

Board Games

Gary Lu

1/25/2022

The data is sourced from R for Data Science Github

Loading TidyTuesday Data

(Note: Data can be joined together by the id column)

Cleaning data

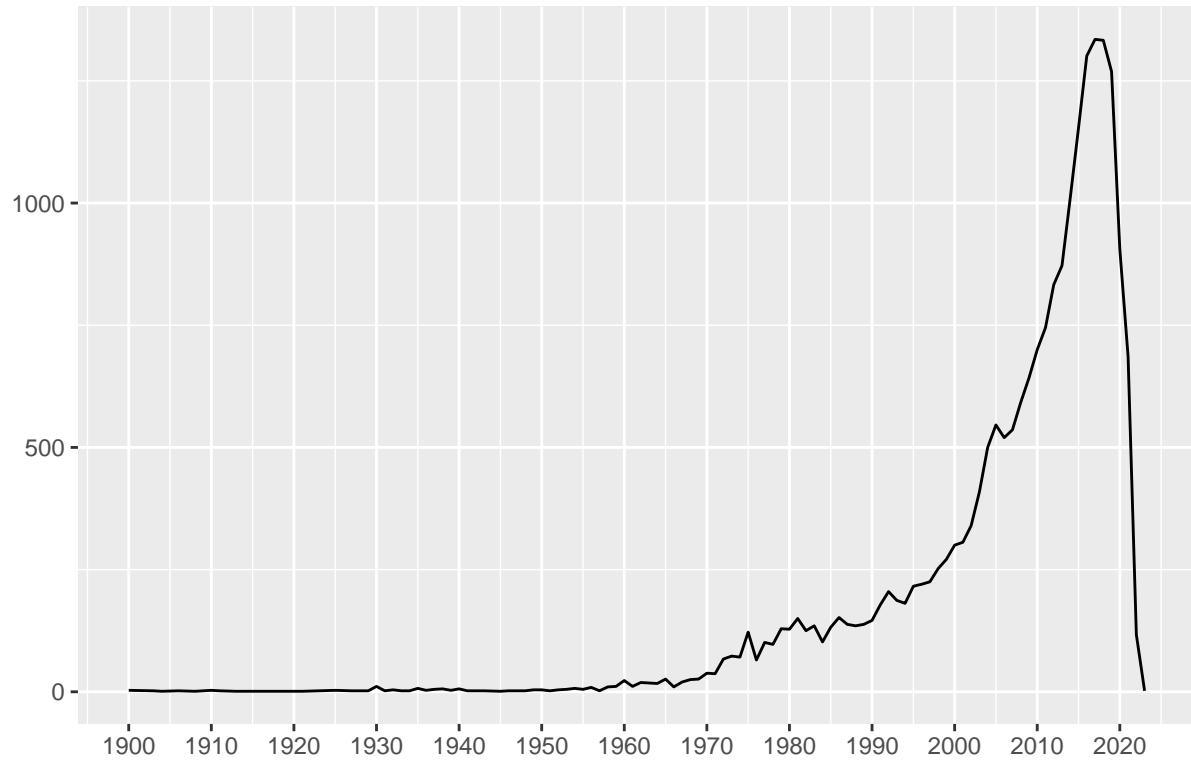
Looking at the highest ranked games

```
## # A tibble: 20 x 10
##       num      id name     year   rank average bayes_average users_rated url
##     <dbl>    <dbl> <chr>    <dbl> <dbl>    <dbl>        <dbl>    <dbl> <chr>
## 1      0 174430 Gloomhav~  2017     1    8.74        8.51    47827 /boardg~
## 2      1 161936 Pandemic~  2015     2    8.59        8.44    45041 /boardg~
## 3      2 224517 Brass: B~  2018     3    8.66        8.42    25484 /boardg~
## 4      3 167791 Terrafor~  2016     4    8.42        8.27    74216 /boardg~
## 5      4 233078 Twilight~  2017     5    8.68        8.26    16025 /boardg~
## 6      5 291457 Gloomhav~  2020     6    8.68        8.26    15918 /boardg~
## 7      6 220308 Gaia Pro~  2017     7    8.47        8.18    19169 /boardg~
## 8      7 187645 Star War~  2016     8    8.42        8.17    25586 /boardg~
## 9      8 182028 Through ~  2015     9    8.38        8.15    25605 /boardg~
## 10     9 115746 War of t~  2012    10    8.51        8.14    15498 /boardg~
## 11    10 162886 Spirit I~  2017    11    8.36        8.13    32199 /boardg~
## 12    11 193738 Great We~  2016    12    8.29        8.12    32960 /boardg~
## 13    12 12333 Twilight~  2005    13    8.28        8.11    43052 /boardg~
## 14    13 169786 Scythe     2016    14    8.22        8.06    64569 /boardg~
## 15    14 84876 The Cast~  2011    15    8.13        8.01    50549 /boardg~
## 16    15 173346 7 Wonder~  2015    16    8.11        7.98    69472 /boardg~
## 17    16 167355 Nemesis    2018    17    8.39        7.98    17688 /boardg~
## 18    17 120677 Terra My~  2012    18    8.12        7.97    41806 /boardg~
## 19    18 124361 Concordia  2013    19    8.12        7.97    31519 /boardg~
## 20    19 28720 Brass: L~  2007    20    8.18        7.96    20915 /boardg~
## # ... with 1 more variable: thumbnail <chr>
```

We see that the website ranks are based on Bayes average and not average (The average rating on the website the data was from)

