Analise_ED.Caso_estudo_vendas_jogos

Adilson

07/04/2020

```r
#Analisamos as variáveis de vendas, primeiro observamos qual a proporção que as vendas representam em r
vendas<- dts%>% select("NA_Sales","EU_Sales","JP_Sales","Other_Sales","Global_Sales")
apply(vendas, 2, sum)/sum(dts$Global_Sales)*100
```

```
##      NA_Sales     EU_Sales     JP_Sales  Other_Sales Global_Sales
##     49.245889    27.287107    14.472604     8.942945   100.000000
```

```r
#Diferença nas preferências de mercado no Japão em relação a outros países por data:

global<-dts%>%group_by(Year, Genre)%>%
  summarize(Revenue = sum(Global_Sales))%>%
  top_n(1)
```

```
## Selecting by Revenue
```

```r
  jp<-dts%>%
  group_by(Year, Genre) %>%
  summarize(Revenue = sum(JP_Sales)) %>%
  top_n(1)
```

```
## Selecting by Revenue
```

```r
  global
```

```
A tibble: 40 x 3
Groups: Year [40]
Year Genre   Revenue
<ord> <fct>   <dbl>
1 1980 Shooter   7.07
2 1981 Action   14.8
3 1982 Puzzle   10.0
4 1983 Platform   6.93
5 1984 Shooter   31.1
6 1985 Platform   43.2
7 1986 Action   13.7
8 1987 Fighting   5.42
9 1988 Platform   27.7
10 1989 Puzzle   37.8
