

NLSY97

Autogenerated data summary from dataReporter

2024-04-10 14:04:45.665962

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	1168
Number of variables	21

Checks performed

The following variable checks were performed, depending on the data type of each variable:

	character	factor	labelled	haven labelled	numeric	integer	logical	Date
Identify miscoded missing values	×	×	×	×	×	×		×
Identify prefixed and suffixed whitespace	×	×	×	×				
Identify levels with < 6 obs.	×	×	×	×				
Identify case issues	×	×	×	×				
Identify misclassified numeric or integer variables	×	×	×	×				
Identify outliers					×	×		×

Please note that all numerical values in the following have been rounded to 2 decimals.

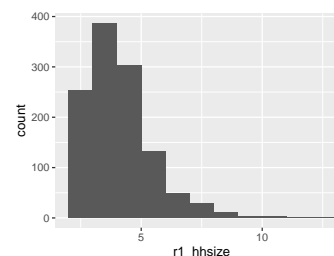
Summary table

	Variable class	# unique values	Missing observations	Any problems?
r1_hhsize	integer	11	0.00 %	×
r1_hhchildren	integer	9	0.00 %	×
r1_parents	character	2	0.00 %	
r1_urban	character	2	0.00 %	
r1_race	character	2	0.00 %	
r1_mcollege	character	2	0.00 %	
r1_mage	integer	32	0.00 %	×
r1_age	integer	3	0.00 %	×
r1_sex	factor	2	0.00 %	
r1_veg	integer	8	0.00 %	
r1_exercise	integer	8	0.00 %	
r1_health	factor	2	0.00 %	
r1_depressed	factor	2	0.00 %	
r6_veg	numeric	7	0.00 %	×
r6_exercise	integer	8	0.00 %	
r6_health	factor	2	0.00 %	
r6_depressed	factor	2	0.00 %	
r6_docvisits	numeric	5	0.00 %	
r12_health	factor	2	0.00 %	
r12_depressed	factor	2	0.00 %	
r12_docvisits	numeric	5	0.00 %	

Variable list

r1_hhsize

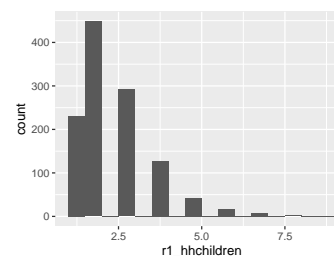
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	11
Median	4
1st and 3rd quartiles	4; 5
Min. and max.	2; 13



- Note that the following possible outlier values were detected: "2", "3", "13".

r1_hhchildren

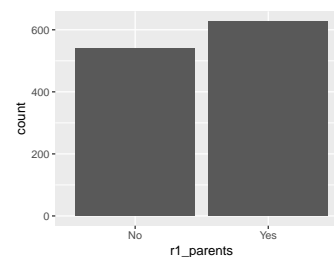
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	9
Median	2
1st and 3rd quartiles	2; 3
Min. and max.	1; 9



- Note that the following possible outlier values were detected: "1".

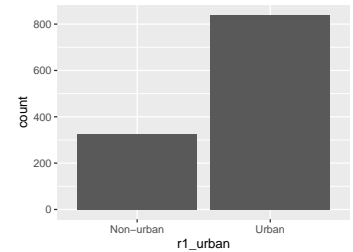
r1_parents

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"Yes"



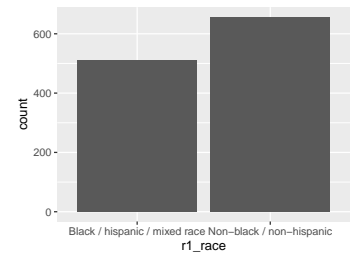
r1__urban

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Urban”



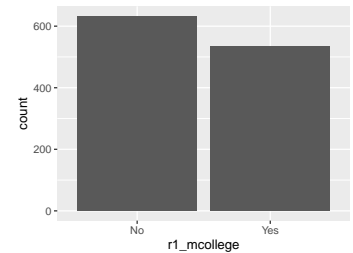
r1__race

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Non-black / non-hispanic”



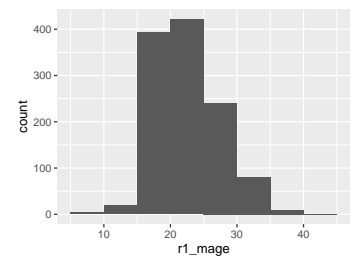
r1__mcollege

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“No”



r1__mage

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	32
Median	22
1st and 3rd quartiles	19; 26
Min. and max.	5; 42

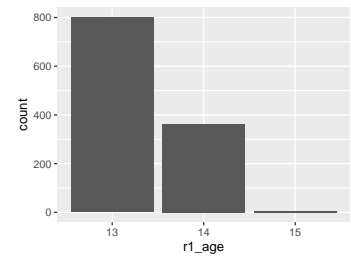


- Note that the following possible outlier values were detected: "5", "8", "9", "11", "12", "13", "14".

r1__age

- Note that this variable is treated as a factor variable below, as it only takes a few unique values.

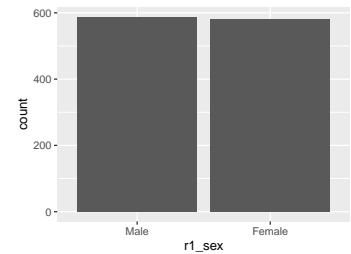
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"13"
Reference category	13



- Note that the following levels have at most five observations: "15".

