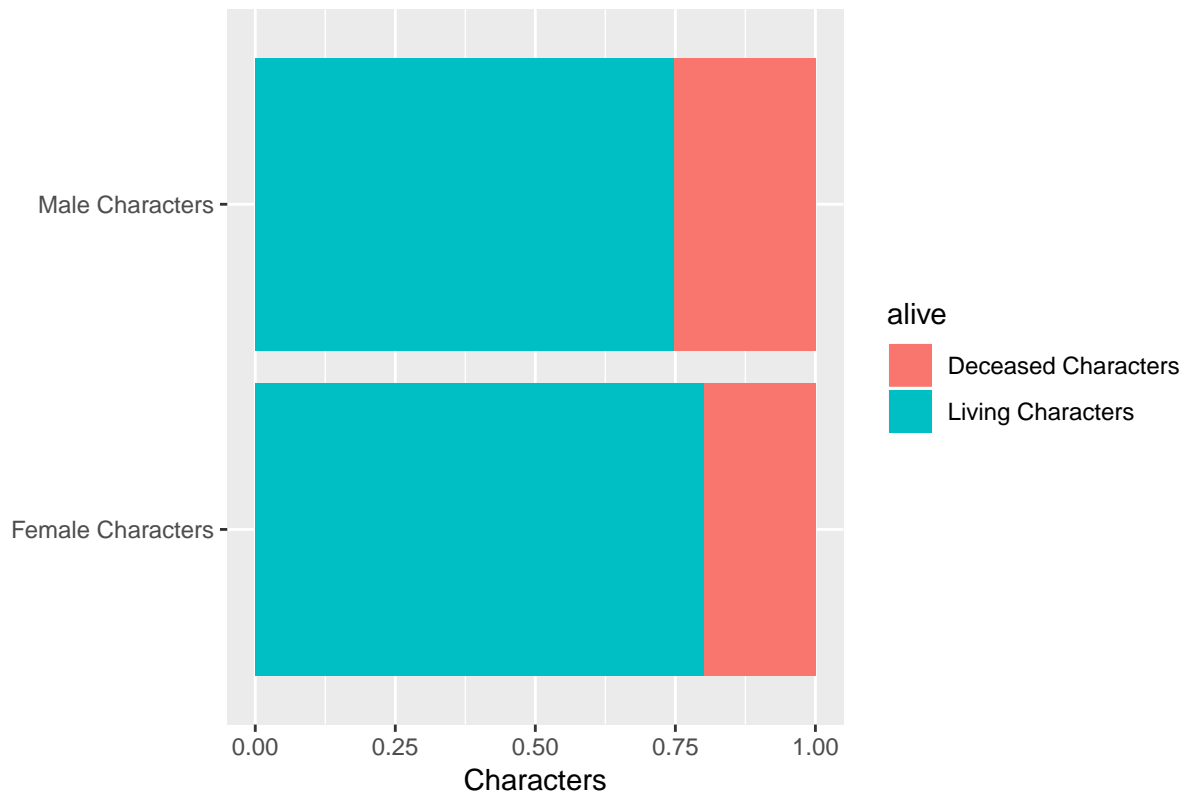


4.2.) ໙ Bar Chart

```
p2 <- ggplot(data=by_sex, aes(x = sex, y = n)) +
  # map "alive" of beverage to aes "fill" of the bar
  geom_col(aes(fill = alive), position = position_fill()) +
  # flip coordinate
  coord_flip() +
  # add labels
  labs(x="", y="Characters", title="Comic characters status by character gender (In percentage)")

p2
```

