Be Well Program: Baseline And 12 Months Post Program Measurements Analysis

The Data

Variable list:

1. HbA1C: hba1c
2. Fasting Blood Glucose: fast_blood_gluc
3. Blood Pressure (Systolic): blood_press_sys
4. Weight
5. BMI
6. LDL
7. Triglycerides (indicators 13, 14, 17, 18, 20, 22, 25)
8. Vegetable Dervings : pa_21
9. Fruit Servings :pa_22
10. Soda Consumption: pa_23
11. Sugar-Sweetened Beverages Consumption: pa_24
12. Attitudes Toward Healthy Foods : healthy_eating_important

We will be renaming the time points in the 'time' variable as follows only for more convenience when analyzing the data:

'1' for baseline measurement

'2' for the end of program assessment

'3' for the 6-month mark after the program

'4' for the 12-month mark after the program

'5' for the 24-month mark after the program

In this analysis, we will only be focusing on the baseline and 12 months after the program measurements.

Quick Inventory Check

The data has been filtered to include only: - People who have data for both baseline and EOP - Observations, for each health metric, that were recorded in both baseline and EOP. (Eg: if someone had only baseline triglycerides levels recorded, that observation is disregarded in this analysis by being turned into NA)

Total sample size:

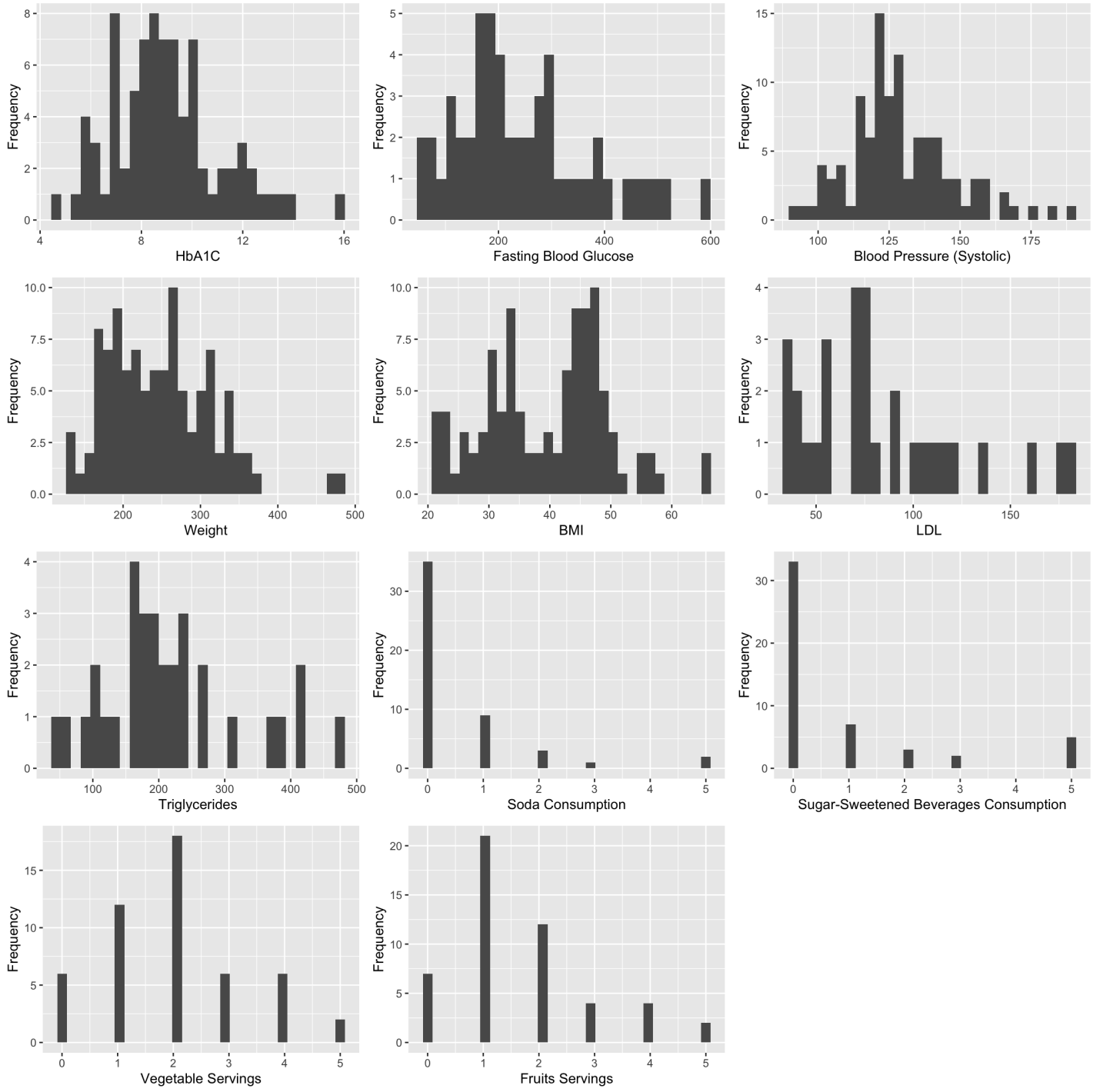
[1] 110

Sample size for high participation: [1] 72

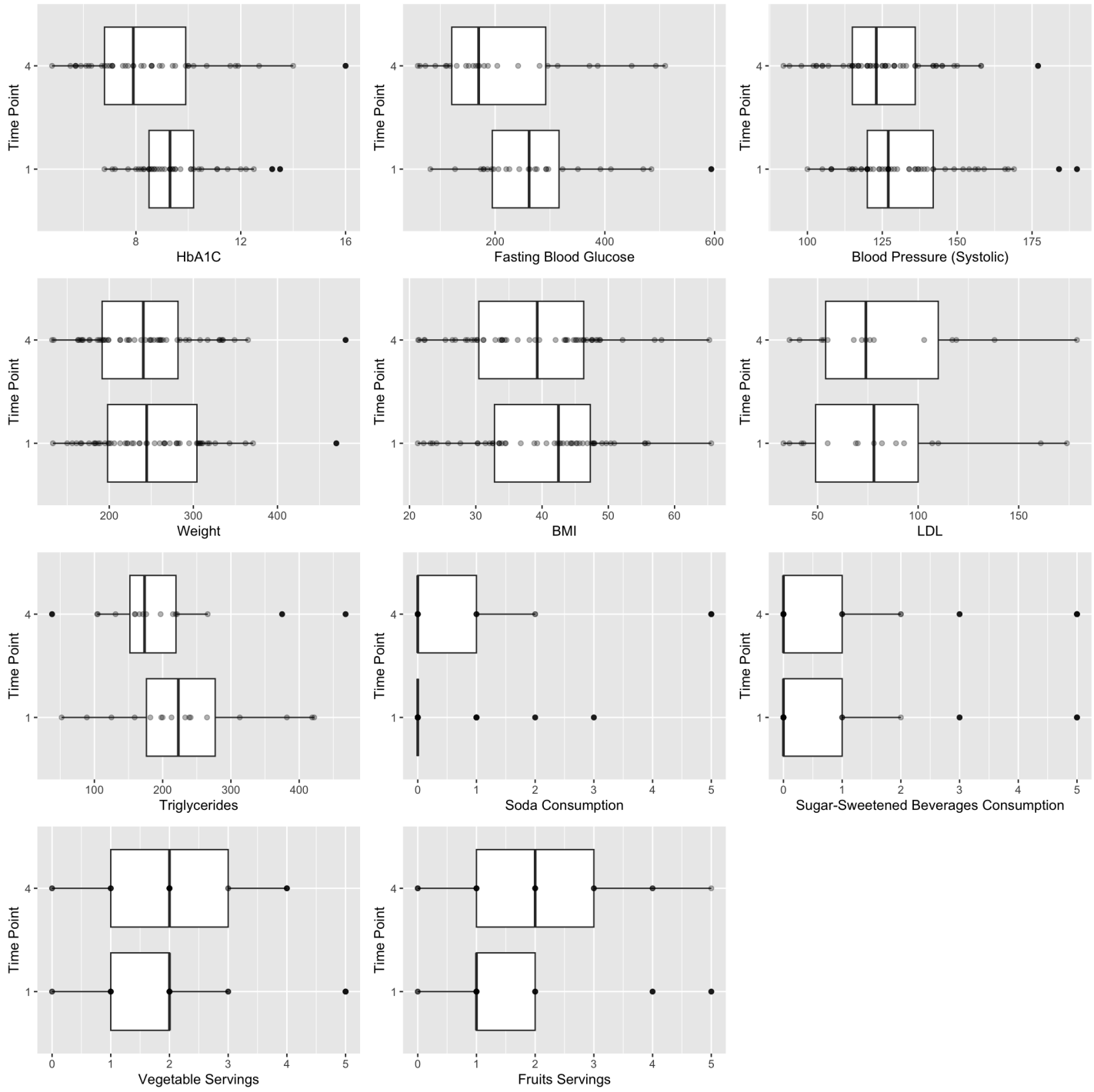
Plots and Graphs

Distributions of The Variables

In assessing the health program's impact, distribution plots were generated for key variables, combining data from both Time 1 (baseline) and Time 4 (12 months after the program). This approach offers a comprehensive view of participant characteristics, facilitating a nuanced understanding of the program's potential influence. The distribution plots can aid in visually narrating the dataset's overall evolution, laying the groundwork for further detailed explorations of specific variables across the program's timeline.



Box Plots: The Distribution of Each Variable at Each Time Point:

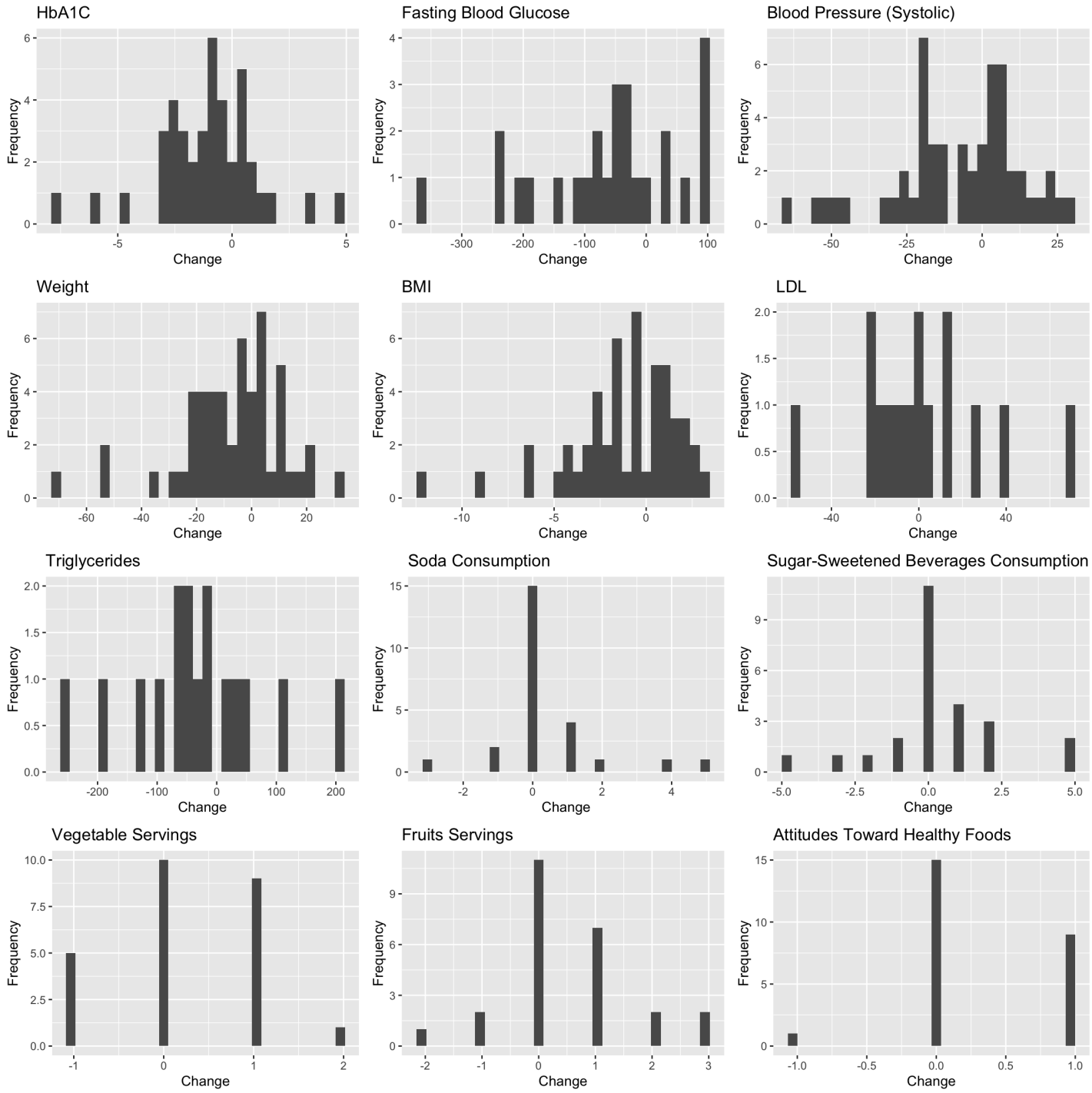


The box plots reaffirm the trends observed in mean values, with the addition of quartiles trends, offering a detailed portrayal of the data distribution.

