HEALTH FREAK · METHODOLOGY PAPER · V4.0

The *Restaurant Leaf Score*

A four-pillar, data-driven framework for rating restaurant healthiness across the UK eat-out market. Built on real menu data, independent web research, and automated NLP analysis across 452 London restaurants.

VERSION	RESTAURANTS	RATING SCALE	DATA SOURCES
4.0 — Production	452	1–4 stars	Supabase · Perplexity · NLP

PUBLISHED

April 2026

01 — OVERVIEW

Measuring the conditions for informed eating

LIVE DATA — 452 RESTAURANTS — APRIL 2026

> *The Restaurant Leaf Score rates how easy it is to eat well at a restaurant — not just whether the food is healthy. It rewards transparency, variety, and preparation clarity alongside menu quality, because* **an informed choice is always better than an uninformed one.**

Health Freak scores individual dishes using the Health Freak Score. The Restaurant Leaf Score extends this upward: it aggregates dish-level nutritional quality with restaurant-level behaviours to produce a single 1–4 star rating that can be explained to a consumer in one sentence per pillar.

Restaurants are not judged on whether they serve indulgent food. A chain that is transparent about a less healthy menu earns more stars than one that obscures a nominally healthier one. A steakhouse, a burger chain, and a salad bar can all earn high scores — if they are honest about what they serve, offer meaningful choice, and help customers understand what they are eating.

The score measures *the conditions for informed eating*, not the food itself. That distinction is deliberate, commercially defensible, and central to everything that follows.

02 — ARCHITECTURE

Four pillars, one score

The composite is built from four pillars, each graded 0–3. Grades sum to a raw total (maximum 12) which maps to a 1–4 star rating via fixed thresholds.

```
raw_total = P1 (quality) + P2 (transparency) + P3 (availability) + P4 (prep clarity)
stars = ★★★★ if total ≥ 9 · ★★★ if ≥ 6 · ★★ if ≥ 3 · ★ if < 3
```

Each pillar contributes equally on a 0–3 scale. No pillar can be gamed in isolation — a restaurant needs consistent performance across all four dimensions to reach ★★★★.

P1 · FREAK_GRADE_QUALITY

40%

Menu health quality

Derived directly from Health Freak's dish-level scoring pipeline. The highest-weighted pillar — anchors the restaurant score to real nutritional data rather than brand claims.

- Median score_protein_prime across all non-zero dishes
- Aggregated at chain level for chains, restaurant level for independents
- Resistant to snack and drink drag via median (not mean)

P2 · FREAK_GRADE_TRANSPARENCY

25%

Nutritional transparency

Rewards restaurants that publish nutritional information clearly. Transparency is treated as independently valuable — disclosing an unhealthy menu honestly scores higher than silence about a healthier one.

- Full macros (protein/carbs/fat) published online
- Allergen information complete and accessible
- Calorie display in-store (chains with 250+ UK employees)

P3 · FREAK_GRADE_AVAILABILITY

20%

Healthy options availability

Measures whether a health-conscious consumer can find meaningful options. The question is not whether all dishes are healthy — it is whether someone who wants to eat well has something to choose from.

- At least one vegan main dish available
- At least one gluten-free main dish available
- At least one main dish confirmed under 500 kcal

P4 · FREAK_GRADE_PREP_CLARITY

15%

Preparation clarity

Scores how clearly a restaurant communicates how its food is prepared. Cooking method is one of the strongest predictors of caloric density — yet it is rarely surfaced at the point of ordering.

- Cooking method language in dish descriptions (NLP)
- Light-to-heavy method ratio across the menu
- Cold-prep path for sushi/poke/salad concepts
- Format floor for inherently transparent formats

SCORING LEVEL

Chains (restaurants sharing a `chain_id`) are scored once at chain level — all dishes and locations aggregated. A single score row is written to `restaurant_scores` keyed by `chain_id`, which individual locations inherit via JOIN. Independents are scored and stored at `restaurant_id` level. The `score_level` field makes this explicit in the database.

