



Universitat
de les Illes Balears

21746 - Data Mining

Final Project

Steam Successful Indie Games Study

Iván Pulgar Rodas
Jordi Sevilla Marí
Nahuel Vazquez
Yelyzaveta Denysova
Xiaozhe Cheng
Gabriel Oliver Artigues
Arturo Mus Mejías

Contents

1	Introduction	1
1.1	Explanation of the Attributes	1
1.2	Objectives	1
2	Procesing de data	1
2.1	Handling of NA values	4
3	Exploratory Data Analysis	6
4	Requeriments for Successful Indie Games	12
4.1	Obtain Unique values for different columns	12
5	Can a single game have enough influence to make other games have its tag?	18
5.1	Game study: Slay the Spire (Roguelike Deckbuilder)	18
5.2	Game Study: The Binding Of Isaac + The Binding Of Isaac Rebirth (Roguelike)	21

1 Introduction

1.1 Explanation of the Attributes

The dataset we will be working with contains a total number of 94948 observations and 47 columns or variables. The columns that will be used are described below:

- **appid**: Unique identifier of the game on Steam. [num]
- **name**: Name of the game. [text]
- **'released_date'**: Represents the date where the game was released. [time]
- **'required_age'**: Corresponds to the minimum age required to play the game. [num]
- **price**: How much the game costs. If its 0 it means that the game is Free to Play. [num]
- **dlc_count**: Amount of DLCs (Downloadable Contents) the game has. [num]
- **support_url**: URL to the support page of the game. [text]
- **windows**: Determines if the game runs in windows. [categorical]
- **mac**: Determines if the game runs in mac. [categorical]
- **linux**: Determines if the game runs in linux. [categorical]
- **metacritic_score**: Metacritic score based on critical reviews (reviews from professionals). By performing an investigation we think that a score of 0 means that when the scraping of the data was done there were no reviews for that game yet. [num]
- **achievements**: Number of achievements the game has. [num]
- **'recommendations'**: Amount of user recommendations. [num]
- **supported_languages**: List of languages that the game supports. **[NO SÉ QUE CATEGORÍA DARLE]**
- **packages**: Available packages for the game. It contains the name and a description of the package and the names, descriptions and subprices of the subpackages. **[LO MISMO QUE ARRIBA]**
- **Developers**: List of developers associated with the game. **[LO MISMO QUE ARRIBA]**
- **publishers**: List of publishers associated with the game. **[LO MISMO QUE ARRIBA]**
- **categories**: List of categories that the game has. **[LO MISMO QUE ARRIBA]**
- **genres**: List of genres that the game belongs to. **[LO MISMO QUE ARRIBA]**
- **positive**: Amount of positive votes the game has. [num]
- **negative**: Amount of negative votes the game has. [num]
- **estimated_owners**: Estimated owners of the game. [text]
- **average_playtime_forever**: Average playtime since March 2009 measured in minutes. [num]
- **average_playtime_2weeks**: Average playtime in the last two weeks measured in minutes. [num]
- **median_playtime_forever**: Median playtime since March 2009 measured in minutes. [num]
- **median_playtime_2weeks**: Median playtime in the last two weeks measured in minutes. [num]
- **peak_ccu**: Number of current users playing the day before the data was scrapped. [num]
- **tags**: List of tags the game has with its name and its key. **[NO SE QUE CATEGORIA DARLE]**
- **pct_pos_total**: Percentage of all reviews that are positive. [num]
- **num_reviews_total**: Number of the total reviews the game has. [num]

1.2 Objectives

TODO: ADD OBJECTIVES OF STUDY

2 Procesing de data

General look of the dataset

```
summary(steam)
```

```
##      appid          name       release_date   required_age
##  Min.   :    20  Length:94948   Length:94948   Min.   :-1.0000
##  1st Qu.: 887338 Class :character  Class :character  1st Qu.: 0.0000
##  Median :1591145 Mode  :character  Mode  :character  Median : 0.0000
##  Mean   :1707530                    Mode  :character  Mean   : 0.1783
##  3rd Qu.:2491702                    Mode  :character  3rd Qu.: 0.0000
##  Max.   :3570420                    Mode  :character  Max.   :21.0000
##
##      price          dlc_count  detailed_description about_the_game
##  Min.   : 0.000  Min.   : 0.0000  Length:94948   Length:94948
##  1st Qu.: 0.990  1st Qu.: 0.0000  Class :character  Class :character
##  Median : 3.990  Median : 0.0000  Mode  :character  Mode  :character
##  Mean   : 6.911  Mean   : 0.5632
##  3rd Qu.: 9.990  3rd Qu.: 0.0000
##  Max.   :999.980  Max.   :3427.0000
##
##      short_description reviews      header_image      website
##  Length:94948   Length:94948   Length:94948   Length:94948
##  Class :character Class :character Class :character Class :character
##  Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##      support_url      support_email      windows      mac
##  Length:94948   Length:94948   Length:94948   Length:94948
##  Class :character Class :character Class :character Class :character
##  Mode  :character  Mode  :character  Mode  :character  Mode  :character
##
##      linux          metacritic_score metacritic_url      achievements
##  Length:94948   Min.   : 0.000  Length:94948   Min.   : 0.00
##  Class :character 1st Qu.: 0.000  Class :character  1st Qu.: 0.00
##  Mode  :character Median : 0.000  Mode  :character  Median : 2.00
##                      Mean   : 2.764                    Mean   : 19.54
##                      3rd Qu.: 0.000                    3rd Qu.: 19.00
##                      Max.   :97.000                    Max.   :9821.00
##
##      recommendations      notes      supported_languages full_audio_languages
##  Min.   :    0  Length:94948   Length:94948   Length:94948
##  1st Qu.:    0  Class :character  Class :character  Class :character
##  Median :    0  Mode  :character  Mode  :character  Mode  :character
##  Mean   : 1022
##  3rd Qu.:    0
##  Max.   :4401572
##
##      packages      developers      publishers      categories
##  Length:94948   Length:94948   Length:94948   Length:94948
##  Class :character Class :character  Class :character  Class :character
```

```

##   Mode :character   Mode :character   Mode :character   Mode :character
##
##
##
##
##      genres           screenshots          movies        user_score
##  Length:94948       Length:94948       Length:94948      Min.   : 0.00000
##  Class :character   Class :character   Class :character  1st Qu.: 0.00000
##  Mode  :character   Mode  :character   Mode  :character  Median : 0.00000
##                                         Mean   : 0.03097
##                                         3rd Qu.: 0.00000
##                                         Max.   :100.00000
##
##      score_rank      positive      negative estimated_owners
##  Min.   : 98.00   Min.   :     0   Min.   :     0   Length:94948
##  1st Qu.: 99.00   1st Qu.:     0   1st Qu.:     0   Class :character
##  Median : 99.00   Median :     8   Median :     2.0  Mode  :character
##  Mean   : 99.13   Mean   : 1218   Mean   : 202.1
##  3rd Qu.:100.00   3rd Qu.:    51   3rd Qu.:    15.0
##  Max.   :100.00   Max.   :7480813  Max.   :1135108.0
##  NA's   :94909
##  average_playtime_forever average_playtime_2weeks median_playtime_forever
##  Min.   : 0.0       Min.   : 0.000   Min.   : 0.0
##  1st Qu.: 0.0       1st Qu.: 0.000   1st Qu.: 0.0
##  Median : 0.0       Median : 0.000   Median : 0.0
##  Mean   : 108.6    Mean   : 4.757   Mean   : 108.4
##  3rd Qu.: 0.0       3rd Qu.: 0.000   3rd Qu.: 0.0
##  Max.   :1462997.0  Max.   :18568.000  Max.   :1462997.0
##
##      median_playtime_2weeks   discount      peak_ccu
##  Min.   : 0.000   Min.   : 0.000   Min.   :0.000e+00
##  1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.:0.000e+00
##  Median : 0.000   Median : 0.000   Median :0.000e+00
##  Mean   : 5.018   Mean   : 4.307   Mean   :9.285e+01
##  3rd Qu.: 0.000   3rd Qu.: 0.000   3rd Qu.:0.000e+00
##  Max.   :18568.000  Max.   :100.000  Max.   :1.212e+06
##
##      tags           pct_pos_total num_reviews_total pct_pos_recent
##  Length:94948       Min.   :-1.00  Min.   : -1   Min.   :-1.000
##  Class :character   1st Qu.:-1.00  1st Qu.: -1   1st Qu.:-1.000
##  Mode  :character   Median : 58.00  Median : 15   Median : -1.000
##                                         Mean   : 44.63  Mean   :  5.328
##                                         3rd Qu.: 84.00  3rd Qu.: 80   3rd Qu.:-1.000
##                                         Max.   :100.00  Max.   :8632939  Max.   :100.000
##
##      num_reviews_recent
##  Min.   : -1.00
##  1st Qu.: -1.00
##  Median : -1.00
##  Mean   : 16.88
##  3rd Qu.: -1.00
##  Max.   :96473.00
##

```

Explanation of first looks of it, bad formatting, NA's, negative values...