One noteworthy observation from the box plots is the reduction in extreme outliers at 12 months after the program compared to the baseline. These outliers, indicative of poorer health conditions, appear to diminish post-program participation. This reduction underscores a potential mitigation of severe health issues among participants, aligning with the overarching goal of the health program.

This visual confirmation strengthens our earlier findings and reinforces the notion that the health program may be instrumental in fostering healthier conditions among participants.

The Distribution of Rates of Change in the Variables Before and 12 months after the program



This analysis reveals a favorable trend, with the majority of variables showcasing decreases rather than increases.

For health attributes that we aim to decrease, such as HbA1c and blood sugar levels, the prevalence and extremity of decreases in the rates of change are particularly encouraging. Although there are instances of participants recording higher measurements in certain health attributes, there are also even more notable decreases observed in the majority of variables.

This asymmetry in the distribution suggests that, on the whole, participants are experiencing improvements in key health indicators rather than deteriorations.

Summary Tables

These concise and generalized tables offer convenient summaries of the aforementioned graphs, encapsulating both the mean values of the variables and their respective distributions.

The Means of Each Variable Before and 12 months after the program:

Variables legend:

fast_blood_gluc: Fasting Blood Glucose

blood_press_sys: Blood Pressure (Systolic)

pa_23: Soda Consumption

pa_24: Sugar-Sweetened Beverages Consumption

pa_21: Vegetable Servings

pa_22: Fruits Servings

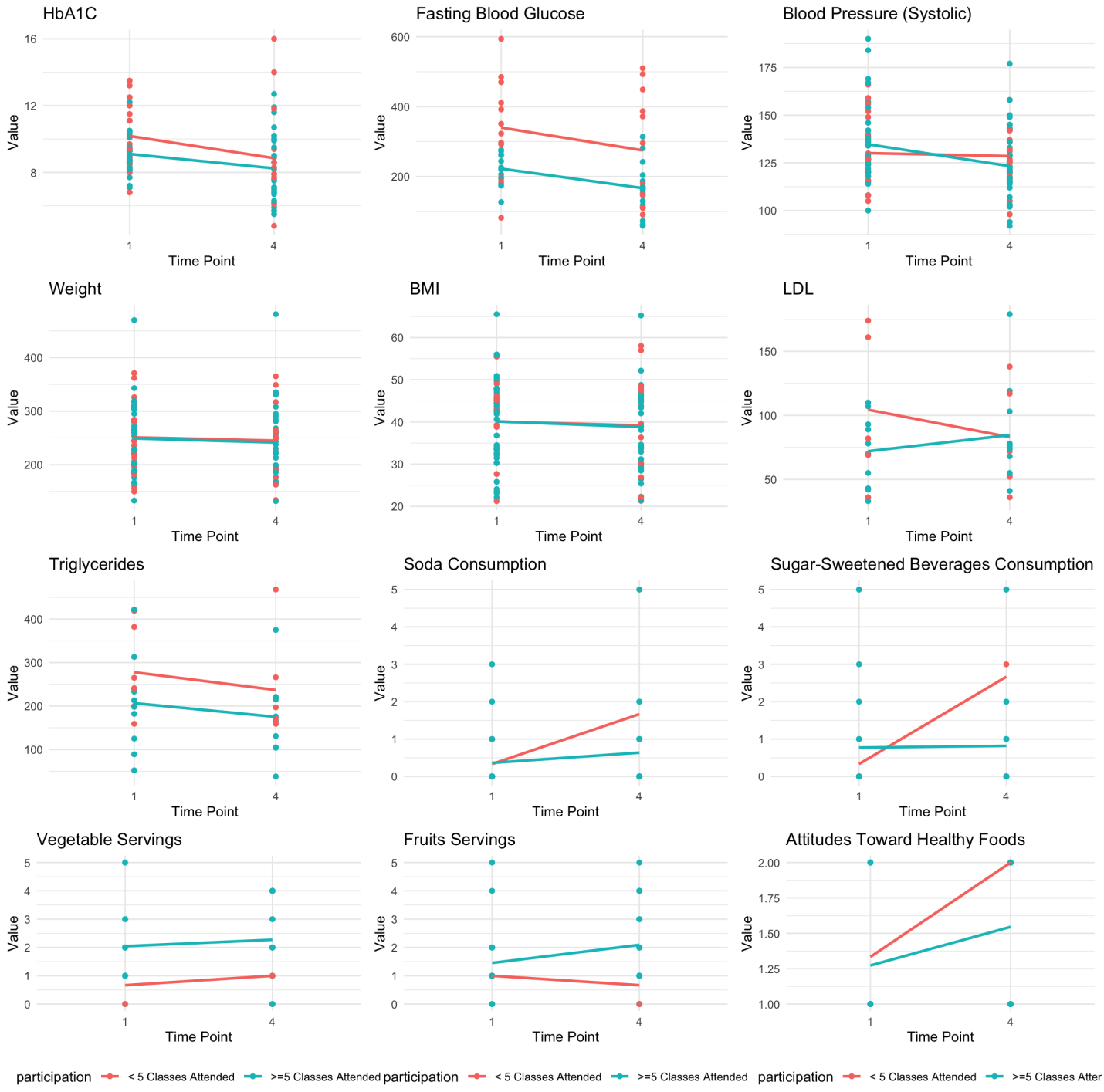
healthy_eating_important: Attitudes Toward Healthy Foods

time	hba1c	fast_blood_gluc	blood_press_sys	weight	bmi	ldl	triglycerides	pa_23	pa_24	pa_21	pa_22
1	9.50	276.79	133.08	249.85	40.13	82.80	233.25	0.36	0.72	1.88	1.40
4	8.47	216.73	125.17	242.75	38.94	84.07	198.19	0.76	1.04	2.12	1.92