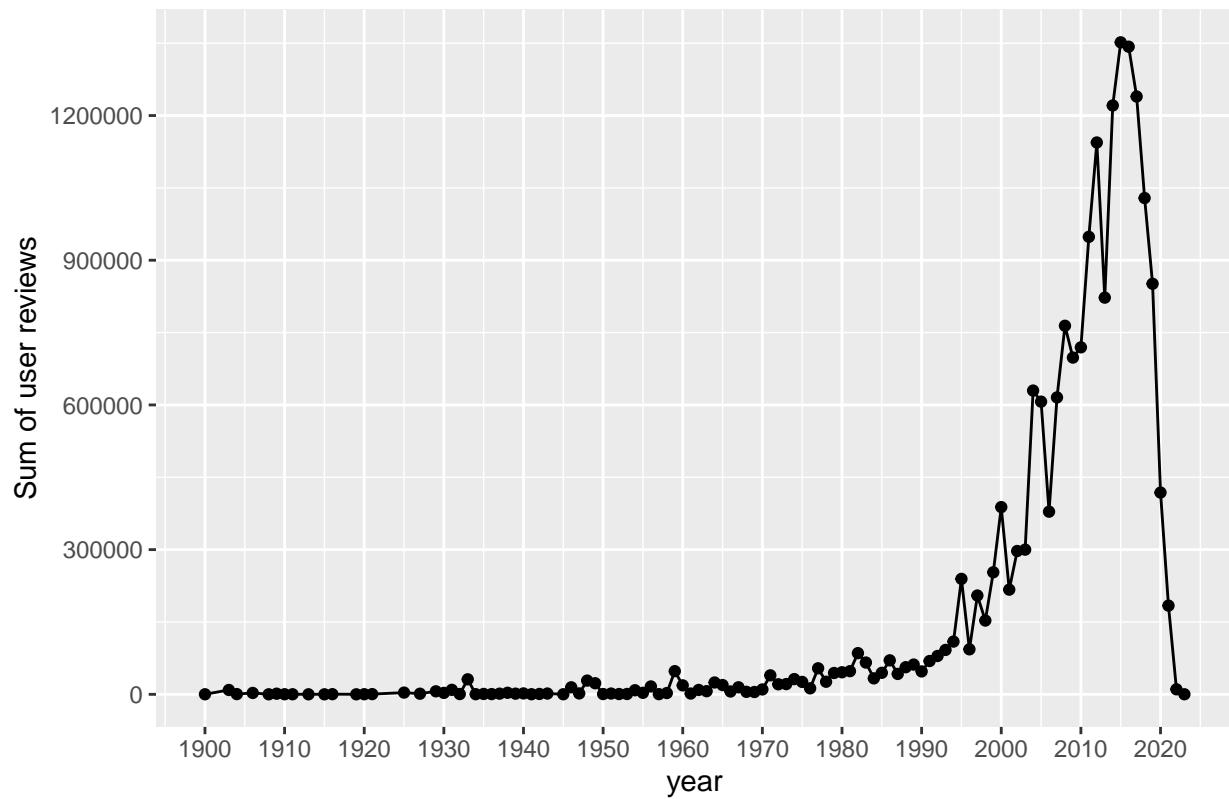
Checking to see how things have changed over the years

Number of games released per year



Plotting number of ratings per year

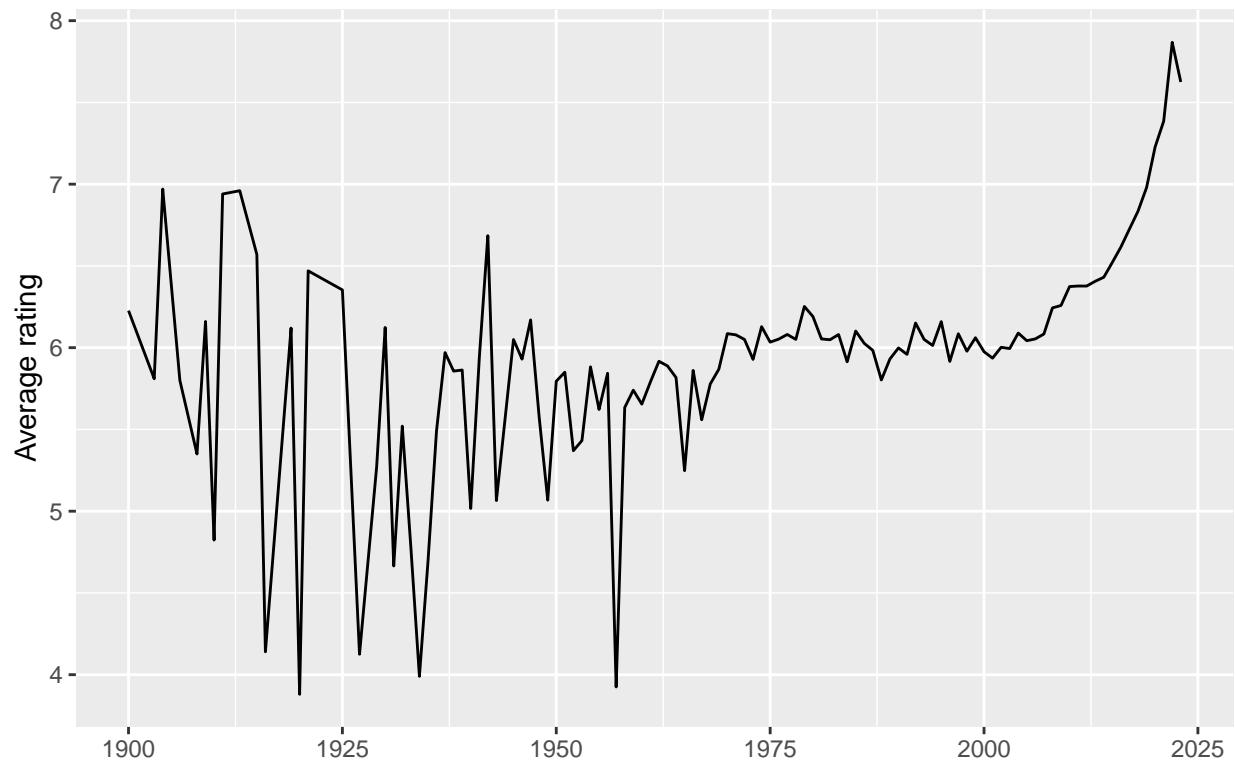
Number of user reviews per game's release year



Here we see board games gaining popularity over time, up until 2015. This correlates to the previous plot where there are less games that were made after 2015

Checking the ratings over time

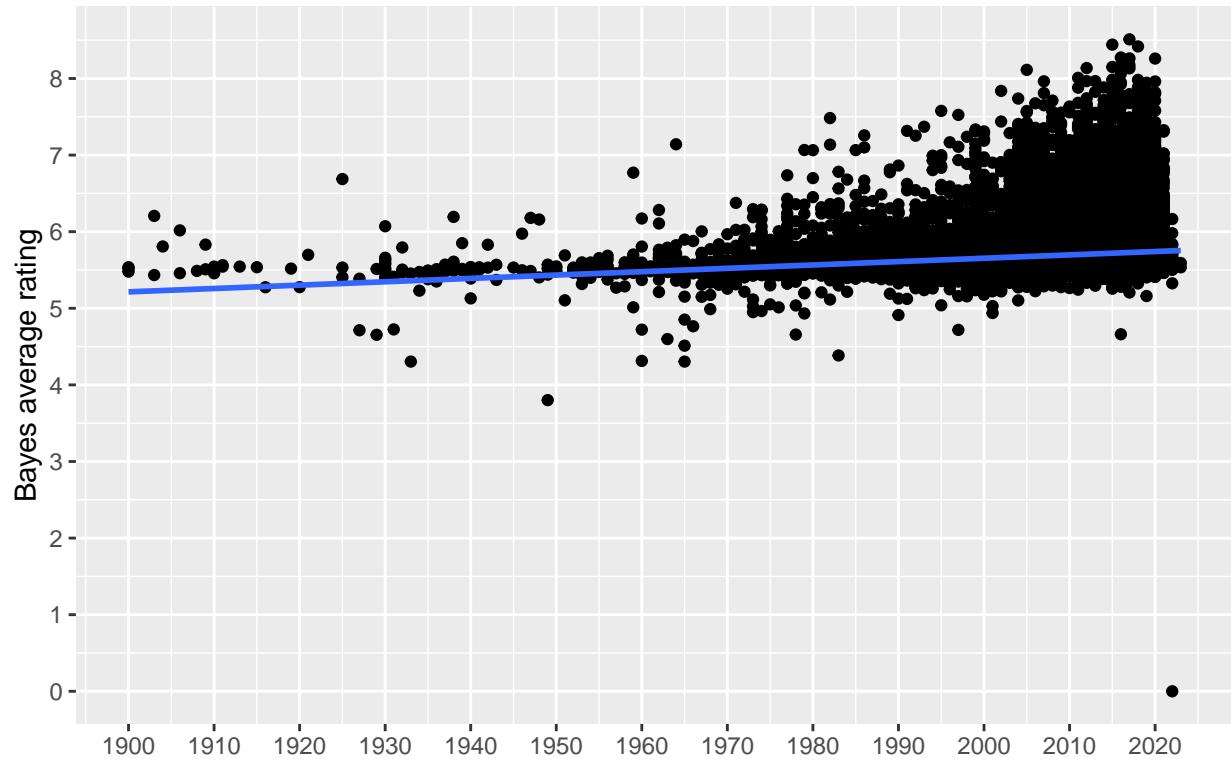
Average user rating of board games per release year