2.1 Handling of NA values

The attributes with missing values are:

```
na_counts <- steam %>% summarise_all(~ sum(is.na(.)))
```

```
print(na_counts)
```

```
##   appid name release_date required_age price dlc_count detailed_description
## 1      0     0           0          0     0       0                   0
##   about_the_game short_description reviews header_image website support_url
## 1            0           0         0         0       0                   0
##   support_email windows mac linux metacritic_score metacritic_url achievements
## 1            0     0   0   0           0           0                   0
##   recommendations notes supported_languages full_audio_languages packages
## 1            0     0           0           0       0                   0
##   developers publishers categories genres screenshots movies user_score
## 1            0     0           0           0       0       0                   0
##   score_rank positive negative estimated_owners average_playtime_forever
## 1      94909      0     0           0           0           0
##   average_playtime_2weeks median_playtime_forever median_playtime_2weeks
## 1            0           0           0           0
##   discount peak_ccu tags pct_pos_total num_reviews_total pct_pos_recent
## 1            0     0   0           0           0           0
##   num_reviews_recent
## 1            0
```

The columns with missing values are as follows:

Los juegos con número de reseñas = -1 creemos que son porque el scrappear ha fallado durante su ejecución. Si vamos a trabajar con las reseñas podríamos decir en la presentación que vamos a probar de volver a intentar scrappear la información

```
## # tibble [88,982 x 30] (S3: tbl_df/tbl/data.frame)
## $ appid                  : int [1:88982] 2556940 449940 1287250 866510 ...
## $ name                   : chr [1:88982] "! Shakabula *" "! That Bastard Is Trying To Steal Our Go...
## $ release_date            : Date[1:88982], format: "2023-10-13" "2016-03-03" ...
## $ required_age            : int [1:88982] 0 0 0 0 0 0 0 0 0 ...
## $ price                  : num [1:88982] 14.99 2.99 19.99 1.99 0.99 ...
## $ dlc_count               : int [1:88982] 0 0 0 39 0 0 0 1 3 0 ...
## $ windows                 : chr [1:88982] "True" "True" "True" "True" ...
## $ mac                     : chr [1:88982] "False" "False" "False" "False" ...
## $ linux                   : chr [1:88982] "False" "True" "False" "False" ...
## $ metacritic_score        : int [1:88982] 0 0 0 0 0 0 0 0 0 ...
## $ achievements             : int [1:88982] 0 0 9 4997 2021 0 0 19 13 5 ...
## $ recommendations          : int [1:88982] 0 0 0 495 0 0 0 108 0 0 ...
## $ supported_languages      : chr [1:88982] "[{'English'}]" "[{'English'}]" "[{'English', 'Simplified Chinese'}]" ...
## $ packages                : chr [1:88982] "[{'title': 'Buy ! Shakabula *', 'description': '', 'subs...
## $ developers               : chr [1:88982] "[{'Skermunkel'}]" "[{'WTFOMGames'}]" "[{'Andreev Worlds'}]" "[{'...
## $ publishers              : chr [1:88982] "[{'Skermunkel'}]" "[{'WTFOMGames'}]" "[{'Andreev Worlds'}]" "[{'...
## $ categories              : chr [1:88982] "[{'Single-player', 'Full controller support', 'Steam Cloud...]
```

```

## $ genres : chr [1:88982] "[‘Action’, ‘Indie’, ‘RPG’, ‘Early Access’]" "[‘Action’, “Indie”, “RPG”, “Early Access”]"
## $ positive : int [1:88982] 0 57 45 410 25 83 37 126 0 0 ...
## $ negative : int [1:88982] 4 78 34 180 32 18 102 10 0 0 ...
## $ average_playtime_forever: int [1:88982] 0 312 0 360 0 0 244 0 0 0 ...
## $ average_playtime_2weeks : int [1:88982] 0 0 0 0 0 0 0 0 0 0 ...
## $ median_playtime_forever : int [1:88982] 0 391 0 378 0 0 244 0 0 0 ...
## $ median_playtime_2weeks : int [1:88982] 0 0 0 0 0 0 0 0 0 0 ...
## $ peak_ccu : int [1:88982] 0 0 0 6 0 0 0 0 0 0 ...
## $ tags : chr [1:88982] “{‘Early Access’: 213, ‘Action’: 193, ‘RPG’: 187, ‘JRPG’: 165, ‘Strategy’: 152, ‘Puzzle’: 148, ‘Sports’: 145, ‘Horror’: 142, ‘Adventure’: 139, ‘Fiction’: 136, ‘Science Fiction’: 133, ‘Mystery’: 130, ‘Thriller’: 127, ‘Drama’: 124, ‘Romance’: 121, ‘Fantasy’: 118, ‘Horror’: 115, ‘Science Fiction’: 112, ‘Mystery’: 109, ‘Thriller’: 106, ‘Drama’: 103, ‘Romance’: 100, ‘Fantasy’: 97, ‘Strategy’: 94, ‘Puzzle’: 91, ‘Sports’: 88, ‘Adventure’: 85, ‘Fiction’: 82, ‘Science Fiction’: 80, ‘Mystery’: 77, ‘Thriller’: 74, ‘Drama’: 71, ‘Romance’: 68, ‘Fantasy’: 65, ‘Strategy’: 62, ‘Puzzle’: 60, ‘Sports’: 57, ‘Adventure’: 54, ‘Fiction’: 51, ‘Science Fiction’: 48, ‘Mystery’: 45, ‘Thriller’: 42, ‘Drama’: 40, ‘Romance’: 37, ‘Fantasy’: 34, ‘Strategy’: 31, ‘Puzzle’: 28, ‘Sports’: 25, ‘Adventure’: 22, ‘Fiction’: 20, ‘Science Fiction’: 18, ‘Mystery’: 15, ‘Thriller’: 13, ‘Drama’: 11, ‘Romance’: 9, ‘Fantasy’: 7, ‘Strategy’: 6, ‘Puzzle’: 5, ‘Sports’: 4, ‘Adventure’: 3, ‘Fiction’: 2, ‘Science Fiction’: 1}”
## $ pct_pos_total : int [1:88982] -1 55 61 71 55 82 55 91 95 66 ...
## $ num_reviews_total : int [1:88982] -1 68 62 495 18 101 20 108 281 12 ...
## $ estimated_owners_min : int [1:88982] 0 50000 0 100000 0 0 100000 20000 0 0 ...
## $ estimated_owners_max : int [1:88982] 20000 100000 20000 200000 20000 20000 200000 50000 0 0 ...

##      appid          name       release_date   required_age
## Min. : 20  Length:88982    Min. :1997-06-30  Min. :-1.0000
## 1st Qu.: 852782 Class :character  1st Qu.:2018-12-05  1st Qu.: 0.0000
## Median :1522535 Mode  :character  Median :2021-10-29  Median : 0.0000
## Mean   :1655079                  Mean   :2021-04-04  Mean   : 0.1826
## 3rd Qu.:2429338                  3rd Qu.:2023-12-14  3rd Qu.: 0.0000
## Max.  :3542350                  Max. :2025-03-10  Max. :21.0000

##      price        dlc_count       windows        mac
## Min. : 0.000  Min. : 0.0000  Length:88982  Length:88982
## 1st Qu.: 0.990  1st Qu.: 0.0000  Class :character  Class :character
## Median : 4.990  Median : 0.0000  Mode  :character  Mode  :character
## Mean   : 7.349  Mean   : 0.5957
## 3rd Qu.: 9.990  3rd Qu.: 0.0000
## Max.  :999.980  Max. :3427.0000

##      linux      metacritic_score achievements recommendations
## Length:88982  Min. : 0.00  Min. : 0.00  Min. : 0
## Class :character  1st Qu.: 0.00  1st Qu.: 0.00  1st Qu.: 0
## Mode  :character  Median : 0.00  Median : 5.00  Median : 0
## Mean   : 2.91  Mean   : 20.67  Mean   : 1013
## 3rd Qu.: 0.00  3rd Qu.: 20.00  3rd Qu.: 0
## Max.  :97.00  Max. :9821.00  Max. :4401572

##      supported_languages packages      developers      publishers
## Length:88982  Length:88982  Length:88982  Length:88982
## Class :character  Class :character  Class :character  Class :character
## Mode  :character  Mode  :character  Mode  :character  Mode  :character
## 

##      categories      genres      positive      negative
## Length:88982  Length:88982  Min. : 0  Min. : 0.0
## Class :character  Class :character  1st Qu.: 1  1st Qu.: 0.0
## Mode  :character  Mode  :character  Median : 10  Median : 2.0
## Mean   : 1267  Mean   : 208.3
## 3rd Qu.: 59  3rd Qu.: 17.0
## Max.  :7480813  Max. :1135108.0

##      average_playtime_forever average_playtime_2weeks median_playtime_forever
## Min. : 0.0  Min. : 0.000  Min. : 0.0
## 1st Qu.: 0.0  1st Qu.: 0.000  1st Qu.: 0.0
## Median : 0.0  Median : 0.000  Median : 0.0
## Mean   :115.6  Mean   : 4.998  Mean   : 115.5

```

```

## 3rd Qu.: 0.0          3rd Qu.: 0.000          3rd Qu.: 0.0
## Max.   :1462997.0      Max.   :18568.000      Max.   :1462997.0
## median_playtime_2weeks peak_ccu           tags           pct_pos_total
## Min.   : 0.000         Min.   :0.000e+00    Length:88982     Min.   :-1.00
## 1st Qu.: 0.000         1st Qu.:0.000e+00   Class :character 1st Qu.:-1.00
## Median : 0.000         Median :0.000e+00   Mode  :character Median : 60.00
## Mean   : 5.277         Mean   :9.778e+01
## 3rd Qu.: 0.000         3rd Qu.:0.000e+00
## Max.   :18568.000      Max.   :1.212e+06   Mean   : 45.39
##                                         Max.   :100.00
## num_reviews_total estimated_owners_min estimated_owners_max
## Min.   :-1             Min.   : 0            Min.   : 0
## 1st Qu.:-1             1st Qu.: 0            1st Qu.: 20000
## Median : 15            Median : 0            Median : 20000
## Mean   : 1320          Mean   : 59038        Mean   : 143056
## 3rd Qu.: 81             3rd Qu.: 0            3rd Qu.: 20000
## Max.   :8632939        Max.   :2000000000    Max.   :5000000000

