

Is It Possible to Satisfy the RDA for All Nutrients?

Justin Sharber

Draft: 9/3/2019

Executive Summary

The nutrient Recommended Daily Allowances (RDA) represent the U.S. government's recommendations for the daily intake of various nutrients. The nutrients of food consist of macronutrients (protein, carbohydrates, and fat), vitamins, and minerals.

This report takes interest in satisfying RDA using natural foods and without the help of nutritional supplementation or fortification. It asks two questions.

1. Is it possible for a young man to satisfy his RDA across all nutrients in a natural, healthy way (without supplementation, etc.)? Or are some RDA set too high to be satisfied in this way?
2. What is the ideal diet of natural, healthy food toward meeting all RDA?

The hypothesis was that it is *not* possible to satisfy all RDA under the given constraints. Practically speaking, the hypothesis is thoroughly refuted. But technically, one nutrient, Fluoride, does prove impossible to meet.

An analysis of foods for nutritional value by calorie (or “nutrient density”) finds that leafy green vegetables, mushrooms, and oysters to be the most nutritious. But in terms of nutritional value by mass, seafood, seeds, and nuts are the most nutritious.

The report also constructs ideal diets for a few different types of diets (ex: vegetarian, vegan). It approaches the construction of the ideal diet as a special kind of knapsack problem, where the value function changes with each iteration of food choice. The choice of loss function is an important component of the calculation. This report favors a modified sum-of-squares loss function, here called a “tapered SOS” loss function.

The report is an exercise in data analysis. It does not represent mature nutritional research, nor does it represent nutritional advice. Its findings are totally bland and consistent with general nutritional advice and conventional wisdom: eat a variety of natural foods.

Please don't die of shock.

Specifying the Problem

Background

Nutrients fall into several categories. *Macronutrients* are the nutrients people need in large amounts: protein, carbohydrates, and fats. Fiber is a class of complex carbohydrates and is not technically a separate macronutrient, but it does have a separate RDA, so is treated separately here. *Micronutrients* are those needed only in small amounts, consisting of vitamins and minerals. *Vitamins* are complex, organic compounds. Vitamins divide into the *water-soluble* and *fat-soluble* subgroups, depending on solubility. Minerals are elements or inorganic compounds. Minerals divide into the *major* and *trace* groups. *Major minerals* are needed in larger amounts, on the order of a gram per day. *Trace minerals* are needed in much smaller amounts, on the order of a milligram per day. *Electrolytes* are ions necessary for the conduction of the electrical impulses of neural activity.

The U.S. government has made dietary recommendations to the public since 1894, and during World War II, the recommendations were first cast as RDA. RDA are set to meet the dietary requirements for 97% of the population. The calculation starts with the Estimated Average Requirement, the point at which 50% of the population shows some sign of deficiency. Nutrition scientists use a bell-curve model and add two standard deviations of intake to result in the RDA. (Smolin and Grosvenor, 34-8)

Nutrition science involves multiple layers of variability. One important variable is “bioavailability,” the proportion of an ingested nutrient that the body absorbs and uses. For minerals, bioavailability varies by food, meal, and person. For example, minerals from animal sources are better absorbed than minerals from plants. If multiple minerals of similar kinds are ingested at the same time, they can compete for absorption. And the body can be selective about which minerals to absorb. (Smolin and Grosvenor, 442-3)

Data

This project will use the USDA National Nutrient Database for Standard Reference, Release 28. ([My Food Data: Nutrition Facts Database Tools and Spreadsheet](#)) RDA values and calorie goal come from Smolin and Grosvenor (2013). A list of RDA is also available online, although several values differ ([Health Supplements Nutritional Guide: Recommended Daily Allowances](#)).

Caveats

This report is an exercise in data analysis. This project does not represent mature nutritional research, *nor does it represent nutritional advice*. This report's analysis of nutrients and ideal diet are based solely on the data.

Some nutrients have Upper Limits set as well, but this project will not factor these in. It will make some note that one can have too much of a nutrient, but rely on the RDA for a target.

Analyzing the Data

Curating the Data

The USDA's Standard Reference 28 dataset contains records for over 6,300 foods. Particularly for meats, the same basic food could have hundreds of different records, corresponding to different cuts and preparation methods, representing a fair amount of redundancy for the purposes of this project. The dataset also includes many foods that are not well suited for the project, including the following.

Egg, whole, raw, fresh	– inedible
Dairy cream substitute, flavored, liquid	– artificial
Ice cream, soft serve, chocolate	– not nutrient dense, unnecessary
Spices, chili powder	– cannot be eaten in significant portions, could skew results
KRAFT FREE Singles American	– name-brand, beyond scope

