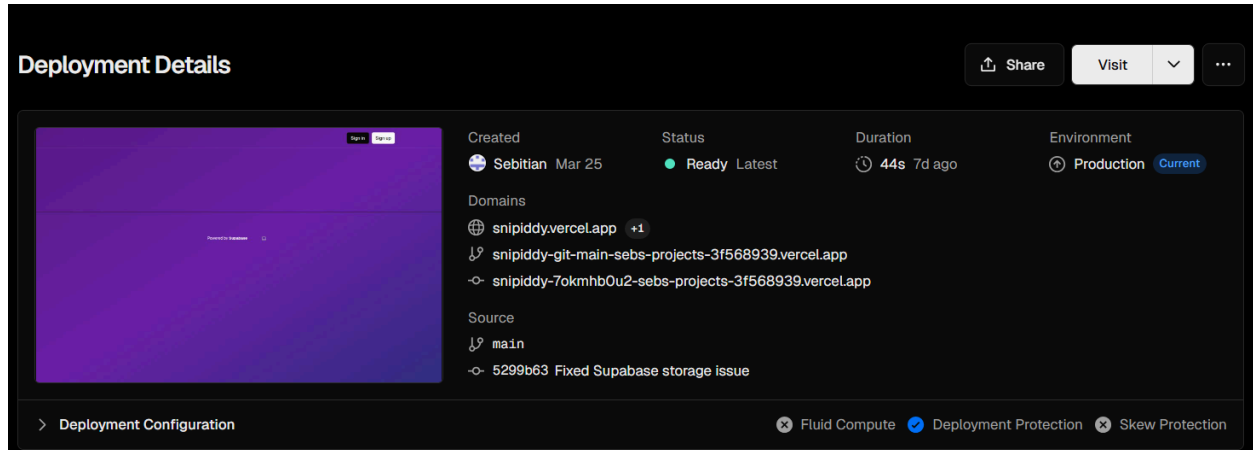


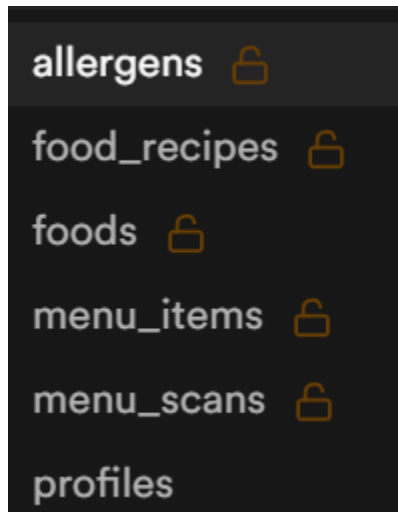
Database Design

Hosted server with supabase



Part 1

I. Tables :



1. Allergens

- Used to store a list of known well-known allergens as well as the ingredients associated with it.
- DDL

```
CREATE TABLE allergens (  
  id UUID DEFAULT uuid_generate_v4() PRIMARY KEY,  
  class TEXT,
```

```
type TEXT,  
group_name TEXT,  
food TEXT,  
allergy TEXT,  
created_at TIMESTAMP WITH TIME ZONE DEFAULT  
timezone('utc'::text, now()) NOT NULL);
```

2. Foods

- a. Used to store a list of ingredients to and known allergens.
- b. DDL

```
CREATE TABLE foods (  
    id UUID DEFAULT uuid_generate_v4() PRIMARY KEY,  
    food_product TEXT NOT NULL,  
    main_ingredient TEXT,  
    sweetener TEXT,  
    fat_oil TEXT,  
    seasoning TEXT,  
    allergens TEXT[],  
    prediction TEXT,  
    created_at TIMESTAMP WITH TIME ZONE DEFAULT now()  
    NOT NULL);
```

3. Food_recipes

- a. Use to store recipes and known ingredients
- b. DDL

```
CREATE TABLE food_recipes (  
    id UUID DEFAULT uuid_generate_v4() PRIMARY KEY,  
    title TEXT,  
    ingredient TEXT,  
    instructions TEXT,  
    image_name TEXT,  
    cleaned_ingredients TEXT);
```