We see the average review go up as well as the number of reviews going up as newer board games come out

Spread of ratings

Spread of average ratings

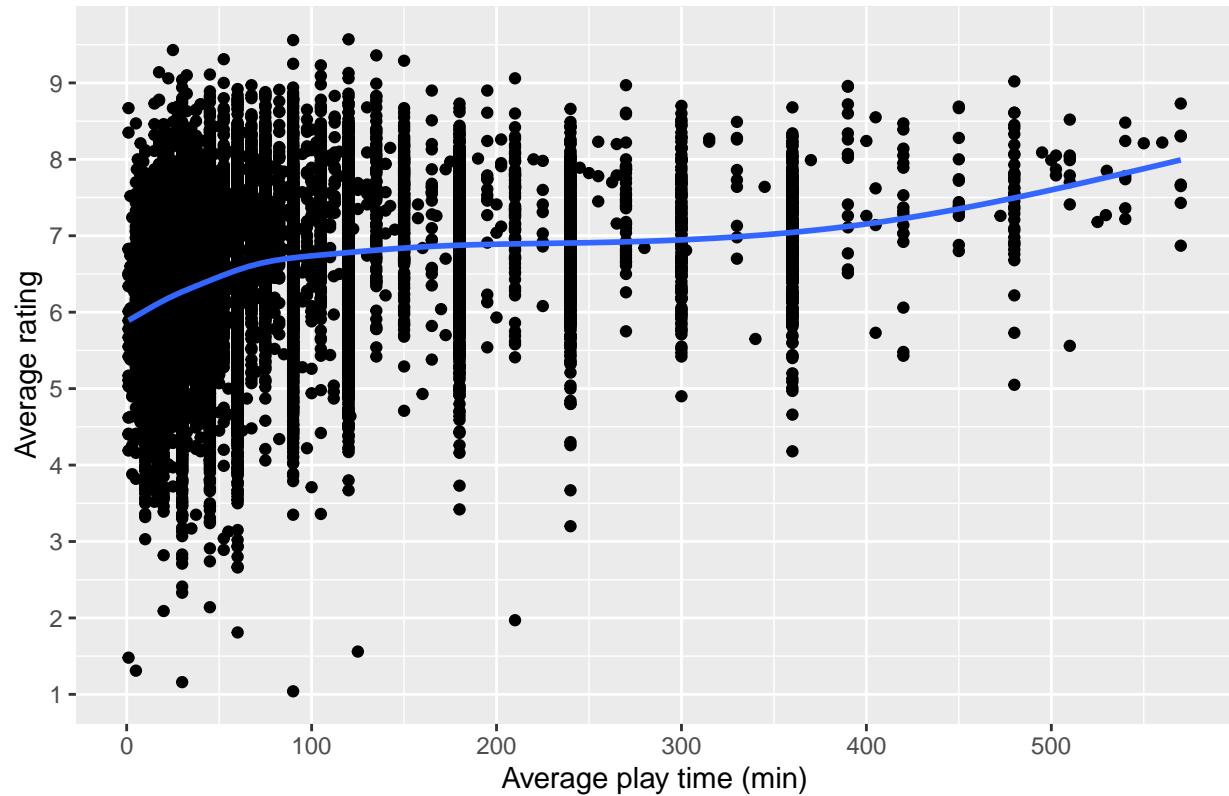


Notes so far:

- Games released in the later years are:
 - More popular (more users are reviewing them)
 - Higher rated
- Less games released from 2016 and on

Joining the datasets

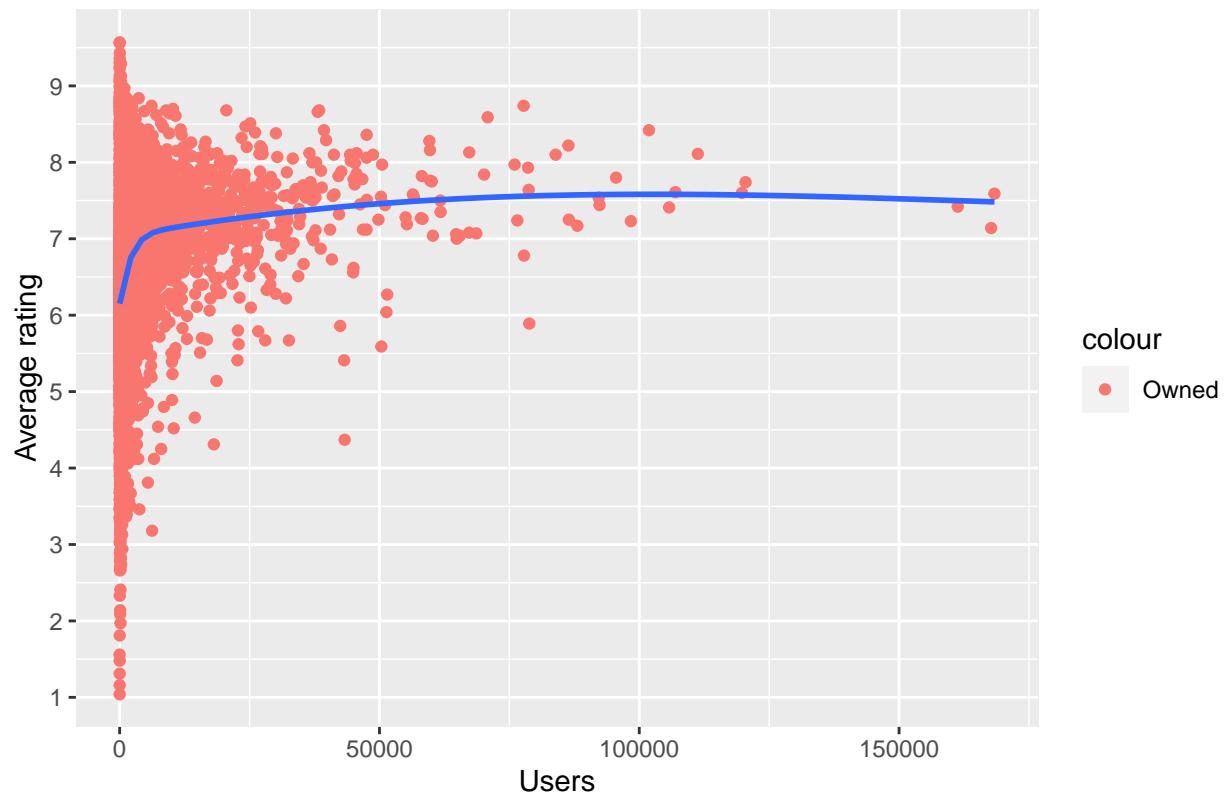
Average play time vs User Average Rating



