Appendix C — Food DB

Appendix C1 — Food DB Schema & Data Dictionary

This appendix defines the schema, allowed value ranges, and logical dependencies for the Mealyze structured Food Database (`meal_advisor_food_db_v1.csv`, 300 rows).

1. Identity and Serving Columns

Column Name	Type	Description	Value Range / Example
id	integer	Unique numeric identifier (1–300).	1 … 300
name_en	string	English name of the food item (for UI and retrieval).	“Kimchi Fried Rice”
name_ko	string	Korean name (for reference only).	“김치볶음밥”
category	string (enum)	High-level grouping of cuisine.	Korean / Japanese / Chinese / Southeast Asian / Indian / Middle Eastern / Western / Mexican / Dessert / Drink
serving_desc	string	Unit definition of 1 serving.	“1 bowl”, “1 plate”, “100 g”, “1 piece”

---

2. Nutritional Value Columns (per 1 serving)

Column	Type	Unit	Definition / Source
kcal	numeric	kcal	Caloric energy per serving. Representative guide value standardized to 1 serving.
protein_g	numeric	g	Protein grams per serving.
fat_g	numeric	g	Fat grams per serving.

carbs_g	numeric	g	Carbohydrate grams per serving.

---

3. Taste & Type Attributes (0 or 1)

Column	Meaning of 1	Dependencies / Notes
spicy	Food is spicy (≥ typical ramen level).	Often implies soul_food = 1, stomach_sensitive = 0.
sweet	Dominant sweet taste (e.g., desserts, sugar/syrup base).	—
salty	High sodium level (soups, pickles, fermented sides).	—
oily	High oil usage or fatty cooking.	Must be 1 if deep_fried = 1.
deep_fried	Cooked by deep-frying.	Implies oily = 1.
grilled	Cooked by grilling/pan-grilling.	—
noodle	Noodle-based dish.	—
rice_based	Rice is the main carbohydrate.	—
soup_or_stew	Has substantial broth or stew base.	—
one_dish_meal	Complete single-dish meal (bowl / plate / pizza / burger).	—

---

4. Context-Aware Attributes (0 or 1)

Column	Meaning of 1	Criteria / Examples
soul_food	Comfort / mood-lifting dish (spicy or rich “soul food”).	떡볶이, 라면, 치킨, 삼겹살, 피자 etc.
stomach_sensitive	Gentle on stomach (non-spicy, non-oily, soft texture).	죽, 계란찜, 맑은국, 누룽지 etc.

hangover_friendly	Helpful for hangover relief (brothy stimulating soups).	콩나물국, 북어국, 순대국, 쌀국수 etc.
late_night	High-risk as late-night snack (high fat/sodium).	치킨, 족발, 라면, 떡볶이 etc.

---

5. Goal / Nutrition-Based Attributes (0 or 1)

Column	Logic Formula	Purpose
cut_friendly	1 if (kcal ≤ 450 AND fat_g ≤ 15 AND category ≠ “Dessert”)	Weight-cut diet option.
bulk_friendly	1 if (kcal ≥ 600 AND protein_g ≥ 20)	Muscle gain / high-calorie option.
high_protein	1 if (protein_g ≥ 20)	Protein-dense food.
low_fat	1 if (fat_g ≤ 10)	Low-fat item.
low_carb	1 if (carbs_g ≤ 40)	Low-carbohydrate item.
low_sugar	1 if dessert with reduced sugar content or natural sweetness only.	Auxiliary for sugar control.

---

6. Safety & Allergen Attributes (0 or 1, Conservative Rule)

Column	Meaning of 1	Notes
contains_egg	May contain egg ingredients.	Even trace possibility = 1.
contains_milk	May contain milk or dairy.	—
contains_wheat	May contain wheat or gluten.	—
contains_peanut	May contain peanuts or nuts.	—
contains_seafood	May contain fish or shellfish.	—
contains_cucumber	May contain cucumber (including sauce / garnish).	—

---

7. Logical Dependencies (Summary)

Relationship	Explanation
deep_fried → oily	Every deep-fried food is classified as oily.
soul_food ↔ stomach_sensitive (mutual exclusion)	Normally mutually exclusive (exceptions allowed for mild but satisfying dishes).
cut_friendly and bulk_friendly cannot both be 1	Opposite nutritional targets.
category = “Dessert” → cut_friendly = 0	Desserts excluded from cut foods.

---

8. Provenance Policy Summary

Item	Policy
Serving standard	Nutrition values are entered as 1-serving guide values; serving_desc states the unit.
Source types	public_food_db (MFDS/RDA/USDA), manufacturer_label, recipe_average
Traceability	Provenance samples (≥ 50 items) documented in Appendix E.

Appendix C2 — Food DB Labeling Rules (Final Data Specification)

This document defines the finalized labeling rule-set (“Final Data Specification”) used to construct the Mealyze structured Food DB (300 foods, meal_advisor_food_db_v1.csv).
Its goals are to (1) reduce hallucination risks in LLM-generated explanations,
(2) reflect Korean food context tags (e.g., hangover-friendly, late-night risk, comfort food, stomach-sensitive)
with consistent rules, and (3) make nutrition/goal/safety tags reproducible and auditable.

