



Gráficos de dispersión

Visualización Científica

19 de septiembre de 2018

Eduardo Castro, PhD

Center for Bioinformatics and Integrative Biology

www.cbib.cl

www.castrolab.org

Revisar tarea

- Instalar R y R Studio
- Realizar el tutorial: <http://www.datacarpentry.org/R-ecology-lesson/> de la lección 1 a la 6
- Averiguar lo que es un Tufte Handout
- Instalar el paquete de R “ggplot2”

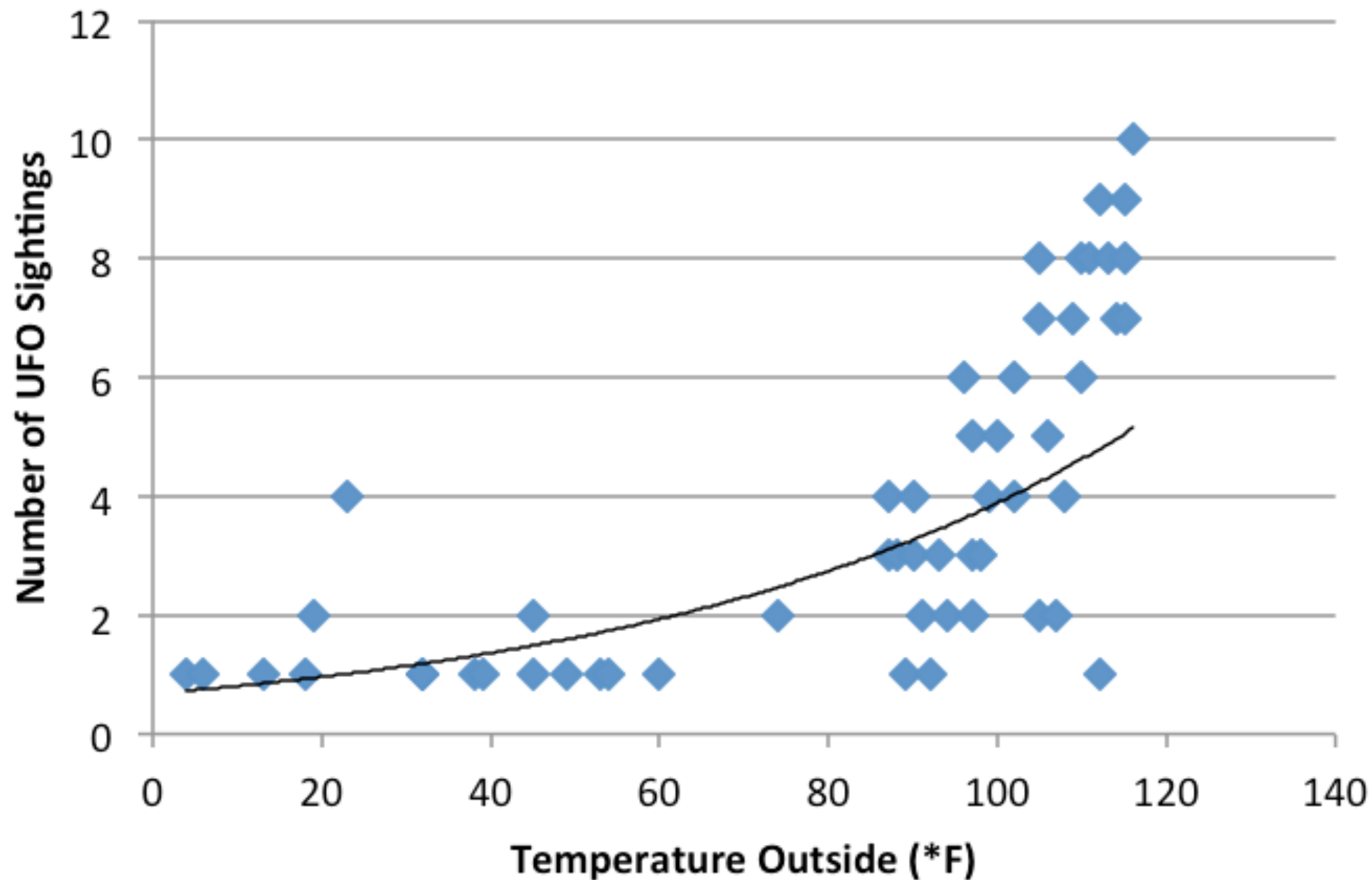
Diseñados para mostrar la relación entre dos variables

- Scatterplot en inglés
- Gráfico XY también
- ¿Cuál es la relación entre X e Y? ¿Cómo X influye en Y?

Variable dependiente e independiente

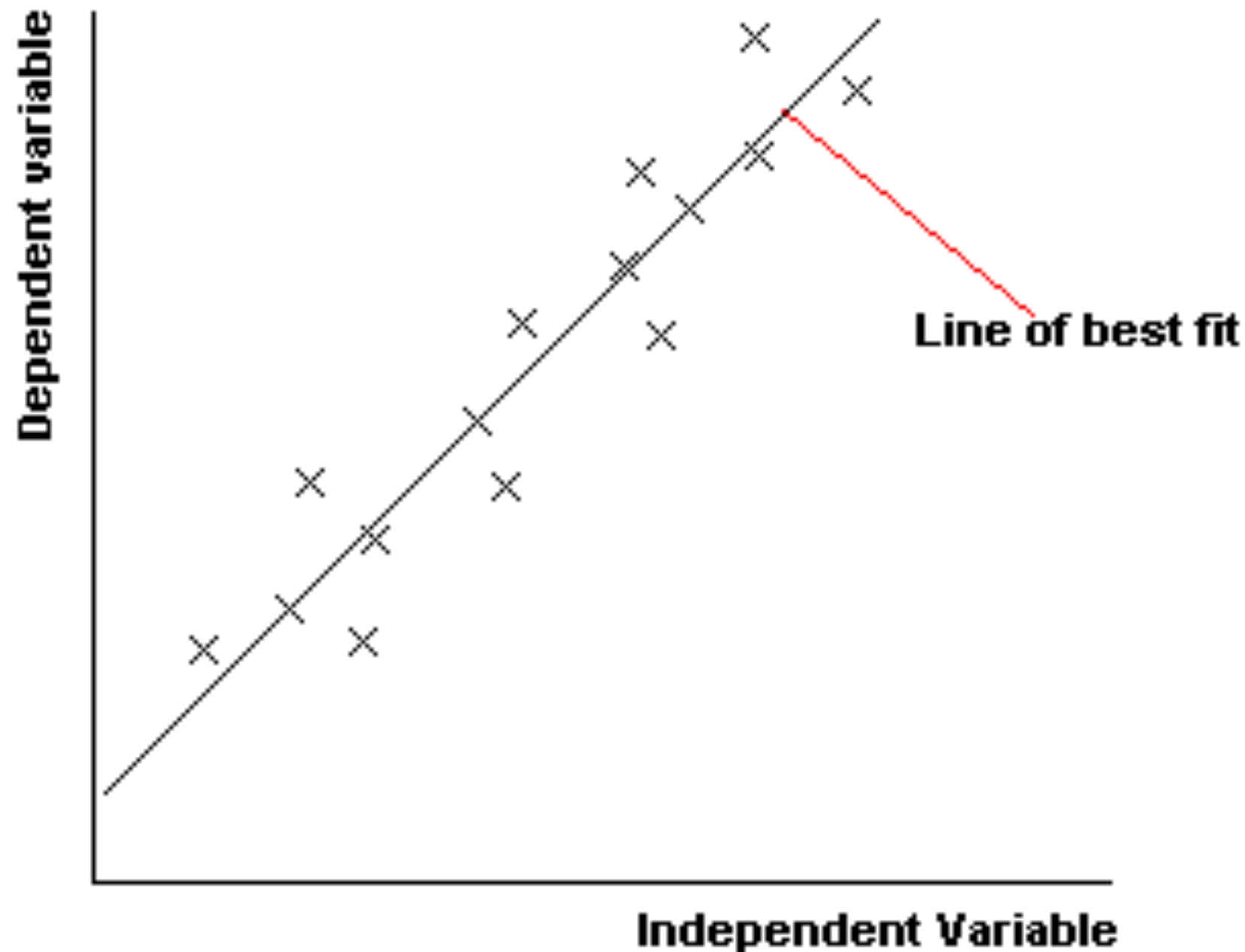
- V. dependiente = cambia o es controlada por el experimento —> en el eje Y
- V. independiente = no cambia con el experimento —> en el eje X. Ejemplos = tiempo, temperatura

Variable dependiente e independiente



Variable dependiente e independiente

- V.
e)
- V.
—



el

iento
ura

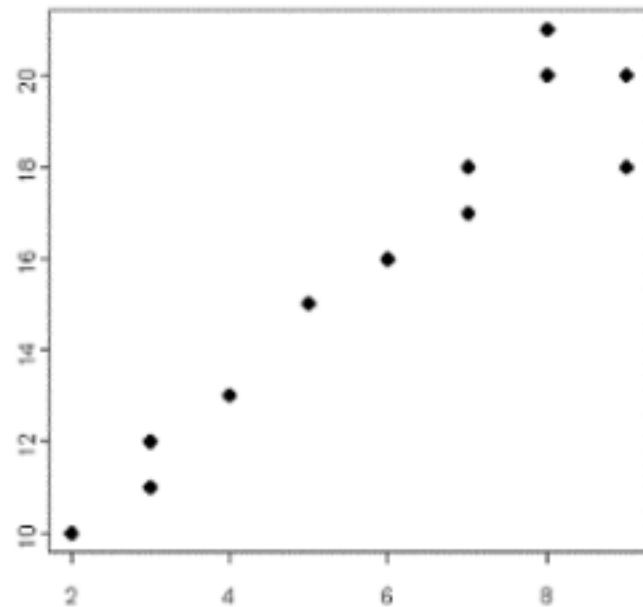
Variable de respuesta y explicatoria

- V. de respuesta es equivalente a la dependiente, en tanto la variable explicatoria es equivalente a la independiente

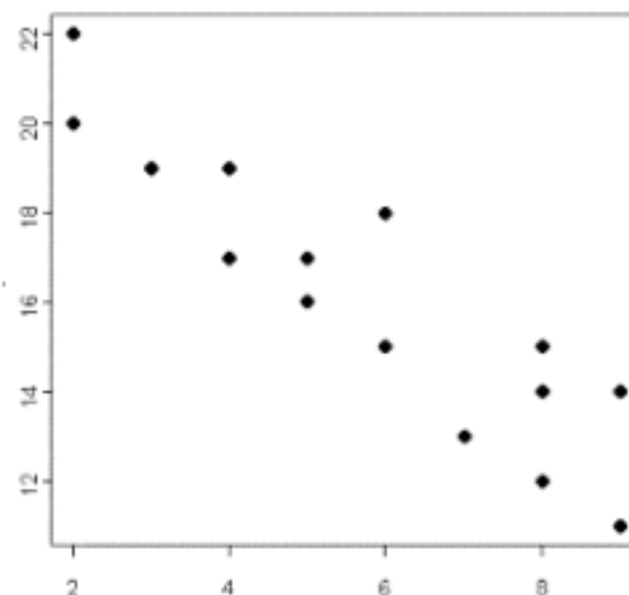
¿Cuándo es apropiado usar un gráfico de dispersión?

