

## U1173 ANÀLISI DE DADES COMPOSICIONALS EN CIÈNCIES SOCIALS

### PROFESSORAT

**Maribel Ortego**, Universitat Politècnica de Catalunya

**Berta Ferrer-Rosell**, Universitat de Lleida

### OBJECTIUS

L'anàlisi de dades composicional ha començat a ser usada en varis camps de les ciències socials, que sovint afronten preguntes de recerca similars. Per exemple en marketing, les preguntes de recerca típiques estan relacionades o bé amb la distribució d'un tot (e.g., quotes o assignacions), o amb la importància relativa (e.g., dominància, perfil, prevalença). En aquesta línia de recerca, Morais et al. (2018) estudia els determinants de la quota de mercat en la indústria automobilística, Ferrer-Rosell & Coenders (2018) segmenten els turistes en base a com distribueixen el pressupost del viatge, i Blasco-Duatis et al. (2019) estudien la importància relativa dels continguts en comunicació en línia a través de Twitter.

La motivació d'aquest curs és augmentar la consciència de les peculiaritats de les dades composicionals (CoDa) i ajudar els investigadors a evitar els problemes més habituals en l'anàlisi de CoDa. El tema és oportú, donada la gran presència de conjunts de dades composicionals en tots els àmbits de les ciències socials, des de marketing i economia, passant per urbanisme o turisme

El curs busca que els participants siguin capaços de reconèixer bases de dades que són de naturalesa composicional i que puguin explorar dades composicionals tant numèricament com gràficament. També es treballaran eines de representació de dades en coordenades log-ratio així com utilitzar aquestes coordenades en anàlisis multivariades posteriors. Finalment, el curs també busca tractar les dades irregulars en dades composicionals i introduir als participants al software especialitzat CoDaPack.

### CONTINGUTS

1. Introducció: definició de composició, preguntes de recerca composicionals i transformacions (logaritmes de ràtios)
2. CoDa a la pràctica: anàlisi exploratòria (centre, matriu de variació, biplot)
3. Balanços i CoDa dendrograma
4. Processament de dades: tractament dels zeros i altres irregularitats en les dades
5. Cluster composicional i regressió amb variable predita composicional / predictor composicional
6. Introducció al software CoDaPack

## **METODOLOGIA**

El curs proporcionarà una introducció als aspectes teòrics i, sobretot, pràctics de l'anàlisi estadística de CoDa, i s'enfocarà sobretot a l'àmbit de l'anàlisi turística.

El curs constarà d'una part de conferència (introducció de contingut de caire teòric) i una altra part pràctica. La sessió pràctica es farà amb el programa gratuït CoDaPack (<http://ima.udg.edu/codapack/>). Es demanarà a tots els participants que se l'instal·lin abans de la segona sessió.

**Durada:** 8 hores

**Lloc:** No presencial a través de l'eina de videoconferència del campus virtual

**Data:** 18 i 24 de novembre de 2020

**Horari:** De 9:00 h a 13:00 h

**Inscripció:** Telemàticament omplint el formulari que apareix en aquest enllaç:  
<http://www.formacioprofessorat.udl.cat/upu/activitats.php>

**Observacions:**

Aquesta activitat no es durà a terme si no hi ha un mínim de 15 persones inscrites.

Nombre màxim d'inscrits: 20 persones.