关于创建综合食品数据库

Lexington Whalen 计算与工程学院 南卡罗来纳大学 南卡罗来纳州哥伦比亚，美国 LA WHALEN@email.sc.edu

Brie Turner-McGrievy 阿诺德公共卫生学院 南卡罗来纳大学 南卡罗来纳州哥伦比亚，美国 brie@sc.edu

Matthew McGrievy 阿诺德公共卫生学院 南卡罗来纳大学 南卡罗来纳州哥伦比亚，美国 mjm@sc.edu

Andrew Hester 阿诺德公共卫生学院 南卡罗来纳大学 南卡罗来纳州哥伦比亚，美国 hesteras@mailbox.sc.edu

Homayoun Valafar 计算与工程学院 南卡罗来纳大学 南卡罗来纳州哥伦比亚，美国 homayoun@cse.sc.edu

摘要——以解决饮食失调为主要目标的研究，通常侧重于评估食品的营养成分含量，但只关注热量摄入。这类研究存在两个主要障碍。第一个障碍在于，每种食品的热量摄入量都是根据现有数据库计算得出的。第二个障碍则涉及将热量摄入作为营养成分含量唯一衡量指标的科学意义。由于研究必须依赖现有数据库，研究人员不得不寻找能够提供全面食品项目及其相应营养素的数据来源。光是这项搜寻就十分困难，而且即便找到，往往也意味着必须使用付费 ***API*** 服务。这些服务价格昂贵且无法定制，耗费了原本可用于研究其他环节的经费，最终提供的数据库既难以使用，也无法修改或贡献内容。在本研究中，我们推出了美国农业部食品数据库的新版本，涵盖杂货店以及餐厅或快餐店中的食品。目前，我们已收集约 ***150*** 万条食品条目，其中包括美国约 ***18,000*** 个品牌和 ***100*** 家餐馆。这些食品还关联了大量营养数据，从热量到饱和脂肪含量均有涵盖。数据以 ***MySQL*** 格式存储，分布在五个主要数据表中。我们还收集了这些食品条目在可获取时的图像，并将所有数据和程序脚本放入开源代码库，供任何人免费访问。

关键词—***MySQL***、表设计、美国农业部、食品数据库

I. 引言

美国农业部（USDA）提供了一个全面的数据库，收录了美国境内销售的食品 [1]。该数据库名为 FoodData Central，可在 usda.gov 免费下载，其中包含多个子数据库。每个数据库的描述可参见《美国农业部数据库》。截至目前（2022 年 6 月），美国农业部提供 JSON 和 CSV 格式的数据库文件。过去也支持 Microsoft Access 文件格式，但该支持已于 2021 [2] 年 10 月终止。尽管其中包含有用的数据，但结构相当复杂，整体上比实际所需的复杂得多。这种不便于用户使用的特性令研究人员难以快速、轻松地从数据库中获取食品或营养素信息。具体而言，该数据库包含已不再受支持的旧版内容、由不同供应商报告的营养成分，以及许多非技术领域的研究人员通常不会使用的其他组成部分。对这些附加信息进行适当规范化，使数据库模式变得更加复杂，也妨碍了数据库的便捷使用。最后，数量庞大的数据表之间存在一些关键关系，但这些关系并未明确定义，严重降低了数据库的实际使用效果。

虽然该数据库全面收录了杂货店出售的食品，但仍缺少许多常见餐馆食品条目；这是一个重大缺陷，因为众所周知，在美国，约有 40% 的成年人每天都会吃快餐 [3]，，且有 60% 的美国人承认每周至少在外吃一次晚餐 [4]。本文阐述了不同的数据来源和数据收集方法，并比较了它们各自的优缺点。促成本论文的原始研究选择了 Menustat，这是一个最初由纽约市卫生与心理卫生局建立的数据库 [5]。关于 Menustat 的更多详情，请参见 MENUSTAT 数据库一节。

II. 背景

许多由国际、国家和地区机构资助的研究旨在帮助参与者减重或调整饮食。

在许多研究中，研究人员会使用智能设备等技术，通过记录参与者每天的膳食来追踪他们的营养摄入。为此，这些研究的核心数据库需求是提供一系列可供选择的食品。SocialPOD 就是此类研究的一个近期案例。SocialPOD 将社交网络与一项互动机制结合起来，用户可以通过彼此互动、发布体重