```

3 Exploratory Data Analysis

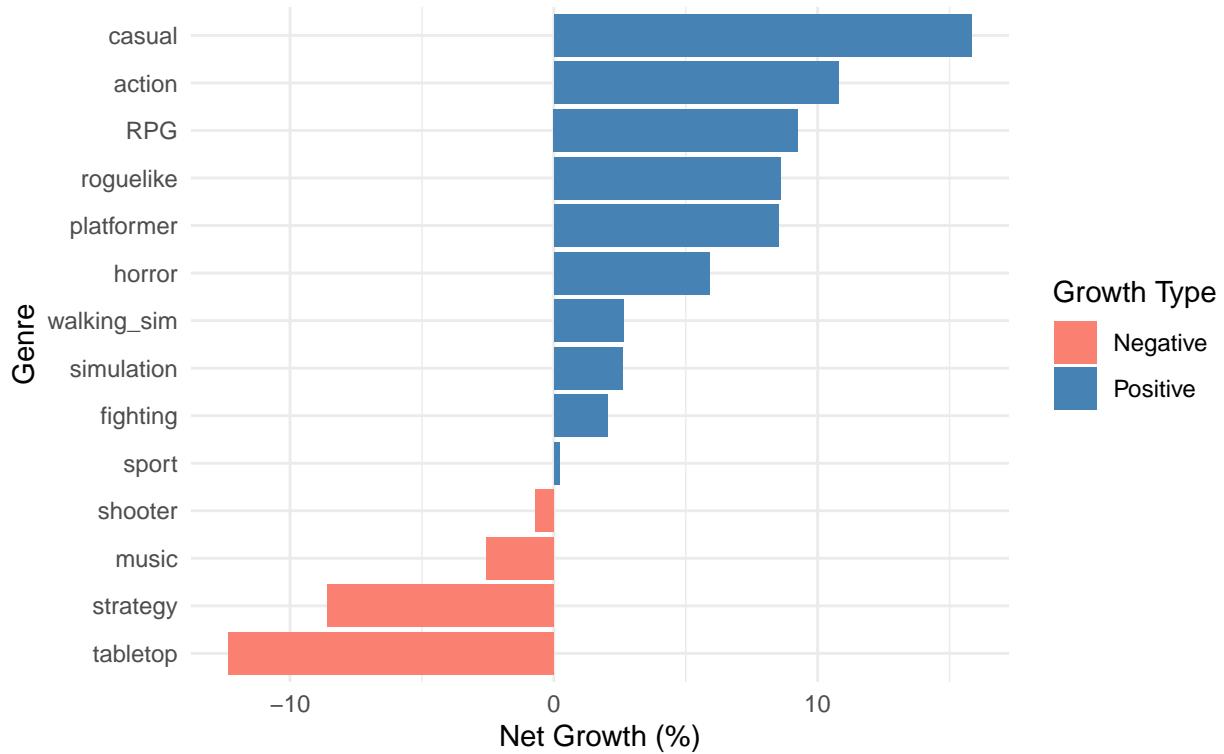
```

## # A tibble: 14 x 3
##   Genre      Total_Volatility Net_Growth
##   <chr>          <dbl>       <dbl>
## 1 action        54.0        10.8
## 2 casual        43.6        15.8
## 3 platformer    41.4        8.51
## 4 tabletop      36.6       -12.3
## 5 strategy      28.8       -8.60
## 6 shooter        27.6       -0.703
## 7 RPG            23.7        9.25
## 8 fighting       19.3        2.04
## 9 simulation    18.1        2.60
## 10 horror        15.4        5.90
## 11 sport          13.8       0.224
## 12 walking_sim   13.7        2.64
## 13 roguelike     13.2        8.59
## 14 music          10.3       -2.58

```

Top 5 and Bottom 5 Genres by Net Growth

Comparing highest vs. lowest performance



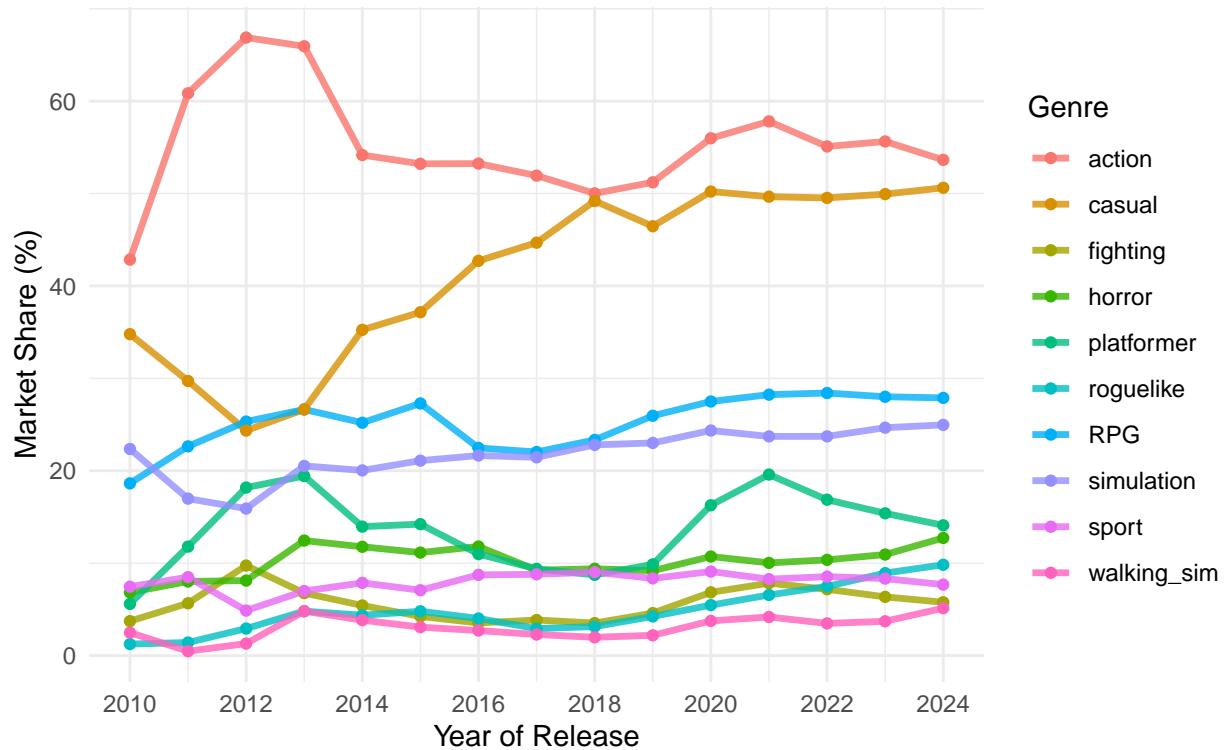
```
names(market_share)
```

```
## [1] "release_year"      "Genre"           "Count"
## [4] "Total_Games_Yearly" "Market_Share"

positive_genres <- genre_grow %>% filter(Net_Growth > 0) %>% pull(Genre)
negative_genres <- genre_grow %>% filter(Net_Growth < 0) %>% pull(Genre)
```

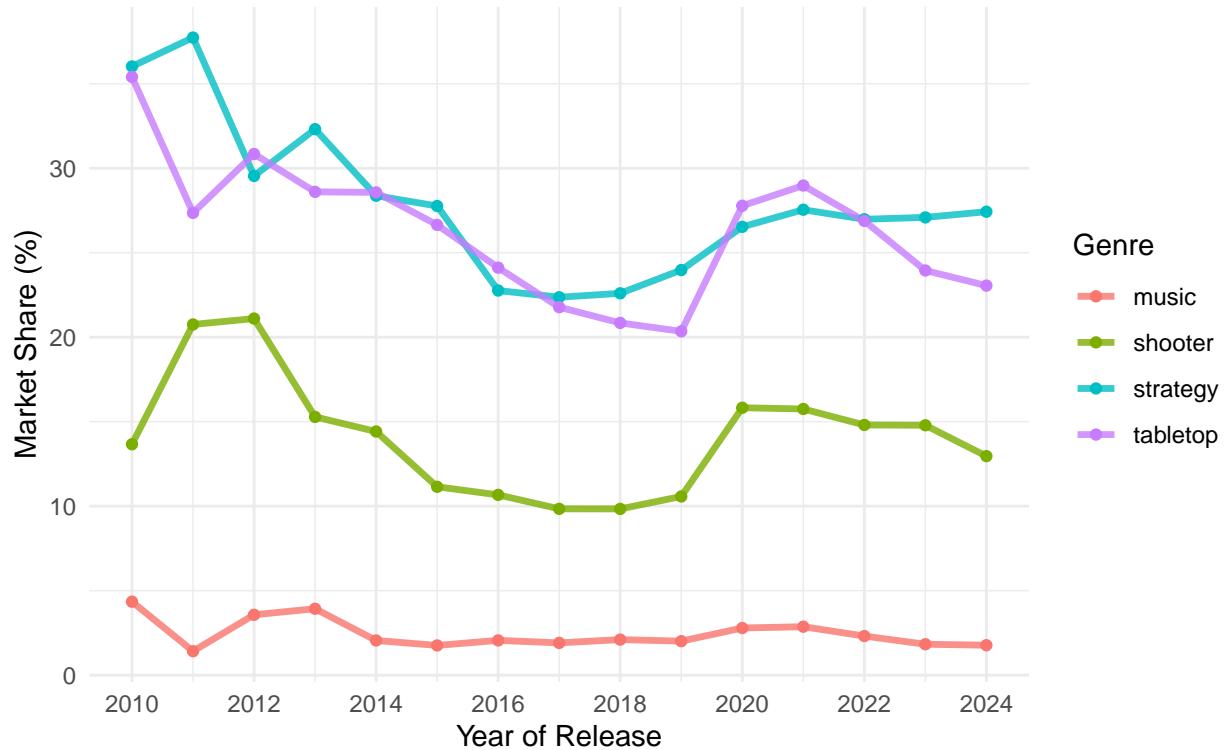
Market Share Trend: Positive Growth Genres

