

# 电子游戏市场发展原因探究及发展方向

## 数据集及研究目标

### 数据集

本数据集收集了1980年至2020年电子游戏的相关信息，包括名称，平台，发行年份，类型，发行商，各地区销量等。

### 研究目标

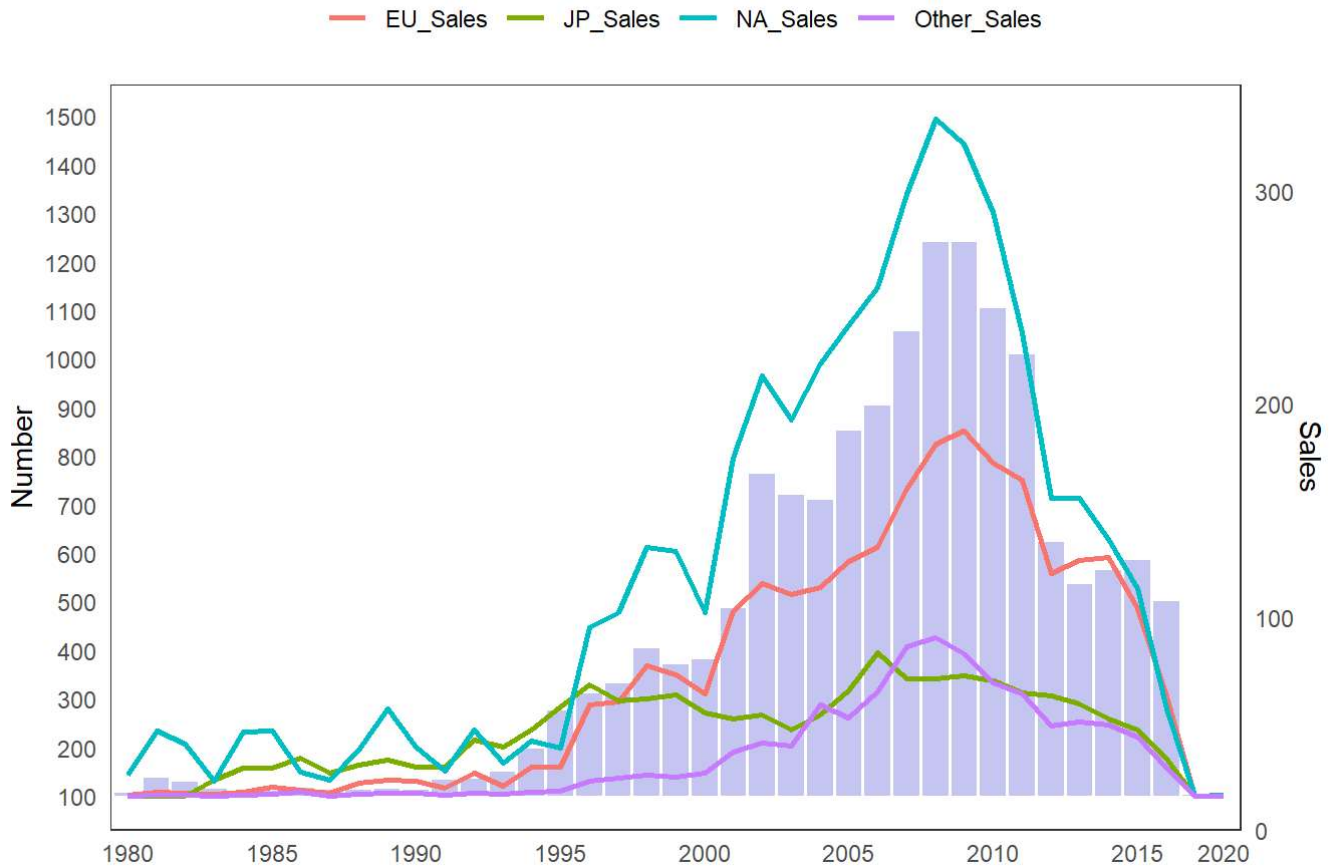
本研究主要有以下几个研究目标：电子游戏市场整体发展情况；游戏市场迅速发展与衰退的原因；今后游戏内容发展的方向。旨在通过可视化分析来帮助游戏发行商和开发商明确行业发展趋势，从而调整产品以适应市场。