更新、膳食或评论来“用磅数换积分” [6]。因此，SocialPOD 需要一个供用户选择的食品数据库，但找不到任何可行的免费数据集，只得花钱购买昂贵的应用程序接口。随后，SocialPOD 项目的研究人员表示有意创建一个开源的综合性食品数据库，由此促成了本文的撰写。由于缺乏实用且易于使用的数据库，这些研究只得转而使用 Nutritionix 等昂贵的商业数据库 [7], ——一个广受欢迎的食品数据库应用程序接口，因此也耗费了原本可以更有效地用于其他运营成本的资金。

III. 美国农业部数据库

美国农业部提供多种数据库下载方式，也提供应用程序接口，可查询其在线数据库。然而，由于数据十分复杂，要以可用形式下载并重新构建数据库是一项艰巨的任务。美国农业部提供的应用程序接口相对容易使用，但可扩展性有限，因为每天的应用程序接口请求数量受到限制，而且每个请求的响应速度相对较慢。美国农业部提供的数据库可下载版本采用两种常见的数据库文件格式：JSON 和 CSV。

截至2021年10月，所有可下载内容中共有37个 CSV 文件，但只有少数文件与计算大量营养素含量有关 [1]。这些文件列于表1中。请注意，文件名以 <FILE_NAME> 表示。

文件名	Size
<品牌食品_食品.csv>	745 M
<food.csv>	140 M
<food _nutrient.csv>	1.3 G
<食品_portion.csv>	5.8 M
<foundation food.csv>_	9 K
<nutrient.csv>	21 K
<sr_legacy_food.csv>	128 K

表1：文件及其大小。

每个 CSV 文件都经过解析，并使用 *--local-infile* 标志加载到 MySQL 中；此选项允许加载客户端计算机上的文件，而无需将文件放在服务器上，从而节省大量时间。需要注意的是，文件的加载是通过 MySQL 命令行完成的，因为图形用户界面版本会导致加载时间大幅增加。

表定义是使用 MySQL Workbench 的导入向导生成的。导入向导只需运行大约一秒钟，之后即可安全关闭，因为我们只需要表定义；实际数值将通过命令行导入。当然，表定义也可以手动创建，但由于列通常很多，使用 MySQL Workbench 会简单得多。使用导入向导时，务必注意系统有时会建议错误的数据类型；例如，fdc_id 的类型应为 BigInt，而不应只是 int（如果只用 int，添加数据时就会发生溢出）。图 1 展示了数据导入向导；请注意，用户可以从已识别的列中进行选择，并更改其数据类型。

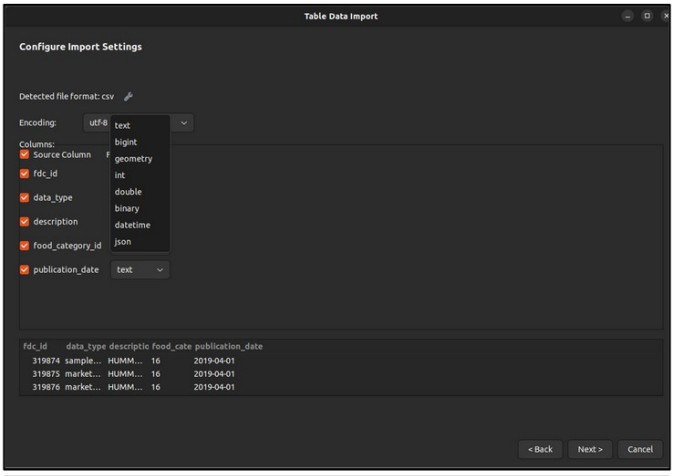

Figure 1: The Table Data Import Wizard

例如，以下命令是在 Ubuntu 20.04 的 MySQL 终端中运行的，用于载入 food_nutrient.csv，其中方括号内的值应替换为相应机器上的对应值：

```
从本地文件 [PATH_TO_CSV] 载入数据到表中
[TABLE_NAME] 字段以‘,’分隔，并用“”括起
行以‘\n’结束，忽略首行
```

对话 *1*。使用 *MySQL* 命令行将 *CSV* 文件加载到数据库中的命令。这样做是为了加快较大文件（例如 *<food_nutrient.csv>*）的加载速度。

虽然该命令相对容易理解，但一些常见错误包括忘记将 -- load-infile 标志设置为 1，或忽略第一行（因为第一行仅包含列名，而不是值）。

一个值得注意且有趣之处是，采用这种方法是因为 MySQL Workbench 比使用终端慢得多；MySQL Workbench 加载大型文件通常需要数小时甚至数天，而命令行版本即使处理最大的文件，耗时也不超过一小时。

