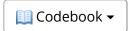# Data Cleaning Example

📖 Codebook ▾

📖 Data Dictionary Codebook   04/08/2018 6:04pm

#	Variable / Field Name	Field Label *Field Note*	Field Attributes (Field Type, Validation, Choices, Calculations, etc.)
	Instrument: **Demographics** (demographics)		
1	study_id	Study ID	text
2	date_enrolled	Section Header: *Consent Information* Date subject signed consent *YYYY-MM-DD*	text (date_ymd)
3	patient_document	Upload the patient's consent form	file
4	family_name	Section Header: *Contact Information* Family Name	text, Identifier
5	given_name	Given Name	text, Identifier
6	street_address	Street Address	text
7	city_prefecture	City, Prefecture	notes, Identifier
8	postal_code	Postal Code *Format: xxx-xxxx*	text
9	phone	Phone number	text (phone), Identifier
10	email	E-mail	text (email), Identifier
11	dob	Date of birth	text (date_ymd, Min: 1900-05-31, Max: 2000-05-31), Identifier
12	age	Age (years)	calc Calculation: rounddown(datediff([dob],'today','y'))
13	age_warning Show the field ONLY if: [age] < 18	WARNING: If patient is < 18 years old, he or she should not be enrolled in this study.	descriptive
14	gender	Gender	radio | 0	Female
1	Male	
15	given_birth Show the field ONLY if: [gender] = "0"	Has the patient given birth before?	yesno | 1	Yes
0	No	

16	num_children Show the field ONLY if: [gender] = "0" and [given_birth] = "1"	How many times has the patient given birth?	text (integer, Min: 0)		
17	gym	Section Header: *Please provide the patient's weekly schedule for the activities below.* Gym (Weight Training)	checkbox		
			0	gym___0	Monday
			1	gym___1	Tuesday
			2	gym___2	Wednesday
			3	gym___3	Thursday
			4	gym___4	Friday
			5	gym___5	Saturday
			6	gym___6	Sunday
18	aerobics	Aerobics	checkbox		
			0	aerobics___0	Monday
			1	aerobics___1	Tuesday
			2	aerobics___2	Wednesday
			3	aerobics___3	Thursday
			4	aerobics___4	Friday
			5	aerobics___5	Saturday
			6	aerobics___6	Sunday
19	eat	Eat Out (Dinner/Lunch)	checkbox		
			0	eat___0	Monday
			1	eat___1	Tuesday
			2	eat___2	Wednesday
			3	eat___3	Thursday
			4	eat___4	Friday
			5	eat___5	Saturday
			6	eat___6	Sunday
20	drink	Drink (Alcoholic Beverages)	checkbox		
			0	drink___0	Monday
			1	drink___1	Tuesday
			2	drink___2	Wednesday
			3	drink___3	Thursday
			4	drink___4	Friday
			5	drink___5	Saturday
			6	drink___6	Sunday

21	mood	Section Header: *Other information* Specify the patient's mood.	slider Slider labels: Very sad, Indifferent, Very happy Custom alignment: RH
22	any_antidep	Is the patient taking any antidepressants?	checkbox 1 | any_antidep___1 | Yes 0 | any_antidep___0 | No
23	which_antidep Show the field ONLY if: [any_antidep(1)] = '1'	Which antidepressant(s) is the patient taking? (check all that apply)	checkbox 1 | which_antidep___1 | Lexapro 2 | which_antidep___2 | Celexa 3 | which_antidep___3 | Prozac 4 | which_antidep___4 | Paxil 5 | which_antidep___5 | Zoloft
24	height	Height (cm)	text (number, Min: 130, Max: 215)
25	weight	Weight (kilograms)	text (integer, Min: 35, Max: 200)
26	bmi	BMI	calc Calculation: round((([weight]*10000)/(([height])^(2)),1)
27	demographics_complete	Section Header: *Form Status* Complete?	dropdown 0 | Incomplete 1 | Unverified 2 | Complete

Instrument: **Baseline Data** (baseline_data)