4. Menu_scans

- a. Used to record scans from user and raw text from after performing optical character recognition (OCR)
- b. DDL

```
CREATE TABLE menu_scans (  
    id SERIAL PRIMARY KEY,
```

```
raw_text TEXT NOT NULL,  
created_at TIMESTAMP WITH TIME ZONE DEFAULT now()  
NOT NULL);
```

5. Menu_items

- a. Used to store all the items identified in the menu_scan
- b. DDL

```
CREATE TABLE menu_items (  
    id SERIAL PRIMARY KEY,  
    menu_scan_id INTEGER REFERENCES menu_scans(id),  
    dish_name TEXT NOT NULL,  
    description TEXT,  
    ingredients TEXT[],  
    allergens TEXT[],  
    price DECIMAL(10,2),  
    created_at TIMESTAMP WITH TIME ZONE DEFAULT now()  
NOT NULL);
```

6. Profiles

- a. Used to store the profile information and user settings.
- b. DDL

```
CREATE TABLE profiles (  
    id UUID DEFAULT uuid_generate_v4() PRIMARY KEY,  
    user_id UUID REFERENCES auth.users(id) ON DELETE  
CASCADE NOT NULL,  
    full_name TEXT,  
    email TEXT,  
    phone_number TEXT,  
    allergens TEXT[] DEFAULT '{}',  
    diets TEXT[] DEFAULT '{}',  
    other_info TEXT,  
    created_at TIMESTAMP WITH TIME ZONE DEFAULT now(),  
    updated_at TIMESTAMP WITH TIME ZONE DEFAULT now(),  
    UNIQUE(user_id));
```

II. Advanced SQL Queries

Description	Query
<p>1. menu_items → foods</p> <p>This query relates the items extracted from a scan to a list of known food items. We will use this query to identify items from a scanned menu, which we will then use to display menu items according to the user's preference.</p>	<pre>SELECT mi.id AS menu_item_id, mi.dish_name, mi.description, f.id AS food_id, f.food_product, f.main_ingredient, f.sweetener, f.fat_oil, f.seasoning, f.allergens, f.prediction FROM menu_items AS mi INNER JOIN foods AS f ON mi.dish_name = f.food_product;</pre>
<p>2. menu_items → food_recipes</p> <p>This query relates the menu_items to food_recipes using the dish_name. We then use this information to extract potential ingredients from a recipe, which we can relate to known allergens</p>	<pre>SELECT DISTINCT mi.id AS menu_item_id, mi.dish_name, mi.description, fr.id AS recipe_id, fr.title, fr.ingredients, fr.instructions, fr.image_name, fr.cleaned_ingredients FROM menu_items AS mi INNER JOIN food_recipes AS fr ON TRIM(LOWER(mi.dish_name)) = TRIM(LOWER(fr.title));</pre>
<p>3. profiles → allergens</p> <p>This query relates a user's allergy preferences to known food items.</p>	<pre>WITH user_allergens AS (SELECT p.id AS profile_id, p.full_name, LOWER(TRIM(unnested)) AS user_allergen FROM profiles p CROSS JOIN LATERAL unnest(p.allergens) AS unnested) SELECT DISTINCT u.profile_id, u.full_name,</pre>

