

1 In the current dataset, missing values occurred for all variables except for year, age,
2 gender, urbanization and province. There were 68 different patterns of missing variables in the
3 dataset (see Table S3). Looking at the first two rows of Table S3, there were 19936 subjects
4 with all variables observed and 3189 subjects with an unobserved income. Table S3 reveals
5 that 582 participants did not report their past-month cannabis use. Due to the illicit nature of
6 cannabis use, it might be the case that users are more likely to not report their cannabis use as
7 compared to non-users. Finally, 1610 participants did not report their physical activity level.
8 As physical activity is the outcome variable, we do not expect participants with missing data
9 in any of the predictor variables to have a different distribution of physical exercise compared
10 to complete cases. The reason for this assumed independence is that we do not see any
11 reasons why participants with specific higher or lower values in certain covariates would be
12 more/less likely to report their physical activity (e.g. why people with higher education would
13 be more reluctant to report their physical activity level). Of course, this is still an assumption
14 based on intuition and subject matter knowledge, there still might be a level of dependence
15 between missingness in physical activity and any of the predictors.

Table S3

Patterns of missing data with respective frequencies

Cannabis	Physical	Education	Income	Depression	Anxiety	GALI	Frequency
X	X	X	X	X	X	X	19936
X	X	X		X	X	X	3189
X		X	X	X	X	X	1610
X		X		X	X	X	277
	X	X	X	X	X	X	582
	X	X		X	X	X	116
		X	X	X	X	X	173
		X		X	X	X	35
X	X	X	X	X		X	34
X	X	X		X		X	6
X		X	X	X		X	12
X		X		X		X	1
	X	X	X	X		X	10
	X	X		X		X	1
		X	X	X		X	12
		X		X		X	2
X	X	X	X		X	X	10
X	X	X			X	X	1
X		X	X		X	X	6
X		X			X	X	1
	X	X	X		X	X	2
		X	X		X	X	2
X	X	X	X			X	197
X	X	X				X	28
X		X	X			X	65
X		X				X	12
	X	X	X			X	87
	X	X				X	20
		X	X			X	84
		X				X	16
X	X	X	X	X	X		323
X	X	X		X	X		47
X		X	X	X	X		39
X		X		X	X		3
	X	X	X	X	X		26
	X	X		X	X		2
		X	X	X	X		16
X	X	X	X	X			1
X		X	X	X			2
	X	X	X	X			1
		X	X	X			2
X	X	X	X		X		1
X	X	X			X		1

	X	X	X		X		2
	X	X			X		1
		X	X		X		1
X	X	X	X				25
X	X	X					4
X		X	X				4
X		X					1
	X	X	X				26
	X	X					4
		X	X				37
		X					14
X	X		X	X	X	X	240
X	X			X	X	X	57
X			X	X	X	X	18
X				X	X	X	6
	X		X	X	X	X	13
	X			X	X	X	2
			X	X	X	X	3
X	X		X			X	1
X	X					X	1
	X		X			X	2
	X					X	1
						X	2
X	X		X	X	X		7
			X				1

Note. Cells containing an 'X' refers to the variable being observed, whereas empty cells reflect that the corresponding variable was not observed.