



# Modelado de Simulación en R commander para ciencias sociales y biosanitarias

# R commander

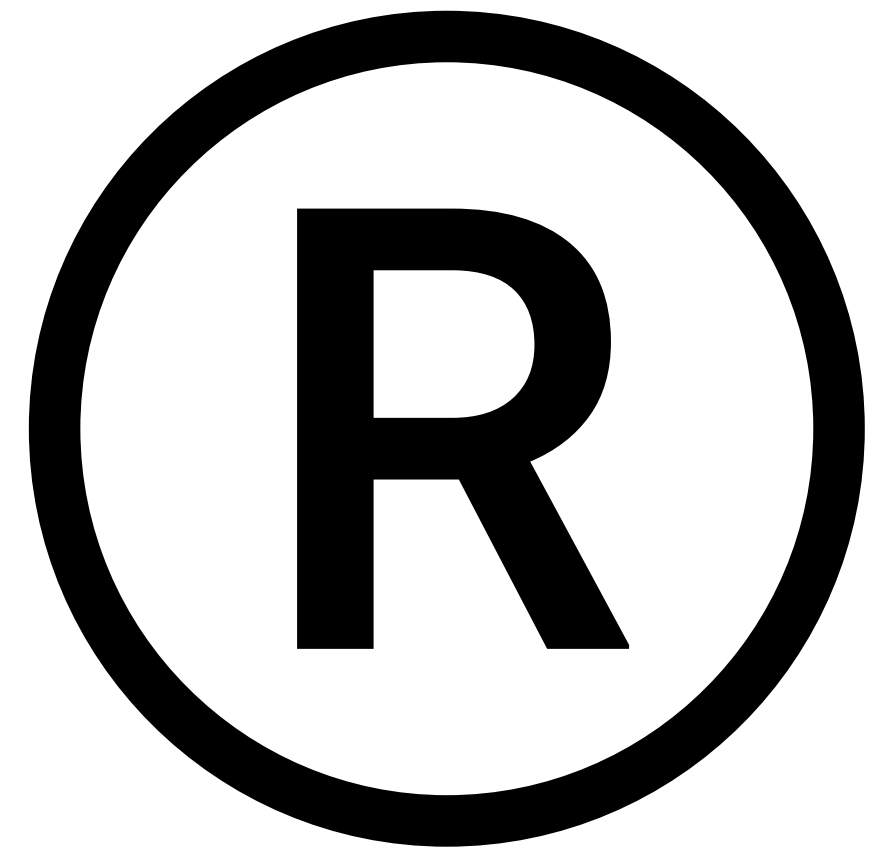
MODELADO DE DATOS CUANTITATIVOS

## **R COMMANDER**

Interfaz gráfica de usuario para el lenguaje de programación R

## **CONFIRMAR O NEGAR LA HIPÓTESIS ALTERNATIVA**

Surge de la pregunta de investigación y tiene por objetivo contrastar mediante un modelo lógico matemático la relación entre variables (cuantitativas o cualitativas cuantificadas dummy)



