Forenic on Gacha Game

Uyen Doan

2025-06-24

Set up

Load Data

The data set used for this is named: "gacha database - 2024.csv"

Goal(s):

To test the accuracy of the notions that: Gacha games developed based on established IPs (Intellectual Properties) have shorter lifespan than stanalone gacha games.

About the data set:

- The data set has 9 variables, but notables are:
    1. Title: The name of the game. All using English names or the romanization of the Languages.
    2. Time: Total time the game had been in service. Counted in years.
    3. Region: The Region the game has servers in.
    4. IP: Whether or not the game was using/based on an already existed IP (Intellectual Property)
    5. Franchise: If the answers to Variable #4 is Y (Yes), what franchise does the game belonged to.
- All chosen games are those that announced End of Service (EoS) in 2024.

Plotting:

The dataset - due to the nature of live-service, multiserver games - contain multiple duplicates as certain games may have different servers open up as

different time. The approach chosen for this is to compare:

- The original with duplicates
- The one grouped together using the mean time of all group's elements
- The one grouped together using the max time of all group's elements

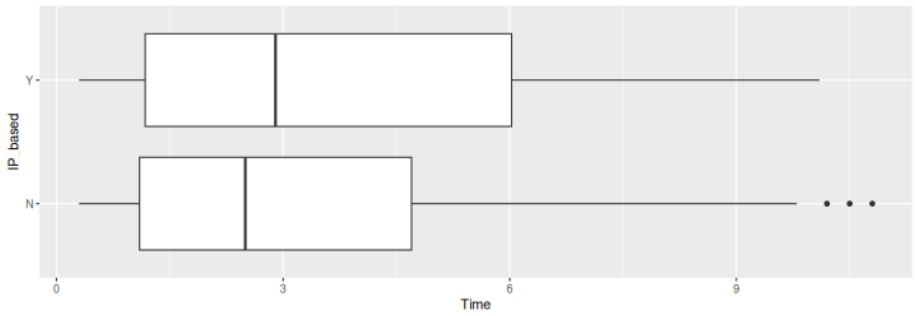

Figure 1: plot of chunk Before cleaning duplicates

```
# A tibble: 2 × 3
  IP_based  Time Standard
  <fct>    <dbl>    <dbl>
1 N         3.39     2.96
2 Y         3.81     2.75
```

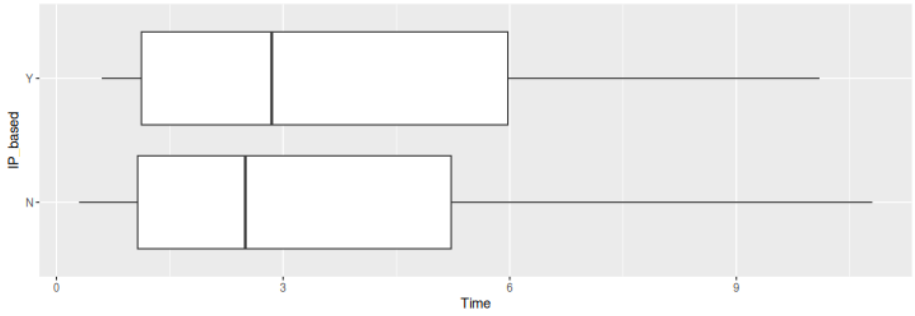

Figure 2: plot of chunk Cleaned up and using the mean Time for duplicates

```
# A tibble: 2 × 3
  IP_based  Time Standard
  <fct>    <dbl>    <dbl>
1 N         3.46     3.20
2 Y         3.73     2.83
```

```
# A tibble: 2 × 3
  IP_based  Time Standard
  <fct>    <dbl>    <dbl>
```