## 电子游戏市场发展

随着手机与PC的技术不断进步，电子游戏在上世纪八十年代逐渐兴起，再到逐渐没落。

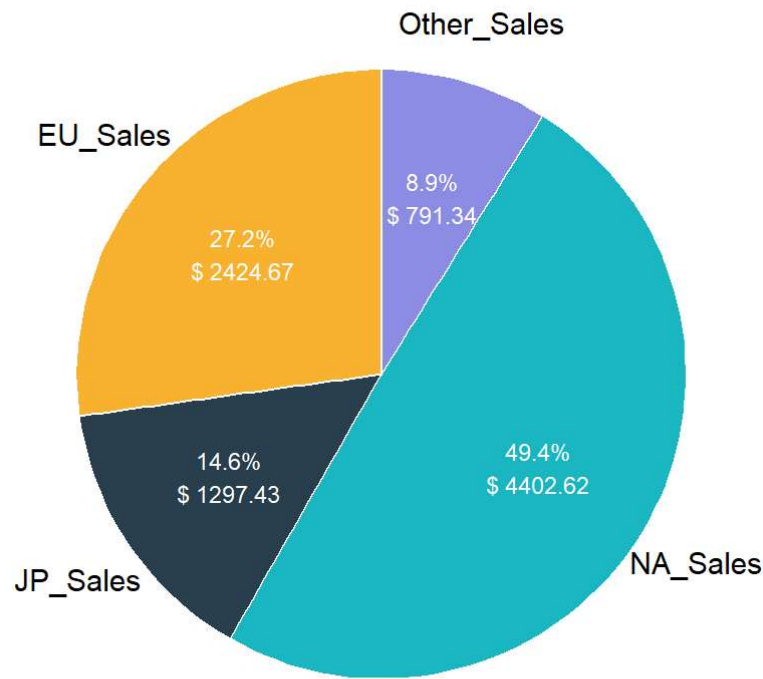
```
Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure") %>%
  add_count(Year_of_Release) %>%
  group_by(Year_of_Release) %>%
  summarise(number=n(), NA_Sales = sum(NA_Sales), EU_Sales= sum(EU_Sales),
            JP_Sales= sum(JP_Sales), Other_Sales= sum(Other_Sales)) %>%
  pivot_longer(NA_Sales:Other_Sales, values_to = "value", names_to = "sales") %>%
  ggplot() +
  geom_col(aes(x = Year_of_Release, y = number/20), fill="#8C8CE3", alpha = 0.5) +
  geom_line(aes(x = Year_of_Release, y =value, colour=sales, group=sales), size=1) +
  scale_y_continuous(n.breaks=15, labels= c(seq(0, 1500, 100)), name="Number",
                    sec.axis = sec_axis( ~rescale(., c(0, 350)), name = "Sales"))+
  scale_x_discrete(name = NULL, breaks= c(seq(1980, 2020, 5))) +
  labs(title = "Number and Sales of Video Games, 1980 to 2020") +
  theme_bw() +
  theme(legend.position = "top", axis.ticks = element_blank(),
        legend.title = element_blank(), panel.grid = element_blank())
```

## Number and Sales of Video Games, 1980 to 2020



```
Video_Games_Sales %>%
  select(NA_Sales:Other_Sales) %>%
  summarise(NA_Sales = sum(NA_Sales), EU_Sales= sum(EU_Sales),
            JP_Sales= sum(JP_Sales), Other_Sales= sum(Other_Sales)) %>%
  pivot_longer(NA_Sales:Other_Sales, values_to = "value", names_to = "Region") %>%
  mutate(proportion = percent(value/sum(value), suffix = "%"),
         label = paste0(proportion, "\n$ ", value)) %>%
  ggplot(aes(x="", y=value, fill=Region)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar("y", start = 0) +
  geom_text(aes(label = Region, x=1.7),
            position = position_stack(vjust = 0.5), size = 4) +
  geom_text(aes(label = label, x=1.1), color = "white",
            position = position_stack(vjust = 0.5), size = 3) +
  scale_fill_manual(values = c("#f8b230", "#293f4d", "#1ab6c2", "#8C8CE3")) +
  labs(title = "The distribution of Video Games Sales") +
  theme_void() +
  theme(legend.position = "none",
        plot.title = element_text(hjust = 0.5))
```

The distribution of Video Games Sales



从电子游戏整体来看，从1980年开始电子游戏的数量不断增加，2005年开始增速加快，在2008与2009年新发布的电子游戏数量达到顶峰，从2010年开始迅速下降，电子市场急剧衰退，2012年开始在550左右波动。

从不同地区来看，北美市场一直是电子游戏的主要消费市场，而后是欧洲，最后是日本和其他地区。1980到2020年间，北美地区的总销量约占全球销量的一半。在电子游戏市场发展迅速的2002到2011年，北美的销量甚至可以达到欧洲的两倍左右。比较游戏数量以及四大地区销量的图线，我们可以发现四大地区销量的起伏基本与游戏数量趋势相同。其中，北美和欧洲地区销量波动与游戏数量在2000到2008年高速增长期与2009年到2012年快速下降期的趋势完全相同。所以，我们可以发现，电子游戏数量直接影响了北美地区和欧洲地区的销量。