- Cuando queremos explorar la potencial relación entre dos **variables numéricas continuas**
- Cuántas veces sale cara al tirar una moneda?
- La estatura de un grupo de personas?

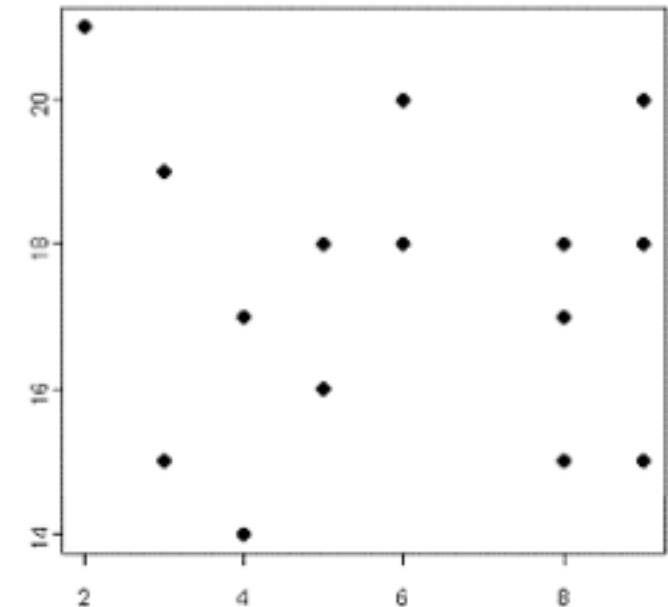
Los gráficos de dispersión nos ayudan a explorar potenciales asociaciones entre variables



Asociación positiva

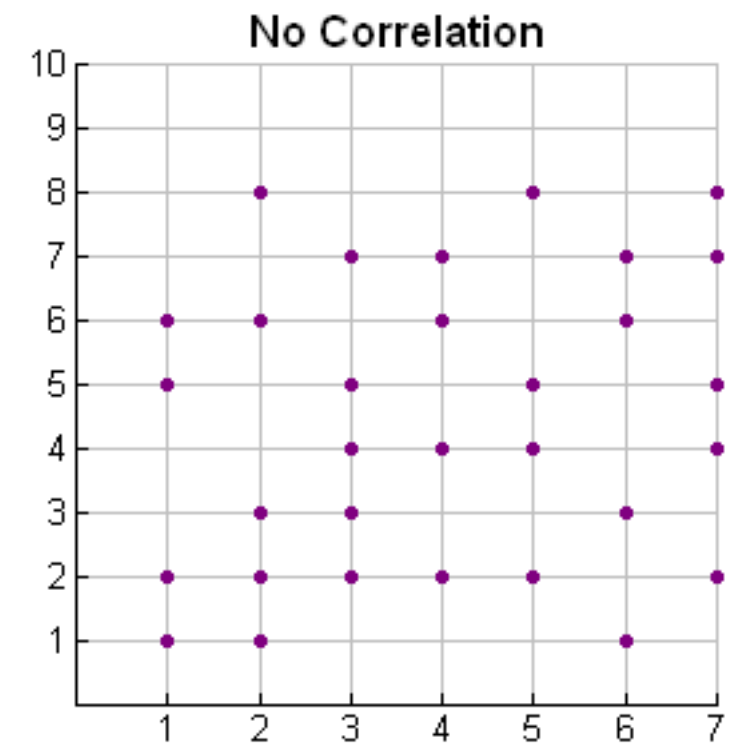
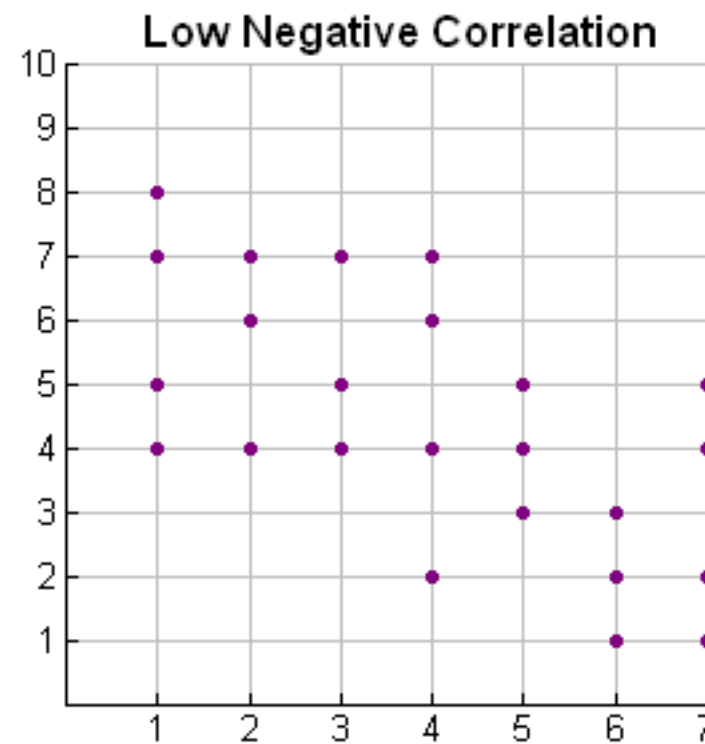
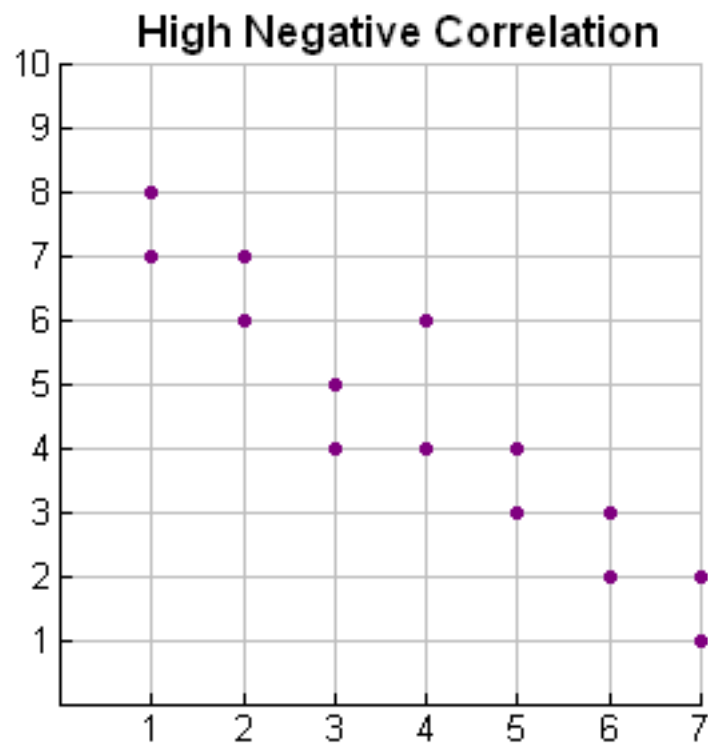
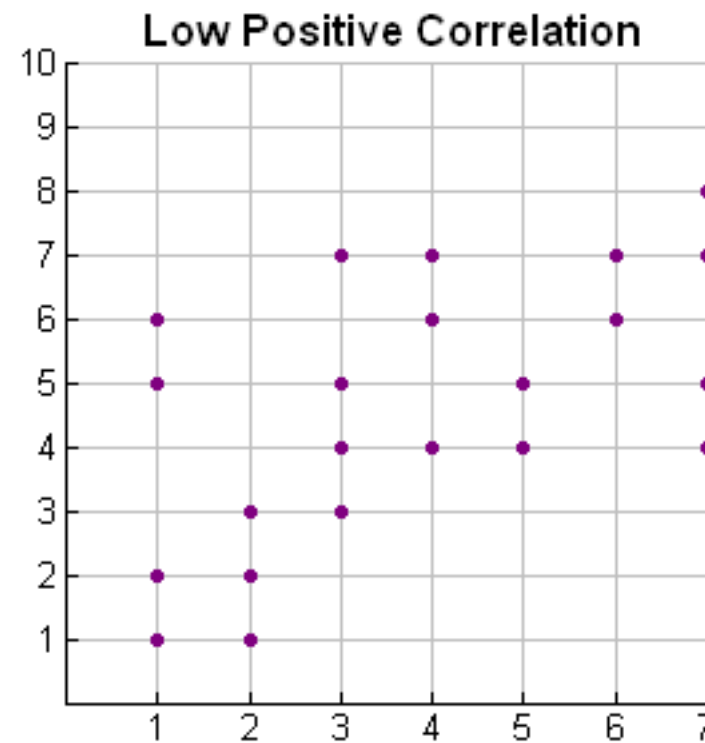
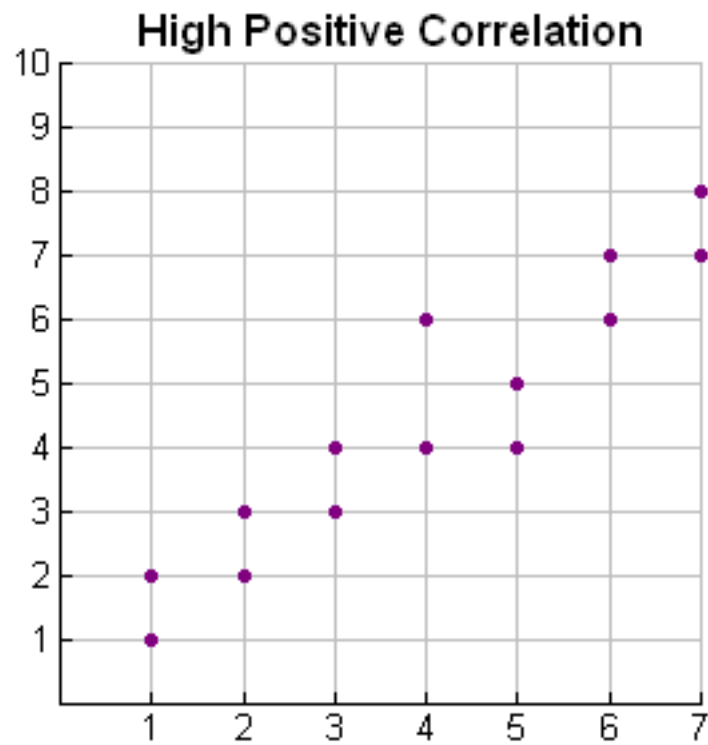


