| **Updated data dictionary (after multivariate analysis, before final feature selection)** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **#** | **Variable name** | **Variable** | **Data type** | **Measurement units** | **Values/categories** | **Exploratory data analysis results (univariate)** | **Initial interpretation of EDA** |
| 1 | districtName  *Originally district\_name* | District name | Character |  |  | Balaka: 282  Chikwawa: 343  Chiradzulu: 290  Machinga: 310  Mulanje: 229  Nsanje: 546  Thyolo: 301  Zomba: 397  NA: 0 | Samples from the different regions are reasonably similar, with a larger sample from Nsanje |
| 2 | childSex  *Originally child\_sex* | Sex of child | Character (categorical) |  | M, F | F: 1353  M: 1345  NA: 0 | Almost perfectly balanced |
| 3 | bioMom  *Originally q902* | If the respondent is the child’s biological mother | Character (categorical) |  | biological mother, caregiver, other | Biological mother: 2478  Caregiver: 213  Other: 7  NA: 0 | Imbalances, with the majority of respondents being the child’s biological mother |
| 4 | disposeFaeces  *Originally q909* | The last time the child passed stool, where were the faeces disposed of? | Character (categorical) |  | Dropped into toilet facility, dropped into rubbish pit, rinsed/washed away, buried in yard, throw in the compound, throw outside the compound, did nothing/left it there, other, NA | Dropped into toilet facility: 2106  Dropped into rubbish pit: 25  Rinsed/washed away: 462  Buried in yard: 41  Throw in the compound: 2  Throw outside the compound: 20  Did nothing/left it there: 16  Other: 26  NA: 0 | Imbalanced, the majority dropped faeces into toilet facilities |
| 5 | neverHW  *Compiled from q911\_1-4* | The respondent never washes their hands | Logical |  | TRUE, FALSE | TRUE: 258  FALSE: 2440  NA: 0 | Imbalanced, majority do wash their hands |
| 6 | bfFoodPrepHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents before food preparation | Logical |  | TRUE, FALSE | TRUE: 569  FALSE: 2129  NA: 0 | Imbalanced, majority do not wash before food prep |
| 7 | aftFoodPrepHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents after food preparation | Logical |  | TRUE, FALSE | TRUE: 122  FALSE: 2576  NA: 0 | Imbalanced, majority do not wash after food prep |
| 8 | bfFeedChildHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents before feeding children | Logical |  | TRUE, FALSE | TRUE: 2446  FALSE: 252  NA: 0 | Imbalanced, majority do wash before feeding child |
| 9 | aftFeedChildHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents after feeding children | Logical |  | TRUE, FALSE | TRUE: 107  FALSE: 2591  NA: 0 | Imbalanced, majority do not wash after feeding child |
| 10 | bfCleanChildHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents before cleaning children | Logical |  | TRUE, FALSE | TRUE: 119  FALSE: 2579  NA: 0 | Imbalanced, majority do not wash before cleaning child |
| 11 | aftToiletHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents after using the toilet | Logical |  | TRUE, FALSE | TRUE: 2108  FALSE: 590  NA: 0 | Imbalanced, majority do wash after using the toilet |
| 12 | aftChildDefHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents after attending to a child who has defecated | Logical |  | TRUE, FALSE | TRUE: 1679  FALSE: 1019  NA: 0 | Slightly imbalanced, though more wash after cleaning a child that has defecated than not |
| 13 | bfFoodHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents before eating food | Logical |  | TRUE, FALSE | TRUE: 728  FALSE: 1970  NA: 0 | Imbalanced, just under 3/4 do not wash before eating food |
| 14 | aftFoodHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents after eating food | Logical |  | TRUE, FALSE | TRUE: 259  FALSE: 2439  NA: 0 | Imbalanced, majority do not wash after eating food |
| 15 | otherHW  *Compiled from q911\_1-4* | The respondent washes their hands with soap/ash/cleaning agents other times | Logical |  | TRUE, FALSE | TRUE: 41  FALSE: 2657  NA: 0 | Imbalanced, majority do not wash any other time |