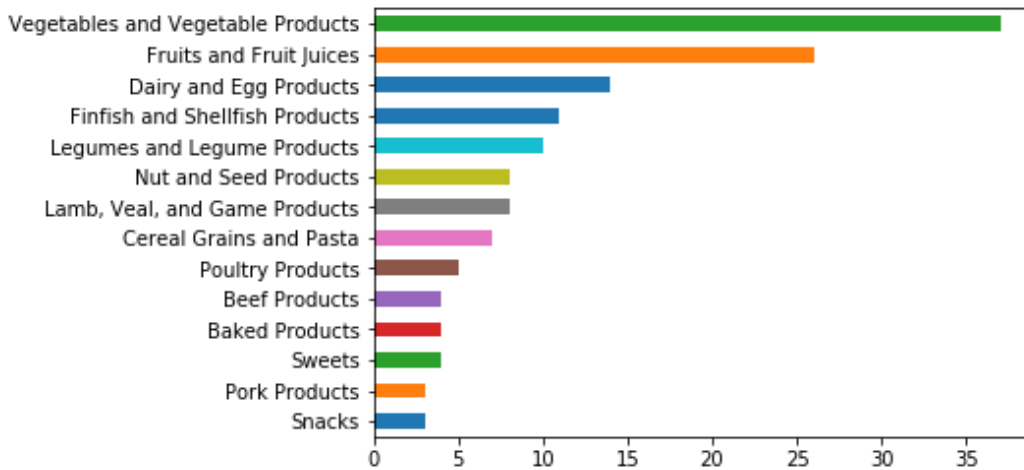
- *Egg, whole, raw, fresh*
(inedible)
- *Dairy cream substitute, flavored, liquid*
(artificial)
- *Ice cream, soft serve, chocolate*
(not nutrient dense, unnecessary)
- *Spices, chili powder*
(cannot be eaten in significant portions)
- *KRAFT FREE Singles American*
(name-brand, beyond scope)

It proved easier to choose good foods directly, rather than to develop rules that would filter out every unhelpful record. Thus, a subset of the foods was chosen on an individual basis. Foods were chosen on their most common preparation methods. For meats, baked or roasted preparations were preferred. Fruits were taken raw. Vegetables were taken raw, unless not usually eaten in that form; carrots were taken as raw, but broccoli was taken as cooked.

The curated subset contains 158 core foods. The vegetable and fruit categories represented the most real variety in food type—37 different vegetables were included. Out of 961 records for beef, 4 were selected:

- Braised beef brisket
- Braised arm pot roast
- Roasted bottom round
- Broiled top sirloin

The counts of food groups in the curated data set are shown below.



Two extra groups of foods were set aside:

- “Unusual” foods, including tea and lard, which might be interesting to study, but might give weird results if incorporated as foods for the “ideal diet” project.
- “Exotic” foods, including eel and seaweed, which are perfectly natural foods, but are not local (to the author), so do not fit the spirit of the project.

After setting these food groups aside, 138 foods remained.

This project examines all nutrients that are both included as features in the SR28 dataset and have set RDA. There are 26 of these, below.

Nutrients in the Project	
Macronutrients	Protein, carbohydrates, fiber.
Vitamins	Vitamin A, thiamin (B1), riboflavin (B2), niacin (B3), vitamin B5, vitamin B6, folate (B9), vitamin B12, vitamin C, vitamin D, vitamin E, choline.
Minerals	Calcium, copper, fluoride, iron, magnesium, manganese, phosphorus, potassium, selenium, sodium, zinc.

Best Sources of Nutrients

The main questions of this project revolve around nutrient density: how much nutrition each food has, relative to the calories that food contains. Thus the analysis of nutrients will be in these terms.

Which foods are the best sources of each nutrient, per calorie? A table was generated connecting each nutrient to its best food source, below.

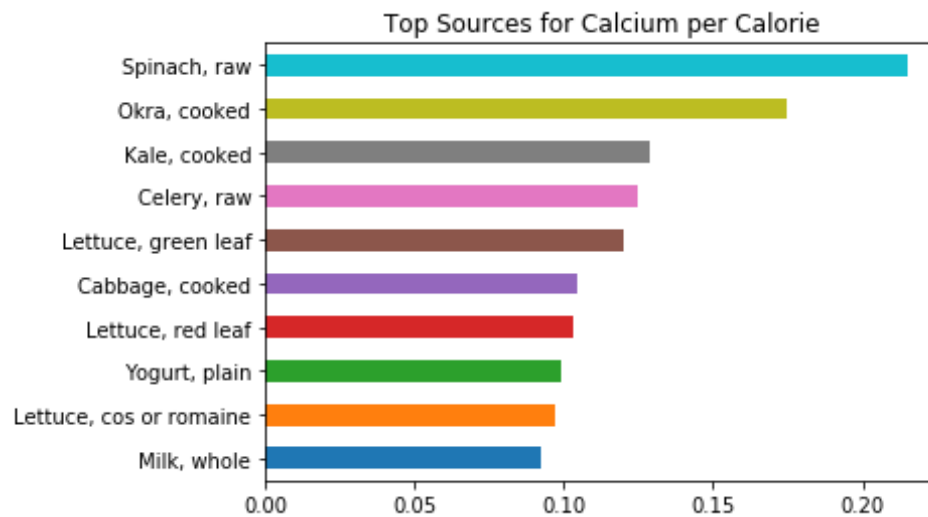
Best Sources of Nutrients						
Nutrient	Food	Proportion RDA		Nutrient	Food	Proportion RDA
Protein	Fish, haddock	0.2		Vitamin A	Lettuce, cos or romaine	8.54
Carbohydrates	Limes, raw	0.14		Vitamin B1	Asparagus, cooked	0.31
Fiber	Raspberries, raw	0.16		Vitamin B2	Mushrooms, white	0.41
Calcium	Spinach, raw	0.22		Vitamin B3	Mushrooms, white	0.5
Copper	Mushrooms, morel	1.12		Vitamin B5	Mushrooms, white	0.77
Fluoride	Asparagus, cooked	0.01		Vitamin B6	Okra, cooked	0.33
Iron	Mushrooms, morel	2.46		Vitamin B9	Spinach, raw	1.05
Magnesium	Spinach, raw	0.43		Vitamin B12	Mollusks, oyster	6.41
Manganese	Spinach, raw	0.85		Vitamin C	Peppers, sweet	1.48
Phosphorus	Mushrooms, morel	0.45		Vitamin D	Mushrooms, morel	0.55
Potassium	Spinach, raw	0.26		Vitamin E	Spinach, raw	0.29
Selenium	Mollusks, oyster	0.89		Choline	Egg, whole	0.17
Sodium	Asparagus, cooked	0.36				
Zinc	Mollusks, oyster	2.6				

