

Lab5_Daniela_Canabal.R

Perfil 1

2023-03-11

```
# Daniela Alexandra Canabal Valdes
# 10/03/2023
# Maestria Ciencias Forestales primer semestre

# HW_05
# Laboratorio 5: Correlacion

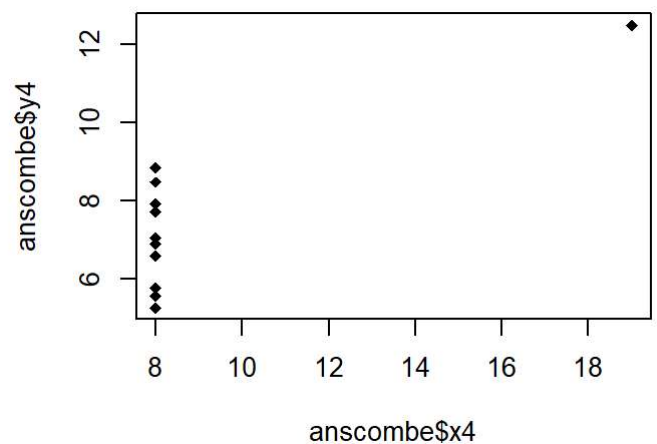
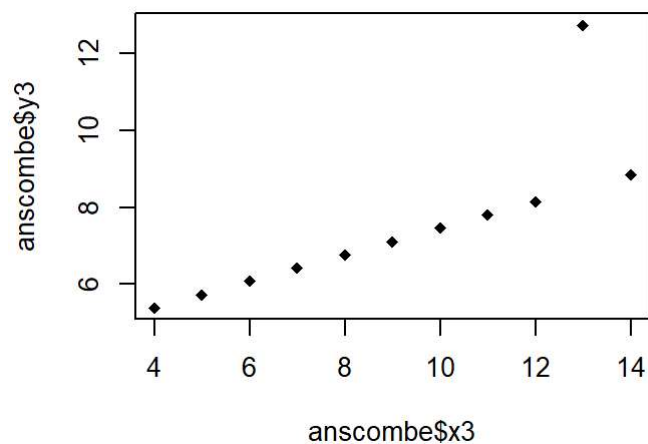
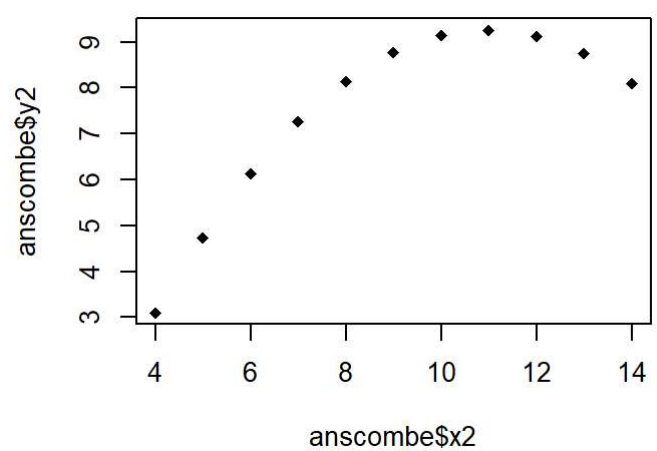
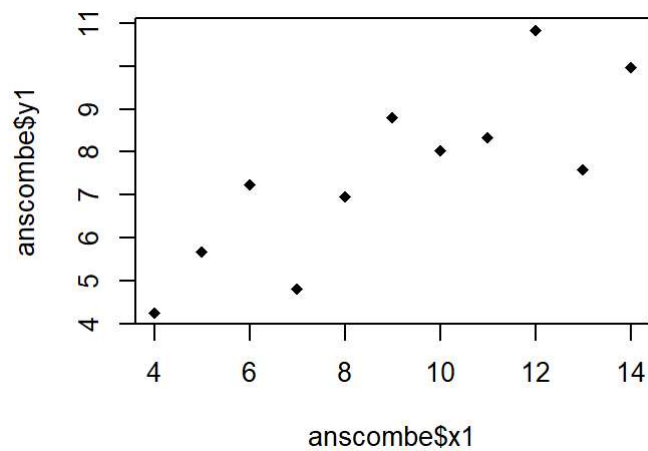
# Para cada ejercicio
# - Examinar la relación que existe entre dos muestras mediante una correlación,
# - Explore los datos gráficamente y explique,
# - Establezca la Hipótesis nula y la Hipótesis alternativa,
# - Aplique la prueba correspondiente,
# - Reporte los datos (indicar valor de r, grados de libertad y probabilidad, así como la signi-
ficancia de la correlación).

# Ejercicio 1. El cuarteto de Anscombe -----
anscombe
```

```
##      x1 x2 x3 x4      y1  y2  y3  y4
## 1   10 10 10  8   8.04 9.14  7.46 6.58
## 2    8  8  8  8   6.95 8.14  6.77 5.76
## 3   13 13 13  8   7.58 8.74 12.74 7.71
## 4    9  9  9  8   8.81 8.77  7.11 8.84
## 5   11 11 11  8   8.33 9.26  7.81 8.47
## 6   14 14 14  8   9.96 8.10  8.84 7.04
## 7    6  6  6  8   7.24 6.13  6.08 5.25
## 8    4  4  4 19   4.26 3.10  5.39 12.50
## 9   12 12 12  8  10.84 9.13  8.15 5.56
## 10   7  7  7  8   4.82 7.26  6.42 7.91
## 11   5  5  5  8   5.68 4.74  5.73 6.89
```

```
# Generar los gráficos de distribución de puntos para cada par de datos.
# Gratifica en un cuadro de 2x2
```

```
op = par(mfrow = c(2, 2), mar = c(4.5, 4, 1, 1))
plot(anscombe$x1, anscombe$y1, pch = 18)
plot(anscombe$x2, anscombe$y2, pch = 18)
plot(anscombe$x3, anscombe$y3, pch = 18)
plot(anscombe$x4, anscombe$y4, pch = 18)
```



```
par(op)
```

```
# Coeficiente de correlación -----
```

```
