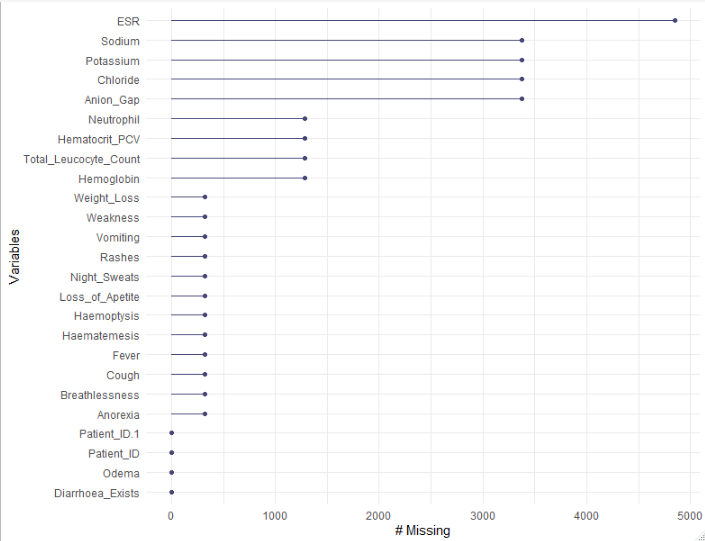
Predicting growth in children

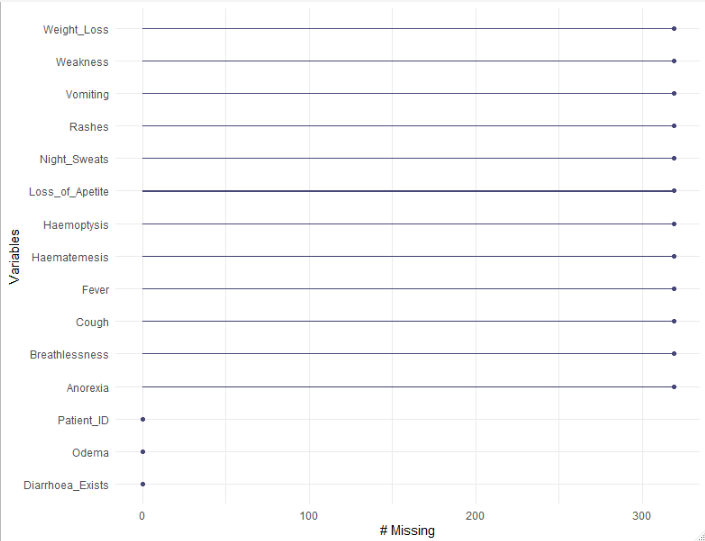
Objective 1, First dataset exploration

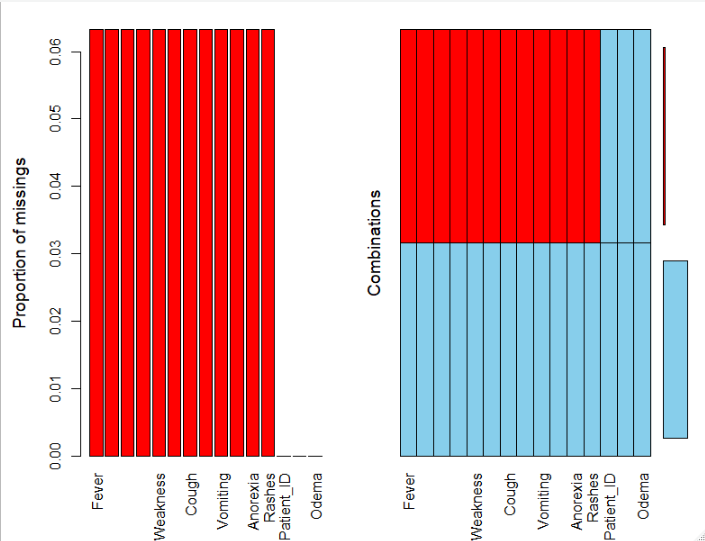
1. Create a smaller dataset with growth measurements (growth dataset), edema, date of birth, age and sex.

