# Flashpoint Analytics

## Flashpoint Analytics

Flashpoint is a web-game preservation project, made in 2018 in an effort to save as many games as possible from the then upcoming Flash End-Of-Life, while also making them playable for everyone. Today, it hosts more than 170 000 games and thousands of active users all around the world.

This notebook contains a descriptive statistical analysis about the games available in Flashpoint, with an emphasis on categorical data, such as the technology that was used to make them or the publisher who used to host them in the past.

The Flashpoint database, which keeps all the data that will be used in the analysis, can be found here.

### Import modules

```
options(warn = -1)
library(RSQLite)
library(dplyr, warn.conflicts=FALSE)
library(ggplot2)
library(lubridate, warn.conflicts=FALSE)
library(tidyr)
```

These are the modules that we are going to use for our analysis. Notably: * *RSQLite* allows us to communicate with a SQLite database; * *dplyr*, *ggplot2*, *lubridate* and *tidyr* are part of the *tidyverse*, a set of very useful packages for data science.

### Retrieve data

```
dir.create("data")
url <- "http://infinity.unstable.life/Flashpoint/Data/flashpoint.sqlite"
destfile <- "data/flashpoint.sqlite"

# by default, there is a 60 seconds timeout limit when downloading a file
# this can be a problem when downloading large files, so we increase that limit
options(timeout = 300)

download.file(url, destfile, mode = "wb")

con <- dbConnect(SQLite(), "data/flashpoint.sqlite")
games <- dbReadTable(con, "game")
dbDisconnect(con)
```

## Explore data

Let's have a first look at our data.

```
str(games)
```

```
## 'data.frame':    177469 obs. of  27 variables:
##  $ id                 : chr  "6db72888-6aa5-34c9-0ff3-ffe4cfe0fc61" "3bba3af6-8e76-b2c8-b423-2d2d8bd
##  $ parentGameId       : chr  "6db72888-6aa5-34c9-0ff3-ffe4cfe0fc61" "3bba3af6-8e76-b2c8-b423-2d2d8bd
##  $ title              : chr  "All Grown Up: Krazy Karts" "Showdown: The Gunfighting Game" "Hamsterba
##  $ alternateTitles    : chr  "" "" "" "" ...
##  $ series             : chr  "" "" "" "" ...
##  $ developer          : chr  "Ezone" "3D Groove" "Ezone" "Ezone" ...
##  $ publisher          : chr  "Nickelodeon" "3D Groove" "atv.Disney.go.com" "atv.Disney.go.com" ...
##  $ dateAdded          : chr  "2019-06-23 04:12:17.521" "2019-06-16 23:56:03.150" "2019-06-16 23:56:0
##  $ dateModified       : chr  "2021-03-07 02:06:08" "2021-03-07 02:06:08" "2022-07-24 14:10:21" "2022
##  $ platform           : chr  "3D Groove GX" "3D Groove GX" "3D Groove GX" "3D Groove GX" ...
##  $ broken             : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ extreme            : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ playMode           : chr  "Single Player" "Single Player" "Single Player" "Single Player" ...
##  $ status             : chr  "Playable" "Playable" "Playable" "Playable" ...
##  $ notes              : chr  "" "" "" "" ...
##  $ source             : chr  "static.nickjr.com" "www.3dgroove.com" "Tomysshadow" "Tomysshadow" ...
##  $ applicationPath    : chr  "FPSoftware\\startGroove.bat" "FPSoftware\\startGroove.bat" "FPSoftware
##  $ launchCommand      : chr  "\"http://www.nick.com/games/data/rrgrownup/rrgp_krazykarts/code.grv\" a
##  $ releaseDate        : chr  "" "" "" "" ...
##  $ version            : chr  "" "" "" "" ...
##  $ originalDescription: chr  "" "" "" "" ...
##  $ language           : chr  "" "" "" "" ...
##  $ library            : chr  "arcade" "arcade" "arcade" "arcade" ...
##  $ orderTitle         : chr  "all grown up: krazy karts" "showdown: the gunfighting game" "hamsterba
##  $ activeDataId       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ activeDataOnDisk   : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ tagsStr            : chr  "Racing" "Shooter" "Arcade" "Arcade" ...
```

There is a total of 27 variables and almost all of them belong to the *object* data type. We are going to need only some of them, so let's keep only the relevant ones.

```
games <- games %>% select(id, title, developer, publisher, platform,
                          releaseDate, language, library, tagsStr)

str(games)
```

```
## 'data.frame':    177469 obs. of  9 variables:
##  $ id         : chr  "6db72888-6aa5-34c9-0ff3-ffe4cfe0fc61" "3bba3af6-8e76-b2c8-b423-2d2d8bdfdd50" "
##  $ title      : chr  "All Grown Up: Krazy Karts" "Showdown: The Gunfighting Game" "Hamsterball Bowli
##  $ developer  : chr  "Ezone" "3D Groove" "Ezone" "Ezone" ...
##  $ publisher  : chr  "Nickelodeon" "3D Groove" "atv.Disney.go.com" "atv.Disney.go.com" ...
##  $ platform   : chr  "3D Groove GX" "3D Groove GX" "3D Groove GX" "3D Groove GX" ...
##  $ releaseDate: chr  "" "" "" "" ...
##  $ language   : chr  "" "" "" "" ...
##  $ library    : chr  "arcade" "arcade" "arcade" "arcade" ...
##  $ tagsStr    : chr  "Racing" "Shooter" "Arcade" "Arcade" ...
```

To complete our preliminary analysis, let's print the first rows of our dataframe.

```
head(games)
```

```
##                                     id
## 1 6db72888-6aa5-34c9-0ff3-ffe4cfe0fc61
## 2 3bba3af6-8e76-b2c8-b423-2d2d8bdfdd50
## 3 fb479276-2325-4dbb-bafd-64fcc8aeb684
## 4 7cdff5f4-11cb-b1de-51bc-bca13ef78adb
## 5 1d9ff021-2404-9785-bb7d-0fceda67a55d
## 6 65caa1b2-6702-4aad-8108-790be1ef86e2
##                                          title developer
## 1                       All Grown Up: Krazy Karts     Ezone
## 2                  Showdown: The Gunfighting Game 3D Groove
## 3                             Hamsterball Bowling     Ezone
## 4                                       Dunk Tank     Ezone
## 5                                    Baby Knight  Pepworks
## 6 Yin Yang Yo: The Dangerous Comic Book of Dread     Ezone
##               publisher     platform releaseDate language library
## 1           Nickelodeon 3D Groove GX                        arcade
## 2             3D Groove 3D Groove GX                        arcade
## 3     atv.Disney.go.com 3D Groove GX                        arcade
## 4     atv.Disney.go.com 3D Groove GX                        arcade
## 5              Pepworks 3D Groove GX                        arcade
## 6 a.media.global.go.com 3D Groove GX                   en   arcade
##                     tagsStr
## 1                    Racing
## 2                   Shooter
## 3                    Arcade
## 4                    Arcade
## 5                 Platformer
## 6 Platformer; Yin Yang Yo!
```

## Analyze data

### Developers and Publishers

It would be interesting to know which are the most prolific developers and publishers. Let's find out by creating a frequency table for each variable and looking at the first ten entries.

```
top_developers <- games %>% count(developer) %>% arrange(desc(n)) %>% head(10)
top_developers
```

```
##          developer     n
## 1                  46151
## 2           123Bee  2730
## 3       Games2Rule  2441
## 4  Games2Jolly.com  1915
## 5      Selfdefiant  1603
## 6    WowEscape.com  1436
## 7    Top10NewGames  1144
## 8    PalmarianFire  1059
```

```
## 9  Ena Game Studio  1031
## 10         Neopets   957
```

The first row is blank because some games (actually, most of them) do not have a developer value associated in the database. Let's filter out those entries.

```
top_developers <- games %>% filter(developer != "") %>% count(developer) %>%
  arrange(desc(n)) %>% head(10)
top_developers
```

```
##          developer    n
## 1           123Bee 2730
## 2       Games2Rule 2441
## 3  Games2Jolly.com 1915
## 4      Selfdefiant 1603
## 5    WowEscape.com 1436
## 6    Top10NewGames 1144
## 7    PalmarianFire 1059
## 8  Ena Game Studio 1031
## 9          Neopets  957
## 10    Mirchi Games  812
```

These are the most represented developers in the database. It is impressive to notice that almost all of them are known to be specialised in escape games, which we can suppose to be a very popular genre; we will dig into this later. Lastly, a special mention for Neopets, which managed to build a passionate community still active after over 20 years.

```
top_publishers <- games %>% filter(publisher != "") %>% count(publisher) %>%
  arrange(desc(n)) %>% head(10)
top_publishers
```

