MathWorks Math Modeling Challenge 2018

Los Altos High School–

Team # 10387 Los Altos, California

Coach: Carol Evans

Students: Ryan Huang, Michael Vronsky, Daniel Wang, Justin Yu, Joanne Yuan

MathWorks Math Modeling Challenge Champions

$20,000 Team Prize

MathWorks Math Modeling Challenge

Executive Summary

Throughout the world, there is an incredible amount of wasted food. Much of this food is perfectly edible yet wasted due to poor decisions: impulse buying, poor planning, unreasonable quality standards, overstocking, etc. [1]. This food could be used to feed the 13% of people in the US who are food-insecure [2]. Instead, it is tossed away, wasting the valuable land, water, and labor resources used to produce it.

To determine if a state can feed its food-insecure population with its wasted food, we first consider whether a state even wastes enough food to feed its food-insecure population. If there is not enough wasted food, we need not consider issues such as the collection, transportation, and distribution of this food. This turned out to be the case in our analysis of Texas. After first defining a production waste vector, $\mathbf{W}_p$, we calculated what percentage of different food types people wasted. Then, after a series of conversions, we converted these values to kilocalories (kcal) per dollar. Once the total kcal values were calculated for each type of food, we summed them up, and divided by the average caloric needs per person per day to get the total number of people the excess food waste could serve. Out of the 4.32 million food-insecure individuals in Texas, only 1.95 million could be served by food waste at 2,000 kcal/day. Alternatively, the entire food-insecure population could be sustained at just 905 kcal/day.

To establish a baseline for the types of households in the United States for the second part of the problem, we used government data to find the average number of calories needed per day for people by age, gender, and activity level. Then, to determine how income affects what food types households eat, we used a nonlinear model fit to predict the proportion of income spent on given food types based on annual income. This allowed us to calculate how many pounds of food are wasted for each of the example households. For the single parent with a toddler, family of four, elderly couple, and single 23-year-old, the total amount of food wasted per year was 256.4, 839.9, 366.8, and 217.2 pounds, respectively.

For the third part of the problem, we realized that the primary issue with food waste was its delivery to the needy. We decided to experiment with various solutions to the problem of food delivery. Our three strategies were: one central distribution center for our county, multiple distribution centers, and one central distribution center with mobile distribution centers arranged with freezer trucks. We used a computational simulation to estimate the efficacy of these models on the side of consumers. Using basic economic principles, we assumed that individuals would only make the trip from their houses to centers if the value of the food they receive is greater than the cost of travelling and the opportunity cost of not working during the time spent collecting food. Using this, we found the percentage of food-insecure individuals in our simulation who would actually go and get food from centers to be about 70% for the one-center model and 90% for the multi-hub model. Ultimately, we found a model with multiple distribution centers to be the most effective in the long-run (after 4.8 years).

Global Assumptions

G.1 Calories are an accurate measure of the nutritional value of food.
G.2 A year has 365 days (as opposed to 366).

Global Definitions

Food waste: "Food waste is part of food loss and refers to discarding or alternative (non-food) use of food that is safe and nutritious for human consumption along the entire food supply chain, from primary production to end household consumer level" [3]
Recoverability: Food waste is recoverable if it can be repurposed from waste to human-grade food.

Part I: Just Eat It!

1.1 Restatement of Problem

We are asked to create a model that lets a state calculate if the wasted food generated by its inhabitants is enough to feed its food-insecure population and to apply this model to Texas.

1.2 Local Assumptions

1. Consumption waste is irrecoverable.
    a. **Justification**: Once a consumer has bought food, it is either practically impossible to recover it (due to spoilage after thawing, transportation, etc.), or doing so will come at a significant risk (e.g. spread of pathogens, as evidenced by the recent passage of the Good Samaritan Food Donation Act) [4].
2. Production waste is only recoverable in postharvest handling, storage, processing and packaging, and distribution (e.g. supermarkets).
    a. **Justification**: The food waste produced in the agricultural production phase is "losses due to mechanical damage and/or spillage," "animal death during breeding," "decreased milk production due to mastitis," etc. [5]. These losses are inherent to the production food and cannot be recovered for human consumption.
3. The problem statement asks us to "determine if [a state] could feed its food-insecure population using the wasted food generated in that state." We interpret "using the wasted food" to mean directly feeding food-insecure people with the wasted food, and not indirect solutions such as using the food waste as biofuel and exporting energy in exchange for high-quality food.

1.3 Variables

Symbol	Definition	Units
m	The number of food types	
$\mathbf{W}_p$	The production waste vector, whose i^{th} element is the percentage	%

	of food wasted at the production level for the i^th^ food type	
$V_\$$	The caloric value per dollar vector, whose i^{th} element is the number of calories per dollar of the i^{th} food type.	kcal/USD
TPW_i	The total production waste for the i^{th} food type,	kcal
R_i	The state receipts for the i^{th} food type	USD

1.4 Solution & Results

Since this part of the problem statement asks specifically if a state "could" use its food waste, we first consider whether there is in fact enough food waste to feed the food-insecure population, *regardless* of its proximity to distribution centers, expiration date, etc. (all factors which would reduce the amount of food available for redistribution). If this preliminary result indicates that there is insufficient food waste to nominally feed the food-insecure population, a further analysis is not necessary.

First, the total waste percentage produced during the phases outlined in assumption 1.2.2 was calculated by summing the percentage waste for each type of food from the data provided. We call this table the production waste vector W_p, where each entry in the vector corresponds to the percentage of production waste for a specific food type.