* "Length\_\_cm\_ time point 1,2,3,4,5 and 6 "
* "Weight\_\_cm\_ time point 1,2,3,4,5 and 6 "
* "MUAC\_\_cm\_ time point 1,2,3,4,5 and 6 "

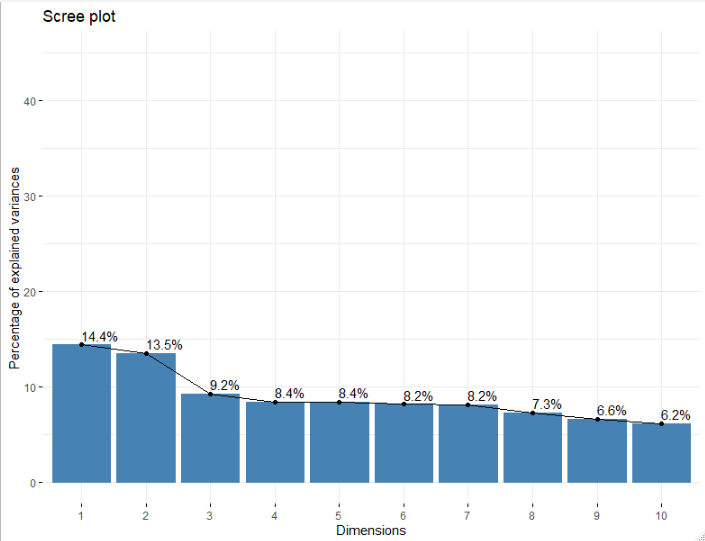
1. Explore the dataset: Calculate of rate of weight gain and missing data