# Modelización en R

**01**

Regresión lineal o no lineal: logística

**02**

Modelos de ANOVA

**03**

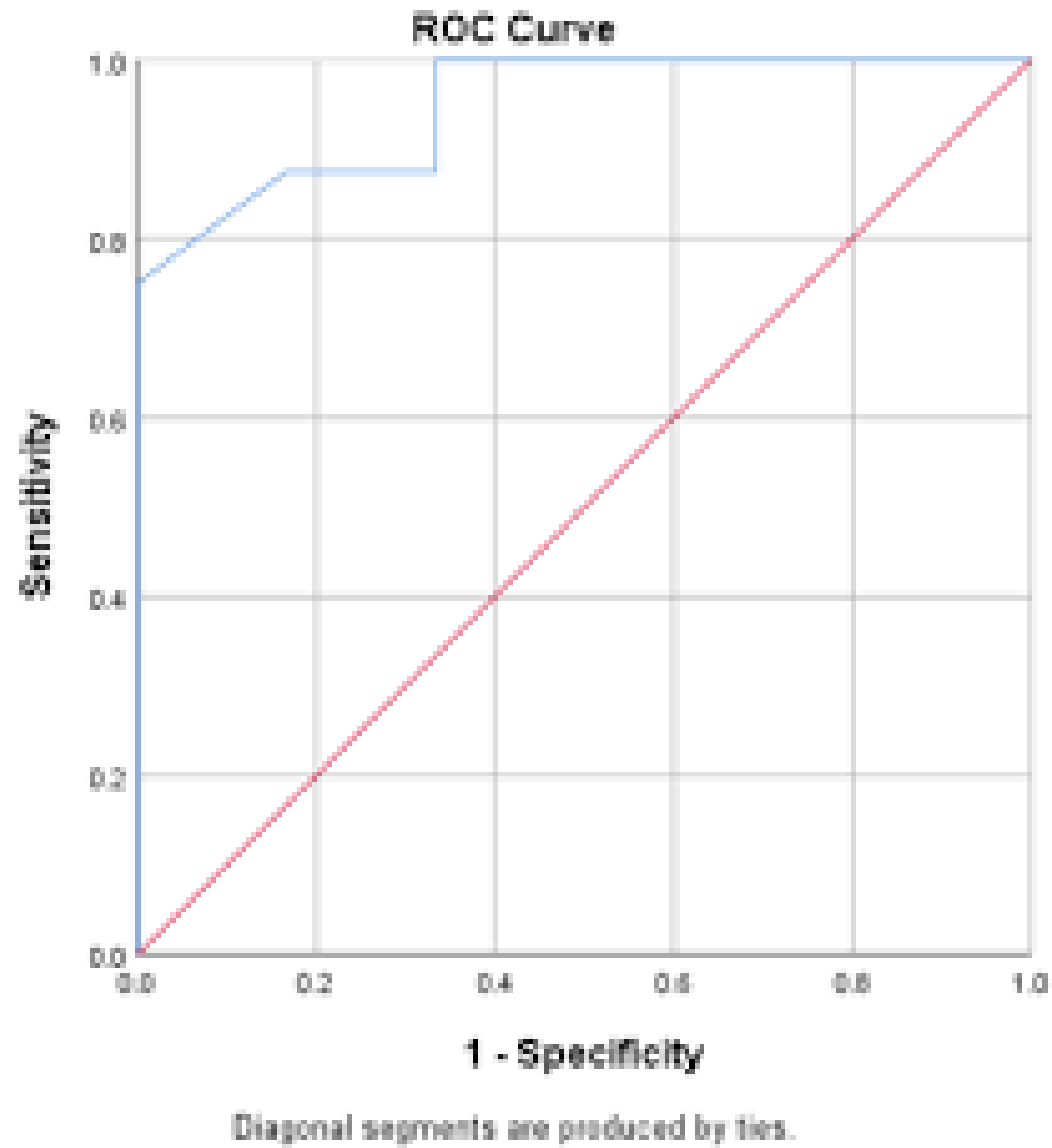
Simulación de Monte Carlo

**04**

Análisis de series de tiempo

**05**

Simulaciones de eventos discretos



# Modelización en R commander

01

## PREGUNTA DE INVESTIGACIÓN

El tipo de base de datos y su hipótesis determina la modelización y visualización de los mismos

