



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO®

TECNOLÓGICO NACIONAL DE MÉXICO

INSTITUTO TECNOLÓGICO DE TIJUANA

SUBDIRECCIÓN ACADÉMICA

DEPARTAMENTO DE SISTEMAS Y COMPUTACIÓN

NOMBRE DE LOS ALUMNOS:

RAYMUNDO HIRALES LAZARENO (N. CONTROL: 17212339)

GALAVIZ LONA OSCAR EDUARDO (N.CONTROL: 17212993)

Carrera: Ingeniería Informática

Semestre: 9no

MATERIA: Minería de datos

PROFESOR: JOSE CHRISTIAN ROMERO HERNANDEZ

Practica Evaluatoria U2

FECHA: 7/11/21

Introducción

En el presente documento se expondrá a detalle la práctica evaluatoria de la unidad 2, donde veremos la utilización de las graficas asi como filtrar o buscar campos especificos o palabras en especifico que se requieren en el dataframe para la creacion de graficas con datos que deseemos, asi como la utilizacion de los archivos CSV para la creacion de nuevos dataframes que se requieran. En cuanto a las graficas se utilizan para crear una imagen o un panorama que represente los datos que estamos analizando para detectar posibles patrones dentro de los datos que nos puedan dar una idea de donde va el asunto.

Desarrollo

```
#Librerias Necearias
library(ggplot2)

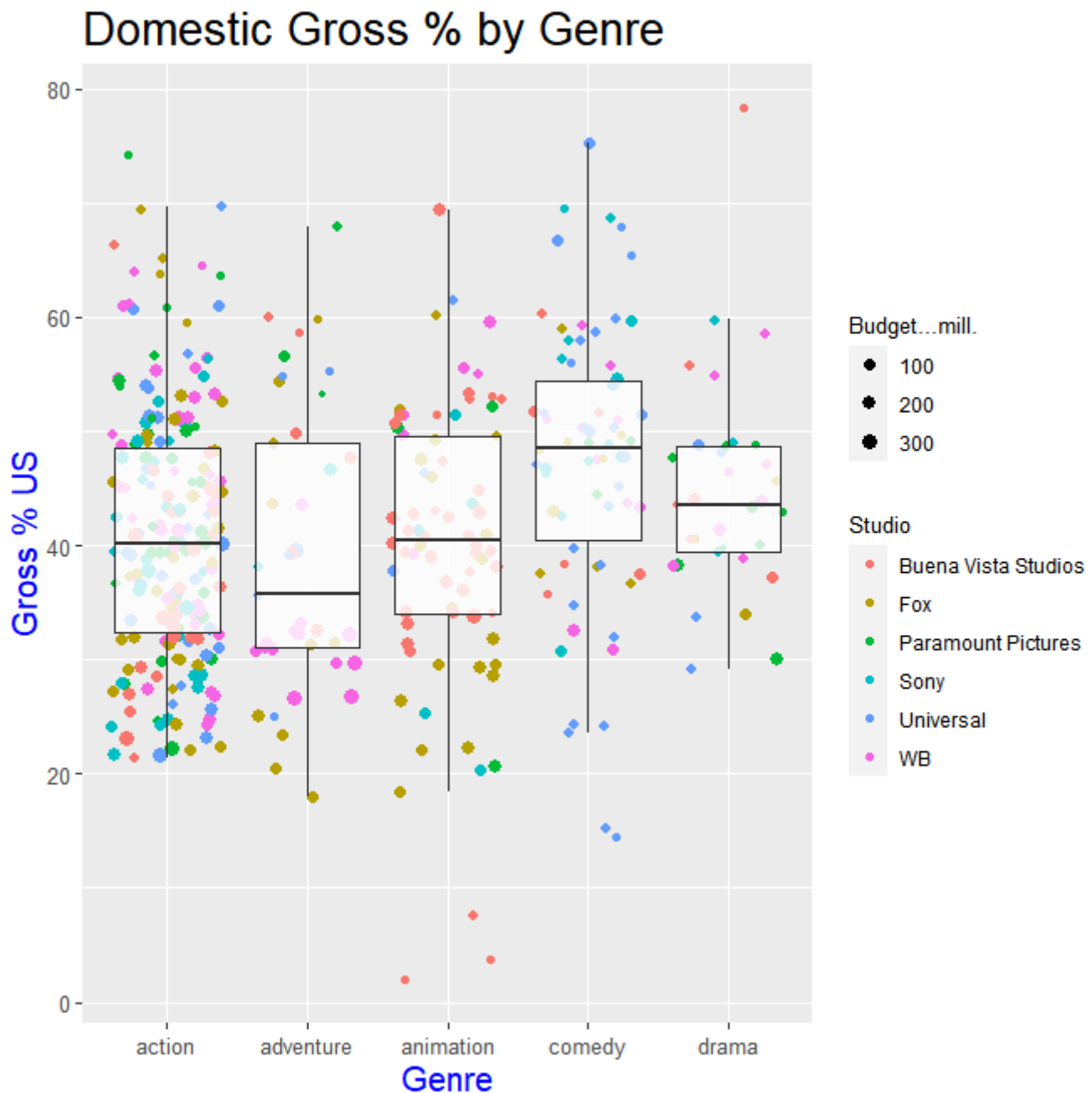
# Importar Dataset
ourdata <- read.csv(file.choose())

#Filtrar el genero del dataframe
#filtro1
filtro <- (ourdata$Genre == "action") | (ourdata$Genre == "adventure") |
(ourdata$Genre == "animation") | (ourdata$Genre == "comedy") |
(ourdata$Genre == "drama")
filtro

#Filtrar los estudios del dataframes
#filtro2
filtro2 <- ourdata$Studio %in% c("Buena Vista Studios", "WB", "Fox", "Universal",
"Sony", "Paramount Pictures")
filtro2

#Creacion del Dataframe Filtrado
ourdata2 <- ourdata[filtro & filtro2,]
ourdata2

# Creacion de la grafica
grafica <- ggplot(data=ourdata2,aes(x=Genre,
y=Gross...US))+geom_jitter(aes(size=Budget...mill.,colour=Studio))+
geom_boxplot(alpha = 0.8, outlier.color = NA)+ scale_size_continuous(range = c(1,
3))+xlab("Genre")+ ylab("Gross % US")+ ggtitle("Domestic Gross % by Genre")+
theme(axis.title.x = element_text(colour = "Blue", size=15), axis.title.y =
element_text(colour = "Blue", size=15), axis.text.x = element_text(size=10),
axis.text.y = element_text(size=10), plot.title = element_text(size=20),
legend.title = element_text(size=10), text = element_text(family = "sans") )
grafica
```



Conclusion

En pocas palabras la creacion de las graficas nos resulto util ya que podemos filtrar datos que ocupamos sin necesidad de tomar un archivo de datos completa y a partir de ahi generar graficas mas especificas y pequeñas ademas de que si se requiere podemos tomar diferentes datos para la creacion, asi como poder generar diferentes graficas que nos aporten un panorama diferente de los datos que estamos utilizando. y por ultimo la identificacion de patrones que son los mas importantes para saber en que situaciones x empresa esta teniendo una ganancia.