

## **Estructura del material a entregar**

**Baixeu-vos les dades que us indicarem a l'aula i realitzeu la pràctica**

### **Portada**

1. Títol
2. Nom complert
3. Data d'entrega

### **Definició del projecte i assignació.**

1. Anàlisi descriptiva univariant inicial de totes les variables. (inclou resum numèric i gràfic i comentaris corresponents)
  - a. Per les variables numèriques: summary, histograma i/o boxplot (eventualment altres representacions adients)
  - b. Per les variables categòriques: Table, barplot i/o pie (eventualment altres representacions adients)
2. Descripció detallada del procés de preprocessament de dades seguit i justificació de totes les decisions preses (inclou tractament de missings, etc)
3. Anàlisi descriptiva univariant de les dades preprocessades i discussió sobre l'aleatorietat de les dades mancants si n'hi ha.
4. Clustering jeràrquic amb el Mètode de Ward i distància de Gower al quadrat per totes les variables. Mostrar l'arbre jeràrquic resultant i seleccionar el número de classes.
5. Profiling dels clusters: Utilitzar la variable de cluster per distribucions condicionades i estadístiques descriptives per grups. Eventualment visualitzar la relació entre parells de variables contínues i les classes. Determinar les particularitats de cada classe i elaborar una descripció prototípica de cada classe. Si les dades són georeferenciades presentar visualitzacions en mapes de les classes.
6. ACP de les variables numèriques: Resultats de l'anàlisi, screeplot i selecció de components factorials retingudes. Pel primer mapa factorial: projecció dels individus; projecció comú de variables numèriques i modalitats de les variables qualitatives. Interpretació de les relacions i oposicions entre variables i interpretació dels eixos.
7. ACM de les variables qualitatives: Resultats de l'anàlisi, scree plot i selecció de components factorials a retenir. Per al primer mapa factorial: projecció de modalitats de les variables qualitatives (eventualment amb les il·lustratives qualitatives que es vulguin afegir) Interpretació de relacions i oposicions entre variables. Anàlisi sobre les coincidències i discrepàncies amb l'ACP
8. Muntar una nova matriu de dades amb tots els factors retinguts de l'ACP i de l'ACM. Clustering jeràrquic sobre aquesta nova matriu. Criteri de Ward. Aquest cop amb mètrica euclidiana. Mostrar l'arbre jeràrquic resultant i seleccionar el número de classes.
9. Taula creuada de les dues classificacions i discussió sobre l'estabilitat de les classes identificades en un cas i altre
10. Profiling dels clusters: Repetir el mateix procés d'interpretació de les classes per aquesta segona classificació (descriu al punt 4) i comparar les dues classificacions. Representar la visualització de les classes sobre el primer pla factorial de l'ACP
11. Afegir una representació de la classificació obtinguda al punt 4 sobre els primers plans factorials i discutir l'aportació
12. Alguna de les dues classificacions retroba resultats obtinguts a l'ACP o a l'ACM? Discutir
13. Entregueu l'informe i els Scripts d'R utilitzats per elaborar les anàlisis (opcionalment es poden incrustar les parts rellevants dels scripts al llarg dels capítols anteriors).