02

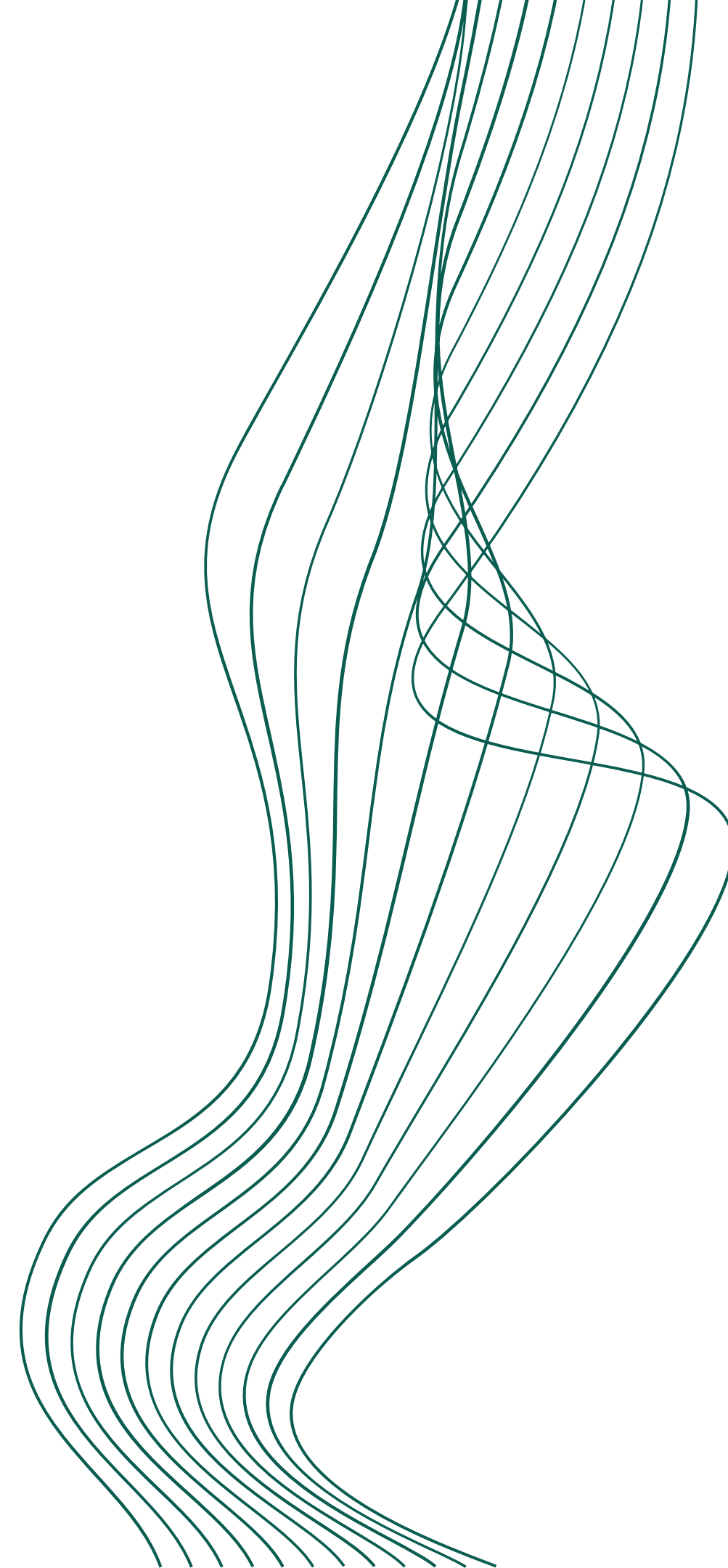
## ESTADÍSTICAS DESCRIPTIVAS

Depende del tipo de variable (cuantitativa o cualitativa) implica una medidas de tendencia central  
Media, desviación estándar e IC al 95% con histogramas o boxplots