## 2000到2008年高速增长期

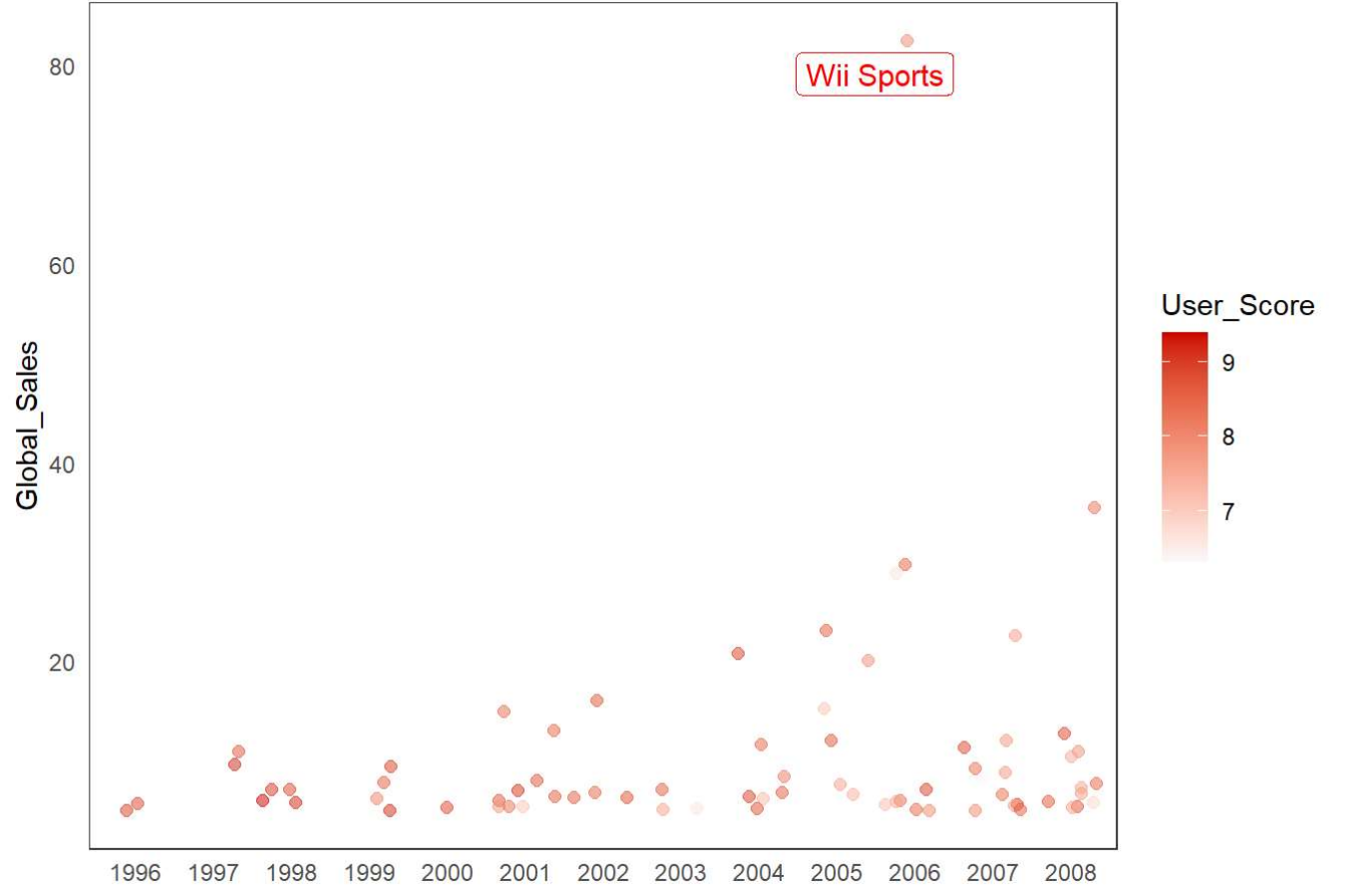
2000到2008年是电子游戏市场的高速增长期，其中新发布游戏数量从2000年的350个增长到2008年的1427个，增加了将近四倍。但是快速发展的原因是什么呢？我们主要考虑以下因素：

