**MATKONIM – Software Documentation**

**EER DIAGRAM:**

**INSERT DIAGRAM HERE**

**DATA-BASE DESIGN PROCESS**

From the beginning of the database design process, we knew that we needed different tables for food recipes, cocktail recipes, and ingredients. While progressing with the application's functionality and database designing, we made several decisions regarding the database's structure:

1. We realized food recipes and cocktail recipes had some common attributes. Therefore, we decided to add a more general table, ALL\_RECIPES, that will include those common attributes and a unique recipe ID. COCKTAIL\_RECIPES and FOOD\_RECIPES tables are referring to that table using foreign keys, and each of those tables include attributes that relate to the specific type of recipe (food/cocktail).
2. We understood that each recipe contains a non-constant number of ingredients, and also each ingredient can appear in many recipes. Therefore, to avoid a many-to-many relation, we decided on a new table, RECIPE2INGREDIENTS, which connects each recipe to its ingredients and amounts of it.
3. At the beginning of the design process, we considered storing for each ingredient in each recipe its quantity and unit. Later on we realised that the conversion between different units is going to be harder than we thought, so instead we did this conversion before inserting the data into the DB, and stored the amounts as a factor of the serving size of each ingredient.
4. At first we thought about storing the nutritional values amounts of each ingredient in the INGREDIENTS table, but then we realized this will not be efficient. We then decided on 2 new tables – NUTRITIONS and INGREDIENT\_NUTRITION. The first table includes the details of each nutritional value, while the second connects each ingredient to its nutritional values amounts. INGREDIENT\_NUTRITION refers to both INGREDIENTS and NUTRITIONS tables by foreign keys. The reason for this 2 separate tables is keeping the database's consistency, and enabling simple updates and inserts to the database, if needed.
5. While writing queries that use NUTRITIONS table, we noticed that some of the nutritional values are considered "bad" (e.g. cholesterol, saturated fat), while others are considered "good" (e.g. iron, potassium). When the user inputs an amount of a nutritional value, we needed a way to know if this input is the maximal or minimal amount of the value. Meaning, for a "bad" nutritional value we would like recipes with **at most** user's input, and for a "good" nutritional value, we would like recipes with **at least** that amount. To solve this problem we added a "max\_or\_min" column which defines whether user's input is the maximal or minimal amount of this nutritient.
6. We contemplated on the best way to save the recommended daily intake for each age and gender, and finally decided on creating a new table – RECOMMENDED\_BY\_AGE\_GENDER, which includes data we manually retrieved from the internet. We thought about what would be the best way to store different age groups, and eventually decided on creating a "virtual dictionary", in which every value 0-6 is representing a different age group.

**DATABASE SCHEME STRUCTURE, TABLES AND OPTIMIZATIONS**

**ALL\_RECIPES(recipe\_id, recipe\_name, picture)**This table lists information regarding all recipes in the database (both food and cocktail).  
This table is referred to by a foreign key in FOOD\_RECIPES, COCKTAIL\_RECIPES and RECIPE2INGREDIENTS tables.

Columns:

* recipe\_id: A surrogate primary key used to uniquely identify each recipe in the table.
* recipe\_name: The recipe’s name.
* picture: A URL to the recipe's picture.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| recipe\_id | SMALLINT | V | V | V | V | - |
| recipe\_name | VARCHAR(255) | V | - | - | - | - |
| picture | VARCHAR(255) | V | - | - | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| recipe\_id | PRIMARY | - |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| recipe\_index\_ar | PRIMARY | recipe\_id |

**FOOD\_RECIPES(food\_id, recipe\_id, course, prep\_time\_in\_minutes, food\_details)**This table lists all details regarding food recipes.  
This table refers to ALL\_RECIPES table by a foreign key.

Columns:

* food\_id: A key used to identify each food recipe in the table.
* recipe\_id: A foreign key identifying the recipe in ALL\_RECIPES table.
* course: The course of the food recipe.
* prep\_time\_in\_minutes: The amount of minutes required to prepare this recipe.
* food\_details: A URL to the full recipe preparation instructions.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| food\_id | INT | V | - | V | - | - |
| recipe\_id | SMALLINT | V | - | V | V | - |
| course | VARCHAR(27) | V | - | - | - | - |
| prep\_time\_in\_minutes | SMALLINT | V | - | V | - | - |
| food\_details | VARCHAR(2500) | V | - | - | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| food\_id, recipe\_id | PRIMARY | - |
| recipe\_id | FOREIGN | ALL\_RECIPES.recipe\_id |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| food\_id\_fr | INDEX | food\_id |
| recipe\_id\_fr | INDEX | recipe\_id |
| course\_fr | INDEX | course |

**COCKTAIL\_RECIPES(cocktail\_id, recipe\_id, is\_alcoholic, serving\_glass, cocktail\_details)**This table lists all details regarding cocktail recipes.  
This table refers to ALL\_RECIPES table by a foreign key.