Asociación negativa

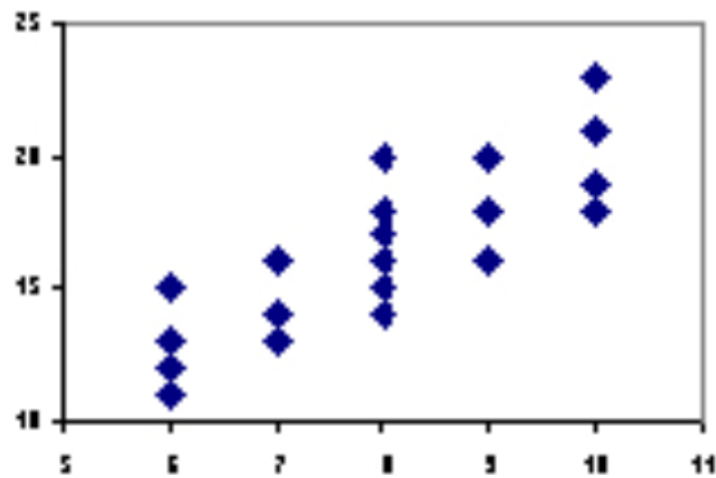


Ninguna tendencia

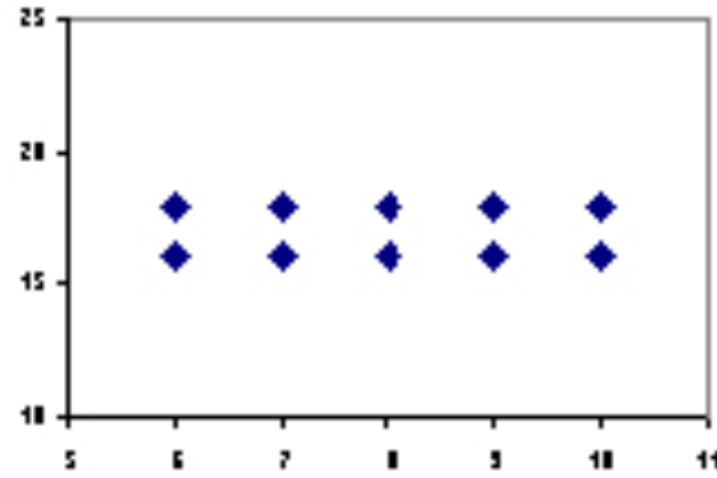
Tipos de asociaciones



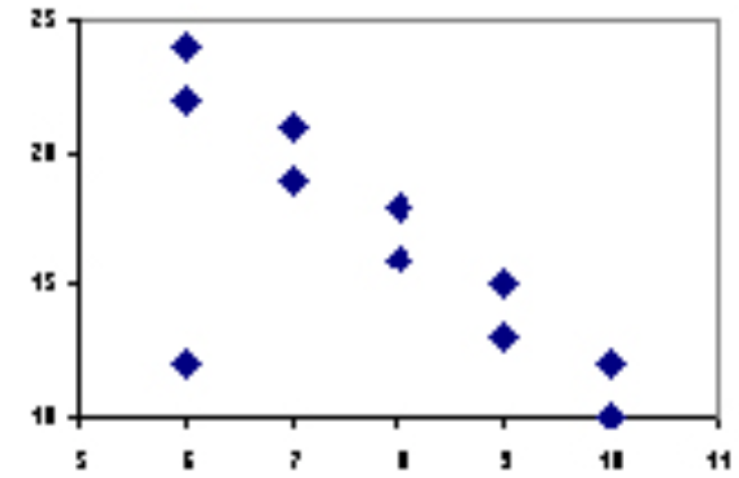
Tipos de asociaciones



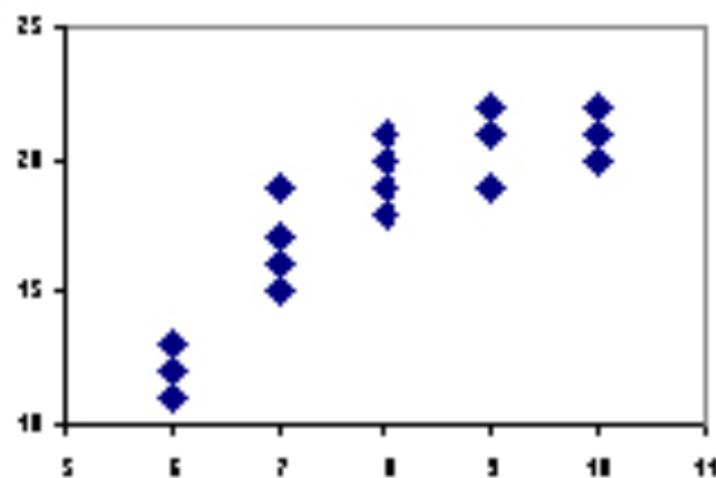
Linear, positive slope, weak



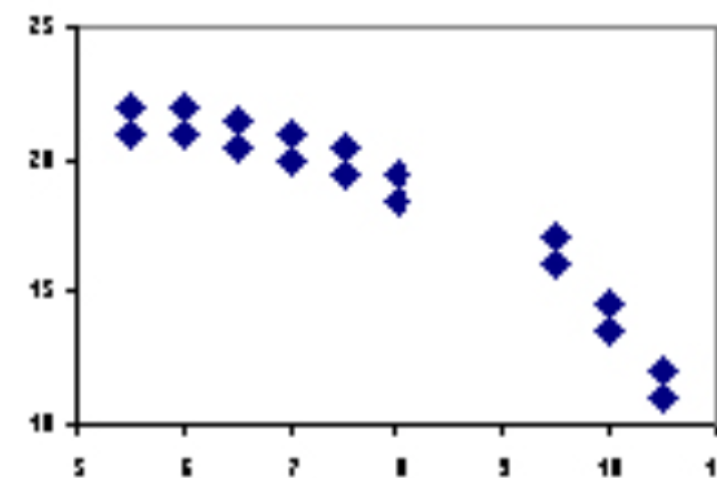
Linear, zero slope, strong



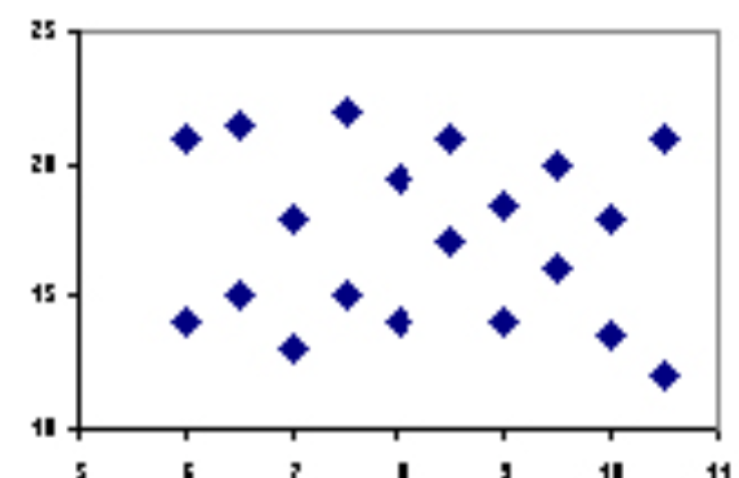
Linear, negative slope, strong,
with outlier



