

Sugarbowl CGM Analyzer

Glycemic Increment

- ▶ JEN BALASI
 - ▶ LA MONTE YARROLL
- 2025-09-14



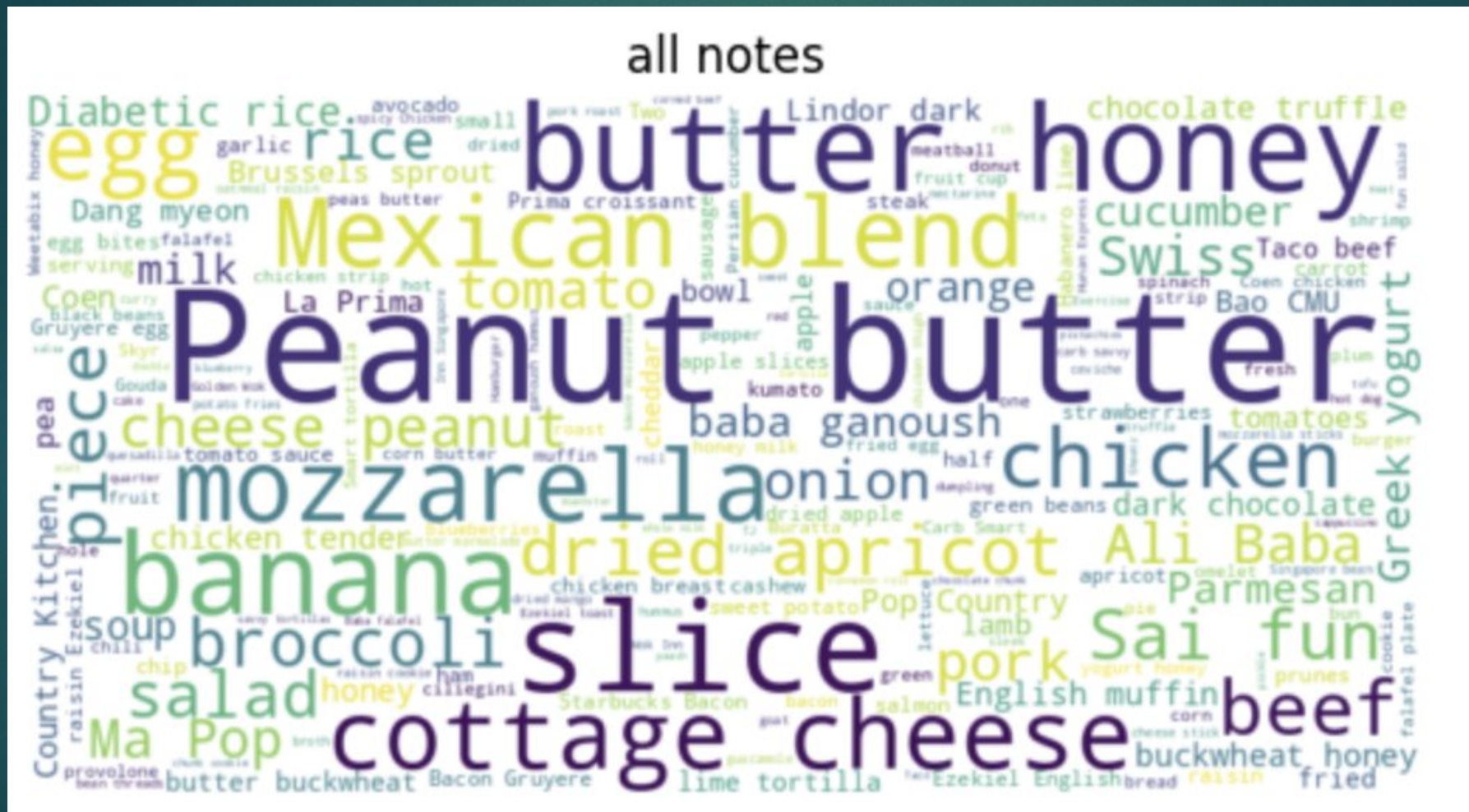
What is a Continuous Glucose Monitor?



