

## Example: Data Cleaning with R (Data Entry)

[Codebook ▾](#)

### Data Dictionary Codebook

04/10/2018 9:31pm

	#	Variable / Field Name	Field Label <i>Field Note</i>	Field Attributes (Field Type, Validation, Choices, Calculations, etc.)				
Instrument: <b>Demographics</b> (demographics)								
	1	study_id	Study ID	text				
	2	date_enrolled	Section Header: <i>Consent Information</i> Date subject signed consent <i>YYYY-MM-DD</i>	text (date_ymd)				
	3	patient_document	Upload the patient's consent form	file				
	4	family_name	Section Header: <i>Contact Information</i> 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 <i>Format: xxx-xxxx</i>	text				
	9	phone	Phone number	text, 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 <table><tr><td>0</td><td>Female</td></tr><tr><td>1</td><td>Male</td></tr></table>	0	Female	1	Male
0	Female							
1	Male							
	15	given_birth  Show the field ONLY if: [gender] = "0"	Has the patient given birth before?	yesno <table><tr><td>1</td><td>Yes</td></tr><tr><td>0</td><td>No</td></tr></table>	1	Yes	0	No
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: <i>Please provide the patient's weekly schedule for the activities below.</i>  Gym (Weight Training)	checkbox <table border="1"> <tr><td>0</td><td>gym__0</td><td>Monday</td></tr> <tr><td>1</td><td>gym__1</td><td>Tuesday</td></tr> <tr><td>2</td><td>gym__2</td><td>Wednesday</td></tr> <tr><td>3</td><td>gym__3</td><td>Thursday</td></tr> <tr><td>4</td><td>gym__4</td><td>Friday</td></tr> <tr><td>5</td><td>gym__5</td><td>Saturday</td></tr> <tr><td>6</td><td>gym__6</td><td>Sunday</td></tr> </table>	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
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 <table border="1"> <tr><td>0</td><td>aerobics__0</td><td>Monday</td></tr> <tr><td>1</td><td>aerobics__1</td><td>Tuesday</td></tr> <tr><td>2</td><td>aerobics__2</td><td>Wednesday</td></tr> <tr><td>3</td><td>aerobics__3</td><td>Thursday</td></tr> <tr><td>4</td><td>aerobics__4</td><td>Friday</td></tr> <tr><td>5</td><td>aerobics__5</td><td>Saturday</td></tr> <tr><td>6</td><td>aerobics__6</td><td>Sunday</td></tr> </table>	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
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 <table border="1"> <tr><td>0</td><td>eat__0</td><td>Monday</td></tr> <tr><td>1</td><td>eat__1</td><td>Tuesday</td></tr> <tr><td>2</td><td>eat__2</td><td>Wednesday</td></tr> <tr><td>3</td><td>eat__3</td><td>Thursday</td></tr> <tr><td>4</td><td>eat__4</td><td>Friday</td></tr> <tr><td>5</td><td>eat__5</td><td>Saturday</td></tr> <tr><td>6</td><td>eat__6</td><td>Sunday</td></tr> </table>	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
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 <table border="1"> <tr><td>0</td><td>drink__0</td><td>Monday</td></tr> <tr><td>1</td><td>drink__1</td><td>Tuesday</td></tr> <tr><td>2</td><td>drink__2</td><td>Wednesday</td></tr> <tr><td>3</td><td>drink__3</td><td>Thursday</td></tr> <tr><td>4</td><td>drink__4</td><td>Friday</td></tr> <tr><td>5</td><td>drink__5</td><td>Saturday</td></tr> <tr><td>6</td><td>drink__6</td><td>Sunday</td></tr> </table>	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
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: <i>Other information</i> Specify the patient's mood.	slider Slider labels: Very sad, Indifferent, Very happy Custom alignment: RH																												
	22	any_statis	Is the patient taking any statins?	radio <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>		0	No	1	Yes																							
0	No																															
1	Yes																															
	23	which_statis	Which statin(s) is the patient taking?	checkbox <table><tr><td>1</td><td>which_statis__1</td><td>Atorvastatin</td></tr><tr><td>2</td><td>which_statis__2</td><td>Cerivastatin</td></tr><tr><td>3</td><td>which_statis__3</td><td>Fluvastatin</td></tr><tr><td>4</td><td>which_statis__4</td><td>Lovastatin</td></tr><tr><td>5</td><td>which_statis__5</td><td>Mevastatin</td></tr><tr><td>6</td><td>which_statis__6</td><td>Pitavastatin</td></tr><tr><td>7</td><td>which_statis__7</td><td>Pravastatin</td></tr><tr><td>8</td><td>which_statis__8</td><td>Rosuvastatin</td></tr><tr><td>9</td><td>which_statis__9</td><td>Simvastatin</td></tr></table>		1	which_statis__1	Atorvastatin	2	which_statis__2	Cerivastatin	3	which_statis__3	Fluvastatin	4	which_statis__4	Lovastatin	5	which_statis__5	Mevastatin	6	which_statis__6	Pitavastatin	7	which_statis__7	Pravastatin	8	which_statis__8	Rosuvastatin	9	which_statis__9	Simvastatin
1	which_statis__1	Atorvastatin																														
2	which_statis__2	Cerivastatin																														
3	which_statis__3	Fluvastatin																														
4	which_statis__4	Lovastatin																														
5	which_statis__5	Mevastatin																														
6	which_statis__6	Pitavastatin																														
7	which_statis__7	Pravastatin																														
8	which_statis__8	Rosuvastatin																														
9	which_statis__9	Simvastatin																														
	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: <i>Form Status</i> Complete?	dropdown <table><tr><td>0</td><td>Incomplete</td></tr><tr><td>1</td><td>Unverified</td></tr><tr><td>2</td><td>Complete</td></tr></table>		0	Incomplete	1	Unverified	2	Complete																					
0	Incomplete																															
1	Unverified																															
2	Complete																															
Instrument: <b>Baseline Data</b> (baseline_data)																																
	28	date_visit_b	Section Header: <i>Baseline Measurements</i> Date of baseline visit YYYY-MM-DD	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 <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>		0	No	1	Yes																							
0	No																															
1	Yes																															