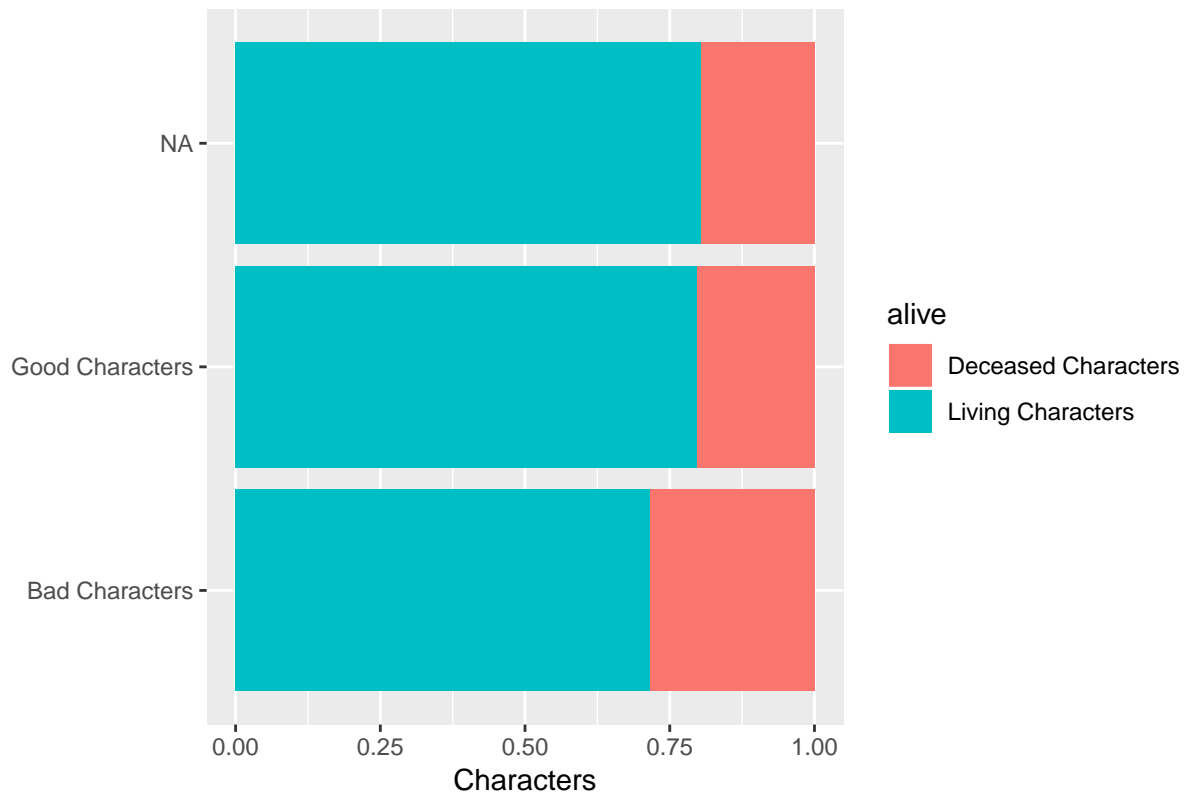
Comic characters status by character gender (In percentage)



```
p3 <- ggplot(data=by_align, aes(x = align, y = n)) +
  # map "alive" of beverage to aes "fill" of the bar
  geom_col(aes(fill = alive), position = position_fill()) +
  # flip coordinate
  coord_flip() +
  # add labels
  labs(x="", y="Characters", title="Comic characters status by character alignment (In percentage)")
```

p3

Comic characters status by character alignment (In percentage)



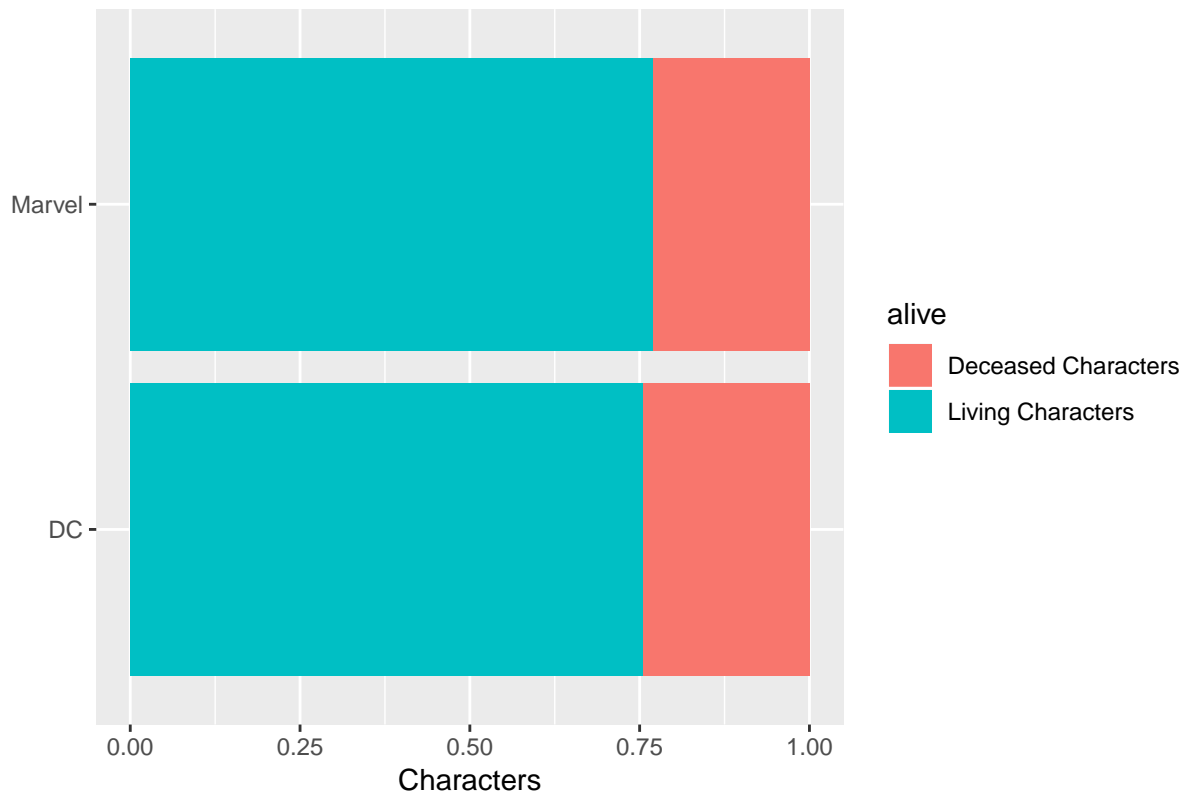
สรุปผล

1.)