导入所需的数据表后，下一步是为它们创建索引。索引本质上是一种数据结构，可由一个或多个列构成，并能提高操作速度。

可以使用以下 MySQL 命令为品牌食品数据表的 fdc ID 列添加索引：_ _

```
ALTER TABLE 品牌食品_
添加索引 fdc_id(fdc_id)
```

对话 *2*。为数据表添加索引。

需要建立索引的重要列包括所有含有 fdc ID 列的数据表中的 fdc_id，以及 food_nutrient 中的营养素 ID。其他可建立索引的列也可以建立索引，以进一步提升性能；不过，这两列最为重要，因为后续创建的数据表会大量使用它们。

美国农业部食品数据库可分为四类：品牌食品、基础食品、*SR* 旧版食品和实验食品。要将数据拆分到这四类数据表中，可以采用以下通用方法：

1. 使用 SELECT 从 food 或其他数据表（即：品牌食品）中选择所需的列 2. 根据 fdc ID_连接数据表 3. 根据相应的 ID 与 food_nutrient 连接

例如，要创建一个仅包含 SR 旧版食品条目，并包含 fdc_id、描述、千卡、份数和克_重量这些列的数据表，可以执行以下命令：

```
CREATE TABLE sr_legacy_食品 AS
    select fs.fdc_id, fs.描述, fn.amount as 千卡,
    fp.amount as 份数, fp.gram_weight
来自 food_sr_legacy_food fs
内连接 food_nutrient fn
    条件是 fn.fdc    _id = fs.fdc    _id
内连接 food_portion fp
    关于 fs.fdc_id = fp.fdc_id

其中
    -- id 1008 是千卡的 id
    fn.nutrient id = 1008_
```

对话 *3*。如何创建食品表。示例展示了如何访问特定数据表中的食品营养素。

请注意，要获取特定的营养素值，需要检查 food_nutrient（参见上述命令中的 1008）。

IV. MENUSTAT

Menustat 是一个专门收录餐厅数据的食品数据库。图 2 展示了该数据库；请注意其中包含餐厅、食物类别、份量和常量营养素信息。还要注意，可用的常量营养素不止能量。数据每年记录一次，可下载为 Excel 或 CSV 文件 [5]。这些文件的导入方式与美国农业部数据类似：先使用导入向导，再使用命令行。导入数据时一个常见问题是，某些行的格式可能不规范，导致行信息缺失。可以在 Linux 机器上运行 cat -v [NAME OF FILE].csv > [NEW NAME OF FILE].csv，将文件转换为已知编码来解决此问题。运行此命令后，部分行尾会变成 ^M；可使用*Vim*或任意文本编辑器将其删除。

在确定各列的数据类型时，最简单的做法是将所有列都指定为 TEXT 类型；这样可以确保所有行都能正确加载，因为尝试将数据加载到其他类型的列中时，可能会由于数据过大而导致行被跳过。加载完成后，可以更改这些列的数据类型。