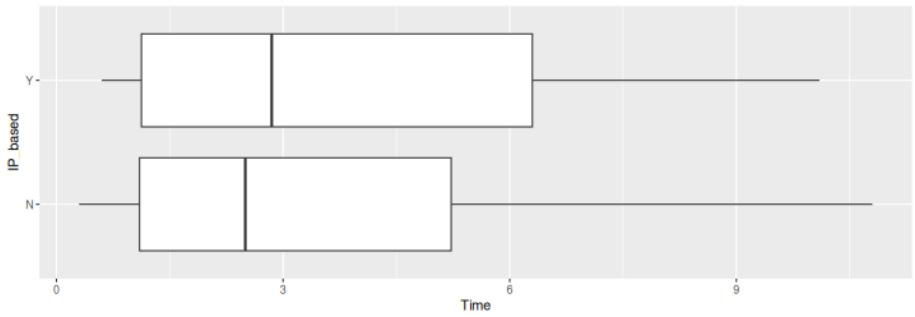

Figure 3: plot of chunk Cleaned up and using the max Time for duplicates

```
1 N          3.53     3.21
2 Y          3.80     2.87
```

- We can see that, in contrast to the notion, in all case:
  - IP-based gacha games either have higher or almost the same minimum values.
  - IP-baseds on average have higher life span.
  - IP-baseds have higher median.
- It should be noted that:
  - The original data frame have the 3 outliers to the right on the Standalones.
  - Standalones have much longer Q3 and higher maximum values in all cases.
  - IP-baseds have longer boxes in all 3 cases.
  - All cases, both catagories have almost the same length on Q1.

Proportion plot

Hypothesis:

Gacha games developed based on established IPs do not generally have shorter lifespan than stanalone gacha games. This notion may stem from the fact that: Standalone games have a fews with long life span. These games could earn fame throughout their services, leading to some sort of survivorship bias.

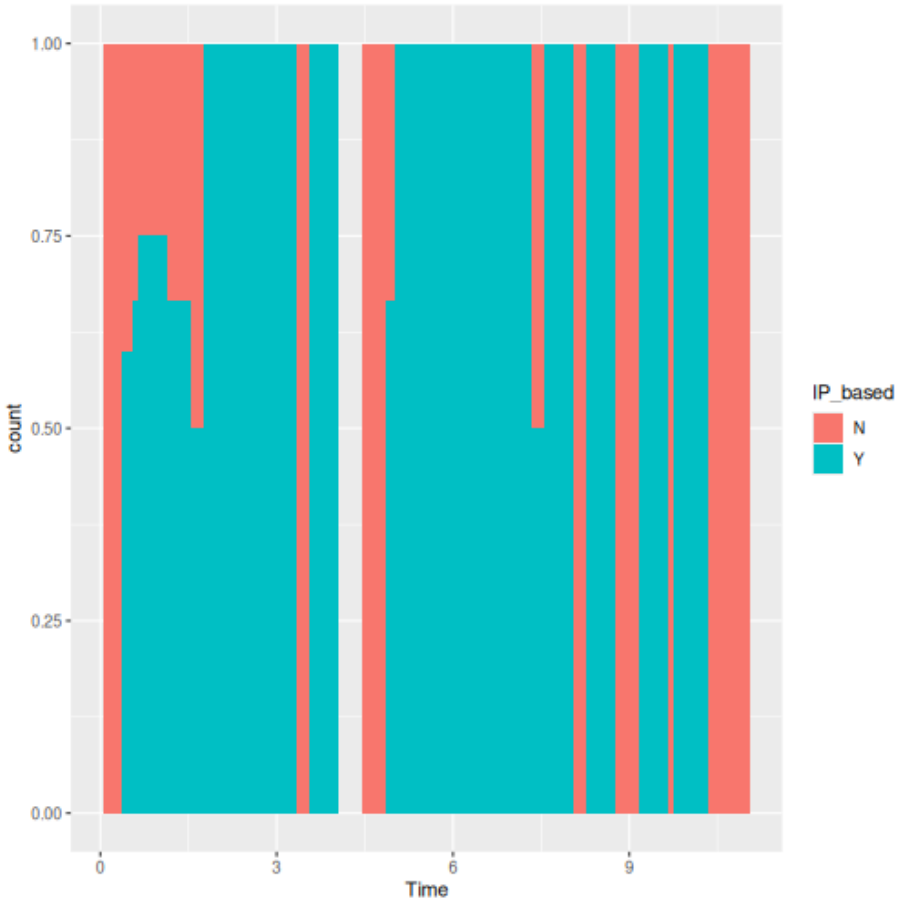

Figure 4: plot of chunk unnamed-chunk-1