Genres that increased their market footprint

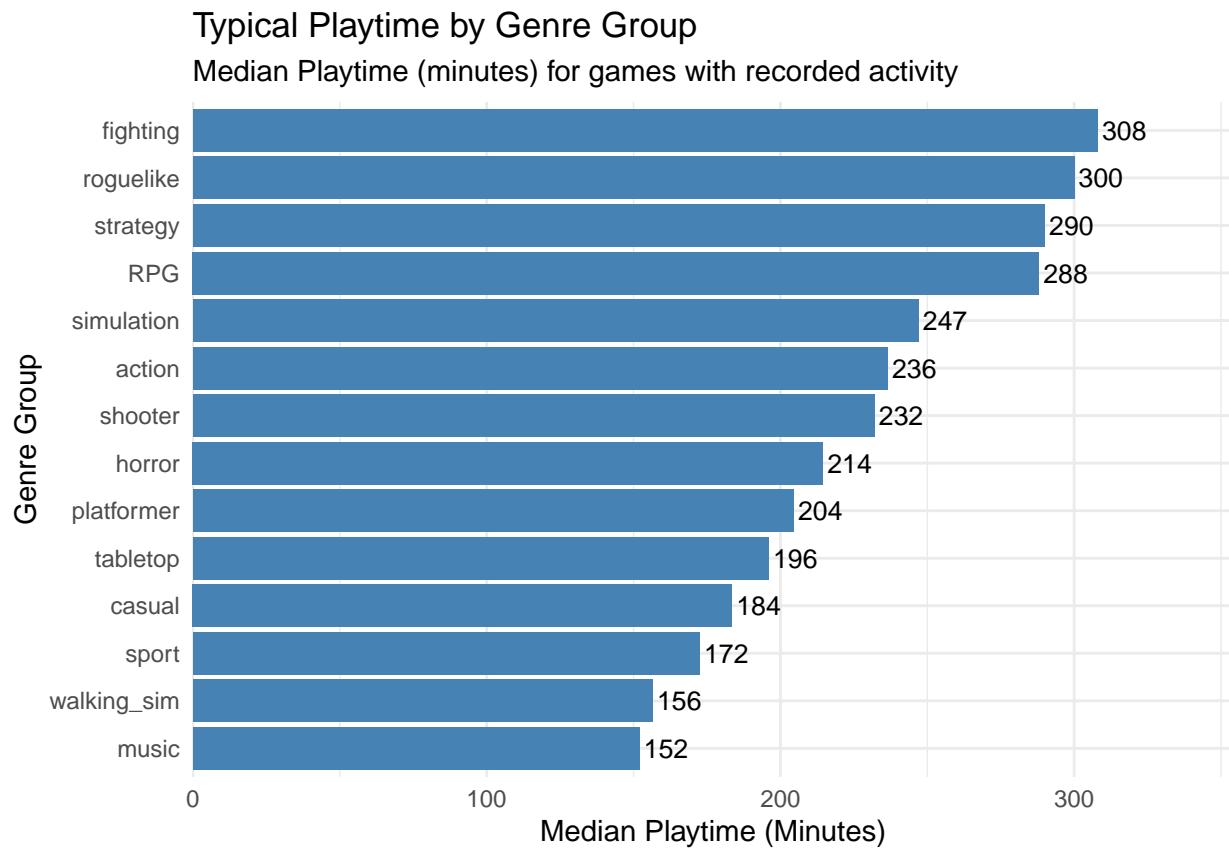


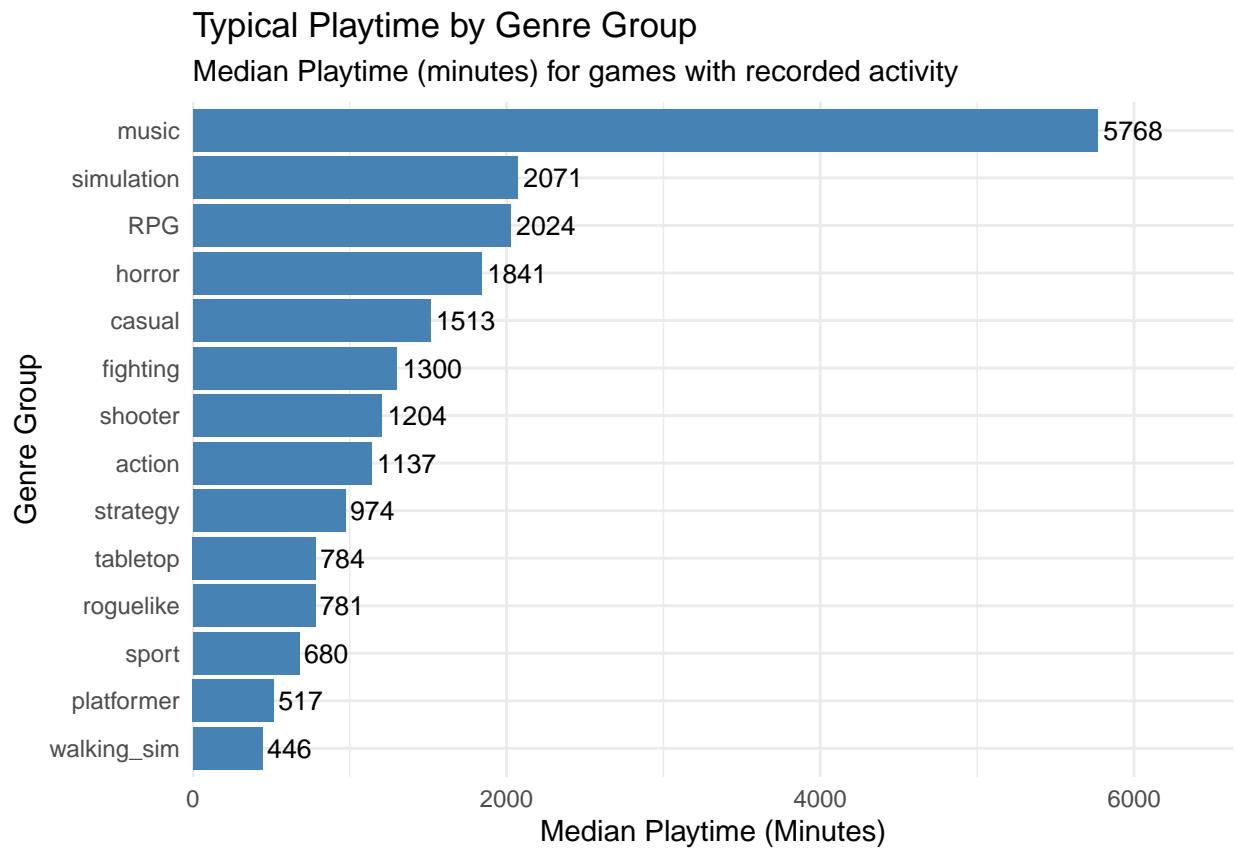
Market Share Trend: Negative Growth Genres

Genres that decreased their market footprint

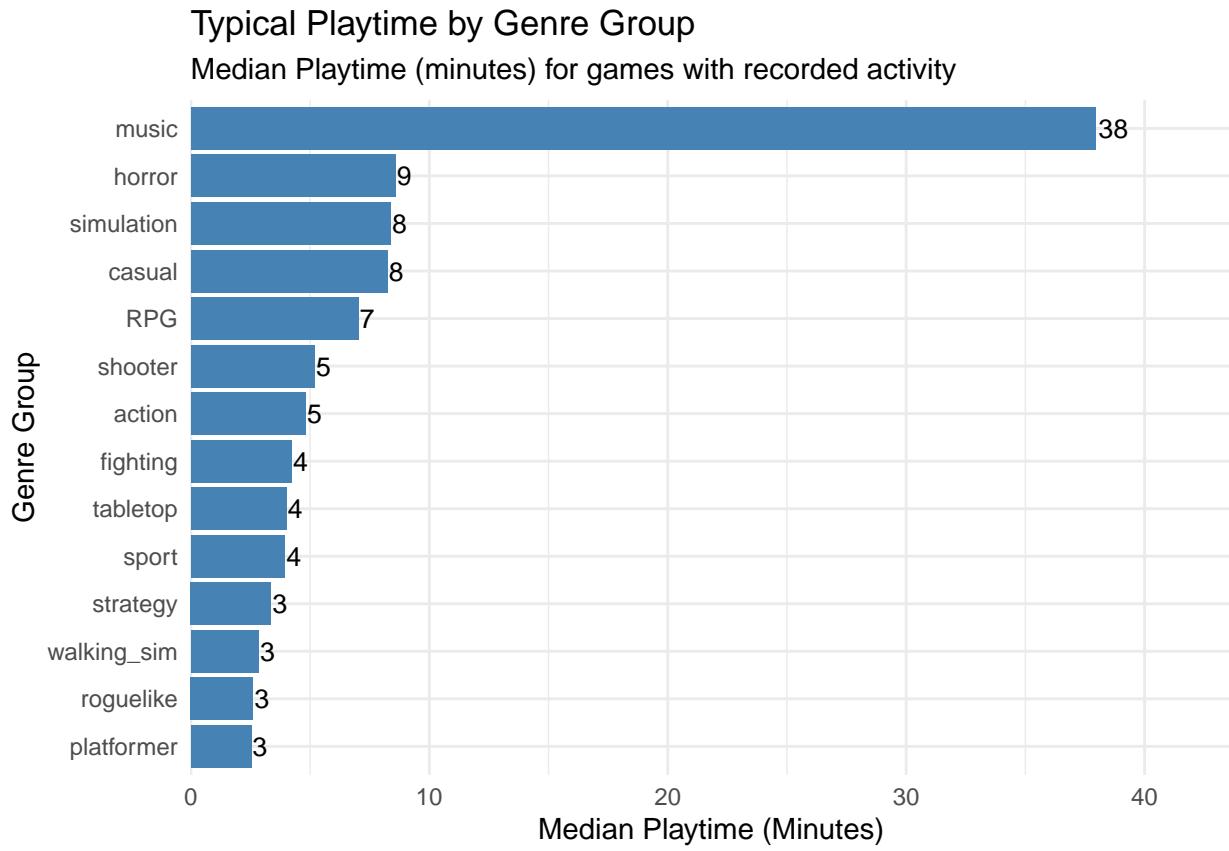


```
## # A tibble: 14 x 4
##   Genre_Group Mean_Playtime Median_Playtime Total_Games
##   <chr>          <dbl>            <dbl>        <int>
## 1 fighting       1300.           308          689
## 2 roguelike      781.            300          624
## 3 strategy        974.           290         2526
## 4 RPG             2024.           288         2851
## 5 simulation     2071.           247         2292
## 6 action           1137.           236         5062
## 7 shooter          1204.           232         1397
## 8 horror            1841.           214         1238
## 9 platformer       517.            204          1194
## 10 tabletop         784.            196         2033
## 11 casual           1513.           184          3164
## 12 sport             680.            172          700
## 13 walking_sim      446.            156          340
## 14 music            5768.           152          196
## #> [1] 7981
```





```
genre_playtime_stats$Skewness_Ratio <- genre_playtime_stats$Mean_Playtime / genre_playtime_stats$Median
```