Nonlinear, positive slope, weak

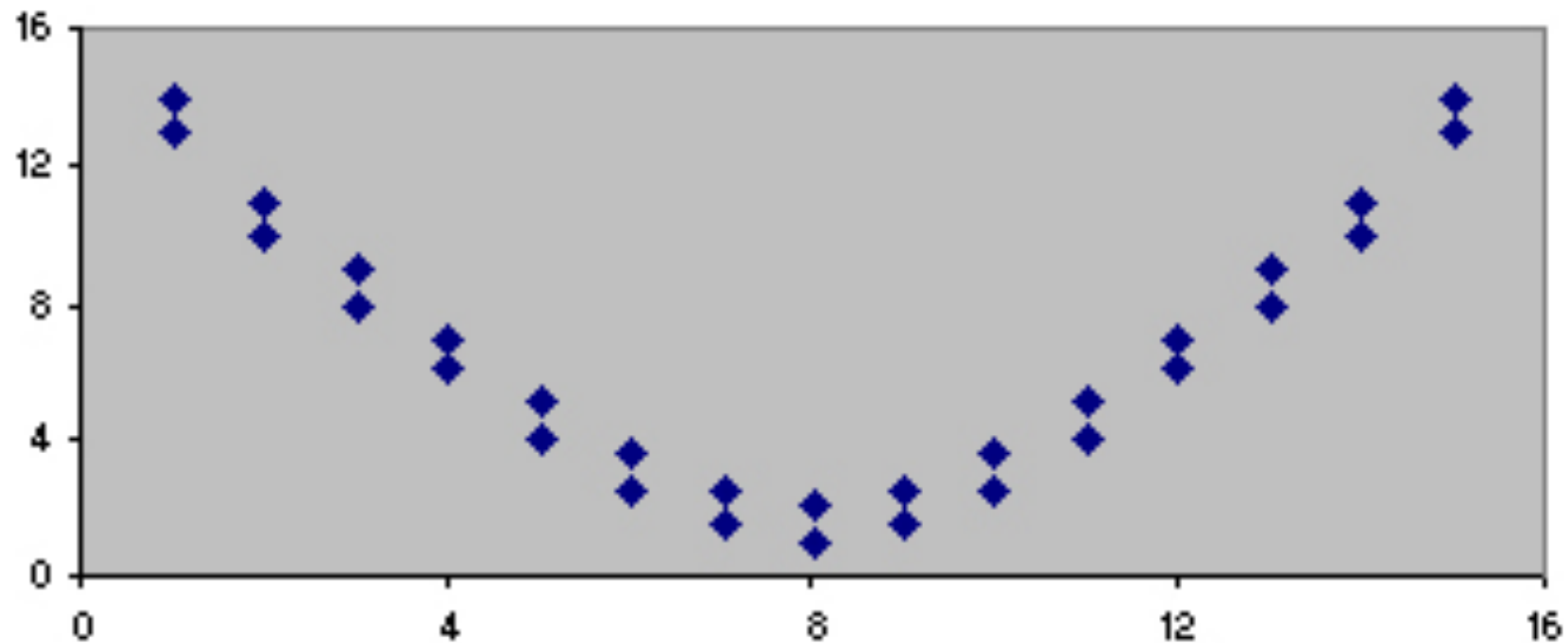


Nonlinear, negative slope, strong,
with gap



Nonlinear, zero slope, weak

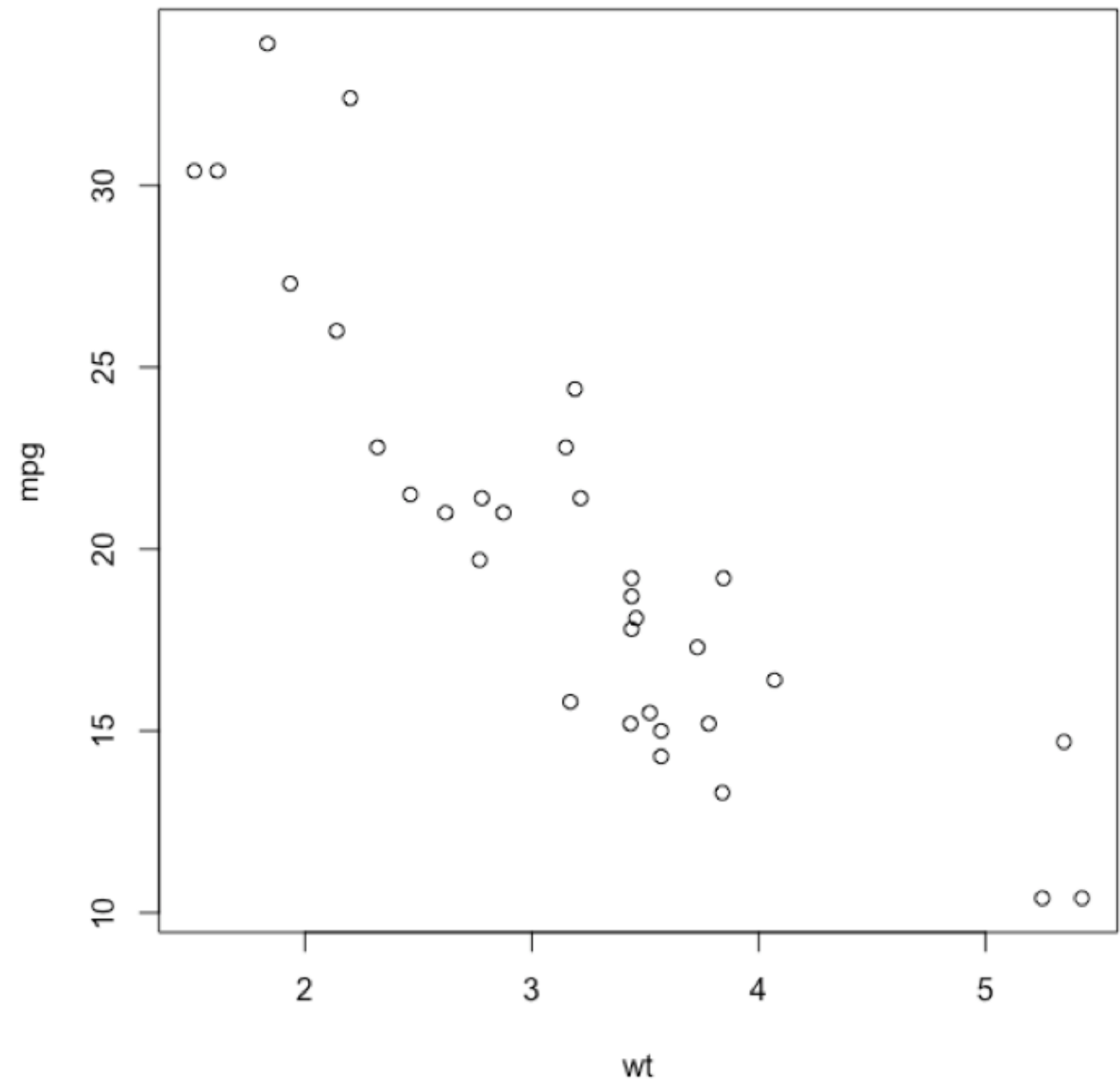
Pregunta?



