

Analiza datelor în R

Laborator C5

1. Setul de date *cabinet* din pachetul *UsingR* conține sumele economisite la plata taxelor de către membrii cabinetului George W. Bush în urma adoptării unei legi de scutire de taxe.
 - a) Să se ordoneze setul de date descrescător după economii.
 - b) Să se determine media și mediana sumelor economisite. Comentarii.
2. Setul de date *npdb* din pachetul *UsingR* conține informații din National Practitioner Data Bank cu privire la despăgubirile plătite de doctori pentru malpraxis în America în perioada 2000-2003.
 - a) Să se determine câte despăgubiri s-au plătit în fiecare an. Să se construiască diagrama pie-chart.
 - b) Să se determine câte despăgubiri s-au plătit în fiecare stat. Să se construiască diagrama barplot.
 - c) Să se determine care sunt statele și care sunt medicii cu cele mai multe despăgubiri acordate pentru malpraxis.
 - d) Să se determine care sunt cele mai mari 10 despăgubiri acordate.
 - e) Să se construiască histograma sumelor plătite ca despăgubire. Comentarii.
3. Setul de date *Cars93* din pachetul *MASS* conține informații cu privire la 93 modele de automobile.
 - a) Să se stabilească de ce tip sunt variabilele *Manufacturer*, *Type*, *AirBags* și *Origin* și care este cea mai adecvată metodă de sumarizare a valorilor în fiecare caz.
 - b) Să se determine pentru fiecare dintre variabilele *Price*, *Horsepower* și *Weight* media, mediana, quartilele, deviația standard și coeficientul de asimetrie. Comentarii.
 - c) Să se construiască pentru fiecare dintre variabilele *Price*, *Horsepower* și *Weight* histograma, curba de densitate și diagrama boxplot. Să se analizeze existența valorilor atipice.
 - d) Să se studieze dacă variabila *Luggage.room* are sau nu distribuție aproximativ normală.