	<pre> u.user_allergen, a.id AS allergen_id, a.food, a.allergy, a.class, a.type, a.group_name FROM user_allergens u JOIN allergens a ON LOWER(TRIM(a.allergy)) ILIKE '%' u.user_allergen '%' OR LOWER(TRIM(a.food)) ILIKE '%' u.user_allergen '%'; </pre>
<p>4. menu_items → allergens</p> <p>This query relates each ingredient found in a dish to a known allergen.</p>	<pre> WITH unnested_ingredients AS (SELECT mi.id AS menu_item_id, mi.dish_name, trim(regex_replace(unnest(string_to_array(regex_replace(regex_replace(mi.ingredients::text, '[\n"]', ", ", 'g'), '\", ', 'g'), '), '\s+ \s+\$', ", ", 'g'))) AS ingredient FROM menu_items mi WHERE mi.ingredients IS NOT NULL) SELECT DISTINCT ui.menu_item_id, ui.dish_name, ui.ingredient, a.food AS allergen_food, a.class AS allergen_class, a.type AS allergen_type, a.group_name AS allergen_group, a.allergy FROM unnested_ingredients ui JOIN allergens a ON ui.ingredient ILIKE a.food OR ui.ingredient ILIKE '%' a.food '%' OR a.food ILIKE '%' ui.ingredient '%' ORDER BY ui.dish_name, a.food; </pre>

4

576 rows

Results	Chart	Export					
menu_item_id	dish_name	ingredient	allergen_food	allergen_class	allergen_type	allergen_group	allergy
188	Artichoke and Sweet Potato Bake	(artichoke	Artichoke	Plant origin	Vegetable	Composite vegetable	Insulin Allergy
188	Artichoke and Sweet Potato Bake	lemon)	Lemon	Plant origin	Fruit	Citrus fruit / citric acid	Citrus Allergy
188	Artichoke and Sweet Potato Bake	sweet potato	Potato	Plant origin	Vegetable	Potato	Potato Allergy
188	Artichoke and Sweet Potato Bake	sweet potato	Sweet potato	Plant origin	Vegetable	Potato	Potato Allergy
44	Aubergine Moutabal	olive oil)	Olive oil	Plant origin	Oil	Oil	NULL
44	Aubergine Moutabal	yogurt	Yogurt	Animal origin	Dairy	Dairy	Milk allergy / Lactose intolerance
40	Aubergine Mussaka'a	onions	Onion	Plant origin	Vegetable	Liliaceous vegetable	Allium Allergy
40	Aubergine Mussaka'a	(plum tomatoes	Tomato	Plant origin	Vegetable	Solanceous vegetable	Nightshade Allergy
197	Blackened Cod Fillet	onion	Onion	Plant origin	Vegetable	Liliaceous vegetable	Allium Allergy
16	Boneless Chicken Wings	celery	Celery	Plant origin	Vegetable	Umbelliferous vegetable	Hypersensitivity
26	Boneless Chicken Wings	celery	Celery	Plant origin	Vegetable	Umbelliferous vegetable	Hypersensitivity
135	Boneless Chicken Wings	celery	Celery	Plant origin	Vegetable	Umbelliferous vegetable	Hypersensitivity
16	Boneless Chicken Wings	blue cheese)	Cheese	Animal origin	Dairy	Dairy	Milk allergy / Lactose intolerance
26	Boneless Chicken Wings	blue cheese)	Cheese	Animal origin	Dairy	Dairy	Milk allergy / Lactose intolerance
135	Boneless Chicken Wings	blue cheese)	Cheese	Animal origin	Dairy	Dairy	Milk allergy / Lactose intolerance
16	Boneless Chicken Wings	(chicken	Chicken	Animal origin	Poultry	Offal	Poultry Allergy
26	Boneless Chicken Wings	(chicken	Chicken	Animal origin	Poultry	Offal	Poultry Allergy
135	Boneless Chicken Wings	(boneless chicken	Chicken	Animal origin	Poultry	Offal	Poultry Allergy
53	Breakfast Pastries	(Butter	Butter	Animal origin	Dairy	Dairy	Milk allergy / Lactose intolerance

Part 2 : Indexing

Analysis

- FinalTried

#	Reasoning	SQL
1		
2	These indices precompute the value of TRIM(LOWER(dish_name)) and TRIM(LOWER(title)) for each row, which speed up the process slightly.	<pre>CREATE INDEX idx_menu_items_dish_name_norm ON menu_items ((TRIM(LOWER(dish_name)))); CREATE INDEX idx_food_recipes_title_norm ON food_recipes ((TRIM(LOWER(title))));</pre>
3	These indices precompute and store the first word of the allergen field so instead of using wildcard searches every time, we can do equality comparisons which are way faster.	<pre>CREATE INDEX idx_allergens_first_word_allergy ON allergens ((split_part(lower(trim(allergy)), ' ', 1))); CREATE INDEX idx_allergens_first_word_food</pre>