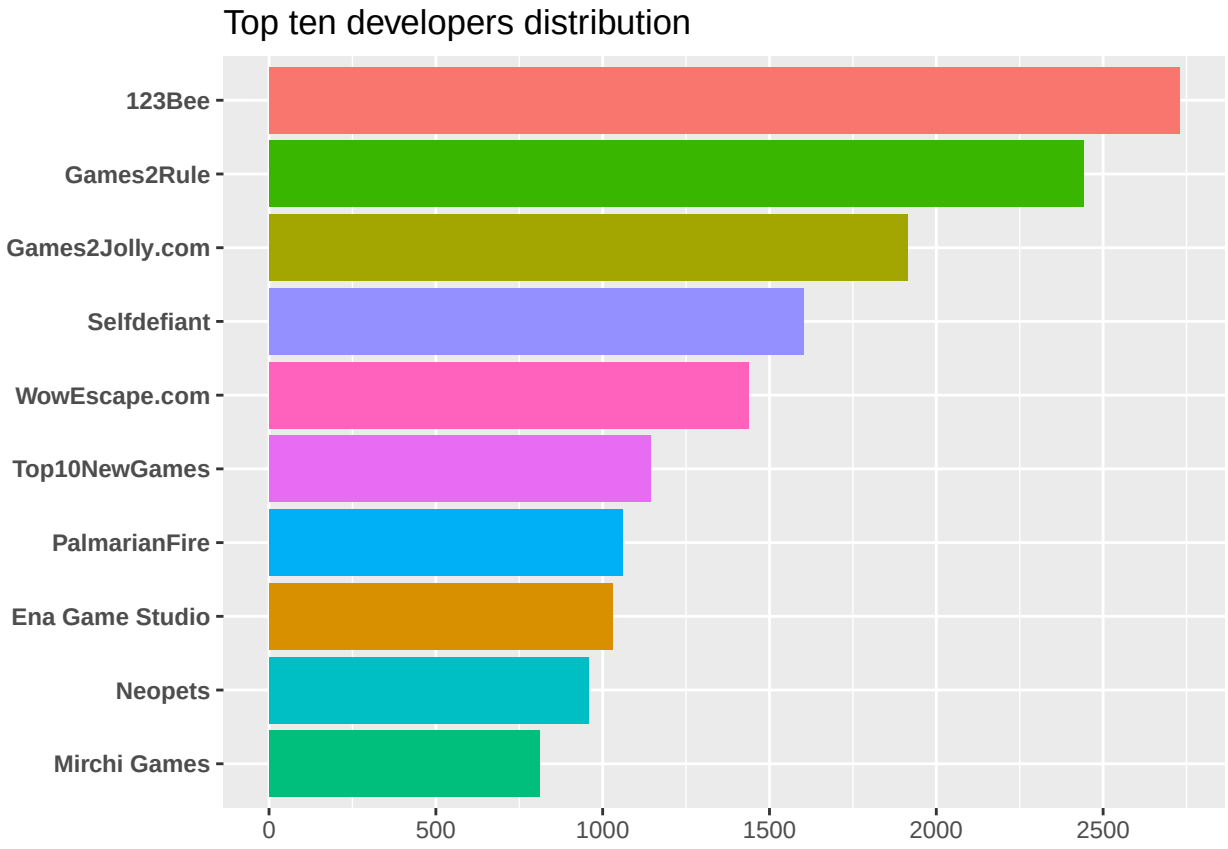
```
##          publisher    n
## 1       DeviantArt 7655
## 2       Newgrounds 6550
## 3           Disney 2428
## 4      Nickelodeon 1659
## 5      Armor Games 1437
## 6     GameMonetize 1374
## 7     Eka's Portal 1261
## 8  Cartoon Network 1243
## 9       Kongregate 1219
## 10   Melting-Mindz 1214
```

Among the publishers, we can see some very renowned names, at least in the gaming community, like *Newgrounds*, *Armor Games* and *Kongregate*. There is also a considerable amount of games published by TV broadcasters, such as *Disney*, *Nickelodeon* and *Cartoon Network*, supposedly to promote their shows.

Now let's look at a visualization of the same data, by making use of bar plots and pie charts.

```
top_developers %>% as.data.frame() %>%
  ggplot(aes(x = reorder(developer, n), y = n, fill = developer)) +
  geom_bar(stat = "identity") + coord_flip() +
```

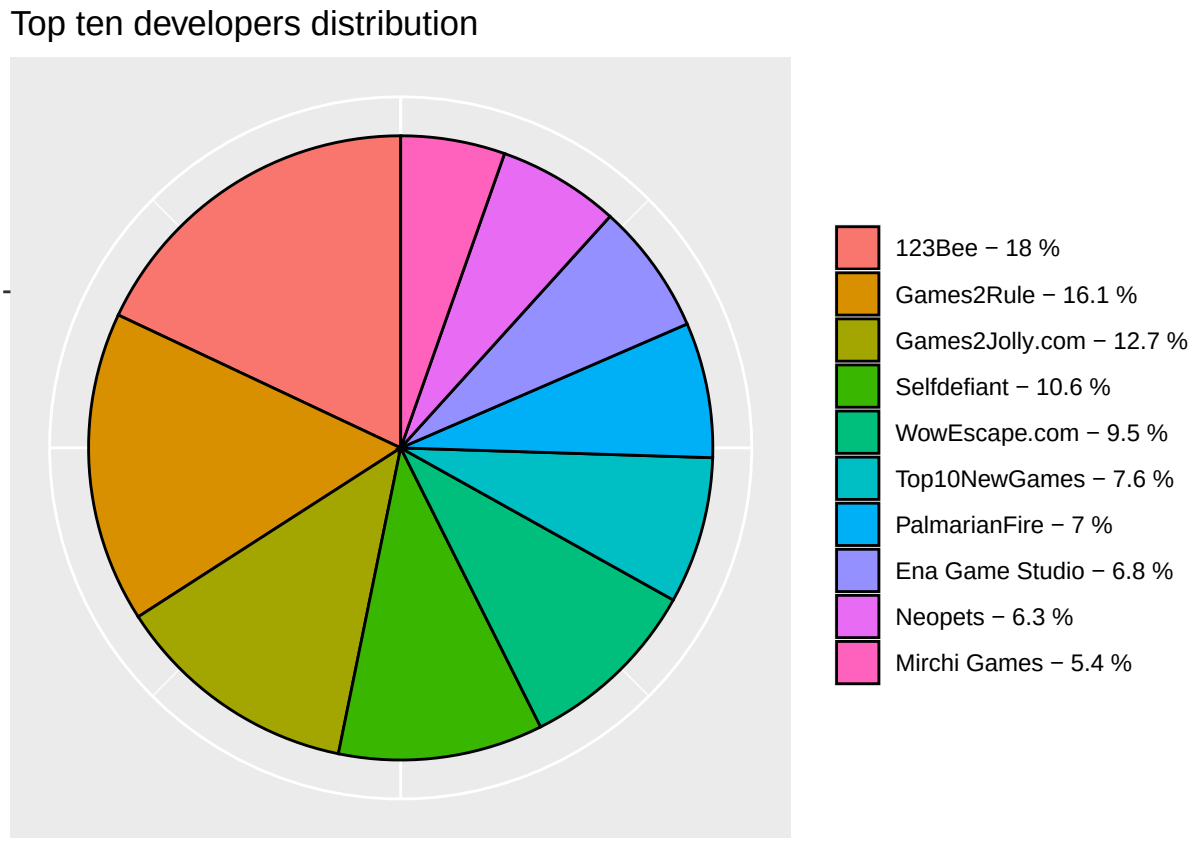
```
  scale_y_continuous(breaks = seq(0, 2500, by = 500)) +
  ggtitle("Top ten developers distribution") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.y = element_text(face = "bold"))
```



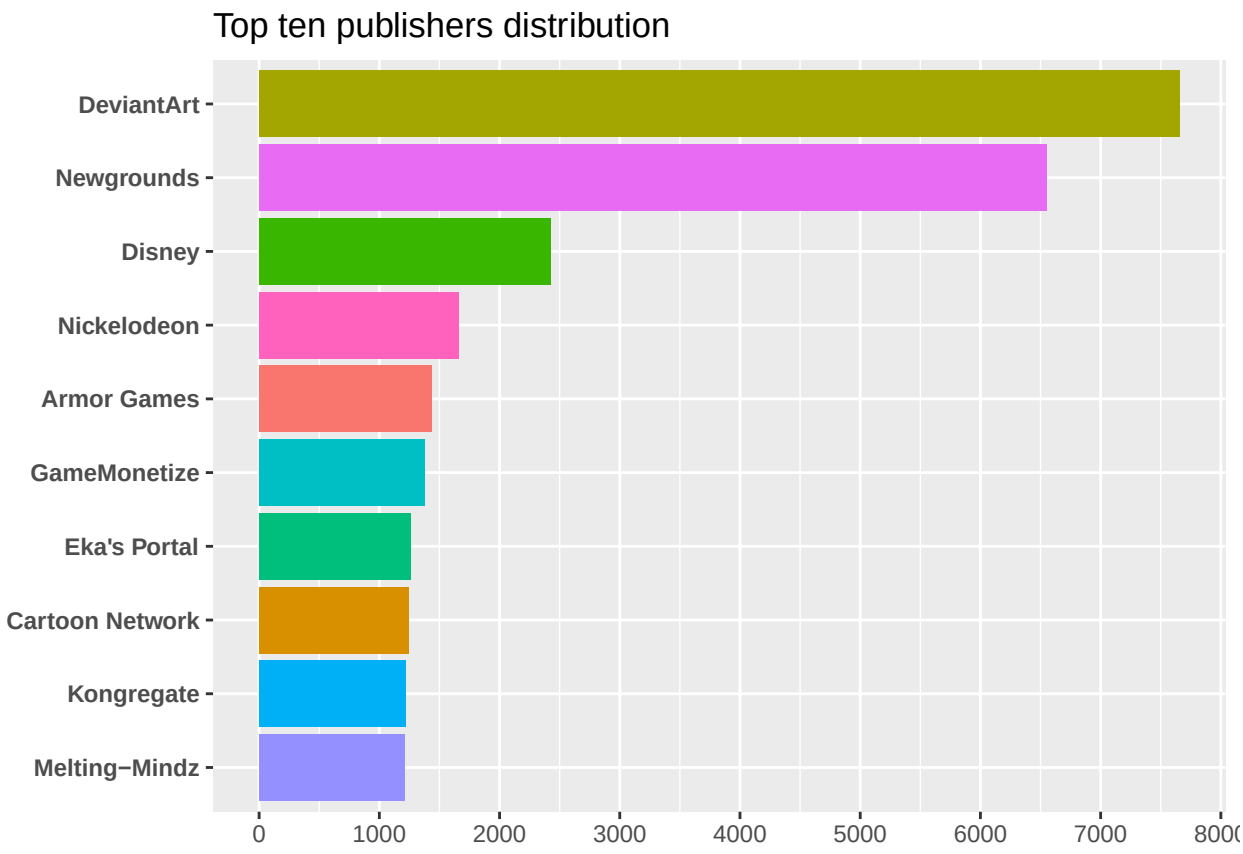

```
sizes <- round(top_developers$n / sum(top_developers$n) * 100, 1)
labels <- vector()

for(i in 1:(length(sizes))){
  labels[i] <- paste(top_developers$developer[i], "-", toString(sizes[i]), "%")
}

top_developers %>% as.data.frame() %>%
  ggplot(aes(x = "", y = sizes, fill = reorder(developer, -n))) +
  geom_bar(stat = "identity", width = 1, color = "black") +
  coord_polar(theta = "y", start = 0) +
  theme(axis.text = element_blank(), axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.title = element_blank()) +
  scale_fill_discrete(labels = labels) +
  ggtitle("Top ten developers distribution")
```



