

Results for the independent dataset of Swedish only (47107 subjects), comparing the results from the original variables and from the filtered variables.

In filtered variables, the extremes are removed based on the MAD method, taking  $3 \times \text{MAD}$  as cutpoint. All the variables, original and filtered are then log and sd transformed.

Beta coefficients are calculated as before and for several different models:

- raw continuous variables(log transformed and standardized), fitted together or separately,
- continuous variables expressed in % of total energy intake, fitted together or separately,
- variables categorized based on guidelines, also fitted together or separately (most of variables have to be expressed in % of TEI to be able to apply recommendations, except for fiber and salt).

Regarding the variables that have only two categories, we discussed with Alaitz and decided it is best and used in other papers, that we set the highest score, so just 1 and 2.

Same as before, I have marked some numbers/variables:

	non-significant
	wrong direction of association
	direction of association different in other data or model

I had added the vif value for each of the nutrient variables. Might be a problem with looking just at vif, since it will not explain the relationship between nutrients and bmi, but only between nutrients, it is only providing us with the information on how much of the variation in one independent variable can be explained by the variation in others, similar as PCA, where only the variability in the independent variables is considered. Although it shows a lot of information, we might need an evaluation of how much does one independent variable influence on the effect of another independent variable on bmi. Sugar always has the lowest vif, always below 4, so the main reason why we went to look at vif is not solved, I made a specific analysis just for sugar, with carbs and fiber, but it seems that the “real” model is far more complicated as it is expected with diet.

We had noticed a quite high correlation of year with sugar, so we will check on that as well, if that might be the confounding variable.

We will also examine the sugar variable more, since the variable is sucrose that could be extracted from items like fruit that also contain a lot of fiber, vitamins and the sugar they contain is not the added sugar, so we might need to extract our own sugar variable, which will be made up of only added sugar.

Number of subjects excluded in filtering(value replaced with NA):

	number of subjects
bmi	1067
POLYsum1	1892
MONOsum1	986
mfetsum1	1151
fettsum1	968
sacksum1	2069
kolhsum1	788
FA	2185
protsum1	806
fibesum1	643
NATRsum1	975

Pairwise Pearson correlation coefficient for all variables:

**Original:**

-continuous

	bmi	POLYsum1	MONOsum1	mfetsum1	fettsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.07	1									
MONOsum1	0.09	0.80	1								
mfetsum1	0.05	0.63	0.91	1							
fettsum1	0.07	0.80	0.97	0.96	1						
sacksum1	-0.02	0.23	0.29	0.30	0.32	1					
kolhsum1	0.02**	0.43	0.46	0.45	0.49	0.70	1				
FA	0.06	0.98	0.74	0.57	0.75	0.24	0.43	1			
protsum1	0.10	0.60	0.73	0.69	0.74	0.34	0.69	0.56	1		
fibesum1	0.00 n.s.	0.35	0.24	0.20	0.26	0.41	0.78	0.37	0.52	1	
NATRsum1	0.11	0.68	0.81	0.71	0.79	0.32	0.63	0.64	0.88	0.48	1

all p-values < 2e-16, except where marked

- % of TEI

	bmi	POLYsum1	MONOsum1	mfetsum1	fettsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.04	1									
MONOsum1	0.07	0.56	1								
mfetsum1	0.01 n.s.	0.16	0.74	1							
fettsum1	0.04	0.53	0.90	0.89	1						
sacksum1	-0.07	-0.29	-0.36	-0.30	-0.36	1					
kolhsum1	-0.07	-0.48	-0.82	-0.77	-0.88	0.50	1				
FA	0.02	0.97	0.45	0.08	0.44	-0.24	-0.40	1			
protsum1	0.09	-0.05	0.05	-0.02	-0.02**	-0.38	-0.23	-0.08	1		
fibesum1	-0.06	-0.14	-0.58	-0.62	-0.62	0.08	0.63	-0.08	-0.01 .	1	
NATRsum1	0.09	0.21	0.37	0.08	0.22	-0.37	-0.34	0.17	0.54	-0.10	1

all p-values < 2e-16, except where marked

- % of TEI and categorized:

	bmi	POLYsum1	MONOsum1	mfetsum1	fettsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.03	1									
MONOsum1	0.04	0.32	1								
mfetsum1	0.00 n.s.	0.07	0.45	1							
fettsum1	0.04	0.38	0.44	0.42	1						
sacksum1	-0.02**	-0.15	-0.13	-0.05	-0.12	1					
kolhsum1	-0.05	-0.36	-0.46	-0.36	-0.70	0.21	1				
FA	0.00 n.s.	-0.24	-0.27	-0.09	-0.18	0.12	0.18	1			
protsum1	0.05	0.01 n.s.	0.04	0.03	0.03	-0.12	-0.12	-0.01 .	1		
fibesum1	0.00 n.s.	0.04	0.20	0.18	0.21	0.02**	-0.25	-0.02	0.02	1	
NATRsum1	0.06	0.08	0.13	0.07	0.11	-0.05	-0.13	-0.06	0.07	-0.32	1

all p-values < 2e-16, except where marked

**Filtered:****-continuous**

	bmi	POLYsum1	MONOsum1	mfetsum1	fettsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.06	1									
MONOsum1	0.07	0.78	1								
mfetsum1	0.04	0.61	0.90	1							
fettsum1	0.05	0.77	0.97	0.96	1						
sacksum1	-0.03	0.21	0.25	0.26	0.28	1					
kolhsum1	0.01 .	0.41	0.43	0.42	0.46	0.66	1				
FA	0.05	0.98	0.72	0.55	0.72	0.22	0.41	1			
protsum1	0.08	0.58	0.71	0.67	0.72	0.31	0.66	0.55	1		
fibesum1	-0.01 n.s.	0.33	0.22	0.18	0.24	0.40	0.77	0.34	0.50	1	
NATRsum1	0.08	0.67	0.79	0.68	0.77	0.29	0.61	0.63	0.87	0.45	1

all p-values &lt; 2e-16, except where marked

**- % of TEI**

	bmi	POLYsum1	MONOsum1	mfetsum1	fettsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.02	1									
MONOsum1	0.05	0.54	1								
mfetsum1	0.00 n.s.	0.17	0.74	1							
fettsum1	0.02	0.50	0.90	0.89	1						
sacksum1	-0.08	-0.25	-0.34	-0.28	-0.34	1					
kolhsum1	-0.07	-0.45	-0.81	-0.76	-0.87	0.50	1				
FA	0.01 n.s.	0.96	0.43	0.09	0.41	-0.19	-0.37	1			
protsum1	0.07	0 n.s.	0.06	0 n.s.	0 n.s.	-0.31	-0.22	-0.02	1		
fibesum1	-0.06	-0.11	-0.57	-0.60	-0.60	0.15	0.63	-0.06	-0.01 *	1	
NATRsum1	0.07	0.25	0.37	0.09	0.22	-0.33	-0.32	0.22	0.54	-0.07	1

all p-values &lt; 2e-16, except where marked

- % of TEI and categorized:

	bmi	POLYsum1	MONOsum1	mfetsum1	fetsum1	sacksum1	kolhsum1	FA	protsum1	fibesum1	NATRsum1
bmi	1										
POLYsum1	0.01**	1									
MONOsum1	0.03	0.31	1								
mfetsum1	-0.01 n.s.	0.07	0.45	1							
fetsum1	0.03	0.35	0.44	0.43	1						
sacksum1	-0.02	-0.11	-0.11	-0.04	-0.11	1					
kolhsum1	-0.05	-0.34	-0.45	-0.36	-0.69	0.17	1				
FA	0.00 n.s.	-0.25	-0.27	-0.10	-0.18	0.09	0.18	1			
protsum1	0.04	0.03	0.04	0.03	0.02	-0.08	-0.10	-0.01 .	1		
fibesum1	0.00 n.s.	0.07	0.19	0.17	0.19	0.05	-0.22	-0.03	0.03	1	
NATRsum1	0.05	0.03	0.11	0.06	0.10	-0.10	-0.14	-0.05	0.04	-0.28	1

all p-values < 2e-16, except where marked

**Original:**

Continuous, together:

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.1449063983	0.000406	1.67e-05	54.699620
MONOsum1	0.1215381828	0.000491	2.19e-06	31.793346
mfetsum1	-0.3574987856	0.00288	< 2e-16	46.786709
fetsum1	0.2139980916	0.000402	1.82e-05	120.277687
acids	-0.2130136014	0.001221	8.26e-14	39.280782
kolhsum1	-0.0890035751	0.001064	3.16e-12	7.862449
sacksum1	-0.0040728002	7e-06	0.57803	2.590965
protsum1	0.1370827090	0.003443	< 2e-16	5.751999
NATRsum1	0.0710509980	0.000705	1.61e-07	7.565204
fibesum1	-0.0467451715	0.000601	1.40e-08	3.842574

Continuous, separately:

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0423381	0.001463	< 2e-16
MONOsum1	0.0621248	0.002711	< 2e-16
mfetsum1	0.0067159	3.4e-05	0.2136436
fettsum1	0.0352077	9e-04	< 2e-16
acids	0.0271136	0.000618	1e-07
kolhsum1	-0.0169606	0.000252	0.0006936
sacksum1	-0.0361812	0.001207	< 2e-16
protsum1	0.0788836	0.005522	< 2e-16
NATRsum1	0.0893801	0.006565	< 2e-16
fibesum1	-0.0158257	0.000245	0.0008174

In % of TEI, together :

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.1134072125	0.000494	2.06e-06	27.552329
MONOsum1	0.0616900900	0.000386	2.70e-05	10.433744
mfetsum1	-0.2050340285	0.002588	< 2e-16	17.149478
fettsum1	0.1495413001	0.000617	1.12e-07	38.351869
acids	-0.1603678491	0.001274	2.41e-14	21.337461
kolhsum1	-0.0013896031	0	0.909342	7.196853
sacksum1	-0.0001253127	0	0.983514	1.777254
protsum1	0.0826511361	0.003505	< 2e-16	2.056356
NATRsum1	0.0388060272	0.000758	3.99e-09	2.100448
fibesum1	-0.0355652554	0.00054	6.80e-07	2.475774

In % of TEI, separately :

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0376752	0.001284	< 2e-16
MONOsum1	0.073881	0.00412	< 2e-16
mfetsum1	-0.0110457	0.000107	0.0273104
fettsum1	0.0310584	0.000794	< 2e-16
acids	0.0185701	0.000323	0.0001232
kolhsum1	-0.0673984	0.003576	< 2e-16
sacksum1	-0.0527268	0.002633	< 2e-16
protsum1	0.1163856	0.013094	< 2e-16
NATRsum1	0.119998	0.013742	< 2e-16
fibesum1	-0.0416894	0.001384	< 2e-16



Categorized , together:

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.0268636052	0.000177	0.00444	1.332826
MONOsum1	0.0671402558	0.000449	6.01e-06	1.677917
mfetsum1	-0.1431796539	0.001017	9.48e-12	1.426825
fettsum1	0.0270826357	9.3e-05	0.03965	2.216237
acids	0.0064082111	2e-06	0.78175	1.129169
kolhsum1	-0.0284284415	0.000106	0.02753	2.355930
sacksum1	0.0070363480	3e-06	0.70226	1.102640
protsum1	0.2789089026	0.003151	< 2e-16	1.051544
NATRsum1	0.1024416459	0.001552	< 2e-16	1.322318
fibesum1	0.0632974719	0.000416	1.32e-05	1.292019

Categorized, separately:

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0528024	0.000816	< 2e-16
MONOsum1	0.081492	0.00094	< 2e-16
mfetsum1	-0.0400857	0.000108	0.0261377
fettsum1	0.0580404	0.000805	< 2e-16
acids	-0.0514248	0.000118	0.0203299
kolhsum1	-0.0754001	0.001401	< 2e-16
sacksum1	-0.0501038	0.000172	0.0051182
protsum1	0.3054257	0.003901	< 2e-16
NATRsum1	0.1031498	0.001835	< 2e-16
fibesum1	0.0377511	0.000186	0.0035724

**Filtered:**

Continuous, together:

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.1332515333	0.000432	4.21e-05	41.139349
MONOsum1	0.1256428552	0.000554	3.49e-06	27.254756
mfetsum1	-0.2789078574	0.001851	< 2e-16	41.977475
fettsum1	0.0894252781	7.2e-05	0.09439	104.673059
acids	-0.1720493110	0.001034	2.31e-10	29.617289
kolhsum1	-0.1290649867	0.002277	< 2e-16	6.990843
sacksum1	-0.0051017559	1.2e-05	0.49221	2.256766
protsum1	0.1531503124	0.004523	< 2e-16	5.027202
NATRsum1	0.0501441523	0.000377	0.00013	6.531625
fibesum1	-0.0145896449	6e-05	0.12796	3.648175

Continuous, separately:

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0205667	0.000354	9.96e-05
MONOsum1	0.0238687	0.000415	2.08e-05
mfetsum1	-0.0204346	0.000324	0.0001726
fettsum1	-0.0011675	1e-06	0.8328029
acids	0.0082973	5.9e-05	0.1135637
kolhsum1	-0.0405721	0.001492	< 2e-16
sacksum1	-0.0541652	0.002829	< 2e-16
protsum1	0.0465586	0.001975	< 2e-16
NATRsum1	0.051625	0.002255	< 2e-16
fibesum1	-0.0281763	0.000792	< 2e-16

In % of TEI, together :

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.0967068190	0.000429	4.42e-05	22.426197
MONOsum1	0.0707355453	0.000498	1.10e-05	10.286581
mfetsum1	-0.1714636535	0.001798	< 2e-16	17.256575
fettsum1	0.1049341231	0.000291	0.000767	38.413457
acids	-0.1247529004	0.000962	9.78e-10	17.339819
kolhsum1	-0.0066147720	6e-06	0.642304	7.264042
sacksum1	-0.0029175219	5e-06	0.646735	1.630010
protsum1	0.1013027267	0.004966	< 2e-16	1.984517
NATRsum1	0.0297053013	0.000447	3.09e-05	1.999623
fibesum1	-0.0107351267	4.8e-05	0.172433	2.434628

In % of TEI, separately :

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0349117	0.001114	< 2e-16
MONOsum1	0.062117	0.002974	< 2e-16
mfetsum1	-0.0153618	0.000211	0.0024376
fettsum1	0.0211722	0.000376	5.03e-05
acids	0.0161282	0.000247	0.0011986
kolhsum1	-0.0706291	0.003938	< 2e-16
sacksum1	-0.0616909	0.003655	< 2e-16
protsum1	0.1139682	0.012661	< 2e-16
NATRsum1	0.1113482	0.01198	< 2e-16
fibesum1	-0.0368214	0.001081	< 2e-16

Categorized , together:

	Regression coefficient	Variance explained	p-value	vif
POLYsum1	0.0291771254	0.000197	0.00539	1.277787
MONOsum1	0.0472704385	0.000223	0.00307	1.666736
mfetsum1	-0.1424103586	0.00106	1.09e-10	1.374794
fettsum1	0.0107470581	5.3e-05	0.14958	1.671534
acids	-0.0160848910	1e-05	0.52957	1.120706
kolhsum1	-0.0356343562	0.000175	0.00874	2.220868
sacksum1	-0.0329351395	4.9e-05	0.16723	1.063103
protsum1	0.2992861487	0.003034	< 2e-16	1.040594
NATRsum1	0.0401826567	0.000217	0.00353	1.223322
fibesum1	0.0672866265	0.000408	6.18e-05	1.240388

Categorized, separately:

	Regression coefficient	Variance explained	p-value
POLYsum1	0.0485437	0.000638	2e-07
MONOsum1	0.065841	0.000624	2e-07
mfetsum1	-0.0446384	0.000139	0.013827
fettsum1	0.0340113	0.000802	< 2e-16
acids	-0.073732	0.000257	0.0009471
kolhsum1	-0.0795254	0.001555	< 2e-16
sacksum1	-0.0818461	0.00029	0.000441
protsum1	0.2792612	0.003042	< 2e-16
NATRsum1	0.070242	0.000859	< 2e-16
fibesum1	0.0512166	0.000348	8.2e-05