| 16 | bedNet  *Compiled from q912 and q913* | If the child slept under a long-lasting insecticide treated bed net last night | Logical |  | TRUE, FALSE, NA | TRUE: 2338  FALSE: 358  NA: 2 | Imbalanced, majority of children slept under a bed net last night |
| 17 | weighIn  *Compiled from q920 and q920b* | If the child was ever weighed at a Ministry of Health or CHAM facility GMP session | Logical |  | TRUE, FALSE, NA | TRUE: 2629  FALSE: 69  NA: 0 | Imbalanced, majority of children have been weighed at a GMP session |
| 18 | childDoc  *Originally q921* | If the child has a health passport or GMP card | Character (categorical) |  | Yes seen, yes not seen, no, NA | No: 148  Yes, not seen: 263  Yes, seen: 2221  NA:66 | Imbalanced, majority have documents and had them the day of the survey |
| 19 | drankBreastmilk  *Compiled from q1001a and q1001b* | If the child was breastfed or fed breastmilk in other ways | Logical |  | TRUE, FALSE, NA | TRUE: 2691  FALSE: 5  NA: 2 | Almost all children were breastfed or received breastmilk |
| 20 | startBreastfeed  *Originally q1002a* | How long after birth the child started breastfeeding | Character (categorical) |  | Same day, 1 day, 2 days, 3 days, NA | 1 day: 80  2 days: 17  3 days: 19  Same day: 2543  NA: 39 | Imbalanced, majority started breastfeeding the day of birth |
| 21 | other3Day  *Originally q1003a* | If the child was given anything other than breastmilk during the first three days of life | Logical |  | TRUE, FALSE, NA | TRUE: 48  FALSE: 2612  NA: 38 | Imbalanced, majority received only breastmilk |
| 22 | other6Month  *Originally q1003b* | If the child was given anything other than breastmilk during the first six months | Logical |  | TRUE, FALSE, NA | TRUE: 1121  FALSE: 1515  NA: 62 | More closely balanced, though more only received breastmilk |
| 23 | breastfedYdayNum  *Originally q1006* | How many times the child breastfed yesterday (day or night) | Integer |  |  | Min: 0  1st Quartile: 0  Median: 0  Mean: 4.012  3rd Quartile: 8  Max: 37  NA: 365 | When 0 not included (only children still breastfed), approaches a normal distribution, with peaks at the “whole” numbers (10, 15, 20) which would be easier to estimate by mothers  When 0 included (all children), skewed to right due to number of 0 |
| 24 | liquidSuppl  *Originally q1008a* | Liquid consumption - did the child consume vitamin or mineral supplements or medicines (including Multiple Micronutrient Powder) yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 216  FALSE: 2480  NA: 2 | Imbalanced, majority did not drink liquid supplements yesterday |
| 25 | liquidORS  *Originally q1008b* | Liquid consumption - did the child consume ORS yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 40  FALSE: 2655  NA: 3 | Imbalanced, majority did not drink oral rehydration solution yesterday |
| 26 | liquidWater  *Originally q1008c* | Liquid consumption - did the child consume plain water yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 2438  FALSE: 258  NA: 2 | Imbalanced, majority did drink water yesterday |
| 27 | liquidFormulaNum  *Originally q1008d\_no* | Number of times the child consumed infant formula yesterday | Integer |  |  | Min: 0  1st Quartile: 0  Median: 0  Mean: 0.009  3rd Quartile: 0  Max: 3  NA: 4 | Imbalanced, with the majority of responses not drinking formula. When only considering children who drank formula, the data leans towards drinking formula one time (skewed to the right) |
| 28 | liquidMilkNum  *Originally q1008e\_no* | Number of times the child consumed animal milk yesterday | Integer |  |  | Min: 0  1st Quartile: 0  Median: 0  Mean: 0.18  3rd Quartile: 0  Max: 3  NA: 37 | Imbalanced, with the majority of responses as not drinking animal milk yesterday. However, when those are removed the data leans towards drinking milk one time (skewed to the right) |
| 29 | liquidJuice  *Originally q1008f* | Liquid consumption - did the child consume Juice or juice drinks yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 231  FALSE: 2465  NA: 2 | Imbalanced, majority did not drink juice yesterday |
| 30 | liquidBroth  *Originally q1008g* | Liquid consumption - did the child consume clear broth yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 19  FALSE: 2677  NA: 2 | Imbalanced, majority did not drink broth yesterday |