[Data Entry Policy]
1) Scope & Units
- All nutrition values (kcal, protein_g, fat_g, carbs_g) are entered as a representative “guide value” per 1 serving.
- serving_desc explicitly describes the unit of 1 serving (e.g., 1 bowl, 1 plate, 1 roll, 100g).

- category stores the high-level cuisine/category (e.g., Korean / Japanese / Chinese / Dessert / Drink).

2) Value Types
- Nutrition columns are non-negative numeric values.
- Tag/attribute columns are binary indicators (0/1):
  (0 = not applicable, 1 = applicable)

[Rule Set 1: Context-Aware Attributes]
- soul_food = 1:
  Foods commonly perceived as mood-lifting/comfort foods, often spicy/oily/high-calorie or culturally “soul food”.
  (e.g., tteokbokki, ramen, fried chicken, samgyeopsal, gopchang, malatang, pizza)
- stomach_sensitive = 1:
  Foods that are gentle on digestion (not spicy + not oily + soft texture).
  (e.g., porridge, steamed egg, clear soups, nurungji)
- hangover_friendly = 1:
  Broth-based dishes commonly associated with hangover relief.
  (e.g., kongnamul soup, dried pollack soup, sundae soup, jjamppong, pho)
- late_night = 1:
  Foods categorized as higher risk after 10 PM due to high fat/sodium and heavy digestion burden.
  (e.g., chicken, jokbal, ramen, tteokbokki)

[Mutual Exclusion Rule]
- In principle, soul_food and stomach_sensitive cannot both be 1.
- Exception: a small set of dishes can be both satisfying and gentle (e.g., mild hotpots such as dak-hanmari, shabu-shabu, mille-feuille nabe).

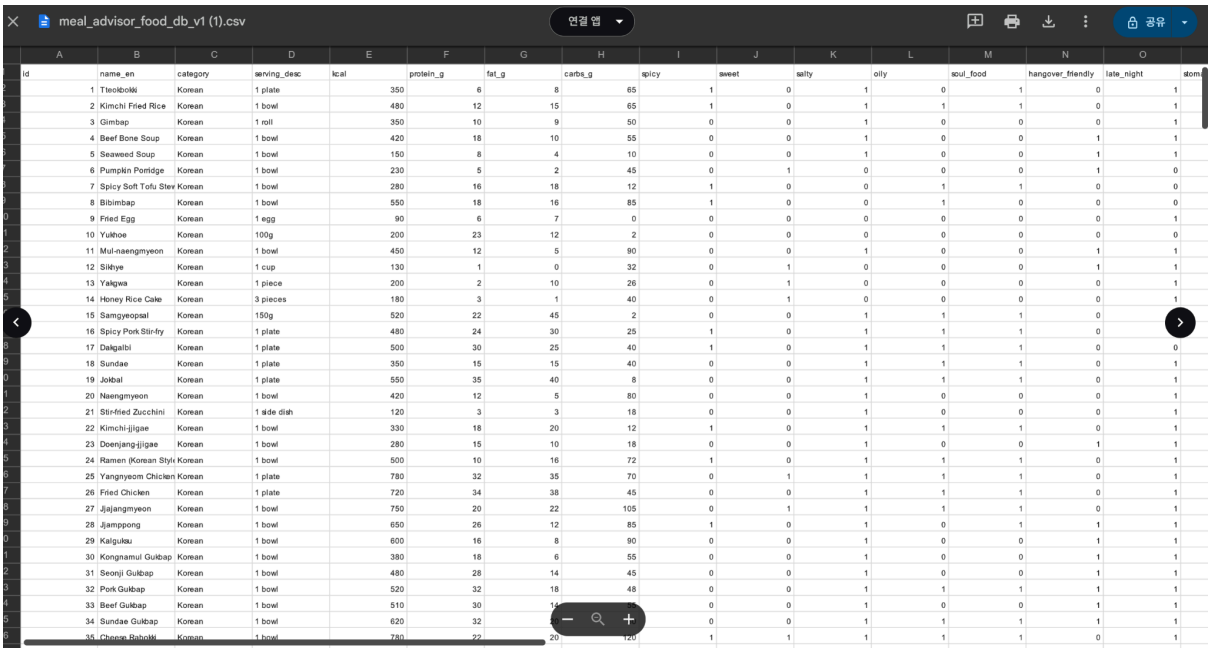
[Rule Set 2: Nutrition & Goal Attributes (Quantitative Thresholds)]
- cut_friendly = 1 if (kcal ≤ 450) AND (fat_g ≤ 15) AND (category ≠ 'Dessert')
- bulk_friendly = 1 if (kcal ≥ 600) AND (protein_g ≥ 20)
- high_protein = 1 if (protein_g ≥ 20)
- low_fat = 1 if (fat_g ≤ 10)
- low_carb = 1 if (carbs_g ≤ 40)

[Rule Set 3: Taste & Type Attributes]
- spicy = 1:
  Classified as spicy at or above a typical ramen-level baseline.
- oily = 1:
  High oil usage or fat-forward dishes (fried foods, fatty cuts, creamy sauces, etc.).
- deep_fried = 1:
  Deep-frying is the core cooking method.
  (Dependency: if deep_fried = 1, then oily must be 1.)
- sweet = 1:
  Dessert-like or sugar/syrup-based sweetness is central.
- salty = 1:
  High-sodium foods (soups, pickled/fermented sides, processed foods).

