List of known inconsistencies arising from DREAMS data cleaning

Date: 11 July 2020
Author: Guy Harling

These notes based on data provided by Jaco Dreyer on July 10th 2020.

	** General/intro
Flag	Preloaded sex was 'male' for 4 people (1 interviewed) and '.' for 87 (9 interviewed). Flag added to dataset as c_pd_sex_flag
Flag	Rapid HIV tests were not done in Round 2, but consent obtained (ic_indiv_consent_hiv11 and ic_indiv_consent_hiv_1).
	** Section 3. Education, Economics and Nutrition
Flag	ie_reas_leave_school_* hidden in round 3, so missing
Flag	25 early round 1 respondents have data on who decides about spending, despite saying they don't make such spending: *tab ca_visitdate if !missing(ie_spenton_decide) & ie_dailyneeds_spend > 1*
	** Section 4. DREAMS
Flag	One interviewer entered 'ahri' for almost all interviews in dm_source_other in round 3, including 39 who did not say 'yes' to dm_source_96. *bys dm_source_96: tab dm_source_other round if dm_source_other != "", miss*
Flag	For all DREAMS loops, 2016: awareness pre participation; 2017/2018: participation pre awareness
Flag	Questions about schooledu were asked to everyone in round 2, but only those in school in waves 1 and 3. I have hidden responses from those out of school in round 2 (.n) to be consistent.
Flag	Questions in dm*finserv* required ie_inschool == 1 in round 2, but not in round 1 or 3. I have **not** accounted for this when building the created variables, but you might well want to.

Flag	Many changes in question sets for DREAMS across rounds. Merged from round 2:
	    finskill: merge of finlitera + vocatskill
	    finserv: merge of savegroup + microfin
	    cashtran: merge of cashagyw + cashtran (family) + schoolsup
	    hivtest: merge of mobiletest + hometest + familytest
	    genorm: merge of genorm + dm_*violprevnt* (comm) + *violprev_* (schools)
	Truly new in round 2: democond: pepserv
	Dropped in round 2: shuga (gets own section); linkart; artadhere
	New in round 3: thethana

	** Section 5. Mediating variables
Flag	To date we have not recoded several scales, including at minimum: self-efficacy, Hope, Youth importance, Happiness, Life satisfaction, GOALS, Chance to Achieve, GEMS violence
	** Section 7. General health
Flag	45 respondents in R1 say they have never received a test result, but saw the follow-up question asking if ever tested positive. Fortunately all say 'no'. Now recoded to fix. tab hq_pos_hivresult hq_recv_hivresult if round == 1, miss
Flag	hq_pep_taken has two years given instead of time. Left as-is for now
Flag	Skip pattern for hq_prep_currtaken differ by wave: R1 based on hq_prep_knowof; R2/3 based on hq_prep_taken12mths
	** Section 8. Sexual reproductive health
Flag	Missing data for 10 women in round 1 due to pd_sex != 2 – see c_pd_sex_flag
Flag	Six extra people were skipped for much of the SRH section in R1 – source of skip unclear *tab ca_visitdate if sr_currpregnant == .*
Flag	Four women give a year for sr_times_pregnant
	** Section 9. CASI questionnaire
Flag	sr_12mth_newsexpartn was created as a text field in error; now cleaned into c_ variable
Flag	Due to a REDCap issue, all of srs_recnt1* are missing for round 3
Flag	srs_recnt3_personhivstatus only exists for 3rd most recent partner