### 现象级的产品

往往现象级的产品可以很大程度上促进某个行业的发展，例如耳机行业的Airpods Pro，主动降噪功能的引入使得人们摆脱对以往头戴式无线蓝牙降噪耳机的认识，销量不断攀升；在一体式VR设备领域Oculus Quest2 压缩硬件设备成本，但体验感不打折扣，使得VR设备的入门成本降低成为一代热门产品，也带动了虚拟游戏的发展。所以，我们猜测高速增长契机可能是某些爆款游戏的示范效应。下面我们只需要考虑2008年以前发布的游戏。因为数据中存在发行年份缺失的数据，约占到总数据的1.6%，且我们关注其对应全球销量，发现稳定在较低水平的0~5左右，我们合理猜测是因为早年数据收集遗漏所致，且数量不多删去不会影响我们对销量较高游戏的分析。

```
set.seed(1)
Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure"
         & Year_of_Release < 2009 & User_Score >5.0 & Global_Sales> 5) %>%
  mutate(label1= ifelse(Global_Sales == max(Global_Sales), Name,"z"),
         label = ifelse(Year_of_Release == 2006 & label1 != "z", Name, NA_character_)) %>%
  ggplot(aes(x = Year_of_Release,y = Global_Sales,color=User_Score)) +
  geom_jitter(alpha = 0.5,size=2) +
  scale_colour_gradient(low = "#fafafa", high = "#CC0000",
                        breaks = seq(4, 10, 1)) +
  geom_label_repel(aes(label = label), size = 4, color="red") +
  labs(title = "Global Sales and User Scores of Popular Video Games,1980 to 2008",x=NULL) +
  theme_bw() +
  theme(axis.ticks = element_blank(),panel.grid = element_blank())
```

Global Sales and User Scores of Popular Video Games,1980 to 2008



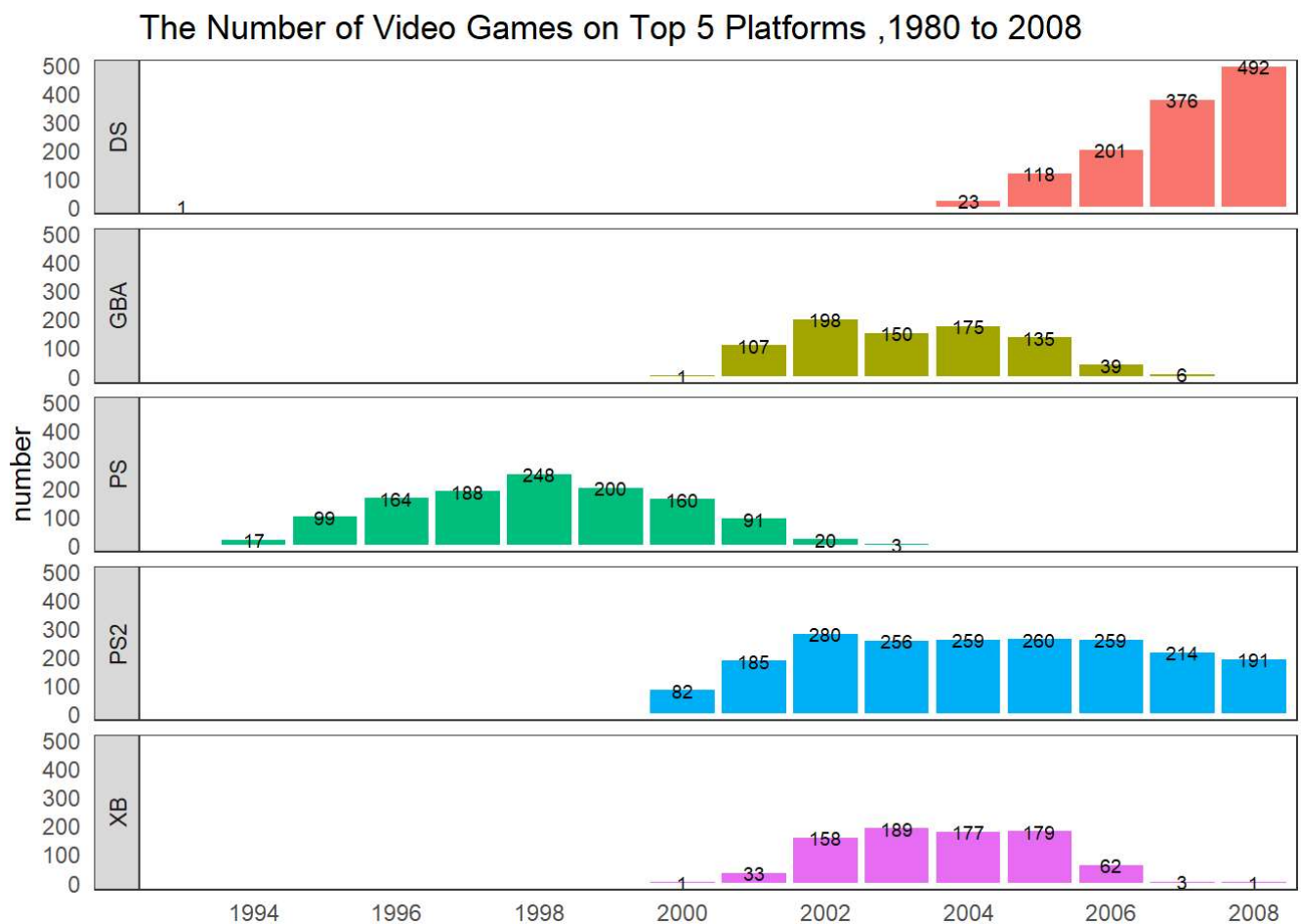
在去除缺失值之后，我们筛选出用户评分大于5，全球销量大于5个单位的游戏，并对其评分以及销量关于年份进行展示。首先，从整体来看，用户评分大于5，全球销量大于5个单位的游戏的数量随着年份在逐渐在增加。其次，从图一中我们可以发现，电子游戏市场增长速度最快的两个时间段分别为2002年以及2007年。而2001年的受欢迎游戏数量明显多于2000年，这可能使得2002年发布游戏数量的增加。2006年上线的爆款游戏Wii Sports，在全球销量超过80个单位，可能对于之后游戏市场的快速发展形成示范效应。