[Rule Set 4: Safety & Allergen Attributes (Conservative Policy)]
- All contains_* allergen tags follow a conservative labeling policy:
  if an ingredient has any plausible presence in common recipes, it is marked as 1 (True).
- Examples: contains_egg, contains_milk, contains_wheat, contains_peanut, contains_seafood, contains_pork, etc.

[Disclosures]
- Nutrition values in the Food DB are per-serving guide values; real meals may differ by recipe, brand, and portion size.
- Allergen tags may over-warn (false positives) because they are designed to prioritize safety over recall.

Appendix C3 — meal_advisor_food_db_v1.csv
https://drive.google.com/drive/u/3/folders/1BsEpLKr7FObXB2jqxgw2I5prwWADBKB7



id	name_en	category	serving_desc	kcal	protein_g	fat_g	carbs_g	spicy	sweet	salty	oily	soul_food	hangover_friendly	late_night	stomach
1	Tteokbokki	Korean	1 plate	350	6	8	65	1	0	1	0	1	0	1	
2	Kimchi Fried Rice	Korean	1 bowl	480	12	15	65	1	0	1	1	1	0	1	
3	Gimbap	Korean	1 roll	350	10	9	50	0	0	1	0	0	0	1	
4	Beef Bone Soup	Korean	1 bowl	420	18	10	55	0	0	1	0	0	1	1	
5	Seaweed Soup	Korean	1 bowl	150	8	4	10	0	0	1	0	0	1	1	
6	Pumpkin Porridge	Korean	1 bowl	230	5	2	45	0	1	0	0	0	1	0	
7	Spicy Soft Tofu Stew	Korean	1 bowl	280	16	18	12	1	0	0	1	1	0	0	
8	Bibimbap	Korean	1 bowl	550	18	16	85	1	0	0	1	0	0	0	
9	Fried Egg	Korean	1 egg	90	6	7	0	0	0	0	0	0	0	1	
10	Yukhoe	Korean	100g	200	23	12	2	0	0	0	0	0	0	0	
11	Mul-naengmyeon	Korean	1 bowl	450	12	5	90	0	0	1	0	0	1	1	
12	Sikhye	Korean	1 cup	130	1	0	32	0	1	0	0	0	1	1	
13	Yakgwa	Korean	1 piece	200	2	10	26	0	1	0	0	0	0	1	
14	Honey Rice Cake	Korean	3 pieces	180	3	1	40	0	1	0	0	0	0	1	
15	Samgyeopsal	Korean	150g	520	22	45	2	0	0	1	1	1	0	1	
16	Spicy Pork Stir-fry	Korean	1 plate	480	24	30	25	1	0	1	1	1	0	1	
17	Dalgalbi	Korean	1 plate	500	30	25	40	1	0	1	1	1	0	0	
18	Sundae	Korean	1 plate	350	15	15	40	0	0	1	1	1	0	1	
19	Jokbal	Korean	1 plate	550	35	40	8	0	0	1	1	1	0	1	
20	Naengmyeon	Korean	1 bowl	420	12	5	80	0	0	1	0	0	0	1	
21	Stir-fried Zucchini	Korean	1 side dish	120	3	3	18	0	0	1	0	0	0	1	
22	Kimchi-jjigae	Korean	1 bowl	330	18	20	12	1	0	1	1	1	0	1	
23	Doenjang-jjigae	Korean	1 bowl	280	15	10	18	0	0	1	0	0	1	1	
24	Ramen (Korean Style	Korean	1 bowl	500	10	16	72	1	0	1	1	1	0	1	
25	Yangnyeom Chicken	Korean	1 plate	780	32	35	70	0	1	1	1	1	0	1	
26	Fried Chicken	Korean	1 plate	720	34	38	45	0	0	1	1	1	0	1	
27	Jjajangmyeon	Korean	1 bowl	750	20	22	105	0	1	1	1	1	0	1	
28	Jjamppong	Korean	1 bowl	650	26	12	85	1	0	1	0	1	1	1	
29	Kalguksu	Korean	1 bowl	600	16	8	90	0	0	1	0	0	1	1	
30	Kongnamul Gukbap	Korean	1 bowl	380	18	6	55	0	0	1	0	0	1	1	
31	Seonji Gukbap	Korean	1 bowl	480	28	14	45	0	0	1	0	0	1	1	
32	Pork Gukbap	Korean	1 bowl	520	32	18	48	0	0	1	1	1	1	1	
33	Beef Gukbap	Korean	1 bowl	510	30	14	50	0	0	1	0	0	1	1	
34	Sundae Gukbap	Korean	1 bowl	620	32	22	60	0	0	1	1	1	1	1	
35	Cheese Rabokki	Korean	1 bowl	780	22	20	120	1	1	1	1	1	0	1	