Dataset

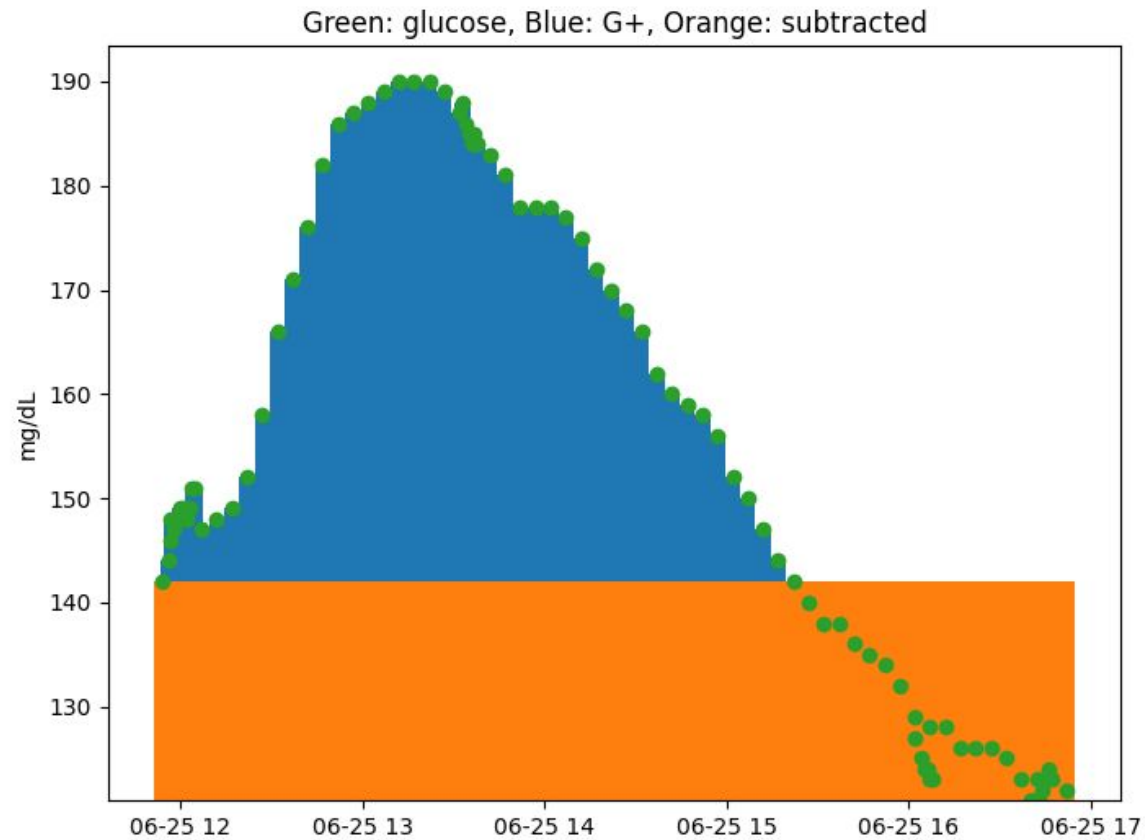
- Our dataset is a near-continuous record of glucose levels every 5 minutes for 2 years, along with free-form text descriptions of most of the meals.

```
Glucose Data,Generated on,07-20-2025 10:36 PM UTC,Generated by,La Monte Yarroll
Device,Serial Number,Device Timestamp,Record Type,Historic Glucose mg/dL,Scan Glucose mg/dL,Non-numeric Rapid-Acting Insulin,Rapid-Acting Insulin (units),Non-numeric Food,Carbohydrates (grams),Carbohydrates (servings),Non-numeric Long-Acting Insulin,Long-Acting Insulin (units),Notes,Strip Glucose mg/dL,Ketone mmol/L,Meal Insulin (units),Correction Insulin (units),User Change Insulin (units)
...
[brand],[sensor id],06-24-2024 03:42 PM,0,194,,,,,,,,,,,,,
[brand],[sensor id],06-24-2024 03:47 PM,0,195,,,,,,,,,,,,,
[brand],[sensor id],06-24-2024 03:52 PM,0,194,,,,,,,,,,,,,
[brand],[sensor id],06-24-2024 03:57 PM,0,186,,,,,,,,,,,,,
[brand],[sensor id],06-24-2024 04:02 PM,0,178,,,,,,,,,,,,,
...
[brand],[sensor id],07-16-2025 12:01 PM,6,,,,,,,,,"Konjac pad Thai, mozzarella.",,,,,,
[brand],[sensor id],07-16-2025 01:39 PM,6,,,,,,,,,"Ma & Pop's Country Kitchen: cup of chili, Reuben omelet.",,,,,,
[brand],[sensor id],07-16-2025 07:13 PM,6,,,,,,,,,"2 dosa, 3 pieces of lamb, spiced masoor dal.",,,,,,
[brand],[sensor id],07-16-2025 10:14 PM,6,,,,,,,,,"1 slice of prosciutto, large bowl of cottage cheese, honey.",,,,,,
[brand],[sensor id],07-17-2025 12:00 AM,6,,,,,,,,,"5 baby bananas, 6 dried apricots.",,,,,,
[brand],[sensor id],07-17-2025 11:53 AM,6,,,,,,,,,"Ma & Pop's Country Kitchen: Capi omelette, strawberries.",,,,,,
...
```