The Standard Deviations of Each Variable Before and 12 months after the program:

time	hba1c	fast_blood_gluc	blood_press_sys	weight	bmi	ldl	triglycerides	pa_23	pa_24	pa_21	pa_22
1	1.60	117.47	19.53	66.94	9.97	42.23	108.32	0.76	1.49	1.30	1.12
4	2.46	132.09	17.06	68.39	10.22	39.82	104.21	1.42	1.70	1.33	1.41

Comparison Graphs Between Participation Levels

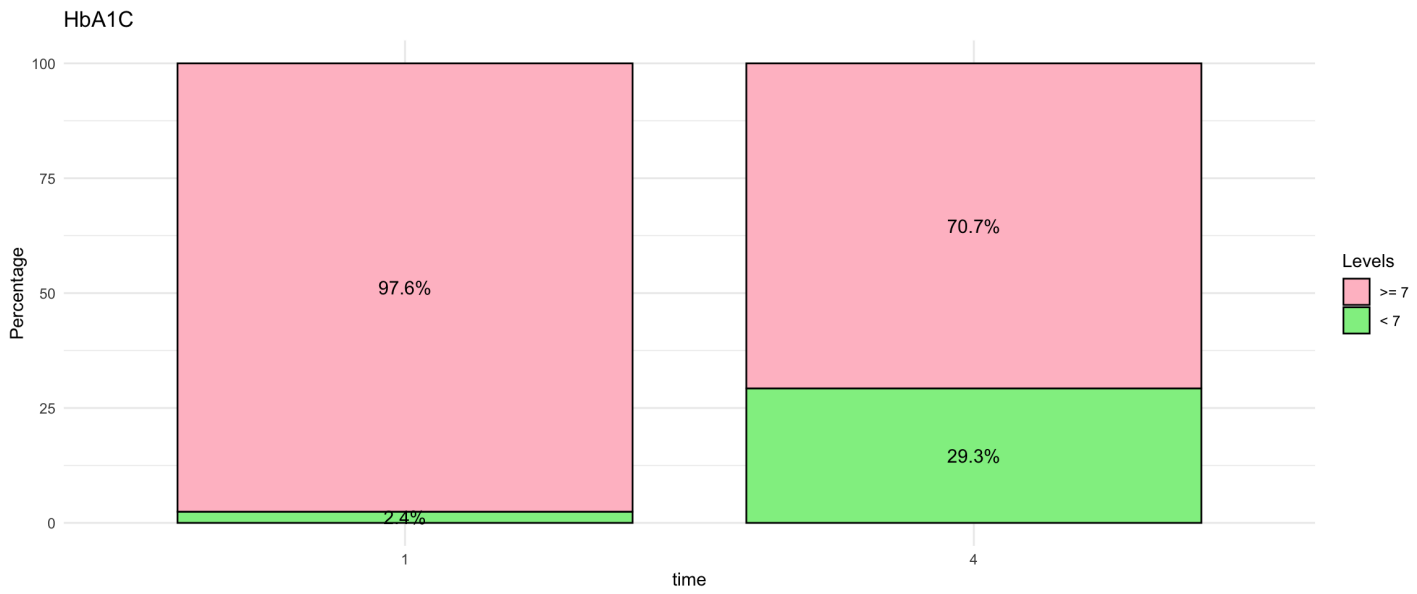


Investigating With Thresholds

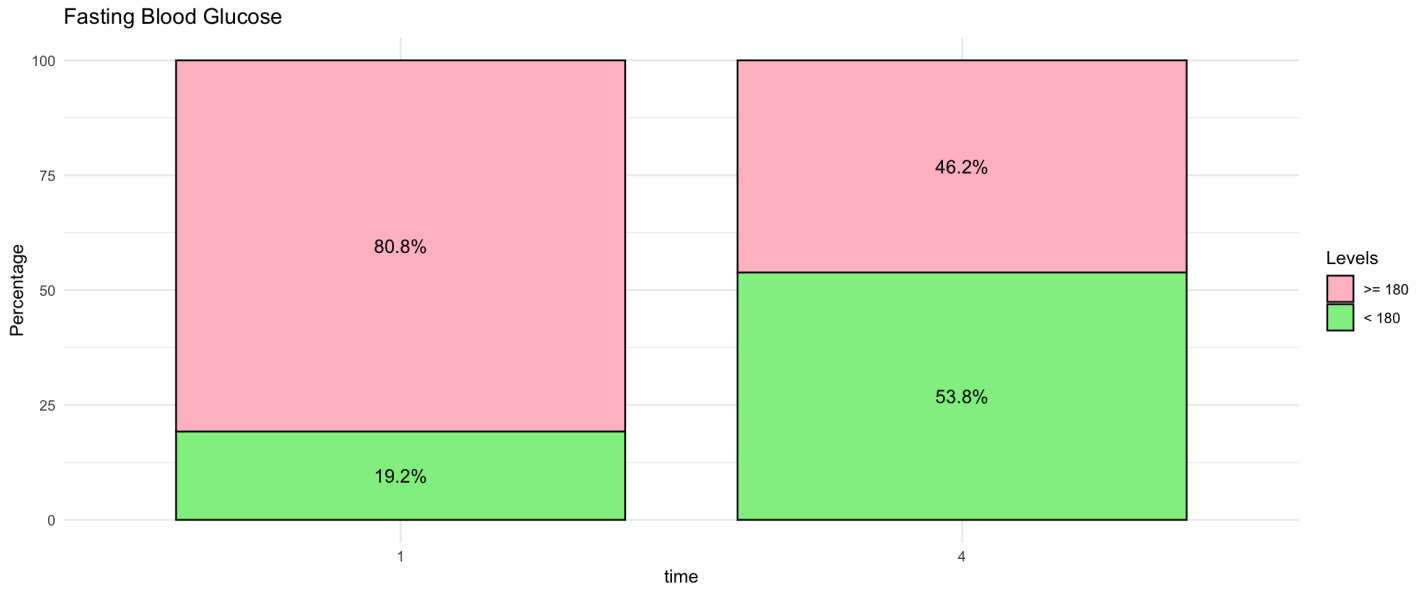
This section will visualize the number of people, if observed, whose certain health metrics are under or over (and equal to) a certain threshold at baseline and 12 months after the program. We will be using bar graphs for this purpose. Due to the nature of the function used in creating these graphs, there will be p-values and equations at the top of the graphs. This information will not be one of our focus points because this analysis only involves data visualization and not predictive modeling.