The best source for protein per calorie is haddock. Raspberries provide the most fiber. Spinach, not milk, provides the most calcium. The egg appears, but as the best source for only one nutrient: Choline. And boring white mushrooms top three B vitamins.

Some foods are the best source for multiple nutrients. Clearly, the property of *being a best source* is clearly not uniformly spread across foods. Just 5 foods represent best sources for 18 of the 26 nutrients, shown below.

Best Source Food	# Nutrients
Spinach, raw	6
Mushrooms, morel	4
Mushrooms, white	3
Mollusks, oyster	3
Asparagus, cooked	2

The top sources for calcium per calorie are shown below. Milk may be a tasty source of calcium, but per calorie it falls well below other sources.

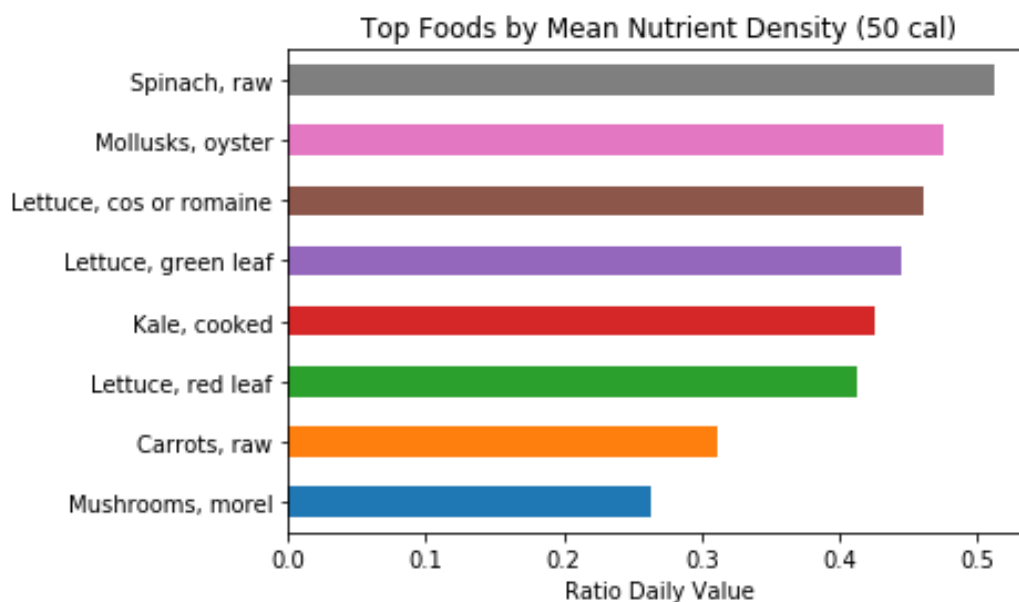


The most difficult nutrient to find naturally is Fluoride. Asparagus is the best source for this nutrient, but a decent portion of 50 calories only provides 1% of the RDA.

The Most Nutrient-Dense Foods

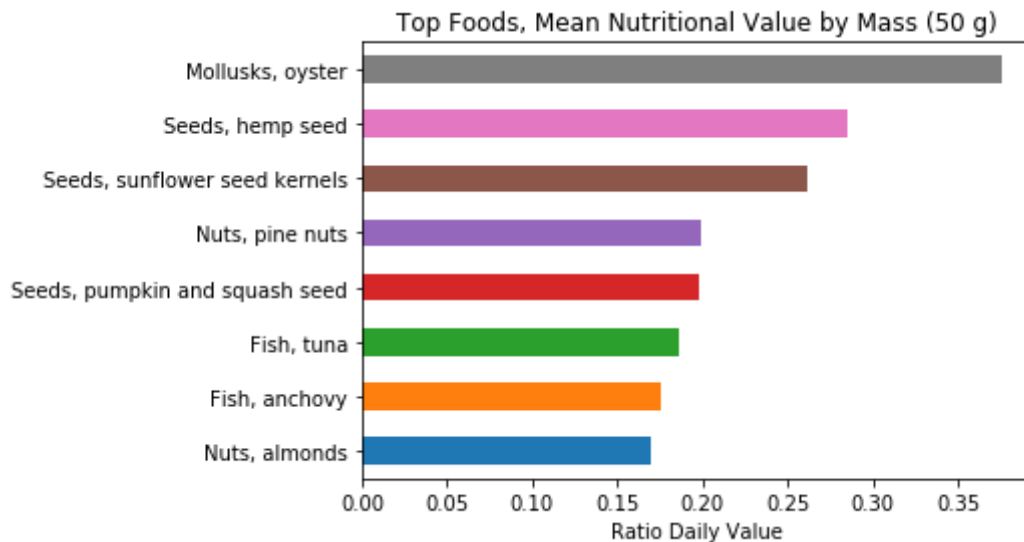
Which foods are the most nutrient-dense? For this analysis, a serving size of 50 calories was chosen, or about an egg's worth of calories (60). Proportions of RDA values were tallied for all nutrients without weights. This is to say, all nutrients were treated as equal: selenium had the same import as protein.