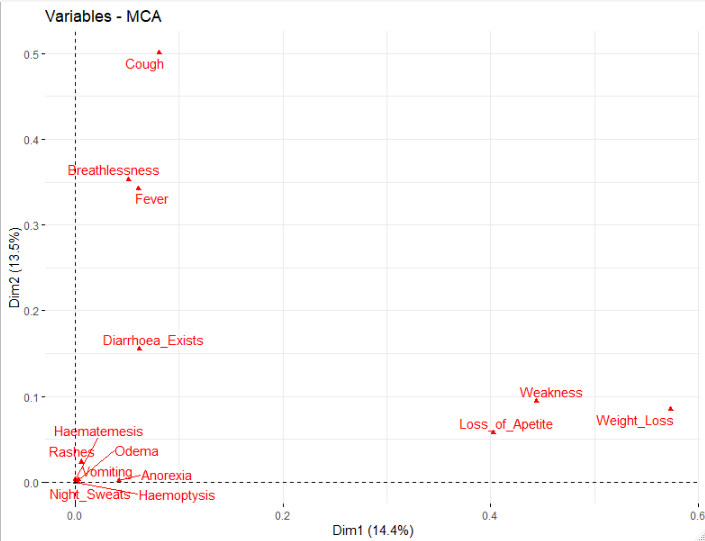


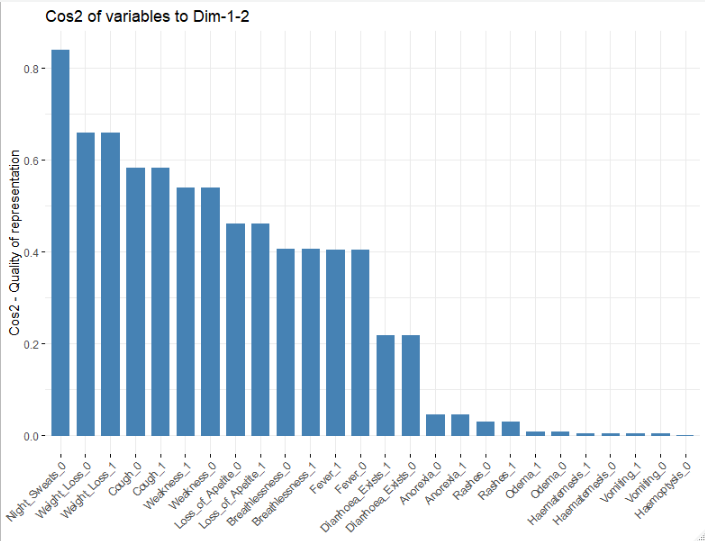


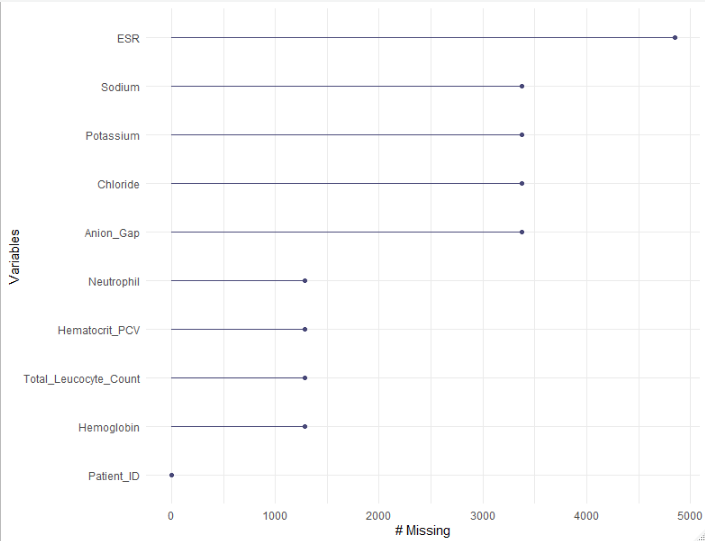


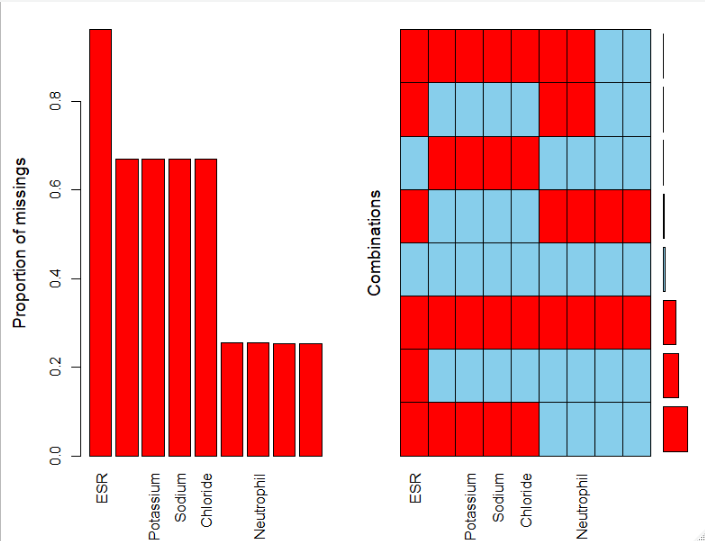
Mca: Groups with similar clinical profile