Columns:

* cocktail\_id: A primary key used to uniquely identify each cocktail recipe in the table.
* recipe\_id: A foreign key identifying the recipe in ALL\_RECIPES table.
* is\_alcoholic: Identifying if the cocktail is alcoholic or not.
* serving\_glass: The recommended serving glass of the cocktail recipe.
* cocktail\_details: Full cocktail preparation instructions.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| cocktail\_id | SMALLINT | V | V | V | - | - |
| recipe\_id | SMALLINT | V | V | V | - | - |
| is\_alcoholic | TINYINT | V | - | V | - | - |
| serving\_glass | VARCHAR(100) | - | - | - | - | NULL |
| cocktail\_details | VARCHAR(3500) | V | - | - | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| cocktail\_id | PRIMARY | - |
| recipe\_id | FOREIGN | ALL\_RECIPES.recipe\_id |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| cocktail\_id\_cr | PRIMARY | cocktail\_id |
| recipe\_id\_cr | INDEX | recipe\_id |

**INGREDIENTS(ingredient\_id, ingredient\_name, serving\_quantity, serving\_unit, serving\_weight\_grams)**This table lists all details regarding all ingredients that appear in food or cocktail recipes in the DB.  
This table is referred to by foreign keys in RECIPE2INGREDIENTS and INGREDIENT\_NUTRITION tables.

Columns:

* ingredient\_id: A surrogate primary key used to uniquely identify each ingredient in the table.
* Ingredient\_name: The name of the ingredient.
* serving\_quantity: The quantity of "units" in a single serving of the ingredient.
* serving\_unit: The size of a "unit" in a single serving of the ingredient.
* serving\_weight\_grams: The weight (in grams) of a single serving of the ingredient.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| ingredient\_id | INT | V | V | V | V | - |
| ingredient\_name | VARCHAR(55) | V | - | - | - | - |
| serving\_quantity | SMALLINT | V | - | V | - | - |
| serving\_unit | VARCHAR(55) | V | - | - | - | - |
| serving\_weight\_grams | SMALLINT | V | - | V | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| ingredient\_id | PRIMARY | - |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| ingredient\_id\_ing | PRIMARY | ingredient\_id |
| ingredient\_name\_ing | FULLTEXT | ingredient\_name |

**RECIPE2INGREDIENTS(recipe\_id, ingredient\_id, servings, full\_ingredient\_line)**This table connects each recipe to its ingredients and amounts.  
This table refers to ALL\_RECIPES and INGREDIENTS tables by foreign keys.

Columns:

* recipe\_id: A foreign key identifying the recipe in ALL\_RECIPES table.
* ingredient\_id: A foreign key identifying the recipe in INGREDIENTS table.
* servings: How many servings of the ingredient are present in the recipe.
* full\_ingredient\_line: The full text line of the ingredient to present to the user.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| recipe\_id | SMALLINT | V | - | V | - | - |
| ingredient\_id | INT | V | - | V | - | - |
| servings | FLOAT(7) | V | - | V | - | - |
| full\_ingredient\_line | VARCHAR(2500) | V | - | - | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| recipe\_id | FOREIGN | ALL\_RECIPES.recipe\_id |
| ingredient\_id | FOREIGN | INGREDIENTS.ingredient\_id |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| recipe\_id\_r2i | INDEX | recipe\_id |
| ingredient\_id\_\_r2i | INDEX | ingredient\_id |

**NUTRITIONS(nutirition\_id, nutrition\_name, max\_or\_min)**This table lists all details regarding all nutritional values present in the DB.

This table is referred to by foreign keys in INGREDIENT\_NUTRITION and RECOMMENDED\_BY\_AGE\_GENDER tables.

Columns:

* nutrition\_id: A surrogate primary key uniquely identifying the nutritional value in the DB.
* nutrition\_name: The name of the nutritional value.
* max\_or\_min: Defines whether the user's input should be the maximum amount or the minimum amount of this nutritional value. For example if the nutritional value is a bad one (e.g. cholesterol), user's input will be the maximum amount of this nutrient. If the nutritional value is a good one (e.g. iron), user's input will be the minimum amount of this nutrient.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| nutrition\_id | SMALLINT | V | V | V | V | - |
| nutrition \_name | VARCHAR(30) | V | - | - | - | - |
| max\_or\_min | VARCHAR(7) | V | - | - | - | - |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| nutrition\_id | PRIMARY | - |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| nutrition\_id\_nut | PRIMARY | nutrition\_id |
| nutrition\_name\_nut | INDEX | nutrition\_name |

**INGREDIENT\_NUTRITION(ingredient\_id, nutrition\_id, weight\_mg\_from\_ingredient)**This table connects each ingredient to the amount of each nutritional value in one serving of it.  
This table refers to INGREDIENTS and NUTRITIONS tables by foreign keys.