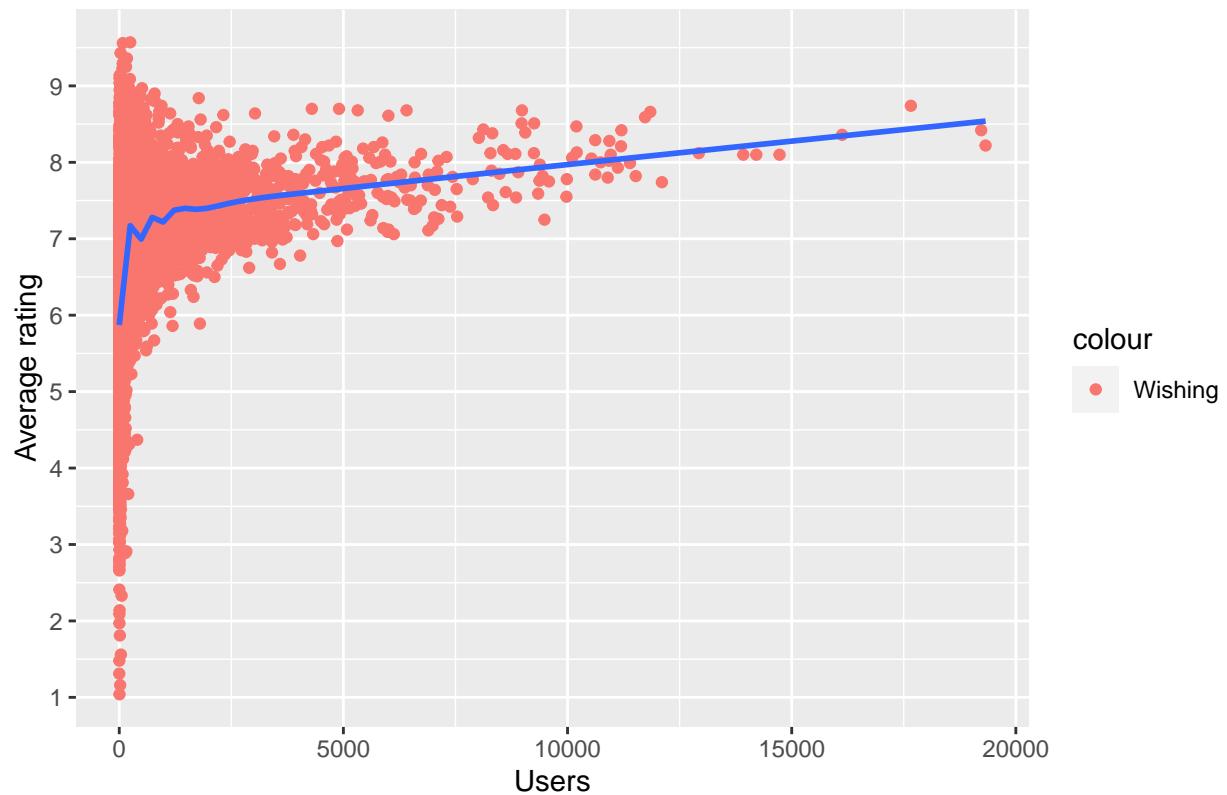
We see that there are a lot more shorter games than longer games, which brings the average rating down

Is there a difference in ratings depending on who owns the game compared to those that want or wish for it?

