

PAC 1: Mostreig i Estadística Descriptiva

Format: El nom i cognoms de l'estudiant s'han d'indicar clarament a la primera pàgina de l'informe. Els exercicis 1 i 2 s'ha de resoldre usant R, és important mostrar cada un dels passos realitzats per arribar al resultat final.

Indiqueu en cada cas quines instruccions heu fet servir i les sortides que s'obtenen (amb copiar-enganxar directament al fitxer de resposta).

L'informe final es lliurarà en un fitxer amb format "pdf" o "html".

EXERCICI 1 (30 punts)

Al fitxer 'nhtemp.rda' hi trobem els mesuraments de la temperatura en graus Fahrenheit a New Haven, Connecticut, entre els anys 1912-1971. Feu ?nhtemp si voleu més informació sobre aquestes dades.

Podeu carregar les dades utilitzant una de les dues instruccions:

```
data(nhtemp)
```

```
load("nhtemp.rda")
```

- Calculeu els resums numèrics d'aquestes dades.
- Feu un histograma per explicar la variable.
- A partir dels resum numèrics i de l'histograma descriuiu com són les dades.
- Feu la comanda `plot(nhtemp)`. Expliqueu el gràfic que surt. Comenteu si és un gràfic útil per aquestes dades

EXERCICI 2 (40 punts)

Al fitxer PIS_MUN.csv hi trobareu informació sobre el nombre de contractes de lloguer i el lloguer mitjà contractual (mitjana dels preus dels lloguers per contracte) als municipis de Catalunya i als districtes de la ciutat de Barcelona. Concretament trobareu les variables.

COD = codi postal del municipi (o "08019" en cas dels districtes de Barcelona),

MUNDIS = Nom del municipi o districte de la ciutat de Barcelona,

MUN = Nom del municipi o "Barcelona" en el cas dels seus districtes,

PROV = Província la la que pertany el municipi/districte.

NAny = Nombre de contractes signats el corresponent any al municipi/districte (on Any pot ser 2019, 2017, 2016 o 2015),

MAny = Lloguer mitjà mensual dels contractes signats el corresponent any al municipi/districte.

Les dades han estat extretes de la web de la Generalitat de Catalunya, concretament de (http://habitatge.gencat.cat/ca/dades/estadistiques_publicacions/indicadors_estadistiques/estadistiques_de_construccio_i_mercat_immobiliari/mercat_de_lloguer/), i contenen “l’exploració estadística de les fiances de lloguer de l’INCASÒL relacionada amb el nombre de lloguers constituïts i la renda mitjana mensual a Catalunya.” De les taules “Anuals per municipis” s’ha eliminat el registre de la ciutat de Barcelona i s’ha substituït per les corresponents dades dels districtes.

Per importar les dades us pot ser útil la instrucció (adaptant el nom del fitxer si cal)

```
dades <- read.csv("PIS_MUN.csv", sep = ";") o bé
```

```
library(rio)
```

```
dades <- import("PIS_MUN.csv")
```

Es demana

- Trobeu els resums numèrics (mitjana, mediana, quartils, desviació típica mostral) del nombre de contractes signats l’any 2019 i comenteu els resultats.
- Obteniu una taula de quants municipis de cada província hi ha a la base de dades. Feu un gràfic adient per mostrar la distribució. Comenteu els resultats.
- Feu un boxplot de la mitjana dels preus mitjans dels contractes signats l’any 2019 distingint per províncies i comenteu quines diferències s’hi aprecien.

EXERCICI 3 (30 punts)

Imagineu que volem estudiar a qui votaran els estudiants de la UOC a les properes eleccions. Podem fer-ho de diverses maneres:

- Enviant un missatge a través del campus a tots els estudiants i analitzant les respostes.
- Seleccionem 9 comarques a l’atzar i seleccionem a l’atzar 50 estudiants d’aquestes comarques als quals truquem per telèfon.
- Assignem un número a cada estudiant i truquem als que el seu número és de la forma 12324*n fins que obtinguem 450 respostes.

Digueu quin tipus de mostreig fem servir en cada cas i quins són els seus inconvenients.

Quin mostreig dels proposats escolliríeu i per què?