		<pre>ON allergens ((split_part(lower(trim(food)), ' ', 1)));</pre>
	<p>Assumes data is normalized (which is not), so it does direct comparisons. However, incorrect data is returned.</p>	<pre>CREATE INDEX idx_allergens_allergy_norm ON allergens ((LOWER(TRIM(allergy)))); CREATE INDEX idx_allergens_food_norm ON allergens ((LOWER(TRIM(food))));</pre>
4	<p>These indices precompute and store the first word of the allergen field so instead of using wildcard searches every time, we can do equality comparisons which are way faster. This allows use to get rid of the ILIKE computations which are expensive.</p>	<pre>CREATE INDEX idx_allergens_first_word_food ON allergens ((split_part(lower(trim(food)), ' ', 1)));</pre>

Performance

#	Before	After
1	<p>Cost : 35.46</p> <pre> QUERY PLAN Hash Join (cost=15.98..35.46 rows=373 width=165) (actual time=2.810 Hash Cond: (mi.dish_name = f.food_product) -> Seq Scan on menu_items mi (cost=0.00..13.29 rows=329 width=77) -> Hash (cost=10.99..10.99 rows=399 width=88) (actual time=2.774 Buckets: 1024 Batches: 1 Memory Usage: 57kB -> Seq Scan on foods f (cost=0.00..10.99 rows=399 width=88) Planning Time: 1.161 ms Execution Time: 3.080 ms </pre>	
2	<p>Cost : 11708.57</p> <pre> QUERY PLAN Unique (cost=9201.47..11708.57 rows=18371 width=1425) (actual time= -> Gather Merge (cost=9201.47..11295.23 rows=18371 width=1425) (Workers Planned: 1 Workers Launched: 1 -> Sort (cost=8201.46..8228.48 rows=10806 width=1425) (act Sort Key: mi.id, mi.dish_name, mi.description, fr.id, Sort Method: quicksort Memory: 32kB Worker 0: Sort Method: quicksort Memory: 37kB -> Hash Join (cost=17.40..3539.44 rows=10806 width=1 Hash Cond: (TRIM(BOTH FROM lower(fr.title)) = TR -> Parallel Seq Scan on food_recipes fr (cost= -> Hash (cost=13.29..13.29 rows=329 width=77) Buckets: 1024 Batches: 1 Memory Usage: 4 -> Seq Scan on menu_items mi (cost=0.00. Planning Time: 0.831 ms Execution Time: 17.060 ms </pre>	<p>Cost : 11469.72</p> <pre> QUERY PLAN Unique (cost=8779.62..11469.72 rows=19712 width=1425) (actual time= -> Gather Merge (cost=8779.62..11026.20 rows=19712 width=1425) (Workers Planned: 1 Workers Launched: 1 -> Sort (cost=7779.61..7808.59 rows=11595 width=1425) (acti Sort Key: mi.id, mi.dish_name, mi.description, fr.id, Sort Method: quicksort Memory: 34kB Worker 0: Sort Method: quicksort Memory: 35kB -> Merge Join (cost=0.43..2771.82 rows=11595 width=1. Merge Cond: (TRIM(BOTH FROM lower(fr.title)) = TI -> Parallel Index Scan using idx_food_recipes_t -> Materialize (cost=0.15..19.53 rows=353 width= -> Index Scan using idx_menu_items_dish_n Planning Time: 1.018 ms Execution Time: 20.999 ms </pre>
3	Cost : 26532.31	Cost : 6140

