



TECNOLÓGICO NACIONAL DE MEXICO
INSTITUTO TECNOLÓGICO DE TIJUANA

SUBDIRECCIÓN ACADÉMICA
DEPARTAMENTO DE INGENIERÍA EN SISTEMAS
COMPUTACIONALES

SEMESTRE FEBRERO-JUNIO 2022

MATERIA:

Minería de datos.

UNIDAD 2

Practica 1.

DOCENTE:

JOSE CHRISTIAN ROMERO HERNANDEZ

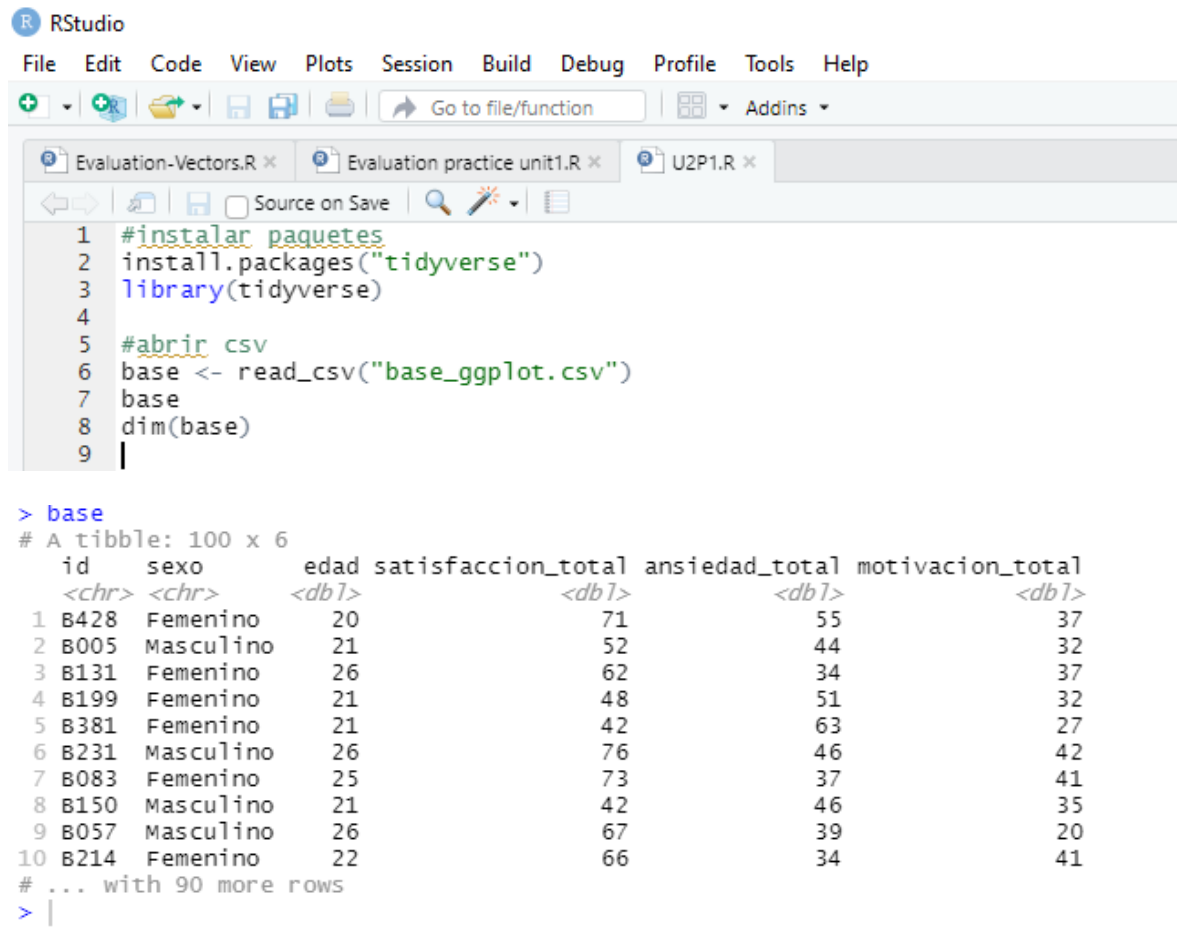
INTEGRANTES:

López Higuera Saúl Alfredo #18210493

Munguía silva Edgar Geovanny #17212344

Tijuana BC 26 de abril del 2022

1. Buscar una fuente de datos con formato csv(Tema libre)
2. Leer el csv y analizar los datos con R



```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ Go to file/function Addins
Evaluation-Vectors.R x Evaluation practice unit1.R x U2P1.R x
Source on Save
1 #instalar paquetes
2 install.packages("tidyverse")
3 library(tidyverse)
4
5 #abrir csv
6 base <- read_csv("base_ggplot.csv")
7 base
8 dim(base)
9 |

> base
# A tibble: 100 x 6
  id      sexo      edad satisfaccion_total ansiedad_total motivacion_total
  <chr>   <chr>    <dbl>          <dbl>          <dbl>          <dbl>
1 B428 Femenino    20             71             55             37
2 B005 Masculino    21             52             44             32
3 B131 Femenino    26             62             34             37
4 B199 Femenino    21             48             51             32
5 B381 Femenino    21             42             63             27
6 B231 Masculino    26             76             46             42
7 B083 Femenino    25             73             37             41
8 B150 Masculino    21             42             46             35
9 B057 Masculino    26             67             39             20
10 B214 Femenino    22             66             34             41
# ... with 90 more rows
> |

```

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).