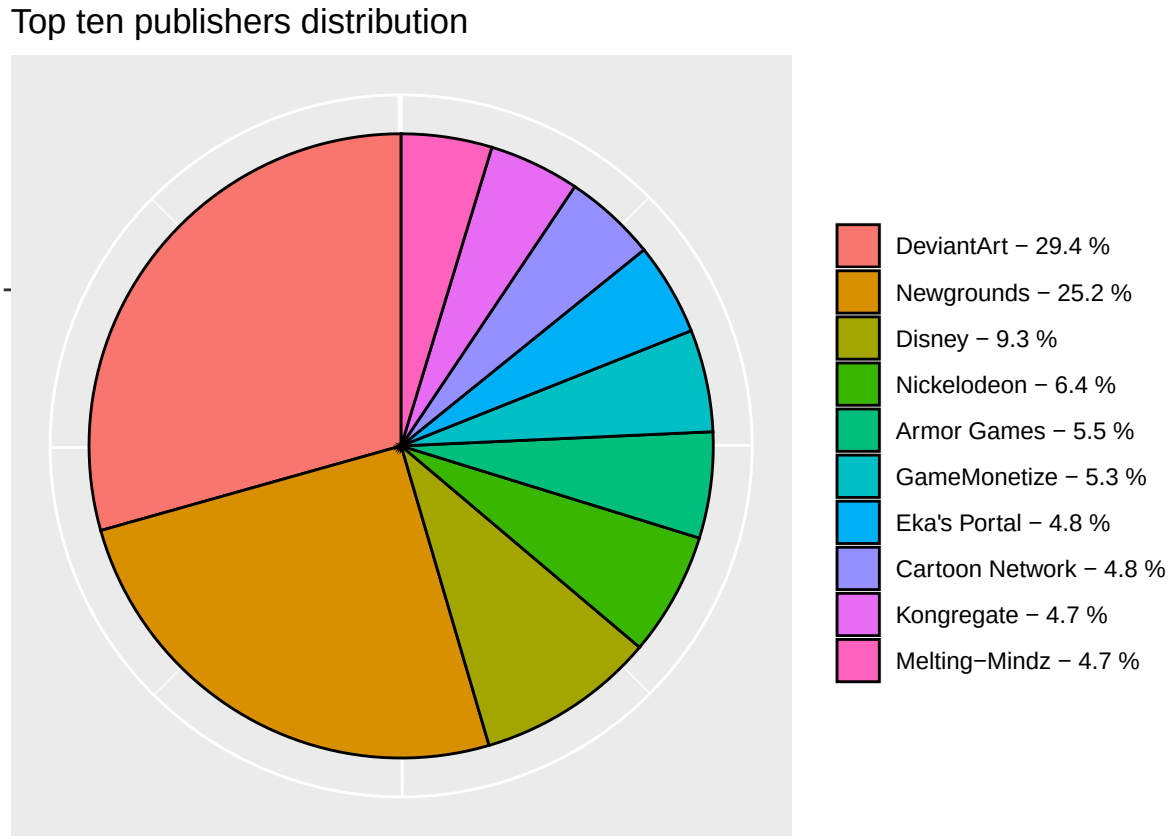

```
top_publishers %>% as.data.frame() %>%
  ggplot(aes(x = reorder(publisher, n), y = n, fill = publisher)) +
  geom_bar(stat = "identity") + coord_flip() +
  scale_y_continuous(breaks = seq(0, 8000, by = 1000)) +
  ggtitle("Top ten publishers distribution") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.y = element_text(face = "bold"))
```




```
sizes <- round(top_publishers$n / sum(top_publishers$n) * 100, 1)
labels <- vector()

for(i in 1:(length(sizes))){
  labels[i] <- paste(top_publishers$publisher[i], "-", toString(sizes[i]), "%")
}

top_publishers %>% as.data.frame() %>%
  ggplot(aes(x = "", y = sizes, fill = reorder(publisher, -n))) +
  geom_bar(stat = "identity", width = 1, color = "black") +
  coord_polar(theta = "y", start = 0) +
  theme(axis.text = element_blank(), axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.title = element_blank()) +
  scale_fill_discrete(labels = labels) +
  ggtitle("Top ten publishers distribution")
```




### Release Dates and Platforms

Flash games started to appear towards the end of the twentieth century and became popular in the next decade. Let's observe the release dates we have got here, being aware that they are not specified for all games.

```
dates <- games %>% filter(releaseDate != "") %>%
  select(title, releaseDate, platform, library) %>%
  arrange(releaseDate)
head(dates)
```

```
##                   title releaseDate  platform library
## 1           TankTrouble  16-12-2007     Flash  arcade
## 2               Blastar        1984     HTML5  arcade
## 3           Idle Johnny        1993 Shockwave theatre
## 4          QP-Shot 1000        1994 Shockwave  arcade
## 5    The Health Checkup        1994 Shockwave theatre
## 6 ZZZ...I want to sleep        1994 Shockwave theatre
```

```
tail(dates)
```

```
##                                       title releaseDate  platform library
## 75992 T-Mobile Tuesdays: Win $2,300 for 2023!  2022-12-27     HTML5  arcade
## 75993         GWL Hayley Footjob (Commission)  2080-10-05     Flash  arcade
```

```
## 75994                  Havok Xtra Marble Demo  21/08/2001 Shockwave  arcade
## 75995                              Ray Cast Car  27/06/2001 Shockwave  arcade
## 75996            (Gift) Luna's Christmas gift  2917-12-26     Flash theatre
## 75997                  Havok Xtra RC Car Demo    7/1/2002 Shockwave  arcade
```

There seems to be a problem with the data. Entries should follow the "YY-MM-DD" date format as per Flashpoint guidelines, but some games come in a different one. In addition, if the exact day or month of release is unknown, it is allowed to specify the year only. Let's clean up our data for consistency.

```
dates$releaseDate <- dates$releaseDate %>% ymd(truncated = 2)
dates <- dates %>% na.omit() %>% arrange(releaseDate)
head(dates)
```

```
##                   title releaseDate  platform library
## 1               Blastar  1984-01-01     HTML5  arcade
## 2           Idle Johnny  1993-01-01 Shockwave theatre
## 3          QP-Shot 1000  1994-01-01 Shockwave  arcade
## 4    The Health Checkup  1994-01-01 Shockwave theatre
## 5 ZZZ...I want to sleep  1994-01-01 Shockwave theatre
## 6         Dangerous Two  1994-01-01 Shockwave theatre
```

```
tail(dates)
```

```
##                                      title releaseDate platform library
## 75982                        Awesome Game  2022-12-23    HTML5  arcade
## 75983                 Christmas in Vienna  2022-12-23    HTML5  arcade
## 75984                   Chrysler Building  2022-12-23    HTML5  arcade
## 75985 T-Mobile Tuesdays: Win $2,300 for 2023!  2022-12-27    HTML5  arcade
## 75986         GWL Hayley Footjob (Commission)  2080-10-05    Flash  arcade
## 75987            (Gift) Luna's Christmas gift  2917-12-26    Flash theatre
```

*ymd* automatically assigned January 1 as month and day for those games whose only date information was the year. There are still a couple of odd observations, the last two, which are most likely typos.

```
dates <- dates %>% slice(1:(nrow(dates)-2))
head(dates, 20)
```

```
##                           title releaseDate  platform library
## 1                       Blastar  1984-01-01     HTML5  arcade
## 2                   Idle Johnny  1993-01-01 Shockwave theatre
## 3                  QP-Shot 1000  1994-01-01 Shockwave  arcade
## 4            The Health Checkup  1994-01-01 Shockwave theatre
## 5          ZZZ...I want to sleep  1994-01-01 Shockwave theatre
## 6                 Dangerous Two  1994-01-01 Shockwave theatre
## 7        Virtual Banana Original  1994-02-01      VRML  arcade
## 8  Virtual University of Auckland  1994-02-01   Hyper-G  arcade
## 9                         Clock  1994-11-17   Hyper-G  arcade
## 10                       Office  1994-11-17   Hyper-G  arcade
## 11   Model of the IICM institute  1994-11-17      VRML  arcade
## 12                        Office  1994-11-17      VRML  arcade
## 13                  Clock Tower  1994-11-17      VRML  arcade
## 14                     Corvette  1994-11-17      VRML  arcade
```

```
## 15                          Cubes  1994-11-17      VRML  arcade
## 16                         Engine  1994-11-17      VRML  arcade
## 17                 Graz 3D Model  1994-11-17      VRML  arcade
## 18                   Fancy Cubes  1994-11-17   Hyper-G  arcade
## 19  The Austrian National Library  1994-11-17      VRML  arcade
## 20                       LED Sign  1995-01-01      Java  arcade
```

Finally, we have got our correct release dates. We see that the oldest game in the list is *Blastar*, which was released in 1984. Actually, the game present in Flashpoint is a HTML5 version, which was developed and released much more recently. Moving on, starting from 1993 we recognize some old technologies, such as *Shochwave*, *VRML* and *Hyper-G*.

We can actually distinguish between proper games and animations by looking at the *library* column: the former are labeled with *arcade*, the latter with *theatre* values. Thus, the oldest animation featured is *Idle Johnny* from 1993, while the first "true" game (not counting *Blastar*) could be either *QP-Shot 1000* (which came out at some time in 1994), or *Virtual Banana Original* and *Virtual University of Auckland*, both from February 1st, 1994.

```
dates <- dates %>% slice(2:nrow(dates))
tail(dates, 20)
```

```
##                                       title releaseDate         platform
## 75965                   Um Conto da Cloe  2022-12-08            HTML5
## 75966                        Pad of Time  2022-12-08            HTML5
## 75967                           /f/ 101  2022-12-09            Flash
## 75968             Vault of the Pineapples  2022-12-09            HTML5
## 75969    Snowsgiving Choose Your Own Adventure  2022-12-10            HTML5
## 75970      look_into_my_eyes_its_open_wide.swf  2022-12-10            Flash
## 75971                    The Smurfs Cooking  2022-12-13            HTML5
## 75972                      Mcdonalds Final  2022-12-13            HTML5
## 75973 In the Name of Freedom: Black Apocalypse  2022-12-13            HTML5
## 75974             Cartoon Network Winter Games  2022-12-14            HTML5
## 75975                          Monkey Mart  2022-12-14            HTML5
## 75976                    Maptroid: Worlds  2022-12-16            HTML5
## 75977                Why is This a Curation?  2022-12-17 Lightning Strike
## 75978                  Monster Bar Simulator  2022-12-18            HTML5
## 75979        Wubbzy explains facts about Mars.  2022-12-19            HTML5
## 75980                  Defender of Ukraine  2022-12-20            HTML5
## 75981                         Awesome Game  2022-12-23            HTML5
## 75982                  Christmas in Vienna  2022-12-23            HTML5
## 75983                    Chrysler Building  2022-12-23            HTML5
## 75984  T-Mobile Tuesdays: Win $2,300 for 2023!  2022-12-27            HTML5
##       library
## 75965  arcade
## 75966  arcade
## 75967  arcade
## 75968  arcade
## 75969  arcade
## 75970 theatre
## 75971  arcade
## 75972  arcade
## 75973  arcade
## 75974  arcade
## 75975  arcade
```

```
## 75976  arcade
## 75977  arcade
## 75978  arcade
## 75979 theatre
## 75980  arcade
## 75981  arcade
## 75982  arcade
## 75983  arcade
## 75984  arcade
```

On the other side, here are the 20 most recent games. As expected, we find out that nowadays *HTML5* is the standard technology to make flash games, though some exceptions arise and there is even a *Flash* entry!