The data sets used to generate these graphs will be them same ones used for the mean plots to ensure consistent samples sizes for the 2 time points.

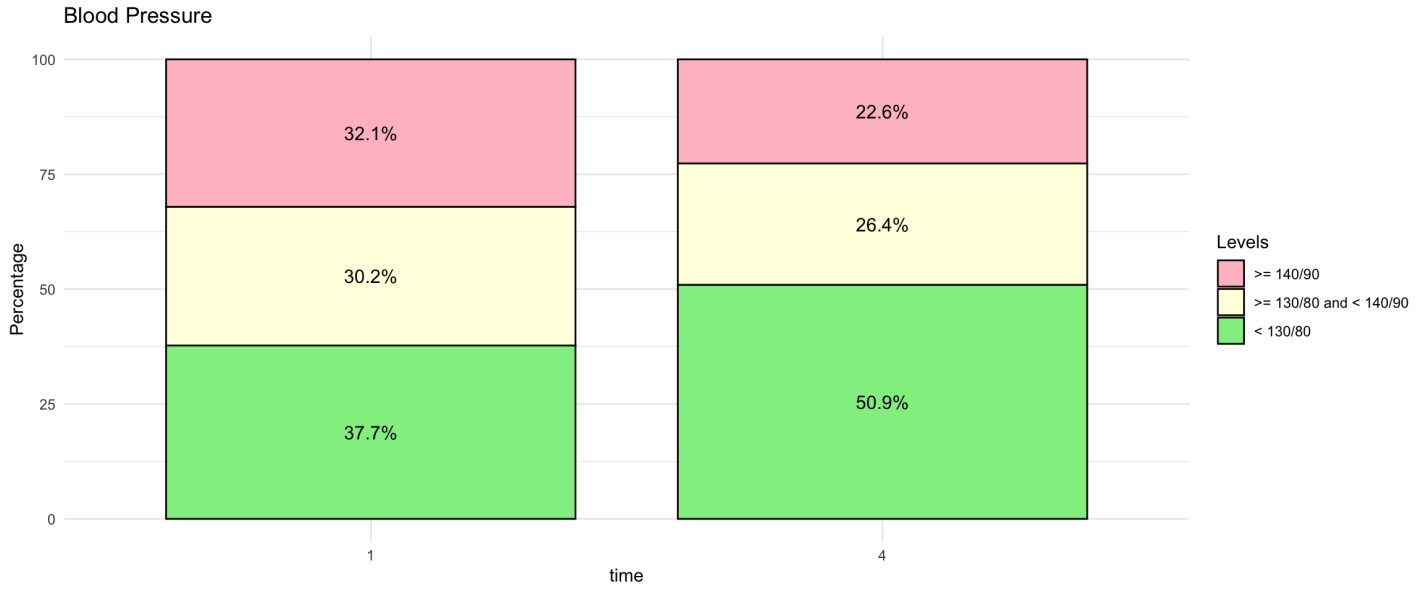
HbA1c



Fasting Blood Glucose



Blood Pressure



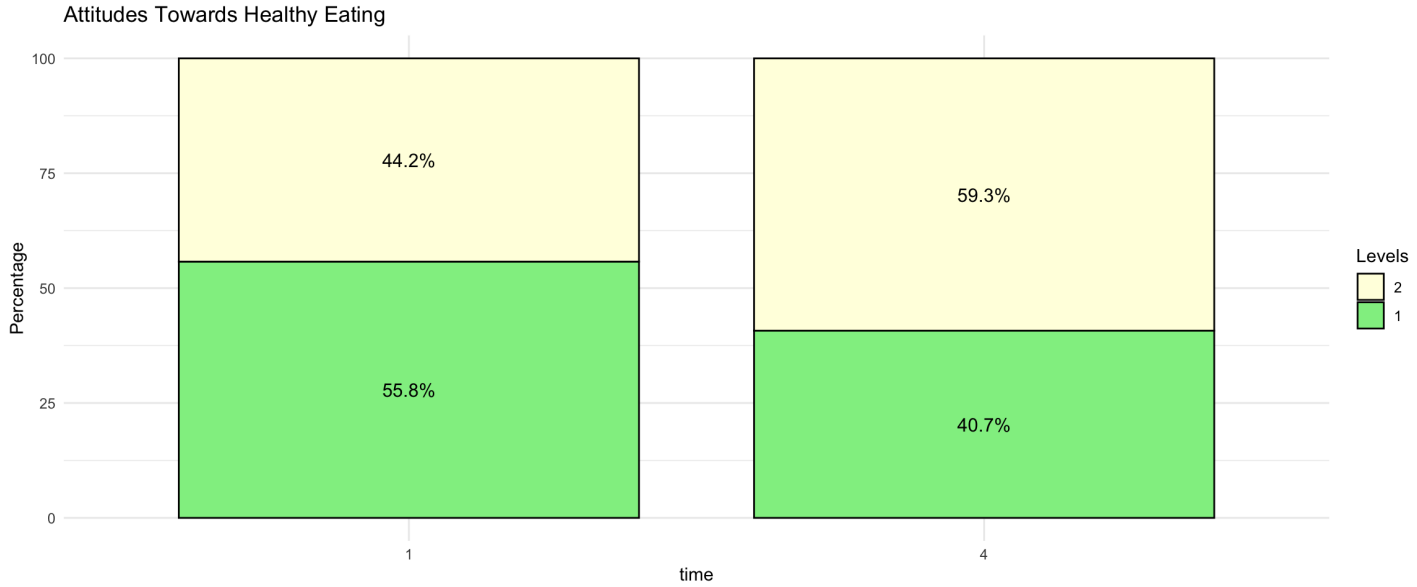
Note: 130/80 is the threshold for diabetics with pre-existing cardiovascular risk.

All three graphs depict a positive indication that the program potentially contributed to an increase in the proportion of individuals whose health metrics align with healthier thresholds.

