

CONTENIDO POR CAPÍTULOS Y ACCESO A LOS CUADROS DE CÓDIGO R

Para facilitar la reproducibilidad del libro, concretamente los cuadros de código **R**, se ha dejado a disposición del lector una versión *online* de este manual en el siguiente enlace:

<https://cdr-book.github.io/>

Parte I. Ciencia, datos, software... y científicos

- 1 ¿Es la ciencia de datos una ciencia?
- 2 Metodología en ciencia de datos
- 3 R para ciencia de datos
- 4 Ética en la ciencia de datos

Parte II. Bienvenidos a la jungla de datos

- 5 Gestión de bases de datos relacionales
- 6 Gestión de bases de datos NoSQL
- 7 Gobierno, gestión y calidad del dato
- 8 Integración y limpieza de datos
- 9 Selección y transformación de variables
- 10 Herramientas para el análisis en ciencia de datos
- 11 Análisis exploratorio de datos

Parte III. Fundamentos de estadística

- 12 Probabilidad
- 13 Inferencia estadística
- 14 Muestreo y remuestreo

Parte IV. Modelización estadística

- 15 Modelización lineal
- 16 Modelos lineales generalizados
- 17 Modelos aditivos generalizados
- 18 Modelos mixtos
- 19 Modelos *sparse* y métodos penalizados de regresión
- 20 Modelización de series temporales
- 21 Análisis discriminante
- 22 Análisis conjunto
- 23 Análisis de tablas de contingencia

Parte V. Machine learning supervisado

- 24 Árboles de clasificación y regresión
- 25 Máquinas de vector soporte
- 26 Clasificador *k*-vecinos más próximos
- 27 Naive Bayes
- 28 Métodos ensamblados: *bagging* y *random forest*
- 29 *Boosting* y el algoritmo XGBoost

Parte VI. Machine learning no supervisado

- 30 Análisis cluster: clusterización jerárquica
- 31 Análisis cluster: clusterización no jerárquica
- 32 Análisis de componentes principales
- 33 Análisis factorial
- 34 Escalamiento multidimensional
- 35 Análisis de correspondencias

Parte VII. Deep learning

- 36 Redes neuronales artificiales
- 37 Redes neuronales convolucionales

Parte VIII. Ciencia de datos de texto y redes

- 38 Minería de textos
- 39 Análisis de grafos y redes sociales

Parte IX. Ciencia de datos espaciales

- 40 Trabajando con datos espaciales
- 41 Geoestadística
- 42 Modelos econométricos espaciales
- 43 Procesos de puntos

Parte X. Comunica y colabora

- 44 Informes reproducibles con R Markdown y Quarto
- 45 Creación de aplicaciones web interactivas con Shiny
- 46 Git y GitHub R
- 47 Geoprocesamiento en nube

Parte XI. Casos de estudio en ciencia de datos

- 48 Análisis de una red criminal
- 49 Optimización de inversiones publicitarias
- 50 ¿Cómo tuitea Elon Musk?
- 51 Análisis electoral: de RStudio a su periódico favorito
- 52 El impacto de las crisis financiera y de la COVID-19 en el paro de CLM
- 53 Segmentación de clientes en el comercio minorista
- 54 Análisis de datos en medicina
- 55 Messi y Ronaldo: dos ídolos desde la perspectiva de los datos
- 56 Una nota sobre el cambio climático
- 57 Implementación de un sistema experto en el ámbito pediátrico
- 58 Predicción de consumo eléctrico con redes neuronales artificiales
- 59 El procesamiento del lenguaje natural para tendencias de moda en textil
- 60 Detección de fraude de tarjetas de crédito