



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO®



Tecnológico Nacional De México

Instituto Tecnológico De Tijuana

Subdirección Académica

Departamento de Sistemas y Computación

Semestre Enero - Junio 2022

Ingeniería Informática

Minería De Datos

Práctica 1

Unidad 2

Díaz Ruiz Uriel

No.18210839

Perez Ortega Victoria Valeria

No.18210718

JOSE CHRISTIAN ROMERO HERNANDEZ

Tijuana, B.C. a 26 de Abril de 2022.

1. Buscar una fuente de datos con formato csv(Tema libre)

```
#Abrir el CSV  
videogames <- read.csv(file.choose())
```

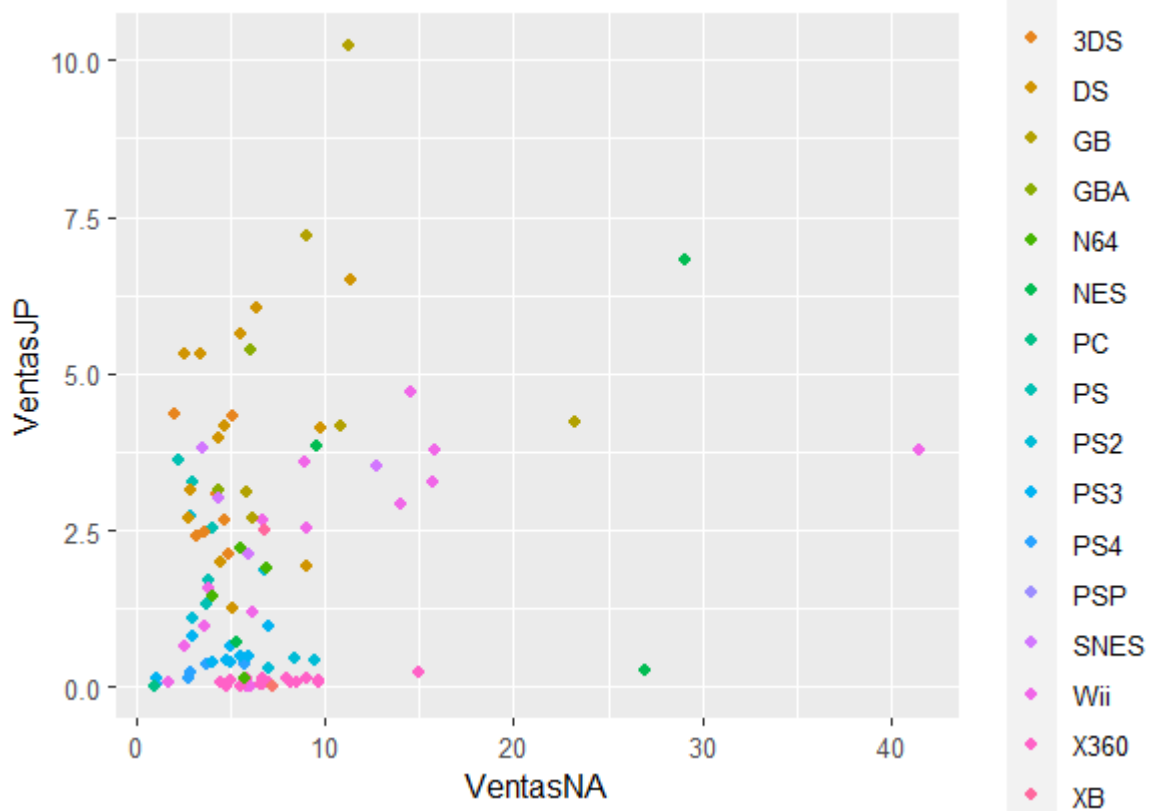
2. Leer el csv y analizar los datos con R

```
#Datos del CSV  
head(videogames)
```

3. Generar tres gráficas con R que cuente la historia de los datos, la primera que sea una gráfica de dispersión de puntos, la segunda que sea una gráfica en facetas y la tercera una gráfica que nos diga algo estadístico como la distribución que tienen los datos y que contenga la capa temas (theme).

#1. Gráfica de dispersión de Puntos

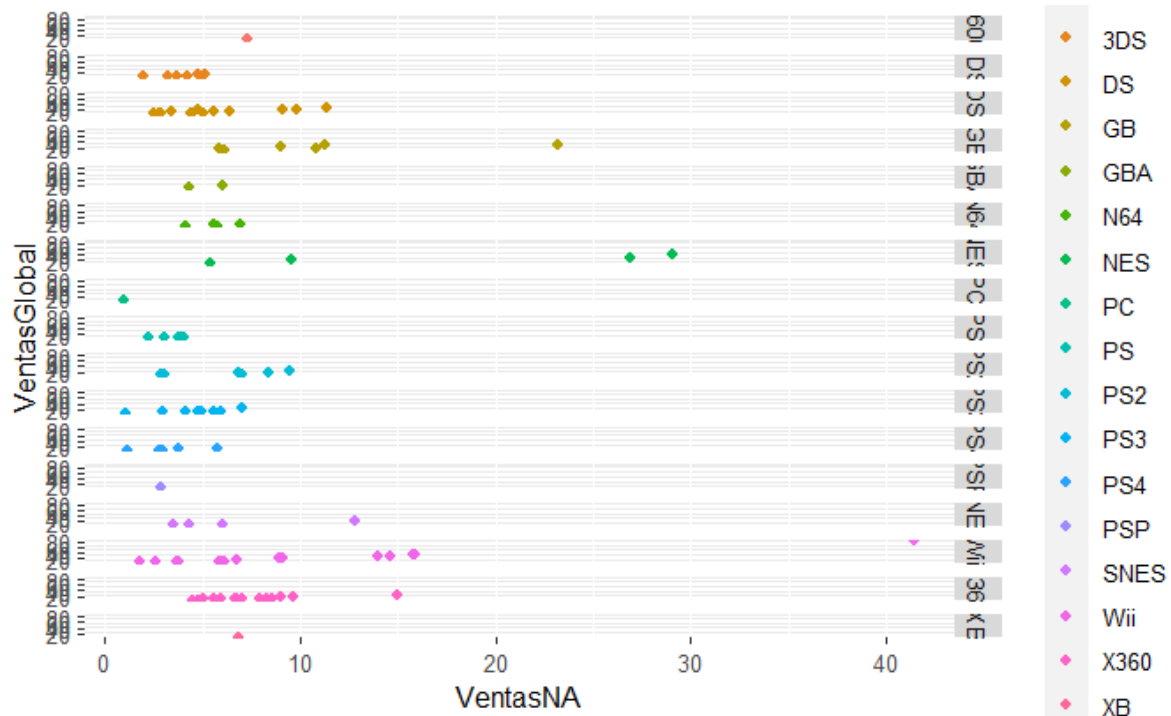
```
ggplot(videogames,aes(x=VentasNA, y=VentasJP,color=Plataforma)) + geom_point()
```





#2. Gráfica de facetas

```
x <- ggplot(videogames, aes(x=VentasNA, y=VentasGlobal,color=Plataforma))
```



#3. Gráfica de Distribución

```
a <- ggplot(videogames, aes(x=VentasNA))
```

```
b <- a + geom_histogram(binwidth = 10, aes(fill=Plataforma),color="Black")
```

```
b + xlab("VentasNA x Genero") + ylab("Number of VentasNA") +
```

```
ggtitle("VentasNA by team Distribution") +
```

```
theme(axis.title.x = element_text(color = "Black", size=10),axis.title.y = element_text(color = "Black", size=10),)
```