Attitudes Toward Healthy Foods

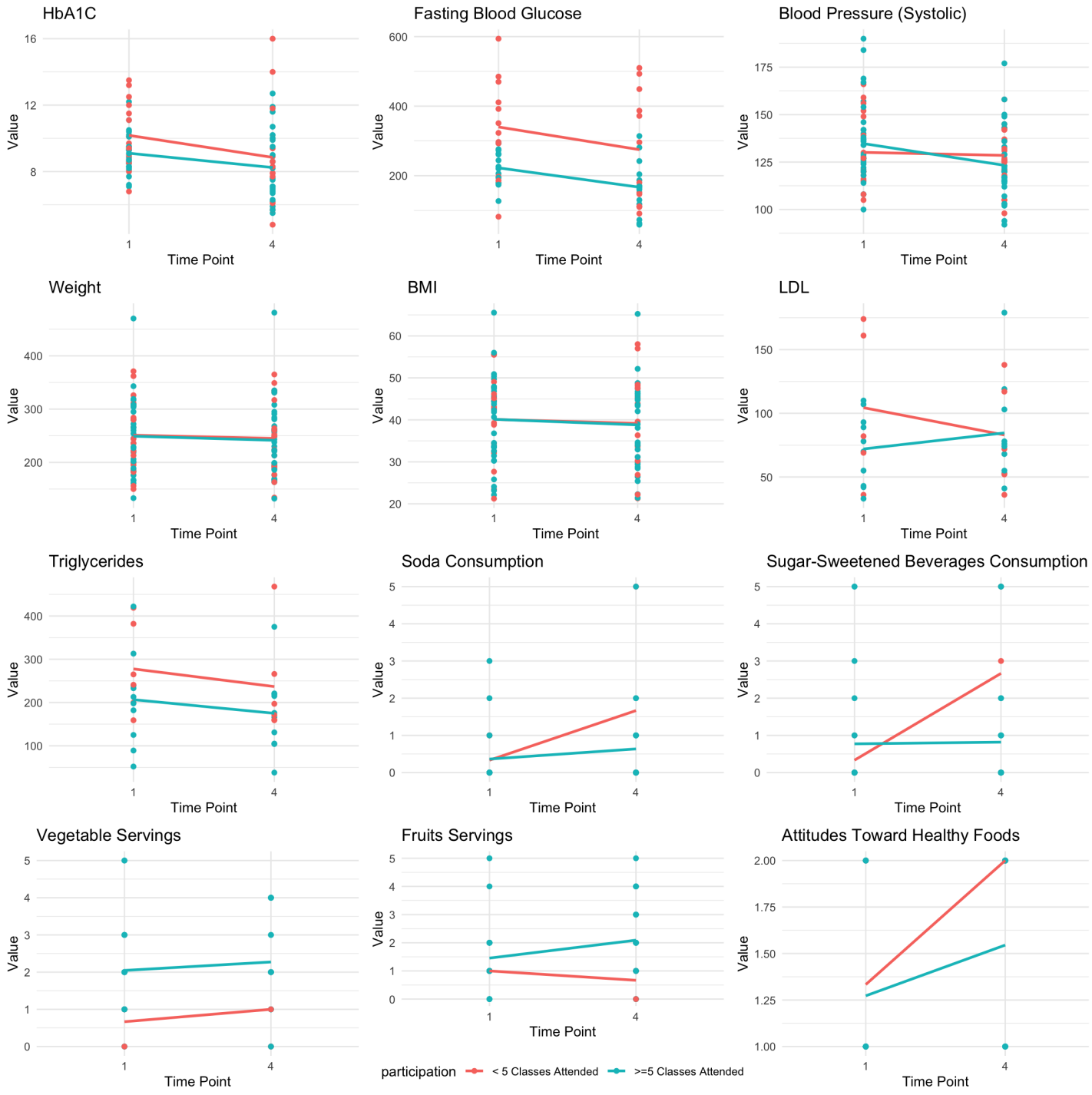
We have delved into various health attributes and dietary intake variables, uncovering notable trends and shifts. As we pivot our focus, the next step involves examining a variable that captures individuals' perspectives on healthy eating.

Given that "Attitude Towards Healthy Eating" is represented by only two values, a transition to a visual representation, such as a bar graph, becomes particularly relevant. This graphical approach allows us to illustrate the proportions of each attitude value both at the baseline and the End-of-Program (EOP) marks. By doing so, we aim to provide a clear and insightful depiction of how participants' attitudes toward healthy foods may have evolved after the health program.



Comparison Between Participation Levels (Number of

Classes Attended)



Linear Mixed Effects Models

The models aim to provide rates of change in the response variable from the baseline to the target time point.

We have chosen to treat the ‘time’ variable as a categorical factor rather than a numeric variable, even though it has been transformed into sequential numbers (1, 2, 3, 4, 5), for several reasons. First, our decision is rooted in the recognition that the rates of change in the response variables over time may not be linear and continuous. Representing ‘time’ as a numeric variable could potentially mislead by implying a linear relationship between time points and response changes that might not accurately reflect the underlying dynamics.

Secondly, treating ‘time’ categorically helps avoid extrapolation or interpolation beyond the provided time points. Using ‘time’ as a factor acknowledges that the observations at our specific time points are the only reliable data we have, and we should refrain from assuming relationships outside these intervals. This approach can be helpful in maintaining the integrity of our findings and prevent unwarranted assumptions about the program’s effects at unobserved time points.

In our model, we will have time points 1 and 2 for baseline and EOP. The summaries of our models will be presented in forms of tables that include the following columns that we will be analyzing further with each variable: - effect: this column indicates whether a variable is a fixed or random effect. In our analysis, the investigated variables will be the fixed effect and the participant id will be the random effect, accounting for the relationship between different observations of the same participant. - term: the coefficient, or in this case, the time point (baseline and EOP as Intercept and time2) - estimate: the average difference in the investigated variable between baseline and EOP - p.value: suggests whether there is evidence that the observed change is based on pure chance or potential effects from the program.

Interaction Terms and Their General Interpretations

- ‘time’ and ‘participation’: the `participation` variable was derived from the `classes_attended` variable, with values categorized as ‘>=5 Classes Attended’ and ‘<5 Classes Attended’. The interaction term between the `time` variable and the `participation` variable captures whether the effect of time on the outcome differs depending on the level of participation in the program. A significant interaction would imply that the relationship between time and the outcome varies depending on the level of participation.
- ‘time’ and ‘classes_attended’: the interaction term between the `time` variable and the `classes_attended` variable quantifies whether the effect of time on the outcome changes based on the number of classes attended. Essentially, it examines whether individuals who attended more classes experienced different changes in their outcomes over time compared to those who attended fewer classes.