Production waste vector W_p	
Food type	**Production waste**
Cereals	0.09
Roots and tubers	0.32
Oilseeds and pulses	0.06
Fruits and vegetables	0.18
Meat	0.10
Fish and seafood	0.16
Milk	0.02

Ultimately, our model will evaluate the quantity of food waste compared to the necessary food for food-insecure individuals. As per assumption G.1, we will be using calories to compare the net amount of food that can be reallocated. However, this presents a problem, as food production data is provided in dollars, and must be compared to calories. Thus, we use the conversion:

$$\frac{kcal}{USD} = \frac{USD}{Lb.} * \frac{Lb.}{kcal} = Price * \frac{1}{\frac{kcal}{Lb}}$$

Further research was conducted in order to research both the price of foods and their caloric value per lbs. The USDA provides many of these values in its databases [6]. The results of the calculation are as follows:

Caloric values per dollar (Vector $\mathbf{V_\$}$)	
Food type	$kcal/(US\ Dollar)$
Cereals	2770
Roots and tubers	3494
Oilseeds and pulses	16351
Fruits and vegetables	405
Meat	253.33
Fish and seafood	142.19
Milk	97.22

Ultimately, the total number of calories wasted for a specific food type is calculated by:

$$TPW_i = R_i * V_{\$,i} * W_{p,i}$$

where TPW_i is the total production waste for the i[th] food type, R_i is the state receipts for the i[th] food type, $(\frac{kcal}{US\ Dollar})_i$ is the ratio of kilocalories to dollars for the i[th] food type, and $\mathbf{W}_{p,i}$ is the i[th] element of the production waste vector.

Knowing TPW_i, we can calculate the total production waste as:

$$TPW_{tot} = \sum_{i=1}^{m} TPW_i$$

where m is the total number of food types.

While this was the general method used to calculate values, often caloric values varied significantly by type of food within the subdivisions created in the vector $\mathbf{W_p}$. For example, the number of calories per USD for wheat was calculated as 3,076 kcal/USD, whereas that for rice was just 842.86 kcal/USD. For these categories (cereals, roots and tubers, and roots and tubers), the same procedure was carried out, despite the fact that these more specific subdivisions (ex. wheat, rice, oats, peanuts) are not formally part of $\mathbf{W_p}$.

Performing this calculation for the state of Texas, the results for TPW_i, as well as the total are shown below.

Total Production Waste		
Food type	TPW_i (Billions kcal)	People fed per annum assuming a 2,000 kcal/day diet
Cereals	229	314,227
Roots and tubers	147	202,031
Oilseeds and pulses	644	882,411

Fruits and vegetables	116	158,903
Meat	285	390,923
Fish and seafood	0.391	535
Milk	3.95	5,415
SUM	**1426.745**	**1,951,445**

The final column was calculated by dividing the second column by (2000*365)=730,000. Comparing to the cited number of food-insecure individuals in the state of Texas, it is clear that even if all food waste that was fit for human consumption was to be repurposed for consumption by food-insecure individuals, it would ***not*** be sufficient to guarantee nutrition for those individuals [7]. Optimally, the food waste in question could provide for all individuals, but only at a level of 905 kcal/day (calculated by taking the sum of TPW_{tot} and dividing by the total number of people multiplied by 365).

1.5 Validation

To validate our model, we converted the values of TPW_i back into lbs using the vector **D** from part II §2.4. Using this calculation, we obtained the total production waste in Texas, in units of lbs. This value was approximately 1.81 billion lbs. Dividing this by the population of Texas, which is 27.86 million people [8], we get the per capita production waste (by our definition) to be 65 lbs.

According to the FAO data [5], 401 lbs of food are wasted per capita annually in North America and Oceania *in production*. The FAO cites this figure to include losses in agricultural production, which we chose to exclude. To reconcile this difference, we calculated the ratio of losses which we defined as production losses to the ratio that the FAO defined as production losses. Dividing the per capita production loss by this ratio yields 116 lbs of annual per capita food waste. While this result is still substantially lower the the FAO data suggests, it does not account for the fact that foods produced in higher quantities may have higher or lower percentages of agricultural waste. More investigation is necessary to determine the exact causes of this error.

1.6 Strengths & Weaknesses

Strengths

- Simplicity
    - Our model assumes maximum efficiency, and only looks at whether there is *enough* food waste to feed the food-insecure population, regardless of whether it can feasibly reach them or not. By taking a simplistic and general approach, we are able to show that under no circumstances would the food waste be enough to feed the entire food-insecure population

- Realism of Recovery

- Our model only considers production waste. This is important as this is the only waste that can feasibly be reallocated to those in need. This plays a key role in reducing the theoretical maximum number of people fed by food waste per annum.

Weaknesses

- Price estimates
    - The price estimates we used were generalizations based on indicative data researched thoroughly. Despite this, prices may vary by region, store, distance to store, etc. In general, these are factors which would increase food prices, and further lower our estimate for the number of people fed by food waste.
- Caloric intake
    - As opposed to our models in parts II and III, we did not factor in different caloric intakes. For the sake of simplicity, generality, and "safety" of the model, we used a conservative estimate of 2,000 kcal/day.

Part II: Food Foolish?

2.1 Restatement of Problem

We are asked to create a model that estimates the the amount of food different households waste in a year and then apply this model to four households types.

2.2 Local Assumptions

1. Activity levels can be divided into sedentary (involving only physical activity required for daily life), moderately active (comparable to walking at 3 to 4 miles per hour for 1.5 to 3 miles), and active (equivalent to walking at 3 to 4 miles per hour for more than 3 miles).
2. If the activity level of an individual is not specified, we assume that they are "moderately active."
    a. **Justification:** The CDC estimates that 52.5% of adults (aged 18 and older) meet federal guidelines for leisure-time aerobic activity. These guidelines match our definition of "moderately active." [9]
3. People in the same income bracket spend equal proportions of money in the food types.
4. The food consumption of each individual is independent of the company they keep.
    a. **Justification:** For example, the food consumption of a toddler living with his mother does not differ from that of living with his grandmother.
5. The caloric needs for a person are calculated by averaging the values for male and females of that person's age and activity levels.
    a. **Justification:** This is due to a lack of specificity on the part of the problem statement, since the example households do not provide genders.
6. We assume that the average age of the single parent of the toddler is 26 (average first-time mother/father age) + 2 (average age of a toddler) = <u>28 years old</u>. Further, we assume

the average age of the teenagers' two parents is 42 years old. It is calculated by: 26 (average first-time mother/father age) + 16 (average age of teenager) = 42 years old. [10]

7. "The amount of food waste a household generates in a year" refers to *only* the consumer waste, not the production waste necessary to bring the food to the household.
8. The percentage of food wasted (by food type) is constant and independent of household income.
9. Food's cost is regressive, and its demand is that of a normal good.
    a. **Justification:** As income increases, food consumption will increase, but eventually level off. Thus, the proportion of income spent on food eventually decreases with income in an inverse relationship.
10. Differences between the `ConsumerBehaviorBasedonIncome` and `Texas_food_data` datasets (such as a lack of data on "roots and tubers" in the former) are negligible.
11. Texas' expenditures on different food types are representative of the US's food spending.