4 Requirements for Successful Indie Games

Create helper function for extracting unique elements from entries with a list of values. All lists follow the same pattern: ['el1', 'el2', ...] being [] for empty lists.

Create helper function for extracting unique tags. These are formatted differently and follow dictionary-like pattern where each element is a key-value pair being [] for empty entries.

4.1 Obtain Unique values for different columns

```
## [1] "Captions available"      "Co-op"
## [3] "Commentary available"    "Cross-Platform Multiplayer"
## [5] "Family Sharing"          "Full controller support"
## [7] "HDR available"           "In-App Purchases"
## [9] "Includes level editor"   "Includes Source SDK"
## [11] "LAN Co-op"               "LAN PvP"
## [13] "MMO"                     "Multi-player"
## [15] "Online Co-op"            "Online PvP"
## [17] "Partial Controller Support" "PvP"
## [19] "Remote Play on Phone"     "Remote Play on Tablet"
## [21] "Remote Play on TV"        "Remote Play Together"
## [23] "Shared/Split Screen"      "Shared/Split Screen Co-op"
## [25] "Shared/Split Screen PvP"   "Single-player"
## [27] "Stats"                   "Steam Achievements"
```

```

## [29] "Steam Cloud"           "Steam Leaderboards"
## [31] "Steam Timeline"        "Steam Trading Cards"
## [33] "Steam Turn Notifications" "Steam Workshop"
## [35] "SteamVR Collectibles"   "Tracked Controller Support"
## [37] "Valve Anti-Cheat enabled" "VR Only"
## [39] "VR Support"            "VR Supported"

## [1] "360 Video"              "Accounting"          "Action"
## [4] "Adventure"               "Animation & Modeling" "Audio Production"
## [7] "Casual"                  "Design & Illustration" "Documentary"
## [10] "Early Access"            "Education"           "Episodic"
## [13] "Free To Play"           "Game Development"    "Gore"
## [16] "Indie"                  "Massively Multiplayer" "Movie"
## [19] "Nudity"                 "Photo Editing"       "Racing"
## [22] "RPG"                    "Sexual Content"      "Short"
## [25] "Simulation"             "Software Training"  "Sports"
## [28] "Strategy"                "Tutorial"            "Utilities"
## [31] "Video Production"        "Violent"             "Web Publishing"

## [1] "1980s"                  "2.5D"
## [3] "2D"                      "2D Fighter"
## [5] "2D Platformer"           "360 Video"
## [7] "3D"                      "3D Fighter"
## [9] "3D Platformer"           "3D Vision"
## [11] "4 Player Local"         "4X"
## [13] "6DOF"                   "8-bit Music"
## [15] "Abstract"                "Action"
## [17] "Action Roguelike"        "Action RPG"
## [19] "Action RTS"              "Action-Adventure"
## [21] "Addictive"                "Adventure"
## [23] "Agriculture"             "Aliens"
## [25] "Alternate History"       "Ambient"
## [27] "America"                  "Animation & Modeling"
## [29] "Anime"                    "Arcade"
## [31] "Archery"                  "Arena Shooter"
## [33] "Artificial Intelligence" "Assassin"
## [35] "Asymmetric VR"           "Asynchronous Multiplayer"
## [37] "Atmospheric"              "ATV"
## [39] "Audio Production"         "Auto Battler"
## [41] "Automation"                "Automobile Sim"
## [43] "Base-Building"             "Baseball"
## [45] "Based On A Novel"         "Basketball"
## [47] "Battle Royale"             "Beautiful"
## [49] "Benchmark"                 "Bikes"
## [51] "Birds"                     "Blood"
## [53] "BMX"                       "Board Game"
## [55] "Boomer Shooter"            "Boss Rush"
## [57] "Bowling"                   "Boxing"
## [59] "Building"                  "Bullet Hell"
## [61] "Bullet Time"                "Capitalism"
## [63] "Card Battler"                "Card Game"
## [65] "Cartoon"                   "Cartoony"
## [67] "Casual"                     "Cats"
## [69] "Character Action Game"     "Character Customization"

```

```

## [71] "Chess"                                "Choices Matter"
## [73] "Choose Your Own Adventure"            "Cinematic"
## [75] "City Builder"                          "Class-Based"
## [77] "Classic"                             "Clicker"
## [79] "Co-op"                               "Co-op Campaign"
## [81] "Coding"                               "Cold War"
## [83] "Collectathon"                         "Colony Sim"
## [85] "Colorful"                            "Combat"
## [87] "Combat Racing"                        "Comedy"
## [89] "Comic Book"                           "Competitive"
## [91] "Conspiracy"                           "Controller"
## [93] "Conversation"                         "Cooking"
## [95] "Cozy"                                 "Crafting"
## [97] "Creature Collector"                  "Cricket"
## [99] "Crime"                                "Crowdfunded"
## [101] "CRPG"                                "Cult Classic"
## [103] "Cute"                                 "Cyberpunk"
## [105] "Cycling"                             "Dark"
## [107] "Dark Comedy"                          "Dark Fantasy"
## [109] "Dark Humor"                           "Dating Sim"
## [111] "Deckbuilding"                        "Demons"
## [113] "Design & Illustration"              "Destruction"
## [115] "Detective"                            "Dice"
## [117] "Difficult"                            "Dinosaurs"
## [119] "Diplomacy"                           "Documentary"
## [121] "Dog"                                  "Dragons"
## [123] "Drama"                               "Driving"
## [125] "Dungeon Crawler"                    "Dungeons & Dragons"
## [127] "Dwarf"                                "Dynamic Narration"
## [129] "Dystopian"                           "e-sports"
## [131] "Early Access"                         "Economy"
## [133] "Education"                            "Electronic"
## [135] "Electronic Music"                    "Elf"
## [137] "Emotional"                           "Epic"
## [139] "Episodic"                            "Escape Room"
## [141] "Experience"                           "Experimental"
## [143] "Exploration"                         "Extraction Shooter"
## [145] "Faith"                                "Family Friendly"
## [147] "Fantasy"                             "Farming"
## [149] "Farming Sim"                          "Fast-Paced"
## [151] "Feature Film"                         "Female Protagonist"
## [153] "Fighting"                            "First-Person"
## [155] "Fishing"                             "Flight"
## [157] "FMV"                                  "Football"
## [159] "Football (American)"                 "Football (Soccer)"
## [161] "Foreign"                             "Fox"
## [163] "FPS"                                  "Free to Play"
## [165] "Funny"                               "Futuristic"
## [167] "Gambling"                            "Game Development"
## [169] "GameMaker"                            "Games Workshop"
## [171] "Gaming"                              "God Game"
## [173] "Golf"                                 "Gore"
## [175] "Gothic"                             "Grand Strategy"
## [177] "Great Soundtrack"                   "Grid-Based Movement"