- La correlación es fuerte
- La pendiente es negativa
- La pendiente es cero

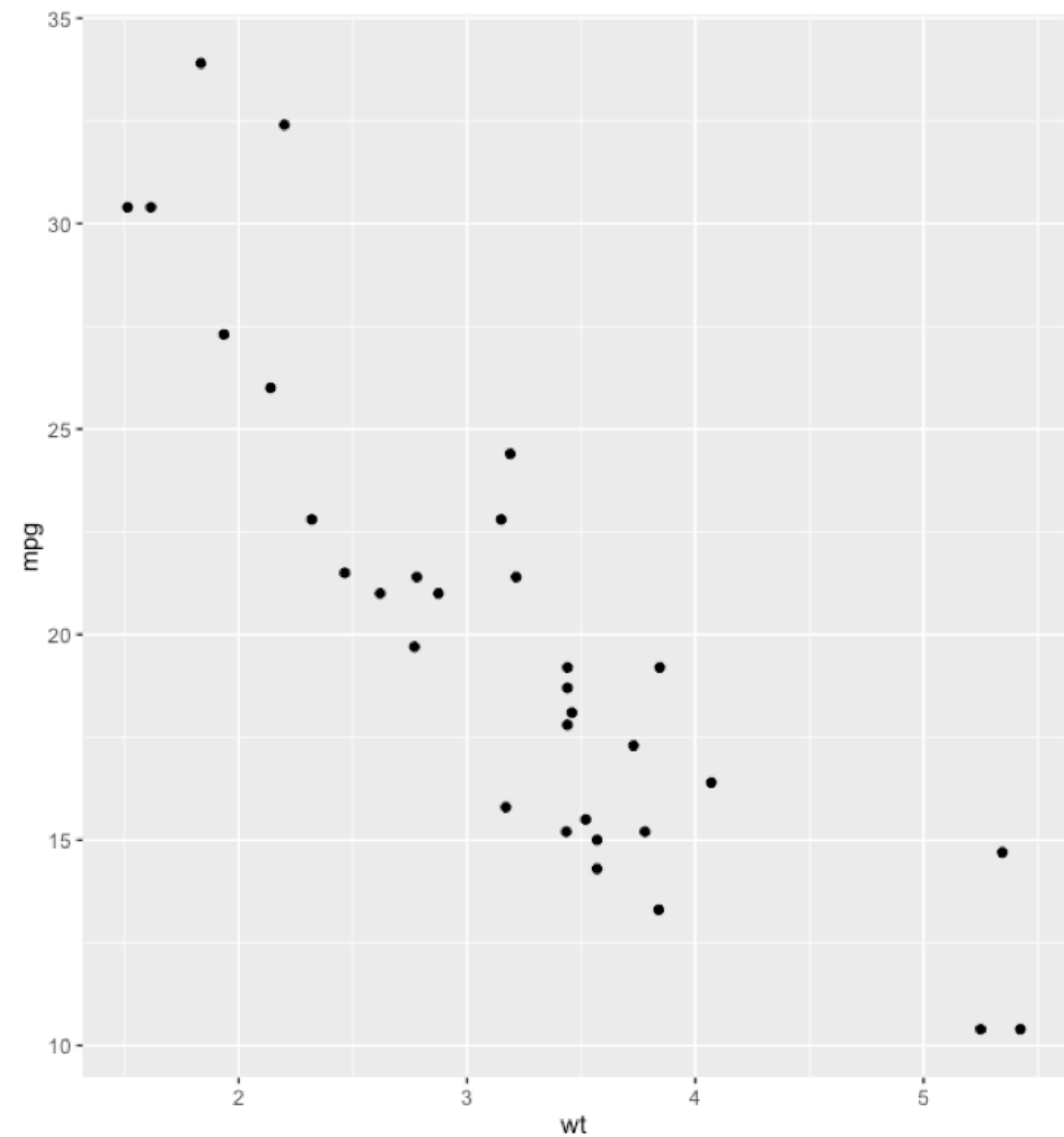
Ejemplo en R

- Abran R Studio
- `attach(mtcars)`
- `plot(wt, mpg)`



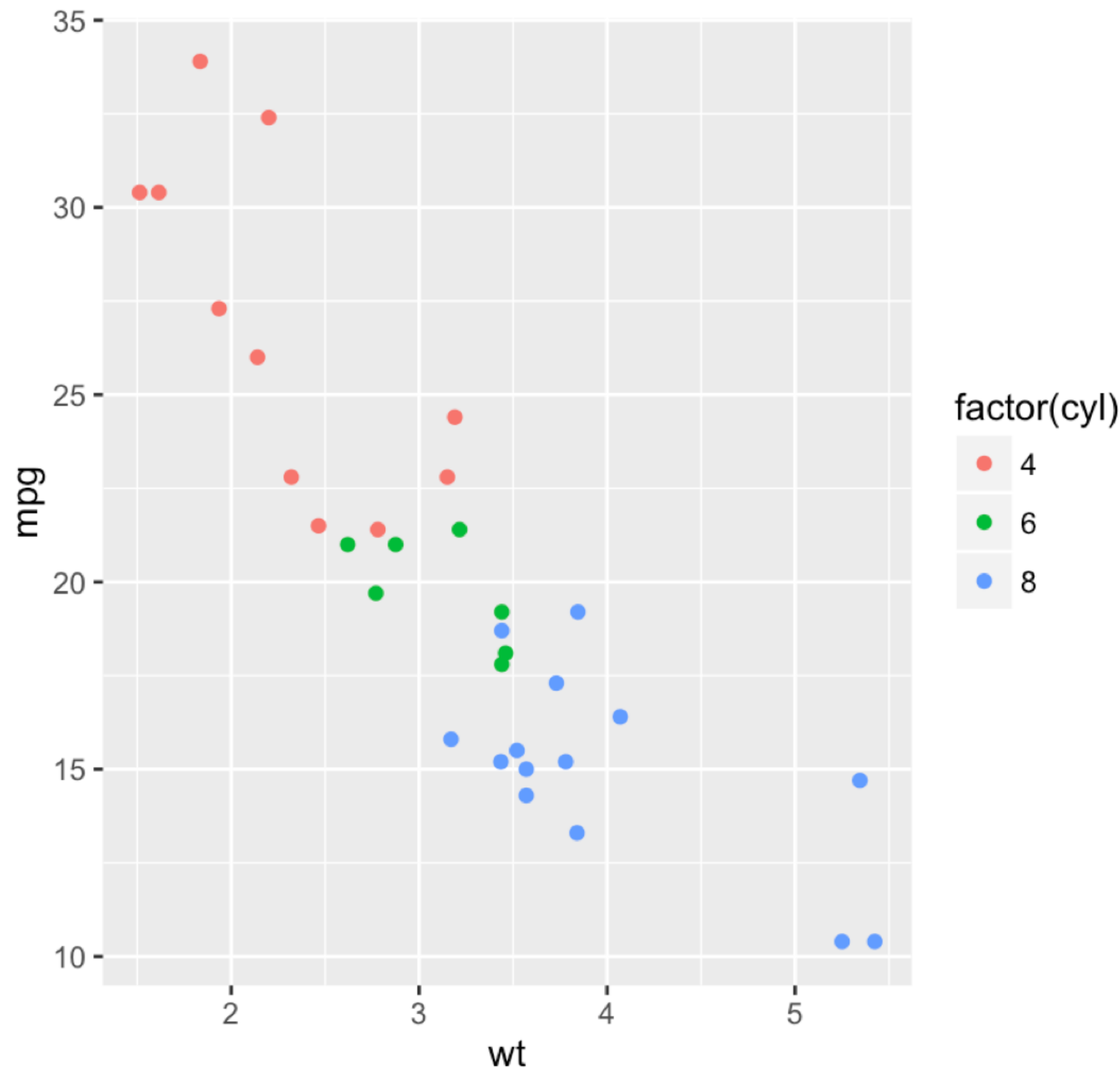
Ejemplo en R con ggplot2

- `library(ggplot2)`
- `p <- ggplot(mtcars, aes(wt, mpg))`
- `p + geom_point()`

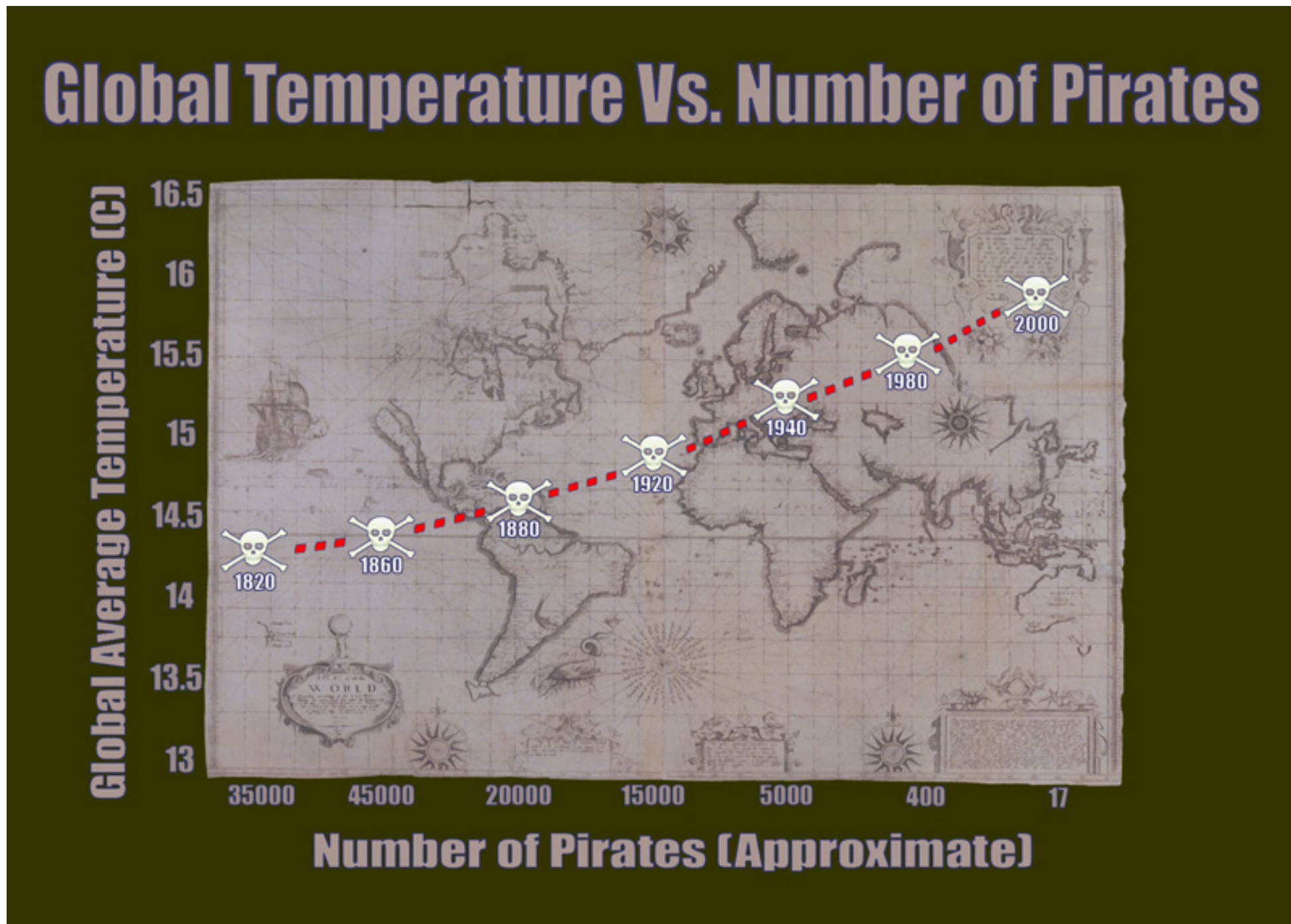


Ejemplo en R con ggplot2

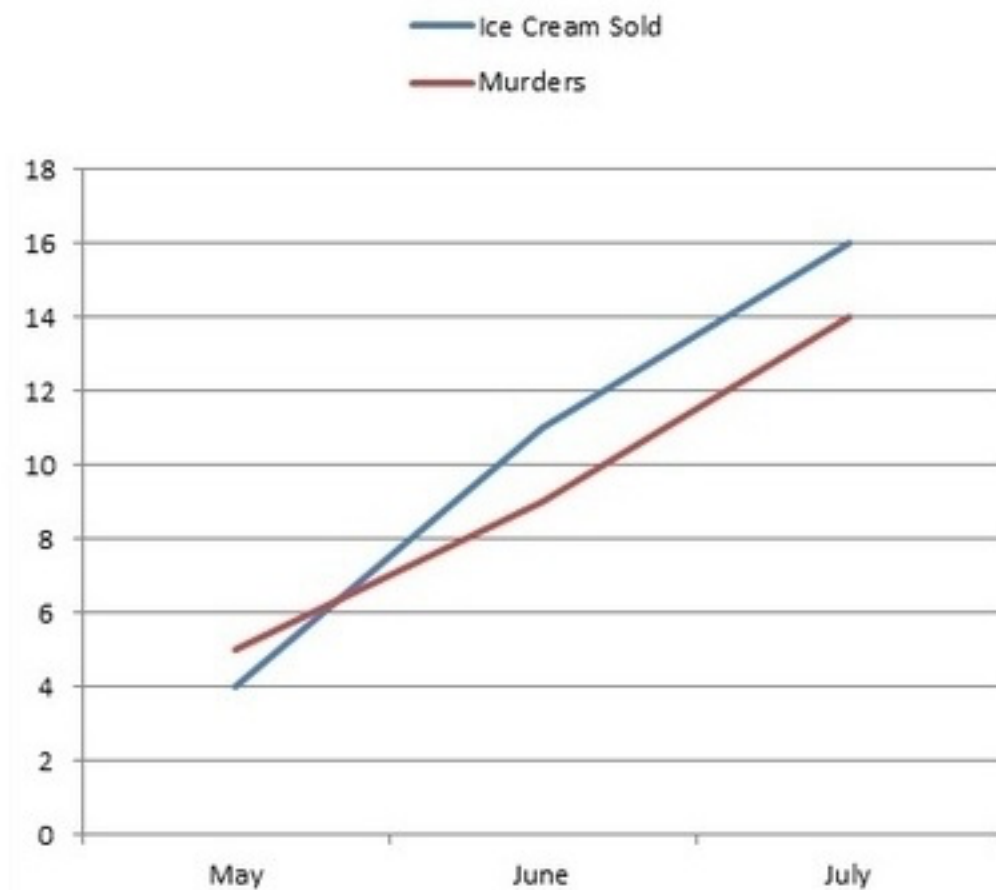
- `p +
geom_point(aes(colour = factor(cyl)))`



Correlación no implica causa/efecto



Correlación no implica causa/efecto



Correlación no implica causa/efecto

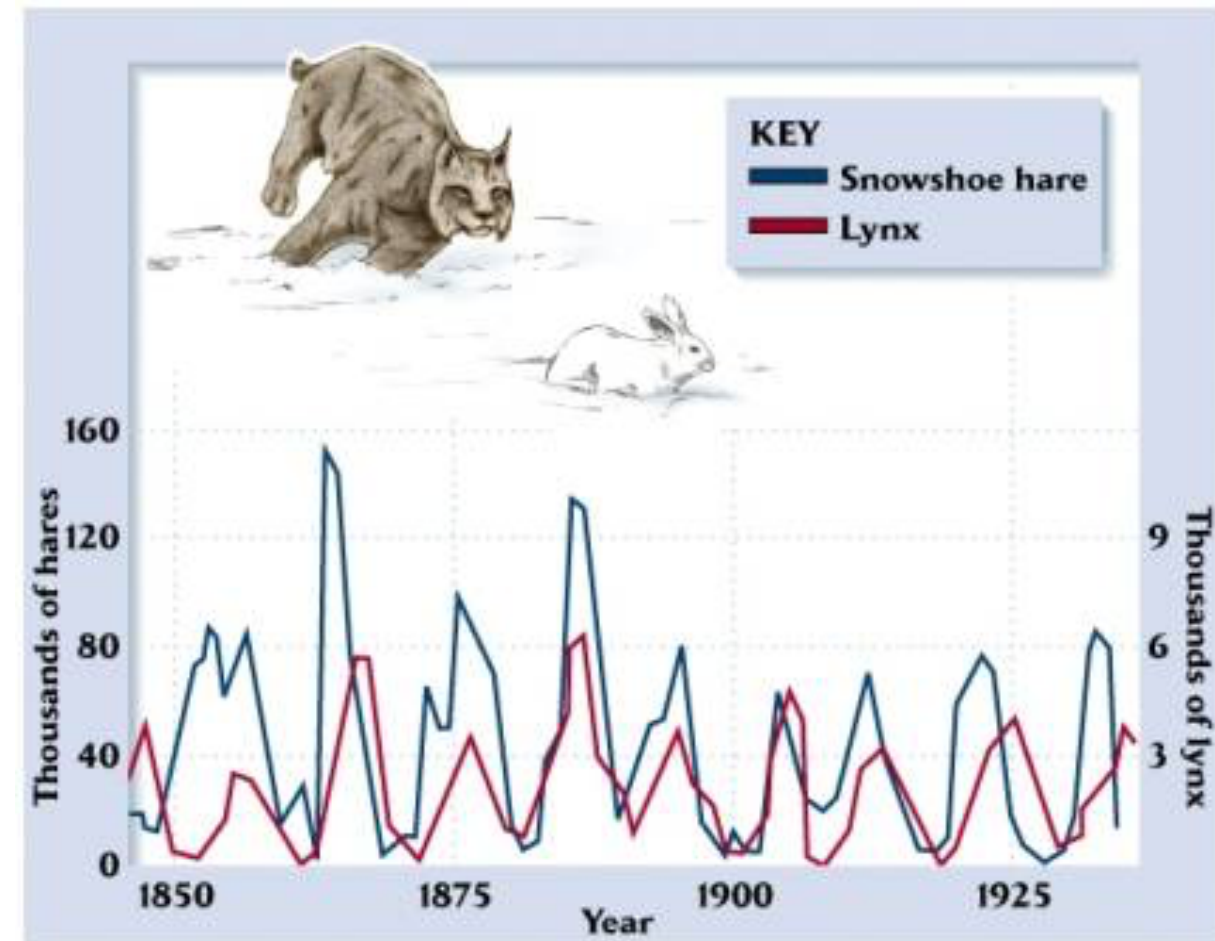
- X causa Y, caso ideal
- Y causa X, causalidad reversa. Cuando los molinos de viento rotan más fuerte, hay más viento.
- Países con alta deuda externa, crecen más lento?

Correlación no implica causa/efecto

- Z causa X e Y
- Consumo de helados y ataques de tiburones
- ¿Qué cosa causa ambos fenómenos?