03 — PILLAR 1

Menu health quality

P1 is the only pillar computed entirely from Health Freak's internal dish data. It uses the `score_protein_prime` field — a pre-computed per-dish nutritional score on a 0–10 scale, based on calorie proximity to target, protein density, and fibre density under the Protein Prime dietary profile.

COMPUTATION

```
scores = dishes WHERE restaurant IN group AND score_protein_prime > 0
P1 = 3 if median ≥ 6.70 · 2 if ≥ 6.00 · 1 if ≥ 5.00 · 0 otherwise
```

Median is used rather than mean. Mean scores are pulled downward by incidental items (drinks, snacks, desserts) that inflate dish counts without representing the core menu. Median is resistant to this drag. Dishes with score_protein_prime = 0.0 are excluded entirely — these are non-food items.

> Profile note. P1 is anchored to the Protein Prime profile specifically — the most data-rich profile in the Health Freak pipeline at the time of scoring. As additional profiles mature (Fibre First, Lean Cut, Bio Hacker), a blended median across profiles will be evaluated for future versions. The current single-profile approach is disclosed here so restaurant partners understand which dietary lens is being applied.

DISTRIBUTION (452 RESTAURANTS)

GRADE	THRESHOLD	INTERPRETATION	COUNT	%
3	≥ 6.70	Top tier — menu consistently strong	89	19.7%
2	6.00–6.69	Solid — decent healthy options	320	70.8%
1	5.00–5.99	Below average — limited healthy choices	31	6.9%
0	< 5.00	Outlier — dessert/drink-focused	12	2.7%

The distribution is well-calibrated: the majority of eat-out restaurants in London cluster in Grade 2 (solid but not exceptional), with a meaningful top tier at Grade 3 and a small tail of dessert and drink-focused outlets at Grade 0.

04 — PILLAR 2

Nutritional transparency

P2 was scored using Perplexity web research. All 13 chains and 107 independent restaurants were researched individually. The remaining 332 independents were assigned a default score of P2=0 based on an overwhelming observed pattern: approximately 70% of researched independents returned no nutritional information at all, and a further 25% had only partial allergen data on delivery platforms.

GRADE CRITERIA

GRADE	CRITERIA
3	Full macros (protein, carbs, fat) published online + allergens complete. In-store calorie display is a bonus but not required — publishing full macros online demonstrates equal or greater commitment to transparency.
2	Calories published (website or PDF) + allergens accessible, but no full macro breakdown. OR full macros available but allergen info missing.
1	Partial disclosure only: allergens listed on delivery platforms (Deliveroo/Uber Eats), or calories available via third-party sites but not officially published.
0	No meaningful nutritional disclosure found online.

> Design decision — in-store display. UK law requires businesses with 250+ employees to display calories at the point of choice in-store. The P2 rubric does not penalise smaller operators for non-compliance with a rule that does not apply to them. All qualifying chains (Pret, Nando's, Wagamama, Itsu, LEON, Wasabi, Subway, Tossed) have in-store calorie display confirmed. However, the grade does not require in-store display for Grade 3 — publishing equivalent information online is treated as equally transparent.

DISTRIBUTION (452 RESTAURANTS)

GRADE	COUNT	%	SOURCE
3	11	2.4%	RESEARCHED
2	26	5.8%	RESEARCHED
1	50	11.1%	RESEARCHED
0	365	80.8%	332 DEFAULTED + 33 researched

The 80.8% P2=0 finding is not a judgment on these restaurants — it is an accurate reflection of the UK independent dining market. Small restaurants have no regulatory incentive to publish nutritional data and no infrastructure to produce it. This is the market gap Health Freak exists to fill. Publishing this finding transparently is important: it explains why P2 is the single largest driver of star ratings across the full dataset.