จะสังเกตได้ว่า ทั้ง Marvel และ DC นั้นมีการเขียนให้ตัวละครเสียชีวิตในอัตราที่ใกล้เคียงกัน โดยอัตราการเสียชีวิตของตัวละครจาก Marvel อยู่ที่ 23.0% และตัวละครจาก DC อยู่ที่ 24.56%

p1

Comic characters status from each publisher (In percentage)

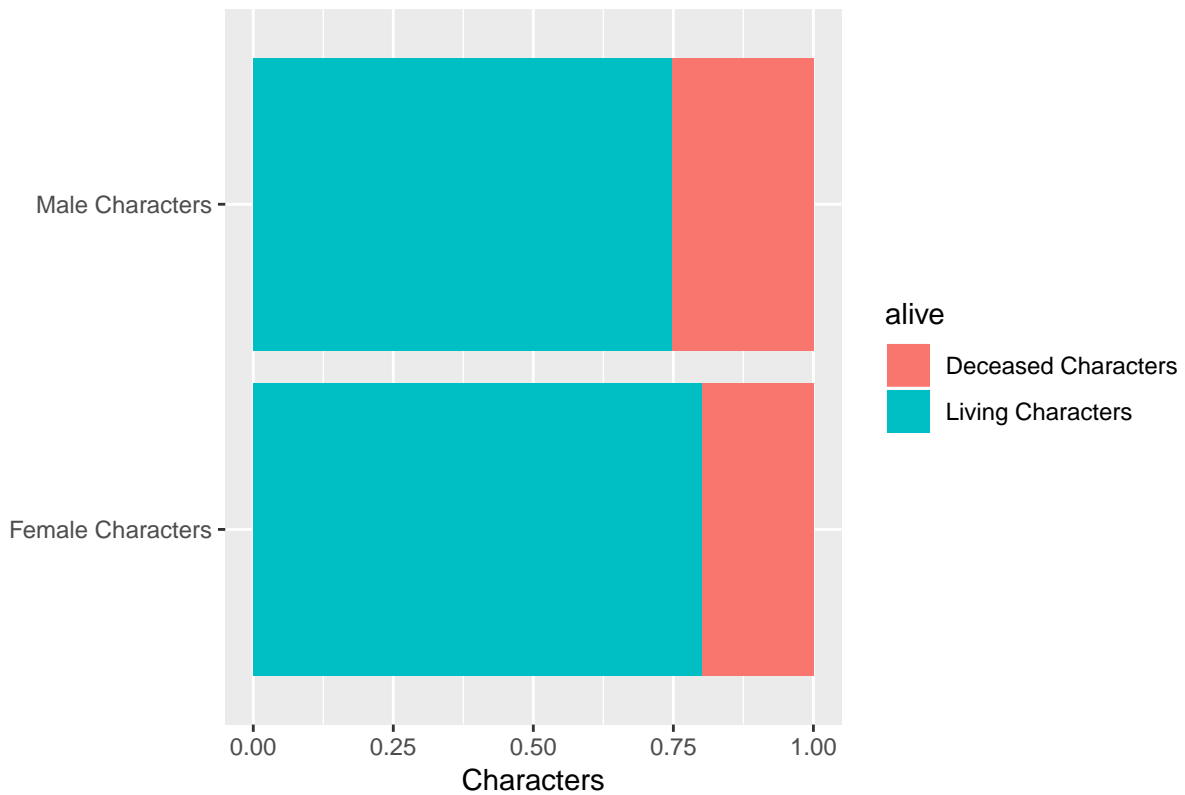


2.)

ต่อมาหากดูจากเพศของตัวละคร จะสังเกตได้ว่า อัตราการเสียชีวิตของตัวละครเพศชาย มีมากกว่าตัวละครเพศหญิงเล็กน้อย โดยตัวละครเพศหญิงมีอัตราการเสียชีวิตอยู่ที่ 19.9% และตัวละครเพศชายมีอัตราการเสียชีวิตที่ 25.21%

p2

Comic characters status by character gender (In percentage)



สุดท้ายหากดูจากฝ่ายของตัวละครก็จะพบว่า ตัวละครที่ไม่อยู่ฝั่งใดฝ่ายหนึ่งหรือฝ่ายดี มีอัตราการเสียชีวิตน้อยกว่าตัวละครฝ่ายร้ายระดับหนึ่ง โดยตัวละครที่ไม่อยู่ฝ่ายใดฝ่ายหนึ่งมีอัตราการเสียชีวิตอยู่ที่ 19.61% ตัวละครฝ่ายดีอยู่ที่ 20.26% และตัวละครฝ่ายร้ายอยู่ที่ 28.42%

p3

Comic characters status by character alignment (In percentage)