```

```

## [179] "Gun Customization"
## [181] "Hacking"
## [183] "Hardware"
## [185] "Hentai"
## [187] "Hex Grid"
## [189] "Historical"
## [191] "Hockey"
## [193] "Horses"
## [195] "Idler"
## [197] "Immersive"
## [199] "Indie"
## [201] "Intentionally Awkward Controls"
## [203] "Inventory Management"
## [205] "Isometric"
## [207] "Job Simulator"
## [209] "Jump Scare"
## [211] "LEGO"
## [213] "Level Editor"
## [215] "Life Sim"
## [217] "Local Co-Op"
## [219] "Logic"
## [221] "Looter Shooter"
## [223] "Lovecraftian"
## [225] "Mahjong"
## [227] "Mars"
## [229] "Massively Multiplayer"
## [231] "Match 3"
## [233] "Mechs"
## [235] "Medieval"
## [237] "Metroidvania"
## [239] "Mini Golf"
## [241] "Minimalist"
## [243] "MMORPG"
## [245] "Mod"
## [247] "Modern"
## [249] "Motorbike"
## [251] "Movie"
## [253] "Multiple Endings"
## [255] "Music-Based Procedural Generation"
## [257] "Mystery"
## [259] "Mythology"
## [261] "Narrative"
## [263] "Naval"
## [265] "Ninja"
## [267] "Nonlinear"
## [269] "NSFW"
## [271] "Offroad"
## [273] "On-Rails Shooter"
## [275] "Open World"
## [277] "Otome"
## [279] "Parkour"
## [281] "Party"
## [283] "Party-Based RPG"
## [285] "Philosophical"
## [287] "Physics"
## [289] "Puzzle"
## [291] "Quest"
## [293] "Racing"
## [295] "Rhythm"
## [297] "Roleplay"
## [299] "RPG"
## [301] "Sci-Fi"
## [303] "Shooter"
## [305] "Simulation"
## [307] "Sims"
## [309] "Sports"
## [311] "Strategy"
## [313] "Survival"
## [315] "Suspense"
## [317] "Tactical"
## [319] "Tech"
## [321] "Thriller"
## [323] "Time Travel"
## [325] "Top-Down"
## [327] "Tower Defense"
## [329] "Trickshot"
## [331] "Underwater"
## [333] "Videogame"
## [335] "Virtual Reality"
## [337] "Visual Novel"
## [339] "Western"
## [341] "Zombie"

```

```

## [287] "Physics"
## [289] "Pirates"
## [291] "Platformer"
## [293] "Political"
## [295] "Politics"
## [297] "Post-apocalyptic"
## [299] "Procedural Generation"
## [301] "Psychedelic"
## [303] "Psychological Horror"
## [305] "Puzzle-Platformer"
## [307] "PvP"
## [309] "Racing"
## [311] "Real-Time"
## [313] "Realistic"
## [315] "Relaxing"
## [317] "Replay Value"
## [319] "Retro"
## [321] "Robots"
## [323] "Rogue-like"
## [325] "Roguelike Deckbuilder"
## [327] "Romance"
## [329] "RPG"
## [331] "RTS"
## [333] "Runner"
## [335] "Sandbox"
## [337] "Sci-fi"
## [339] "Score Attack"
## [341] "Sexual Content"
## [343] "Shop Keeper"
## [345] "Side Scroller"
## [347] "Simulation"
## [349] "Skateboarding"
## [351] "Skiing"
## [353] "Snooker"
## [355] "Snowboarding"
## [357] "Social Deduction"
## [359] "Software Training"
## [361] "Solitaire"
## [363] "Soundtrack"
## [365] "Space Sim"
## [367] "Spectacle fighter"
## [369] "Split Screen"
## [371] "Stealth"
## [373] "Steampunk"
## [375] "Strategy"
## [377] "Stylized"
## [379] "Superhero"
## [381] "Surreal"
## [383] "Survival Horror"
## [385] "Tabletop"
## [387] "Tactical RPG"
## [389] "Team-Based"
## [391] "Text-Based"
## [393] "Third-Person Shooter"
"Pinball"
"Pixel Graphics"
"Point & Click"
"Political Sim"
"Pool"
"Precision Platformer"
"Programming"
"Psychological"
"Puzzle"
"PvE"
"Quick-Time Events"
"Real Time Tactics"
"Real-Time with Pause"
"Reboot"
"Remake"
"Resource Management"
"Rhythm"
"Rock Music"
"Rogue-lite"
"Roguevania"
"Rome"
"RPGMaker"
"Rugby"
"Sailing"
"Satire"
"Science"
"Sequel"
"Shooter"
"Short"
"Silent Protagonist"
"Singleplayer"
"Skating"
"Sniper"
"Snow"
"Soccer"
"Software"
"Sokoban"
"Souls-like"
"Space"
"Spaceships"
"Spelling"
"Sports"
"Steam Machine"
"Story Rich"
"Strategy RPG"
"Submarine"
"Supernatural"
"Survival"
"Swordplay"
"Tactical"
"Tanks"
"Tennis"
"Third Person"
"Thriller"

```

```

## [395] "Tile-Matching"
## [397] "Time Management"
## [399] "Time Travel"
## [401] "Top-Down Shooter"
## [403] "Tower Defense"
## [405] "Trading"
## [407] "Traditional Roguelike"
## [409] "Transhumanism"
## [411] "Trivia"
## [413] "Turn-Based Combat"
## [415] "Turn-Based Tactics"
## [417] "Twin Stick Shooter"
## [419] "Underground"
## [421] "Unforgiving"
## [423] "Vampire"
## [425] "Video Production"
## [427] "Villain Protagonist"
## [429] "Visual Novel"
## [431] "Volleyball"
## [433] "VR"
## [435] "Walking Simulator"
## [437] "Wargame"
## [439] "Web Publishing"
## [441] "Werewolves"
## [443] "Wholesome"
## [445] "World War I"
## [447] "Wrestling"
## [395] "Time Attack"
## [397] "Time Manipulation"
## [399] "Top-Down"
## [401] "Touch-Friendly"
## [403] "TrackIR"
## [405] "Trading Card Game"
## [407] "Trains"
## [409] "Transportation"
## [411] "Turn-Based"
## [413] "Turn-Based Strategy"
## [415] "Tutorial"
## [417] "Typing"
## [419] "Underwater"
## [421] "Utilities"
## [423] "Vehicular Combat"
## [425] "Vikings"
## [427] "Violent"
## [429] "Voice Control"
## [431] "Voxel"
## [433] "VR Only"
## [435] "War"
## [437] "Warhammer 40K"
## [439] "Well-Written"
## [441] "Western"
## [443] "Word Game"
## [445] "World War II"
## [447] "Zombies"

```

From steam's Categories, Genres and Tags, 3 subsets have been created with samples interesting for this study. Given steam's laxity at classifying these values, the 3 subsets have been created with our own interpretation. Additionally, tags that refer to non-game software have been ignored.

Thankfully the task of filtering indie games is already provided by steam with the Tags and Genres Indie, Crowdfunded and Kickstarter. From the original dataset, only Indie games are obtained.

Prepare data for determining successful games

Cluster the data:

```

## [1] "appid"                      "name"
## [3] "price"                       "dlc_count"
## [5] "windows"                     "mac"
## [7] "linux"                       "achievements"
## [9] "recommendations"             "supported_languages"
## [11] "positive"                    "negative"
## [13] "average_playtime_forever"   "median_playtime_forever"
## [15] "peak_ccu"                   "pct_pos_total"
## [17] "num_reviews_total"          "estimated_owners_min"
## [19] "estimated_owners_max"       "estimated_revenue"
## [21] "success_cluster"

```

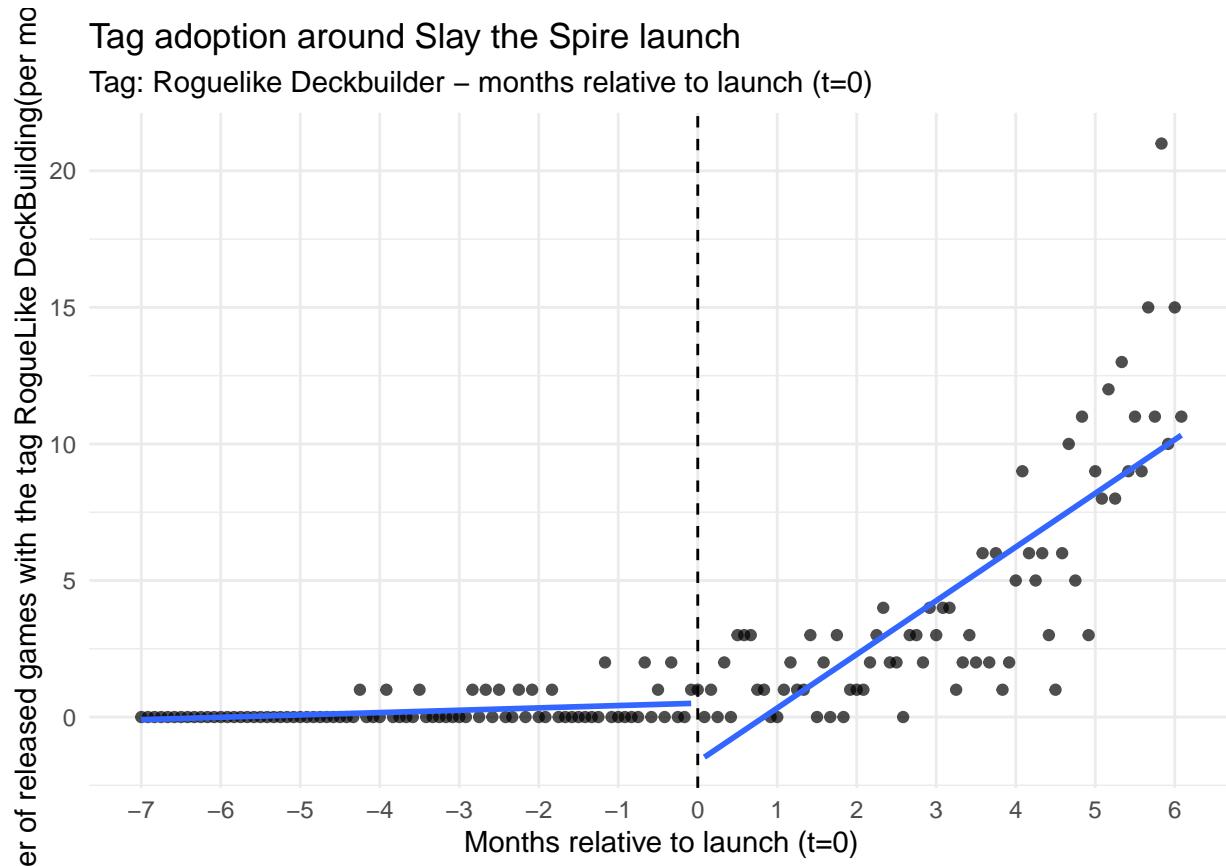
5 Can a single game have enough influence to make other games have its tag?