For the sake of completeness, let's restrict our search to *Flash*-only games.

```
dates %>% filter(platform == "Flash") %>% head(20)
```

```
##                                       title releaseDate platform library
## 1                                 Claus.com  1995-01-01    Flash  arcade
## 2             2 Design's Navigational Demo  1996-01-01    Flash  arcade
## 3                           CHAOS Website  1996-01-01    Flash  arcade
## 4        FutureWave Software, Inc. Website  1996-01-01    Flash  arcade
## 5               Good Music Company Website  1996-01-01    Flash  arcade
## 6                         The Silicon Slip  1996-01-01    Flash  arcade
## 7                        Zygomedia Website  1996-01-01    Flash  arcade
## 8                   First MouseOver Button  1996-01-01    Flash  arcade
## 9                    Simple, Tasty Buttons  1996-01-01    Flash  arcade
## 10                       Discrete Keyboard  1996-01-01    Flash  arcade
## 11                   Animated Screen Beans  1996-01-01    Flash theatre
## 12     HK Media Association website banner  1996-01-01    Flash theatre
## 13      The Hole in the Wall website intro  1996-01-01    Flash theatre
## 14                    Volleynerd home page  1996-01-01    Flash theatre
## 15                     Past*Present*Future  1996-01-01    Flash theatre
## 16                The Simpsons 1996 website  1996-01-01    Flash theatre
## 17 The Simpsons 1996 website normal version  1996-01-01    Flash  arcade
## 18                                  Artcore  1996-01-01    Flash  arcade
## 19                          Advent Calendar  1997-01-01    Flash  arcade
## 20                       Ray of Light Button  1997-01-01    Flash  arcade
```

The first *Flash* game is *Claus.com* from 1995. We notice that most of these are actually websites built in *Flash* and not exactly games or animations.

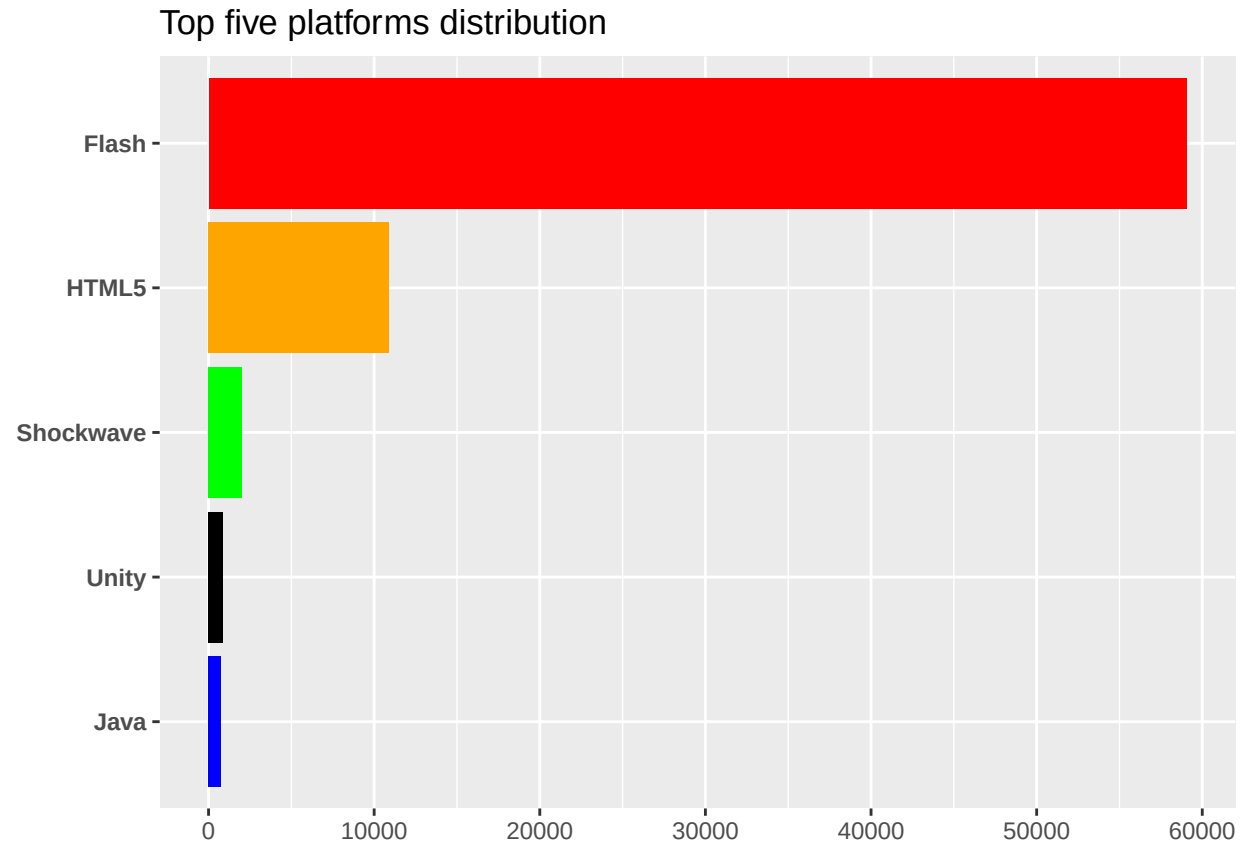
To take an overall view, let's compare the various platforms by games count, considering the top five.

```
top_platforms <- dates %>% count(platform) %>% arrange(desc(n)) %>% head(5)
top_platforms
```

```
##    platform     n
## 1     Flash 59047
## 2     HTML5 10904
## 3 Shockwave  1996
## 4     Unity   875
## 5      Java   751
```

```
colors <- c("red", "orange", "blue", "green", "black")

top_platforms %>% as.data.frame() %>% arrange(platform) %>%
  ggplot(aes(x = reorder(platform, n), y = n, fill = platform)) +
  geom_bar(stat = "identity") + coord_flip() +
  scale_y_continuous(breaks = seq(0, 60000, by = 10000)) +
  ggtitle("Top five platforms distribution") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(), axis.text.y = element_text(face = "bold")) +
  scale_fill_manual(values = colors)
```



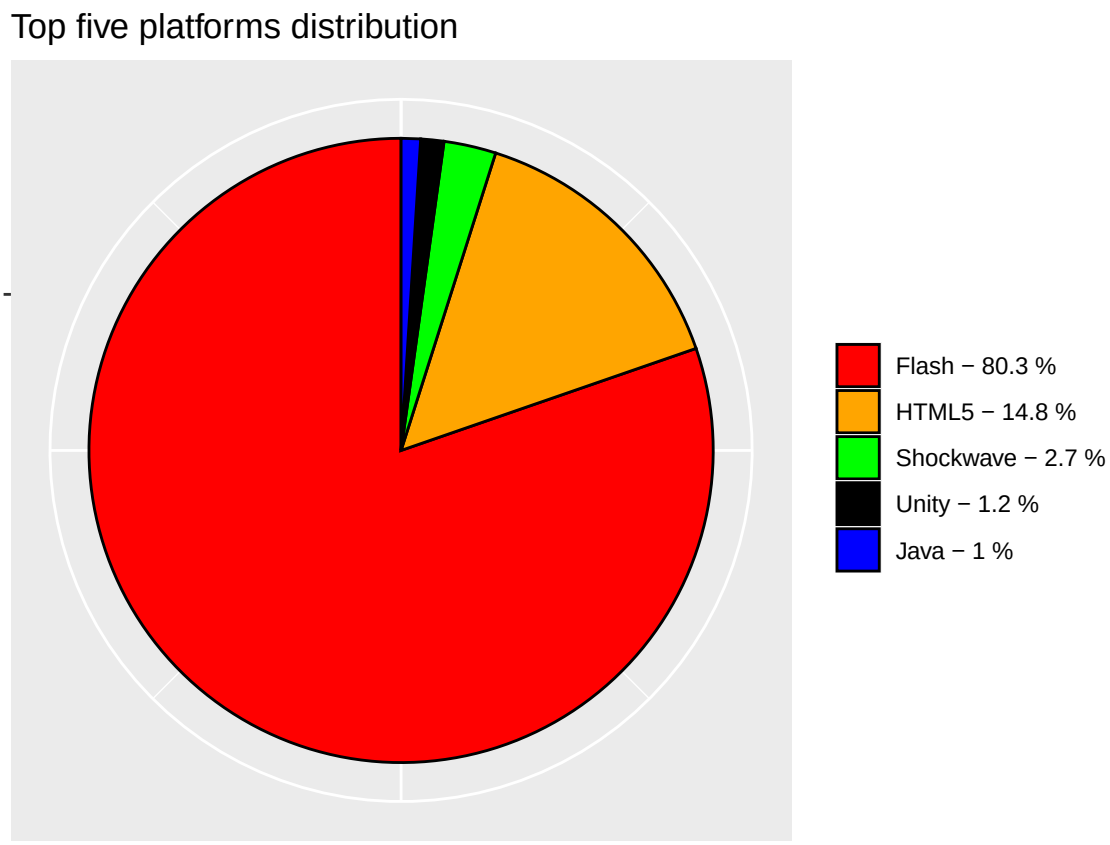

```
sizes <- round(top_platforms$n / sum(top_platforms$n) * 100, 1)
labels <- vector()

for(i in 1:(length(sizes))){
  labels[i] <- paste(top_platforms$platform[i], "-", toString(sizes[i]), "%")
}

colors <- c("red", "orange", "green", "black", "blue")

top_platforms %>% as.data.frame() %>%
  ggplot(aes(x = "", y = sizes, fill = reorder(platform, -n))) +
  geom_bar(stat = "identity", width = 1, color = "black") +
  coord_polar(theta = "y", start = 0) +
  theme(axis.text = element_blank(), axis.title.x = element_blank(),
```

```
      axis.title.y = element_blank(), legend.title = element_blank()) +
  scale_fill_manual(labels = labels, values = colors) +
  ggtitle("Top five platforms distribution")
```