2.3 Symbols Used

Symbols referenced in Part I §1.3 may be used.

Symbol	**Definition**	**Units**	**Notes**
K(age, gender, activity)	The daily caloric intake function.	kcal	$\mathbb{R}^1$
I	Household annual income.	USD	
Dist(I)	A vector-valued function from income to the proportion of calories by food type consumed per household.	%	$\mathbb{R}^+ \rightarrow \mathbb{R}^{m+1}$
C	The vector whose ith term is the number of calories of the ith food type consumed by the household, per annum.	kcal	
D	The vector whose ith element is the inverse energy density of the ith food type.	lbs/kcal	
$\mathbf{C}_{lbs}$	The vector whose ith term is the number of pounds of the ith food type consumed by the household, per annum.	lbs	
$\mathbf{W}_C$	The consumption waste vector, whose ith element is the percentage of food wasted at the consumption level for the ith food type.	%	
$TCW_{per\ annum}$	The total consumption waste per annum for the household.	lbs	

2.4 Solution

In our model, we consider three main factors to significantly affect required calories: age, gender, and activity level. Using data from the Center for Nutrition Policy and Promotion in the United States Department of Agriculture, we can construct this table to find the calories of consumption per day [11]:

	Activity Level					
	Male			Female		
Age (years)	Sedentary	Moderately Active	Active	Sedentary	Moderately Active	Active
2	1,000	1,000	1,000	1,000	1,000	1,000
3	1,200	1,400	1,400	1,000	1,200	1,400
4	1,200	1,400	1,600	1,200	1,400	1,400
5	1,200	1,400	1,600	1,200	1,400	1,600
6	1,400	1,600	1,800	1,200	1,400	1,600
7	1,400	1,600	1,800	1,200	1,600	1,800
8	1,400	1,600	2,000	1,400	1,600	1,800
9	1,600	1,800	2,000	1,400	1,600	1,800
10	1,600	1,800	2,200	1,400	1,800	2,000
11	1,800	2,000	2,200	1,600	1,800	2,000
12	1,800	2,200	2,400	1,600	2,000	2,200
13	2,000	2,200	2,600	1,600	2,000	2,200
14	2,000	2,400	2,800	1,800	2,000	2,400
15	2,200	2,600	3,000	1,800	2,000	2,400
16	2,400	2,800	3,200	1,800	2,000	2,400
17	2,400	2,800	3,200	1,800	2,000	2,400
18	2,400	2,800	3,200	1,800	2,000	2,400
19–20	2,600	2,800	3,000	2,000	2,200	2,400
21–25	2,400	2,800	3,000	2,000	2,200	2,400
26–30	2,400	2,600	3,000	1,800	2,000	2,400
31–35	2,400	2,600	3,000	1,800	2,000	2,200
36–40	2,400	2,600	2,800	1,800	2,000	2,200
41–45	2,200	2,600	2,800	1,800	2,000	2,200
46–50	2,200	2,400	2,800	1,800	2,000	2,200
51–55	2,200	2,400	2,800	1,600	1,800	2,200
56–60	2,200	2,400	2,600	1,600	1,800	2,200
61–65	2,000	2,400	2,600	1,600	1,800	2,000
66–70	2,000	2,200	2,600	1,600	1,800	2,000
71–75	2,000	2,200	2,600	1,600	1,800	2,000
76+	2,000	2,200	2,400	1,600	1,800	2,000

This table defines the daily calorie consumption function we call *K(age, gender, activity)*.

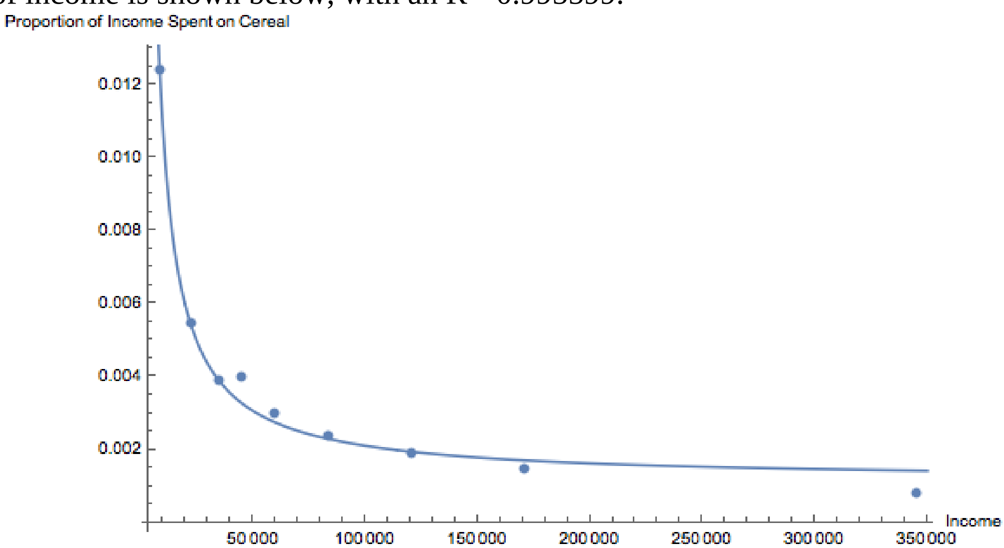
Using n as the number of people in a household we consider, we can calculate the total calories C_{total} consumed in a household each year as

$$C_{total} = 365 * \sum_{i=1}^{n} K(age, gender, activity)$$

Next, we determine the distribution of food types consumed by a family of a certain income. The rationale behind this is that different income brackets may consume different foods in different proportions (for example, lower income households may consume foods with a higher caloric value per dollar, such as fatty foods and simple carbohydrates derived from cereals). Since the data was reported in dollars, we used the dollars to calories conversion outlined in Part I §1.4 ($\boldsymbol{V_\$}$) to derive the caloric distribution for each household. This operation is at its core is a mapping from $\mathbb{R}^+$ (positive real numbers) to a distribution vector in $\mathbb{R}^{m+1}$ (Euclidean space of dimension m+1), where m is the number of food types. The one is added to *m* to account for eating out of the home. The output of this function is a vector whose i^{th} entry is the percentage of calories consumed by the household from the i^{th} food type per annum. We call this mapping the function ***Dist(I)***, where *I* is the income of a household.