# Determinar para cada par de datos Los coeficientes de correlación r.
```

```
cor.test(anscombe$x1, anscombe$y1)
```

```
##
## Pearson's product-moment correlation
##
## data: anscombe$x1 and anscombe$y1
## t = 4.2415, df = 9, p-value = 0.00217
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.4243912 0.9506933
## sample estimates:
## cor
## 0.8164205
```

```
cor.test(anscombe$x2, anscombe$y2)
```

```
##
## Pearson's product-moment correlation
##
## data:  anscombe$x2 and anscombe$y2
## t = 4.2386, df = 9, p-value = 0.002179
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4239389 0.9506402
## sample estimates:
##          cor
## 0.8162365
```

```
cor.test(anscombe$x3, anscombe$y3)
```

```
##
## Pearson's product-moment correlation
##
## data:  anscombe$x3 and anscombe$y3
## t = 4.2394, df = 9, p-value = 0.002176
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4240623 0.9506547
## sample estimates:
##          cor
## 0.8162867
```

```
cor.test(anscombe$x4, anscombe$y4)
```

```
##
## Pearson's product-moment correlation
##
## data:  anscombe$x4 and anscombe$y4
## t = 4.243, df = 9, p-value = 0.002165
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4246394 0.9507224
## sample estimates:
##          cor
## 0.8165214
```

¿Alguna sorpresa? Como puedes ver, Los cuatro pares de Las variables xy tienen básicamente La misma correlación de 0.816. Pero no todos tienen diagramas de dispersión en Los que Los puntos s e agrupan alrededor de una Línea.

El mensaje para llevar a casa es que el coeficiente de correlación puede ser engañoso en presencia de valores atípicos o asociación no lineal. Debido a esto, siempre es importante revisar Los datos con un gráfico de dispersión.