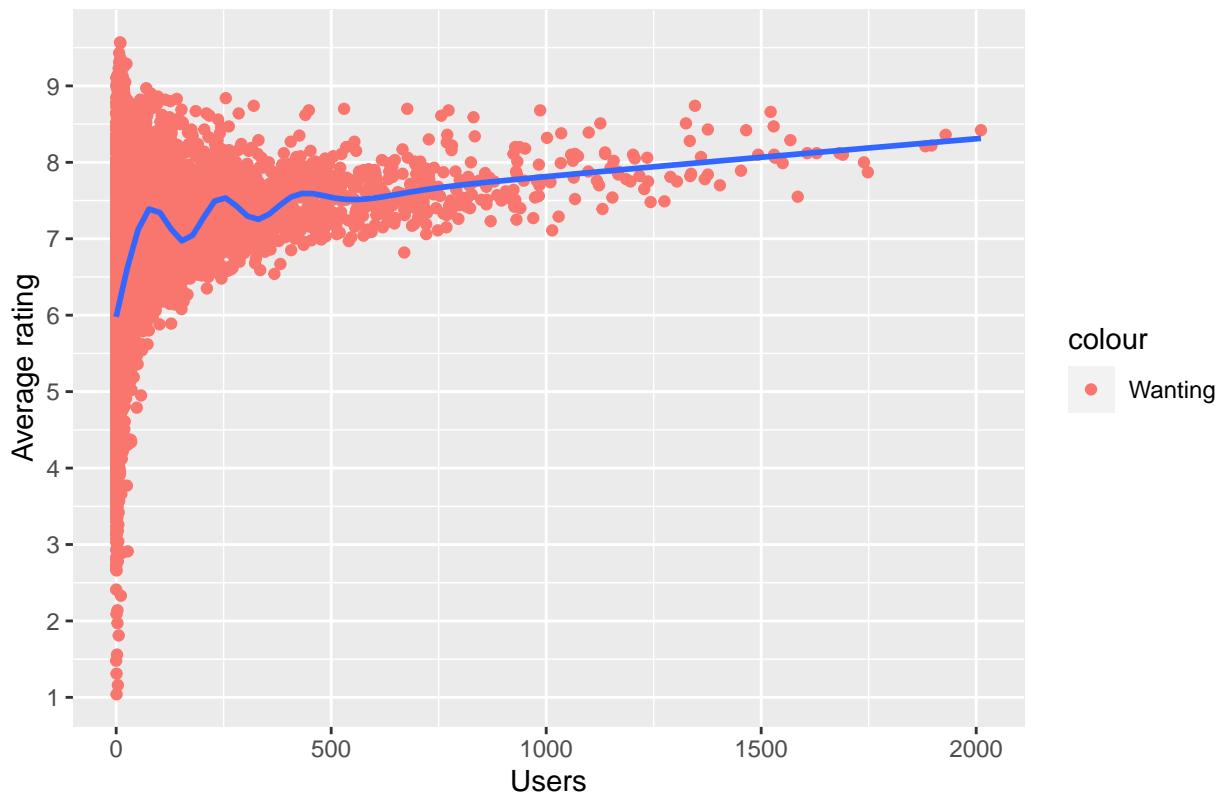
Game ownership vs User Average Rating



Game ownership vs User Average Rating



Game ownership vs User Average Rating



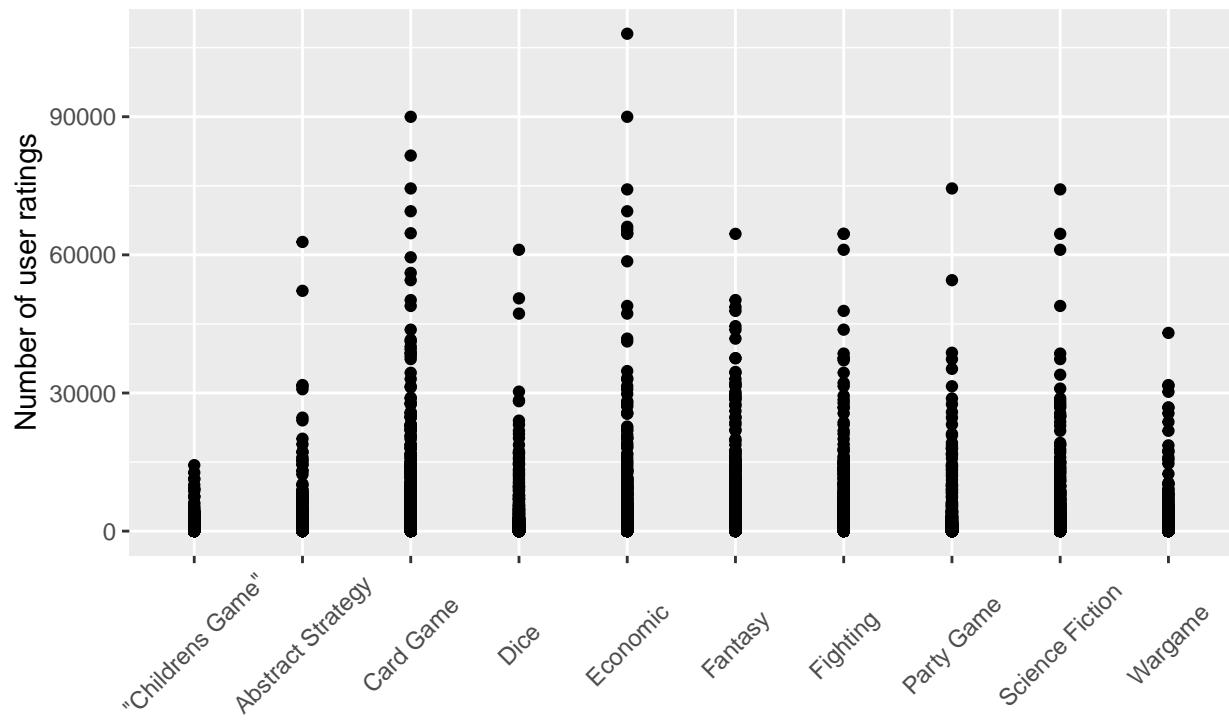
People that do not own the game seem to rate games higher than those that own it

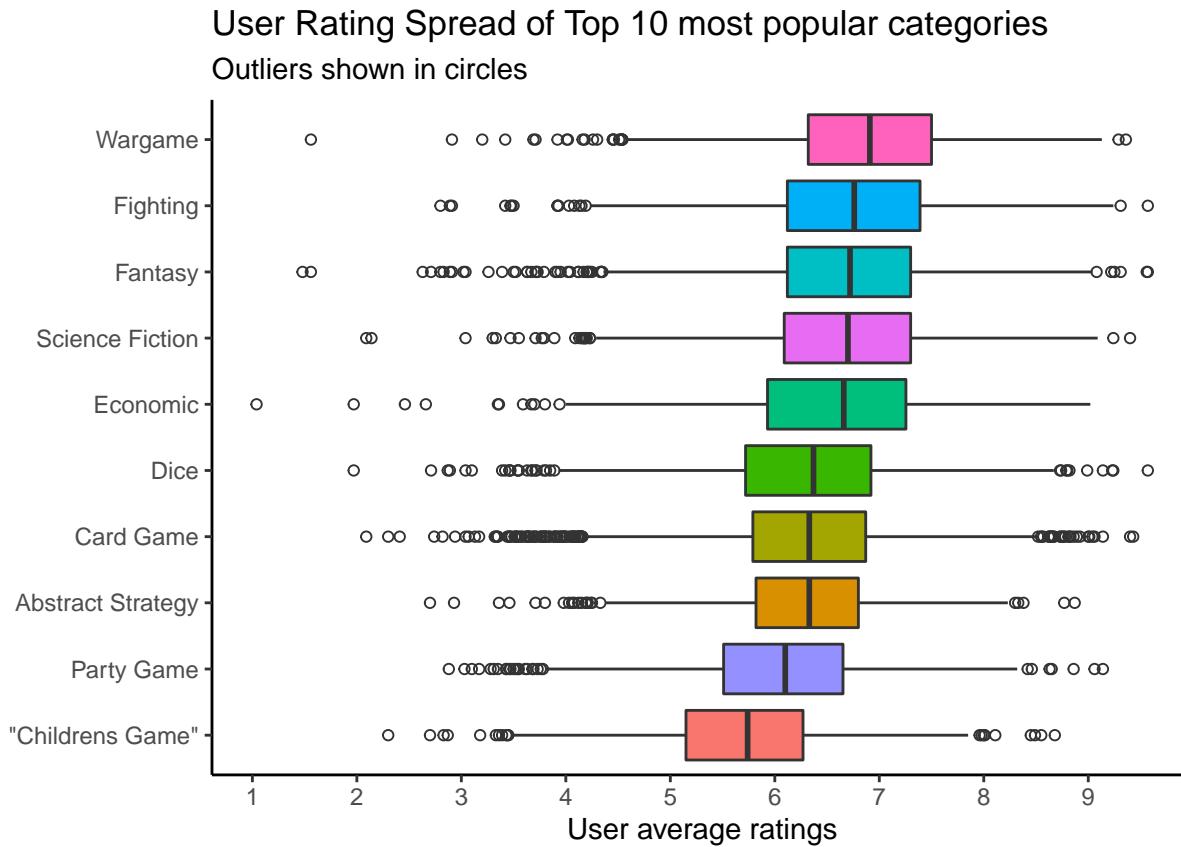
Checking if category has an effect on ratings