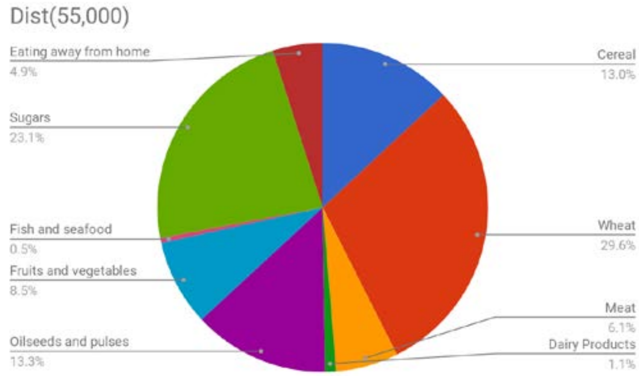
To calculate ***Dist***(I), we first use the `ConsumerBehaviorBasedonIncome` dataset to find the proportion of a household's annual income that is used to buy foods in the eight different categories: cereal, wheat, meat, dairy products, oilseeds and pulses, fruits and vegetables, fish and seafood, sugars, and eating away from home. We calculate this proportion for each range of incomes in the dataset. This gives us a fixed set of values for each food type, which we turn into a set of continuous functions by performing a nonlinear curve fit *a+b/x* (justified by assumption 2.2.9).

A sample curve fit showing the proportion of income spent on a food type (cereal) as a function of income is shown below, with an R^2=0.995599.



After multiplying by our annual household income *I*, this gives us the expenditures of a household on each food type given their income. We then compute the component-wise multiplication of these expenditures with an updated $\mathbf{V}_\$$, which gives us the number of calories consumed per food group by a household with income *I*. Finally, we divide by the total number of calories consumed to give us the proportion of calories consumed per food type. This is ***Dist***(I).

A sample distribution is included below for an income of $55,000:



We then compute:

$$\mathbf{C} = C_{total}\mathbf{Dist}(I)$$

However, there is an issue with multiplying this vector **C** by $\boldsymbol{W_c}$ to get the food waste. Doing so would assume that food waste percentages are calorie-based, while, in fact, they are weight-based. To account for this, we compute the Hadamard product:

$$\mathbf{C}_{lbs} = \mathbf{C} \circ \mathbf{D}$$

where **D** is the inverse energy density vector, whose i^{th} component is the number of lbs per calorie of the i^{th} food type. The Hadamard product is defined as:

$$(\mathbf{A} \circ \mathbf{B})_{i,j} = (\mathbf{A})_{i,j}(\mathbf{B})_{i,j}$$

Thus,

$$\mathbf{C}_{lbs,i} = \mathbf{C}_i\mathbf{D}_i = (\text{Calories of food type i})\left(\frac{lbs}{kcal}\right)_i = \text{Lbs of ith food consumed per annum}$$

Now, taking the dot product of the vector $\mathbf{C}_{\text{lbs}}$ with the consumption waste vector $\mathbf{W}_{\text{C}}$ yields the total food waste, in lbs, for the family, per annum.

In summary,

$$C_{total} = 365 * \sum_{i=1}^{n} K(age, gender, activity)$$
$$\mathbf{C} = C_{total}\ \mathbf{Dist}(I)$$
$$\mathbf{C}_{lbs} = \mathbf{C} \circ \mathbf{D}$$
$$\mathbf{C}_{lbs} \cdot \mathbf{W}_C = TCW_{per\ annum}$$

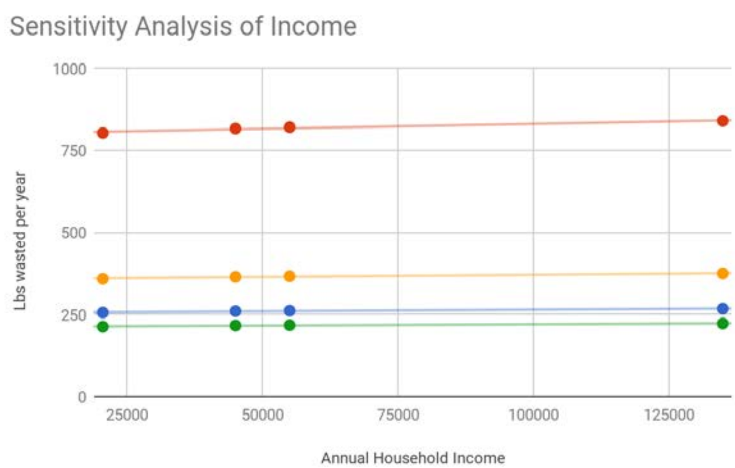
2.5 Results

Description of Household	**$\text{TWC}_{\text{per annum}}$ (lbs)**
Single parent with a toddler, annual income of $20,500	256.4
Family of four (two parents, two teenage children), annual income of $135,000	839.9
Elderly couple, living on retirement, annual income of $55,000	366.8
Single 23-year-old, annual income of $45,000	217.2

2.6 Validation

A basic check of our model reveals that it is reasonable. According to the FAO report [5], the food waste per capita in the United States is found to be 211.6 lbs per year. This is in line with our result for the single 23-year-old (217.2 lbs per year), who has an annual income close to $41,655, the national average for his or her age group.

Next, we performed a sensitivity analysis on annual household income. Keeping our methodology the same, we calculate the pounds of food wasted per year for each of the four family types and incomes ($20,500, $45,000, $55,000, and $135,000). The plot below shows the surprising result that, contrary to what we expected, household income has little effect on the amount food wasted. It does, however, support our assumption that the amount of food wasted varies between different households. Thus, our model is successful in correlating food waste to household type.



2.7 Strengths & Weaknesses

Strengths

- The strength of our model is that it is simple in its final form. While the intermediate calculations are not basic, the final model closely resembles a function with the number and types of people in the household as its input and the yearly food waste as its output.
- The model acts as we would expect: seniors waste less than a 23-year-old, and babies waste the least of all. Families waste the most because they are the largest households. It also takes into account the important aspects of these households: income, age, and activity.

