

## **DATA SCIENCE, BIOSTATISTICS AND BIOINFORMATICS**

### **TECHNOLOGY DESCRIPTION**

The Data Science, Biostatistics and Bioinformatics service of Health Research Institute Hospital La Fe offers an expert approach to the data analysis problems that arise during research or technological development, as well as data processing and generation of models for decision making. Its main objective is to promote the application of the most appropriate techniques for each project, thus maximizing the performance and quality of the results and conclusions obtained.

The combination of professionals from different disciplines provides the service with a unique global perspective to perform the most complex experimental designs and analyzes.

### **SECTORS OF BUSINESS APPLICATION**

The objective of the Data Science, Biostatistics and Bioinformatics platform is to provide an expert support service in all phases of a research or technological development project. The functions of the platform range from the provision of support and methodological advice and collaboration in the development and design of studies and experiments, to the treatment, analysis and interpretation of experimental and / or observational data. The service also offers training and courses in data science, biostatistics and bioinformatics.

### **RESOURCES AVAILABLE**

The platform of Data Science, Biostatistics and Bioinformatics has access to the high performance computing cluster of IIS La Fe, which consists of 224 computer cores spread over 8 nodes and 4 TB of RAM.

### **TECHNICAL ADVANTAGES AND BUSINESS BENEFITS**

Available services:

- Design of studies
  - o Methodologic counseling
  - o Drafting of study protocols and statistical analysis plans (SAP)
  - o Optimum sample size
  - o Generation of randomization lists
- Data analysis
  - o Statistical modelling
    - Classic statistics
    - Generalized linear models
    - Generalized additive models
    - Survival analysis
    - Robust models
    - Multivariate analysis
    - Mixed models
    - Penalized models
    - Resampling methods
    - Bayesian inference
  - o Machine learning
    - Artificial neural networks (ANN)
    - Support Vector Machines
    - Random Forest
    - Boosting and bagging
    - Deep learning
  - o Unsupervised analysis
    - Exploratory analysis
    - Clustering techniques
    - Dimension reduction techniques (PCA, t-SNE)
- Courses
  - o Divulgative

## DATA SCIENCE, BIOSTATISTICS AND BIOINFORMATICS

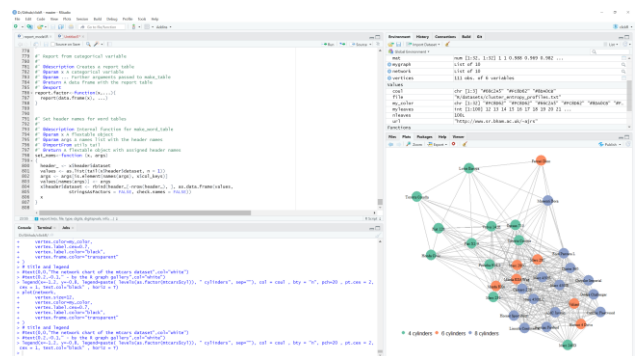
- Formative
- R programming
  - Development of customized scripts for specific analyses
  - Generation and customization of statistical graphs
  - Tasks automation (automatic reports generation, data cleaning, etc.)
- Omic data analysis
  - RNAseq
  - Expression arrays
  - Methylation arrays
  - Metabolomics
  - Etc.

## PARTNER SEARCH

The Data Science, Biostatistics and Bioinformatics Service maintains relationships and collaborations with several research groups from both the Universitat de València and the Universitat Politècnica de València, as well as with several companies in the pharmaceutical and healthcare sectors. As a result of these collaborations, the platform has participated in the development of two registered patents, based on the application of predictive models for the diagnosis of diseases, and is currently in the active development of three other possible patents.

We look for any research group or institution interested in maximizing the quantity and quality of the information obtained from their projects through data analysis, or that have specific needs for generating complex predictive models, application of decision-making systems, statistical validation of technological developments or methodological planning.

## IMÁGES



## CONTACT INFORMATION

Responsible technician/ Technical manager: David Hervás Marín

Tlf: 96 124 66 07

Email: [bioestadistica@iislafe.es](mailto:bioestadistica@iislafe.es)