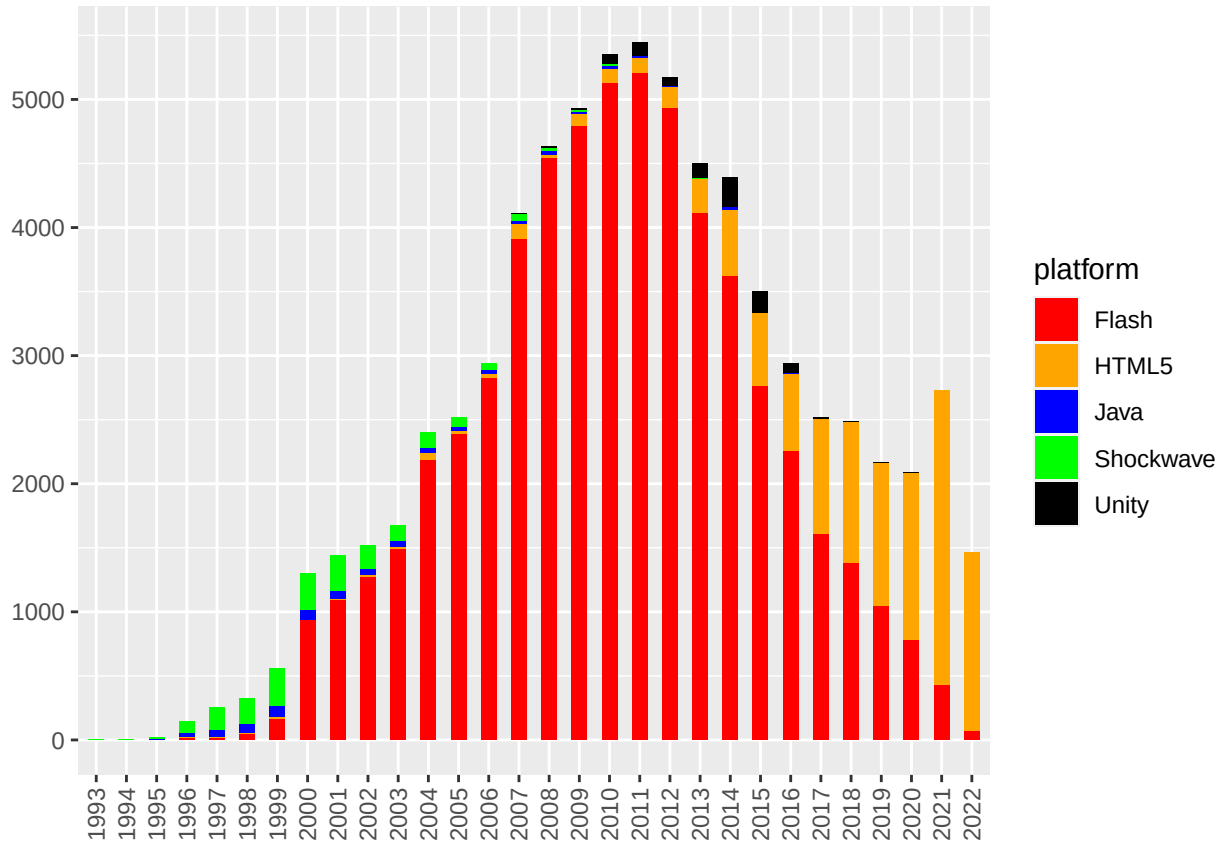

*Flash* is clearly the winner, followed by a rising *HTML5* and its old companion *Shockwave*, with *Unity* and *Java* as outsiders.

Web games were at their peak in the 2000s and many gamers are nostalgic about that decade, which could be considered a golden age. Thus, we expect to see that most of the games in our database have been released between 2000 and 2009. Let's check it out, while also comparing technologies against years.

```
year_platform <- dates %>% filter(platform %in% top_platforms$platform)
years <- year_platform$releaseDate %>% sapply(substr, start = 1, stop = 4)
year_platform$releaseDate <- years

colors <- c("red", "orange", "blue", "green", "black")

year_platform %>% count(releaseDate, platform) %>%
  ggplot(aes(x = releaseDate, y = n, fill = platform, width = 0.5)) +
  geom_bar(position = position_stack(reverse = TRUE), stat = "identity") +
  scale_y_continuous(breaks = seq(0, 6000, by = 1000)) +
  theme(axis.title.x = element_blank(),
        axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1),
        axis.title.y = element_blank()) +
  scale_fill_manual(values = colors)
```




*Et voila*! As we were expecting, web games have steadily risen in popularity in the first decade of the third millennium, reached a peak in 2011 and today their number is slowly decreasing, apart from some fluctuations. This does not mean that fewer games are being made: it could simply be that there is less incentive to curate and preserve a recent game written in a technology which will probably stay on for a long time compared to an old game with a nostalgic value that runs the risk to disappear at any time.

*Flash* dominated the scene between 2000 and 2017 (it's impressive to see that it lasted so long), while *HTML5* started to be relevant around 2013. *Shockwave* was most popular between 1996 and 2000, but continued to be used until 2007. *Unity* had six years of relative notoriety (2010-2016) and lastly, Java, despite being one of the first technologies eligible for making web games, has never known much use and moved off the radar around 2010.

## Most common languages

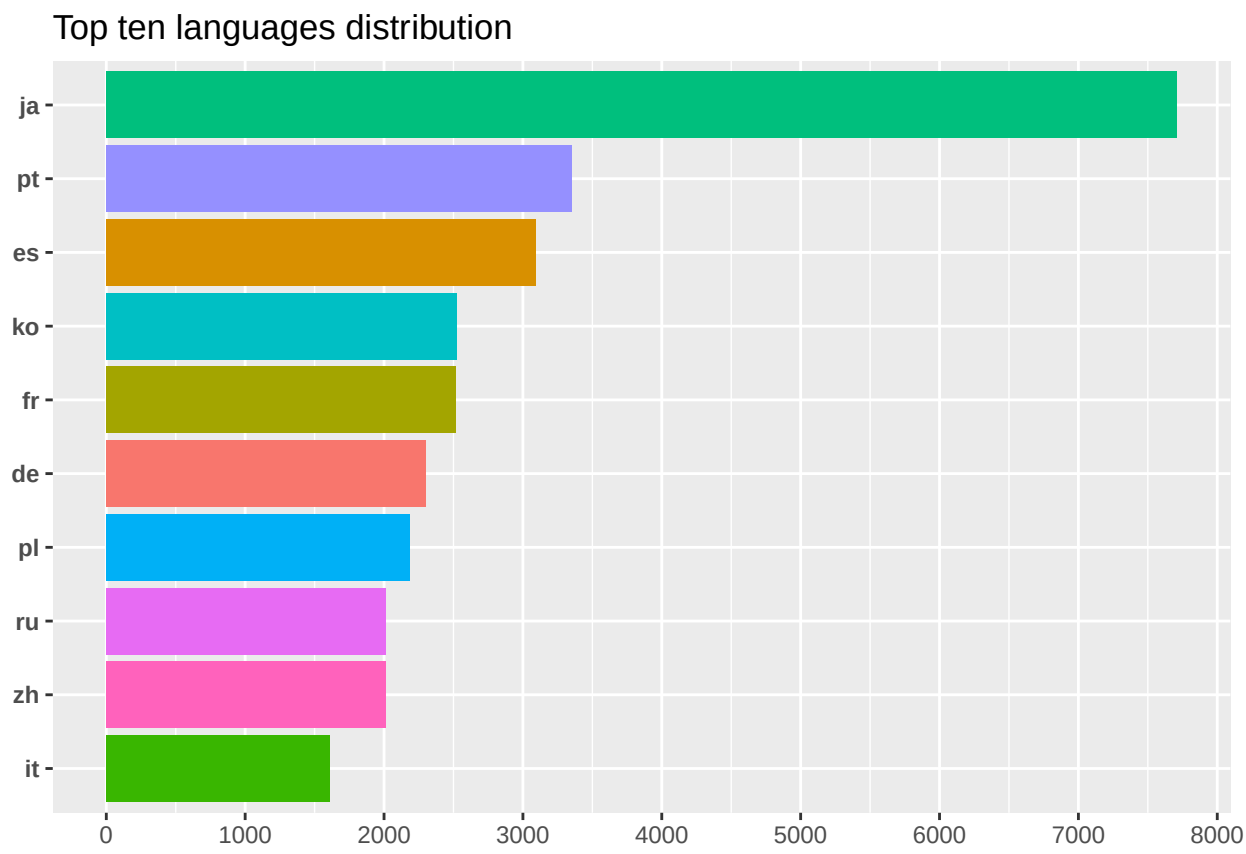
Let's move on to another topic: *Flashpoint* allows non-English content as well, and it can be interesting to know which countries have contributed the most to the world of web games aside from the anglophone ones.

```
top_languages <- games %>% filter(language != "") %>% { . ->> tmp } %>%
  select(language) %>%
  vapply(gsub, pattern = ",", replacement = ";", character(nrow(tmp))) %>%
  as.data.frame() %>%
  separate_rows(language, sep = "; ") %>% count(language) %>%
  arrange(desc(n)) %>%
  filter(language != "en") %>% head(10)
top_languages
```

```
## # A tibble: 10 x 2
##    language     n
##    <chr>    <int>
##  1 ja        7709
##  2 pt        3350
##  3 es        3095
##  4 ko        2525
##  5 fr        2519
##  6 de        2298
##  7 pl        2183
##  8 ru        2015
##  9 zh        2011
## 10 it        1612
```

Since a game can come in different languages, separated by a colon and a space (";"), we had to count each occurrence individually.