The top eight most nutrient-dense foods are shown below. Of these top foods, most are vegetables. Five of the top eight foods are leafy greens. Only one food is an animal.



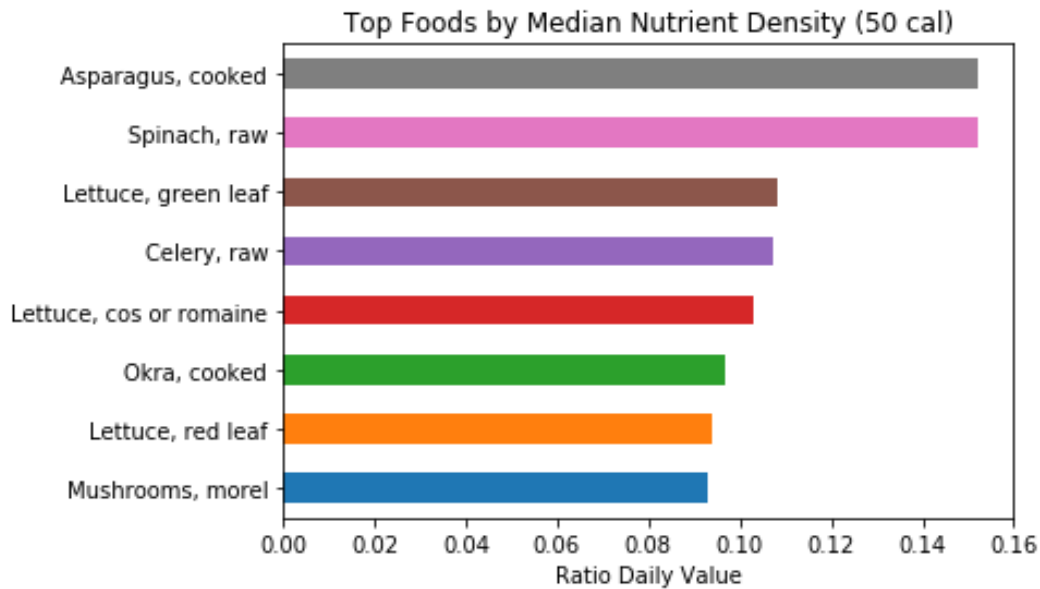
50 calories of spinach provides a huge amount of nutrition with an average of 51% of the RDA across nutrients. But one should note that, given that spinach is a low-calorie food, this is a large helping—about 5 cups chopped.

The prominence of vegetables in this ranking can be attributed to two factors: vegetables are high in nutrients, and they are also low in calories. While nutrient density is the operative notion for this project, it may be helpful to compare this to nutritional value by mass. This ranking is shown below.



This is a very different list. Here the most valuable foods are seafood, seeds, and nuts. Only one food, oysters, appears in both lists.

A potential problem with ranking foods by density is skewness. The top foods tend to have skewed distributions of nutrients: very high amounts in a few nutrients, modest or low values in the rest. In this case, the median can be a more helpful sense of “average” than the mean. The most nutrient dense foods by median are shown below.



The median nutrient densities are much lower than mean, signaling the skew of the nutritional profiles. But the ranking is quite similar, with spinach is still tied for first. The one animal has fallen from the list, leaving only plants and mushrooms.

In the next section, this project will attempt to determine whether all RDA are satisfiable based on natural foods by simulating an idealized diet. Based on this analysis, we would expect vegetables to be the first choices in that diet, and they will be.

Is it Possible to Satisfy All RDA?

Hypothesis

Our ancestors, from cavemen to the people of the 19th Century, should have been able to get all the nutrition they needed from the foods available. If the RDA are appropriately set, if they are correct, then it should be possible to satisfy them with natural foods alone. Let *natural food* denote any food that is grown or raised on a farm, or produced with parochial technology from other natural foods. Thus chicken, apples, and pasta are natural foods, but soda and cheese puffs are not.

The term “natural foods” excludes two groups of nutrition, broadly construed. One is supplements, which includes nutrient supplements like vitamin pills, and foods artificially enriched with nutrients, such as many breakfast cereals. Supplements make it trivial to satisfy the RDA, as a multivitamin contains 100% or more of the RDA for many nutrients. The other group is junk foods. Because junk foods are very low on nutrition, including them would make little difference except to increase the range of foods that would have to be considered. But either way, they do not count as “natural” in the definition of this project and do not need consideration. In short, “natural foods” exclude enriched foods because they could make satisfying the RDA too easy, and junk foods because they will not help.