	<pre> HashAggregate (cost=26449.72..26532.31 rows=5506 width=159) (actual Group Key: p.id, p.full_name, lower(TRIM(BOTH FROM unnested.unnest Batches: 1 Memory Usage: 225kB -> Nested Loop (cost=0.00..26325.83 rows=5506 width=159) (actual Join Filter: ((lower(TRIM(BOTH FROM a.allergy)) ~* ((''::t Rows Removed by Join Filter: 890 -> Nested Loop (cost=0.00..73.00 rows=3000 width=80) (actu -> Seq Scan on profiles p (cost=0.00..13.00 rows=300 -> Function Scan on unnest unnested (cost=0.00..0.10 -> Materialize (cost=0.00..5.76 rows=184 width=79) (actual -> Seq Scan on allergens a (cost=0.00..4.84 rows=184 Planning Time: 0.808 ms Execution Time: 5.083 ms </pre>	<pre> QUERY PLAN HashAggregate (cost=6150.56..6240.32 rows=5984 width=159) (actual t Group Key: p.id, p.full_name, lower(TRIM(BOTH FROM unnested.unnest Batches: 1 Memory Usage: 225kB -> Nested Loop (cost=0.76..6015.92 rows=5984 width=159) (actual -> Nested Loop (cost=0.00..73.00 rows=3000 width=80) (actu -> Seq Scan on profiles p (cost=0.00..13.00 rows=300 -> Function Scan on unnest unnested (cost=0.00..0.10 -> Bitmap Heap Scan on allergens a (cost=0.76..1.95 rows=2 Recheck Cond: ((split_part(lower(TRIM(BOTH FROM unnest Heap Blocks: exact=10 -> BitmapOr (cost=0.76..0.76 rows=2 width=0) (actual -> Bitmap Index Scan on idx_allergens_first_wor Index Cond: (split_part(lower(TRIM(BOTH FRI -> Bitmap Index Scan on idx_allergens_first_wor Index Cond: (split_part(lower(TRIM(BOTH FRI Planning Time: 1.067 ms Execution Time: 1.868 ms </pre>
4	<p>Cost : 19072.72</p> <pre> Unique (cost=18870.04..19072.72 rows=9008 width=115) (actual time=4 -> Sort (cost=18870.04..18892.56 rows=9008 width=115) (actual ti Sort Key: mi.dish_name, a.food, mi.id, (TRIM(BOTH FROM regex Sort Method: quicksort Memory: 101kB -> Nested Loop (cost=0.00..18278.35 rows=9008 width=115) (Join Filter: (((TRIM(BOTH FROM regexp_replace((unnest(Rows Removed by Join Filter: 197592 -> Result (cost=0.00..134.65 rows=3280 width=52) (ac -> ProjectSet (cost=0.00..36.25 rows=3280 widt -> Seq Scan on menu_items mi (cost=0.00. Filter: (ingredients IS NOT NULL) Rows Removed by Filter: 1 -> Materialize (cost=0.00..5.76 rows=184 width=63) (-> Seq Scan on allergens a (cost=0.00..4.84 ro Planning Time: 0.754 ms Execution Time: 469.607 ms </pre>	<p>Cost : 594.53</p> <pre> QUERY PLAN Unique (cost=521.68..594.53 rows=3238 width=115) (actual time=5.958 -> Sort (cost=521.68..529.77 rows=3238 width=115) (actual time=5 Sort Key: mi.dish_name, a.food, mi.id, (TRIM(BOTH FROM regex Sort Method: quicksort Memory: 50kB -> Hash Join (cost=7.14..332.89 rows=3238 width=115) (actu Hash Cond: ((split_part(lower(TRIM(BOTH FROM regexp_re -> Result (cost=0.00..231.77 rows=3520 width=84) (ac -> ProjectSet (cost=0.00..38.17 rows=3520 widt -> Seq Scan on menu_items mi (cost=0.00. Filter: (ingredients IS NOT NULL) Rows Removed by Filter: 1 -> Hash (cost=4.84..4.84 rows=184 width=63) (actual Buckets: 1024 Batches: 1 Memory Usage: 25kB -> Seq Scan on allergens a (cost=0.00..4.84 ro Planning Time: 0.995 ms Execution Time: 6.194 ms </pre>