NOTABLE P2=3 INDEPENDENTS

Six independent (non-chain) restaurants achieved P2=3: Urban Greens, Zambrero, Coco di Mama, Camile Thai, Nusa Kitchen, and Remedy Kitchen. These are the benchmark for what small-operator transparency looks like in practice and represent a meaningful differentiator from the 365 restaurants with no disclosure.

05 — PILLAR 3

Healthy options availability

P3 was scored using the same Perplexity research used for P2, supplemented by cuisine-type defaults for unresearched independents. The three criteria are binary: each confirmed criterion adds one grade point.

THE UNDER-500-KCAL CIRCULARITY

The third criterion — a main dish confirmed under 500 kcal — is intentionally linked to P2. A restaurant that does not publish calorie information cannot have an under-500-kcal dish "confirmed." In practice, P3=3 is only achievable by restaurants that also score P2≥1, because calorie data must exist somewhere to verify. This is by design: it rewards restaurants that enable informed healthy choices, not merely restaurants that happen to serve low-calorie food in silence.

CUISINE-TYPE DEFAULTS FOR UNRESEARCHED INDEPENDENTS

Japanese / Sushi / Poke	Default P3 = 2
Thai / Vietnamese	Default P3 = 2
Indian	Default P3 = 2
Lebanese / Middle Eastern	Default P3 = 2
Greek	Default P3 = 2
Mexican	Default P3 = 2
Healthy / Salad / Acai	Default P3 = 2
Korean	Default P3 = 1
Pizza	Default P3 = 1
Burger	Default P3 = 1
Chicken / Peri Peri	Default P3 = 1
Café / Brunch	Default P3 = 1
Chinese	Default P3 = 1
Dessert / Bakery / Bubble Tea	Default P3 = 0
Other / Unclassified	Default P3 = 1

Cuisine-type inference is a legitimate proxy rather than a shortcut. The connection between cuisine format and dietary variety is structurally determined — sushi and poke concepts reliably offer vegan (tofu) and GF (rice-based) options; dessert concepts reliably do not. Every restaurant scored via cuisine default is flagged as `p3_source = "cuisine_default"` in the database.

DISTRIBUTION (452 RESTAURANTS)

GRADE	COUNT	%
3	29	6.4%
2	194	42.9%

1	222	49.1%
0	7	1.5%

06 — PILLAR 4

Preparation clarity

P4 addresses a gap in conventional restaurant scoring: cooking method is one of the strongest predictors of caloric density, yet it is rarely surfaced to consumers at the point of ordering. Two dishes described only as "chicken" may differ by 400 kcal depending on whether they are grilled or fried. P4 measures how clearly this information is communicated.

P4 was scored using automated NLP analysis of dish name and description columns in the Health Freak database. Three scoring paths were applied depending on restaurant format.

PATH 1 — STANDARD (HOT KITCHEN)

Dish descriptions are parsed for cooking method keywords, classified as light or heavy.

CATEGORY	KEYWORDS
Light methods	grilled, steamed, baked, roasted, poached, pan-seared, slow-cooked, chargrilled, wok-fried, stir-fried
Heavy methods	fried, battered, deep-fried, crispy, breaded, tempura

Grade = 3 if ≥60% of dishes state a method AND light:heavy ratio ≥ 3:1. Grade = 2 if 30–59% state a method OR ratio 2:1–3:1. Grade = 1 if <30% state a method. Grade = 0 if no cooking method language found.

PATH 2 — COLD-PREP (SUSHI / POKE / SALAD)

For restaurants where ≥70% of main dishes require no hot cooking, cooking method analysis is replaced by ingredient specificity scoring. Specific named ingredients ("salmon", "edamame", "mango") in ≥60% of dishes = Grade 3. 30–59% = Grade 2. Generic language only = Grade 1.