Weaknesses

- We do not take into account different activity levels. Instead, we assume that all people are moderately active. While this is a reasonable assumption, the caloric needs between the levels of activity can differ by a few hundred calories. This weakness is due to a lack of specificity on the part of the problem statement.
- We assumed that all people within an income bracket eat the same proportion of food types. However, this is not quite accurate. For example, babies likely consume less protein than bodybuilders in the same income bracket.

Part III: Hunger Game Plan?

3.1 Restatement of Problem

We are asked to use mathematical modeling to find strategies to recover the most wasted food with the least cost in our community. Since the primary issue with food redistribution is its transportation to the needy, we explore various strategies for distributing said food waste to those in need in our community. We define our community to be Santa Clara County, California, although this model could be extended to other counties on the assumption that they lack significant agricultural production facilities.

3.2 Local Assumptions

1. Freezing (and transportation of frozen food via refrigeration trucks) enables food to safely be stored for more than a month [12]. So, for the purposes of our problem, we don't consider that wasted food may expire (in transit or storage, for example) before reaching its food-insecure recipient.
2. The distribution of types of food at a distribution center will be similar to the distribution of food types in an average diet.
3. The population of Santa Clara County is uniformly distributed.
4. The price of gasoline per mile for a car is about $0.15, based on data from 2006, adjusted for inflation ($\kappa_{carTravel}$= $0.19/mile) [13].
5. The price of gasoline per mile for a truck is $0.31. The price of gasoline per hour for a truck is $18.25 [14] [15].
6. Food-insecure people earn about $13.50 an hour, or the Santa Clara minimum wage [16].
7. The truck we will use for the purposes of this part is the 2016 ISUZU NPR refrigerated truck. Its initial cost is $38,000, its maximum load is approximately 9,000 lbs, and its city mileage is 8 mpg with a tank of 30 gallons [17][18][19][20].
8. A hot meal has a monetary value of $1.16 [21].
9. A truck driver earns about $13.65 per hour [22].
10. The food distribution center will be small in scale; thus, it can only service one person at a time and loading takes 12 minutes per person.
11. To incentivize food manufactures to donate food waste, 15% of the cost of the food will be paid for by the food banks through the state. This comes from the markup on wholesale groceries, which is assumed to approximately cover the cost of transportation and distribution [23].
12. People are rational; they will weigh the costs and benefits of each action in terms of monetary amounts. If the benefit exceeds the cost, they will perform the action.
13. A distribution center costs approximately $1,500,000 [24].
14. The operation of distribution centers is free because of state subsidies. Everybody who is operating the distribution center is a volunteer who will not be paid, and the distribution center requires little maintenance.
15. The center is open 8 hours a day, 5 days a week when distribution centers are fixed.
16. The percentage of food-insecure people in the U.S. (12.3%) approximates the percentage of food-insecure people in Santa Clara County [25].

17. According to the U.S. Census Bureau, the population of Santa Clara County in 2015 was 1.918 million; this is a relatively accurate representation of the population today [26].
18. All food at center gets used up.

3.3 Variables

Symbol	Definition	Units	Notes
T	Time in weeks	weeks	Used for distributor-side model
t	Time in hours for a food-insecure individual to make a one way trip to the nearest station	hours	Used for consumer-side model
t_{center}	The time spent at the food distribution center by the consumer	hours	By assumption 3.2.10, equal to 0.2 hours
C_{twage}	Wage in a week for 8 hours per day, 5 days per week	$/week	$13.65 * 8 * 5 = $546
$C_{gasoline}$	Gasoline cost in a week	$/hour	$18.25
N_{fi}	Number of food-insecure people in Santa Clara County	people	.123 * 1,918,000 = 235,914
m	Number of round trips truck makes.		
$t_{driving}$	The average time it takes to complete a round trip to the distribution center for a truck calculated by Mathematica	hours	1.6 hrs
C_{kwage}	Wage of the truck drivers in a week for 20 hours per day, 7 days per week	$/hr	$13.65 * 20 * 7 = $1911
$C_{transport}$	Cost of transporting food, each week, from production center to distribution center	$/week	
$C_{distribution\ center}$	Cost to build a distribution center	$	1,500,000
C_{fixed}	Initial cost of setting up a distribution center	$	Varies depending on the problem
p	The percentage of food-insecure people who will come to a center to obtain food		

W	The total amount of food imported to the center		
$\kappa_{carTravel}$	The cost per mile of private car travel	$/mile	By assumption 3.2.4, this is equal to $0.19/mile
κ_{wages}	The minimum wage opportunity cost of travelling to the nearest center	$/hour	By assumption 3.2.6, this is equal to $13.50/hour
Average Net Gain	For each food-insecure person, the value of the food they will get from the distribution center, subtracting transportation and potential wage losses.	$	
NetFunction	A function showing the net benefit to the people (benefit to food-insecure people minus the operating and fixed costs)		

3.4 Solution

The community that we chose to focus on was Santa Clara County. Since the main issue with repurposing food waste is delivery to food-insecure individuals, we decided to test how various models for food distribution centers would affect the county.

The dollar amount that will be received at each distribution center is derived from assumption 3.2.8, which states that a hot meal has a monetary value of about $1.16. A week's worth of food would thus be:

$$F = 7 * 3 * \$1.16 = \$24.36$$

First, we take into account the cost of transporting food from production centers to distribution centers. We can take the amount of food distributed to people throughout the course of a week, W, to be the total amount of food imported to the center, by assumption 3.2.18. We can then calculate $C_{transport}$ by taking 15% of W, using assumption 3.2.11.

To calculate W, we multiply p[1], the percentage of food-insecure people who will come to a center to obtain food, by 235 to $24.36, which is the value of food each person will get.

$$W = p * 235{,}914\, people * \frac{\$24.36}{person} = \$5{,}746{,}865.04$$

$$C_{transport} = .15W = 862{,}029.76p$$

We will calculate $C_{transport}$ in 3.4.1, 3.4.2, and 3.4.3.

