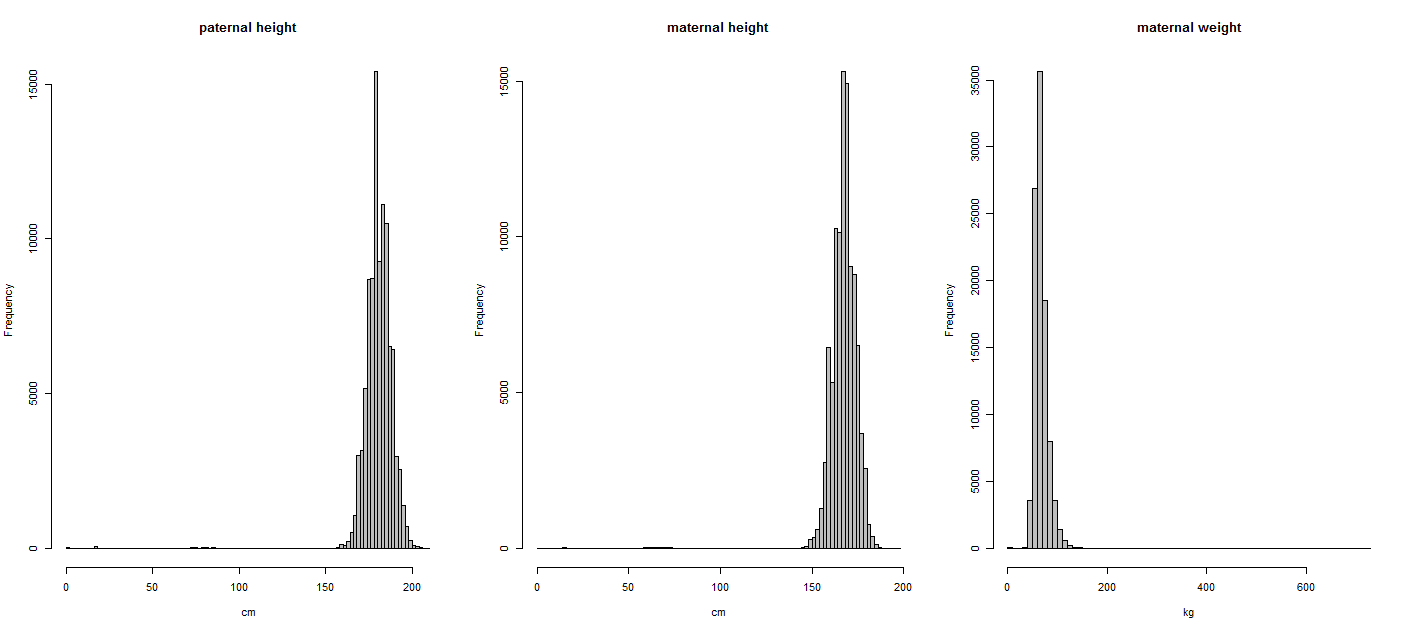
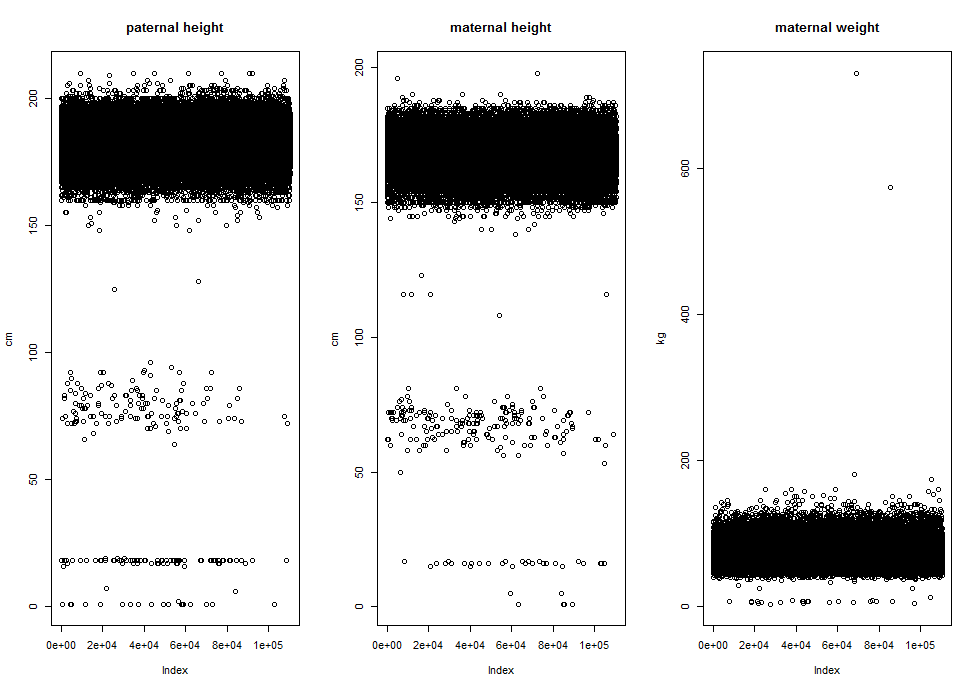