游戏平台的更新

PC和手机功能的不断发展挤压了传统电子游戏市场，也极大地促进了电子游戏的发展。例如，日本世嘉公司推出的游戏主机MD、SS，索尼推出的Play Station（PS），任天堂推出的DS等等。游戏主机的更新能很大程度上提升游戏体验，成为电子游戏市场发展的重要影响因素。

```
Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure"
         & Year_of_Release < 2009 & Platform %in% c("PS2", "DS", "PS", "GBA", "XB")) %>%
  select(Platform, Year_of_Release) %>%
  group_by(Platform, Year_of_Release) %>%
  summarise(number=n()) %>%
  ggplot(aes(x=Year_of_Release, y= number, fill=Platform)) +
  geom_bar(stat = "identity") +
  geom_text(aes(label = number), position = position_dodge(0.9), vjust = 0.5, size= 2.5) +
  scale_x_discrete(name = NULL, breaks= c(seq(1980, 2008, 2))) +
  labs(title = "The Number of Video Games on Top 5 Platforms ,1980 to 2008") +
  theme_bw() +
  theme(axis.ticks = element_blank(), panel.grid = element_blank(), legend.position = "none") +
  facet_wrap(~Platform, ncol = 1, strip.position = "left")
```

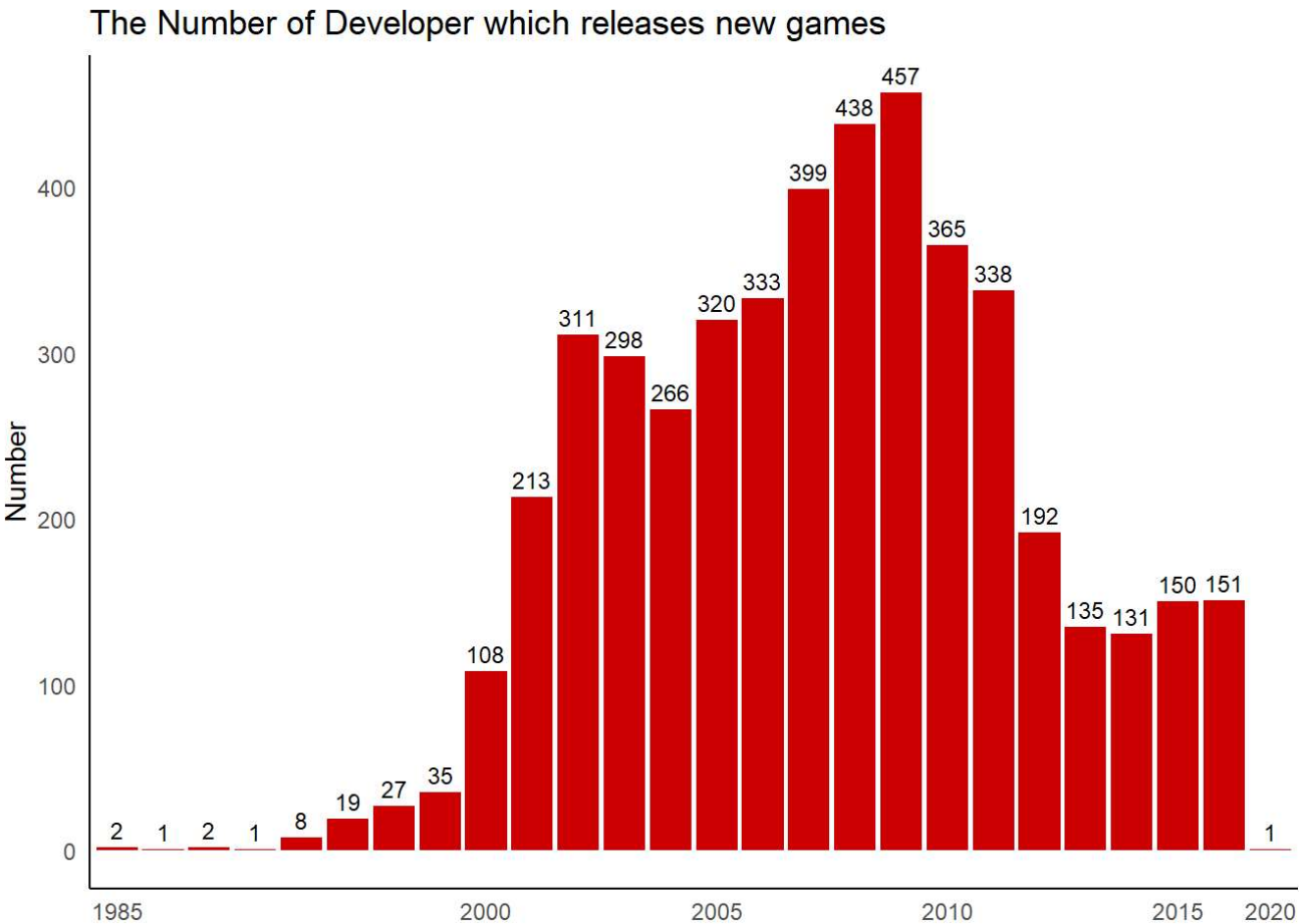
```
## `summarise()` has grouped output by 'Platform'. You can override using the
## `.groups` argument.
```