Columns:

* ingredient\_id: A foreign key identifying the ingreidnet in INGREDIENTS table.
* nutrition\_id: A foreign key identifying the nutritional value in NUTRITIONS table.
* weight\_mg\_from\_ingredient: The weight (in mg) of the nutritional value in one serving of the ingredient.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| ingredient\_id | INT | V | - | V | - | - |
| nutrition\_id | SMALLINT | V | - | V | - | - |
| weight\_mg\_from\_ingredient | FLOAT(3) | - | - | - | - | 0 |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| ingredient\_id, nutrition\_id | PRIMARY | - |
| ingredient\_id | FOREIGN | INGREDIENTS.ingredient\_id |
| nutrition\_id | FOREIGN | NUTRITIONS.nutrition\_id |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| ingredient\_id\_in | INDEX | ingredient\_id |

**RECOM****MENDED\_BY\_AGE\_GENDER(gender, age, nutrition\_id, weight\_mg)**This table lists all recommended daily intake amounts of each nutritional value, per age and gender.   
This table refers to NUTRITIONS table by a foreign key.

Columns:

* gender: The gender for which the recommended amounts are (either male or female).
* age: The age group for which the recommended amounts are, by this dictionary:
  + 0 is ages 14-18
  + 1 is ages 19-30
  + 2 is ages 31-40
  + 3 is ages 41-50
  + 4 is ages 51-60
  + 5 is ages 61-70
  + 6 is ages 71+
* nutrition\_id: A foreign key identifying the nutritional value in NUTRITIONS table.
* weight\_mg: The recommended intake of the nutritional value for this age and gender.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | | | | | | |
| Column Name | **Data Type** | **Not NULL** | **Unique** | **Unsigned** | **Auto Increment** | **Default** |
| gender | VARCHAR(7) | V | - | - | - | - |
| age | TINYINT | V | - | - | - | - |
| nutrition\_id | SMALLINT | V | - | V | - | - |
| weight\_mg | FLOAT(3) | V | - | - | - | 0 |

|  |  |  |
| --- | --- | --- |
| Keys | | |
| Column Name | **Key type** | **Refers to** |
| gender, age, nutrition\_id | PRIMARY | - |
| nutrition\_id | FOREIGN | NUTRITIONS.nutrition\_id |

|  |  |  |
| --- | --- | --- |
| Indexes | | |
| Index Name | **Index Type** | **Index Column** |
| nutrition\_id\_rbag | INDEX | nutrition\_id |
| gender\_age\_rbag | INDEX | gender, age |
| weight\_mg\_rbag | INDEX | weight\_mg |

**COMPLEX QUERIES AND DB OPTIMIZATIONS**

**Query 1 – Recipe by Nutritional Values**

Insert query itself here

Query explanation

Query optimization

**Query 2 – Cocktail by Nutritional Values**

Insert query itself here

Query explanation

Query optimization

**Query 3 – Recipe by Allergies**

Insert query itself here

Query explanation

Query optimization

**Query 4 – Recipe by Recommended Daily Intake**

Insert query itself here

Query explanation

Query optimization

**Query 5 – Daily Meal Plan by Recommended Daily Intake**

Insert query itself here

Query explanation

Query optimization

**Query 6 – First Trivia Quuestion**

Insert query itself here

Query explanation

Query optimization

**Query 7 – Second Trivia Question**

Insert query itself here

Query explanation

Query optimization

**CODE STRUCTURE AND GENERAL APP FLOW**

???

**EXTERNAL PACKAGES / LIBRARIES USED**

**BACK-END**

* Flask – a micro web framework for Python, used for development of the server-side of the application and communication with client-side.
* Python Requests – library for sending HTTP requests, used for API retrieval.
* WSGIServer - library for creating the HTTP server and run it at our server.
* Guy/Or/Yaniv – add here any other important libraries/frameworks you used in backend.

**FRONT-END**

* Bootstrap – a free open-source front-end framework, used for development of the user interface of the application.

**API USAGE, RETRIEVAL AND INTEGRATION PROCESS**

In our application we are using 3 API's: Yummly, CocktailDB and Nutritionix.

