Dataset: voir lasvegas.csv

2.2 Descriptive Statistics

In this part, you are going to make an exploratory analysis of the data set. Calculate descriptive statistics on all the variables. The purpose is for you to get familiarized with this dataset.

>> description des variables input

2.3 Supervised Learning

In this part, you are going to use at least three supervised learning methods studied in the course. You are going to perform Classification and Regression. The target variable is the Hotel Score and the predictors will be the other variables or a subset of them that you consider pertinent for this analysis.

- For regression, you can consider the "Hotel Score" as a continuous variable
- For classification, you will create two classes: *High* (if the score is 4 and 5) and *Low* (otherwise).

Do not forget to evaluate your model performance. For example, by calculating the confusion matrix, accuracy, curve ROC, etc.

>> u**tilisation de trois méthodes supervisées:**

- **régression linéaire: considérer "hotel score" comme variable continue**
- **classification: (à confirmer pour le choix de méthode)**
  - **régression logistique**
  - **LDA**

2.4 Unsupervised Learning

In this part, you are going to apply all the unsupervised learning methods that you learned during the lectures. This time, the variable *Score* will be part of your analysis.

- **clustering**
- **PCA**

ATTENTION: bien expliquer chaque screenshot R

IMPORTANT: Your creativity, your contributions, results and comments will be taken into account for the final mark. Any figure, output code, table or other result with no interpretation will not be considered for the mark.