对于1980年到2008年不同平台上游戏的总数进行分析，前五名依次是PS2、DS、PS、GBA、XB，所以我们进一步分析这五种游戏主机上游戏数量的变化情况。首先，我们可以发现2000年新出现了三种游戏主机分别是PS2、GBA、XB，这可能也是2002年发布游戏数量的显著增加的原因。其次，我们可以发现PS2、GBA、XB三种主机发售后的第3达到新游戏数量顶峰，而后分别维持在175、260、180左右然后逐渐下降；而PS更是在2004年不再发布新游戏。而任天堂DS在发布三年后新游戏数量就超过200，超过部分公司的巅峰数量，而且增速还在增加，到2008年增加到将近500。所以我们有理由认为，DS游戏机能够提供比其他游戏主机更为良好的游戏体验，这也成为促进2006年后电子游戏市场快速发展的原因之一。

## 2009年到2012年快速下降

```
Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure"
    & Developer != "") %>%
  select(Developer, Year_of_Release) %>%
  separate(Developer, c("a","b","c","d","e","f"), sep = ",", ) %>%
  pivot_longer(a:f, names_to = NULL, values_to = "Developer",
    values_drop_na = TRUE) %>%
  count(Developer, Year_of_Release) %>%
  select(Year_of_Release) %>%
  count(Year_of_Release) %>%
  ggplot(aes(x= Year_of_Release, y=n)) +
  geom_bar(stat = "identity", fill = "#CC0000") +
  geom_text(aes(label = n), position = position_dodge(0.9), vjust = -0.5, size= 3) +
  scale_x_discrete(name = NULL, breaks= c(seq(1980, 2020, 5))) +
  labs(title = "The Number of Developer which releases new games", x = NULL, y = "Number") +
  theme_classic() +
  theme(axis.ticks = element_blank(), panel.grid = element_blank())
```



从图一可以发现，电子游戏市场在巅峰时期的2008到2009年并没有维持太久就发生了断崖式下降，从上图中可以看出参与开发的游戏开发商也在减少，且基本趋势与发布游戏数量变化基本相同。所以我们考虑外部因素，北美地区对于全球销量具有较大的影响作用，2012年美国金融危机，并席卷全球其中欧洲也受到严重影响，并且到2015年也没有恢复。



