

Laboratorio I : Desarrollo grupal

1. Usando el archivo de datos `EjemploAlgoritmosRecomendacion.csv` realice lo siguiente:
 - a) Grafique usando los comandos `plot` y `qplot` en el plano XY las variables `Entrega` vs `Precio`.
 - b) Grafique usando comando `scatterplot3d` en 3 dimensiones las variables `Entrega`, `Precio` y `Durabilidad`.
 - c) Usando el comando `cor` calcule la matriz de correlaciones de la tabla `EjemploAlgoritmosRecomendacion.csv` y grafique esta matriz de 4 formas diferentes.
 - d) Usando el comando `Boxplot` encuentre los datos atípicos de la tabla de datos `EjemploAlgoritmosRecomendacion.csv`.
2. Hemos visto diferentes formas de definir vectores con R. Supongamos que queremos definir el vector $x = (1, 2, 3, 4, 5)$. Compruebe que las siguientes formas son equivalentes:

```
x<-c(1,2,3,4,5)
x<-1:5
x<-seq(1,5)
```

3. Defina el vector $y = (1, 3, 5, 7)$ utilizando la función `c()`. ¿Cómo lo harías con la función `seq()`?
4. En muchas ocasiones nos interesa hacer referencia a determinadas partes o componentes de un vector. Defina el vector $x = (2, -5, 4, 6, -2, 8)$, luego a partir de este vector defina instrucciones en R para generar los siguientes vectores:
 - $y = (2, 4, 6, 8)$, así definido y es el vector formado por las componentes positivas de x .
 - $z = (-5, -2)$, así definido z es el vector formado por las componentes negativas de x .
 - $v = (-5, 4, 6, -2, 8)$, así definido v es el vector x eliminada la primera componente.
 - $w = (2, 4, -2)$. así definido w es el vector x tomando las componentes impares.
5. ¿Qué ocurre cuando definimos una matriz en R y sólo especificamos el número de filas o el número de columnas? ¿Qué ocurre cuándo los datos no se corresponden con la dimensión de la matriz que queremos definir? Compruébelo ejecutando los siguientes comandos:

```
matrix(1:6,nrow=2)
matrix(1:6,nrow=4)
matrix(1:6,nrow=4,ncol=4)
```

6. Un grupo de amigos está formado por Ana de 23 años, Luis de 24 años, Pedro de 22, Juan de 24, Eva de 21 y Jorge de 22 años. Cree los vectores correspondientes a nombre, edad y sexo. (Use la codificación M=mujer, H=hombre). Convierta el vector `sexo` en un factor. ¿Cuáles son los niveles de dicho factor? Con los datos anteriores hemos creado el `dataframe` `amigos`, como se muestra seguidamente:

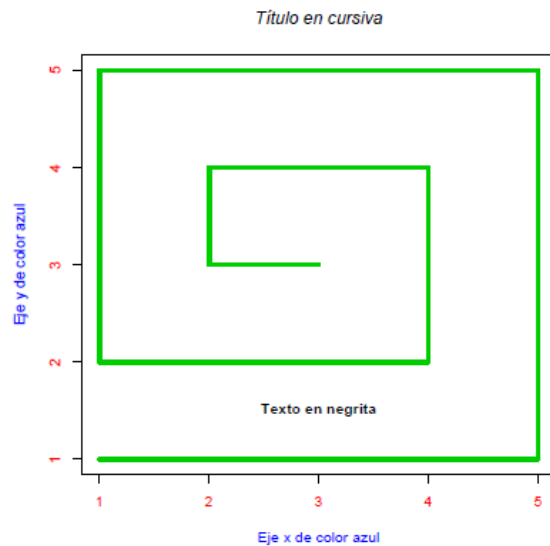
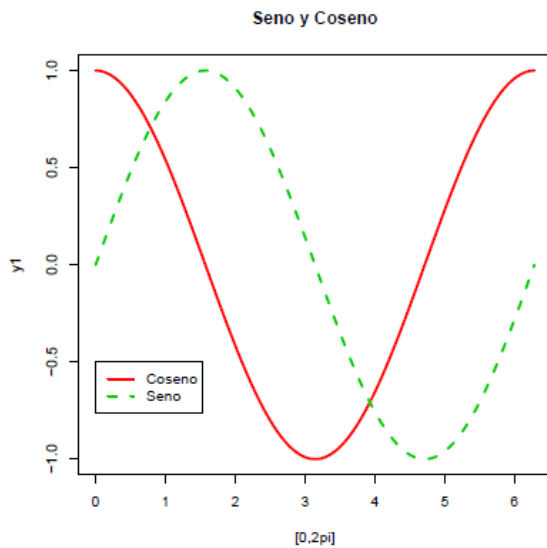
```
> amigos
  nombre edad sexo
1  Ana    23    M
2  Luis   24    H
3  Pedro  22    H
4  Juan   24    H
5  Eva    21    M
6  Jorge  22    H
```