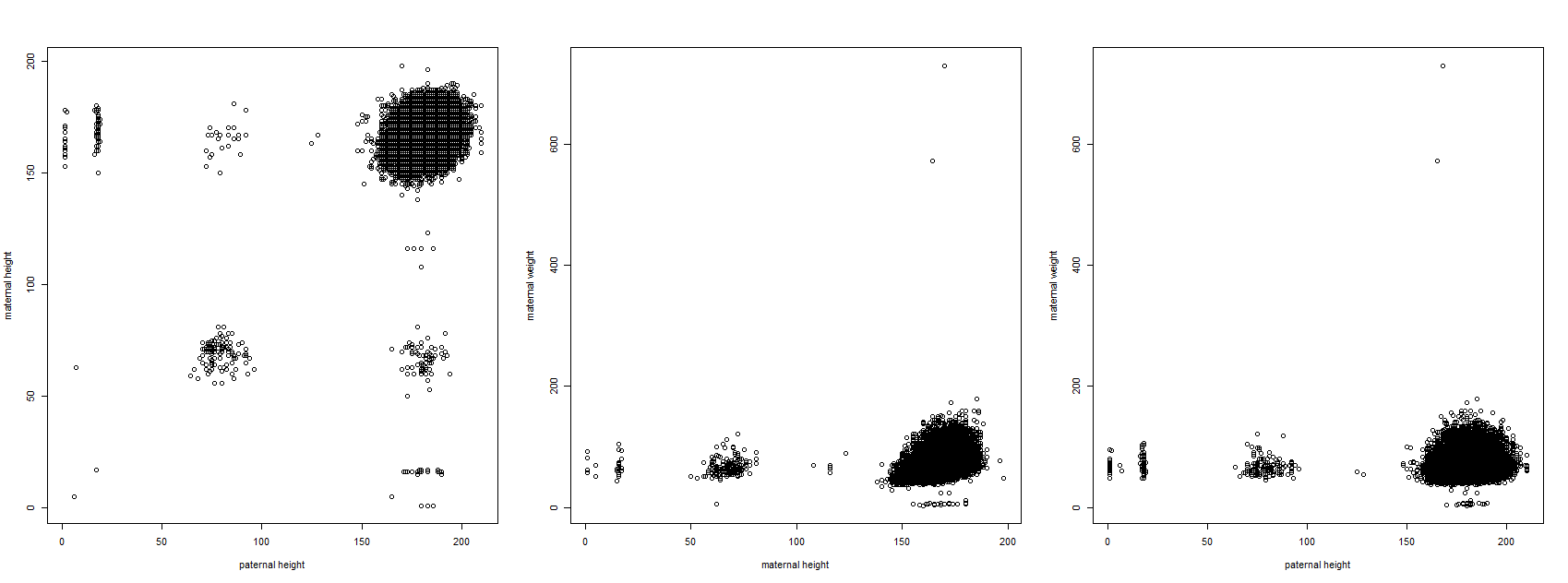
**BEFORE CLEANING**