menustat_id	restaurant	food_categor	description	item_description	serving_size	serving_size_tex	serving_size_un	serving_size_househo	energy_amoun
1	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ Cr…	Hot Coffee, Caramel Mocha, w/ Cr…	20	fl oz	…	…	340
2	Taco Bell	Sandwiches	XXL Grilled Stuft Burrito w/ Steak …	XXL Grilled Stuft Burrito w/ Steak…	434	g	1 Burrito	…	840
3	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ Cr…	Hot Coffee, Caramel Mocha, w/ Cr…	10	fl oz	…	…	170
4	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ Ski…	Hot Coffee, Caramel Mocha, w/ Ski…	…	…	…	…	…
5	Whataburger	Burgers	Double Meat Whataburger	Double Meat Whataburger, Whatab…	401	g	…	…	870
6	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ W…	Hot Coffee, Caramel Mocha, w/ W…	…	…	…	…	…
7	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ Ski…	Hot Coffee, Caramel Mocha, w/ Ski…	…	…	…	…	…
8	Papa John's	Pizza	Garden Fresh Pizza, Thin Crust, La…	Garden Fresh Pizza w/ Baby Porta…	111	g	1 Slice, 1/8 Pizza	…	220
9	Dunkin' Donuts	Beverages	Hot Coffee, Caramel Mocha, w/ W…	Hot Coffee, Caramel Mocha, w/ W…	…	…	…	…	…
10	Dunkin' Donuts	Beverages	Hot Coffee, Cinnamon, w/ Cream, …	Hot Coffee, Cinnamon, w/ Cream …	…	…	…	…	…
11	Taco Bell	Toppings & …	Avocado Ranch Dressing	Avocado Ranch Dressing, Condime…	14	g	…	…	80
12	Dunkin' Donuts	Beverages	Hot Coffee, Cinnamon, w/ Skim Mil…	Hot Coffee, Cinnamon, w/ Skim Mil…	…	…	…	…	…
13	Dunkin' Donuts	Beverages	Hot Coffee, Cinnamon, w/ Whole …	Hot Coffee, Cinnamon, w/ Whole …	…	…	…	…	…
14	Taco Bell	Toppings & …	Border Sauce, Fire	Border Sauce, Fire, Condiments & …	7	g	…	…	0
15	Taco Bell	Toppings & …	Border Sauce, Hot	Border Sauce, Hot, Condiments & …	7	g	…	…	0
16	Dunkin' Donuts	Beverages	Hot Coffee, Cinnamon, w/ Cream, …	Hot Coffee, Cinnamon, w/ Cream …	…	…	…	…	…
17	Dunkin' Donuts	Beverages	Hot Coffee, Cinnamon, w/ Skim Mil…	Hot Coffee, Cinnamon, w/ Skim Mil…	…	…	…	…	…
18	Taco Bell	Toppings & …	Border Sauce, Mild	Border Sauce, Mild, Condiments & …	7	g	…	…	0

Figure 2: A snippet of the Menustat database.

V. 表设计

创建数据表时，根据处理营养素的方式，大致有两种设置方法。第一种是“基于行的方法”：每种食品都有一列，列名为 nutrient_name，并有多行用于描述该食品的各种营养素。例如，西兰花这一条目中，每一行都是一种营养素的新条目；唯一会变化的列是营养素的名称和值，以及不同份量对应的份量。图 3 展示了基于行的方法；请注意，数据是按“行”录入的，每一行都包含下一个营养素类型。与图 4 中的“基于列的方法”形成对比。

这种方法的优点是可以轻松添加新的营养素值；例如，如果西兰花

description	brand_owner	serving_size	serving_size_un	nutrient_name	nutrient_unit	nutrient_amount
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	fatty acids, total…	G	8
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	fatty acids, total…	G	3
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	fatty acids, total…	G	2
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	energy	KCAL	130.05
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	total lipid (fat)	G	14
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	sugars, total inc…	G	1.01
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	sodium, na	MG	830.4
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	energy	KCAL	9.6
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	carbohydrate, b…	G	1.01
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	protein	G	1.99
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	sugars, total inc…	G	1.8
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	fatty acids, total…	G	4.49
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	cholesterol	MG	26.4
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	sodium, na	MG	1597.2
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	potassium, k	MG	286
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	iron, fe	MG	1.06
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	calcium, ca	MG	70.4
CAMPBELL'S SLOW KETTL…	CAMPBELL SOUP COMPANY	440	g	fiber, total dietary	G	1.76

Figure 3: A snippet of the “row-based” approach for the USDA Branded foods table.

测出了维生素D含量，只需添加一个包含该值的新行即可；无需添加新列。无需添加新列这一点的优势，只要想象一下所有营养类别都以列值的形式存储，就很容易看出来。这意味着，如果西兰花新增了一种此前未记录的营养素，就必须为整个数据库新增一列，而不是只添加一个包含这种新营养素的行。主要缺点是，这种方法的使用略有不便；如果事先知道要查询哪些营养素，那么直接返回一行并在其中包含所有可能的营养素或许更简单。这正是第二种方法的出发点。

第二种方法是“基于列的方法”，每一行都包含一种食品的全部数据。这种方法将所有必要数据封装在一行中，因此比基于行的方法更快得到结果。图 4 展示了基于列的方法。请注意，某种食品所需的所有营养素都位于同一行。其优点之一是，如果已知所有食品都需要哪些营养素，数据库的大小就会显著缩小，达到