This project selects a moderately active 28-year-old male of healthy weight as its abstract test subject. The average calorie need per day for this subject to maintain weight is 2600. (Smolin and Grosvenor, 47) The subject is further specified as 5’11” and 160 lb.

Subjectively, this author is suspicious of RDA values. They seem too high, with foods often being counted as good or excellent sources of a nutrient while providing a small fraction of the RDA. Also, despite a thoroughgoing scientific basis for RDA, the specification of these values does not entail that all RDA can be satisfied at the same time on the basis of natural foods. Thus, this project’s hypothesis is framed in the negative.

Basic Hypothesis	It is not possible for a 28-year-old male to satisfy all his nutritional RDA on the basis of natural foods at a maintenance calorie intake.
-------------------------	---

Methodology

In terms of data science, the question of whether a person can satisfy their RDA within a calorie limit is a complex version of the knapsack problem. The basic knapsack problem takes the paradigm of trying to pack a knapsack with valuable things, like gold nuggets and computer chips. The goal is to maximize the value of the knapsack (i.e. its contents) within limited space or weight. This problem is relatively simple. It has a single variable for cost (space or weight), and a single variable for value (price). Both variables remain fixed for each item and never need to be recalculated.

But with the current problem, there are many factors (different nutrients) that contribute to the nutritional value of a food and must be combined explicitly. More importantly, the value of eating a food depends on what nutrients have already been consumed. The value of a nutrient is small or zero once the RDA for that nutrient has been satisfied.

These differences complicate the project over a basic knapsack problem. Overall value is dynamic. But the relative value of a food changes depending on the contents of the "gut." Thus, the value of each food must be recalculated continually throughout the process.

A Python class called a "Gut" was created for the purpose of running the backpack analysis. For each iteration, the Gut evaluates all foods for which has the highest nutritional value for it at that stage, relative to the nutrition the gut already contains. Every time the Gut eats a food, it adds the food's name, its nutrients, and its calories as contents. In the next iteration, the balance of nutrients will have changed, and the "Gut" must re-evaluate foods to decide which to add next.

A serving size of 50 grams was chosen for all foods, about the mass of a medium egg.

Testing the Hypothesis

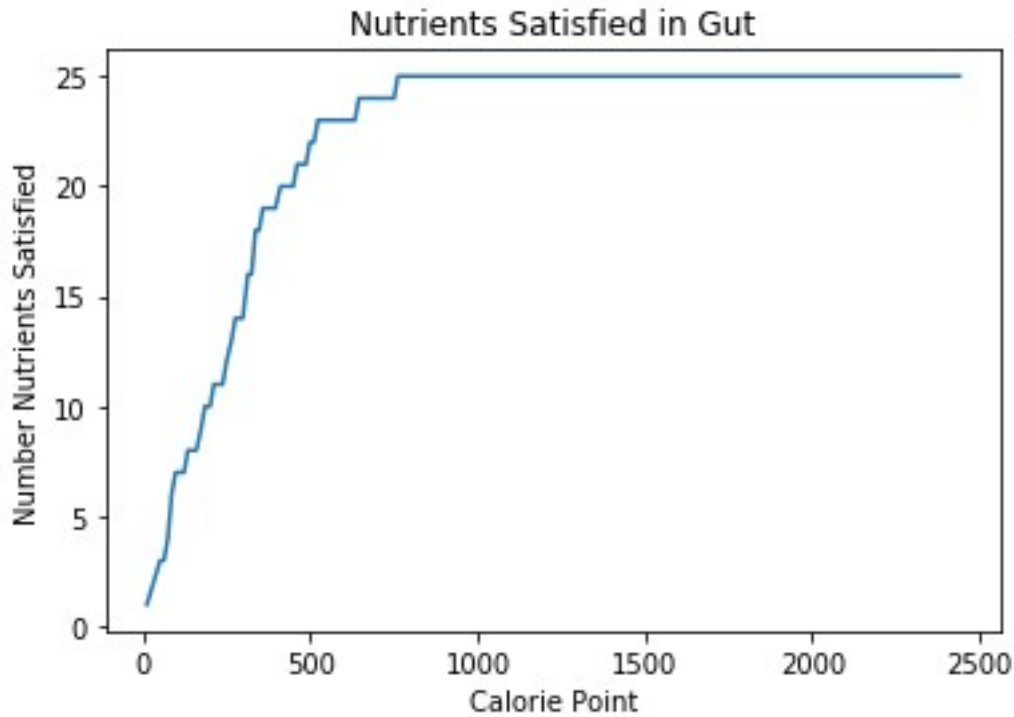
On the initial run of the Gut, almost all nutrients *were satisfied* on the basis of the curated set of natural foods. Only one, Flouride, was not. The median satisfaction of nutrients was 7.68 RDA values, over 7 times the required amount.

The Gut started with spinach, oysters, and mushrooms. But at some point, the Gut started eating asparagus and nothing else. The disconcerting diet is summarized below, in servings of 50 grams.

