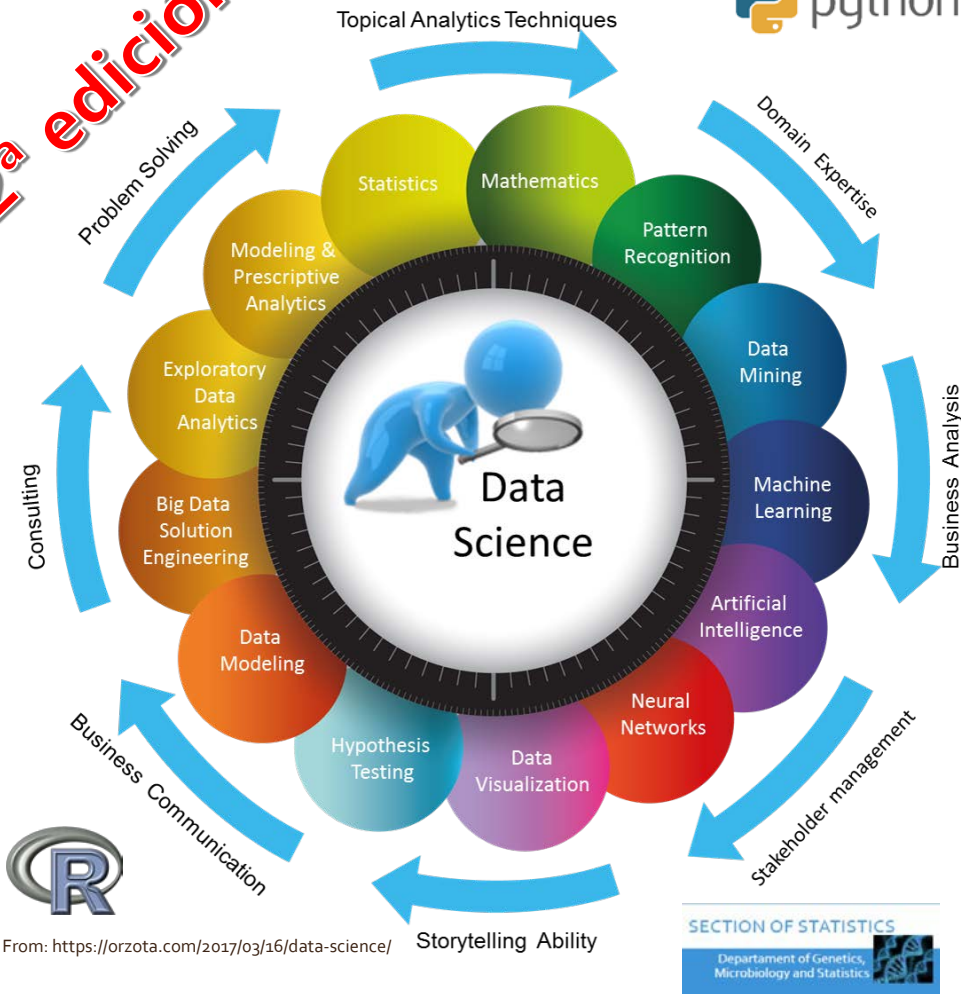


2ª edición



MAYO-JULIO 2019

# DATA SCIENCE (CIENCIA DE LOS DATOS): APLICACIONES A LA BIOLOGÍA Y A LA MEDICINA CON PYTHON Y R

## Data Scientist: The Sexiest Job of the 21st Century !

With biomedical research developing at a rapid pace, biologists and doctors have proved the necessity of adopting concepts and tools from other areas including machine learning, computational chemistry, engineering, mathematics, and physics. The shift of paradigm drives many initiatives in the health and data space, such as precision medicine, value-based healthcare and genomics (based on Ivan Zheng, Data Scientist: "a Profession for Quantitative Biologists")

Con el soporte de:



EXPERTO UNIVERSITARIO 150H



ORGANIZA BIOST3

<http://www.fbg.ub.edu/investigadors/grups-recerca-ub/estadistica-clinica-biodiversitat/>

COORDINACION: DR ANTONIO MONLEON (SECCION DE ESTADISTICA, FACULTAT DE BIOLOGIA, UB)

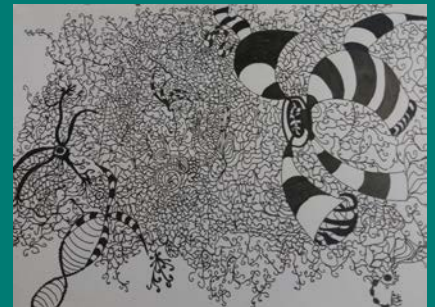


Imagen Sergi Gallardo Masip ("Red Neuronal Daliniana")

[www.ub.edu](http://www.ub.edu)  
[postgraus](http://postgraus)