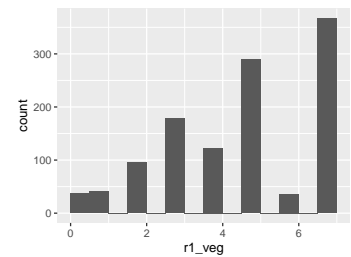
r1_sex

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"Male"
Reference category	Male



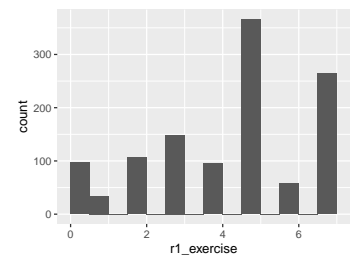
r1_veg

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	8
Median	5
1st and 3rd quartiles	3; 7
Min. and max.	0; 7



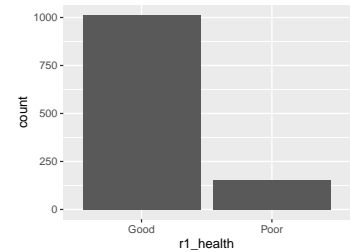
r1_exercise

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	8
Median	5
1st and 3rd quartiles	3; 6
Min. and max.	0; 7



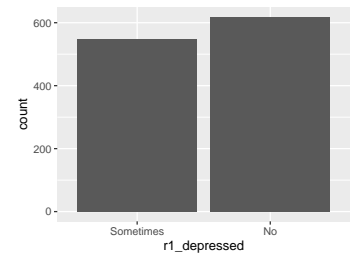
r1__health

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"Good"
Reference category	Good



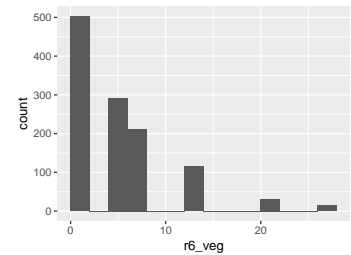
r1__depressed

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	Sometimes



r6__veg

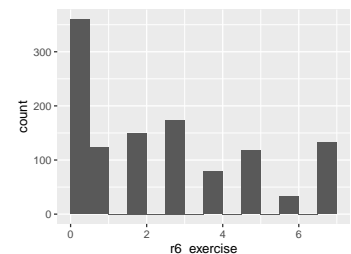
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	7
Median	5
1st and 3rd quartiles	2; 7
Min. and max.	0; 28



- Note that the following possible outlier values were detected: "14", "21", "28".

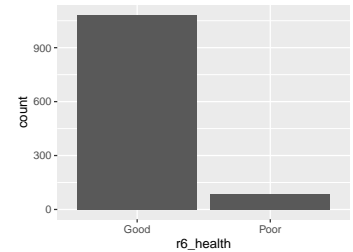
r6__exercise

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	8
Median	2
1st and 3rd quartiles	0; 4
Min. and max.	0; 7



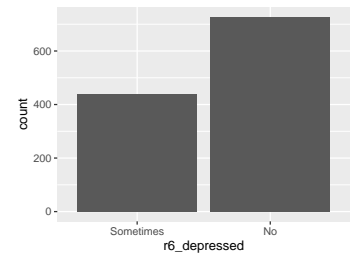
r6__health

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Good”
Reference category	Good



r6__depressed

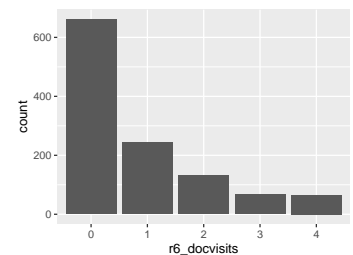
Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“No”
Reference category	Sometimes



r6__docvisits

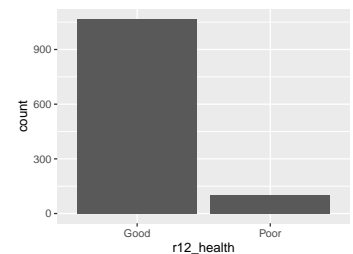
- Note that this variable is treated as a factor variable below, as it only takes a few unique values.

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	“0”
Reference category	0



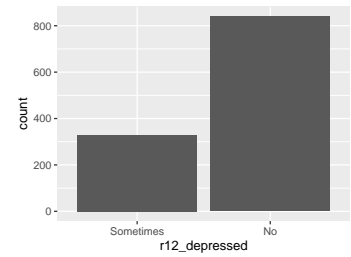
r12__health

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Good”
Reference category	Good



r12_depressed

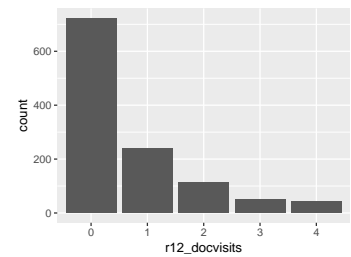
Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	Sometimes



r12_docvisits

- Note that this variable is treated as a factor variable below, as it only takes a few unique values.

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	"0"
Reference category	0



Report generation information:

- Created by: Anne Helby Petersen (username: zms499).
- Report creation time: Wed Apr 10 2024 14:04:45
- Report was run from directory: /Users/zms499/Dropbox/bscopiesDB/causalDisco
- dataReporter v1.0.2 [Pkg: 2021-11-11 from CRAN (R 4.3.0)]
- R version 4.3.2 (2023-10-31).
- Platform: aarch64-apple-darwin20 (64-bit)(Europe/Copenhagen).
- Function call: `makeDataReport(data = NLSY97)`