Food	Servings
Asparagus, cooked	173
Spinach, raw	17
Cauliflower, raw	10
Mushrooms, morel	9
Mushrooms, white	3
Raisins, seedless	1
Mollusks, oyster	1

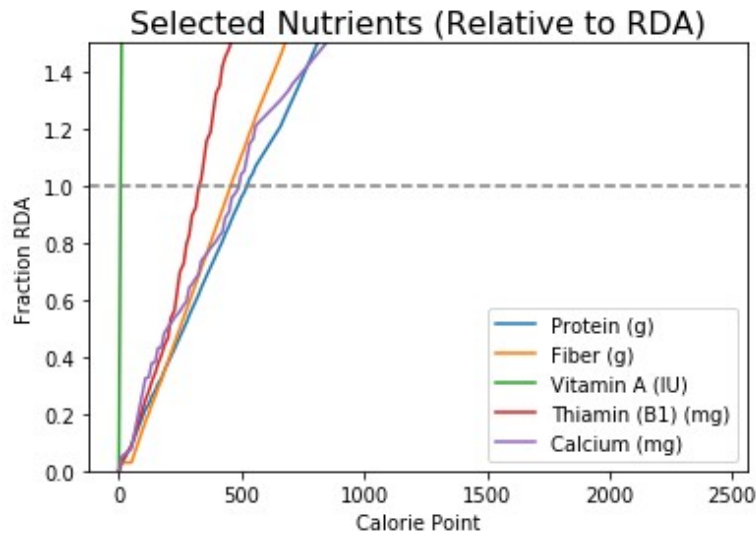
The Gut consumed about 8.5 kilograms, or 19 pounds, of cooked asparagus. Nobody wants that.

One would expect more variety in the diet. However, there is a reason for this behavior. The Gut satisfied *all other nutrients* at only 760 calories (shown below). From that point on, it devoted its attention obsessively to maximizing Fluoride.



Asparagus is the best source of Fluoride per calorie, but only provides 1% of RDA in a significant, 50-calorie serving. The value function for this version of the Gut is rather basic; it focuses only on maximizing unsatisfied nutrients. With no other goals, the Gut obsessively ate asparagus until it had filled its calorie goal.

Several key nutrients were chosen to show a profile for the Gut's progress over time. The initial plot treats the RDA as a ceiling. But with a diet of vegetables, the Gut's nutrients quickly go off the chart. In particular, the Gut's RDA for vitamin A is satisfied in the first 12 calorie snack; this nutrient is shown only as a vertical line.



The Gut's nutrient levels at the end are impressive, even for nutrients with the lowest levels of satisfaction. It satisfied almost half of its requirement for Fluoride. It satisfied roughly 1.5 times its requirement for vitamin D, 3 times the requirement for calcium, and 4.5 times its requirement for protein (almost solely on vegetables). Its most over-satisfied nutrient was vitamin A, 53 times the RDA.

Technically, the hypothesis is falsified by the one outstanding nutrient. But for all intents and purposes, the hypothesis is wrong. If calories are the only limitation, the RDA for almost all nutrients is easily achievable on the basis of natural foods. Thus the best *representation* of the results is a rejection of the hypothesis.

Basic Hypothesis	Rejected
Median Nutrient Satisfaction	7.7
Unsatisfied Nutrients	1 / 26

The Ideal Diet, By the Numbers

Eating 19 pounds of asparagus is hardly ideal. This report will move on to two follow-up questions.

1. Can a more realistic and balanced diet be achieved on the numbers alone, particularly by adjusting the Gut's value function?
2. How realistic is it for a person to satisfy their RDA with less calculated, more random eating?

The Value Function

The basic implementation of the Gut has two shortcomings.

1. The Gut will stop at nothing (short of the calorie goal) to try to satisfy an outstanding nutrient, even eating one food (the asparagus) in huge amounts.
2. The Gut actually gets far too much of many nutrients. Getting several times the RDA for some nutrients is no problem, but a median satisfaction of over 7 times the recommended amount is too much.

In fact, both of these shortcomings can be solved with one feature: penalizing over-satisfaction. The basic value function of the Gut places no penalty on taking in more of a nutrient than is required. Penalizing over-satisfied nutrients naturally solves the second problem.

One could solve the first problem of obsessive eating with a separate factor, penalizing foods that have been eaten previously. But penalizing over-satisfaction also steers the Gut *away from* nutrient dense foods as the Gut's nutrients fill up. In turn, this tends to solve the first problem anyway.

Two parameters were added to the Gut to penalize over-satisfaction. The first is a *target factor*. The target factor specifies what proportion of the RDA is considered ideal. 1.0 times the RDA (or 100%) is the natural choice, but raising that ceiling a little to 1.5 gives the Gut some wiggle room in satisfying nutrients.

The second parameter is an *over-satisfaction factor*, specifying the penalty for over-satisfying nutrients. An over-satisfaction factor of 1.0 means that the Gut penalizes over-satisfaction the same as under-satisfaction, and treats getting too much as just as bad as getting too little. An over-satisfaction factor of 0 (used in the original hypothesis test) means the value function places no penalty on over-satisfied nutrients.

To avoid belaboring the report, the details of this analysis will be skipped. The ideal over-satisfaction factor appears to be about 0.1. The next section analyzes a diet generated on this setting.