... with 30 more rows
```

```
jp
```

```
## # A tibble: 56 x 3
## # Groups:   Year [40]
##     Year Genre    Revenue
##     <ord> <fct>     <dbl>
##  1 1980  Action        0
##  2 1980  Fighting      0
##  3 1980  Misc          0
##  4 1980  Shooter       0
##  5 1980  Sports        0
##  6 1981  Action        0
##  7 1981  Platform      0
##  8 1981  Puzzle        0
##  9 1981  Racing        0
## 10 1981  Shooter       0
## # ... with 46 more rows
```

```r
ggplot(data=global, aes(x=Year,y=Revenue,fill=Genre))+
  geom_bar(stat='identity')+
  ggtitle("Top Genre by Revenue each Year in global") +
  theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4),
        legend.position = "top")
```

Top Genre by Revenue each Year in global
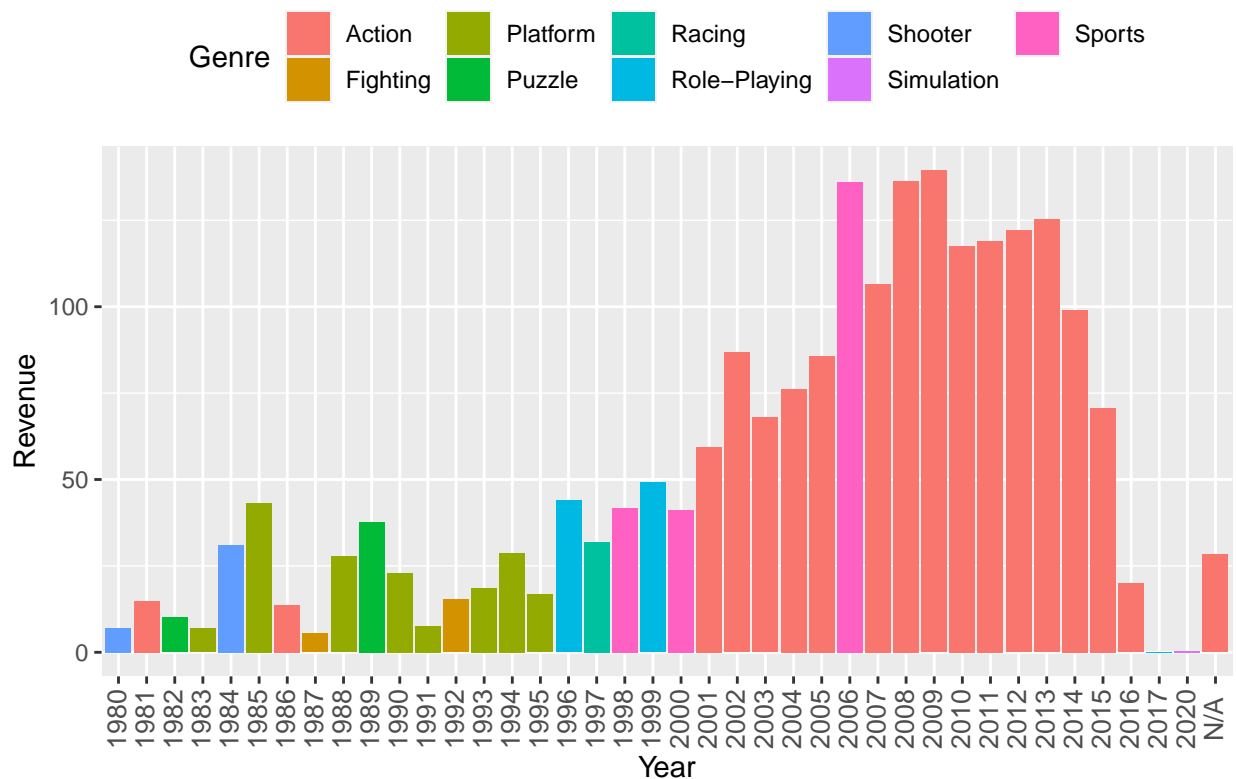
```
ggplot(data=jp, aes(x=Year,y=Revenue,fill=Genre))+
  geom_bar(stat='identity')+
  ggtitle("Top Genre by Revenue each Year in JP") +
  theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4),legend.position = "top")
```

Top Genre by Revenue each Year in JP
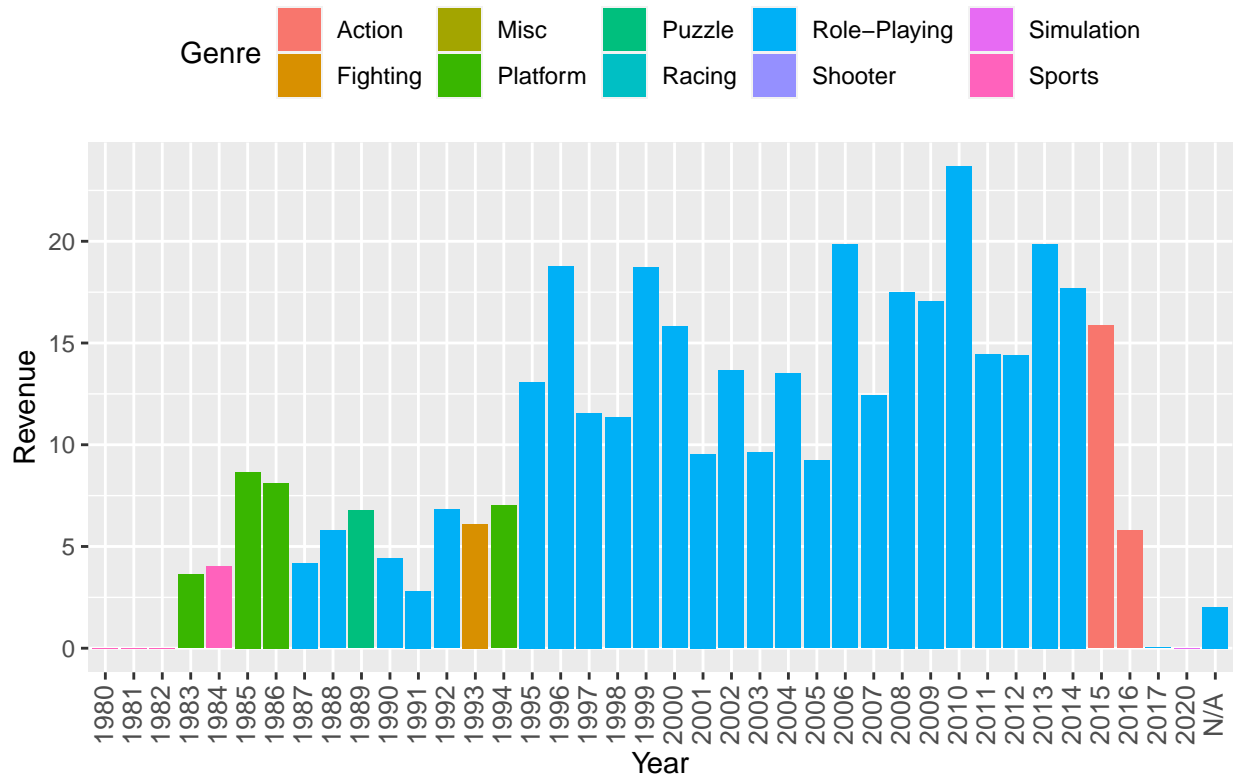