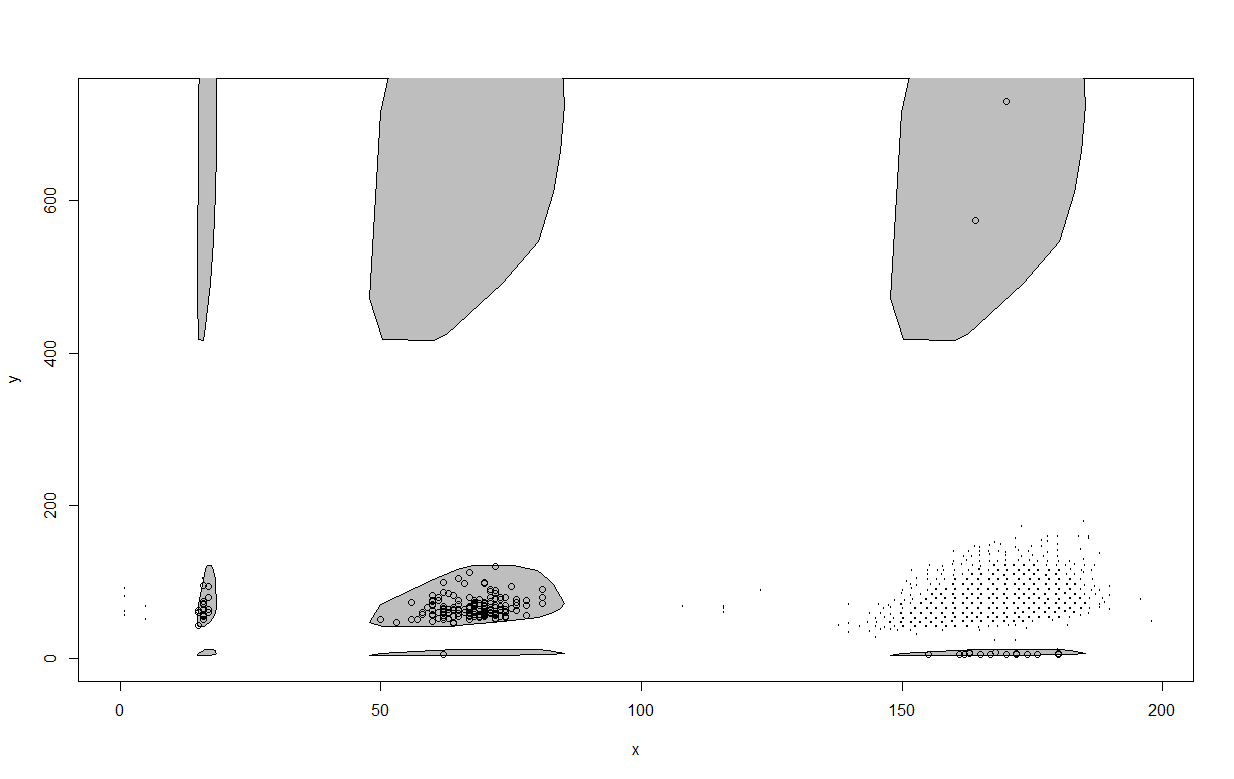




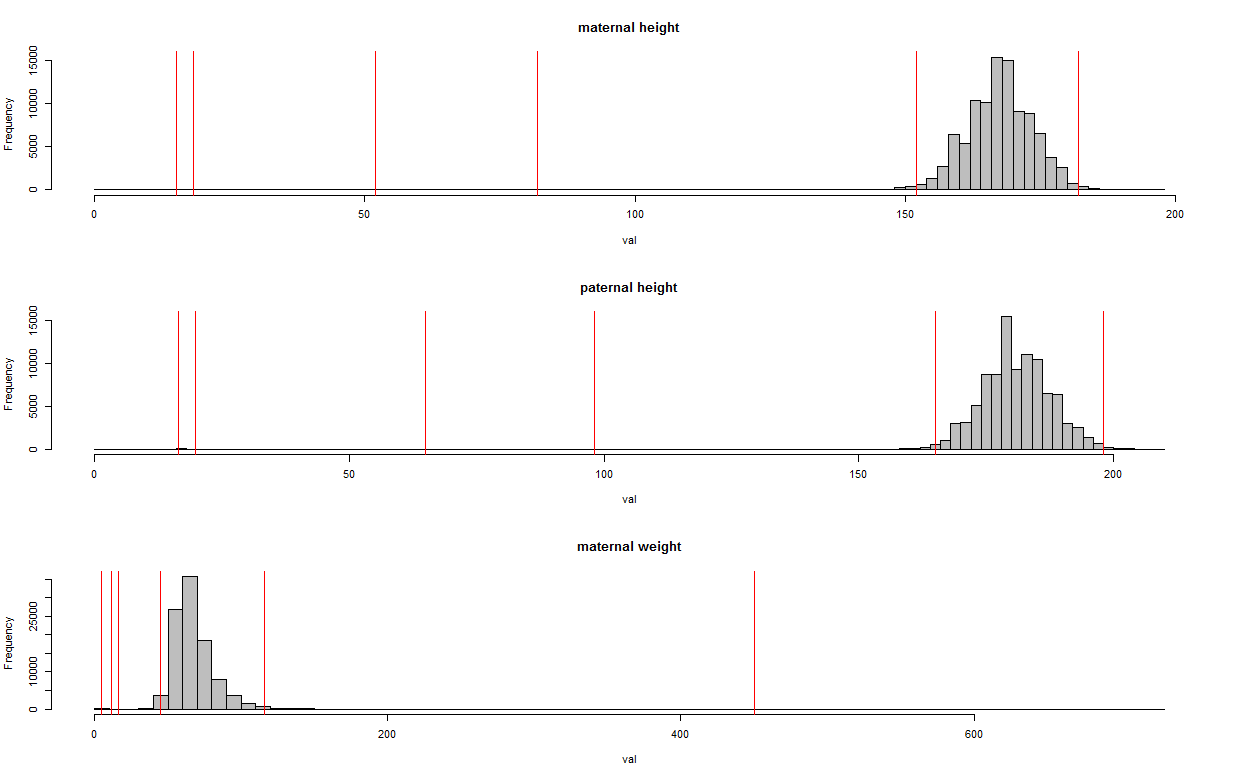
**CORRECTION AREAS-**  based on 99% of most common values

(100 units lower, 10 times lower, 10 times higher and their combinations)

Bivariate plot (maternal height and weight):

