

Avengers ELO Analysis

Emanuele Lena - 142411

Avengers ELO Rating



Figure 1: <https://www.flickr.com/photos/tales2astonish/6976086962>

Inspirations

Dragon Ball Power Levels

TODO...

The ELO Algorithm

$$\begin{aligned}r_i &\leftarrow r_i + \kappa(s_{i,j} - \mu_{i,j}) \\r_j &\leftarrow r_j + \kappa(s_{j,i} - \mu_{j,i})\end{aligned}$$

Building my Dataframe

My sample

Movies:

```
## [1] "Captain America"      "Iron Man"              "L'incredibile Hulk"
## [4] "Thor"                  "Iron Man 2"            "Avengers"
```

Dataset Structure:

```
## # A tibble: 10 x 7
##       id movie    winner    loser    comment    terrain_winner terrain_loser
##   <int> <chr>    <chr>    <chr>    <chr>        <chr>         <chr>
## 1     1  1 Captain~ 1-civilian Steve ~ scena bullo c~ land
## 2     2  2 Captain~ Bucky    1-civi~ scena bullo c~ land
## 3     3  3 Captain~ Red Skull 1-civi~ scena scopert~ land
## 4     4  4 Captain~ 3-Hydra s~ 3-US A~ scena creazio~ land
## 5     5  5 Captain~ 3-Hydra s~ Abrahm~ scena creazio~ land
## 6     6  6 Captain~ 1-US Army~ 1-Hydr~ scena creazio~ land
## 7     7  7 Captain~ Peggy Car~ 1-Hydr~ scena creazio~ land
## 8     8  8 Captain~ Captain A~ 1-Hydr~ scena creazio~ land
## 9     9  9 Captain~ Red Skull 3-nazi~ scena visita ~ land
## 10    10 10 Captain~ Captain A~ 5-Hydr~ scena liberaz~ land
```

The concept of fight

I have different variants of fights:

- 1 vs 1

```
## # A tibble: 1 x 3
##       id winner    loser
##   <int> <chr>    <chr>
## 1    13 Captain America Red Skull
```

- 1 vs many

```
## # A tibble: 1 x 3
##       id winner    loser
##   <int> <chr>    <chr>
## 1    49 Iron Man 11-terrorist
```

- many vs 1 (or many vs many)

```
## # A tibble: 2 x 3
##       id winner    loser
##   <int> <chr>    <chr>
## 1    30 50-Hydra soldier Captain America
## 2    34 50-US Army Soldier WW2 31-Hydra Soldier
```

- more than one winner

```
## # A tibble: 1 x 3
##       id winner          loser
##   <int> <chr>          <chr>
## 1     23 Captain America,Bucky 1-Hydra soldier
```

- partial damage

```
## # A tibble: 1 x 3
##       id winner          loser
##   <int> <chr>          <chr>
## 1     93 5-US Army Soldier 0.1-Hulk
```

Data Manipulations

Tools

Libraries:

Data Tidying

- Separate rows with “charater 1,charater 2,...” into more partial victories (w 1/n “winning rate”)
- Turn values with singular charaters into “1-charater”
- separate pairs “n-charaters” in 2 columns

```
## # A tibble: 409 x 9
##       id winner_n winner_charater loser_n loser_charater movie comment
##   <int>   <dbl> <chr>          <dbl> <chr>          <chr> <chr>
## 1     1       1 civilian          1 Steve Rogers  Capt~ scena ~
## 2     2       1 Bucky          1 civilian    Capt~ scena ~
## 3     3       1 Red Skull        1 civilian    Capt~ scena ~
## 4     4       3 Hydra soldier      3 US Army Soldi~ Capt~ scena ~
## 5     5       3 Hydra soldier      1 Abrahm Eskine Capt~ scena ~
## 6     6       1 US Army Soldie~    1 Hydra soldier Capt~ scena ~
## 7     7       1 Peggy Carter      1 Hydra soldier Capt~ scena ~
## 8     8       1 Captain America    1 Hydra soldier Capt~ scena ~
## 9     9       1 Red Skull        3 nazi soldier Capt~ scena ~
## 10    10       1 Captain America    5 Hydra soldier Capt~ scena ~
## # ... with 399 more rows, and 2 more variables: terrain_winner <chr>,
## #   terrain_loser <chr>
```

