# Question 1

Start with a descriptive analysis that allows:

* 1. Check for any errors (obviously suspicious observations due to data entry errors, should be corrected if possible). Clarify in your report what mistakes you have detected, and how you detected them. All further questions below should be answered based on the cleaned data.

The active dose, denoting the amount of pills taken a day is given for each group. However, summary statistics show that a number of patients didn’t take the prescribed pills for extended periods of time. Before removing observations of patients who didn’t follow the prescribed treatment plan, the relation between the active dose and weight, age and dbp is examined to explore whether the dosage assigned to each patient might vary according to one of these variables. There’s a weak positive linear relationship between mean dbp measure during the run-off period and the dosage in the first experimental group (correlation coefficient of 0.28), which suggests higher dosages might have been prescribed based on the bloodpressure. However, the total dosage taken during the span of the experiment is consistently a whole number, suggesting that each pill was equally dosed and that some patients simply popped a pill extra or less on a few occasions. It is unknown whether patients were told not to take a pill when their dbp was low or an extra pill when it was very high. Based on this, any patients whose total dosage during the experiment was lower than 80% of the prescribed dosage have been removed (bias).

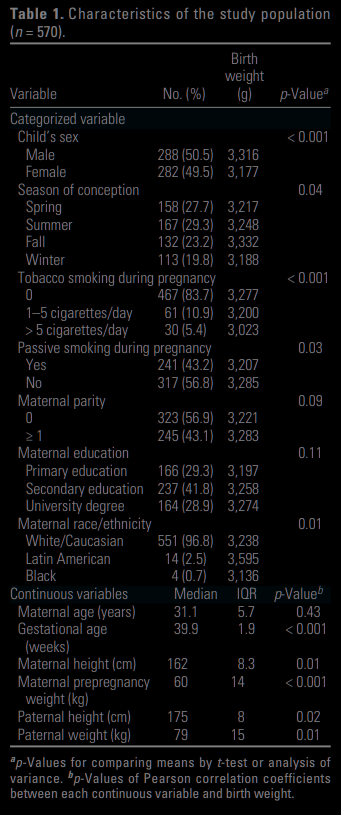
Another oddity found using summary statistics is an outlier dbp3 measurement of 66, which seems unrealistic and thus might have been a measurement error. The outlier has been replace by the mean of the previous two run-in measures (101).

* 1. Describe the study population and hence the scope of the evaluation.

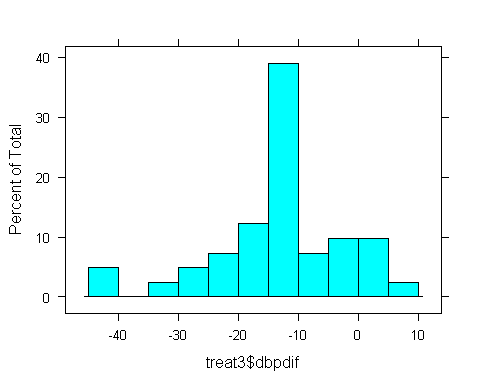
The study population consists of people who suffer from hypertension stage 2 (diastolic bloodpressure above 90). No other criteria seem to have been applied (age, weight, gender).

Characteristics of the study population table (formatted correctly) -> to add (Laurens: still looking into it)

Placeholder example:

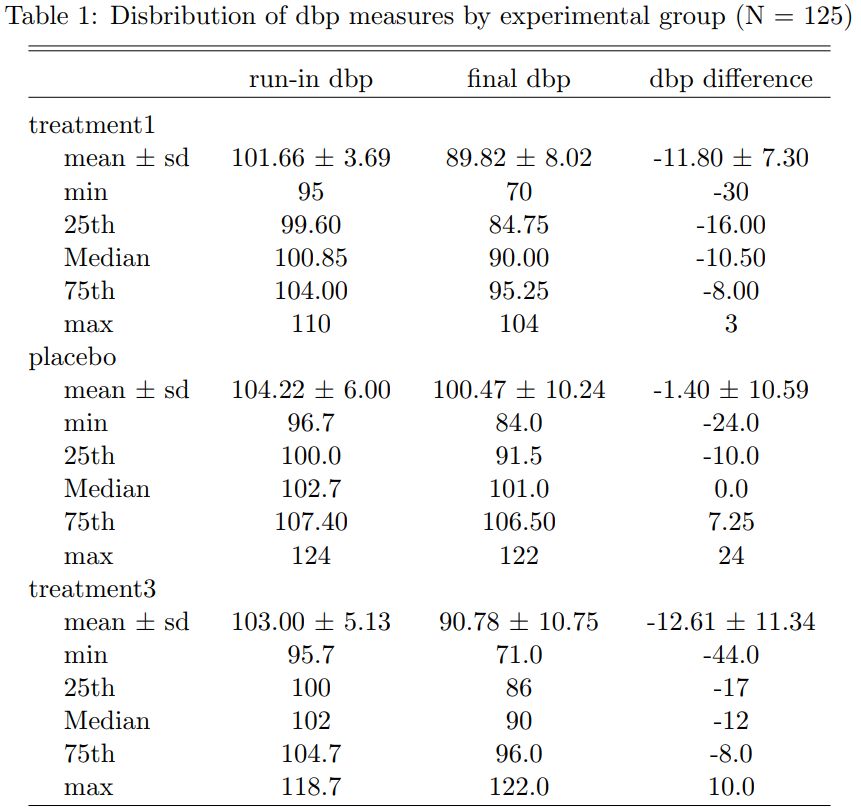


* 1. to help evaluate assumptions which facilitate subsequent analyses.
* Comparable groups
* Population summary table (b) split by experimental group
* Normality: it’s difficult to tell whether dbpdif is normally distributed or not due to a small number of observations



# Question 2

Follow up with a descriptive analysis of your research question. Summarize these results as they would appear in an applied paper.



# Question 3

Compare treatment 3 with placebo. Start by performing a formal primary analysis to evaluate whether the mean change in diastolic blood pressure over the active treatment period depends on the treatment arm. State the assumptions on which your analysis relies and explore/discuss their plausibility in your data. What alternative method do you plan to use if the assumptions are not fulfilled. Motivate your choice for both approaches. Perform them both, consider any differences and draw a conclusion.

* + 1. Assuming normality