Observamos que a preferência no gênero de videogame no Japão é para role-playing games, enquanto glob
#Realizamos um teste do Chi Square de ajuste da qualidade do dinheiro nas vendas em cada gênero, com as

#Nós agrupamos desde 2008, que é a última mudança na série, e fazemos contrastes:

```
EU_gen <- dts%>%
  filter(Year>=2008)%>%
  group_by(Genre) %>%
  summarize(Revenue = sum(EU_Sales))

NA_gen <- dts%>%
  filter(Year>=2008)%>%
  group_by(Genre) %>%
  summarize(Revenue = sum(NA_Sales))

JP_gen <-dts%>%
  filter(Year>=2008)%>%
  group_by(Genre) %>%
  summarize(Revenue = sum(JP_Sales))
```
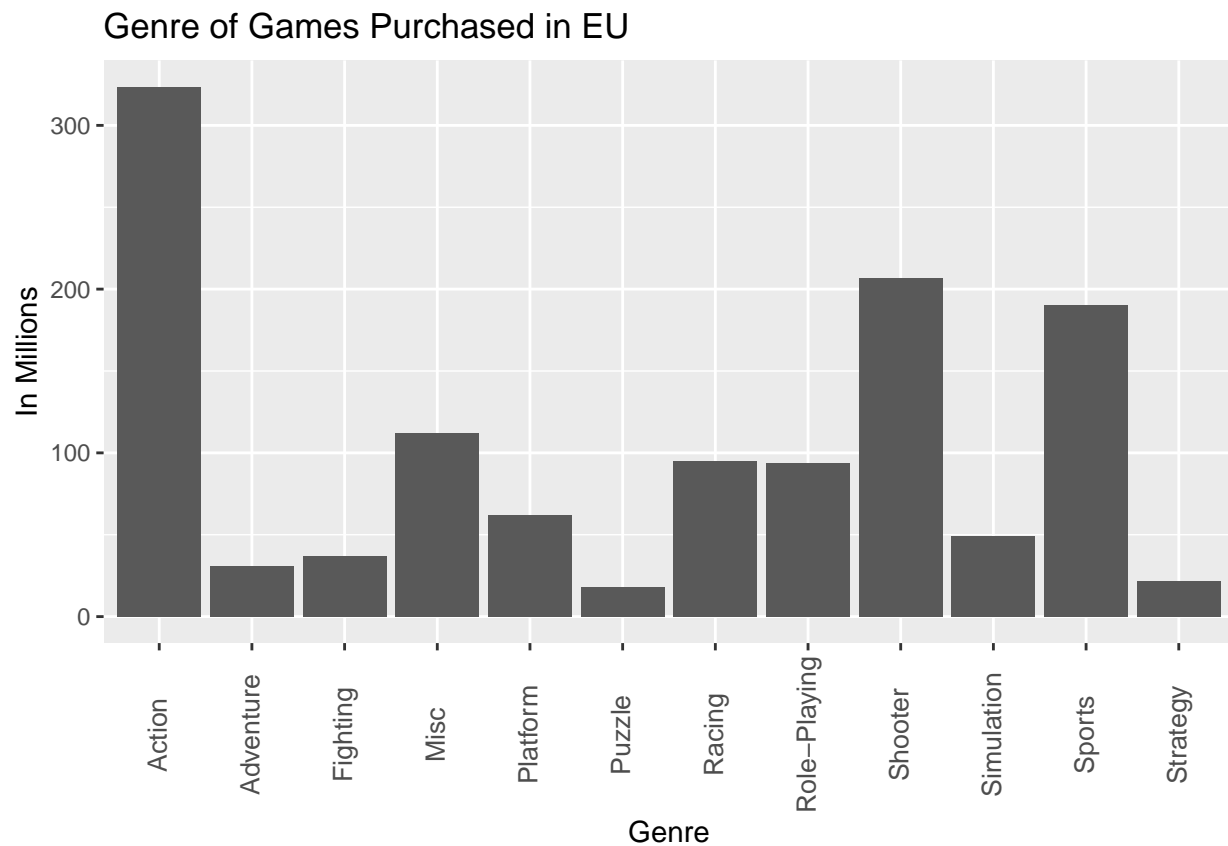
```
Global_gen <- dts%>%
  filter(Year>=2008)%>%
  group_by(Genre) %>%
  summarize(Revenue = sum(Global_Sales))

Other_gen <- dts%>%
  filter(Year>=2008)%>%
  group_by(Genre) %>%
  summarize(Revenue = sum(Other_Sales))

EU_gen %>%
   ggplot(aes(Genre,Revenue))+
   geom_bar(stat = 'identity')+
   ggtitle("Genre of Games Purchased in EU")+
   ylab("In Millions")+
   theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4))
```


Genre of Games Purchased in EU

```
JP_gen %>%
  ggplot(aes(Genre,Revenue))+
