exploratory_data_analysis

Linh Vu & Duc Nguyen

2024-04-18

Before you build any models, you should perform appropriate exploratory (or descriptive) analysis. This might include:

Data wrangling, including joining two or more data sets into a single set, converting quantitative variables to categorical or collapsing categorical variables to ones with fewer levels, renaming variables and/or variable levels, creating new variables from existing ones

Descriptive statistics for all variables you intend to investigate. For quantitative variables, this includes: mean, standard deviation, 5-number summaries, and histograms and/or boxplots; and for categorical variables, this includes: a list of all factor levels, as well as counts and proportions within each level, and bar charts.

Exploration of the relationships between variables, both numerically and graphically. Consider not only the relationship between the response and explanatory variables, but also between two (or more) explanatory variables.

You should not focus on building statistical models at this stage.

```r
student_data <- read.csv('https://archive.ics.uci.edu/static/public/697/data.csv')
```

Data wrangling