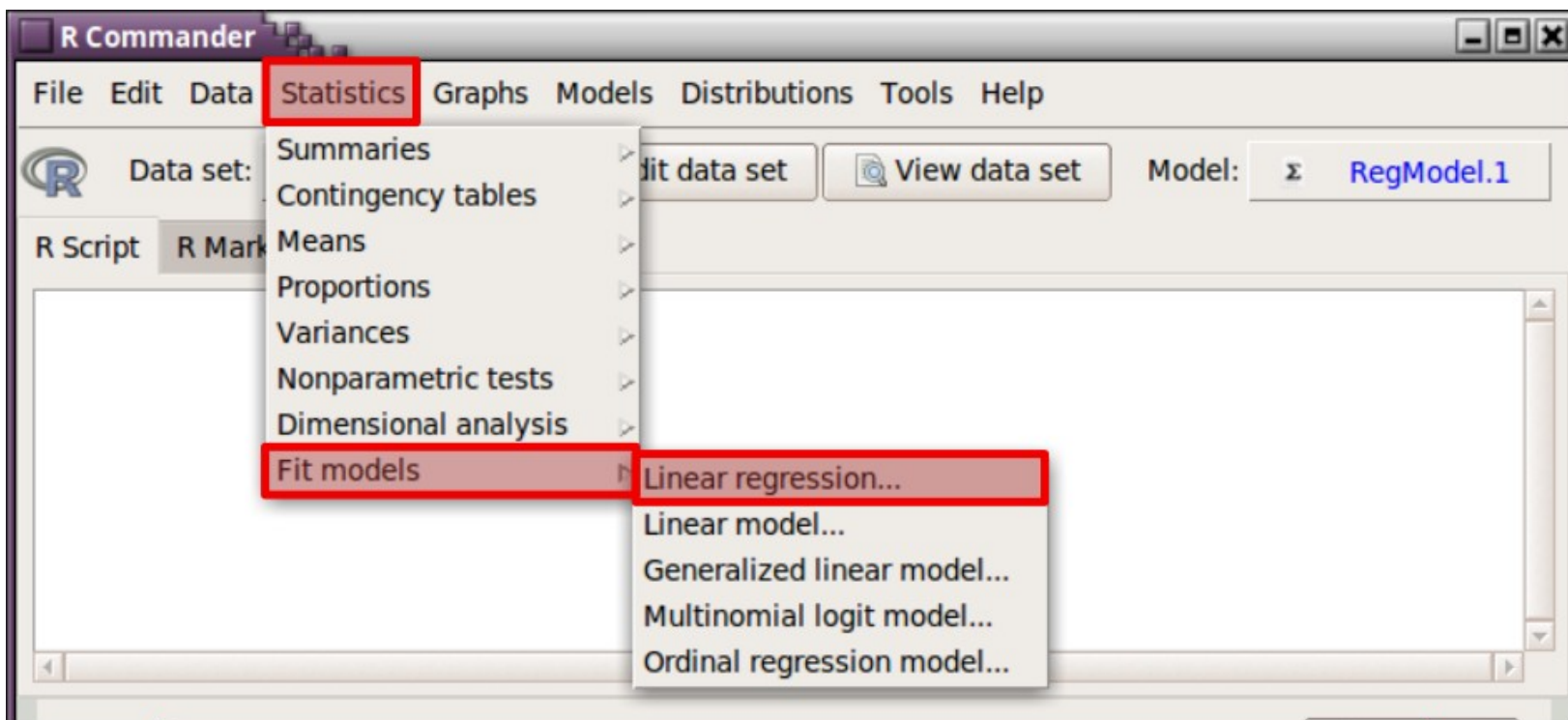
03

## PRUEBAS DE HIPÓTESIS Y MODELOS MULTIVARIANTES

Pruebas de determinación de la hipótesis (chi-cuadrado o T-student) y modelos más complejos como regresión logística