  geom_bar(stat = 'identity')+
  ggtitle("Genre of Games Purchased in Japan")+
  ylab("In Millions")+
  theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4))
```

Genre of Games Purchased in Japan
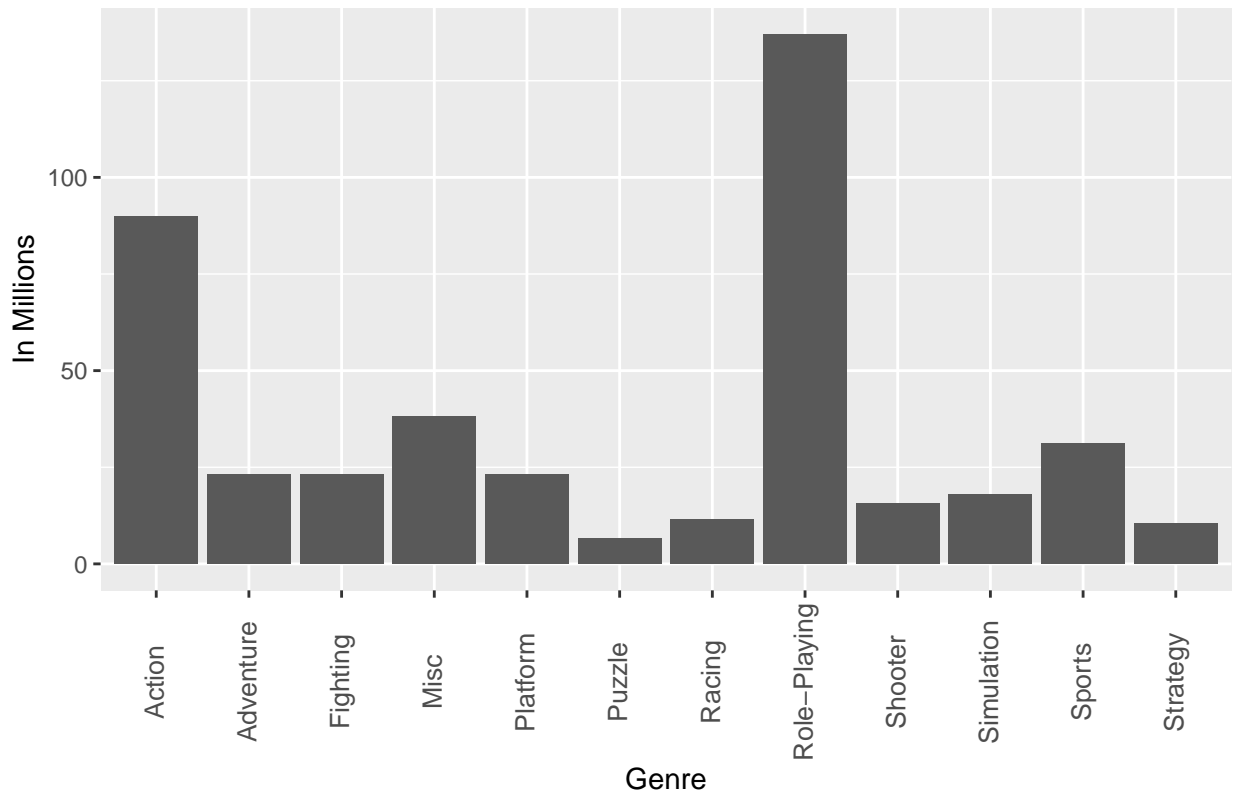

```r
#Evolución de las ventas por región

EU<- dts%>%
     group_by(Year)%>%
     summarise(Total_venda= sum(EU_Sales))

EU%>%
  ggplot(aes(Year, Total_venda))+
  geom_bar(stat = 'identity')+
  ggtitle('Total de Vendas no Estados Unidos')+
  ylab('Total de Vendas')+
  xlab('Ano')+
  theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4))
```

Total de Vendas no Estados Unidos
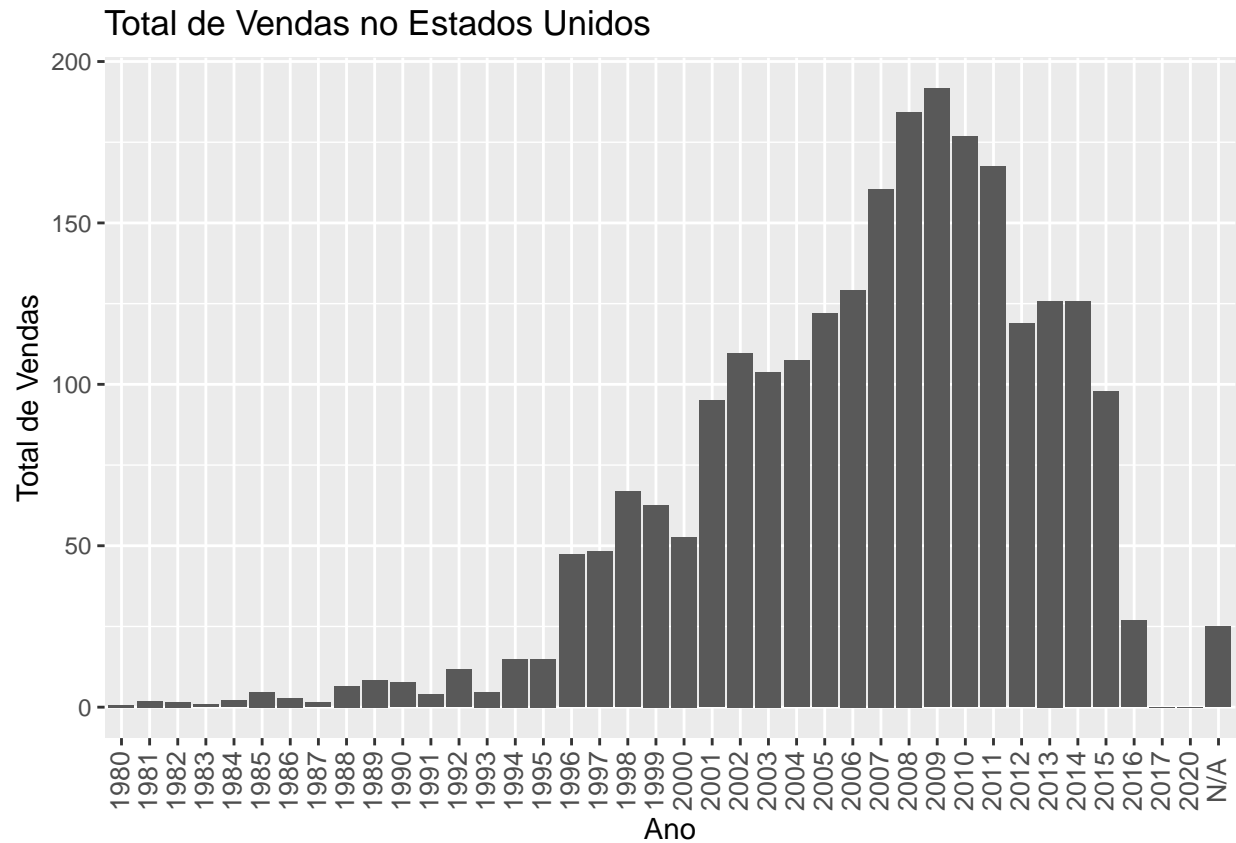

```r
NA_S<- dts%>%
  group_by(Year)%>%
  summarise(Total_venda= sum(NA_Sales))

NA_S%>%
  ggplot(aes(Year, Total_venda))+
  geom_bar(stat = 'identity')+
  ggtitle('Total de Vendas no NA')+
  ylab('Total de Vendas')+
  xlab('Ano')+
  theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4))
```
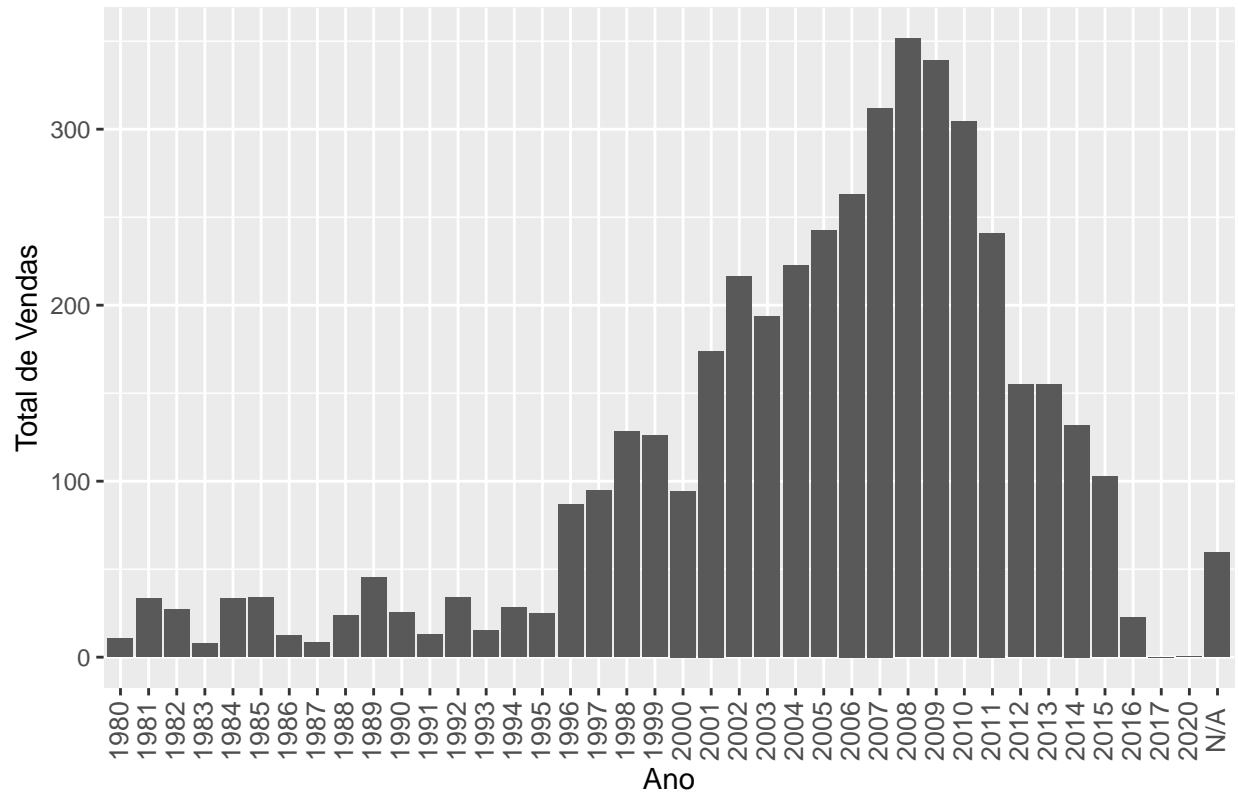
Total de Vendas no NA


```r
JP<- dts%>%
  group_by(Year)%>%
  summarise(Total_venda= sum(JP_Sales))

OTH<- dts%>%