Glycemic Increment

Glycemic Increment (G+) is new for our study. It's the area under the glucose curve for the duration of the measurement period, minus the starting glucose, standardized to 1 hour. The unit is mg-hours/dL.



Derived datasets: Meals

Regular meals (every note is a meal)

- **Total meals: 1874**
- **Duration: 2 hours from note**

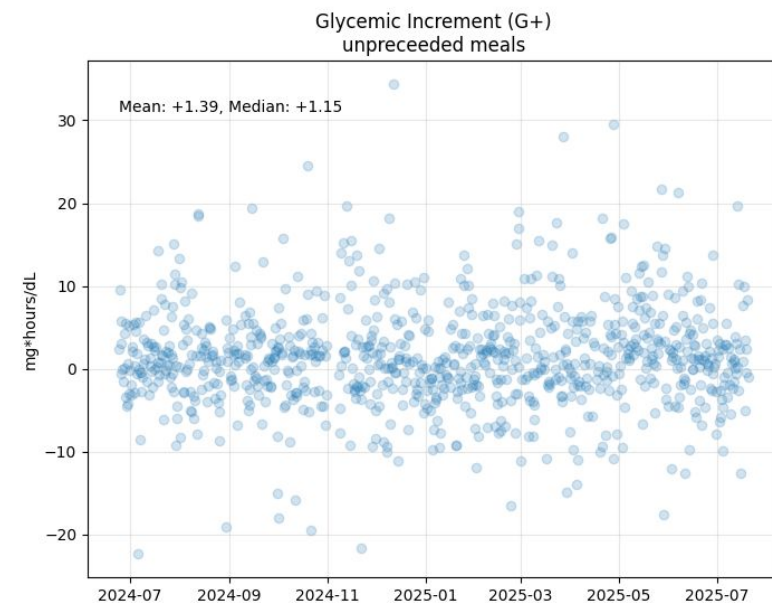
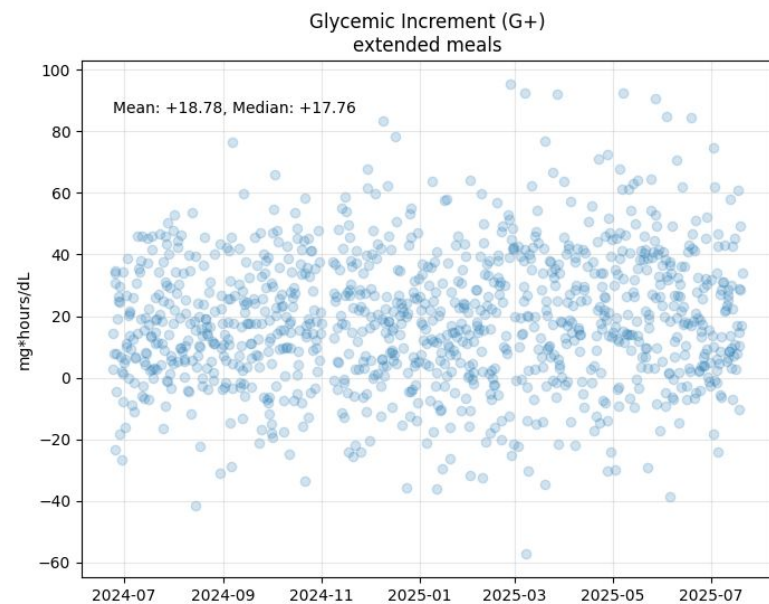
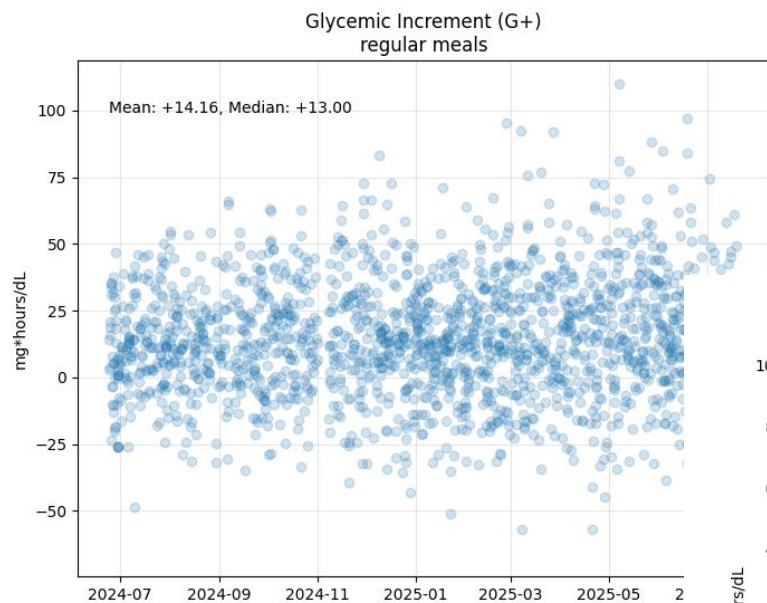
Extended meals (close notes make a meal)