28	date_visit_b	Section Header: *Baseline Measurements* Date of baseline visit	text (date_ymd)
29	creat_b	Creatinine (mg/dL)	text (number, Min: 0.5, Max: 20)
30	hdl_b	HDL Cholesterol (mg/dL)	text (number, Min: 20, Max: 100)
31	ldl_b	LDL Cholesterol (mg/dL)	text (number, Min: 30, Max: 300)
32	drywt_b	Dry weight (kilograms)	text (number, Min: 35, Max: 200)
33	plasma1_b	Collected Plasma 1?	dropdown 0 | No 1 | Yes
34	plasma2_b	Collected Plasma 2?	dropdown 0 | No 1 | Yes

35	serum1_b	Collected Serum 1?	dropdown 0 No 1 Yes
36	serum2_b	Collected Serum 2?	dropdown 0 No 1 Yes
37	sga_b	Subject Global Assessment (score = 1-7)	text (number, Min: 0.9, Max: 7.1)
38	date_supplement_dispensed	Date patient begins supplement	text (date_ymd)
39	baseline_data_complete	Section Header: *Form Status* Complete?	dropdown 0 Incomplete 1 Unverified 2 Complete

Instrument: **Monthly Data** (monthly_data)			
40	date_visit_m	Section Header: *Month 1* Date of Monthly Visit	text (date_ymd)
41	creat_m	Creatinine (mg/dL)	text (number, Min: 0.5, Max: 20)
42	hdl_m	HDL Cholesterol (mg/dL)	text (number, Min: 20, Max: 100)
43	ldl_m	LDL Cholesterol (mg/dL)	text (number, Min: 30, Max: 300)
44	drywt_m	Dry weight (kilograms)	text (number, Min: 35, Max: 200)
45	trt_missed	Number of treatments missed	text (number, Min: 0, Max: 7)
46	compliance	How compliant was the patient in drinking the supplement?	dropdown 0 100 percent 1 99-75 percent 2 74-50 percent 3 49-25 percent 4 0-24 percent
47	hosp_yn	Section Header: *Hospitalization Data* Was patient hospitalized since last visit?	dropdown 0 No 1 Yes
48	hosp_cause	What was the cause of hospitalization?	dropdown 1 Vascular access related events 2 CVD events 3 Other
49	hosp_adm	Date of hospital admission	text (date_ymd)
50	hosp_dis	Date of hospital discharge	text (date_ymd)

51	hosp_summary_binder	Discharge summary in patients binder?	dropdown 0 No 1 Yes
52	monthly_data_complete	Section Header: *Form Status* Complete?	dropdown 0 Incomplete 1 Unverified 2 Complete

Instrument: **Completion Data** (completion_data)			
53	completed_study	Section Header: *Study Completion Information* Has patient completed study?	dropdown 0 No 1 Yes
54	completed_date Show the field ONLY if: [completed_study] = '1'	Date of study completion	text (date_ymd)
55	studywd_yn Show the field ONLY if: [completed_study] = '0'	Did the patient ever withdraw from the study?	dropdown 0 No 1 Yes
56	studywd_date Show the field ONLY if: [studywd_yn] = '1'	Date of study withdrawal	text (date_ymd)
57	studywd_reason Show the field ONLY if: [studywd_yn] = '1'	Reason patient withdrew from study	dropdown 0 Non-compliance 1 Did not wish to continue in study 2 Could not tolerate the supplement 3 Hospitalization 4 Other
58	death_yn Show the field ONLY if: [completed_study] = '0'	Did the patient die during the study?	dropdown 0 No 1 Yes

59	death_date Show the field ONLY if: [death_yn] = '1'	Date of death	text (date_ymd)
60	death_cause Show the field ONLY if: [death_yn] = '1'	Cause of death	dropdown | 0	All cause
1	Cardiovascular event	
61	study_comments	Section Header: *General Comments* Comments	notes
62	completion_data_complete	Section Header: *Form Status* Complete?	dropdown | 0	Incomplete
1	Unverified
2	Complete	