



SGIker Prestakuntza Eskaintza (Ikerkuntzarako Zerbitzu Orokorrak)

Ikastaroaren izenburua:

Datu-bolumen handien aldagai anitzeko tratamendu estatistikoa.

Data	2022ko irailaren 12tik 16ra
Iraupena	25 ordu
Ordutegia	9:00etatik- 14:00etara
Tokia	Zientzia eta Teknologia Fakultatea Bizkaiko Campusa (Leioa) Euskal Herriko Unibertsitatea UPV/EHU

Hizlariak eta irakasleak:

Luís Bartolomé Moro doktorea eta Laura Ortíz-Herrero lizentziaduna.

Ikastaroaren helburuak:

1. Ikasleei datuen tratamenduari eta haren aplikazioei buruzko hastapenak ematea.
2. Osagai nagusien bidezko analisiaren bitartez (PCA) datuen tratamenduan sakontzea.
3. HCA/PCA sailkatze-sistema ezagutzea.
4. Ikastaroko parte-hartzaileek entregatutako datuen tratamendua, adibide gisa (alde praktikoa).
5. Kalibrazioak partziala karratu txikienez (PLS) egitea. Adibide praktikoa batzuk.

Ikastaroaren edukiak:

1. Osagai nagusien bidezko analisisa (PCA). Sarrera
 - a. Datuak matrize gisa adieraztea.
 - b. Osagai nagusien modeloak. Puntuazioak eta kargak.
 - c. PCA baten helburuak.
 - d. PCA bat interpretatzea.
 - e. PCA bidezko sailkapena eta klusterraren analisisa (HCA).
2. Eskalatzea eta normalizazioa.
3. Erlazio faltsuak nola tratatu behar diren.
4. Parte-hartzaileek beraiek jarritako adibide praktikoa.
5. Aldagai anitzeko eskualdeak (PLS/OPLS).
 - a. Erabilera eta adibide praktikoa
 - b. Aldaera anitzeko erregresioen benetako lanen simulazioa eta erakusketa.
6. Adibide praktikoa.



Parte hartzailearen profila:

Aldagai anitzeko tratamendu estatistikoan interesa duten ikertzaile eta profesionalak: biologoak, ekologoak, kimikariak, ingurumen-zientzialariak, botanikariak... ulbisio-tratamenduan esperientziarik ez izatea edo esperientzia gutxi izatea.

Parte hartzaile kopurua (gutxienekoa/gehienezkoa):

5/8

Kontaktua

- Luís Bartolomé Moro doktorea
- Analisirako Zerbitzu Zentrala
- Zientzia eta Teknologia Fakultatea, Euskal Herriko Unibertsitatea
- Sarriena Auzoa, z/g, Leioa, 48940
- Tfno.: 94 601 3473
- luis.bartolome@ehu.eus

ZERBITZUA:

[Analisirako Zerbitzu Zentrala](#)

Prezioa

- UPV/EHUkoek: 150 €
- IEPkoek: 250 €
- Kanpokoek: 500 €

Informazio gehigarria:

- Ikastaro interesgarria da datu-bolumen handia tratatzen **hasten diren** ikertzaileentzat
- **Ikasturtean zehar, Simca softwarearen "demoa" izango da, hilabete bateko erabilgarritasunarekin.** Softwarea ez da konpatiblea MAC ordanagailuekin. UPV/EHUko ikertzaileei ordenagailu korporatiboan soilik instalatuko zaie softwarea.
- Ikastaroa gaztelaniaz ematen da.
- Praktiak egiteko datuen segurtasun neurriak jarraituko dira.
- Ikastaroan bertaratze-ziurtagiria emango da. Ikastaroan parte hartu dutenek ziurtagiri hori jasoko dute, baldin eta ikastaroaren ordu kopuru guztiaren % 80 bete badute.



Oferta formativa SGIker (Servicios Generales de Investigación)

Título del curso:

Tratamiento estadístico multivariante de gran cantidad de datos.

Fechas	Del 12 al 16 de septiembre
Duración	25 horas
Horario	9:00-14:00
Lugar	Plataforma Tecnológica Martina Casiano y/o Facultad de Ciencia y Tecnología Campus de Bizkaia (Leioa) Universidad del País Vasco UPV/EHU

Ponentes y formadores:

Dr. Luís Bartolomé Moro y Lda. Laura Ortíz-Herrero.

Objetivos que se pretenden alcanzar en el curso:

1. Introducir a los alumnos en el tratamiento de datos y sus aplicaciones.
2. Profundización en el tratamiento de datos por medio de Análisis de Componentes Principales (PCA).
3. Conocer el Sistema de clasificación de datos HCA/PCA.
4. Tratamiento de datos entregados por los propios participantes del curso como ejemplo (parte práctica).
5. Realización de calibraciones por mínimos cuadrados parciales (PLS). Aplicación a algunos ejemplos prácticos.