The p-values associated with these interaction terms indicate the strength of evidence for these relationship; lower p-values suggest stronger evidence for an interaction effect between each pair of variables.

P-Value Scale:

For convenience, we can refer to this P-Value scale when determining if an observed trend statistically has strong evidence or is possibly due to random chance.

$p < 0.001$: Extremely strong evidence.

$0.001 \leq p < 0.01$: Very strong evidence.

$0.01 \leq p < 0.05$: Strong evidence.

$0.05 \leq p < 0.1$: Moderate evidence.

$p \geq 0.1$: Not considered statistically significant.

HbA1C

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	10.1800000	0.5303397	19.1952447	66.90838	0.000000
fixed	NA	time4	-1.3266667	0.5774848	-2.2973189	39.00000	0.027052
fixed	NA	participation>=5 Classes Attended	-1.0761538	0.6659772	-1.6159019	66.90838	0.110822
fixed	NA	time4:participation>=5 Classes Attended	0.4651282	0.7251800	0.6413969	39.00000	0.525017

effect	group	term	estimate	std.error	t-value	df	p.value
random	record_id	sd__(Intercept)	1.3106249	NA	NA	NA	NA

$$\text{HbA1c} = 10.18 - 1.3267 \times \text{time4} \quad \text{(for participation < 5 classes attended)}$$
$$\text{HbA1c} = 9.1038 - 0.8616 \times \text{time4} \quad \text{(for participation} \geq \text{5 classes attended)}$$
$$\text{p-value (interaction)} = 0.525$$

For participants who attended fewer than 5 classes, the change in HbA1c over time is represented by the equation $10.18 - 1.3267 \times \text{time4}$. However, for participants who attended 5 or more classes, the change in HbA1c over time is represented by the equation $9.1038 - 0.8616 \times \text{time4}$. The interaction between time4 and participation status (5 or more classes attended) is not statistically significant ($p = 0.525$), indicating that the relationship between time4 and HbA1c does not differ significantly based on the number of classes attended.

Fasting Blood Glucose

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	340.108333	32.71267	10.3968389	39.89092	0.000000
fixed	NA	time4	-65.191667	34.28228	-1.9016143	24.00000	0.069292
fixed	NA	participation>=5 Classes Attended	-117.594048	44.57985	-2.6378295	39.89092	0.011838
fixed	NA	time4:participation>=5 Classes Attended	9.534524	46.71887	0.2040829	24.00000	0.840009
random	record_id	sd__(Intercept)	76.090582	NA	NA	NA	NA

Blood Pressure (Systolic)

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	130.105263	4.211403	30.8935680	89.58241	0.000000
fixed	NA	time4	-1.578947	4.718605	-0.3346216	51.00000	0.739283
fixed	NA	participation>=5 Classes Attended	4.630031	5.258057	0.8805593	89.58241	0.380913
fixed	NA	time4:participation>=5 Classes Attended	-9.862229	5.891313	-1.6740289	51.00000	0.100248
random	record_id	sd__(Intercept)	11.201011	NA	NA	NA	NA

Weight

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	251.052632	15.674777	16.0163452	51.95458	0.000000
fixed	NA	time4	-6.052632	4.341467	-1.3941444	50.00000	0.169439
fixed	NA	participation>=5 Classes Attended	-1.901116	19.676419	-0.0966190	51.95458	0.923401
fixed	NA	time4:participation>=5 Classes Attended	-1.644338	5.449808	-0.3017241	50.00000	0.764114
random	record_id	sd__(Intercept)	67.001601	NA	NA	NA	NA

BMI

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	40.1127778	2.4043193	16.6836320	50.17944	0.000000
fixed	NA	time4	-0.9444444	0.7166277	-1.3179011	48.00000	0.193792
fixed	NA	participation>=5 Classes Attended	0.0215972	3.0053991	0.0071861	50.17944	0.994295
fixed	NA	time4:participation>=5 Classes Attended	-0.3821181	0.8957846	-0.4265736	48.00000	0.671597
random	record_id	sd__(Intercept)	9.9715356	NA	NA	NA	NA

LDL

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	104.40000	18.32560	5.696948	15.73779	0.000035
fixed	NA	time4	-21.40000	11.37041	-1.882079	13.00000	0.082401
fixed	NA	participation>=5 Classes Attended	-32.40000	22.44419	-1.443581	15.73779	0.168468
fixed	NA	time4:participation>=5 Classes Attended	34.00000	13.92585	2.441503	13.00000	0.029683
random	record_id	sd__(Intercept)	36.82286	NA	NA	NA	NA

Triglycerides

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	277.666667	42.65061	6.5102619	23.87815	0.000001
fixed	NA	time4	-40.833333	46.11492	-0.8854690	14.00000	0.390859
fixed	NA	participation>=5 Classes Attended	-71.066667	53.94923	-1.3172878	23.87815	0.200244
fixed	NA	time4:participation>=5 Classes Attended	9.233333	58.33127	0.1582913	14.00000	0.876488
random	record_id	sd__(Intercept)	67.340114	NA	NA	NA	NA

Soda Consumption

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.3333333	0.6567109	0.5075800	45.46027	0.614202
fixed	NA	time4	1.3333333	0.8766728	1.5209020	23.00000	0.141914
fixed	NA	participation>=5 Classes Attended	0.0303030	0.7000562	0.0432866	45.46027	0.965663
fixed	NA	time4:participation>=5 Classes Attended	-1.0606061	0.9345363	-1.1349009	23.00000	0.268106
random	record_id	sd__(Intercept)	0.3754663	NA	NA	NA	NA

Sugar-Sweetened Beverages Consumption

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.3333333	0.9027723	0.3692330	44.17679	0.713718
fixed	NA	time4	2.3333333	1.1396746	2.0473681	23.00000	0.052211
fixed	NA	participation>=5 Classes Attended	0.4393939	0.9623585	0.4565803	44.17679	0.650210
fixed	NA	time4:participation>=5 Classes Attended	-2.2878788	1.2148972	-1.8831872	23.00000	0.072383
random	record_id	sd__(Intercept)	0.7047739	NA	NA	NA	NA

Vegetable Servings

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.6666667	0.7322037	0.9104935	28.69504	0.370148
fixed	NA	time4	0.3333333	0.4894584	0.6810249	23.00000	0.502655
fixed	NA	participation>=5 Classes Attended	1.3787879	0.7805318	1.7664725	28.69504	0.087954
fixed	NA	time4:participation>=5 Classes Attended	-0.1060606	0.5217644	-0.2032730	23.00000	0.840709
random	record_id	sd__(Intercept)	1.1175923	NA	NA	NA	NA

Fruits Servings

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	1.0000000	0.7220111	1.3850204	34.23106	0.175007
fixed	NA	time4	-0.3333333	0.6567109	-0.5075800	23.00000	0.616581
fixed	NA	participation>=5 Classes Attended	0.4545455	0.7696664	0.5905747	34.23106	0.558685
fixed	NA	time4:participation>=5 Classes Attended	0.9696970	0.7000562	1.3851702	23.00000	0.179289
random	record_id	sd__(Intercept)	0.9575991	NA	NA	NA	NA

Interaction Term with Numerical Variable ‘classes_attended’

The interaction between time and classes_attended suggests that the effect of time on the outcome variable varies depending on the number of classes attended.