5.1 Game study: Slay the Spire (Roguelike Deckbuilder)

```
## [1] "Slay the Spire"

## [1] "2019-01-23"

##
## Call:
## lm(formula = freq ~ t + post + t:post, data = ts_slayTheSpire)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -6.2087 -0.4368 -0.1021  0.5213 11.1710 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.529412  0.403337  1.313 0.191278    
## t           0.007563  0.008292  0.912 0.363151    
## post        -2.164115  0.599569 -3.609 0.000414 ***  
## t:post      0.156203  0.013316 11.731 < 2e-16 ***  
## ---        
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.876 on 154 degrees of freedom
## Multiple R-squared:  0.7433, Adjusted R-squared:  0.7383 
## F-statistic: 148.7 on 3 and 154 DF,  p-value: < 2.2e-16
```



```
##  

## Call:  

## lm(formula = freq ~ t + post + t:post, data = ts_slayTheSpire)  

##  

## Residuals:  

##      Min       1Q   Median       3Q      Max  

## -6.2087 -0.4368 -0.1021  0.5213 11.1710  

##  

## Coefficients:  

##             Estimate Std. Error t value Pr(>|t|)  

## (Intercept) 0.529412  0.403337  1.313 0.191278  

## t          0.007563  0.008292  0.912 0.363151  

## post     -2.164115  0.599569 -3.609 0.000414 ***  

## t:post    0.156203  0.013316 11.731 < 2e-16 ***  

## ---  

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  

##  

## Residual standard error: 1.876 on 154 degrees of freedom  

## Multiple R-squared:  0.7433, Adjusted R-squared:  0.7383  

## F-statistic: 148.7 on 3 and 154 DF,  p-value: < 2.2e-16  

##  

## Call:  

## lm(formula = freq ~ t + post + t:post, data = ts_global)  

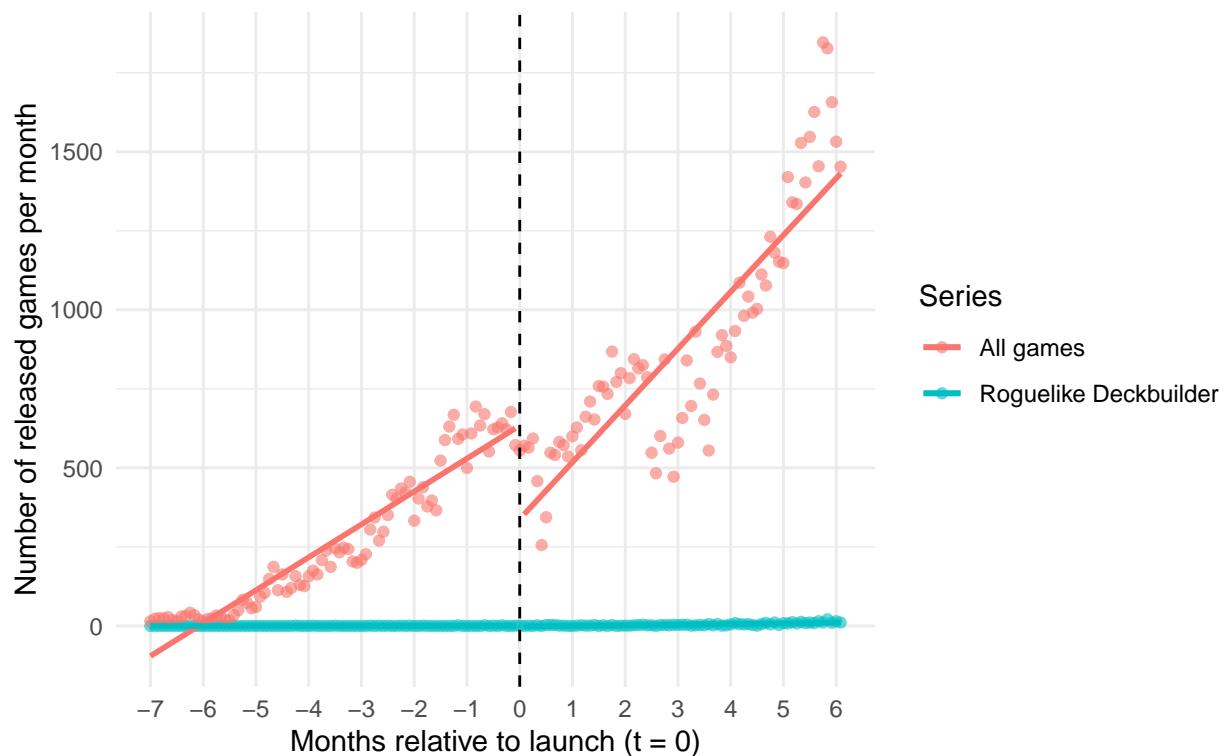
##
```

```

## Residuals:
##      Min      1Q   Median      3Q     Max
## -427.10  -66.85   -1.87  81.84 474.64
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 630.1863   28.6368  22.006 < 2e-16 ***
## t           8.6137    0.5887  14.631 < 2e-16 ***
## post       -291.8553   42.5691 -6.856 1.61e-10 ***
## t:post      6.3578    0.9454   6.725 3.24e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 133.2 on 154 degrees of freedom
## Multiple R-squared:  0.9061, Adjusted R-squared:  0.9042
## F-statistic: 495.2 on 3 and 154 DF, p-value: < 2.2e-16

```

Slope comparison around Slay the Spire launch
Target tag vs overall Steam release trend (monthly bins)



```

##
## Call:
## lm(formula = freq ~ t * post * series, data = combined)
##
## Residuals:
##      Min      1Q   Median      3Q     Max
## -427.10  -4.05   -0.12    4.33  474.64
## 
```

```

## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)            630.1863   20.2513 31.118 < 2e-16 ***
## t                      8.6137    0.4163 20.689 < 2e-16 ***
## post                  -291.8553   30.1039 -9.695 < 2e-16 ***
## seriestag              -629.6569   28.6396 -21.986 < 2e-16 ***
## t:post                 6.3578    0.6686  9.509 < 2e-16 ***
## t:seriestag             -8.6061   0.5888 -14.617 < 2e-16 ***
## post:seriestag         289.6912   42.5733  6.805 5.29e-11 ***
## t:post:seriestag      -6.2015    0.9455 -6.559 2.29e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 94.18 on 308 degrees of freedom
## Multiple R-squared:  0.9488, Adjusted R-squared:  0.9476
## F-statistic: 815.6 on 7 and 308 DF,  p-value: < 2.2e-16