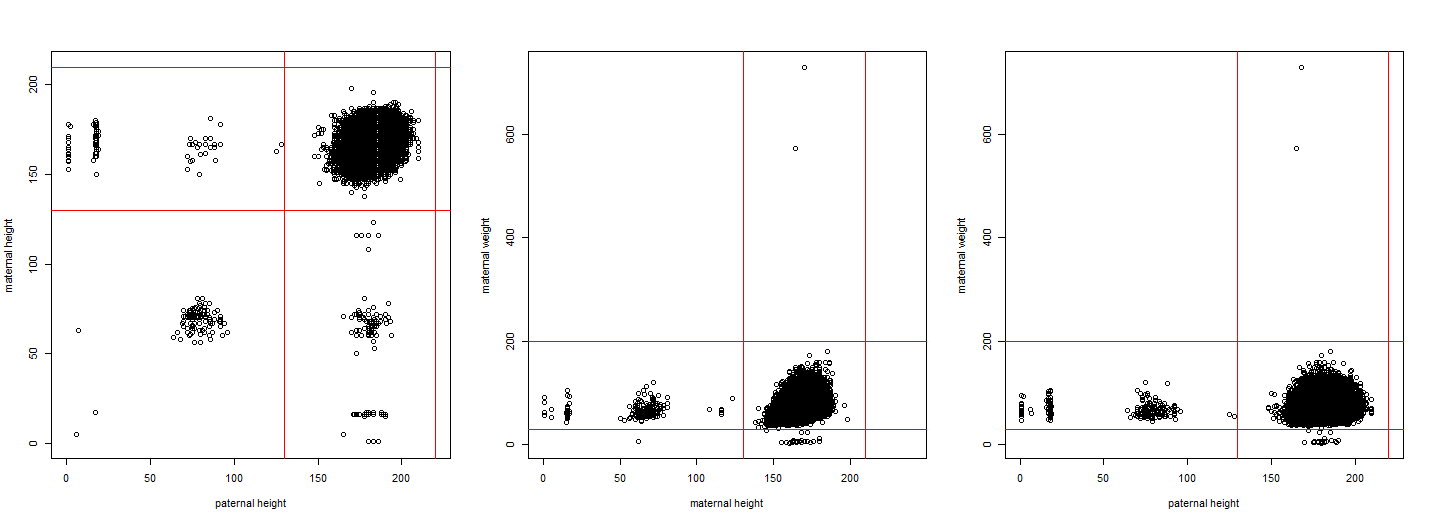


Univariate histograms with correction areas



**ERROR AREAS**

= insane/unreasonable values. Bivariate plots



210cm # maternal height upper threshold

130cm # maternal height lower threshold

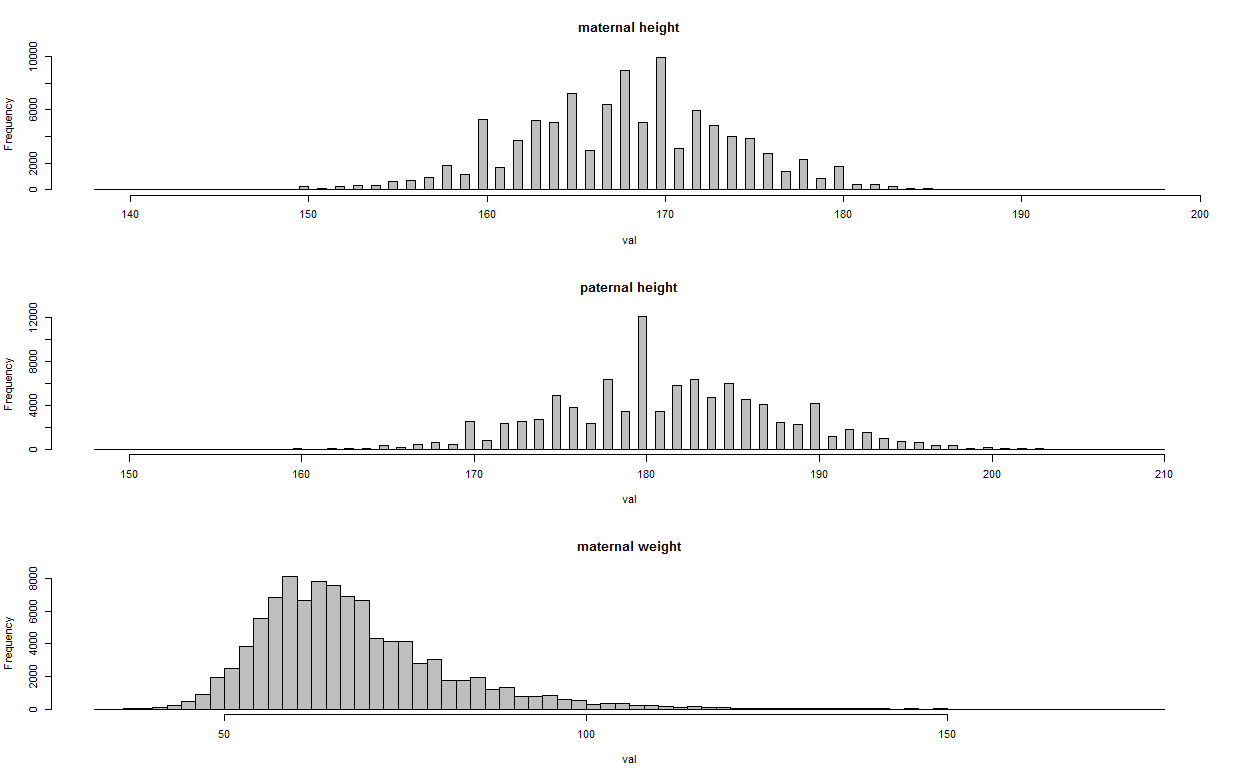
220cm # paternal height upper threshold

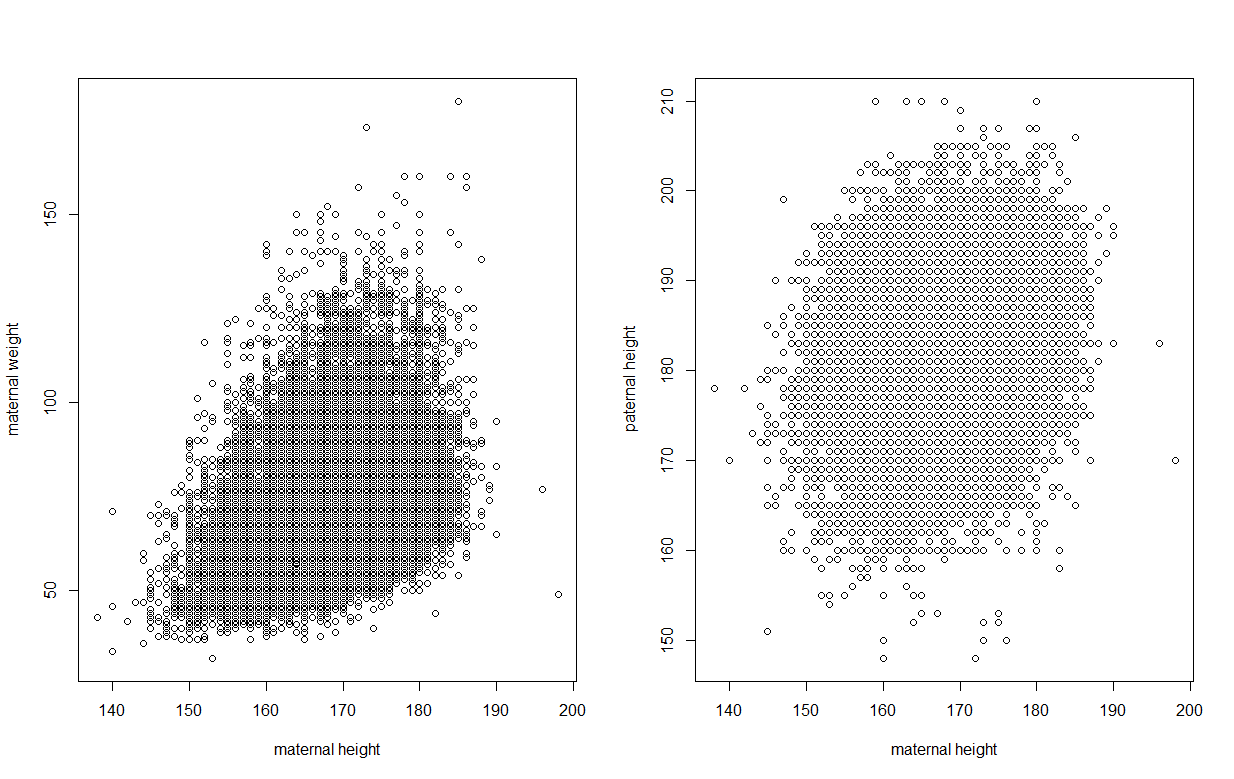
130cm # paternal height lower threshold

200kg # maternal weight upper threshold

30kg # maternal weight lower threshold

**AFTER CLEANING**

****

****

**DESCRIPTIVES:**

**Before:**

> sum(is.na(mh))

[1] 10374

> sum(is.na(mh))/length(mh)

[1] 0.09405173

> sum(is.na(mw))

[1] 11758

> sum(is.na(mw))/length(mw)

[1] 0.1065992

> sum(is.na(ph))

[1] 12028

> sum(is.na(ph))/length(ph)

[1] 0.1090471

**After:**

> sum(is.na(mh))

[1] 10391

> sum(is.na(mh))/length(mh)

[1] 0.09420585

> sum(is.na(mw))

[1] 11765

> sum(is.na(mw))/length(mw)

[1] 0.1066627

> sum(is.na(ph))

[1] 12055

> sum(is.na(ph))/length(ph)

[1] 0.1092918