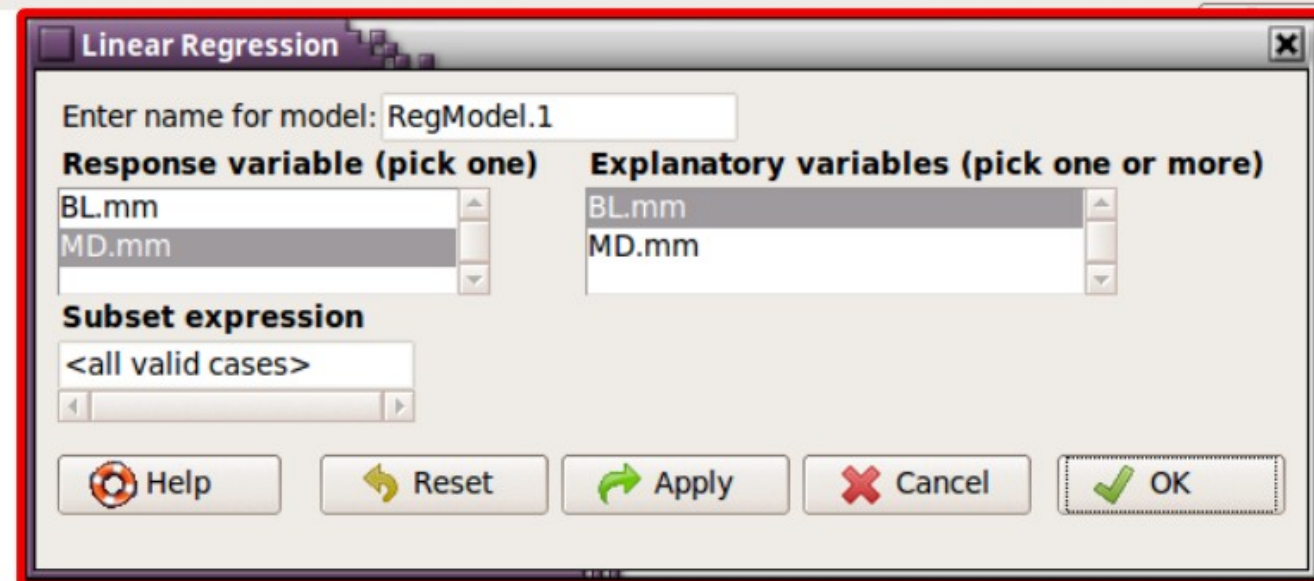
# Modelización de los datos



Es imprescindible entender cómo  
están organizados los datos

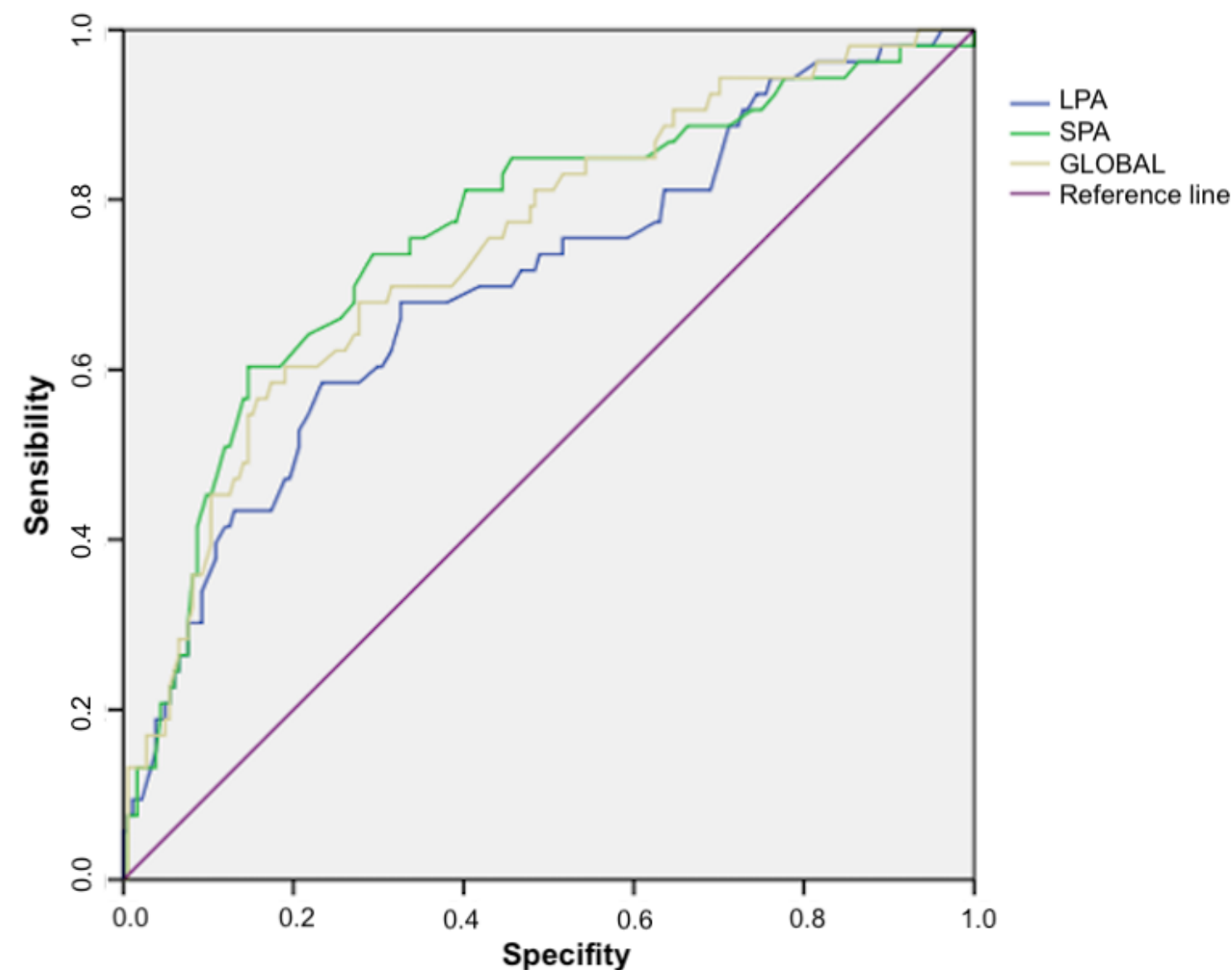


LOS MODELOS MÁS IMPLEMENTADOS EN CIENCIAS  
SOCIALES Y BIOSANITARIAS  
SE CENTRAN EN REGRESIÓN LOGÍSTICA Y  
REGRESIÓN LINEAL MULTIPLE