  group_by(Year)%>%
  summarise(Total_venda= sum(Other_Sales))


data_sales<- data.frame(year=c(1980:2019),EU$Total_venda,JP$Total_venda,OTH$Total_venda,NA_S$Total_venda
data_sales
```

```
##    year EU.Total_venda JP.Total_venda OTH.Total_venda NA_S.Total_venda
## 1  1980           0.67           0.00            0.12            10.59
## 2  1981           1.96           0.00            0.32            33.40
## 3  1982           1.65           0.00            0.31            26.92
## 4  1983           0.80           8.10            0.14             7.76
## 5  1984           2.10          14.27            0.70            33.28
## 6  1985           4.74          14.56            0.92            33.73
## 7  1986           2.84          19.81            1.93            12.50
## 8  1987           1.41          11.63            0.20             8.46
## 9  1988           6.59          15.76            0.99            23.87
## 10 1989           8.44          18.36            1.50            45.15
## 11 1990           7.63          14.88            1.40            25.46
## 12 1991           3.95          14.78            0.74            12.76
```

```
## 13 1992        11.71         28.91          1.65          33.87
## 14 1993         4.65         25.33          0.89          15.12
## 15 1994        14.88         33.99          2.20          28.15
## 16 1995        14.90         45.75          2.64          24.82
## 17 1996        47.26         57.44          7.69          86.76
## 18 1997        48.32         48.87          9.13          94.75
## 19 1998        66.90         50.04         11.03         128.36
## 20 1999        62.67         52.34         10.05         126.06
## 21 2000        52.75         42.77         11.62          94.49
## 22 2001        94.89         39.86         22.76         173.98
## 23 2002       109.74         41.76         27.28         216.19
## 24 2003       103.81         34.20         26.01         193.59
## 25 2004       107.32         41.65         47.29         222.59
## 26 2005       121.94         54.28         40.58         242.61
## 27 2006       129.24         73.73         54.43         263.12
## 28 2007       160.50         60.29         77.60         312.05
## 29 2008       