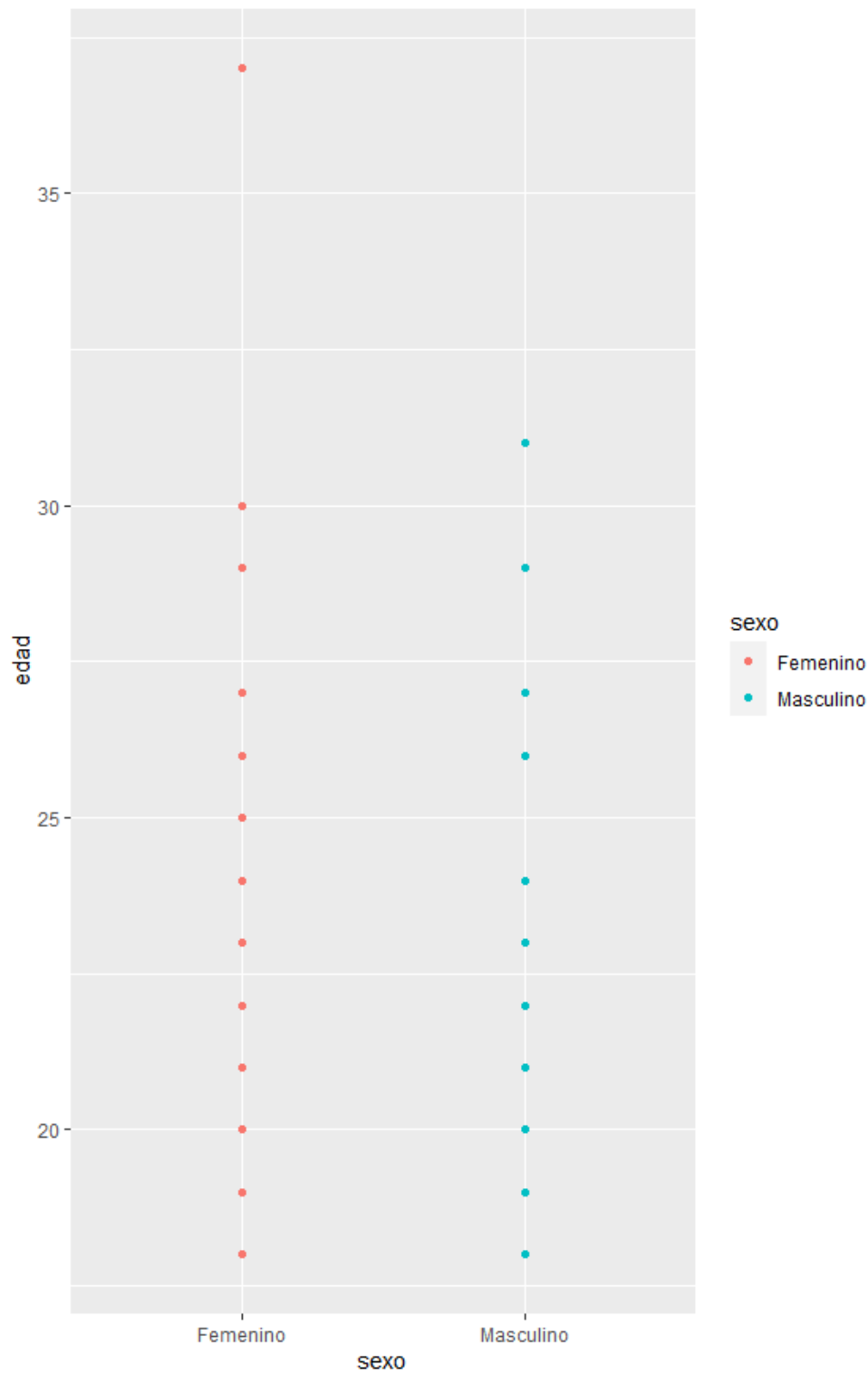
1. Grafica de dispersión de puntos.

```

#grafica de dispersion de puntos
grafica_dispersion1 <-ggplot(data = base)+
  aes(x = sexo, y = edad, color=sexo)+
  geom_point()

grafica_dispersion1
|

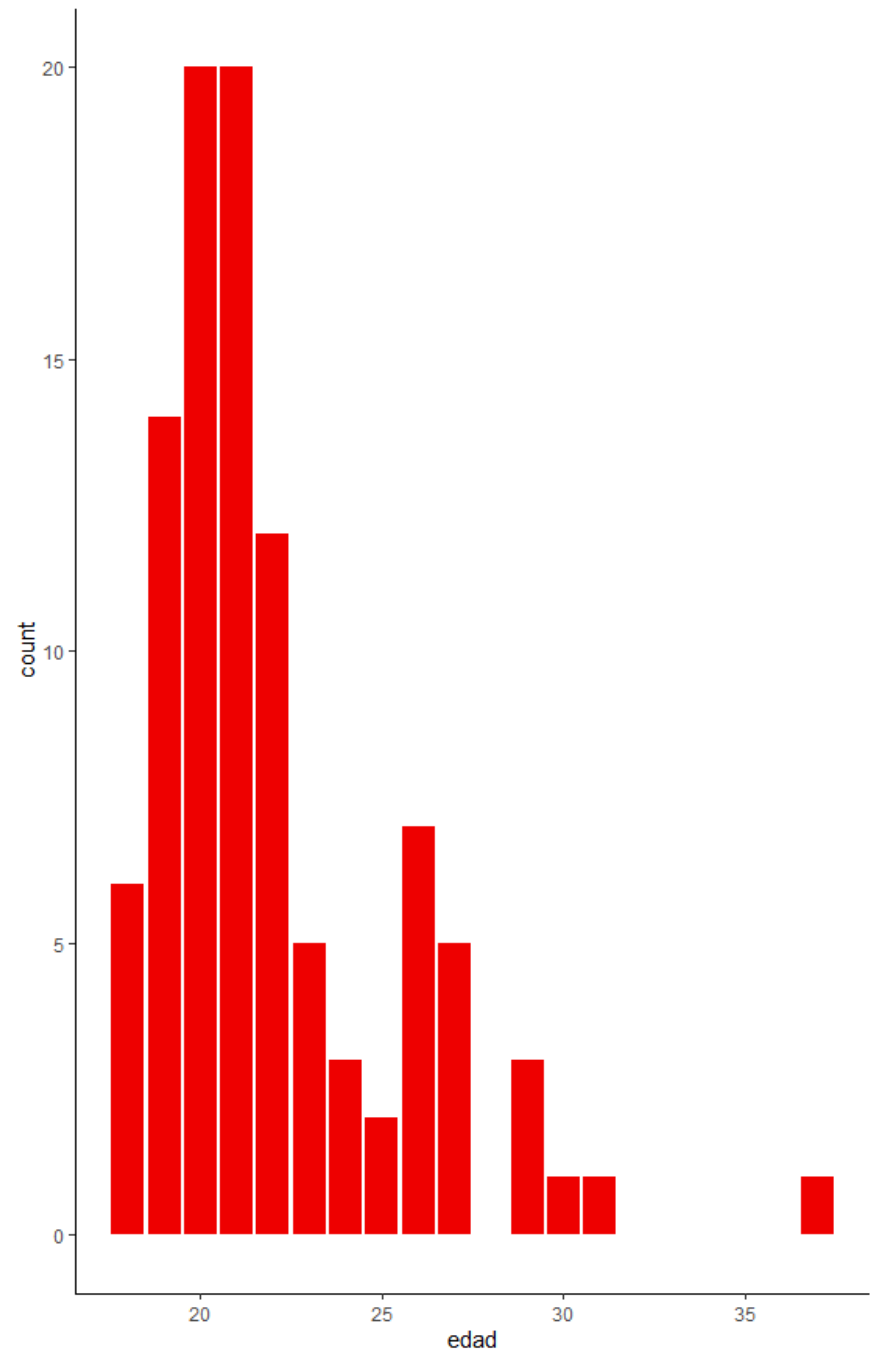
```