This is simulated as the variable `foodDistributionDollars.`

Following from assumption 3.2.12, people will come to the food distribution center to get food if the cost of their transportation and their potential lost wages to and from the food bank is

[1] The explanation for calculating p is shown later in the problem and the results are shown in Section 3.5.

less than the cost of the food they get (i.e. it is a net profit for them). Mathematically speaking, they will get to the distribution center if and only if the following condition is satisfied:

$$F > (\kappa_{carTravel} * d) + (\kappa_{wages} * (2t + t_{center}))$$

where $\kappa_{carTravel}$is the cost of car travel per mile ($0.19/mile under assumption 3.2.4), $\kappa_{wages}$is the cost in lost wages per hour ($13.50/hour under assumption 3.2.6), t is the time spent driving to and from the center, and t_{center} is the time spent at the distribution center (0.2 hours under assumption 3.2.10). This part of the model is referred to as the consumer-side model.

3.4.1 One Distribution Center

The first model we tested was placing a food bank in the geometric centroid of Santa Clara County[2]. The food bank essentially acts as a food distribution center, giving people who can provide evidence that they are food-insecure about one week's worth of food.

The following equations describe the distributor-side model. T is in weeks.

$$C_{transport} = .15W = \$862{,}029.76p = \$862{,}029.768 * .6930 = \$597{,}386.62$$

$$C_{fixed} = C_{distribution\ center} = \$1{,}500{,}000$$

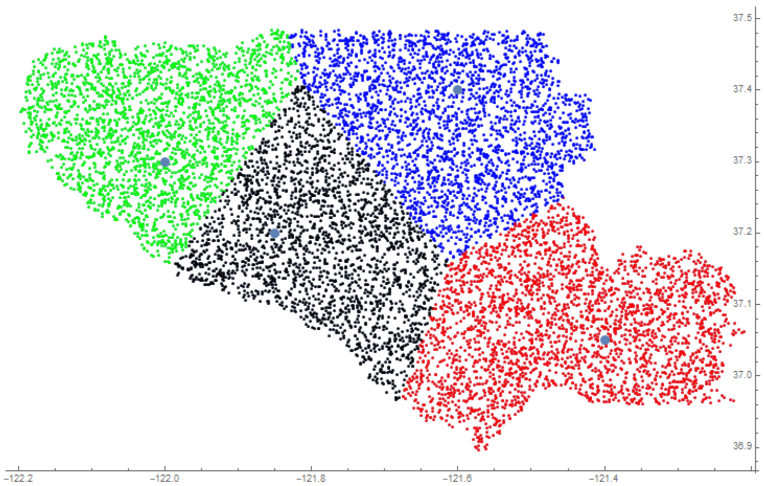
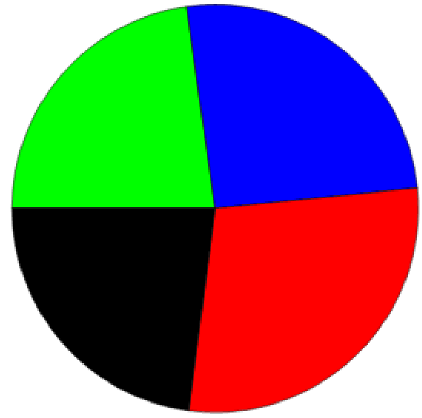
$$netGain = p(N\_fi)(Average\ Net\ Gain) = .6930 * 235914 * \$7.01647 = \$1147111$$

$$netFunction = \left(netGain - C_{transport}\right)T - C_{fixed} = -1{,}500{,}000 + 549{,}724T$$

3.4.2 Multiple Distribution Centers

Next, we ran a similar simulation using four distribution centers in Santa Clara County instead of one. In order to select the locations for these centers, we created a Voronoi partition of the county (using Euclidean distance instead of driving distance to minimize runtime). We then manually moved the distribution centers so that they produced four approximately equal areas. While this heuristic may not be perfect, it was effective given the time constraints, and could be improved with an optimization model (which would require significantly more runtime, see §3.7). The Voronoi partition was generated numerically with Wolfram Mathematica, and its graphical results are shown below, along with a pie chart showing the shares of the population served by each center.

[2] Under assumption 3.2.3 this is also the mean population center of the county, which is the real underlying reason for its selection

The four centers were located at geographic coordinates

{-122, 37.3}, {-121.6, 37.4}, {-121.4, 37.05}, {-121.85, 37.2}

The distributor-side model for this solution is four times as expensive due to there being four times as many distribution centers

$$C_{transport} = .15W = \$862{,}029.76p = \$862{,}029.76 * .896 = \$772{,}378.66$$

$$C_{fixed} = 4C_{distribution\ center} = 4 * \$1{,}500{,}000 = \$6{,}000{,}000$$

$$netGain = p(N_{fi})(Average\ Net\ Gain) = .896 * 235914 * \$11.78 = \$2{,}490{,}044$$

$$netFunction = (netGain - C_{transport})T - C_{fixed} = -6{,}000{,}000 + 1{,}717{,}665T$$

3.4.3 One Food Center with Mobile Distribution Centers

For our last simulation, we set a food bank again in the center of the county. This time, however, the food bank acted as a more conventional one, storing the food and not distributing it. Four trucks, sent out each day to fixed locations on the map, would distribute one week's worth of meals. The same distribution locations were used as in the previous section.

To calculate *m*, the number of trips necessary, we assume that all food-insecure people for whom it is economically viable to come will obtain $24.36 worth of food food from one of the four trucks. From our calculations in Part II, we then use a conversion factor of *447.8 kcal/$* and *.001311 lbs/kcal* to convert the dollar amount to *lbs* needed. Each truck can hold a maximum of *9,000 lbs*, so dividing our previous value by 9,000 will yield the number of trips required for each truck each week. We then round up.

$$.896 * 235914\,people * \frac{1}{4} * \frac{\$24.36}{person} * \frac{447.8\,kcal}{\$} * \frac{.001311\,lbs}{kcals} \times \frac{1}{9000\,lbs}$$
$$= 83.97\,trips m(t_{driving}) = 134.4\,hours$$

The time required to make a trip means that we will have to hire other workers to fill two roles: one role is driving the truck to the destination, unloading, driving back to the food center, and reloading to bring more food to the mobile center. The other will stay at the destination to hand out food. In total, each role requires that the workers at each mobile distribution center work a total of *19.2* hours, *7* days a week. This could easily be divided amongst multiple workers working in shifts without any changes to our model, as hiring additional workers in a competitive labor market implies no fixed costs and constant wages.