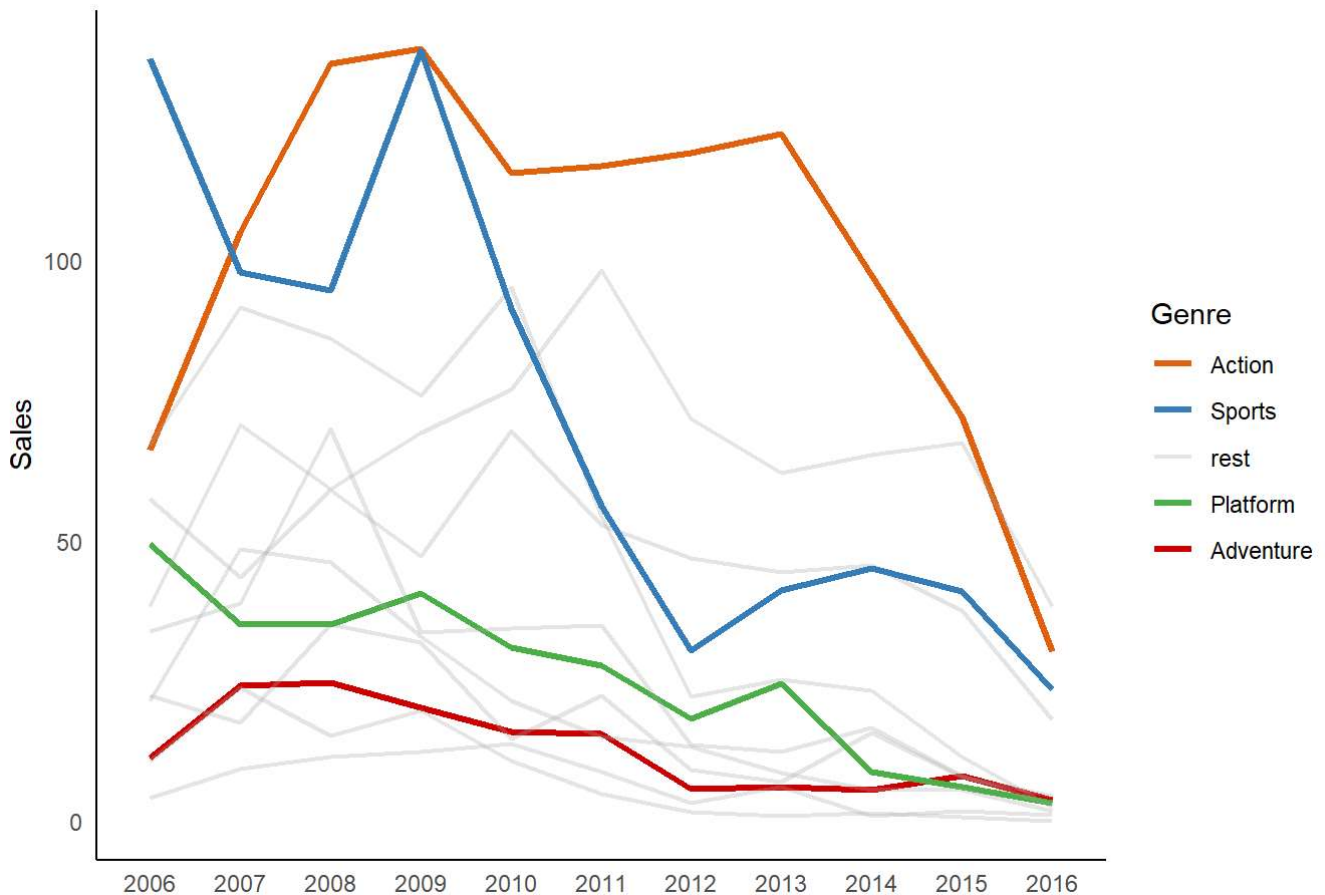
# 电子游戏市场发展趋势

从前面的分析我们可以知道，爆款游戏和游戏主机性能的提升都能促进电子游戏行业的发展，但是金融危机使得行业迅速衰退，在不考虑外部因素的情况下，能够在游戏内容方面做出什么样的改进来促进行业发展呢？

```
Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure"
         & Genre != "" & Year_of_Release > 2005 & Year_of_Release < 2017 ) %>%
  group_by(Year_of_Release, Genre) %>%
  summarise(Sales=sum(Global_Sales)) %>%
  mutate(mcv = ifelse(Genre %in% c("Action", "Adventure", "Sports",
                                   "Platform"), Genre, "rest")) %>%
  ggplot(aes(x = Year_of_Release, y = Sales, color = reorder(mcv, -Sales), group = Genre,
              size = reorder(mcv, -Sales), alpha = reorder(mcv, -Sales))) +
  geom_line() +
  scale_color_manual(values = c("#E0610E", "#377EB8", "dark grey",
                                "#4DAF4A", "#CC0000"), name = "Genre") +
  scale_size_manual(values = c(1.2, 1.2, 0.75, 1.2, 1.2), name = "Genre") +
  scale_alpha_manual(values = c(1, 1, 0.3, 1, 1), name = "Genre") +
  labs(title = "Sales of Different Game Genres", x = NULL, y = "Sales") +
  theme_classic() +
  theme(axis.ticks = element_blank(), panel.grid = element_blank())
```

```
## `summarise()` has grouped output by 'Year_of_Release'. You can override using
## the `.groups` argument.
```