	34	plasma2_b	Collected Plasma 2?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes						
0	No													
1	Yes													
	35	serum1_b	Collected Serum 1?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes						
0	No													
1	Yes													
	36	serum2_b	Collected Serum 2?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes						
0	No													
1	Yes													
	37	sga_b	Subject Global Assessment (score = 1-7)	text (number, Min: 0.9, Max: 7.1)										
	38	date_supplement_dis pensed	Date patient begins supplement <i>YYYY-MM-DD</i>	text (date_ymd)										
	39	baseline_data_compl ete	Section Header: <i>Form Status</i> Complete?	dropdown <table><tr><td>0</td><td>Incomplete</td></tr><tr><td>1</td><td>Unverified</td></tr><tr><td>2</td><td>Complete</td></tr></table>	0	Incomplete	1	Unverified	2	Complete				
0	Incomplete													
1	Unverified													
2	Complete													
Instrument: <b>Monthly Data</b> (monthly_data)														
	40	consent_reminder	Reminder: Patient was consented on [baseline_visit_arm_1][date_enrolled].	descriptive										
	41	date_visit_m	Date of Monthly Visit <i>YYYY-MM-DD</i>	text (date_ymd)										
	42	creat_m	Creatinine (mg/dL)	text (number, Min: 0.5, Max: 20)										
	43	hdl_m	HDL Cholesterol (mg/dL)	text (number, Min: 20, Max: 100)										
	44	ldl_m	LDL Cholesterol (mg/dL)	text (number, Min: 30, Max: 300)										
	45	drywt_m	Dry weight (kilograms)	text (number, Min: 35, Max: 200)										
	46	trt_missed	Number of treatments missed <i>Out of 8 possible</i>	text (number, Min: 0, Max: 8)										
	47	compliance	How compliant was the patient in drinking the supplement?	dropdown <table><tr><td>0</td><td>100 percent</td></tr><tr><td>1</td><td>99-75 percent</td></tr><tr><td>2</td><td>74-50 percent</td></tr><tr><td>3</td><td>49-25 percent</td></tr><tr><td>4</td><td>0-24 percent</td></tr></table>	0	100 percent	1	99-75 percent	2	74-50 percent	3	49-25 percent	4	0-24 percent
0	100 percent													
1	99-75 percent													
2	74-50 percent													
3	49-25 percent													
4	0-24 percent													
	48	hosp_yn	Section Header: <i>Hospitalization Data</i> Was patient hospitalized since last visit?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes						
0	No													
1	Yes													