$$C_{transport} = .15W = \$862{,}029.76p = \$862{,}029.76 * .896 = \$772{,}378.66$$

$$C_{fixed} = 4C_{truck} + C_{distribution\ center} = 4 * \$38{,}000 + \$1{,}500{,}000 = \$1{,}652{,}000$$

$$C_{variable} = 4(C_{kwage} + mC_{gasoline} * t_{driving}) = 4[\$1911 + 84(\$18.25) * 1.6] = \$17455.2$$

$$netGain = p(N_{fi})(Average\ Net\ Gain) = .896 * 235{,}914 * \$11.78 = \$2{,}490{,}044$$

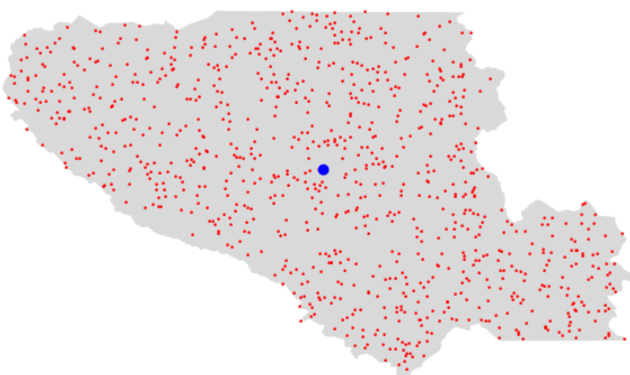
$$netFunction = (netGain - C_{variable} - C_{transport})T - C_{fixed} = -1652000 + 1700210T$$

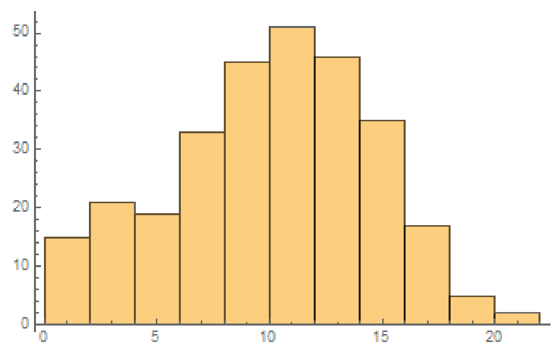
3.5 Results

The results of the consumer-side models were as follows:

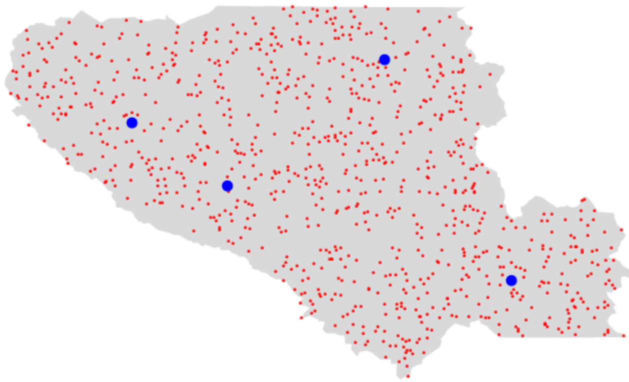
Model	**Percentage of People Served**	**Mean Net Gain per trip**
One Center	69.3%	$7.01
Multiple Centers	89.6%	$11.78
One Center with Trucks	89.6%	$11.78

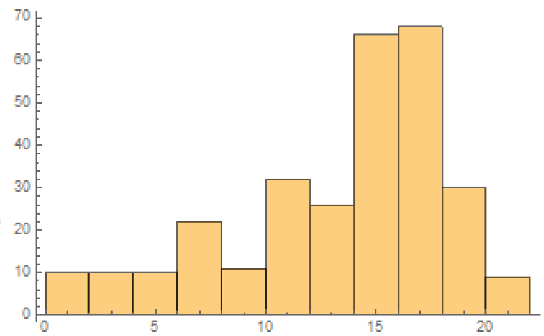
Note that the results for the second and third models are identical due to the fact that on the consumer side, they are identical (their only variation is on the distributor side).





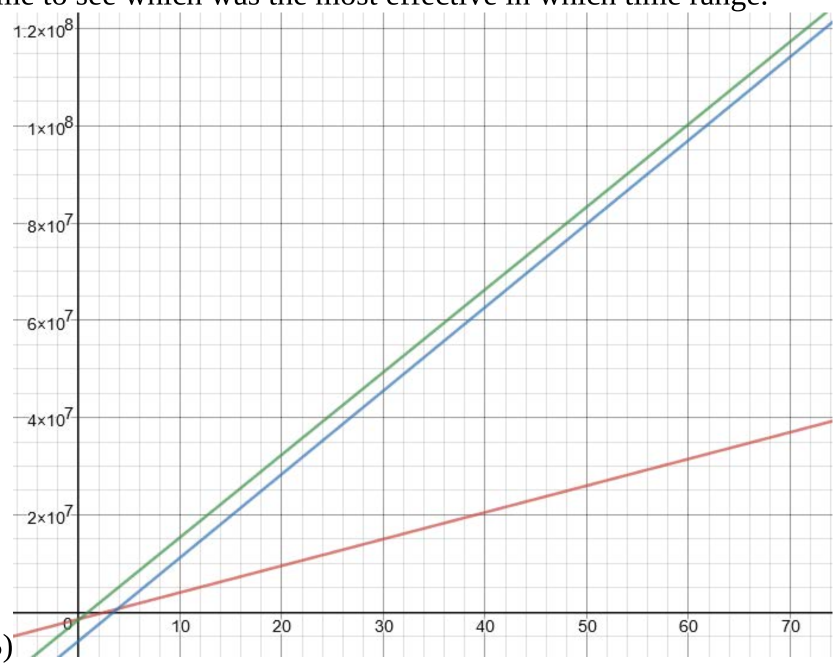
The results for the one-center model are shown above. The model, as expected, generated a lower percentage of people served, and was, on average, more expensive due to its higher mean distance from consumers. The histogram shows the distribution of net gains of individuals who *did* go to the center (it excludes those who chose not to visit the center).





The results for the multi-center models are shown above. As expected, the histogram shows that the net gains for a model with more centers were higher due to the decreased travel time (and distance) of consumers to centers.

For the distributor-side model, we plotted the three linear equations for the net benefit (*netFunction*) over time to see which was the most effective in which time range.