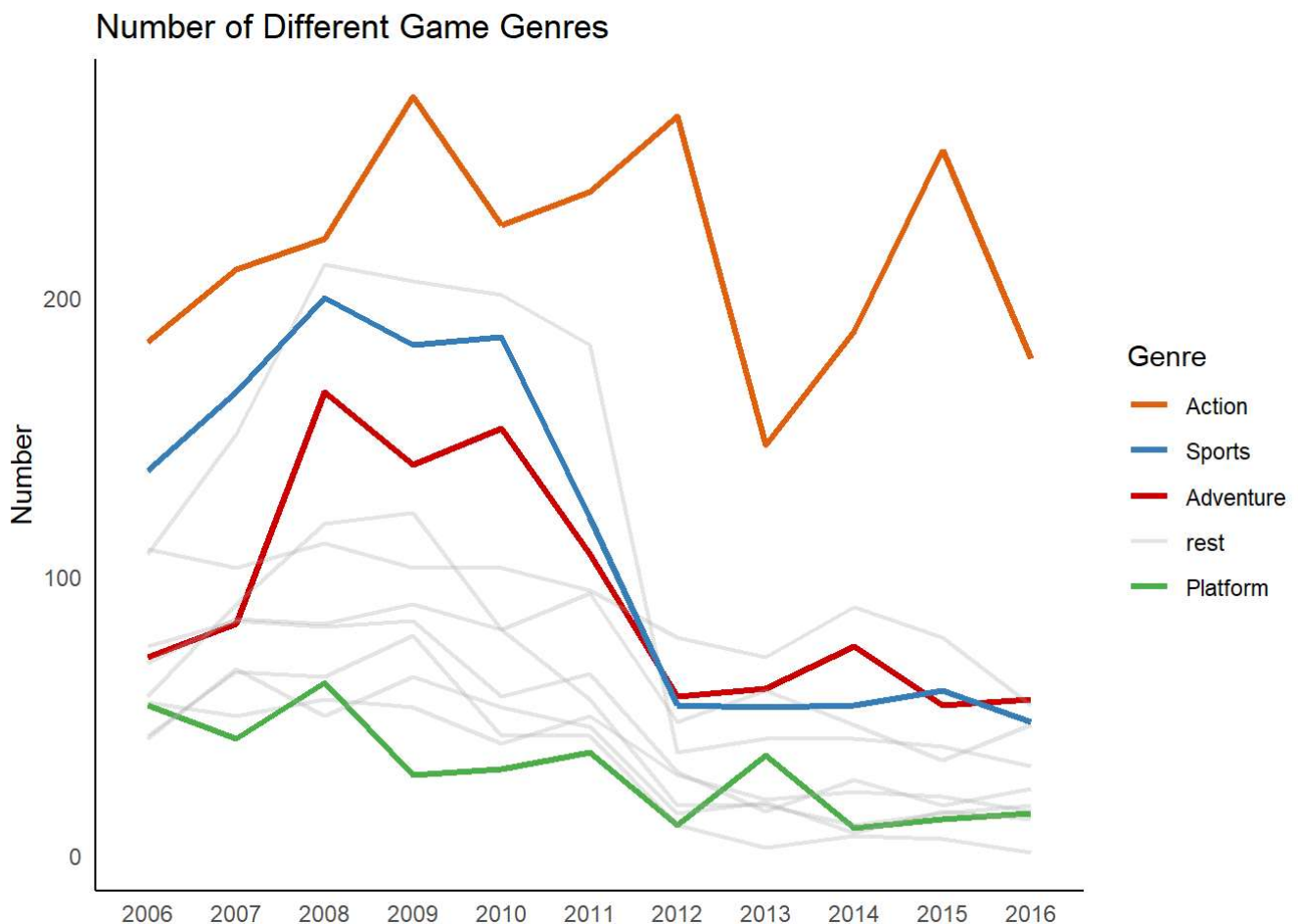
Sales of Different Game Genres



```

Video_Games_Sales %>%
  filter(Year_of_Release != "N/A" & Year_of_Release != "Adventure"
         & Genre != "" & Year_of_Release > 2005 & Year_of_Release < 2017) %>%
  group_by(Year_of_Release, Genre) %>%
  add_count(Genre) %>%
  mutate(mcv = ifelse(Genre %in% c("Action", "Adventure", "Sports",
                                   "Platform"), Genre, "rest")) %>%
  ggplot(aes(x = Year_of_Release, y = n, color = reorder(mcv, -n), group = Genre,
               size = reorder(mcv, -n), alpha = reorder(mcv, -n))) +
  geom_line() +
  scale_color_manual(values = c("#E0610E", "#377EB8", "#CC0000", "dark grey", "#4DAF4A"),
                    name = "Genre") +
  scale_size_manual(values = c(1.2, 1.2, 1.2, 0.75, 1.2), name = "Genre") +
  scale_alpha_manual(values = c(1, 1, 1, 0.3, 1), name = "Genre") +
  labs(title = "Number of Different Game Genres", x = NULL, y = "Number") +
  theme_classic() +
  theme(axis.ticks = element_blank(), panel.grid = element_blank())

```



两幅图第一幅描绘了不同游戏题材全球销量变化，第二幅描绘了不同游戏题材新游戏数量的变化。对比两幅图，我们可以发现，动作题材（Action）无论是数量还是销量都位居第一，但是在2014~2016年期间新游戏数量明显增加，但销量显著下降，今后应当减少相关题材。其次，冒险（Adventure）题材游戏数量较多，但销量低于大部分题材，说明此题材盈利能力低，应当相应减少。第三，平台闯关类游戏（Platform）2014年后游戏数量增加但销量减少，说明盈利能力下降，应当相应减少此类题材。最后，体育类游戏（Sports）2012年到2015年游戏数量保持稳定，但是盈利能力有所增加，所以应该增加体育类游戏。