The Ideal Diet, By the Numbers

The “Ideal Diet” presented here is the best diet this project can produce. It is intended to maximize nutrient satisfaction, based solely on the data provided and a knapsack methodology with a “greedy” algorithm. With a target of 1.5 times the RDA for nutrients and an over-satisfaction factor of 0.1, the Gut chose a diet which is summarized below. With this diet, the median nutrient satisfaction was dramatically reduced to 2.5 times the RDA, while Fluoride remained the only unsatisfied nutrient.

The Ideal Diet			
Food	Serv.	Food	Serv.
Turnips, cooked	9	Mushrooms, morel	2
Honey	8	Mushrooms, Chanterelle	2
Okra, cooked	8	Nuts, almonds	1
Asparagus, cooked	7	Pork, cured	1
Mushrooms, white	4	Mollusks, oyster	1
Spinach, raw	3	Fish, salmon	1
Artichokes, (globe or french)	2	Broccoli, raw	1
Pickles, cucumber	2	Egg, whole	1
Cauliflower, raw	2		

This ideal diet shows more variety than that from the basic hypothesis test. There is a wider variety of vegetables, and the Gut enjoyed small helpings of pork, egg, fish, and nuts.

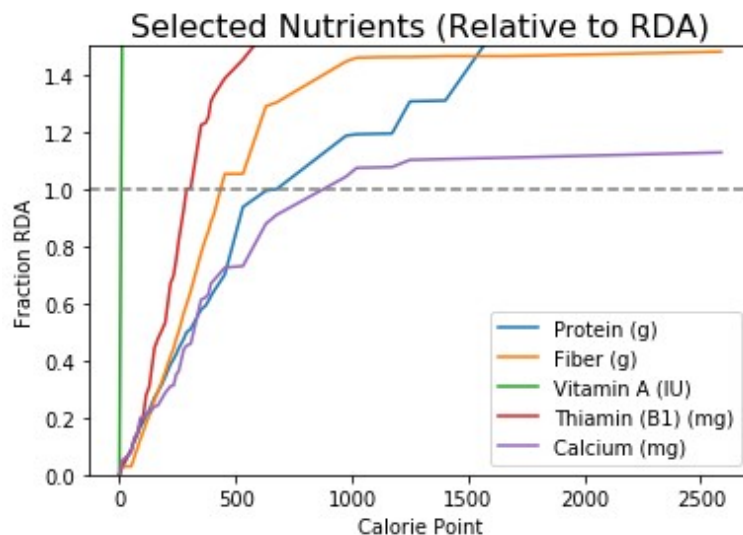
At the outset, the Gut eats exactly what we would expect given the nutritional analysis earlier: a mix of spinach, oysters, asparagus, and mushrooms, alternating

frequently between them as one or the other as the relative importances of the nutrients shift.

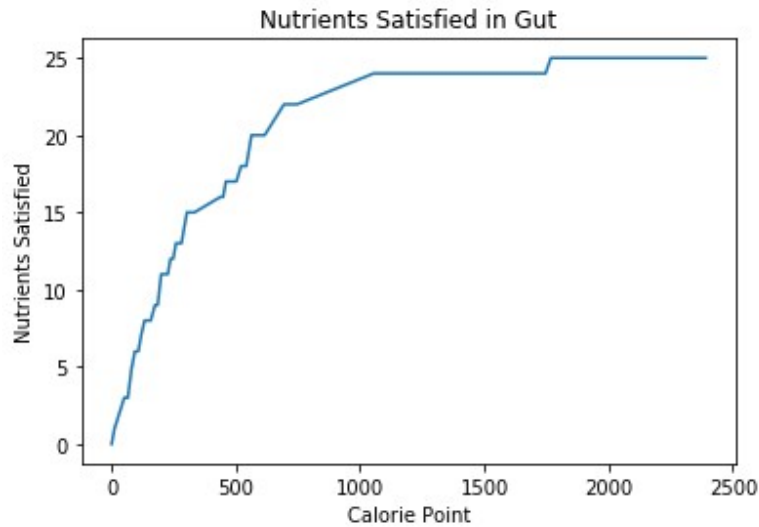
The Gut exhibits four big phases of eating foods.

1. 0 – 230 calories: A mix of spinach, asparagus, oysters, and mushrooms.
Mean relative nutrient density for spinach against 50 calories: 51% RDA.
2. 230 – 350 calories: Mostly okra.
Mean density for okra: 13%.
3. 530 – 630: Mostly turnips, with pickles and almonds.
Mean density for turnips: 5%.
4. 1020 – 2600: Mostly honey, with a piece of pork and an egg.
Mean density for honey: 1%.

As the Gut satisfies more nutrients, it chooses foods with complementary nutritional profiles and with less nutrition overall. At the end of the diet, the Gut takes in honey, which very low nutrient density, effectively filling out its calories to round out the diet.



Plotting selected nutrients shows the Gut's nutrients growing more slowly, and some successfully leveling off near their RDA limits.



Nutrient satisfaction grows more slowly than in the basic version. It still plateaus at 25 nutrients (out of a total of 26).

The “Ideal Diet” presented here is the best this project can do. However, it could be improved upon. The Gut’s greedy approach combined with an oversatisfaction factor means that the Gut satisfies nutrients aggressively in the beginning, and then intentionally chooses less nutritious foods later. Instead, if the Gut were to make more average choices consistently throughout the diet, it may satisfy nutrients just as well, with a more palatable, more variegated diet. The diet with the highest satisfaction and most food variety has yet to be seen. But the next section will investigate increasing variety with an easier approach.