Overview of the top 10 categories

Review Spread of Top 10 most popular categories

Points represent an individual game





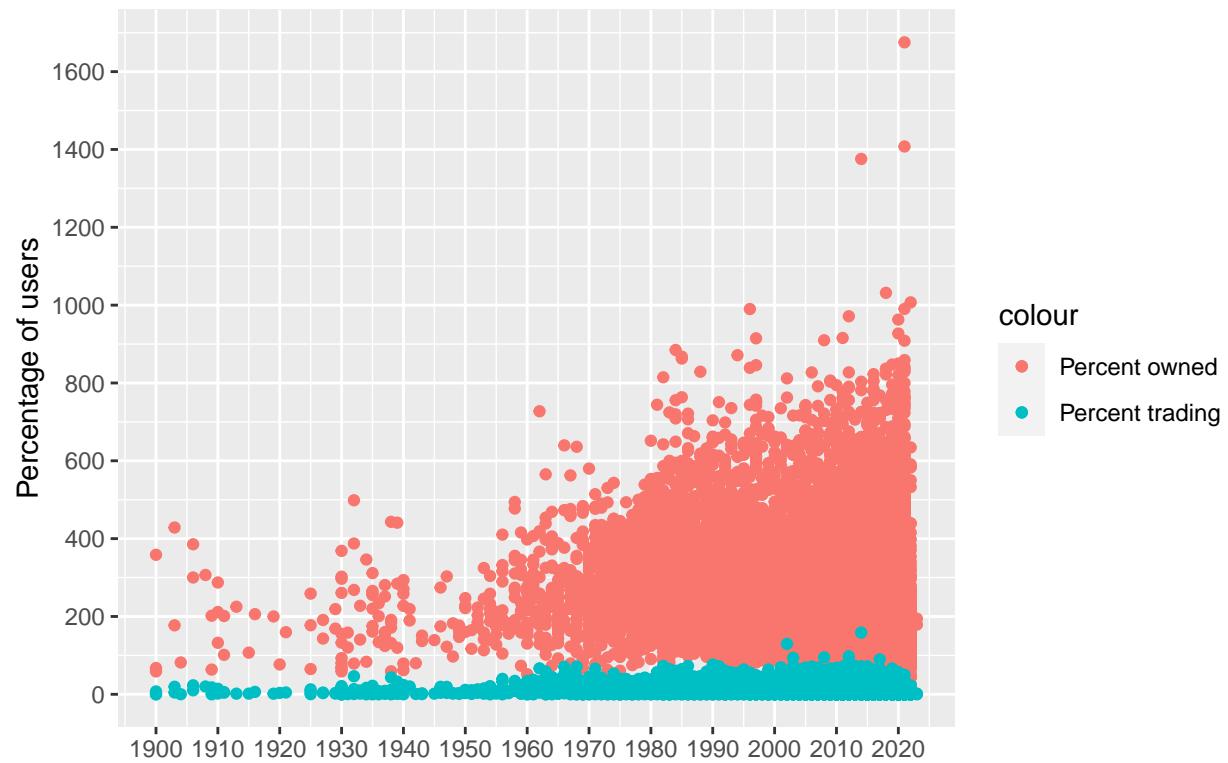
War games, Fighting, Fantasy, Sci-Fi and Economic games average higher than the rest of the top 10 categories

Checking differences based on game ownership

```
## Warning: Removed 191 rows containing missing values (geom_point).
```

```
## Warning: Removed 191 rows containing missing values (geom_point).
```

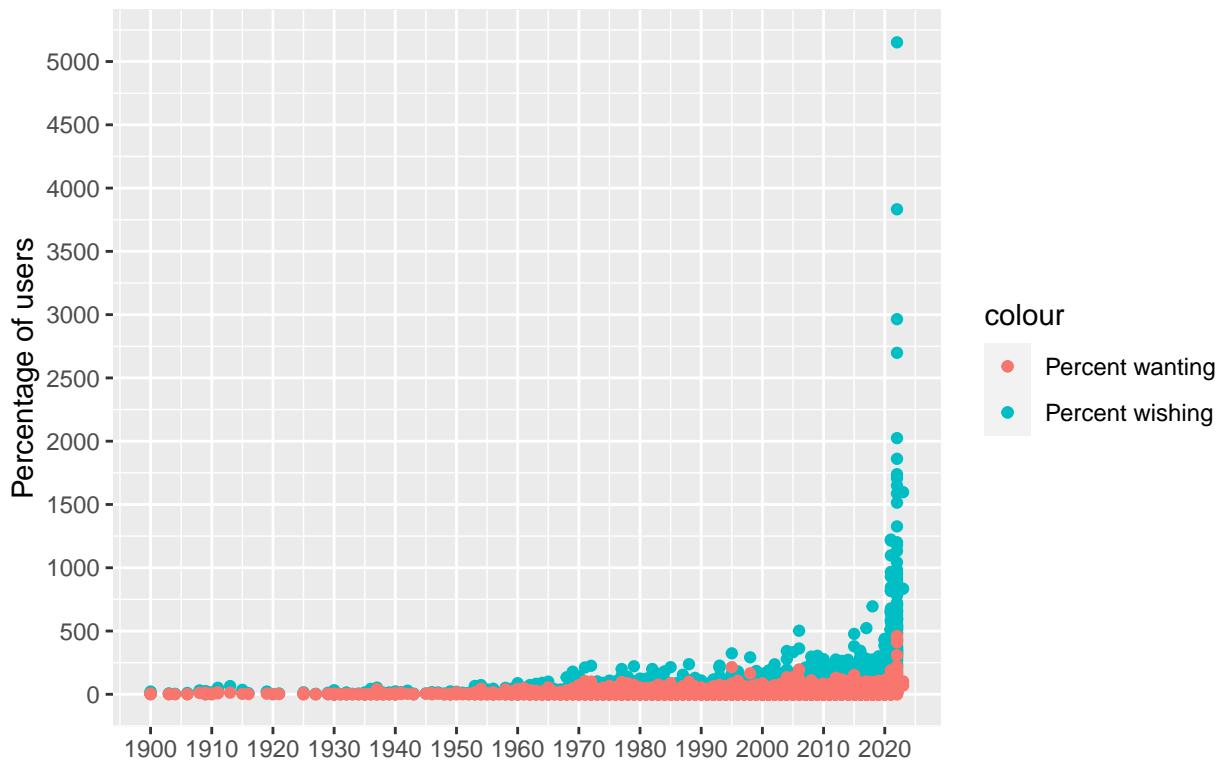
Game ownership / number of ratings



```
## Warning: Removed 191 rows containing missing values (geom_point).
```

```
## Warning: Removed 191 rows containing missing values (geom_point).
```