```

5.2 Game Study: The Binding Of Isaac + The Binding Of Isaac Rebirth (Roguelike)

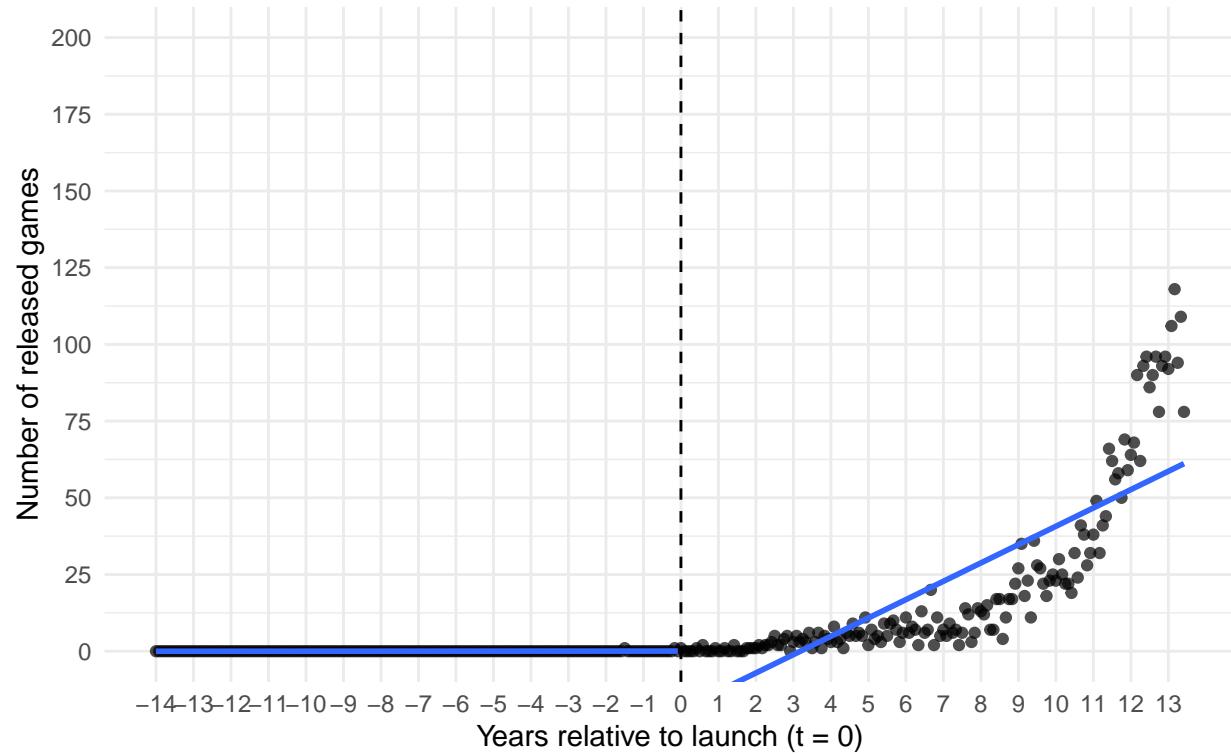
```

## [1] "The Binding of Isaac"
## [1] "2011-09-28"
## [1] "The Binding of Isaac: Rebirth"
## [1] "2014-11-04"

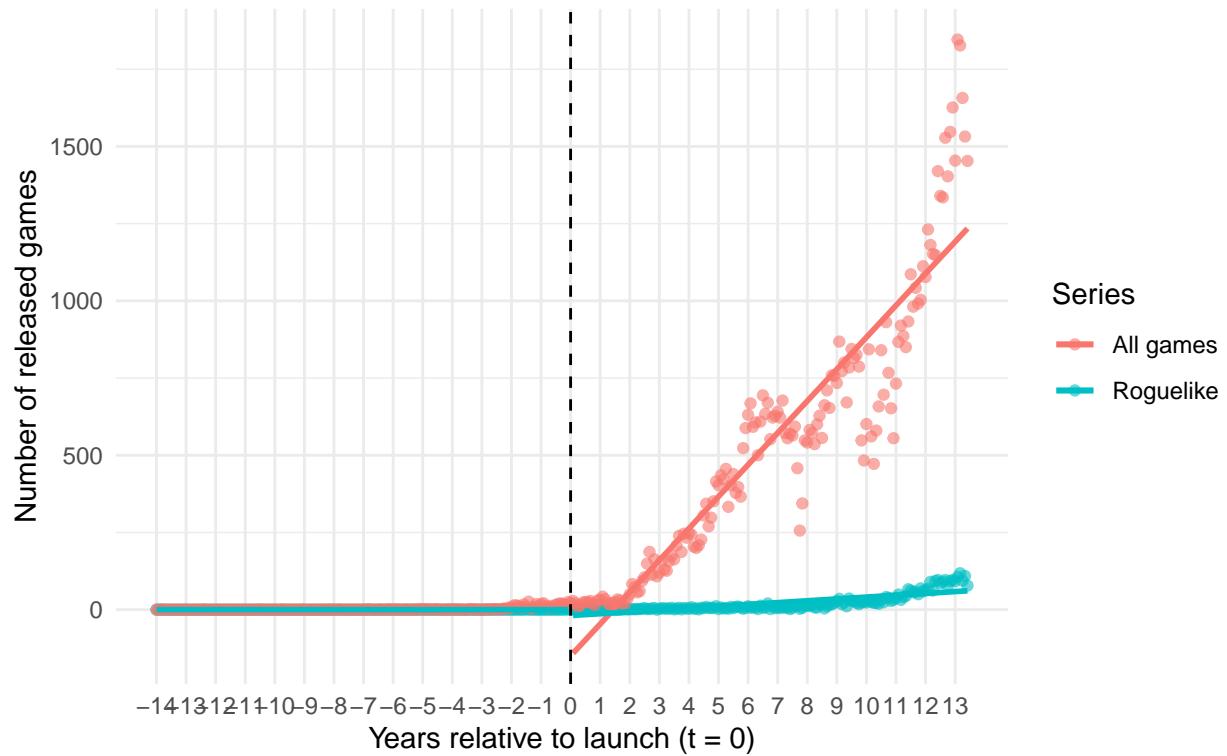
```

Tag adoption around The Binding Of Isaac

Tag: Roguelike – months relative to launch (t=0)

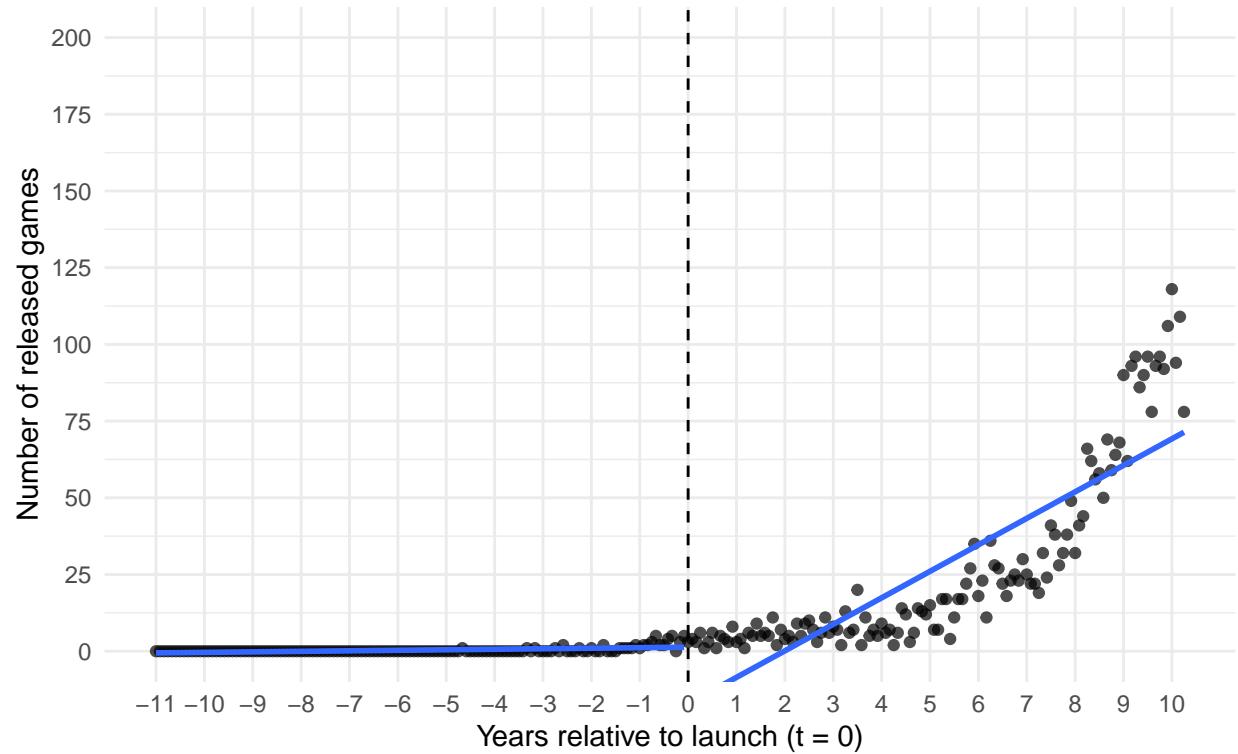


Slope comparison around The Binding of Isaac launch
Target tag vs overall Steam release trend (monthly bins)

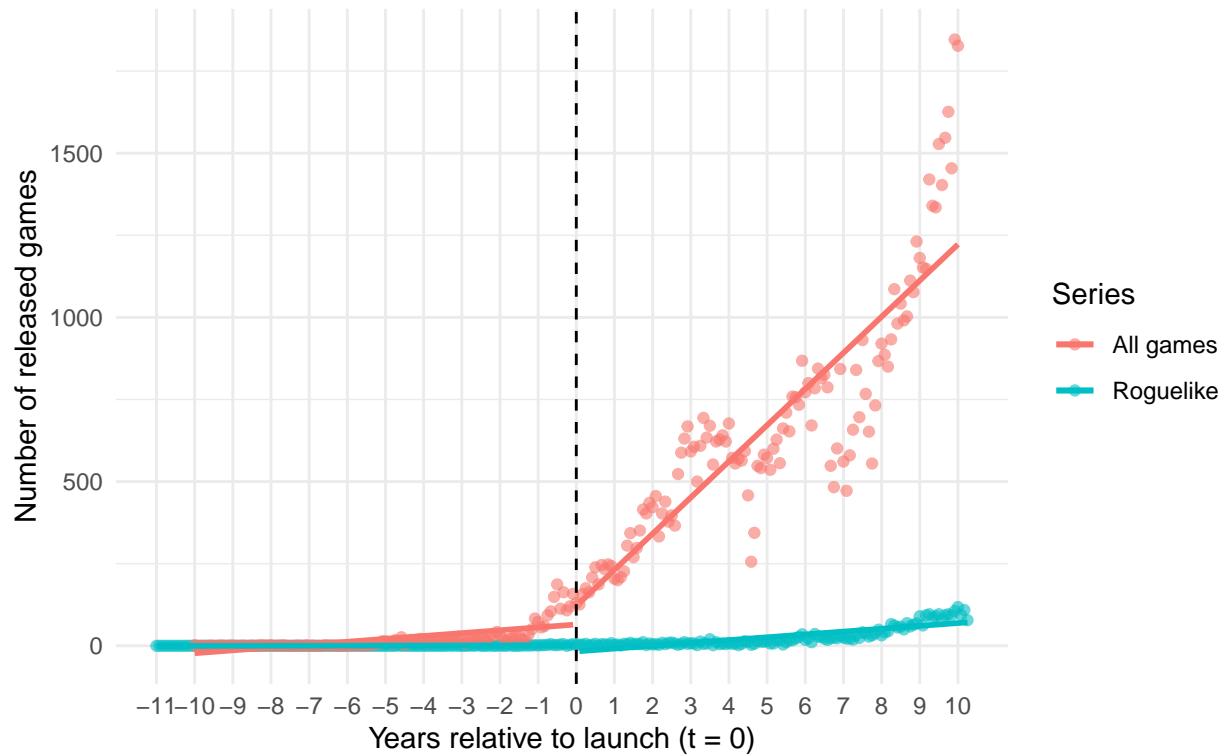


Tag adoption around The Binding Of Isaac REBIRTH

Tag: Roguelike – months relative to launch (t=0)



Slope comparison around The Binding of Isaac Rebirth launch
Roguelike tag vs overall Steam releases



```
## [1] "Terraria"
```

```
## [1] "2011-05-16"
```

Tag adoption around Terraria

Tag: Open World Survival Craft – months relative to launch (t=0)