**R commander**

# Regresión logística



**Curva ROC: visualización del modelo**

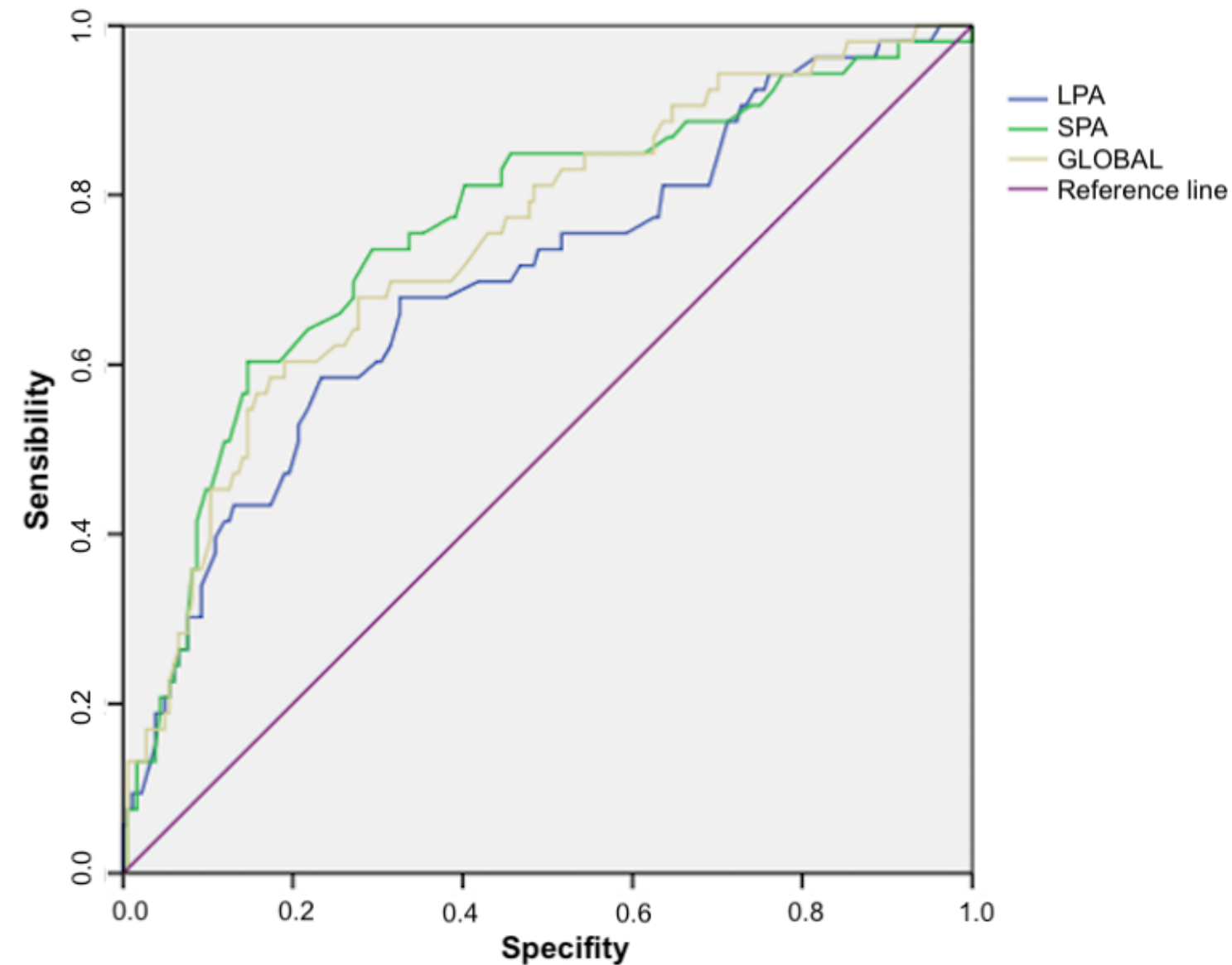
Utilizado para analizar datos binarios, es decir, datos que solo pueden tener dos posibles valores (0 o 1)  
Se tiene que evaluar la variable dependiente. Ejemplo SI o No riesgo



ESTE TIPO DE MODELO SE UTILIZA COMÚNMENTE EN ESTUDIOS DE SALUD PARA EVALUAR LA ASOCIACIÓN ENTRE UNA VARIABLE DE INTERÉS Y UN RESULTADO BINARIO, COMO LA PRESENCIA O AUSENCIA DE UNA ENFERMEDAD



# Consideraciones



Es importante tener en cuenta que antes de ajustar un modelo de regresión logística, se deben cumplir ciertas condiciones



**ES RECOMENDABLE REALIZAR PRUEBAS DE BONDAD DE AJUSTE Y VALIDACIÓN CRUZADA PARA EVALUAR LA CAPACIDAD PREDICTIVA DEL MODELO.**

**Curva ROC: visualización del modelo**

# Pasos en R y análisis

Variables en la ecuación									
		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95,0 % para Exp(B)	
Paso 1	CAFE	-.093	1,304	,283	1	,595	,500	,039	6,439
	MEDIO	1,609	,856	3,532	1	,060	5,000	,933	26,785
	CAFE by MEDIO	1,451	1,390	1,090	1	,296	4,267	,280	65,010
	Constante	-1,009	,775	4,317	1	,038	,200		