| 31 | liquidYogurtNum  *Originally q1008h* | Number of times the child consumed yogurt yesterday | Integer |  |  | Min: 0  1st Quartile: 0  Median: 0  Mean: 0.03  3rd Quartile: 0  Max: 4  NA: 3 | Imbalanced, with the majority of responses as not drinking yogurt yesterday. However, when those are removed the data leans towards drinking yogurt one time (skewed to the right) |
| 32 | liquidPorridge  *Originally q1008i* | Liquid consumption - did the child consume thin porridge yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 427  FALSE: 2269  NA: 2 | Imbalanced, majority did not drink porridge yesterday |
| 33 | liquidTea  *Originally q1008j* | Liquid consumption - did the child consume tea yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 238  FALSE: 2458  NA: 2 | Imbalanced, majority did not drink tea yesterday |
| 34 | liquidOther  *Originally q1008k* | Liquid consumption - did the child consume any other liquids yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 232  FALSE: 2464  NA: 2 | Imbalanced, majority did not drink other liquids yesterday |
| 35 | foodGrain  *Originally q1009a* | Food consumption – did the child consume bread, rice, noodles, or other foods made from grains, including thick grain-based porridge yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 2004  FALSE: 693  NA: 1 | Imbalanced, majority ate grain yesterday |
| 36 | foodYellow  *Originally q1009b* | Food consumption – did the child consume pumpkin, carrots, squash, or orange/yellow fleshed sweet potato yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 93  FALSE: 2604  NA: 1 | Imbalanced, majority did not eat yellow produce yesterday |
| 37 | foodWhite  *Originally q1009c* | Food consumption – did the child consume white fleshed sweet potatoes, white yams, manioc, cassava, or any other foods made from roots yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 573  FALSE: 2124  NA: 1 | Imbalanced, majority did not eat white roots yesterday |
| 38 | foodGreen  *Originally q1009d* | Food consumption – did the child consume dark green leafy vegetables yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 1221  FALSE: 1475  NA: 2 | More closely balanced, half the children ate dark green leafy vegetables yesterday |
| 39 | foodMangoPap  *Originally q1009e* | Food consumption – did the child consume ripe mangoes or papayas yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 1702  FALSE: 994  NA: 2 | Partially imbalanced, ~2/3 ate mango or papaya yesterday |
| 40 | foodProduce  *Originally q1009f* | Food consumption – did the child consume any other fruits or vegetables yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 345  FALSE: 2352  NA: 1 | Imbalanced, majority did not eat other produce yesterday |
| 41 | foodOrgan  *Originally q1009g* | Food consumption – did the child consume liver, kidneys, heart, or other organ meat yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 15  FALSE: 2682  NA: 1 | Imbalanced, majority did not eat organ meat yesterday |
| 42 | foodMeat  *Originally q1009h* | Food consumption – did the child consume any meat, such as beef, pork, lamb, goat, chicken, or duck yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 113  FALSE: 2584  NA: 1 | Imbalanced, majority did not eat meat yesterday |
| 43 | foodEgg  *Originally q1009i* | Food consumption – did the child consume eggs yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 142  FALSE: 2552  NA: 4 | Imbalanced, majority did not eat eggs yesterday |
| 44 | foodFish  *Originally q1009j* | Food consumption – did the child consume fresh or dried fish, shellfish, or seafood yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 987  FALSE: 1710  NA: 1 | Partially imbalanced, ~1/3 ate fish yesterday |
| 45 | foodBean  *Originally q1009k* | Food consumption – did the child consume any foods made from beans, peas, lentils, or nuts yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 1161  FALSE: 1536  NA: 1 | More closely balanced, close to half the children ate beans, peas, lentils, or nuts yesterday |