- **Total meals: 1071**
- **Close: < 2 hours**
- **Duration: time of first note to 2 hours past last note**

Unpreceeded meals (no recent note)

- **Total meals: 889**
- **Fasting: > 4 hours**
- **Duration: 30 minutes from note**

Derived datasets: Meals



Models

Tokenizer

- **CountVectorizer's ngram_range=(1,3)**
- **Full Token (Restaurant Name: food1 no bun, 2 servings of food2.)**

Vectorizer

- **CountVectorizer binary**
- **QuantifierVectorizer (Uses full token parser)**

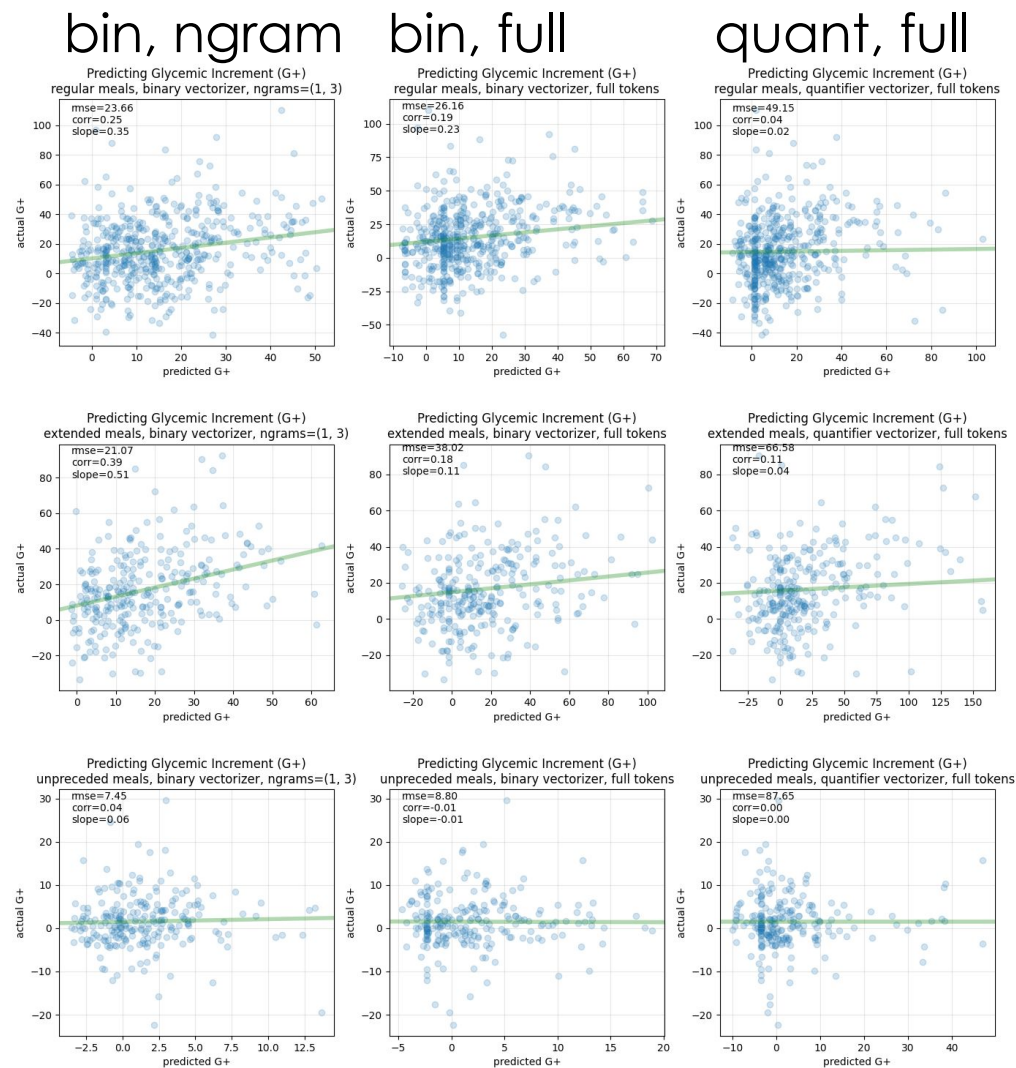
meal type	vectorizer	tokenizer
regular	Binary	ngrams
regular	Binary	full
regular	Quantifier	full
extended	Binary	ngrams
extended	Binary	full
extended	Quantifier	full
unpreceeded	Binary	ngrams
unpreceeded	Binary	full
unpreceeded	Quantifier	full