PATH 3 — FORMAT FLOOR

Certain restaurant formats are inherently transparent about preparation by their nature, regardless of how menu descriptions are written. A minimum floor of Grade 2 is applied to these formats:

FORMAT	FLOOR	RATIONALE
Juice / smoothie bar	2	Ingredients are the product — listed by definition on the menu
Poke / build-your-own bowl	2	Assembly format — components explicitly listed, customer selects each

Grilling concept (peri peri / rotisserie)	2	Cooking method is the brand identity — implied even when not written on every dish

> Why the format floor matters. Without it, Joe & The Juice scores P4=0 — no cooking method keywords in juice and sandwich descriptions. With it, Joe & The Juice scores P4=2, correctly reflecting that a juice bar is definitionally transparent about what goes into every drink. The floor corrects a systematic bias in text-only analysis against formats where preparation is visible or structural rather than described.

> Insufficient data. Seven restaurants had fewer than 5 dish descriptions in the database — insufficient for reliable NLP analysis. These were assigned a conservative default of P4=1 and flagged as `p4_source = "insufficient_data_default"`. They will be rescored when additional menu data is available.

CHAIN P4 RESULTS

CHAIN	DISHES	% METHOD STATED	L:H RATIO	P4	NOTE
Nando's	836	42.9%	4.8:1	2	Closest chain to Grade 3 — falls short on 60% coverage
LEON	222	45.0%	3.4:1	2	Best light:heavy among mid-range chains
The Salad Project	29	44.8%	6.0:1	2	Excellent ratio, just under 60% threshold
Itsu	179	40.8%	1.5:1	2	Heavy tempura/katsu drag on ratio
Wagamama	704	30.5%	0.5:1	2	Heavy-leaning — katsu, fried dishes prominent
Pret	187	25.1%	11:1	2	Excellent ratio, lower % (cold sandwich format)
Pure	115	26.1%	∞	2	All light, zero heavy methods
Tossed	159	24.5%	5.7:1	2	Good ratio, lower method coverage
Farmer J	83	16.9%	∞	2	All light, no heavy methods found
Chipotle UK	31	9.7%	∞	2	Grilled only, low method mention in descriptions
Simple Health	128	9.4%	∞	2	Light only but rare in descriptions
Wasabi	144	27.8%	0:1	1	Only chain scoring P4=1 — all methods are heavy
Joe & The Juice	64	0.0%	–	2	Format floor applied — juice bar inherently transparent

Note: no chain reaches P4 Grade 3. Nando's comes closest (42.9% method mention, 4.8:1 ratio) but falls short of the 60% coverage threshold. This reflects how chain menus are written — flavour and sauce descriptions dominate over cooking method language.

DISTRIBUTION (452 RESTAURANTS)

GRADE	COUNT	%	BREAKDOWN
3	16	3.9%	4 standard path + 12 cold-prep path
2	344	76.1%	303 standard + 17 cold-prep + 8 format floor + 9 other
1	79	17.5%	66 standard + 6 cold-prep + 7 insufficient data default
0	13	2.9%	12 no method language + 1 cold-prep

07 — COMPOSITE SCORING

From grades to stars

```
raw_total = P1 + P2 + P3 + P4 (range: 0–12)
```

All four pillars carry equal maximum weight (3 points each). No normalisation is applied — the raw integer sum maps directly to a star band.

★★★★

Total 9–12

Exemplary. Strong across quality, transparency, and healthy variety. A health-conscious consumer can eat well here with confidence.

★★★

Total 6–8

Good. Reasonable healthy options available and some nutritional transparency. One or two areas with room to improve.

★★

Total 3–5

Mixed. Healthy eating here requires navigation. Limited transparency or variety — often good food, but not designed to help you choose.

★

Total 0–2

Limited. Significant deficiencies across multiple pillars. Typically dessert or drink-focused outlets not primarily serving food.