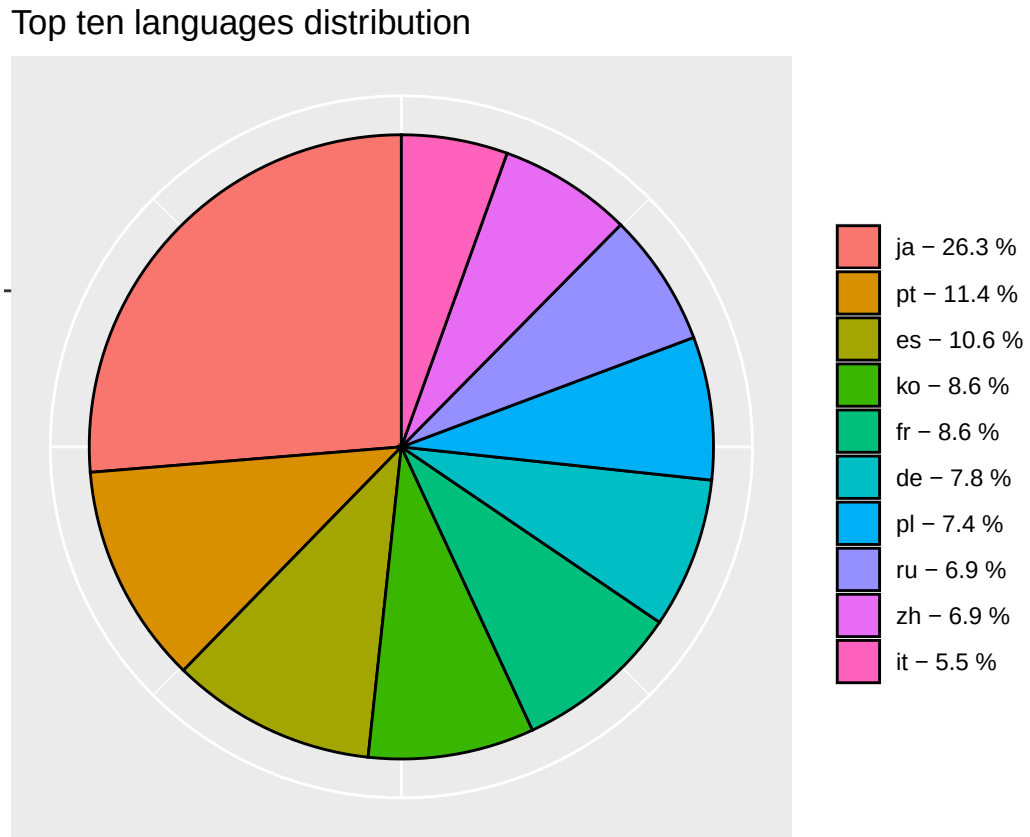
```
top_languages %>% as.data.frame() %>% arrange(language) %>%
  ggplot(aes(x = reorder(language, n), y = n, fill = language)) +
  geom_bar(stat = "identity") + coord_flip() +
  scale_y_continuous(breaks = seq(0, 8000, by = 1000)) +
  ggtitle("Top ten languages distribution") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(), axis.text.y = element_text(face = "bold"))
```



```
sizes <- round(top_languages$n / sum(top_languages$n) * 100, 1)
labels <- vector()

for(i in 1:(length(sizes))){
  labels[i] <- paste(top_languages$language[i], "-", toString(sizes[i]), "%")
}

top_languages %>% as.data.frame() %>%
  ggplot(aes(x = "", y = sizes, fill = reorder(language, -n))) +
  geom_bar(stat = "identity", width = 1, color = "black") +
  coord_polar(theta = "y", start = 0) +
  theme(axis.text = element_blank(), axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.title = element_blank()) +
  scale_fill_discrete(labels = labels) +
  ggtitle("Top ten languages distribution")
```




We can see a strong presence of Asian content, with Japanese, Korean and Chinese among the top ten languages. The rest of the list is completed by European countries, namely Portugal, Spain, France, Germany, Poland and Italy, as well as Russia.

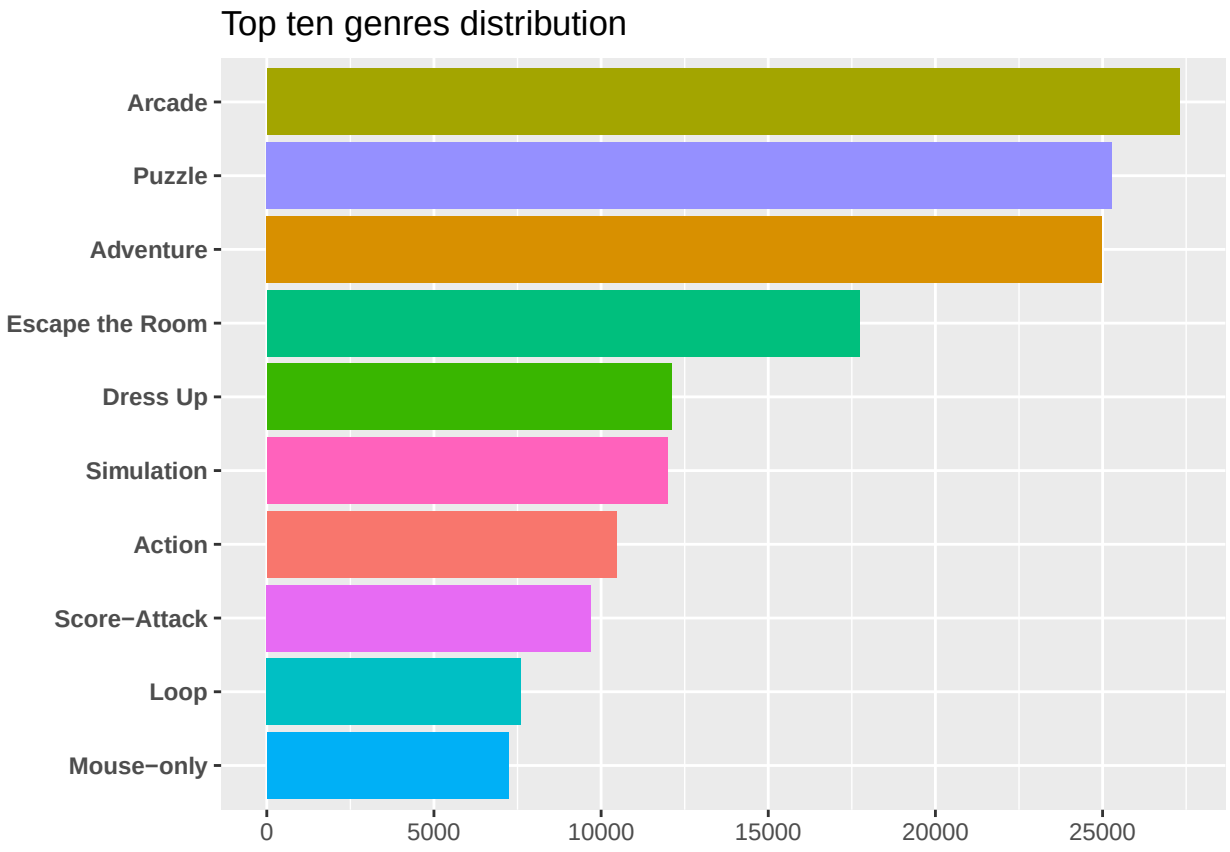
## Most popular genres

Let's now focus on game genres, featured on the *tagsStr* column, to discover the most common ones.

```r
top_genres <- games %>% filter(tagsStr != "") %>% { . ->> tmp } %>%
  select(tagsStr) %>%
  vapply(gsub, pattern = ",", replacement = ";", character(nrow(tmp))) %>%
  as.data.frame() %>%
  separate_rows(tagsStr, sep = "; ") %>% count(tagsStr) %>%
  arrange(desc(n)) %>% head(10)
top_genres
```

```
## # A tibble: 10 x 2
##    tagsStr             n
##    <chr>           <int>
##  1 Arcade          27308
##  2 Puzzle          25284
##  3 Adventure       24981
##  4 Escape the Room 17725
##  5 Dress Up        12101
##  6 Simulation      11993
##  7 Action          10452
##  8 Score-Attack     9703
##  9 Loop             7599
## 10 Mouse-only       7232
```

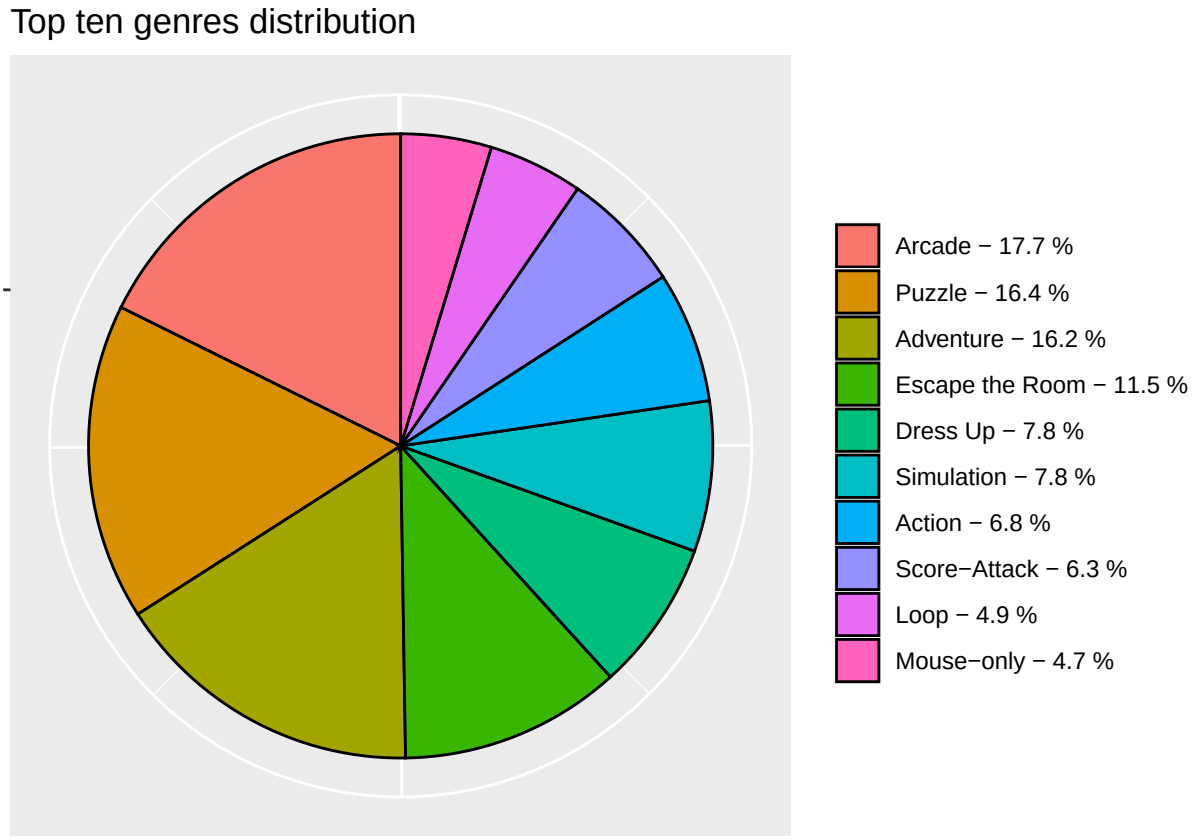
```r
top_genres %>% as.data.frame() %>%
  ggplot(aes(x = reorder(tagsStr, n), y = n, fill = tagsStr)) +
  geom_bar(stat = "identity") + coord_flip() +
  scale_y_continuous(breaks = seq(0, 30000, by = 5000)) +
  ggtitle("Top ten genres distribution") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.y = element_text(face = "bold"))
```




```
sizes <- round(top_genres$n / sum(top_genres$n) * 100, 1)
labels <- vector()

for(i in 1:(length(sizes))){
  labels[i] <- paste(top_genres$tagsStr[i], "-", toString(sizes[i]), "%")
}

top_genres %>% as.data.frame() %>%
  ggplot(aes(x = "", y = sizes, fill = reorder(tagsStr, -n))) +
  geom_bar(stat = "identity", width = 1, color = "black") +
  coord_polar(theta = "y", start = 0) +
  theme(axis.text = element_blank(), axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        legend.title = element_blank()) +
  scale_fill_discrete(labels = labels) +
  ggtitle("Top ten genres distribution")
```