对于较大的数据表，可缩小至吉字节级别。作为比较，美国农业部品牌食品表采用基于列的方法时，文件大小约为 3.47 吉字节，而采用基于行的方法时，文件大小为 6.7 吉字节。不过，这种方法的一个主要缺点是，添加此前未定义的营养素值时，需要为所有行添加并填充整列；这是一项代价高昂的操作。

description	brand_owner	serving_size	serving_size_un	fiber_amoun	fiber_uni	energy_amoun	energy_uni
WESSON Vegetable Oil 1 GAL	Richardson Oilseed Products (US) …	15	ml	0	G	130.05	KCAL
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	240	ml	0	G	9.6	KCAL
CAMPBELL'S SLOW KETTLE SO…	CAMPBELL SOUP COMPANY	440	g	1.76	G	360.8	KCAL
CAMPBELL'S SLOW KETTLE SO…	CAMPBELL SOUP COMPANY	440	g	1.76	G	360.8	KCAL
SWANSON BROTH CHICKEN	CAMPBELL SOUP COMPANY	240	ml	0	G	9.6	KCAL
CAMPBELL'S SOUP BEAN AND …	CAMPBELL SOUP COMPANY	412	g	9.89	G	251.32	KCAL
SWANSON BROTH BEEF	CAMPBELL SOUP COMPANY	411	g	0	G	16.44	KCAL
PREGO SAUCES TOMATO BASIL	CAMPBELL SOUP COMPANY	120	ml	3	G	69.6	KCAL
CAMPBELL'S SOUP TOMATO	CAMPBELL SOUP COMPANY	120	ml	0.96	G	90	KCAL
CAMPBELL'S PASTA SPAGHETTI…	CAMPBELL SOUP COMPANY	443	g	5.32	G	385.41	KCAL
BUSH'S Bourbon and Brown Suga…	Bush Brothers And Company	130	g	4.55	G	162.5	KCAL
BUSH'S Organic Garbanzo Beans …	Bush Brothers And Company	130	g	4.42	G	136.5	KCAL
BUSH'S Smokehouse Tradition Gril…	Bush Brothers And Company	130	g	4.68	G	167.7	KCAL
PEPPERIDGE FARM BREAD TUS…	PEPPERIDGE FARM	57	g	1.03	G	149.91	KCAL
PEPPERIDGE FARM BREAD GA…	CAMPBELL SOUP COMPANY	50	g	1	G	180	KCAL
SWANSON BROTH CHICKEN	CAMPBELL SOUP COMPANY	240	ml	0	G	9.6	KCAL
PEPPERIDGE FARM COOKIES GI…	PEPPERIDGE FARM	28	g	0	G	129.92	KCAL
PEPPERIDGE FARM BREAD TUS…	PEPPERIDGE FARM	57	g	1.03	G	140.22	KCAL

Figure 4: A snippet of the “column-based” approach for the USDA Branded foods table.

由于美国农业部数据库最初采用基于列的形式，因此需要进行一些调整，将数据重组为基于行的形式。具体做法是创建一个数据表，其中的列用于存放确定为必要的营养素，以及基于列的数据表中包含的其他有价值的信息，例如 FDC索引或食品名称。由于*Menustat*数据表原本就采用基于行的形式，因此无需进行此类修改。

VI. 其他数据来源

在发现 *Menustat* 之前，我们曾研究其他收集餐馆相关数据的方法。最有前景的做法

是编写脚本从互联网上收集所需数据，这种方法也称为网页抓取。我们使用 *Beautiful Soup* [8] 解析由 Python 的 *requests* 库 [9]获取的 HTML。一般工作流程是，先对包含食品数据的网站发起初始查询，再逐一访问餐馆，抓取其中所有食品项目的数据，最后转到下一家餐馆。需要注意的是，这种方法要求目标网站的 HTML 结构保持一致；如果网站突然更改数据的组织方式，网页抓取程序就可能失效或返回无效数据。

编写了这样一个脚本来处理网站 *menuwithnutrition.com*。该网站按字母顺序列出餐厅，每家餐厅都列出其提供的食品类别以及具体食品，因此其结构足够一致，便于进行抓取 [10]。下方提供了 *MenuWithNutrition* 餐馆列表的截图

图 5。请注意，该列表按字母顺序排列。点击一家餐厅后，页面会显示该餐厅提供的所有食品；随后可以点击每种食品，获取其营养数据。我们从该网站抓取了约 205,000 种食品，每条食品记录都包含餐厅名称、食品名称以及脂肪、胆固醇、钠、碳水化合物、蛋白质、饱和脂肪、反式脂肪、膳食纤维和糖的含量，并注明了各项的计量单位。

