Jeu de données N°1 :

Sales of Child Car Seats

On cherche à prédire la variable « High » à l’aide des autres variables.

**Description**

A simulated data set containing sales of child car seats at 400 different stores.

**Format**

A data frame with 400 observations on the following 11 variables.

High

Unit sales (in thousands) at each location. “Yes” (unit sales > 8). “No” (<= 8).

CompPrice

Price charged by competitor at each location

Income

Community income level (in thousands of dollars)

Advertising

Local advertising budget for company at each location (in thousands of dollars)

Population

Population size in region (in thousands)

Price

Price company charges for car seats at each site

ShelveLoc

A factor with levels Bad, Good and Medium indicating the quality of the shelving location for the car seats at each site

Age

Average age of the local population

Education

Education level at each location

Urban

A factor with levels No and Yes to indicate whether the store is in an urban or rural location

US

A factor with levels No and Yes to indicate whether the store is in the US or not

Jeu de données N°2 :

Ozone\_complet.txt

On considère le jeu de données ozone\_complet.txt où on cherche à expliquer la concentration maximale en ozone relevée sur une journée (variable maxO3) par d’autres variables essentiellement météorologiques.