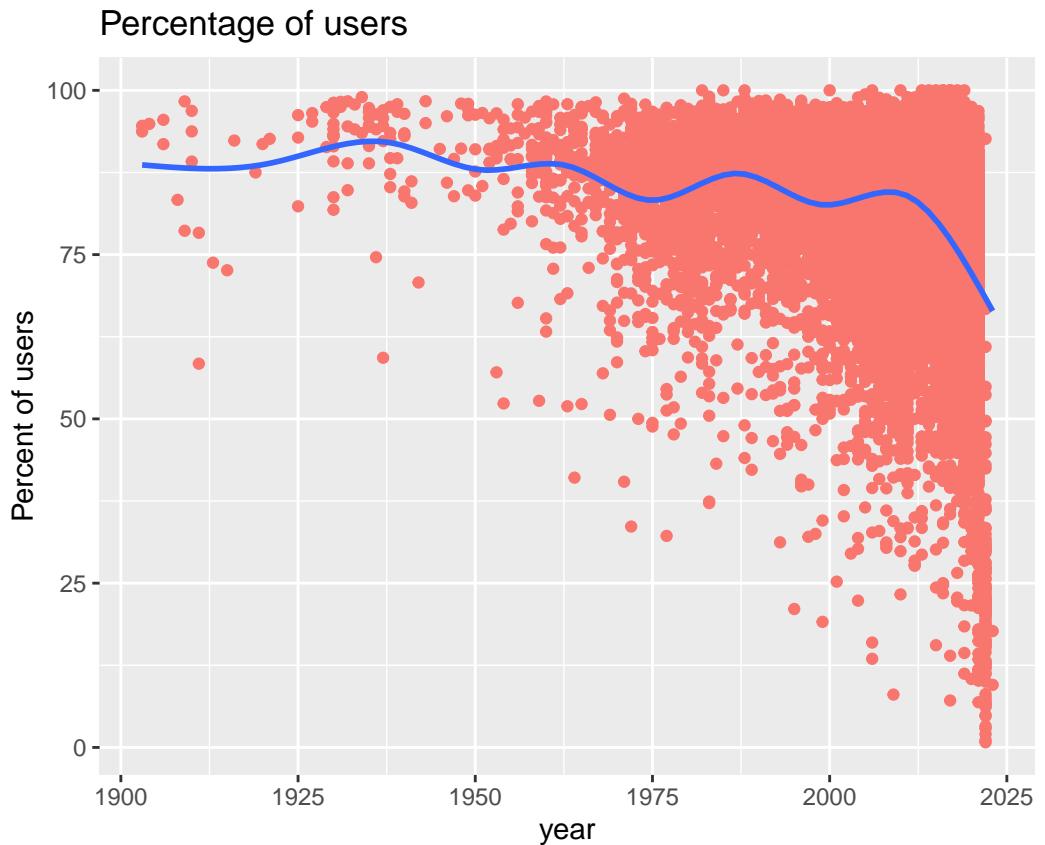
Game ownership / number of ratings



There are a lot points that are over 100 percent especially in the wishing and owning categories which can state that reviewers are more likely to state their ownership of a game rather than rate it

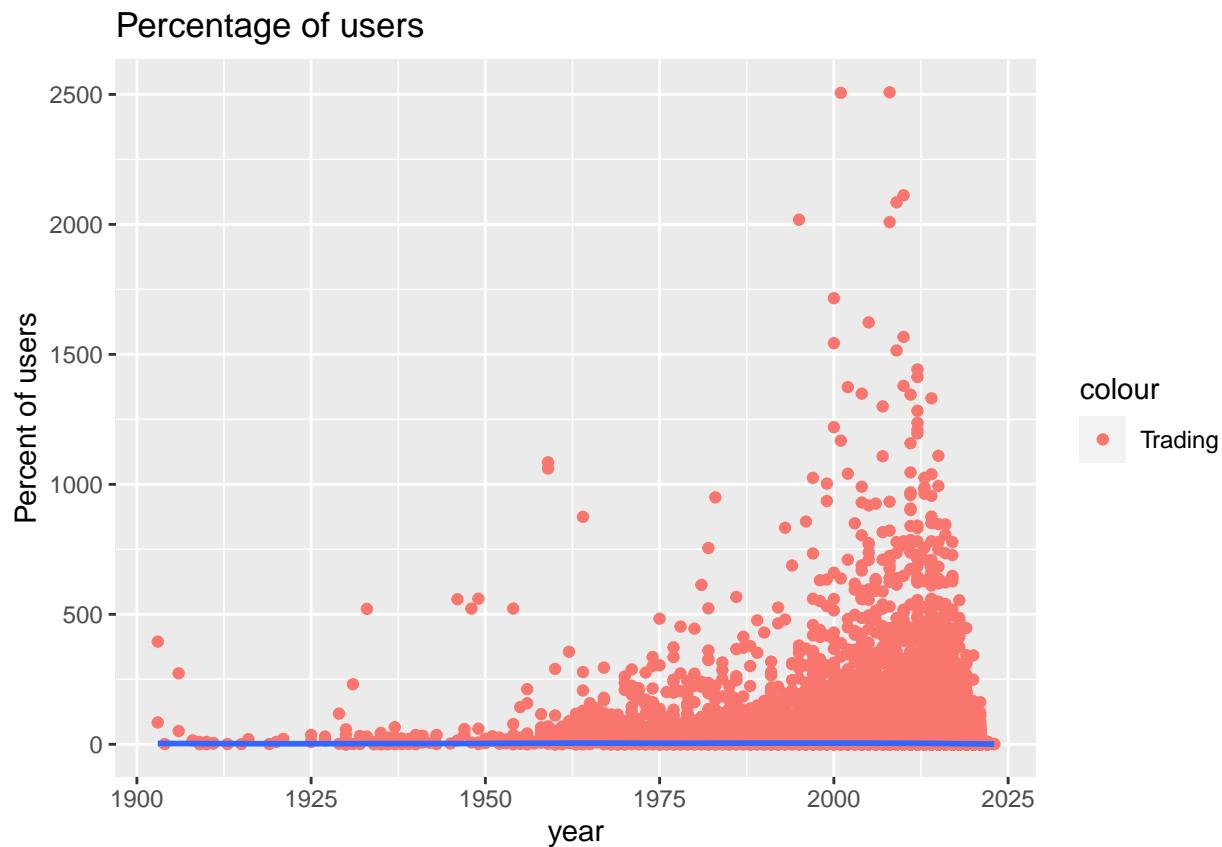
```
## Warning: Removed 191 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 191 rows containing missing values (geom_point).
```