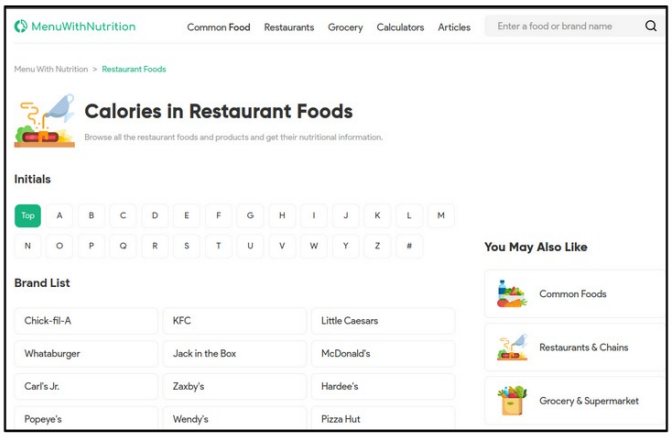

Figure 5: The restaurant listing of *MenuWithNutrition*.

VII. 食品图像

该数据库可与包含数百万张食品图像的文件夹配合使用。每张图像都经过调整大小，以减小文件体积。对于包含餐厅或品牌信息的数据库，图像的命名格式为 [餐厅 / 品牌名称][食品名称].png，其中仅保留字母和数字字符；所有非字母数字字符都会被忽略。

为了获取这些图像，我们创建了一个与上述网页抓取脚本类似的脚本。该程序使用线程，将查询网页以及保存／调整图像大小的任务分配到客户端计算机的各个核心上。需要注意的是，根据用户的网速和核心数量，可能需要进行一些调整，以确保不会因 requests 频率过高而遭到 IP 封锁。可以通过调整每个处理器的休眠时长来实现这一点。

VIII. 数据存储库：

所有数据库文件及生成这些文件的相应代码均可在以下位置找到：

https://github.com/lxaw/ComprehensiveFoodDatabase

IX. 结论

许多研究都会使用食品数据，但可靠且免费的此类信息来源却十分匮乏。本项目探讨了不同数据源和表设计的优点，并说明了如何利用这些数据源创建数据表或抓取图像。本项目旨在提供一种可行的解决方案：建立一个数据库，其中餐馆食品和非餐馆食品的条目数量均足够，同时为每个条目提供图像。不过，项目的主要目标是为未来的研究人员留下一个可用的免费数据库，从而以更低的成本开展更多研究。

X. 致谢

本出版物所报告的研究得到了美国国立卫生研究院的资助。

XI. VIII. 参考文献

[1] Bodner-Montville、Janice、Jaspreet KC Ahuja、Linda A. Ingwersen、Etta Susanne Haggerty、Cecilia Wilkinson Enns 和 Betty P. Perloff。《美国农业部膳食研究食品与营养数据库：在线发布》。《食品成分与分析》19 (2006)：S100-S107。[2] “FoodData Central”，Usda.gov，2013。https://fdc.nal.usda.gov/download-datasets.html[3] C. Fryar、J. Hughes、K. Herrick 和 N. Ahluwalia，“产品 - 数据简报 - 第320期 - 2018年9月”，美国疾病控制与预防中心，2018年10月。https://www.cdc.gov/nchs/products/databriefs/db322.htm [4] 美国疾病控制与预防中心，“《糖尿病患者外出就餐》”，美国疾病控制与预防中心，2019年9月19日。
h
t
t
p
s://www.cdc.gov/diabetes/managing/eat-well/eating-out.html [5] “MenuStat”，*MenuStat*。https://www.menustat.org/（访问于2022年11月29日）。[6] B. Turner-McGrievy 和南卡罗来纳大学，“完善并试点测试促进健康行为的社交网络：Social Pounds Off Digitally（Social POD）研究”，clinicaltrials.gov，2015年6月23日。https://clinicaltrials.gov/ct2/show/NCT02344836（访问于2022年12月6日）。[7] Nutritionix，“Nutritionix”，Nutritionix.com，2019。https://www.nutritionix.com[8] “Beautiful Soup 文档 — Beautiful Soup 4.4.0 文档”，beautiful-soup-4.readthedocs.io。https://beautiful-soup-4.readthedocs.io/en/latest/[9] K. Reitz，“requests：面向人类的 Python HTTP。”，PyPI。https://pypi.org/project/requests/[10] “您的可靠营养数据库 - Menu With Nutrition”，www.menuwithnutrition.com。https://www.menuwithnutrition.com/（访问于2022年11月29日）。