**Yummly API**

This is the main API we used. It was used for retrieving ~40,000 different food recipes. The retrieval from this API was done with the Python script "yummly\_retrival.py" which does the following:

* Sends a single request to the API for a list of all possible courses. The response is received as a JSON dictionary. Overall we used 11 different courses.
* For each course in the abovementioned dictionary, it sends 20 consecutive requests, each request for metadata of 200 recipes of that course. Therefore, metadata of 4000 recipes for each course was retrieved. This metadata was written into a JSON file.
  + The reason this is done in 20 requests instead of one request for 4000 recipes, is that Yummly tends to compress large responses, in a way that Python's Requests library was unable to deal with. Requesting for a relatively small amount of recipes kept the response un-zipped and so we were able to parse it.
* For each recipe's metadata dictionary in the created metadata JSON file, it sends a single request to the API, for the recipe's details, using its ID. For each recipe retrieved, it excludes irrelevant information from the JSON returned, and writes it to a JSON file of all recipe details.
  + Yummly API has a limit of 30,000 API calls per user per lifetime, so we used multiple accounts for the retrieval process, which were switched by the script while running.
* The script parses the recipes details JSON file into a CSV file containing all the relevant information that we need for the DB, for each recipe.
* Other issues we encountered:
  + Some recipes included non-latin UNICODE characters which we had to ASCII-encode.
  + A recipe's metadata JSON includes an array of its ingredients, and the recipe's details JSON includes an array of ingredient full-description line. The amounts of each ingredient needed to be parsed and converted from those 2 arrays by us.
  + Some ingredients' amounts were inconsistent (e.g. "two and a half", "2.5", "2 1/2", etc.) and we had to parse them to eliminate as many inconsistencies as we could.

**CocktailDB API**

This is the second API we used. It was used for retrieving ~500 different cocktail recipes. The retrieval from this API was done with the Python script "cocktailDB\_retrival.py" which does the following:

* Sends a single request to the API for a list of all possible ocktail categories. The response is received as a JSON dictionary.
* For each category in the abovementioned dictionary, it sends a single request for metadata of all cocktails in that category. The results are saved in a JSON array.
* For each cocktail's metadata dictionary in the created metadata JSON array, it sends a single request to the API, for the cocktail's details, using its ID. For each cocktail retrieved, it excludes irrelevant information from the JSON returned, and writes it to a CSV file of all cocktail details.
* Other issues we encountered:
  + Some cocktails included non-latin UNICODE characters which we had to ASCII-encode.
  + The amount of each ingredient in the cocktails had to be parsed and converted. Some ingredient names were written inconsistently, and had to be parsed and replaced.

**Nutritionix API**

This is the third API we used. It was used for retrieving ~3500 different ingredients. The retrieval from this API was done with the Python script "nutritionix\_retrival.py" which does the following:

* It retrieves a list of ingredients from the food and cocktail CSV/JSON files. It eliminates duplicates and left us with only the unique ingredients we needed to retrieve from the API.
* Given the abovementioned array of ingredients, it sends a request to the API for details of 10 ingredients each time. The results are saved in a JSON file of all ingredients details.
  + The reason we retrieved 10 ingredients each time, is that we wanted to retrieve as many ingredients in one request as possible, but we saw that when requesting too many ingredients, the retrieval fails to find results for some of them.
  + Nutritionix API has a limit of 100 API calls per user per day, so we used multiple accounts for the retrieval process, which were switched by the script while running.
* The script parses the ingredients details JSON file into a CSV file containing all the relevant information that we need for the DB, for each ingredient.
* Other issues we encountered:
  + Some nutritional values were in grams, while others while in mg, so we had to convert.
  + The amount of each ingredient in the cocktails had to be parsed and converted.

**Fixing the data**

After retrieving the ingredients from Nutritionix API, we noticed that not all ingredients were found by the API. Therfore, we had to eliminate the recipes that contained ingredients that weren't retrieved. This was done with a Python script called "retrival\_utils.py" which does the following:

* It goes over the cocktails and food CSV files and eliminates recipes that include ingredients not retrieved.
* It fixes the amounts of ingredients in each recipe to be the number of servings, according to the size of a single serving found in the ingredients CSV file.
* It outputs results to fixed cocktails and food CSV files, which are later used to construct the DB.

**BONUS**

1. **INFORMATION SECURITY:**

Our application is immune to SQL Injection attacks. This is true due to the fact most user input is received through buttons and drop-down menus. The only field that allows free-text input is the allergies field, and we applied input validation on this field, which blocks the possibility to input text that is non-latin letters (e.g. digits or symbols). We also implemented auto-complete on this field, and blocked the possibility to input text that is not part of the fixed ingredient list.

1. **USER INTERFACE:**

Our application's user interface is based on Bootstrap libraries. It was originally designed by us and was created from scratch without a template. The use of Bootstrap makes the app comfortable to use in many platforms. Each feature in the application is accompanied by a short explanation paragraph, in order to clarify to the user all the different possibilities. We used JQuery to make our user interface interactive and responsive to user's actions. Furthermore, our user interface is (objectively, of course) **very** pretty.