

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

Repaso final © EDICIONES ROBLE, S.L.

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

Indice

Repaso final	3
Resumen	3

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

campusformacion.imf.com © EDICIONES ROBLE, S.L.
IVAN GARCIA GARCIA

Repaso final

Resumen



A lo largo de las seis unidades, se han conocido y expuesto someramente multitud de conceptos de estadística y programación. Esta combinación de conocimientos adquiridos para afrontar el análisis de datos, tanto en la componente teórica como en la práctica, aporta un nivel de iniciación suficientemente sólido para el estudiante como para poder profundizar en la materia.

La ciencia de datos aglomera tantas disciplinas y es foco de interés por parte de personas con formaciones tan dispares, que el conocimiento global necesario para dominar todos los aspectos de la materia es realmente complejo y vasto.

Se recomienda, por tanto, profundizar en los conocimientos estadísticos, así como en la programación en R para poder seguir creciendo en este campo, que está en constante evolución, ya que cada día se añaden nuevas técnicas y la complejidad es mayor. Para ello adjuntamos una completa bibliografía.

Por otra parte, además de seguir desarrollándose aprendiendo conceptos y técnicas nuevas, es imprescindible aplicarlos, ya que como mejor y más se aprende en ciencia de datos es con la experiencia. Es por esto que adjuntamos, en el siguiente apartado, enlaces a datasets que permitan “jugar” con los datos probando lo aprendido.

Además, para tener datos de uso público con los que se pueda entrenar y probar las técnicas aprendidas, recomendamos los siguientes enlaces:

- UCI Machine Learning Repository. [En línea] URL disponible en <http://archive.ics.uci.edu/ml/datasets.html>

The UCI Machine Learning Repository is a collection of databases, domain theories, and data generators that are used by the machine learning community for the empirical analysis of machine learning algorithms.

- Kaggle Datasets. [En línea] URL disponible en <https://www.kaggle.com/datasets>