

# Topic 1 - Exercise 1 - Multivariate Analysis

The file *Colleges.csv* contains a data set corresponding to information about 777 colleges in USA in 1995. The data set contains 18 variables:

1. *Private*, which is a categorical variable with levels *No* and *Yes* indicating private or public university.
2. *Apps*, which is the number of applications received.
3. *Accept*, which is the number of applications accepted.
4. *Enroll*, which is the number of new students enrolled.
5. *Top10perc*, which is the percentage of new students from top 10% of high school class.
6. *Top25perc*, which is the percentage of new students from top 25% of high school class.
7. *F.Undergrad*, which is the number of fulltime undergraduates.
8. *P.Undergrad*, which is the number of parttime undergraduates.
9. *Outstate*, which is the out-of-state tuition.
10. *Room.Board*, which is the room and board costs.
11. *Books*, which is the estimated book costs.
12. *Personal*, which is the estimated personal spending.
13. *PhD*, which is the percentage of faculty with Ph.D.'s.
14. *Terminal*, which is the percentage of faculty with terminal degree.
15. *S.F.Ratio*, which is the student/faculty ratio.
16. *Perc.alumni*, which is the percentage of alumni who donate.

17. *Expend*, which is the instructional expenditure per student.
18. *Grad.Rate*, which is the graduation rate.

Place the file that can be found in Aula Global in a given working directory, set the working directory and then load the data set using:

```
> Colleges <- read.csv("Colleges.csv")  
> X <- Colleges[, 2 : 19]  
> rownames(X) <- Colleges[, 1]
```

The exercise consists of the following points:

1. Identify the type of all variables.
2. Perform a visual analysis of each of the quantitative variables. Then, Perform a visual analysis of each of the quantitative variables taking into account the variable *Private*. Describe all the possible conclusions from the visual analysis that you can think of.
3. Perform a visual analysis of all quantitative variables together. Then, Perform a visual analysis of each of the quantitative variables taking into account the variable *Private*. Describe all the possible conclusions from the visual analysis that you can think of.