

Datu-iturri berria ADR kalkulatzeko (okupatutako gela bakoitzeko eguneko batez besteko tarifa)

Sarrera

Eustat, *Big Data* terminoaren barruan bildutako hainbat teknika erabiliz informazioa lortzeko aukera handiak daudela jakinik, hainbat proiektutan ari da lanean, teknika horien bidez lortutako datuak bere estimazio-prozesuetan sartzeko. Informazioa biltzeko prozedurak aztertzen ari dira, bai eta horiek estatistika ofizialean duten erabilera ere, bai datu-iturri bakar gisa, bai lagin-eta/edo administrazio-jatorriko datuekin konbinatuta.

Testuinguru horretan, 2017an proiektu bat abiarazi zen, Establezimendu turistiko hartzaileen inkestako gela okupatu bakoitzeko eguneko batez besteko tarifa (ADR) lortzeko, *Big Data* metodologia erabiliz.

2020ko urritik aurrera, aldagai horri (ADR) buruz eta *Gela erabilgarriko diru-sarrerak* aldagai deribatuari (RevPar) buruz zabaldutako datuen jatorria da hotel-establezimenduetako gelak *online* erreserbatzeko plataformen bidez Interneten eskuragarri dagoen informazioa atzematea.

Webeko informazioa atzemateko (*web scraping*) Eurostatek lege-bermeei, lehiari, datuen babesari eta sekretu estatistikoari buruz emandako gomendioei jarraitzen zaie.

ADR modelizatzea

Euskal AEko hotel-establezimenduen ADR-ren azken estimazioa egiteko, estatistika-eredu hau definitu da, online erreserba-plataformetatik lortutako informazioa (*web scraping* teknikak erabiliz) eta inkestaren direktorioan dagoena konbinatzen dituen, bai eta establezimenduetatik bertatik inkesta bidez lortutakoa ere:

$$ADR = \beta_0 + \beta_1 \text{Mediana} + \beta_2 \text{Max} + \beta_3 \text{Min} + \beta_4 \text{Kategoria} \\ + \beta_5 \text{Estratua} + \beta_6 \text{ErreserbaOnline} + \beta_7 \text{Okupazioa} + \beta_8 \text{Tamaina} + \varepsilon$$

non azalpen-aldagaiak hauek diren:

- **Mediana, Max** eta **Min** dira azken 4 hilabeteetan Interneten harrapatutako eguneko prezioen mediana, maximoa eta minimoa
- **Kategoria:** 5 maila hartzen dira kontuan: 4 eta 5 izar, 3 izar, 2 izar, izar 1, ostatuak
- **Estratua:** establezimendua kokatuta dagoen eremu geografikoa
- **ErreserbaOnline:** aldagai dikotomiko horrek adierazten du establezimendua gela online erreserba-plataformetan dagoen ala ez
- **Okupazioa:** establezimenduko gela arabera okupazio-tasa erreferentziako hilabetearen
- **Tamaina:** tamaina-blokea establezimenduko gela-kopuruaren arabera (< 9, 9-15, 16-39, > 40 gela)

Online erreserba-plataformetan lortzen da informazioa hotel-establezimendu bakoitzerako, *web scraping* bidez, egunean behin, aurreko 120 egunetako bakoitzerako. Beraz, establezimendu eta egun bakoitzerako 120 prezio daude erreserba-plataforma bakoitzeko. Beraz, 31 eguneko hilabete batean, guztira 3.720 prezio desberdin lortuko lirateke establezimendu eta plataforma bakoitzerako.

Bi urteko epean, inkestatutako informaziotik lortutako ADR eta zenbatetsitako ADR alderatu dira, datuak atzitzeko eredu eta prozedura desberdinak erabiliz, eta gorago proposatutako eredua da egokiena.

Informazioaren atzematea eta zenbatespen-prozedura Eustaten Establezimendu turistikoa hartzaileen inkestaren ekoizpen-prozesuetan integratuta daude.