A restaurant needs a total of 9 or above to earn ★★★★ — requiring consistent performance across all four pillars simultaneously. A restaurant scoring 3/3 on quality and 0/3 on transparency can reach only ★★★ at best. This design prevents any single pillar from being gamed.

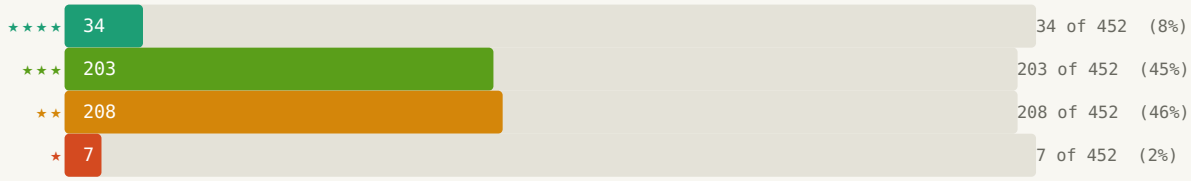
DATA SOURCE FLAGS

Every score row in `restaurant_scores` carries four source fields — one per pillar — so the evidence basis for each grade is transparent in the database and queryable for audit.

PILLAR	SOURCE VALUES
P1	`supabase_dish_data` (always)
P2	`perplexity_researched` · `cuisine_default`
P3	`perplexity_researched` · `cuisine_default`
P4	`menu_text_analysis` · `menu_text_coldprep` · `format_floor_applied` · `insufficient_data_default`

452-restaurant dataset

FINAL STAR DISTRIBUTION



PILLAR-LEVEL DISTRIBUTIONS

PILLAR	GRADE 3	GRADE 2	GRADE 1	GRADE 0
P1 — Menu quality	89 (20%)	320 (71%)	31 (7%)	12 (3%)
P2 — Transparency	11 (2%)	26 (6%)	50 (11%)	365 (81%)
P3 — Availability	29 (6%)	194 (43%)	222 (49%)	7 (2%)
P4 — Prep clarity	16 (4%)	344 (76%)	79 (17%)	13 (3%)

THE 13 CHAINS — FULL SCORECARDS

CHAIN	P1	P2	P3	P4	TOTAL	RATING
★★★★ EXEMPLARY (9–12)						
Itsu	3/3	3/3	3/3	2/3	11/12	★★★★
Nando's	2/3	3/3	3/3	2/3	10/12	★★★★
Pret A Manger	2/3	3/3	3/3	2/3	10/12	★★★★
The Salad Project	3/3	2/3	3/3	2/3	10/12	★★★★
Tossed	2/3	3/3	3/3	2/3	10/12	★★★★
Farmer J	3/3	2/3	2/3	2/3	9/12	★★★★
LEON	2/3	3/3	2/3	2/3	9/12	★★★★
Pure	3/3	2/3	2/3	2/3	9/12	★★★★
Simple Health	2/3	2/3	3/3	2/3	9/12	★★★★
Wagamama	2/3	3/3	2/3	2/3	9/12	★★★★
★★★ GOOD (6–8)						
Chipotle UK	2/3	2/3	2/3	2/3	8/12	★★★
Wasabi	2/3	3/3	2/3	1/3	8/12	★★★
Joe & The Juice	1/3	1/3	2/3	2/3	6/12	★★★

09 — FINDINGS

What the data reveals

P2

Transparency is the great divider

81% of restaurants score P2=0. This single pillar explains why nearly half of all restaurants land at ★★ despite having decent menus (P1=2) and reasonable dietary options (P3=1–2). The path from ★★ to ★★★★ runs almost entirely through transparency — publishing nutritional data is the single highest-leverage action any restaurant can take to improve their Restaurant Leaf Score.

0

No chain reaches P4 Grade 3

Nando's comes closest (42.9% method mention, 4.8:1 light:heavy ratio) but falls short of the 60% coverage threshold. Chain menus are written to sell flavour and occasion — not to communicate preparation method. This is a product insight: chains that deliberately add cooking method language to their descriptions could meaningfully differentiate on this dimension.