HbA1C

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	10.5409355	0.5541310	19.022460	68.20124	0.000000
fixed	NA	time4	-1.3867491	0.6175299	-2.245639	39.00000	0.030468
fixed	NA	classes_attended	-0.1639018	0.0718159	-2.282249	68.20124	0.025599

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	time4:classes_attended	0.0557728	0.0800324	0.696878	39.00000	0.490012
random	record_id	sd__(Intercept)	1.2344589	NA	NA	NA	NA

Fasting Blood Glucose

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	353.9158417	35.189914	10.0573091	40.78148	0.000000
fixed	NA	time4	-56.9746174	37.877188	-1.5041934	24.00000	0.145579
fixed	NA	classes_attended	-14.1219147	5.079226	-2.7803278	40.78148	0.008180
fixed	NA	time4:classes_attended	-0.5645067	5.467101	-0.1032552	24.00000	0.918618
random	record_id	sd__(Intercept)	71.6105542	NA	NA	NA	NA

Blood Pressure (Systolic)

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	129.7883851	4.518909	28.7211753	89.34373	0.000000
fixed	NA	time4	-0.2452210	5.046733	-0.0485900	51.00000	0.961436
fixed	NA	classes_attended	0.5108962	0.583297	0.8758766	89.34373	0.383446
fixed	NA	time4:classes_attended	-1.1906255	0.651428	-1.8277161	51.00000	0.073445
random	record_id	sd__(Intercept)	11.2424999	NA	NA	NA	NA

Weight

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	254.6847144	16.831515	15.1314195	51.96229	0.000000
fixed	NA	time4	-6.9198054	4.670864	-1.4814830	50.00000	0.144753
fixed	NA	classes_attended	-0.7578468	2.179937	-0.3476461	51.96229	0.729510
fixed	NA	time4:classes_attended	-0.0276208	0.604948	-0.0456582	50.00000	0.963765
random	record_id	sd__(Intercept)	66.9265962	NA	NA	NA	NA

BMI

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	41.0100051	2.5934944	15.8126443	50.19396	0.000000
fixed	NA	time4	-1.2688929	0.7755302	-1.6361618	48.00000	0.108347
fixed	NA	classes_attended	-0.1363280	0.3328277	-0.4096052	50.19396	0.683838
fixed	NA	time4:classes_attended	0.0123291	0.0995251	0.1238797	48.00000	0.901928
random	record_id	sd__(Intercept)	9.9548906	NA	NA	NA	NA

LDL

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	119.150160	24.067969	4.950570	15.01894	0.000174
fixed	NA	time4	-40.651757	12.957792	-3.137244	13.00000	0.007862
fixed	NA	classes_attended	-5.242812	3.128085	-1.676045	15.01894	0.114416
fixed	NA	time4:classes_attended	6.045927	1.684109	3.589986	13.00000	0.003294
random	record_id	sd__(Intercept)	37.372436	NA	NA	NA	NA

Triglycerides

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	278.145349	55.396489	5.0209924	23.60755	0.000041
fixed	NA	time4	-21.272093	59.077281	-0.3600723	14.00000	0.724167
fixed	NA	classes_attended	-6.651163	7.212009	-0.9222345	23.60755	0.365736
fixed	NA	time4:classes_attended	-2.043023	7.691207	-0.2656310	14.00000	0.794395
random	record_id	sd__(Intercept)	69.452741	NA	NA	NA	NA

Soda Consumption

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.6495256	0.6571837	0.9883470	45.74709	0.328182
fixed	NA	time4	1.1238139	0.8941788	1.2568112	23.00000	0.221437
fixed	NA	classes_attended	-0.0327518	0.0698184	-0.4690992	45.74709	0.641227
fixed	NA	time4:classes_attended	-0.0818794	0.0949965	-0.8619204	23.00000	0.397629
random	record_id	sd__(Intercept)	0.3077801	NA	NA	NA	NA

Sugar-Sweetened Beverages Consumption

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.6225895	0.9298477	0.6695608	44.7846	0.506575
fixed	NA	time4	1.5958066	1.2018159	1.3278295	23.0000	0.197263
fixed	NA	classes_attended	0.0110193	0.0987859	0.1115471	44.7846	0.911681
fixed	NA	time4:classes_attended	-0.1443220	0.1276795	-1.1303462	23.0000	0.269981
random	record_id	sd__(Intercept)	0.6482052	NA	NA	NA	NA

Vegetable Servings

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	0.7124273	0.7454028	0.9557615	28.55418	0.347215
fixed	NA	time4	0.4104683	0.4925897	0.8332864	23.00000	0.413253
fixed	NA	classes_attended	0.1320784	0.0791907	1.6678520	28.55418	0.106277
fixed	NA	time4:classes_attended	-0.0192837	0.0523321	-0.3684876	23.00000	0.715880
random	record_id	sd__(Intercept)	1.1318767	NA	NA	NA	NA

Fruits Servings

effect	group	term	estimate	std.error	t-value	df	p.value
fixed	NA	(Intercept)	1.5488215	0.7271167	2.1300867	32.2123	0.040898
fixed	NA	time4	-0.9709213	0.6046562	-1.6057411	23.0000	0.121975
fixed	NA	classes_attended	-0.0168350	0.0772480	-0.2179347	32.2123	0.828855
fixed	NA	time4:classes_attended	0.1686563	0.0642379	2.6254926	23.0000	0.015120
random	record_id	sd__(Intercept)	1.0101230	NA	NA	NA	NA