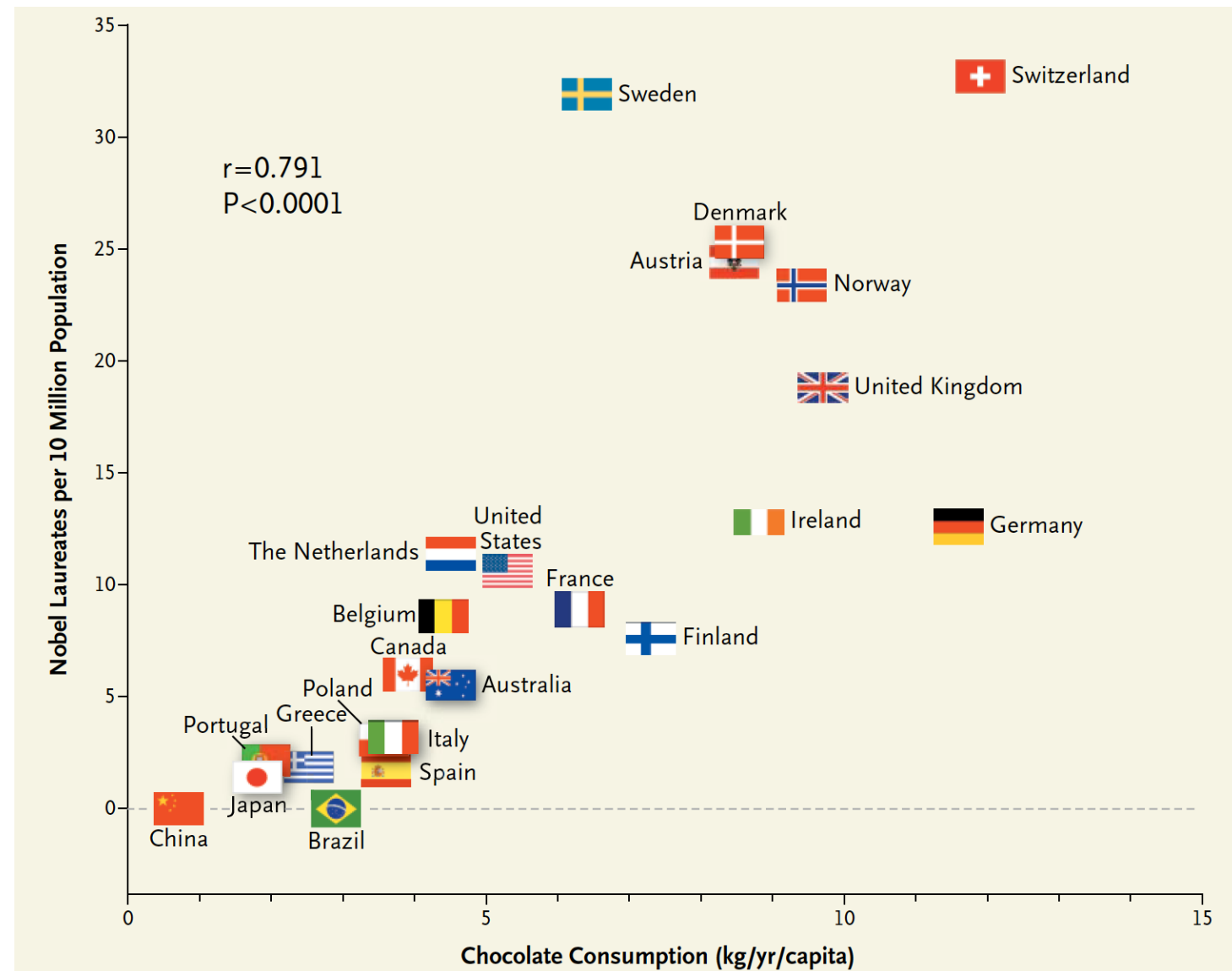
Correlación no implica causa/efecto

- Circularidad, X causa Y, Y causa X
- Aumenta la población de lince y baja la población de liebres. Esto causa que baje la población de lince, y aumente la de liebres



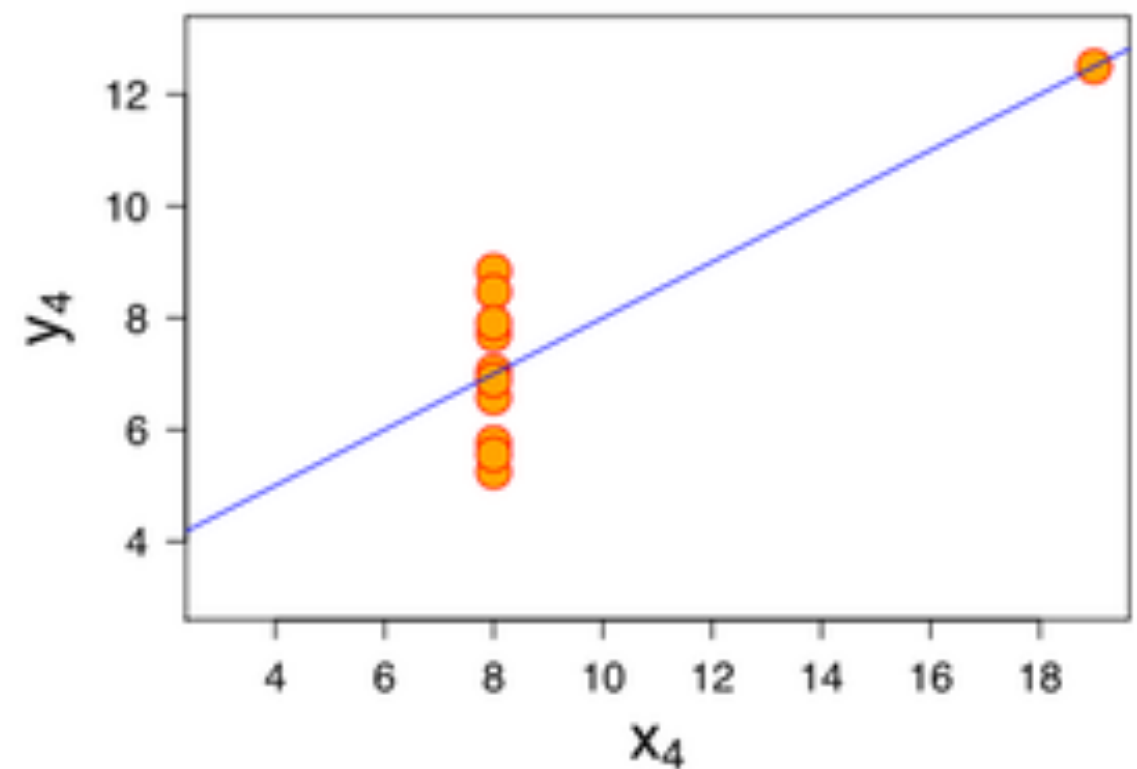
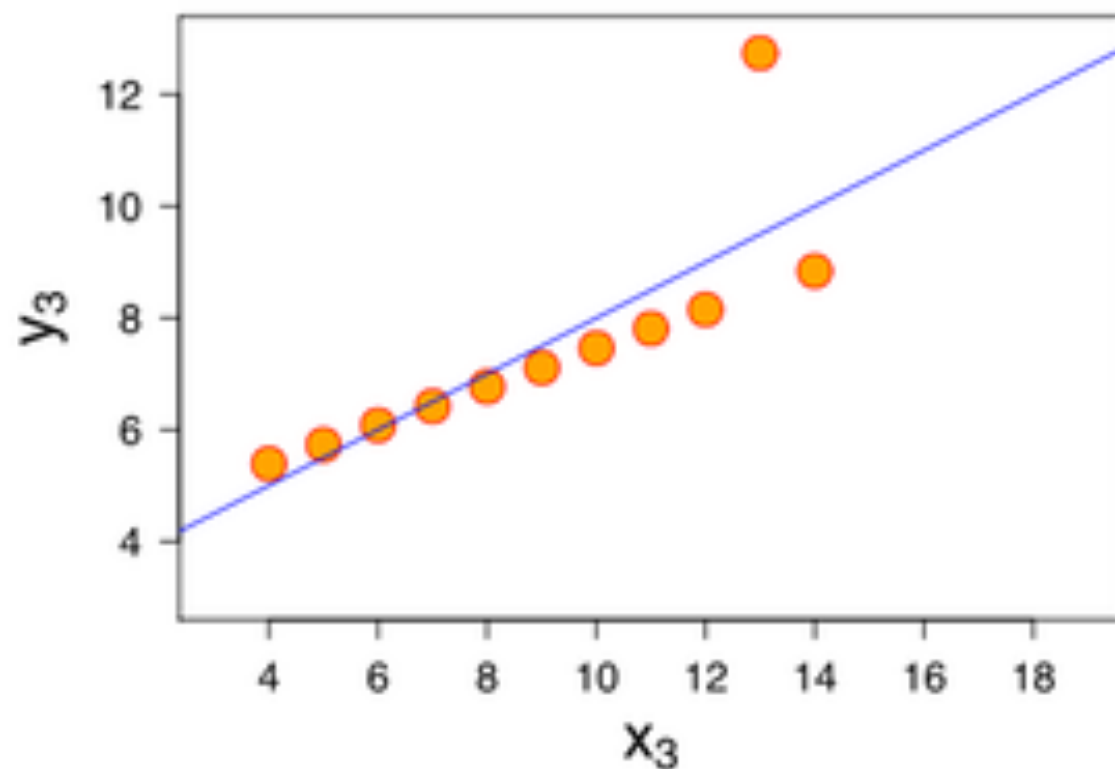
Correlación no implica causa/efecto

- X e Y están correlacionadas por coincidencia
- Cantidad de ganadores de premios Nobel y consumo de chocolate