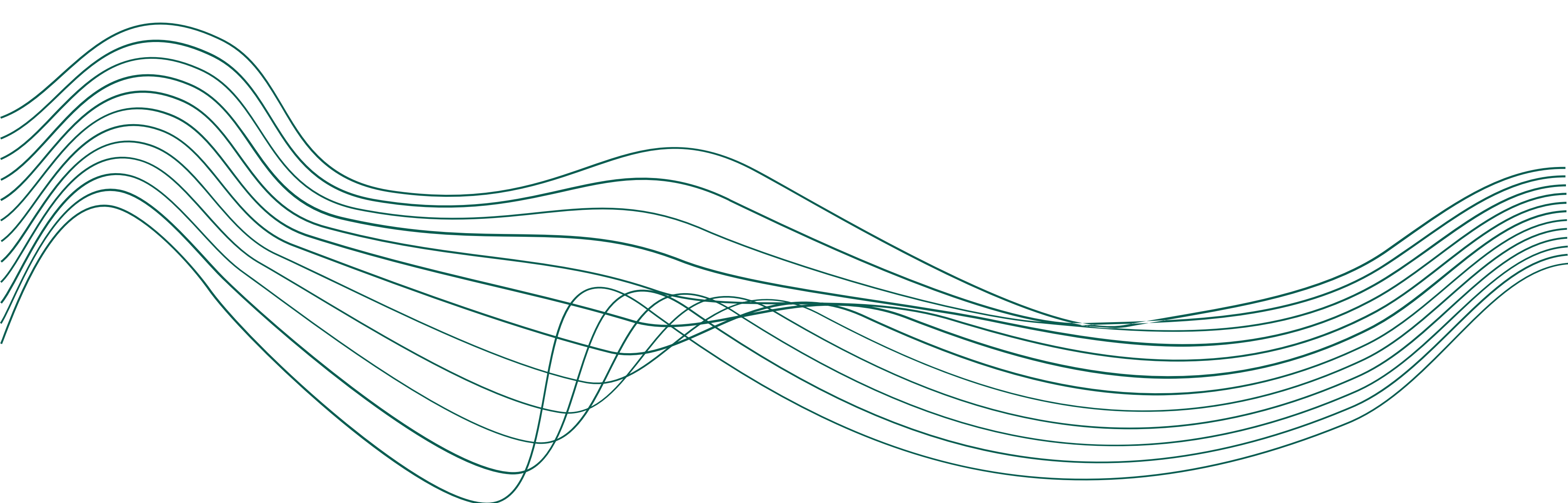
a. Variable(s) introducida(s) en el paso 1: CAFE, MEDIO, CAFE \* MEDIO .

**Es necesario ajustar los datos  
que dan fallo en el modelo.**

**Identificando los factores clave  
según el valor significativo**

1. Abra R Commander y cargue los datos que desea analizar. Puede hacer esto utilizando la opción "Datos" en el menú principal de R Commander.
2. Seleccione "Modelos lineales y generalizados" en el menú "Estadísticas" y luego seleccione "Modelo de regresión logística".
3. En el cuadro de diálogo que aparece, seleccione la variable dependiente (es decir, el resultado binario que desea predecir) y las variables independientes (es decir, las variables que se utilizarán para predecir el resultado).
4. Haga clic en "Aceptar" y R Commander ajustará automáticamente el modelo de regresión logística y mostrará los resultados en una ventana separada.
5. Analice los resultados del modelo, incluyendo los coeficientes de la regresión logística, los valores de la prueba de Wald y los intervalos de confianza.