The big three genres are *Arcade*, *Puzzle* and *Adventure* and honestly it's kind of odd to see *Action* at such a low position. Conversely, as we expected from our previous analysis on developers, *Escape the Room* is fairly popular, along with *Dress Up* and *Simulation* games.

## Most played games

As a final insight, let's find out which are the most played games among the *Flashpoint* users: to do this, we are going to use some official statistics from the platform itself. Visit https://flashpoint-analytics.unstable.life/, scroll down to the corresponding section and download the data in .csv format.

```
most_played <- read.csv("data/most_played.csv") %>% rename(id = category)
most_played
```

```
##                                     id Play.Count
## 1  83e1b5e7-4282-4bbd-868e-dcfa965e4abf      48493
## 2  a94d865c-cb38-4d31-96f3-dda26502c4a3      10228
## 3  617ca7f3-1cff-3f0c-5b53-07498b3b28d8       7454
## 4  8d09fc0d-6f25-4be6-b396-8fcaddad4e5e       7253
## 5  b0ce771e-7c02-4317-8528-ba48139e2688       6922
## 6  fdee4800-b5c9-49e0-b19e-22f2b0ccab68       6919
## 7  1e903a30-5c37-15bb-8e5e-6fea5a8103f2       6678
## 8  07921a2f-26fd-4364-9671-ee0c8d256ec1       6423
## 9  b9a8dbb9-0cd7-434b-b226-13dc9dd07b49       6341
## 10 a8707c0f-6aac-4c94-9d8a-cc397c97cc88       5745
## 11 92ba2d91-e041-2bd4-49ea-21758df711ff       5613
```

```
## 12 f55d6576-1aee-414a-bd1e-b7678d697dcf       5605
## 13 16b04977-f714-4239-b343-b759e16a33af       5579
## 14 5fa91cac-25c9-2a53-f391-09a099cb489b       5203
## 15 ed058412-fa96-4ea5-b2cb-2baff6a24b2a       5197
## 16 3337a4a2-dacf-4027-824c-aaf77623de65       4842
## 17 164e2dc8-31dd-4f59-a6be-46c2087e190f       4760
## 18 92781aea-6ad0-2bdb-a963-7131c49b7d07       4628
## 19 602762ba-a1b6-4810-8093-e5dbf0c33b61       4521
## 20 938f4383-8c92-8e8d-511b-004b5c69999d       4427
## 21 279b23ca-c7f3-4f56-a02f-3752ec7c5c5d       4407
## 22 6b50cb43-ad66-4ec8-9e46-3bb06821f2ab       4332
## 23 228087d3-c291-4a87-82e6-c6ffd3d4d2f6       4176
## 24 dbb21635-b0d5-78d9-a749-c4778a07e698       3958
## 25 de163cff-9dbe-fcbb-164c-53f5d6873fad       3883
## 26 16c27895-f7a6-6c65-18f3-feebcf87d28a       3874
## 27 c3ab2546-a7b0-89b2-82de-4044d61e1cbd       3684
## 28 b0d2b9a9-ab00-465e-b5a3-56031b92f070       3625
## 29 8f7f9fe8-4c55-43b1-a574-046c63712b39       3477
## 30 2ecf56d6-c5e4-a801-bc7f-60374ba1a051       3410
## 31 4423194a-22fe-427a-8eae-1b4d9c42395c       3247
## 32 9525910d-72b8-4e84-b668-43a267e00d9c       3238
## 33 c85ff4e3-1e2d-4b4e-8c4f-958d9db4aff2       3237
## 34 da6b3cb4-78a7-e998-fcf4-cb26a7950754       3227
## 35 2d0071b4-8ea9-40b1-a642-c970ca260cb2       3157
## 36 ab638461-9317-bbaa-9ccf-e7f0360e3b1b       3136
## 37 cf29b65d-e3e9-4da9-8201-d76de38736ea       3036
## 38 15ac0ed0-dbd4-8a95-5eec-4cf3e04fc771       3017
## 39 190c1bf6-4fcc-4b32-8278-f5b26db8eec1       2850
## 40 242dedc0-431a-4e4e-990c-58fe3c8ef740       2822
```

The file contains the *id* for the most 40 played games, along with a play count. Let's use the identifiers to find the titles of these games and their other info.

```
rank <- inner_join(games, most_played, by = join_by(id)) %>% arrange(desc(Play.Count))
rank
```

```
##                                     id
## 1  83e1b5e7-4282-4bbd-868e-dcfa965e4abf
## 2  a94d865c-cb38-4d31-96f3-dda26502c4a3
## 3  617ca7f3-1cff-3f0c-5b53-07498b3b28d8
## 4  8d09fc0d-6f25-4be6-b396-8fcaddad4e5e
## 5  b0ce771e-7c02-4317-8528-ba48139e2688
## 6  fdee4800-b5c9-49e0-b19e-22f2b0ccab68
## 7  1e903a30-5c37-15bb-8e5e-6fea5a8103f2
## 8  07921a2f-26fd-4364-9671-ee0c8d256ec1
## 9  b9a8dbb9-0cd7-434b-b226-13dc9dd07b49
## 10 a8707c0f-6aac-4c94-9d8a-cc397c97cc88
## 11 92ba2d91-e041-2bd4-49ea-21758df711ff
## 12 f55d6576-1aee-414a-bd1e-b7678d697dcf
## 13 16b04977-f714-4239-b343-b759e16a33af
## 14 5fa91cac-25c9-2a53-f391-09a099cb489b
## 15 ed058412-fa96-4ea5-b2cb-2baff6a24b2a
## 16 3337a4a2-dacf-4027-824c-aaf77623de65
## 17 164e2dc8-31dd-4f59-a6be-46c2087e190f
```

```
## 18 92781aea-6ad0-2bdb-a963-7131c49b7d07
## 19 602762ba-a1b6-4810-8093-e5dbf0c33b61
## 20 938f4383-8c92-8e8d-511b-004b5c69999d
## 21 279b23ca-c7f3-4f56-a02f-3752ec7c5c5d
## 22 6b50cb43-ad66-4ec8-9e46-3bb06821f2ab
## 23 228087d3-c291-4a87-82e6-c6ffd3d4d2f6
## 24 dbb21635-b0d5-78d9-a749-c4778a07e698
## 25 de163cff-9dbe-fcbb-164c-53f5d6873fad
## 26 16c27895-f7a6-6c65-18f3-feebcf87d28a
## 27 c3ab2546-a7b0-89b2-82de-4044d61e1cbd
## 28 b0d2b9a9-ab00-465e-b5a3-56031b92f070
## 29 8f7f9fe8-4c55-43b1-a574-046c63712b39
## 30 2ecf56d6-c5e4-a801-bc7f-60374ba1a051
## 31 4423194a-22fe-427a-8eae-1b4d9c42395c
## 32 9525910d-72b8-4e84-b668-43a267e00d9c
## 33 c85ff4e3-1e2d-4b4e-8c4f-958d9db4aff2
## 34 da6b3cb4-78a7-e998-fcf4-cb26a7950754
## 35 2d0071b4-8ea9-40b1-a642-c970ca260cb2
## 36 ab638461-9317-bbaa-9ccf-e7f0360e3b1b
## 37 cf29b65d-e3e9-4da9-8201-d76de38736ea
## 38 15ac0ed0-dbd4-8a95-5eec-4cf3e04fc771
## 39 190c1bf6-4fcc-4b32-8278-f5b26db8eec1
## 40 242dedc0-431a-4e4e-990c-58fe3c8ef740
##                                        title
## 1                                  Poptropica
## 2                                   Jacksmith
## 3                             Papa's Cheeseria
## 4                        Strike Force Heroes 3
## 5                  Super Mario Bros. Crossover
## 6                          Strike Force Heroes
## 7                             Papa's Scooperia
## 8                                  Bloons TD 5
## 9                              Papa's Sushiria
## 10                              Papa's Bakeria
## 11                             Papa's Freezeria
## 12                              Papa's Burgeria
## 13                                Super Mario 63
## 14                       Epic Battle Fantasy 5
## 15                         Super Smash Flash 2
## 16                      Flappy Bird For Dinner
## 17                      Madness: Project Nexus
## 18                         Papa's Hot Doggeria
## 19                            Swords and Souls
## 20                            Papa's Donuteria
## 21                             Papa's Pizzeria
## 22                       Strike Force Heroes 2
## 23                         The Impossible Quiz
## 24                        Ben 10: Battle Ready
## 25                           Papa's Cupcakeria
## 26         Papa Louie 2: When Burgers Attack!
## 27                             Papa's Wingeria
## 28                  Portal: The Flash Version
## 29                The Last Stand: Union City
## 30                             Papa's Pastaria
```