Charater list + general statistics

```
## # A tibble: 81 x 5
##       id charater          n_fights n_win n_lose
##   <int> <chr>          <dbl> <dbl> <dbl>
## 1    11 Iron Man          57    42    15
## 2    17 Hulk            56    37    19
## 3    35 Loki            52    29    23
## 4    31 Thor            52    37    15
## 5     7 Captain America   40    31     9
## 6    57 Chitauri Soldier  40     8    32
## 7     4 Hydra soldier     37     7    30
## 8    27 Ice Giant        34    10    24
## 9    24 Abominio        26    22     4
```

```
## 10      1 civilian                23      4      19
## # ... with 71 more rows
```

Prepare the dataframe for ELO

Expected Format

```
## # A tibble: 409 x 9
##       id winner_n winner_charater loser_n loser_charater movie comment
##   <int>   <dbl> <chr>           <dbl> <chr>           <chr> <chr>
## 1     1     1     1 civilian             1 Steve Rogers   Capt~ scena ~
## 2     2     2     1 Bucky              1 civilian      Capt~ scena ~
## 3     3     3     1 Red Skull          1 civilian      Capt~ scena ~
## 4     4     4     3 Hydra soldier       3 US Army Soldi~ Capt~ scena ~
## 5     5     5     3 Hydra soldier       1 Abrahm Eskine Capt~ scena ~
## 6     6     6     1 US Army Soldie~    1 Hydra soldier Capt~ scena ~
## 7     7     7     1 Peggy Carter       1 Hydra soldier Capt~ scena ~
## 8     8     8     1 Captain America    1 Hydra soldier Capt~ scena ~
## 9     9     9     1 Red Skull          3 nazi soldier  Capt~ scena ~
## 10    10    10     1 Captain America    5 Hydra soldier Capt~ scena ~
## # ... with 399 more rows, and 2 more variables: terrain_winner <chr>,
## #   terrain_loser <chr>
```

Actual Dataset Format

```
## # A tibble: 409 x 9
##       id winner_n winner_charater loser_n loser_charater movie comment
##   <int>   <dbl> <chr>           <dbl> <chr>           <chr> <chr>
## 1     1     1     1 civilian             1 Steve Rogers   Capt~ scena ~
## 2     2     2     1 Bucky              1 civilian      Capt~ scena ~
## 3     3     3     1 Red Skull          1 civilian      Capt~ scena ~
## 4     4     4     3 Hydra soldier       3 US Army Soldi~ Capt~ scena ~
## 5     5     5     3 Hydra soldier       1 Abrahm Eskine Capt~ scena ~
## 6     6     6     1 US Army Soldie~    1 Hydra soldier Capt~ scena ~
## 7     7     7     1 Peggy Carter       1 Hydra soldier Capt~ scena ~
## 8     8     8     1 Captain America    1 Hydra soldier Capt~ scena ~
## 9     9     9     1 Red Skull          3 nazi soldier  Capt~ scena ~
## 10    10    10     1 Captain America    5 Hydra soldier Capt~ scena ~
## # ... with 399 more rows, and 2 more variables: terrain_winner <chr>,
## #   terrain_loser <chr>
```

I decided to apply this algorithm:

- Winner goes always in White, Loser goes always in Black => score will be always be a numbers in (0,1]
- given n winners and m losers, I will have...
 - m/n lines where White=winner, Black=loser, Score=1
 - (eventually) one line with White=winner, Black=loser and