Wasabi

The only chain scoring P4=1

Wasabi scores P4=1 with a 0:1 light-to-heavy ratio — every cooking method mentioned in their dish descriptions is heavy (tempura, katsu). Despite strong P2 transparency, this preparation clarity weakness holds them to ★★★ alongside Chipotle. It is an honest reflection of a menu that leans heavily fried even as it publishes full nutritional data.

80%

The UK independent restaurant transparency gap

Four in five restaurants in this dataset publish no meaningful nutritional information online. This is not a failure — small operators have no regulatory incentive and no infrastructure to produce it. It is, however, the market gap Health Freak exists to fill. The six independent restaurants that do achieve P2=3 (Urban Greens, Zambrero, Coco di Mama, Camile Thai, Nusa Kitchen, Remedy Kitchen) represent a benchmark for what small-operator transparency looks like in practice.

7

Bottom tier is all dessert and drink

All seven ★ restaurants are dessert or drink-focused outlets: bubble tea shops, a doughnut chain, a patisserie, frozen yogurt, and a CBD drink store. None are restaurants in any meaningful sense — they are on delivery platforms but are not primarily food operators. This validates the bottom threshold: it catches genuine non-restaurants without misclassifying any legitimate food concept.

Itsu

Highest-scoring chain overall

Itsu scores 11/12 — the highest total in the chain dataset — with Grade 3 on menu quality (strong dish-level scores), Grade 3 on transparency (full macros, allergens, in-store calories), Grade 3 on availability (vegan, GF, and under-500-kcal all confirmed), and Grade 2 on prep clarity. The only point lost is P4, where tempura and katsu dishes lower the light:heavy ratio.

10 — LIMITATIONS & NEXT STEPS

Honest disclosure

P1 single profile P1 uses score_protein_prime only. Future versions should evaluate a blended median across all active dietary profiles (Protein Prime, Fibre First, Lean Cut, Bio Hacker) once those profiles have comparable data coverage. The single-profile approach may undervalue restaurants optimised for other dietary goals.

P2 defaults 332 of 452 restaurants were assigned P2=0 by default based on observed patterns from 107 researched independents. This is an accurate generalisation but not an individually verified score. Restaurants that believe they should score higher on P2 are encouraged to submit documentation.

P3 cuisine defaults Cuisine-type inference for P3 is structurally well-founded but may misclassify edge cases — a Korean restaurant with a comprehensive vegan menu, or a pizza restaurant with a certified GF base. The `cuisine_default` flag identifies these entries for future manual review.

P4 text-only limitation P4 reads menu descriptions. Google Reviews NLP and Google Images vision model signals are planned for a future version and will add a corroborating signal that reflects customer experience rather than marketing copy. The format floor partially corrects for the text bias but does not fully replace empirical evidence.

P4 insufficient data Seven restaurants were assigned P4=1 as a default due to fewer than 5 dish descriptions in the database. These are identified by `p4_source = "insufficient_data_default"` and will be rescored when additional menu data is available.

Seasonal menus All grades reflect a single scrape point. Menus change seasonally. A quarterly refresh cadence for chains and bi-annual for independents should be defined and automated before scores are surfaced to consumers.

Geographic scope This dataset covers London restaurants only. Expanding to other UK cities may require recalibration of thresholds — particularly P1, where the median score distribution may differ in markets with different cuisine compositions.

Restaurant appeals process No formal appeals process exists yet. Before public launch, Health Freak should define a mechanism by which restaurant operators can submit evidence to request a score review — particularly for P2 and P3 where the score depends on publicly available information that may be incomplete or outdated.

Scores are computed from Health Freak's internal database, Perplexity web research, and automated NLP analysis. They represent Health Freak's independent assessment and have

not been verified or endorsed by the restaurants named herein. Not for redistribution without permission.