184.40         60.26         82.39         351.44
## 30 2009       191.59         61.89         74.77         338.85
## 31 2010       176.73         59.49         59.90         304.24
## 32 2011       167.44         53.04         54.39         241.06
## 33 2012       118.78         51.74         37.82         154.96
## 34 2013       125.80         47.59         39.82         154.77
## 35 2014       125.65         39.46         40.02         131.97
## 36 2015        97.71         33.72         30.01         102.82
## 37 2016        26.76         13.70          7.75          22.66
## 38 2017         0.00          0.05          0.00           0.00
## 39 2018         0.00          0.00          0.02           0.27
## 40 2019        25.01          6.72          8.74          59.52
```

```r
ggplot(data_sales, aes((year))) +
geom_line(aes(y = EU$Total_venda, colour = "Eu_sale"),size=1)+
geom_line(aes(y = JP$Total_venda , colour = "JP_Sales"),size=1)+
geom_line(aes(y = OTH$Total_venda , colour = "OTH_Sales"),size=1)+
geom_line(aes(y = NA_S$Total_venda , colour = "NA_Sales"),size=1)+
scale_x_continuous(breaks = scales::pretty_breaks(n = 20))+
ggtitle("Evolución de ventas a lo largo del tiempo") +
ylab('Total de ventanas')+
xlab('Ano')
```

Evolución de ventas a lo largo del tiempo
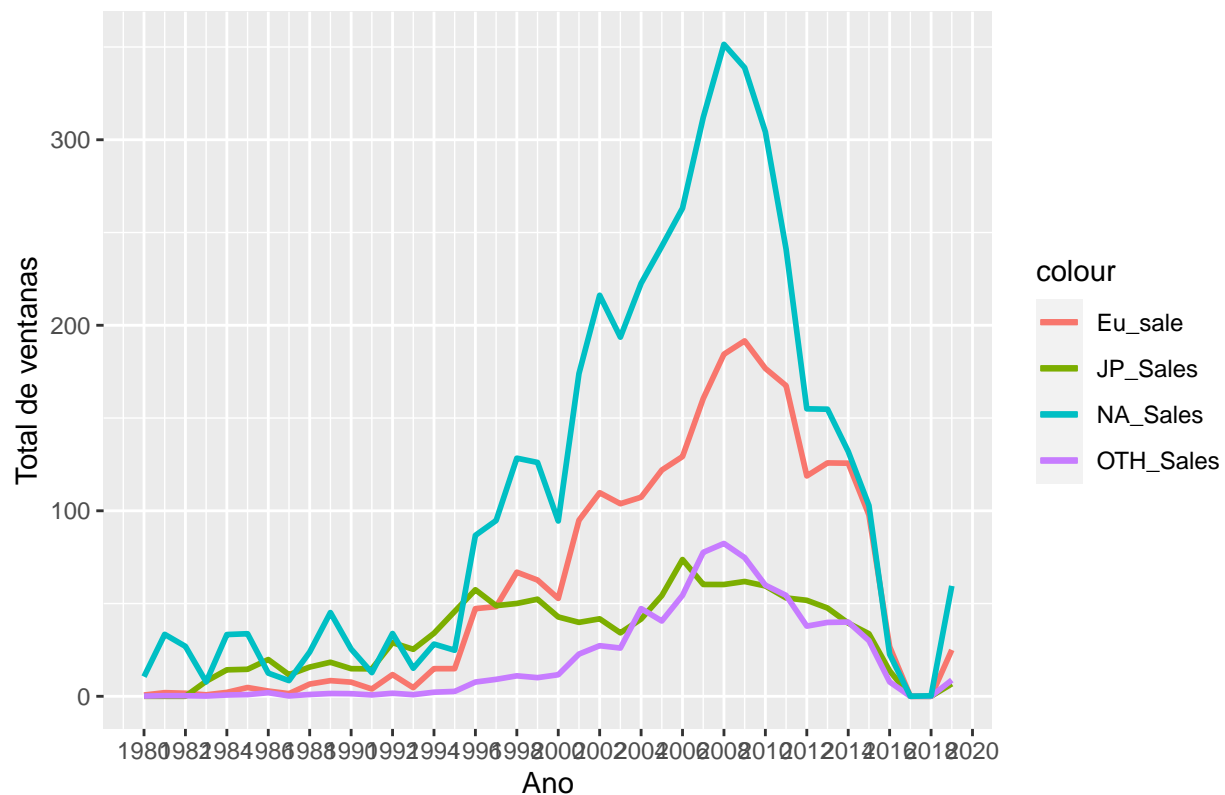

```r
theme(axis.text.x = element_text(angle = 90, size = 10, vjust = 0.4),
      legend.position = "top")
```

```
## List of 2
##  $ axis.text.x    :List of 11
##   ..$ family      : NULL
##   ..$ face        : NULL
##   ..$ colour      : NULL
##   ..$ size        : num 10
##   ..$ hjust       : NULL
##   ..$ vjust       : num 0.4
##   ..$ angle       : num 90
##   ..$ lineheight  : NULL
##   ..$ margin      : NULL
##   ..$ debug       : NULL
##   ..$ inherit.blank: logi FALSE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
##  $ legend.position: chr "top"
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```

```r
#Há um domínio nas vendas de NA entre 2000 e 2010. A partir de 2008 e coincidindo com a crise econômica
#Portanto, existem pontos críticos na evolução das vendas em 1995, 2000 e 2008 (pode-se perguntar por q
#Por sua vez, observa-se que as séries NA e UE estão fortemente ligadas.
```