FACULTAT DE BIOLOGIA – UNIVERSITAT DE BARCELONA (AVDA DIAGONAL 645, BARCELONA)

INFORMACIÓN: [amonleong@ub.edu](mailto:amonleong@ub.edu)  
[eballester@ub.edu](mailto:eballester@ub.edu) telf: 34.93.402.15.60

Mayo a Julio de 2019 (clases y seminarios de especialistas, orientación práctica y aplicada)

Programa y profesorado  
<http://biost3.blogspot.com.es/>

# **Programa del curso (general)**

## **1- Lenguajes de programación y datos**

- Sistema operativo Linux
- Administración en Linux
- Lenguaje Python
- Bases de datos relacionales y no relacionales
- Lenguaje R
- Servidores de datos y super-computación (HPC)
- Data management, Markdown, latex, notebooks
- Aplicaciones a medida web (Shiny)

## **2-Análisis estadístico con R**

- Estadística descriptiva y exploratoria
- Diseño de experimentos
- Análisis multivariante
- Análisis bayesiano

## **3-Machine Learning (ML)**

- Conceptos básicos de Machine-learning (ML)
- Aprendizaje no supervisado (Unsupervised learning)
  - o Clustering (k-means y hierarchical clustering)
  - o Reducción de la dimensión (Principal Component Analysis)
- Aprendizaje supervisado (Supervised Learning) / Clasificación
  - o Generative vs discriminative modelos
  - o Validación de modelos (bias variance trade-off and learning curves)
  - o Clasificadores (Naive Bayes, KNN, Linear SVM)
- Ensemble methods (Boosting, Bagging, y Random Forests)
- Neural nets y Deep learning
- Introduction to Feature Selection Methods: Sequential Searches and Genetic Algorithms

## **Seminarios (a confirmar)**

- Análisis de datos ómicos: metagenómica y transcriptómica (RNAseq). (A. Paytuví. Sequentia-Biotech)
- Arquitectura de datos: High-performance computing en las biociencias. (P. Puigdomènech, HPCNow!)
- Statistical techniques to assess association of novel genetics variants and complex diseases: CNVassoc R package ". (Isaac Subirana, IMIM)
- Computational Discovery of Volatile Biomarker Fingerprints: Opportunities and Risks (Santiago Marco. Facultat de Física UB)
- Deep Learning aplicado al campo biomédico (Petia Radeva, Dept. Mathematics and Computer Science. Facultad de Matemáticas e Informática. UB.)
- Análisis de datos ómicos (metabarcoding y metagenome skimming) en biomonitorización. Cesc Múrria. Secció d'Ecologia, Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals. Fac Biologia. UB
- Ciència ciutadana per a la salut pública (Mosquito Alert) i ambiental (Projecte Riu-Net)

**Trabajo final aplicado: resolución de problema real de ciencia de los datos en el campo biomédico**

## **PROFESORADO:**

- Antonio Monleón Getino (Coordinación). Biost3, GRBIO. Sección Estadística. Dep GME-UB
- Biel Stela Ballester. Bioinformático. Pharmacoinformatics Group. UPF
- Alexandre Perera i Lluna. UPC
- Pere Puigdomènech, HPCNow!
- Pere López. Biost3, Sección Estadística. Dep GME-UB
- Xavier de Pedro. Ajuntament de Barcelona.
- Carlos Crespo. Biost3, Sección Estadística. Dep GME-UB
- Andreu Paytuví. Bioinformático. Sequentia-Biotech
- David Jorquera. Informático.
- Cesc Múrria. Secció d'Ecologia, Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals. Fac Biologia. UB
- Santiago Marco, Departamento de Ingeniería Electrónica y Biomédica. Facultad de Física. UB
- Petia Radeva. Dept. Mathematics and Computer Science. Facultad de Matemáticas e Informática. UB.
- Frederic Bartumeus. CSIC
- Isaac Subirana. IMIM. Sección Estadística. Dep GME-UB
- Pau Fortuño. Secció d' Ecologia. Facultat de Biologia. UB.



## **INFORMACION INSTITUCIONAL, PRECIO Y MAS DETALLES EN:**

[https://www.ub.edu/web/ub/ca/estudis/oferta\\_formativa/masters\\_propis/fitxa/C/201811524/index.html](https://www.ub.edu/web/ub/ca/estudis/oferta_formativa/masters_propis/fitxa/C/201811524/index.html)

- Periodo lectivo: 7 Mayo a 16 julio. Horario de 16 a 20h (tardes) / Martes y jueves (seminarios dentro del periodo del curso, días a decidir)
- Es recomendable que los estudiantes traigan su propio portatil (laptop) a clase, también es recomendable conocimientos de estadística, matemáticas e informática.
- 3-5 descuentos especiales (10%) a miembros de BIB-Bioinformatics Barcelona (CONSULTRAR a eballester@ub.edu telf: 34.93.402.15.60, o amonleong@ub.edu)