$$Score = \frac{rest(\frac{m}{n})}{n} * 0.5 + 0.5$$

Dataset in ELO Format

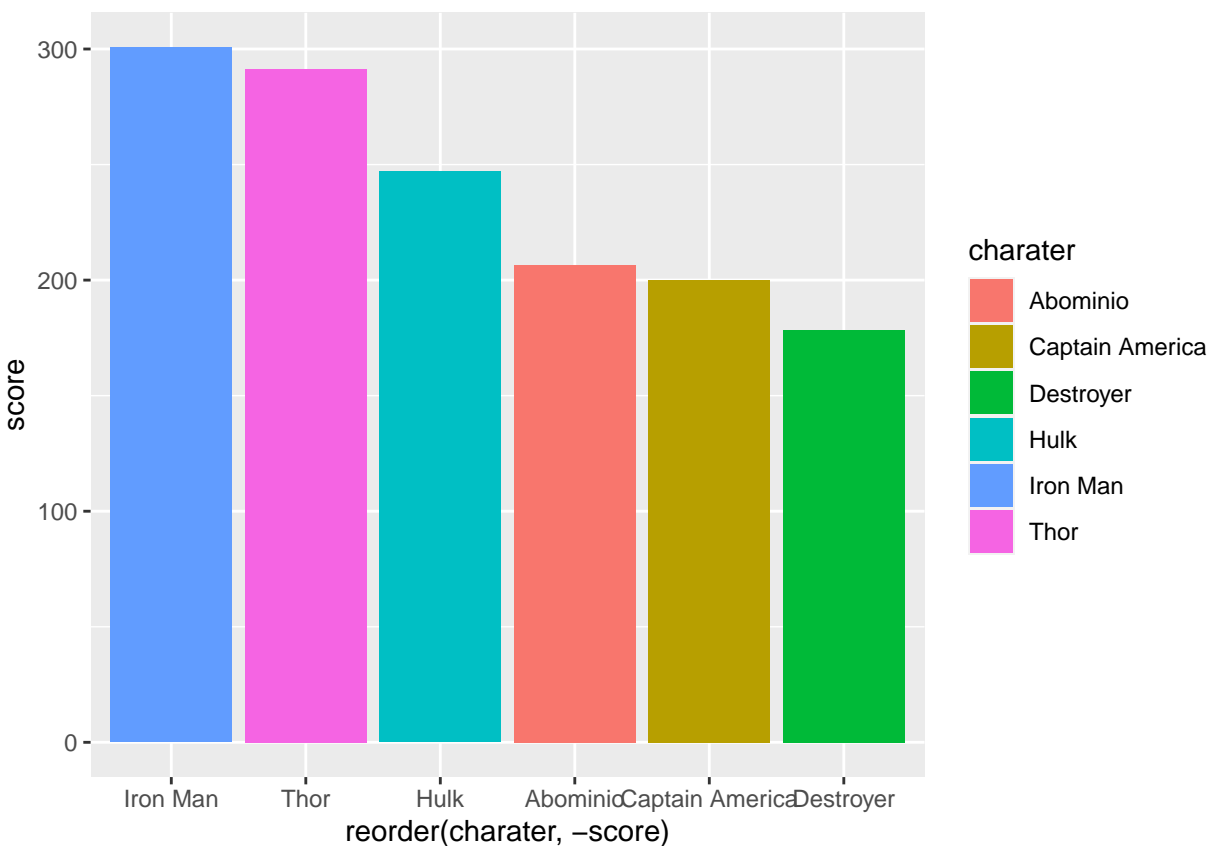
```
## # A tibble: 828 x 9
## # Groups:   id [374]
```

```
##      id count winner      loser      Score movie      comment      White Black
##    <int> <dbl> <chr>      <chr>      <dbl> <chr>      <chr>      <int> <int>
##  1      1      1 civilian Steve Ro~ 1      Captain~ scena bullo cine~      1      59
##  2      2      1 Bucky      civilian 1      Captain~ scena bullo cine~      2      1
##  3      3      1 Red Skull civilian 1      Captain~ scena scoperta c~      3      1
##  4      4      1 Hydra sol~ US Army ~ 1      Captain~ scena creazione ~      4      8
##  5      5      1 Hydra sol~ Abrahm E~ 0.667 Captain~ scena creazione ~      4      60
##  6      6      1 US Army S~ Hydra so~ 1      Captain~ scena creazione ~      5      4
##  7      7      1 Peggy Car~ Hydra so~ 1      Captain~ scena creazione ~      6      4
##  8      8      1 Captain A~ Hydra so~ 1      Captain~ scena creazione ~      7      4
##  9      9      1 Red Skull nazi sol~ 1      Captain~ scena visita uff~      3      61
## 10     9      2 Red Skull nazi sol~ 1      Captain~ scena visita uff~      3      61
## # ... with 818 more rows
```

ELO Classification

Results and comments

Results



```
## # A tibble: 81 x 3
##       id score charater
##   <int> <dbl> <chr>
## 1     11  301. Iron Man
## 2     31  291. Thor
```

```
## 3    17 247. Hulk
## 4    24 206. Abominio
## 5     7 200. Captain America
## 6    30 178. Destroyer
## 7    48 148. Black Widow
## 8    51 145. Hawkeye
## 9    35 139. Loki
## 10   49 108. War Machine
## # ... with 71 more rows
```

Problems

- “Implicit Strengths” (ex. Giant Chitauri, Military veicles, ecc...)

```
## # A tibble: 4 x 3
##   id score charater
##   <int> <dbl> <chr>
## 1    16  9.63 Pepper Potts
## 2    58  9.12 Chitauri Giant
## 3    79 -7.17 Elicopter
## 4    13 -17.1 Fighter
```