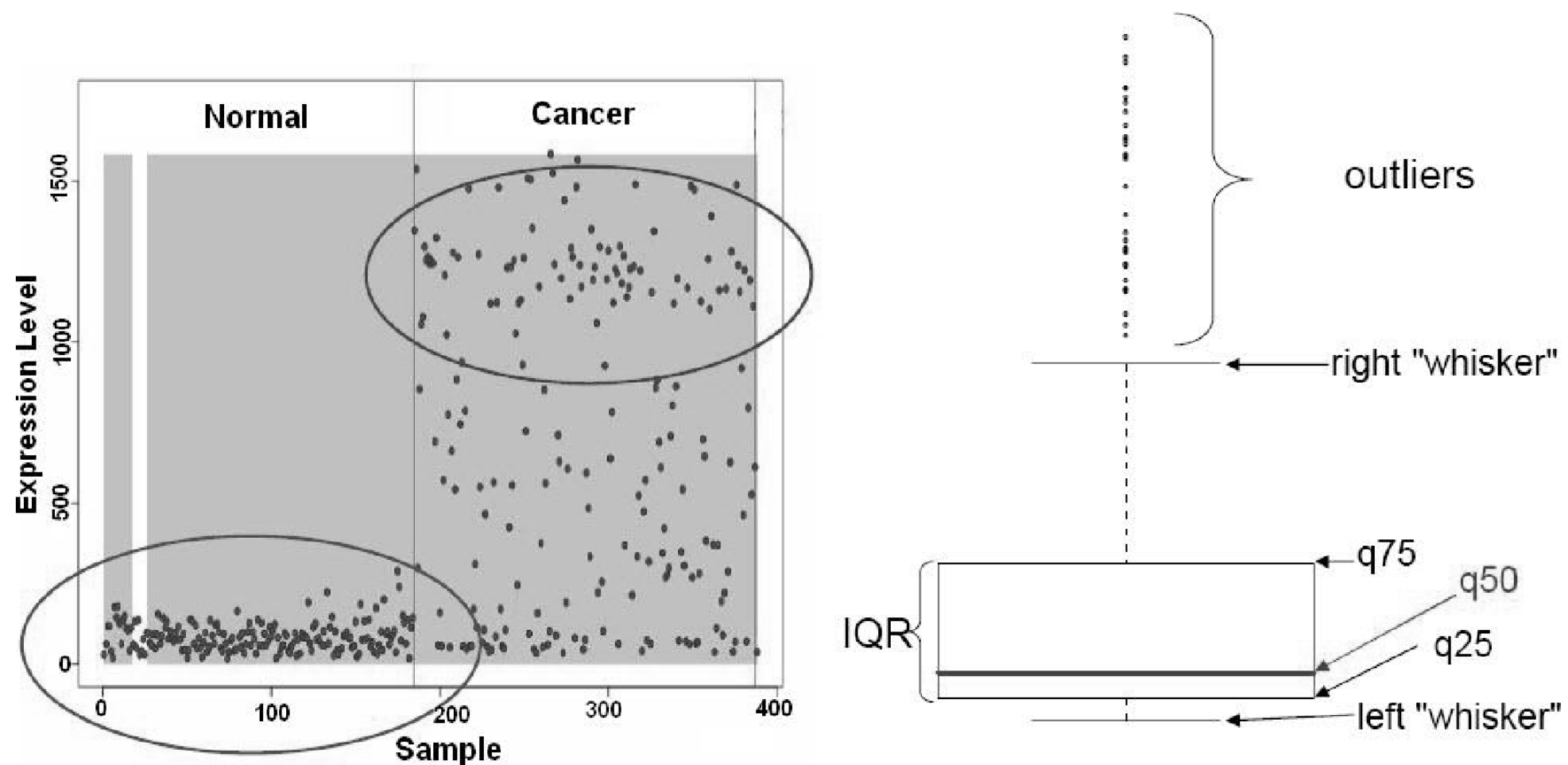
Detección de valores atípicos

- Outliers en inglés
- Es una observación tan desviada de las otras observaciones que uno sospecha que fueron generadas por otro mecanismo



Detección de valores atípicos

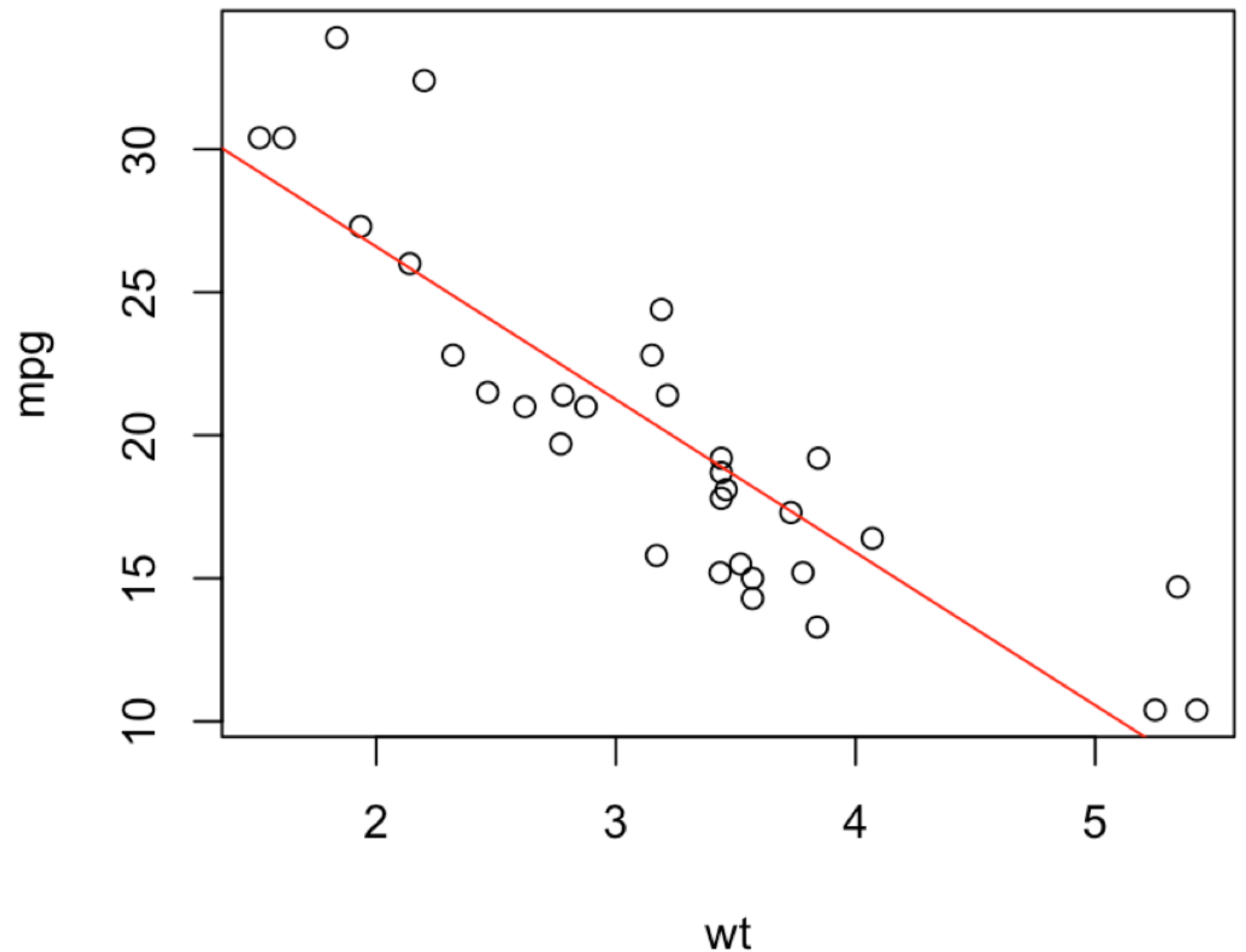
- Valores atípicos o outliers pueden tener significado biológico



Ajustar una recta en R

```
plot(wt, mpg)
```

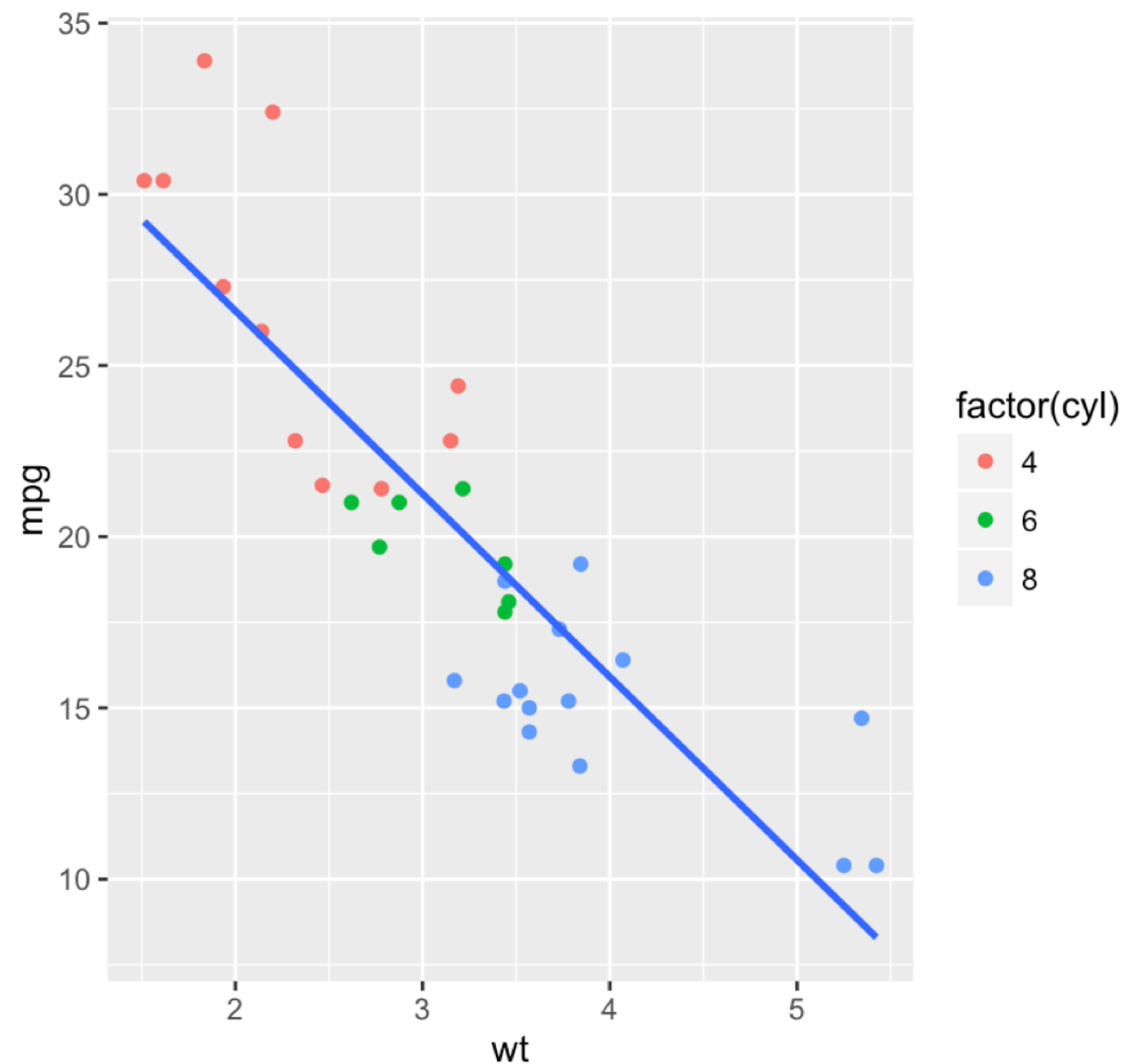
```
abline(lm(mpg~wt),  
col="red")
```



Ajustar una recta en ggplot2

```
p +  
geom_point(aes(colour = factor(cyl))) +
```

