



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Superior d'Enginyeries Industrial,
Aeroespacial i Audiovisual de Terrassa

MUEI, MUEA, and MASE

Advanced Engineering Data Analysis (AEDA) CART & RF

Lab Practice

Prof. Daniel Fernández

daniel.fernandez.martinez@upc.edu

Data on Children who have had Corrective Spinal Surgery

Description: The data 81 rows (children) and 4 variables:

Kyphosis: a factor with levels absent present indicating if a kyphosis (a type of deformation) was present after the operation,

Age (in months)

Number: the number of vertebrae involved

Start: the number of the first (topmost) vertebra operated on.

Location:

```
library(gam)
data(kyphosis)
```

```
head(kyphosis)
  Kyphosis Age Number Start
1  absent  71      3     5
2  absent 158      3    14
3  present 128      4     5
4  absent   2      5     1
5  absent   1      4    15
6  absent   1      2    16
```

Activity: Apply a pre-processing of the data and apply CART and RF.

Solution proposal: Kyphosis_solution1.pdf & Kyphosis_solution2.pdf

Baseball data set

Description: Major League Baseball Data from the 1986 and 1987 seasons.

A data frame with 322 observations of major league players on 20 variables of their performance (e.g. number of hits, number of runs,...).

More information here:

<https://rdrr.io/cran/ISLR/man/Hitters.html>

Location:

```
library(ISLR)
```

```
data(Hitters)
```

Activity: Apply a pre-processing of the data and Regression Tree using Salary as a response variable and Years and Hits as the descriptive variables

Solution proposal: Baseball_solution.pdf

Titanic

Description: Titanic data set with 1046 passengers and 6 variables

Titanic Data

Location:

```
library(rpart)  
data(ptitanic)
```

The data has 1046 observations on 6 variables.

- `pclass`: passenger class
- `survived`: died or survived
- `sex`: male or female
- `age`: age in years
- `sibsp`: number of siblings or spouses aboard
- `parch`: number of parents or children aboard

Activity: Apply a pre-processing of the data and Classification Tree. Check the solution for practicing the pruning.

Solution proposal: Titanic_solution.pdf