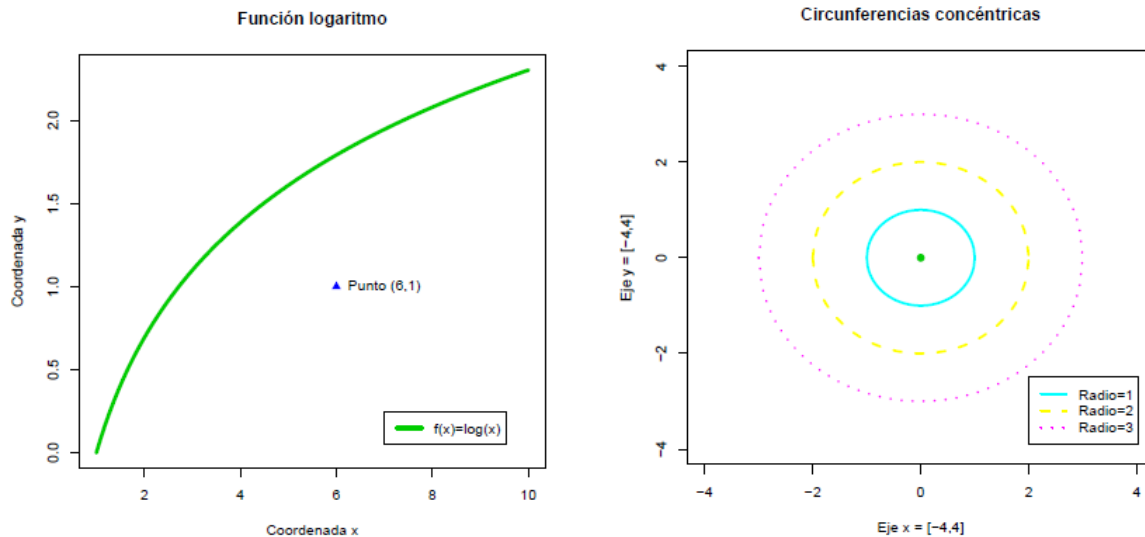
¿Cuál es el código de R da como resultado esta salida?

7. Queremos representar gráficamente la función coseno en el intervalo $[0, 2\pi]$. Para esto creamos el vector x de la siguiente forma `x<-seq(0,2*pi,length=100)`. ¿Cuál es la diferencia entre las gráficas obtenidas por comandos `plot`?

```
x<-seq(0,2*pi,length=100)
plot(cos(x))
plot(x,cos(x),col="red")
```

8. Usando los comandos `plot` y `par` genere en R los siguientes gráficos:





9. Ejecute y explique el siguiente código:

```
library(ggplot2)
library(reshape)
head(tips)
dim(tips)

qplot(tip,total_bill,geom="point", data=tips)
qplot(tip,total_bill,geom="point", data=tips, colour=smoker)
qplot(tip,total_bill,geom="point", data=tips, colour=smoker,
      xlab="Tip (in dollars)", ylab="Total Bill (in dollars)",
      main="Scatterplot of Tip by Total Bill,
      Colored by Smoking Status")

tips$rate <- tips$tip/tips$total_bill
qplot(rate,geom="histogram", data=tips)
qplot(rate,geom="histogram", data=tips, binwidth=.05)
ggplot(tips, aes(x=rate)) + geom_histogram(binwidth=.05,
      colour="black", fill="lightblue")

qplot(sex, rate, geom="boxplot", data=tips,fill = sex)
qplot(sex, rate, geom="boxplot", data=tips,
      xlab="Gender", ylab="Tipping Rate",
      main="Boxplots of tipping rate by gender",fill = sex) + geom_jitter()
```

10. Para tabla de Datos que viene en el archivo `DJTable.csv` el cual contiene los valores de las acciones de las principales empresas de Estados Unidos en el año 2010, usando el comando `plot` de R, grafique (en un mismo gráfico) las series de valores de las acciones de las empresas `CSCO` (Cisco), `IBM`, `INTC` (Intel) y `MSFT` (Microsoft).
11. Repita el ejercicio anterior usando funciones del paquete `ggplot2`.

12. En R para con tabla de datos que está en el archivo `SAheartv.csv` realice 3 gráficos interesantes con `ggplot2`, explique detalladamente como realizó cada uno de estos (cómo uso los parámetros de las funciones de `ggplot2`), además explique cómo se interpreta cada uno de los gráficos.

Entregables: Incluya en un solo archivo PDF todas las instrucciones, códigos R utilizados y los gráficos generados (cuando corresponda) en cada ejercicio.