Ideal Diet	
Median Nutrient Satisfaction	2.5
Unsatisfied Nutrients	1 / 26

MyPlate Eating Recommendations

As previously noted, this project's Ideal Diet does not show much variety, and does not look too appetizing. But nutrition scientists and the US government make much more relaxed dietary recommendations. The government's recommendations used to be summarized by the Food Pyramid, and are now available on [ChooseMyPlate.gov](https://www.choosemyplate.gov). By modifying the Gut class, this project can simulate the effect of randomized eating within those guidelines. How nutritious a diet does this recommendation produce?

This project's abstract subject is 5'11", 160 lb, and moderately active. The dietary recommendations have their own set of serving sizes, and need to be taken with some modification for this project. This project converts the servings given into the best equivalents in terms of calories. The original recommendations and the adaptation for this project are given below.¹

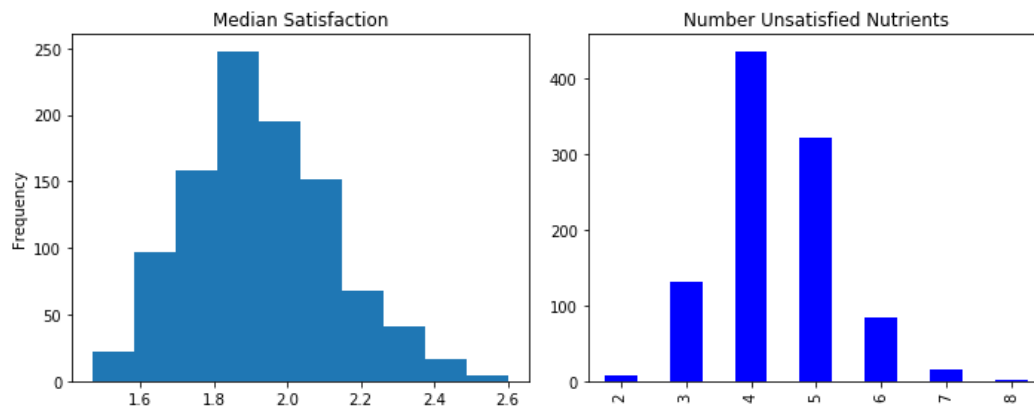
Food Group	MyPlate Servings	Conversion to Calories
Fruits	2.5 cups	300
Veggies	3.5 cups	150
Grains	10 ounce equivalents	1100
Protein	7 ounce equivalents	400
Dairy	3 cups (equivalents)	450
Treats	(None)	200

A new version of the Gut class, RandomGut, was written to simulate a person following these recommendations, choosing from the curated set of foods at random, in serving sizes of 50 calories. In the simulation, the Gut first eats from the fruit group, choosing fruits at random until it has had the recommended amount or more (here, 300 calories). It then moves through the subsequent food groups this way until reaching "Snacks," finishing out the remaining calories.

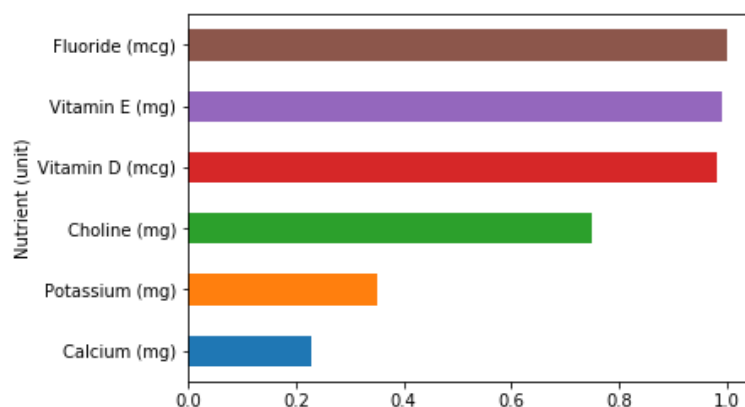
The RandomGut was simulated 1000 times, with surprisingly good results. The average of the median RDA satisfaction across simulated instances was 1.93, a good level of satisfaction seeming to be too high. The average number of unsatisfied nutrients was 4.4, higher than the ideal diet. But out of a total of 26

1 For convenience, this project keeps eggs grouped with dairy following SR28, while MyPlate groups them with proteins. The impact on results should be minor.

nutrients, this represents over 80% satisfaction of the RDA. The frequencies for median satisfaction and unsatisfied nutrients are shown below.



Six nutrients went unsatisfied any significant frequency: vitamin D, vitamin E, choline, calcium, potassium, and fluoride (of course). Some of these nutrients are anticipated by nutritional science. The DRI Workshop listed vitamin D, vitamin E, and potassium among the “unrealistic nutrients” which very few Americans meet the RDA. (DRI Workshop) The rates of various nutrients going unsatisfied are shown below.



It is impressive that eating according to recommendations, without specifically trying to eat the optimum diet, achieves such a high level of RDA satisfaction.

The result shows that either the RDA or the guidelines are imperfect, but generally supports the nutritional advice provided by MyPlate.

Random eating based on MyPlate	Averages
Median Nutrient Satisfaction	1.93
Unsatisfied Nutrients	4.4 / 26

References

...