| 46 | foodFats  *Originally q1009m* | Food consumption – did the child consume oil, fats, butter, or foods made with any of these yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 1342  FALSE: 1355  NA: 1 | Balanced, half of the children ate some sort of fat yesterday |
| 47 | foodSugar  *Originally q1009n* | Food consumption – did the child consume sugary foods such as chocolates, sweets, candies, pastries, cakes, or biscuits yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 551  FALSE: 2146  NA: 1 | Imbalanced, majority did not eat sugar yesterday |
| 48 | foodCondiment  *Originally q1009o* | Food consumption – did the child consume condiments for flavor, such as chilies, spices, herbs, or fish powder yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 158  FALSE: 2539  NA: 1 | Imbalanced, majority did not eat condiments yesterday |
| 49 | foodInsect  *Originally q1009p* | Food consumption – did the child consume grubs, snails, or insects yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 65  FALSE: 2632  NA: 1 | Imbalanced, majority did not eat insects yesterday |
| 50 | foodRedPalm  *Originally q1009q* | Food consumption – did the child consume foods made with red palm oil, red palm nut, and red palm nut pulp sauce yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 9  FALSE: 2688  NA: 1 | Imbalanced, majority did not eat red palm yesterday |
| 51 | foodOther  *Originally q1009r* | Food consumption – did the child consume any other solid or semi-solid food yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 1930  FALSE: 766  NA: 2 | Partially imbalanced, ~2/3 ate other food yesterday |
| 52 | foodSalt  *Originally q1009s* | Food consumption – did the child consume salt yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 2392  FALSE: 305  NA: 1 | Imbalanced, majority did not eat salt yesterday |
| 53 | foodNum  *Originally q1010* | How many times did the child eat solid, semi-solid, or soft foods yesterday (day and night) | Integer |  |  | Min: 0  1st Quartile: 2  Median: 3  Mean: 2.544  3rd Quartile: 3  Max: 9  NA: 4 | Skewed to the right, with the majority of children eating 2-3 times a day |
| 54 | childCode  *Originally anon\_child\_code* | From original survey | Integer |  |  |  |  |
| 55 | childQNum  *Originally anon\_child\_qnum* | From original survey | Integer |  |  |  |  |
| 56 | Treatment | Whether the child was in a WALA program community (2) or not (1) | Integer |  | 0, 1 | 1 (WALA community): 1340  0 (other community): 1358 | Balanced, expected given the survey design (paired communities) |
| 57 | ageGroup | The child’s age group in months | Character (categorical) |  | 00-05 mo, 06-11 mo, 12-23 mo, 24-35 mo, 36-47 mo, 48-59 mo | 00-05 mo: 269  06-11 mo: 316  12-23 mo: 530  24-35 mo: 533  36-47 mo: 517  48-59 mo: 533  NA: 0 | Balanced by year of age (first two groups are 6 months, remaining are one year) |
| 58 | stunted | Dependent variable 1: if the child is stunted (short). | Logical |  | TRUE, FALSE, NA | TRUE: 809  FALSE: 1610  NA: 279 | Imbalanced, with ~1/3 stunted |
| 59 | wasting | Dependent variable 2: if the child is wasting (underweight for height). | Logical |  | TRUE, FALSE, NA | TRUE: 87  FALSE: 2247  NA: 364 | Imbalanced, with ~7.5% wasting |
| 60 | underweight | Dependent variable 3: if the child is underweight or has oedema. | Logical |  | TRUE, FALSE, NA | TRUE: 397  FALSE: 2295  NA: 6 | Imbalanced, with ~15% underweight |
| 61 | malnourished | Dependent variable 4: if the child shows signs of malnourishment (TRUE for dependent variables 1-3) | Logical |  | TRUE, FALSE | TRUE: 1740  FALSE: 958  NA: 0 | Imbalanced, with >1/3 being malnourished in some way |
| 62 | childWeighIn | If the child was weighed at least once in the past three months | Logical |  | TRUE, FALSE, NA | TRUE: 1240  FALSE: 981  NA: 477 | While it was closely split between yes and no, there were also almost 500 NA |
| 63 | anyDairy | Food and liquid consumption – did the child consume cheese, yogurt, or other milk products (solid or liquid) yesterday? | Logical |  | TRUE, FALSE, NA | TRUE: 84  FALSE: 2612  NA: 2 | Imbalanced, majority did not eat diary yesterday |