Models on Test Sets

regular meals

extended meals

unprecedented meals



Low score foods (eat these)

regular meals, binary vectorizer, ngrams=(1, 3)

low_coef	low_food
-0.0024	noodle butter dill
-0.0024	miracle noodle butter
-0.0024	miracle noodle
-0.0024	butter dill eggs
-0.0024	noodle
0.0081	apricots prunes
0.0081	dried apricots prunes
0.0130	wendy dave triple
0.0130	dave triple no
0.0130	triple no

regular meals, binary vectorizer, full tokens

low_coef	low_food
-0.1516	triple burger no bun
-0.1516	chipotle bbq sauce
-0.1205	toasted ezeziel english muffin
-0.1205	buratta
-0.0008	poblano
0.0261	scrambled eggs
0.0261	braised bed
0.0627	fried jalapeno
0.0627	eggs sunny side up
0.0758	whole milk cottage cheese

extended meals, binary vectorizer, ngrams=(1, 3)

low_coef	low_food
-0.0045	mozzarella sticks fried
-0.0045	sticks fried fish
-0.0019	shaking
-0.0016	150g of
-0.0016	150g of greek
0.0004	glaze donuts
0.0004	balsamic glaze
0.0004	truffle balsamic glaze
0.0004	skin
0.0004	skin salad

extended meals, binary vectorizer, full tokens

low_coef	low_food
-0.2822	two cherry tomatoes
-0.2599	sauerkraut
-0.1841	halloumi slices
-0.1280	campari tomatoes
-0.0527	greens
0.0458	tate's bake shop oatmeal raisin cookies
0.0924	pico
0.2293	thai sai fun salad
0.2293	hot chocolate milk sweetened
0.3625	lobster bisque

Extreme score foods (avoid these)

regular meals, binary vectorizer, ngrams=(1, 3)

high_coef	high_food
-71.7892	apricots banana
-36.6472	chicken tenders donut
-36.6472	tenders donut
-25.5440	prunes dried apricots
-21.7564	apple banana
27.5494	apple plum
36.8195	donut
45.2271	croissant
57.8512	slices raisins peanut
97.0728	dried apricots banana

regular meals, binary vectorizer, full tokens

high_coef	high_food
-66.4125	flavor bomb tomatoes
-62.3769	dairy queen dilly bar
-57.0013	RESTAURANT speedway
-56.6239	wine
-50.6041	masoor beef
74.0197	singapore noodles mixed meats
74.3032	maple bear claw
78.3518	cheese powder
83.7768	1/2 spicy chicken tenders
92.9081	apple streusel

extended meals, binary vectorizer, ngrams=(1, 3)

high_coef	high_food
-26.5724	cruller
-13.1386	thai beef
-13.1386	thai beef sai
-12.9476	kumato
-12.8285	milk peanut
14.6522	butter honey butter
14.6718	honey butter
14.7849	cinnamon
27.6110	giordano
51.2691	croissant

extended meals, binary vectorizer, full tokens

high_coef	high_food
-75.5220	corn
-72.7618	pot roast
-67.0276	chicken tenders
-66.3326	peach preserves
-60.1470	cooked carrots
85.5229	gravy
93.3223	streusel
114.8571	biscuits
128.0414	chicken cassava gloop
138.9475	cinnamon roll