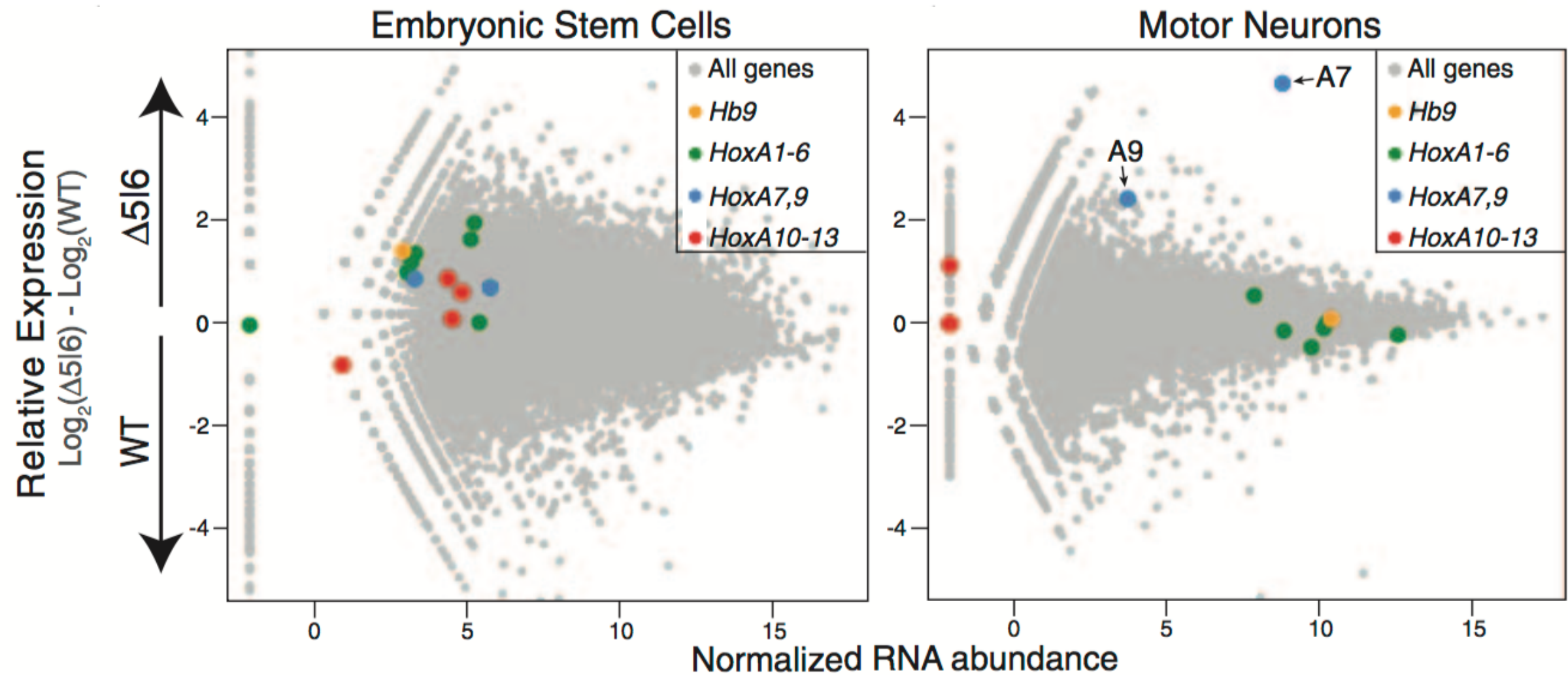
```
geom_smooth(method = "lm", se = FALSE)
```



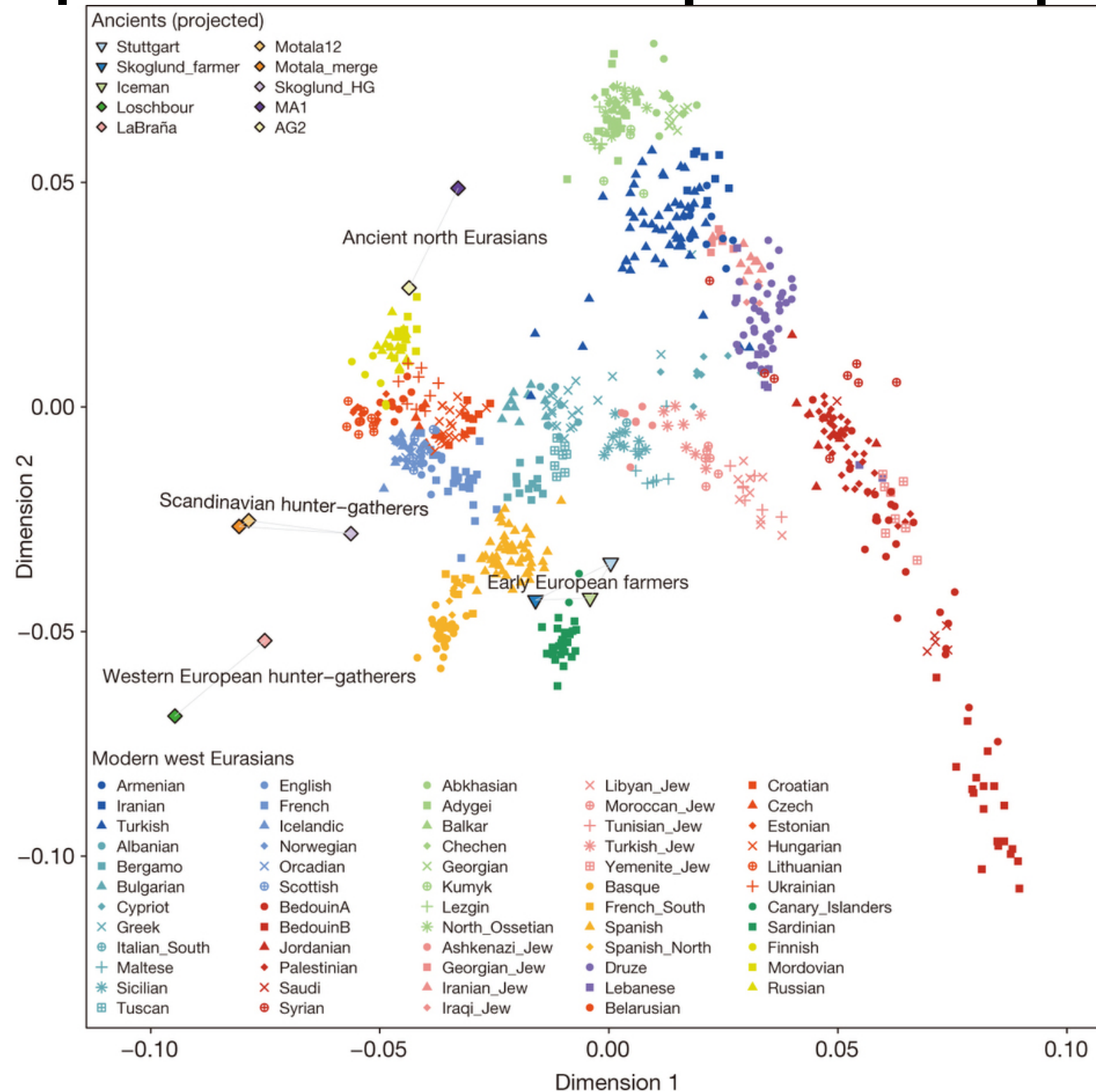
Ejemplos en genómica

- MA plot
- Principal components plot
- Representar clusters en microbiomas
- Prevalencia de grupos taxonómicos

Mean Average plot

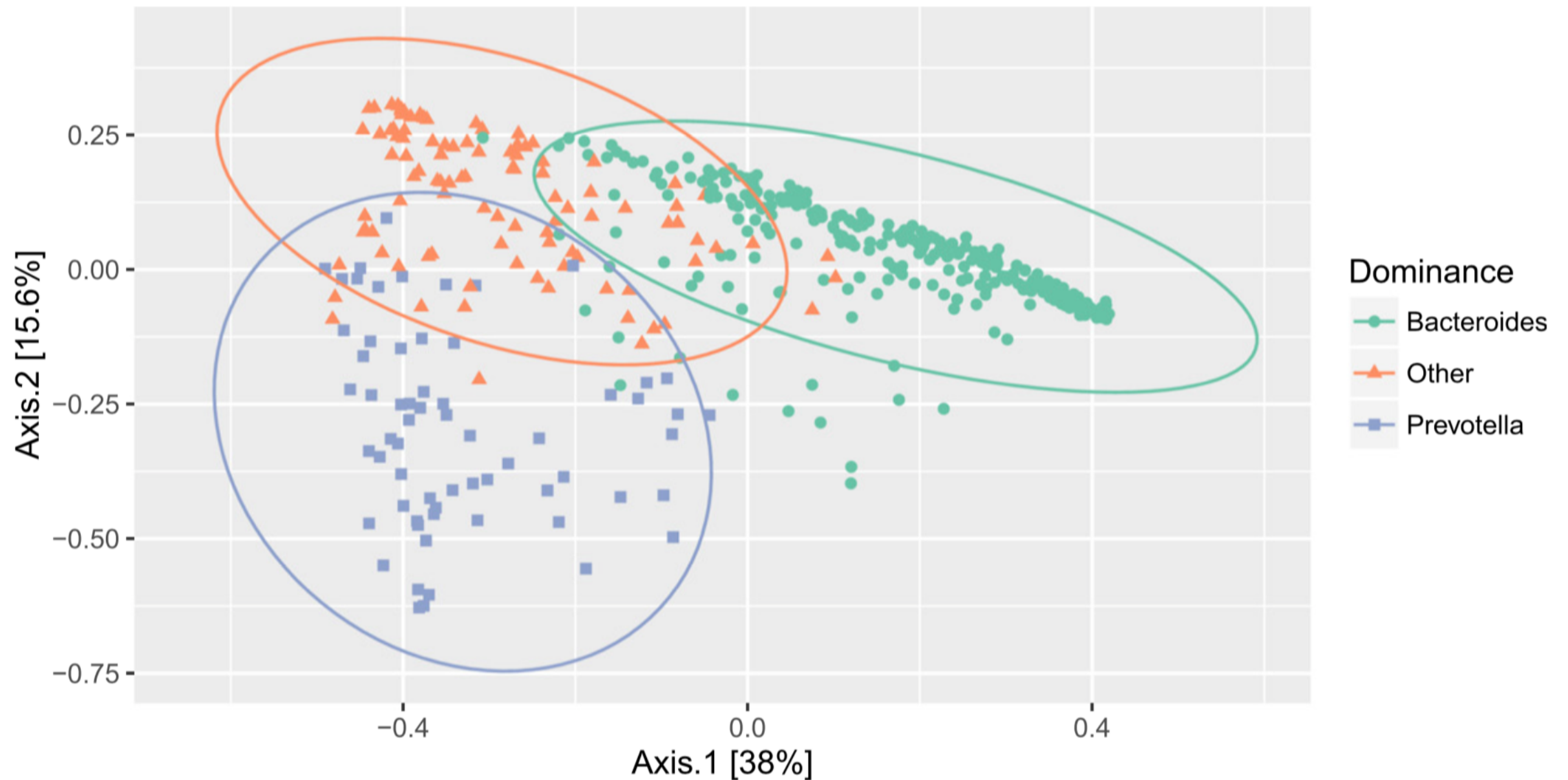


Componentes principales



Representar clusters en microbiomas

a



Prevalencia de grupos taxonómicos