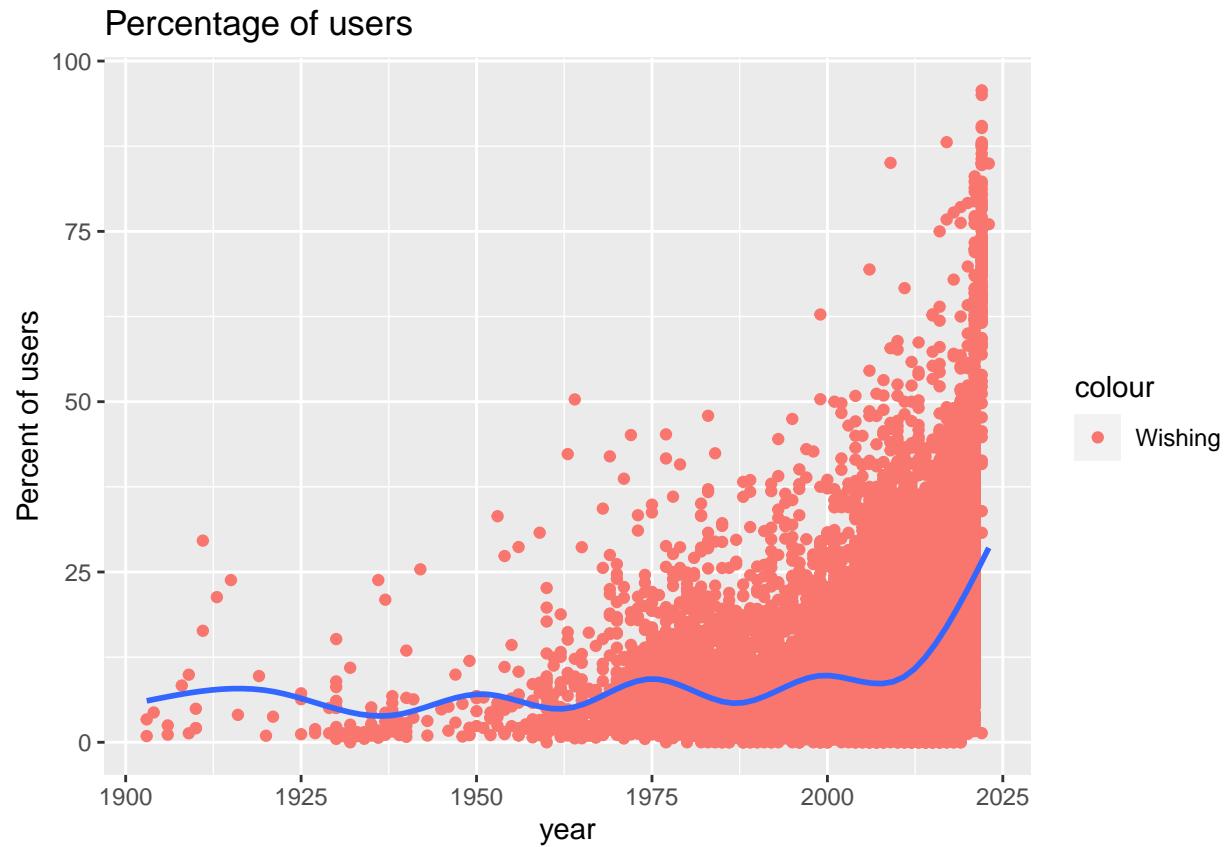
```
## Warning: Removed 191 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 191 rows containing missing values (geom_point).
```



```
## Warning: Removed 191 rows containing non-finite values (stat_smooth).
```

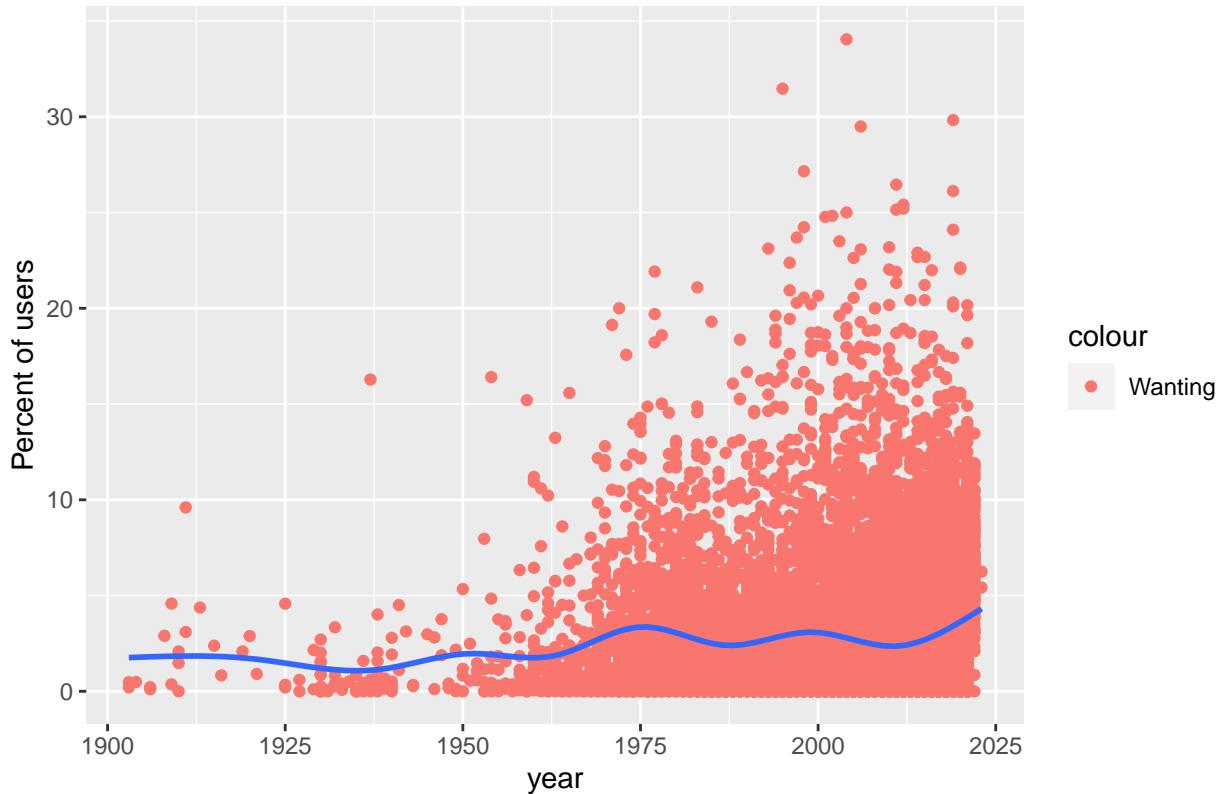
```
## Warning: Removed 191 rows containing missing values (geom_point).
```



```
## Warning: Removed 191 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 191 rows containing missing values (geom_point).
```

Percentage of users



These 4 plots here calculate the percentage of users that own/trading/want/wishing for the game. We see that as newer games come out, less people own them and more people want them i.e. lower supply higher demand

Concluding notes

- Board games have gain increasing popularity
- War games, Fighting, Fantasy, Sci-Fi and Economic games are the most popular genres and highest rated
 - May suggest that the competitive ones are more preferred by consumers
- Most games will have an average play time of < 2 hours
- The number of ratings is very different than the number of ownership status, for people that state that they own the game
 - This suggests that people are far less likely to rate something that they buy