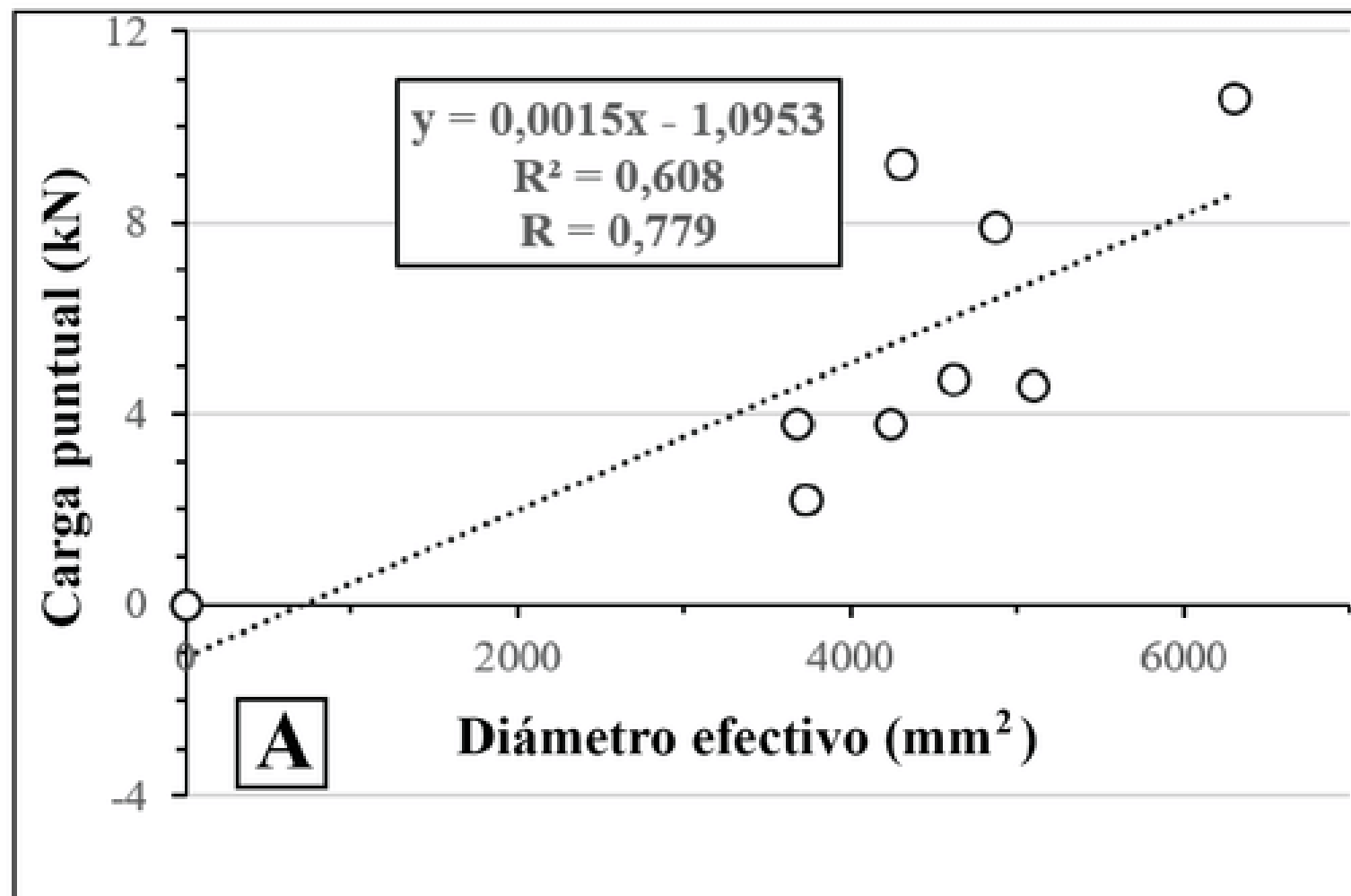


A series of approximately 15 thin, dark teal wavy lines that originate from the top left and flow towards the right, creating a sense of movement and depth. They vary in amplitude and frequency, with some lines crossing each other.

# Regresión lineal multiple

- Evaluar la relación entre una variable dependiente y dos o más variables independientes.
- Este tipo de análisis es común en muchos campos, incluyendo la investigación médica y de salud pública

# Consideraciones



Es importante tener en cuenta que antes de ajustar un modelo de regresión logística, se deben cumplir ciertas condiciones



LA LINEALIDAD ENTRE LAS VARIABLES, LA INDEPENDENCIA ENTRE LAS OBSERVACIONES, LA HOMOGENEIDAD DE VARIANZA Y LA NORMALIDAD DE LOS ERRORES

# Pasos en R y análisis

Resumen del modelo

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	,765 <sup>a</sup>	,585	,481	43,64452

a. Predictores: (Constante), VAR00001

ANOVA<sup>a</sup>

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	10747,456	1	10747,456	5,642	,076 <sup>b</sup>
	Residuo	7619,378	4	1904,844		
	Total	18366,833	5			

a. Variable dependiente: VAR00002

b. Predictores: (Constante), VAR00001

Coefficientes<sup>a</sup>

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	95,0% intervalo de confianza para B	
		B	Error estándar	Beta			Límite inferior	Límite superior
1	(Constante)	-7,061	38,560		-,183	,864	-114,121	99,999
	VAR00001	7,615	3,206	,765	2,375	,076	-1,286	16,516

a. Variable dependiente: VAR00002

1. Abra R Commander y cargue los datos que desea analizar. Puede hacer esto utilizando la opción "Datos" en el menú principal de R Commander.
2. Seleccione "Modelos lineales y generalizados" en el menú "Estadísticas" y luego seleccione "Modelo de regresión lineal".
3. En el cuadro de diálogo que aparece, seleccione la variable dependiente (es decir, la variable que desea predecir) y las variables independientes (es decir, las variables que se utilizarán para predecir la variable dependiente).
4. Haga clic en "Aceptar" y R Commander ajustará automáticamente el modelo de regresión lineal múltiple y mostrará los resultados en una ventana separada.
5. Analice los resultados del modelo, incluyendo los coeficientes de regresión, la estadística F y los valores p.

**Es necesario ajustar los datos  
que dan fallo en el modelo.  
Identificando los factores clave  
según el valor significativo**

# Consideraciones de los modelos

INDEPENDIENTEMENTE DE LA HERRAMIENTA



## MODELO LÓGICO MATEMÁTICO

Es importante seguir las mejores prácticas estadísticas y validar el modelo para garantizar resultados precisos y confiables.

Es imprescindible ejecutar y validar el modelo utilizando datos de validación, independientes a los usados para obtener el modelo.