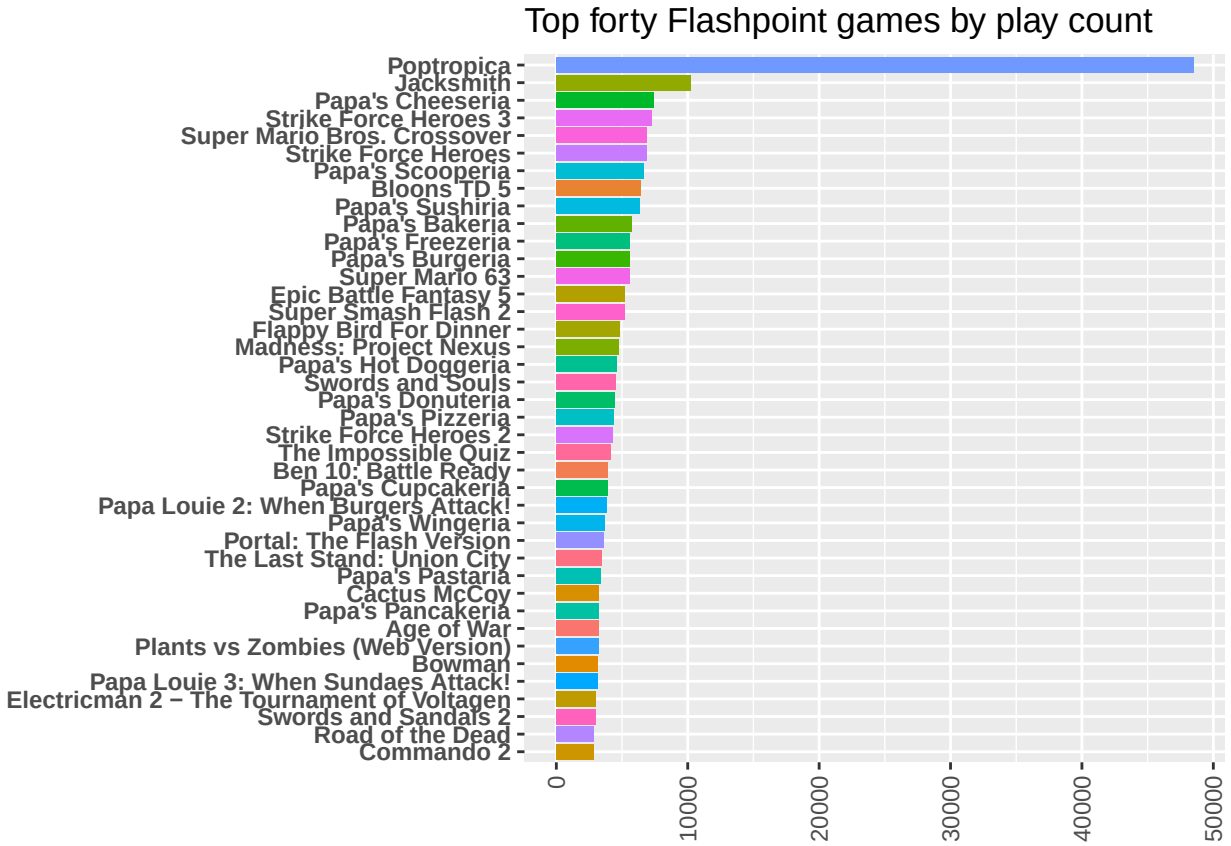
```
## 31                              Cactus McCoy
## 32                          Papa's Pancakeria
## 33                                Age of War
## 34           Plants vs Zombies (Web Version)
## 35                                    Bowman
## 36        Papa Louie 3: When Sundaes Attack!
## 37 Electricman 2 - The Tournament of Voltagen
## 38                       Swords and Sandals 2
## 39                           Road of the Dead
## 40                                 Commando 2
##                                  developer       publisher  platform
## 1  Sandbox Networks Inc.; Pearson Education                      Flash
## 2                           Flipline Studios   PapaLouie.com     Flash
## 3                           Flipline Studios   PapaLouie.com     Flash
## 4                                 Sky9 Games     Armor Games     Flash
## 5                           Exploding Rabbit                     Flash
## 6                                 Sky9 Games     Armor Games     Flash
## 7                           Flipline Studios   PapaLouie.com     Flash
## 8                                 Ninja Kiwi      Ninja Kiwi     Flash
## 9                           Flipline Studios   PapaLouie.com     Flash
## 10                          Flipline Studios   PapaLouie.com     Flash
## 11                          Flipline Studios     Armor Games     Flash
## 12                          Flipline Studios                     Flash
## 13                                    Runouw      Newgrounds     Flash
## 14                                Kupo Games                     Flash
## 15                             McLeodGaming    McLeodGaming     Flash
## 16                                            Cooking Games     Flash
## 17      Michael Swain & Matt Krinkels Jolly      Newgrounds     Flash
## 18                          Flipline Studios   PapaLouie.com     Flash
## 19                                  SoulGame     Armor Games     Flash
## 20                          Flipline Studios   PapaLouie.com     Flash
## 21                          Flipline Studios                     Flash
## 22                                Sky9 Games     Armor Games     Flash
## 23                             Splapp-Me-Do      Newgrounds     Flash
## 24                                           Cartoon Network Shockwave
## 25                          Flipline Studios   PapaLouie.com     Flash
## 26                          Flipline Studios   PapaLouie.com     Flash
## 27                          Flipline Studios   PapaLouie.com     Flash
## 28                           We Create Stuff     Armor Games     Flash
## 29                                ConArtists     Armor Games     Flash
## 30                          Flipline Studios   PapaLouie.com     Flash
## 31                          Flipline Studios                     Flash
## 32                          Flipline Studios   PapaLouie.com     Flash
## 33                                   Louissi       Max Games     Flash
## 34                                    PopCap                     Flash
## 35                                     f1ash                     Flash
## 36                          Flipline Studios   PapaLouie.com     Flash
## 37                             Damien Clarke     Armor Games     Flash
## 38                              3rdsense.com                     Flash
## 39                                  Evil Dog      Newgrounds     Flash
## 40                               Miniclip.com    Miniclip.com     Flash
##    releaseDate language library
## 1      2007-09       en  arcade
## 2   2012-09-27       en  arcade
```

```
## 3   2015-06-10       en  arcade
## 4                     en  arcade
## 5   2010-04-27       en  arcade
## 6   2012-05-31       en  arcade
## 7   2018-07-24       en  arcade
## 8   2011-12-13       en  arcade
## 9   2016-12-13       en  arcade
## 10  2016-03-14       en  arcade
## 11  2011-08-05       en  arcade
## 12  2010-12-06       en  arcade
## 13  2009-06-10       en  arcade
## 14  2020-01-21           arcade
## 15  2007-12-25       en  arcade
## 16                    en  arcade
## 17  2012-03-25       en  arcade
## 18  2012-11-19       en  arcade
## 19  2015-10-18       en  arcade
## 20  2014-06-16       en  arcade
## 21  2007-08-07       en  arcade
## 22                    en  arcade
## 23  2007-02-20       en  arcade
## 24               en; ja  arcade
## 25  2013-08-07       en  arcade
## 26  2013-06-05       en  arcade
## 27  2012-06-13       en  arcade
## 28  2007-11-19       en  arcade
## 29  2011-07-29           arcade
## 30  2013-12-09       en  arcade
## 31  2011-03-10       en  arcade
## 32  2012-03-05       en  arcade
## 33  2008-01-14       en  arcade
## 34                    en  arcade
## 35  2004-12-05       en  arcade
## 36  2015-03-04       en  arcade
## 37  2008-01-08       en  arcade
## 38  2007-01-07       en  arcade
## 39  2010-10-13       en  arcade
## 40  2008-10-03           arcade
##
## 1
## 2
## 3
## 4
## 5
## 6
## 7
## 8
## 9
## 10
## 11
## 12
## 13
## 14
## 15 Bleach; Bomberman; Chibi-Robo!; F-Zero; Fire Emblem; Game & Watch; Golden Sun; Kid Icarus; Rayman
```

```
## 16
## 17
## 18
## 19
## 20
## 21
## 22
## 23
## 24
## 25
## 26
## 27
## 28
## 29
## 30
## 31
## 32
## 33
## 34
## 35
## 36
## 37
## 38
## 39
## 40
##    Play.Count
## 1       48493
## 2       10228
## 3        7454
## 4        7253
## 5        6922
## 6        6919
## 7        6678
## 8        6423
## 9        6341
## 10       5745
## 11       5613
## 12       5605
## 13       5579
## 14       5203
## 15       5197
## 16       4842
## 17       4760
## 18       4628
## 19       4521
## 20       4427
## 21       4407
## 22       4332
## 23       4176
## 24       3958
## 25       3883
## 26       3874
## 27       3684
## 28       3625
```

```
## 29       3477
## 30       3410
## 31       3247
## 32       3238
## 33       3237
## 34       3227
## 35       3157
## 36       3136
## 37       3036
## 38       3017
## 39       2850
## 40       2822
```

```
rank %>% ggplot(aes(x = reorder(title, Play.Count), y = Play.Count, fill = title)) +
  geom_bar(stat = "identity") + coord_flip() +
  scale_y_continuous(breaks = seq(0, 50000, by = 10000)) +
  ggtitle("Top forty Flashpoint games by play count") +
  theme(legend.position = "none", axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1),
        axis.text.y = element_text(face = "bold"))
```




There we go! *Poptropica* is the indisputable winner, with almost fifty thousand play counts. There is a massive presence of *Papa's Gameria* franchise, as well as all-time classics like *Strike Force Heroes*, *Super Mario 63* and *Age of War*. Finally, a special remark about *Ben 10: Battle Ready*, which was thought to be lost forever, before it was restored and made playable again on *Flashpoint*.

## Conclusion

This was a thorough analysis of the *Flashpoint* catalogue, which hopefully gives some insights about the world of web-based games and their significant relevance in the history of the Internet.

The effort to preserve this kind of content has generated amazing results, saving an astounding quantity of material which would have disappeared otherwise. Despite the concrete risk of a digital dark age, we should insist on preserving the stuff that we care about and keep it alive, not only for historical reasons, but also for the nostalgic value we associate with it.