



Laboratoris Anàlisi Multivariant

Grau interuniversitari d'Estadística

Facultat de Matemàtiques i Estadística

B. Kostov^(1,2) B. Sevilla^(1, 3) K. Gibert^(1,3,4)

(1) Department d'Estadística i Investigació Operativa

(2) Institut d'Investigació Biomèdica August Pi i Sunyer, IDIBABS

(3) Knowledge Engineering and Machine Learning group at Intelligent Data Science and Artificial Intelligence Research Center

(4) Institute for Cience and Technology of Sustainability

*belchin.adriyanov.kostov@upc.edu, bea.sevilla@gmail.com,
karina.gibert@upc.edu*

Universitat Politècnica de Catalunya, Barcelona

K. Gibert



Fases de la pràctica

- Formar el grup de pràctiques (6 persones)
- Triar tema i localitzar base de dades
- Elaborar pla de treball
 - Calendari
 - Repartiment de tasques
 - Anàlisi de riscos previs
- Desenvolupament
 - Depuració i descriptiva
 - Clustering
 - ACP
 - ACM
 - AFM o Anàlisi Discriminant o Anàlisi Textual
- Conclusions
- Anàlisi comparativa
- Anàlisi crítica del pla de treball, gestió d'incidències
- Preparar presentacions MINUCIOSAMENT

D1: PreEntrega
21feb2019

D2: Entrega
28 febrer 2019

D3: Primera
Presentació
21 març 2019

Redacció document
PARAL.LELA

D4: Entrega
Final
30 maig 2019

Competències transversals que s'avaluen al treball pràctic

- Organització del treball en grup
- Planificació i gestió d'incidències
- Raonament
- Síntesi i claredat en informe escrit
- Capacitat de comunicació (presentació oral)
- Integració de coneixements (anàlisi comparativa)
- Desenvolupament al llarg del curs en sessions de seguiment

Formació dels grups de pràctiques

- Es repartirà l'aula entre 8 grans projectes
- Presentar la proposta de grup data D1
- Els professors refinaran la proposta de tots els grups amb criteris globals

Assumir la importància de saber treballar bé amb l'equip professional que t'integren, al marge d'aspectes personals.

Elecció del tema de la pràctica

- Determinar un (o uns pocs) temes d'interès general del grup (evitar temes repetits en cursos anteriors, see Campus Virtual)
- Buscar unes dades que aportin informació sobre aquella temàtica
 - Es suggereix baixar-se dades obertes disponibles a Internet
 - Pot ajudar el llistat de webs open source que hi ha al Campus Virtual
 - Assegurar-se que les dades que es baixen tinguin formats directament importables a R

Característiques de la Base de Dades

- Assegurar que els formats son compatibles amb R
- Estructura de BD rectangular
 - Individus en files
 - Variables en columnes
- BD amb 2000-5000 files
 - si en te més feu una selecció per un perfil d'interès o random
- Almenys 7 variables numèriques
- Almenys 7 variables categòriques
- Proposar la BD als professors per la seva aprovació data D1

Elaboració d'un pla de treball (data D2)

- Identificar grans tasques (inclou documentació, etc)
- Seqüenciació temporal (meditar amb cura les precedències entre tasques, guieu-vos amb el temari del curs)
- Assignació de recursos humans a cada tasca (qui fa què i quan)
(Assegurar la viabilitat, és a dir, que tothom tingui càrrega equilibrada i sense puntes de dedicació que puguin col.lapsar al grup).
- Reservar un marge de temps abans de l'entrega per incidències imprevistes. Planifiquen per acabar ABANS
- Reserveu temps per preparar les presentacions també
- Format del pla de treball:
 - Diagrama de Gant + taula d'assignació + anàlisi de riscos
(veure recursos pel treball en equip)

Materials

Campus virtual (<https://campusvirtual2.ub.edu/>)

- Apunts
- Scripts
- Bibliografia
- OpenData Links
- Llista de pràctiques de cursos anteriors