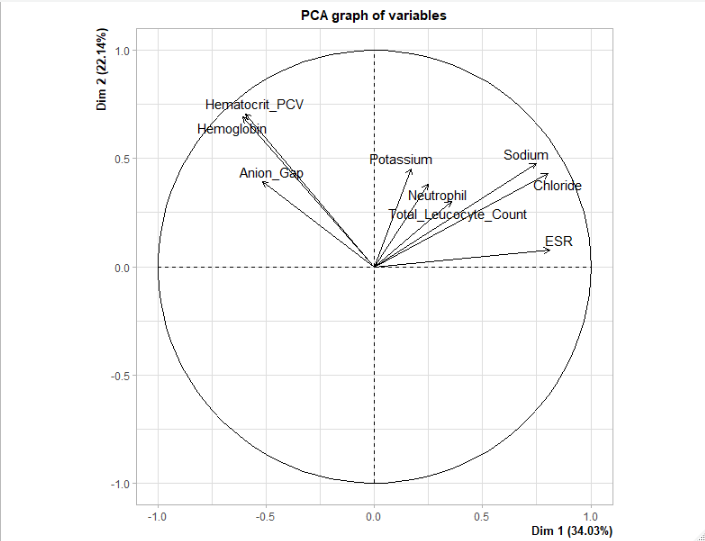


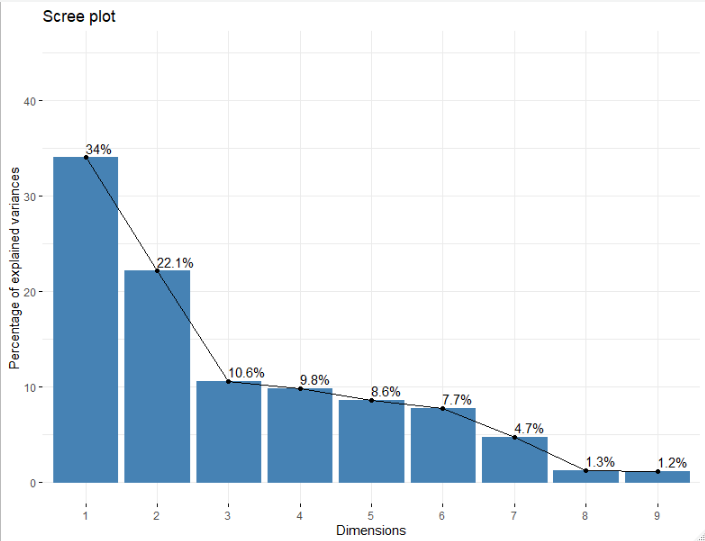


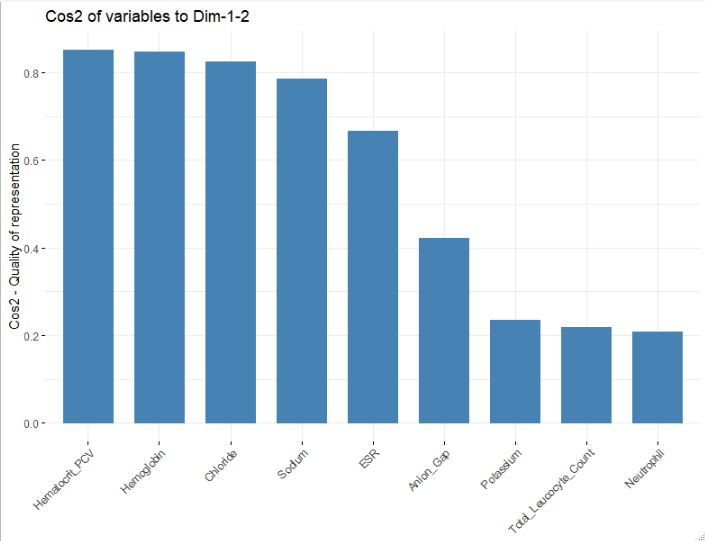












## Clinical data excluding patients with more than 10 NAs (n=319)

|  |  |
| --- | --- |
| **#NAs** | **#Patients** |
| 0 | 147 |
| 1 | 1317 |
| 3 | 3 |
| 4 | 35 |
| 5 | 2164 |
| 9 | 1057 |
| 12 | 8 |
| 13 | 66 |
| 15 | 1 |
| 17 | 145 |
| 21 | 99 |

From the total dataset containing all 21 variables (clinical and biochemical) we identified the “most problematic patients”, meaning those that had NAs in the majority of the variables (>10). The total number of these cases was 319 patients. After excluding these patients, we repeated this study from the other dimension, that of variables. We studied what combinations of variables had the most NAs between them. As it can be seen from the table below, the worst combination was that of {Anion\_gap, ESR, Potassium, Sodium, Chloride}, for which 43% of the patients (n=2044) had only NAs. The second worst combination included only ESR, for which about 28% of the patients (n=1317) had only NAs. Finally, the third worst combination included all biochemical markers, for which about 22% of the patients (n=1057) had only NAs.

|  |  |  |  |
| --- | --- | --- | --- |
| **Index** | **Combinations** | **Count** | **Percent** |
| **1** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0 | 147 | 3.1117697 |
| **2** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:0:0:0:0 | 1317 | 27.8789162 |
| **3** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:0:0:1:1:1 | 35 | 0.7408975 |
| **4** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:1:0:1:1:1 | 2044 | 43.2684166 |
| **5** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:0:1:0:1:0:0:0:0 | 3 | 0.0635055 |
| **6** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:0:1:1:1:0:1:1:1 | 1 | 0.0211685 |
| **7** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:1:1:0:1:1:0:0:0 | 120 | 2.5402202 |
| **8** | 0:0:0:0:0:0:0:0:0:0:0:0:0:0:1:1:1:1:1:1:1:1:1 | 1057 | 22.3751058 |