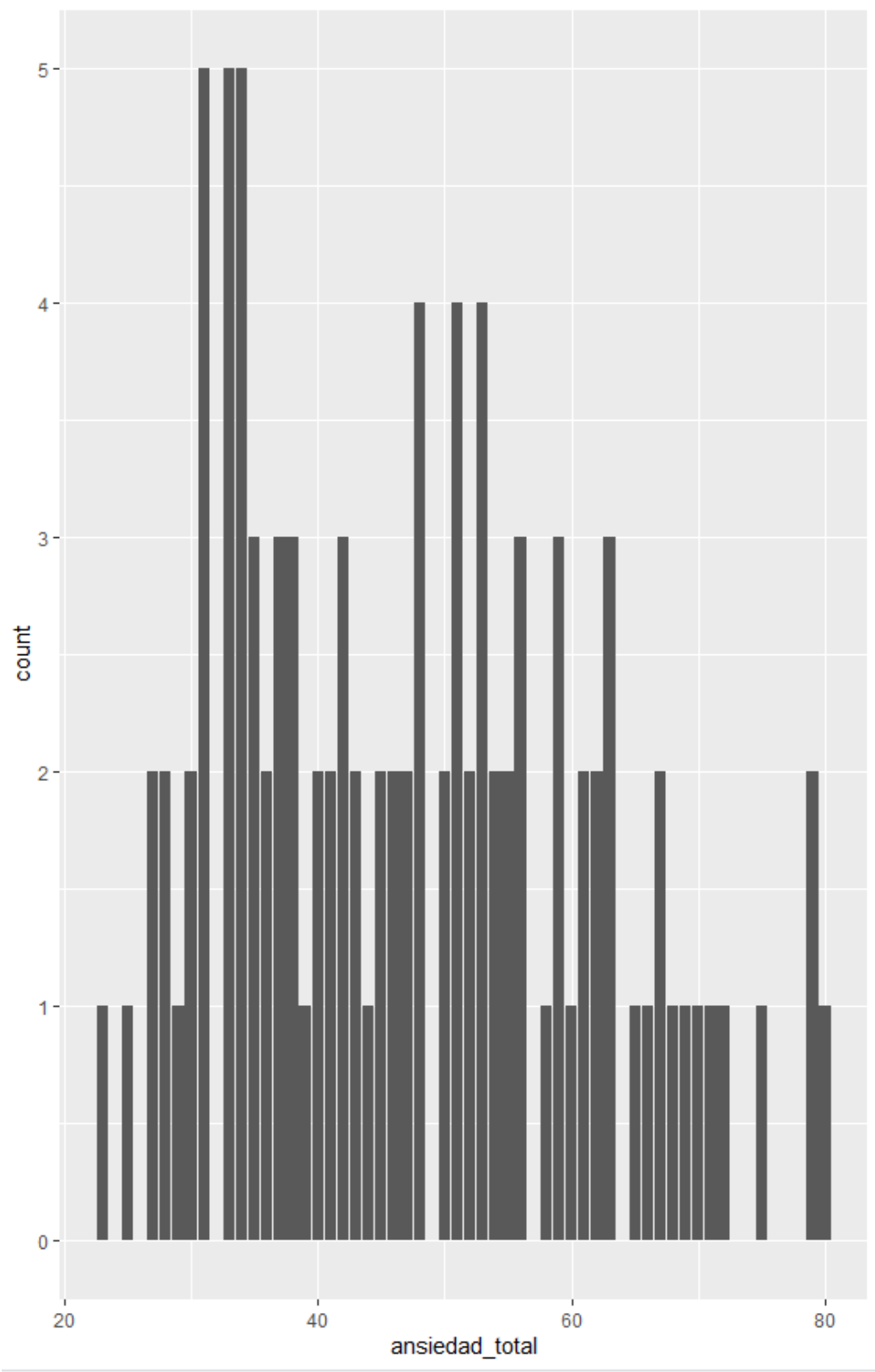
2. Grafica de facetas.

```
#Grafica de facetas
```

```
a <-ggplot(base,aes(x =edad))+ geom_bar(fill= "red2")  
a + theme_classic()
```



```
ggplot (base,aes(x =ansiedad_total ))+ geom_bar(aes(fill = edad), position= "dodge")
```



3. Grafica de distribución.

`#grafica de distribucion`

```
ggplot(base, aes(x = ansiedad_total, y = motivacion_total ))+ geom_point()+ geom_point(aes (color = sexo))+  
  theme_light()
```