Contenidos que se van a trabajar durante el curso:

1. Análisis de Componentes Principales (PCA). Introducción.
 - a. Representación de datos como una matriz.
 - b. Modelos de componentes principales. Scores and Loadings.
 - c. Objetivos de un PCA.
 - d. Interpretar un PCA.
 - e. Clasificación con PCA y análisis de cluster (HCA).
2. Escalado y normalización.
3. Cómo tratar espurios.
4. Ejemplos prácticos traídos por los propios participantes
5. Regresiones multivariantes (PLS/OPLS).
 - a. Usos y ejemplos prácticos
 - b. Simulación y exposición de trabajos reales de regresiones multivariantes.



6. Ejemplos prácticos traídos por el propio alumnado.

Perfil del participante:

Investigadores y profesionales interesados en el tratamiento estadístico multivariante: biólogos, ecólogos, químicos, ambientalistas, botánicos... con poca o nula experiencia en la materia.

Número de participantes (mínimo/máximo):

5/8

Datos de contacto

- Dr. Luís Bartolomé Moro
- Servicio Central de Análisis
- Facultad de Ciencia y Tecnología
Universidad del País Vasco / Euskal Herriko Unibertsitatea
- Bº. Sarriena s/n, Leioa, 48940
- Tfno.: 94 601 3473
- luis.bartolome@ehu.eus

SERVICIO:

[Servicio Central de Análisis.](#)

Precio

- Usuarios de la UPV/EHU: 150€.
- Usuarios de Organismos Públicos de Investigación: 250€
- Usuarios externos: 500 €

Otra información adicional:

- Curso de gran interés para investigadores que se inicien en el manejo de gran cantidad de datos en sus trabajos.
- **Durante el curso se dispondrá de la "demo" del software SIMCA con disponibilidad limitada a un mes.** El software no es compatible con ordenadores MAC. Al personal investigador de la UPV/EHU se le instalará el software únicamente en el ordenador corporativo.
- El curso se imparte en castellano.
- Se seguirán en todo momento las medidas de seguridad necesarias para el tratamiento de datos ajenos.
- Se entregará certificado de asistencia. Las personas asistentes al curso recibirán dicho certificado siempre y cuando hayan completado al menos el 80% de la duración total del mismo.



SGIker training courses offered (Advanced Research Facilities)

Course title:

Multivariate statistical processing of large amounts of data.

Dates	2022, from September 12 to 16
Duration	25 hours
	9:00-14:00
Course venue	Faculty of Science and Technology Bizkaia Campus (Leioa) University of the Basque Country, UPV/EHU

Speakers and trainers:

Phd. Luís Bartolomé Moro and Ld. Laura Ortíz-Herrero.

Objectives to be fulfilled during the course:

1. To provide the participants with an introduction to the processing of data and its applications.
2. To deal in depth with data processing through Principal Component Analysis (PCA).
3. To get to know the HC/PCA data classification System.
4. Processing of data provided by the course participants themselves as an example (practical part).
5. Calibrations by conducting partial least squares (PLS). Application to some practical examples.

Content that is going to be worked on during the course:

1. Principal Component Analysis (PCA). Introduction.
 - a. Representing data as a matrix.
 - b. Principal component models. Scores and Loadings.
 - c. Objectives of a PCA.
 - d. Interpreting a PCA.
 - e. Classification with PCA and cluster analysis (HCA).
2. Scaling and normalisation.
3. How to deal with outliers.
4. Practical examples provided by the participants themselves.
5. Multivariate regressions (PLS/OPLS).
 - a. Uses and practical examples
 - b. Simulation and exposure of real works of multivariate regressions.
6. Practical examples brought by the students themselves.



Participant profile:

Researchers and professionals interested in multivariate statistical processing: biologists, ecologists, chemists, environmentalists, botanists... with no or little experience in multivariate treatment.

Number of participants (minimum/maximum):

5/8 (The capacity will be adapted to the current health regulations at the time of delivery)

Contact

- Dr. Luís Bartolomé Moro
- Central Analysis Service
- Faculty of Science and Technology,
University of the Basque Country
- Bº. Sarriena s/n, Leioa, 48940
- Phone: 94 601 3473
- luis.bartolome@ehu.eus

SERVICE:

[Central Analysis Service.](#)

Course fee

- UPV/EHU users: 150 €
- PRB users: 250 €
- External users: 500 €

Other additional information:

- A course of great interest for researchers who are beginning in the processing of large volumes of data.
- **During the course, the SIMCA software "demo" will be available with availability limited to one month.** This software is incompatible with MAC computers. The UPV / EHU research staff will install the software only on the corporate computer.
- The course is taught in Spanish.
- The necessary security for data will be followed at all times.
- A certificate of attendance will be provided. Those attending courses will receive this certificate provided that they have completed at least 80% of the total duration of the course.