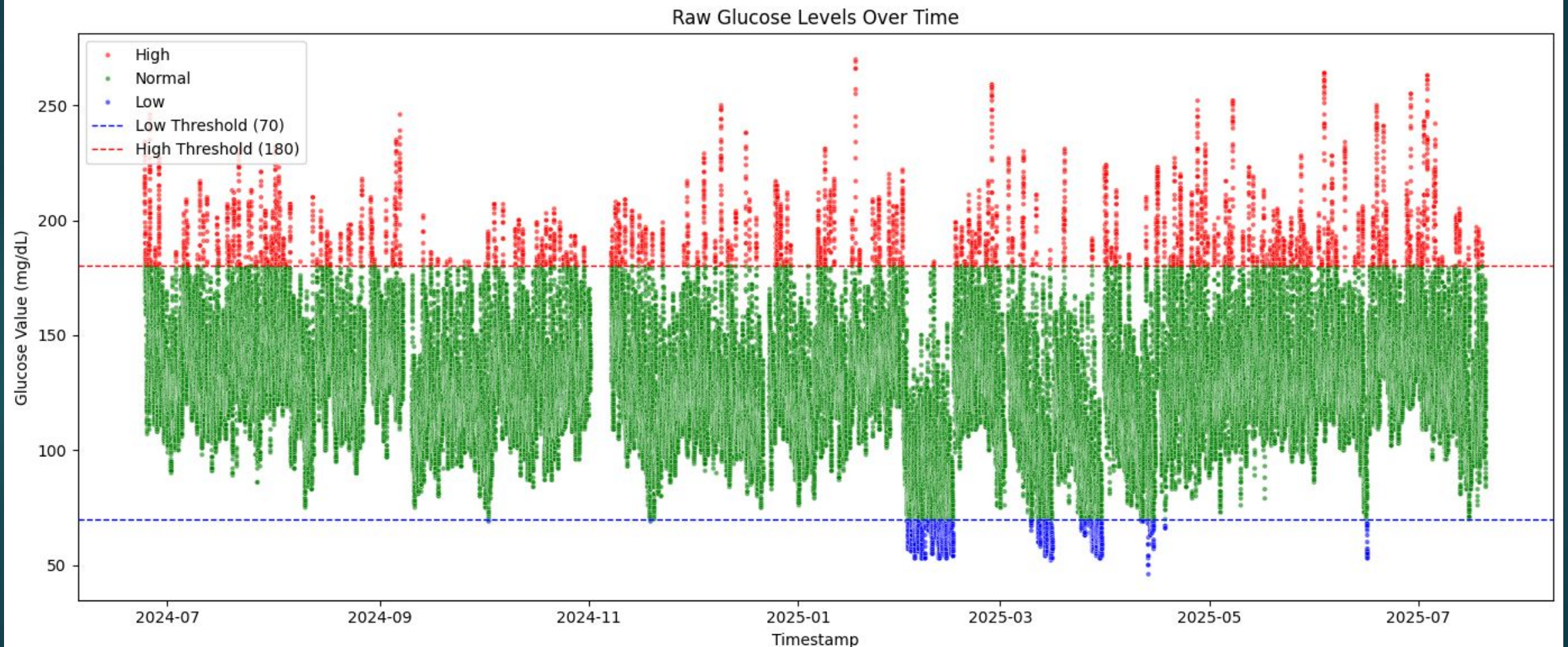
Questions?

Standard Glycemic Measures

Glycemic Index (GI) is a measure of how quickly 50g of carbohydrates from a particular food are absorbed over 2 hours, normalized against 50g of pure glucose.

Glycemic Load (GL) is a measure of how quickly a typical serving of a particular food will affect glucose levels over the next 2 hours. It is $GI * \text{carbs in a serving} / 100$.

ADA Target Levels for Blood Glucose



Glucose over time regressed

