**Aim 1 Data Dictionary:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Units** | **Name in Survey Dataset** | **Coded name in analysis** |
| Household (HH) Visit number (timepoint) | Integer | visit | visit |
| **Neighborhood Variables** | | | |
| Sub-neighborhood code (GEE cluster variable) | Nominal | code | code |
| Sub-neighborhood matching strata identifier | Factor | final\_match\_strata\_3 | final\_match\_strata\_3 |
| **Household Variables (covariates)** | | | |
| SES Score (poverty score of >66) | Integer | Based on I\_2:8, I\_10, I\_11 | high\_poverty |
| Basic sanitation access at enrollment | Categorical | Based on E\_1, E\_4, and E\_2 | san\_basic\_obs\_bl |
| Fixed employment of primary wage earner | Binomial | Based on A\_9, A\_12, and A\_10 | fixed\_emp\_pri |
| Secondary education or above complete at enrollment | Binomial | Based on A\_8 | secondary\_complete\_bl |
| Number of children under 5 in HH | Continuous | Based on A\_19 | num\_lt5 |
| HH density | Continuous | Based on A\_17 | num\_HH |
| Months living in HH | Continuous | Based on A\_15 | months\_in\_hh |
| Human feces in or near the HH | Binomial | Based on M\_13 | human\_feces |
| Animal feces in or near the HH | Binomial | Based on M\_14 | animal\_feces |
| Food insecurity (HFIAS score) | Integer | Based on G1:G10 | HFIAS\_score |
| Handwashing (HW) station in dwelling/yard | Binomial | Based on JMP\_H1:H3 | HW\_dwell |
| Seasonality (rainy vs dry) | Binomial | Based on hh\_date – Dec-April = Rainy | season\_rainy |
| Flooding in home or yard in the last month | Binomial | Based on K\_10 | flooding\_HH\_yard |
| **Intervention** | | | |
| Intervention neighborhood designation (1= intervention, 0 = control) | Binomial | arm | arm |
| Direct connection- active FIPAG connection at the household (1= hh connection, 0 = no hh connection\_ | Binomial | C\_1 & C\_2 | hh\_water |
| **Water Quality and Access – primary** | | | |
| Prevalence of any detectable *E. coli* in source water (primary outcome) – measured by IDEXX | Binomial | ecoli\_bin\_source | ecoli\_bin\_source |
| Prevalence of any detectable *E. coli* in stored water – measured by IDEXX | Binomial | ecoli\_bin\_stored | ecoli\_bin\_stored |
| Prevalence of access to at least basic water source (defined by the JMP) | Binomial | Based on water\_service\_ladder (created by Josh) | basic\_water |
| Prevalence of water that is accessible on premises | Binomial | Based on C\_10 and C\_11 | onpremises |
| Prevalence of HH with sufficient quantities of drinking water when needed | Binomial | Baed on C\_23 | sufficient |
| Prevalence of HH with HWISE score >= 12 | Binomial | Based on D\_1:12 -> HWISE\_scale | HWISE\_insecure |
| Prevalence of ALWAYS being satisfied with water service ( 1= always satisfied, 0 = not always satisfied) | Binomial | C\_17\_bin | always\_sat\_service |
| Prevalence of ALWAYS being satisfied with water affordability (1= always satisfied, 0 = not always satisfied) | Binomial | C\_18\_bin | always\_sat\_afford |
| Prevalence of ALWAYS being satisfied with water availability (1= always satisfied, 0 = not always satisfied) | Binomial | C\_19\_bin | always\_sat\_avail |
| Prevalence of ALWAYS being satisfied with water pressure (1= always satisfied, 0 = not always satisfied) | Binomial | C\_20\_bin | always\_sat\_pressure |
| Prevalence of ALWAYS being satisfied with water color and appearance (1= always satisfied, 0 = not always satisfied) | Binomial | C\_21\_bin | always\_sat\_appear |
| Prevalence of ALWAYS being satisfied with water taste and smell | Binomial | C\_22\_bin | always\_sat\_taste |
| Water usage (L/day)- respondent reported (above or below median use of 80 L/day) | Binomial | Based on C\_15\_D | use\_respond\_bin |
| Sufficient levels of free chlorine in source water (>=0.2 mg/L) - measured | Binomial | Based on O\_W\_7 | free\_chlor\_WHO |