	49	hosp_cause  Show the field ONLY if: [hosp_yn] = '1'	What was the cause of hospitalization?	dropdown <table><tr><td>1</td><td>Vascular access related events</td></tr><tr><td>2</td><td>CVD events</td></tr><tr><td>3</td><td>Other</td></tr></table>	1	Vascular access related events	2	CVD events	3	Other
1	Vascular access related events									
2	CVD events									
3	Other									
	50	hosp_adm  Show the field ONLY if: [hosp_yn] = '1'	Date of hospital admission	text (date_ymd)						
	51	hosp_dis  Show the field ONLY if: [hosp_yn] = '1'	Date of hospital discharge	text (date_ymd)						
	52	hosp_summary_binder  Show the field ONLY if: [hosp_yn] = '1'	Discharge summary in patient's binder?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes		
0	No									
1	Yes									
	53	monthly_data_complete  Section Header: <i>Form Status</i> Complete?		dropdown <table><tr><td>0</td><td>Incomplete</td></tr><tr><td>1</td><td>Unverified</td></tr><tr><td>2</td><td>Complete</td></tr></table>	0	Incomplete	1	Unverified	2	Complete
0	Incomplete									
1	Unverified									
2	Complete									
Instrument: <b>Completion Data</b> (completion_data)										
	54	completed_study  Section Header: <i>Study Completion Information</i> Has patient completed study?		dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes		
0	No									
1	Yes									
	55	completed_date  Show the field ONLY if: [completed_study] = '1'	Date of study completion	text (date_ymd)						
	56	studywd_yn  Show the field ONLY if: [completed_study] = '0'	Did the patient ever withdraw from the study?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>	0	No	1	Yes		
0	No									
1	Yes									
	57	studywd_date  Show the field ONLY if: [studywd_yn] = '1'	Date of study withdrawal	text (date_ymd)						

	58	studywd_reason  Show the field ONLY if: [studywd_yn] = '1'	Reason patient withdrew from study	dropdown <table><tr><td>0</td><td>Non-compliance</td></tr><tr><td>1</td><td>Did not wish to continue in study</td></tr><tr><td>2</td><td>Could not tolerate the supplement</td></tr><tr><td>3</td><td>Hospitalization</td></tr><tr><td>4</td><td>Other</td></tr></table>		0	Non-compliance	1	Did not wish to continue in study	2	Could not tolerate the supplement	3	Hospitalization	4	Other
0	Non-compliance														
1	Did not wish to continue in study														
2	Could not tolerate the supplement														
3	Hospitalization														
4	Other														
	59	death_yn  Show the field ONLY if: [completed_study] = '0'	Did the patient die during the study?	dropdown <table><tr><td>0</td><td>No</td></tr><tr><td>1</td><td>Yes</td></tr></table>		0	No	1	Yes						
0	No														
1	Yes														
	60	death_date  Show the field ONLY if: [death_yn] = '1'	Date of death	text (date_ymd)											
	61	death_cause  Show the field ONLY if: [death_yn] = '1'	Cause of death	dropdown <table><tr><td>0</td><td>All cause</td></tr><tr><td>1</td><td>Cardiovascular event</td></tr></table>		0	All cause	1	Cardiovascular event						
0	All cause														
1	Cardiovascular event														
	62	study_comments	Section Header: <i>General Comments</i> Comments	notes											
	63	completion_data_complete	Section Header: <i>Form Status</i> Complete?	dropdown <table><tr><td>0</td><td>Incomplete</td></tr><tr><td>1</td><td>Unverified</td></tr><tr><td>2</td><td>Complete</td></tr></table>		0	Incomplete	1	Unverified	2	Complete				
0	Incomplete														
1	Unverified														
2	Complete														