- “Cinematic Distorsion” (some charaters are too much “in the spotlight”)
-

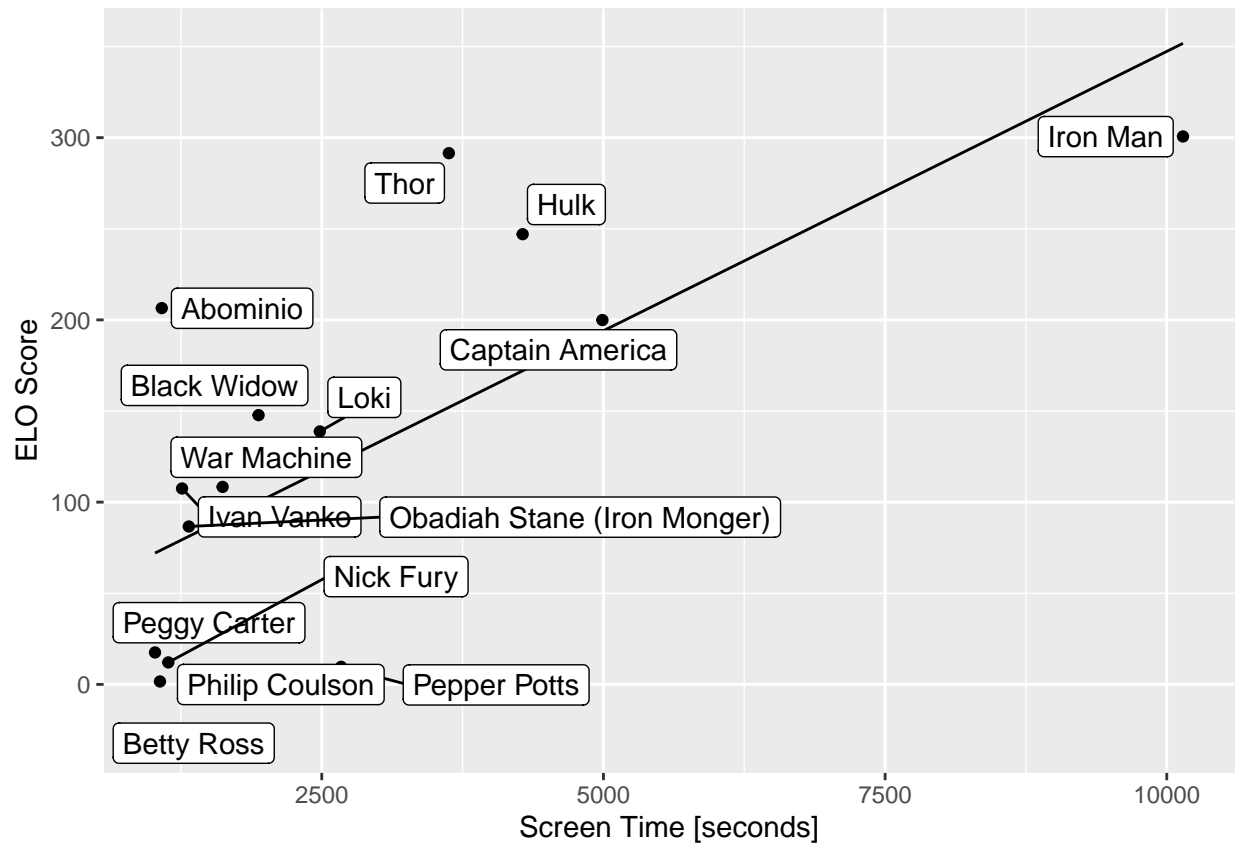
Correlation Score-Fights

Now I try to find a correlation between score and fights results

```
## # A tibble: 1 x 3
##   cor_score_tot cor_score_win cor_score_lose
##   <dbl>         <dbl>         <dbl>
## 1      0.242      0.605      -0.292
```

Correlation Score-Screen Time

```
## # A tibble: 15 x 4
##   id charater          score screen_time
##   <int> <chr>          <dbl>     <dbl>
## 1    11 Iron Man      301.      10144
## 2    31 Thor          291.       3628
## 3    17 Hulk          247.       4283
## 4    24 Abominio      206.       1080
## 5     7 Captain America 200.       4991
## 6    48 Black Widow   148.       1939
## 7    35 Loki          139.       2483
## 8    49 War Machine    108.       1620
## 9    45 Ivan Vanko     107.       1260
## 10   14 Obadiah Stane (Iron Monger) 86.6       1320
## 11     6 Peggy Carter    17.5       1020
## 12   52 Nick Fury       12.1       1138
## 13    16 Pepper Potts     9.63       2672
## 14   54 Philip Coulson     1.60       1064
## 15   20 Betty Ross     -29.5       1620
```



```
## [1] 0.6811509
```

(source screen times: <https://www.youtube.com/watch?v=XE0HwXz43WQ>)

Shiny App

TODO: LINK