Net benefit ($)

Time (weeks)

Red: Model 3.4.1 — Blue: Model 3.4.2 — Green: Model 3.4.3

The results were fascinating, yielding that early on, the third model (with mobile distribution centers) was the most effective. The first model sat far below the others for virtually all time periods. The other two models were very close to each other, with the truck model being more effective in the short-run, but only up to 249 weeks, or approximately 4.8 years. After 1.8 years, the model with more fixed distribution centers is more effective.

3.6 Validation

Our model's results match with what we expected. The truck model has a lower upfront cost due to the reduced number of large centers, but its variable costs are higher due to the operating costs of the trucks. On the other hand, the multi-center model has higher upfront costs but over time becomes more cost-effective. The one-center model, although having a relatively low upfront cost, does not, in comparison to the other two models, become significantly more cost-effective over time.

3.7 Strengths & Weaknesses

Strengths

- Realistic driving time and distances
  - Wolfram Mathematica simulations of travel time and distance are significantly more accurate than estimates based on Euclidean or Minkowski distance estimates for driving.
- Extendability
  - The model could be extended to compute results for many different locations of distribution centers. This would allow for a more precise evaluation of where distribution centers should be placed
- Long-term strategy
  - Our model factors in variable costs in addition to fixed costs, leading to a more insightful long-term analysis of various solutions

Weaknesses

- Uniform population density
  - For a more accurate model, the population density could be modeled probabilistically using the relative frequency of population in census blocks.
- Uniform distribution of food amongst mobile centers
  - Our model does not account for the possibility of mobile centers being smaller than the main distribution centers. This could end up decreasing the costs of fuel for the strategy involving trucks
- Low sample size
  - Due to computational hardware limitations, we were only able to use samples of 500 data points to compute the efficacy of models. Given more time and computational resources, a more precise estimate could be computed. Alternatively, Euclidean distance could be used instead of Mathematica's realistic driving time (significantly reducing runtime), but at the cost of losing the primary strength of the model.

Bibliography

[1]https://www.nrdc.org/sites/default/files/wasted-food-IP.pdf
[2]http://map.feedingamerica.org/
[3]http://www.fao.org/platform-food-loss-waste/food-waste/definition/en/
[4]https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB1219
[5]http://www.fao.org/docrep/014/mb060e/mb060e00.pdf
[6]https://ndb.nal.usda.gov/ndb/
[7]http://map.feedingamerica.org/county/2015/overall/texas
[8]https://www.census.gov/quickfacts/TX
[9]https://www.cdc.gov/nchs/data/nhis/earlyrelease/earlyrelease201705.pdf
[10]https://www.npr.org/sections/health-shots/2016/01/14/462816458/average-age-of-first-time-moms-keeps-climbing-in-the-u-s
[11]https://www.cnpp.usda.gov/sites/default/files/usda_food_patterns/EstimatedCalorieNeedsPerDayTable.pdf
[12]https://www.foodsafety.gov/keep/charts/storagetimes.html
[13]https://archive.nytimes.com/www.nytimes.com/packages/html/business/20060510_LEONHARDT/cost_per_mile.html
[14]https://www.yelp.com/biz/chevron-santa-clara-2?osq=diesel+gas+station
[15]http://www.rtsfinancial.com/guides/trucking-calculations-formulas
[16]http://santaclaraca.gov/government/departments/city-manager/minimum-wage-ordinance
[17]https://www.commercialtrucktrader.com/Isuzu-Npr-Efi-For-Sale/search-results?make=ISUZU%7C2311932&model=NPR+EFI%7C764866628
[18]http://www.nadaguides.com/Cars/2017/Isuzu/NPR-GAS-CREW/150-WB-12000-GVWR-IBT-PWL/Pricing
[19]http://www.practicalmachinist.com/vb/material-handling-and-rigging/2008-isuzu-npr-w-gas-engine-any-idea-highway-mpg-297368/
[20]https://archive.nytimes.com/www.nytimes.com/packages/html/business/20060510_LEONHARDT/cost_per_mile.html
[21]http://www.thesoupkitchen.org/about_us.html
[22]https://www.bls.gov/ooh/transportation-and-material-moving/delivery-truck-drivers-and-driver-sales-workers.htm
[23]https://www.entrepreneur.com/answer/221767
[24]http://www.northvalleyfoodbank.org/architect-explains-food-bank-design/
[25]https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/key-statistics-graphics.aspx
[26]https://www.census.gov/quickfacts/fact/table/santaclaracountycalifornia/PST045216
https://www.livestrong.com/article/308576-how-many-calories-are-in-a-pound-of-sugar/
https://www.ers.usda.gov/data-products/sugar-and-sweeteners-yearbook-tables

Appendix

Note: horizontal lines indicate separate cells evaluated in the Mathematica notebook. All code is written for Wolfram Mathematica Version 11.2. An internet connection is necessary to download the Wolfram language databases used.

Model 2.4 Regression

```
data = Import["~/M3/part2.csv"];
data = Partition[data, 9]
fits = Table[
   NonlinearModelFit[data[[i]], a + b/x, {a, b}, x], {i, 1, 8}];
Table[fits[[i]][45000], {i, 1, 9}]
```

Model 3.4.1 Simulation

```
Clear[distances]
countyshape =
  AdministrativeDivisionData[
   Entity["AdministrativeDivision", {"SantaClaraCounty",
"California",
     "UnitedStates"}], "Polygon"];
countyshape = Polygon[Reverse /@ Flatten[countyshape[[1]][[1]],
1]];
center = {RegionCentroid[countyshape]};
test = Table[RandomPoint[countyshape], {i, 1, 500}];
centerUsed = Flatten[Map[Nearest[center, #] &, test], 1];
distances =
  Quiet[Table[
     TravelDistance[{GeoPosition[Reverse[centerUsed[[i]]]],
     GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}]];
traveltimes =
  Quiet[Table[
     TravelTime[{GeoPosition[Reverse[centerUsed[[i]]]],
     GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}]];
travelcostpermileCar = 0.19;
travelcost = 2*travelcostpermileCar*QuantityMagnitude[distances];
distances = QuantityMagnitude[distances];
errors = Position[distances, QuantityMagnitude[$Failed]];
test = Delete[test, errors];
distances = Delete[distances, errors];
traveltimes = Delete[traveltimes, errors];
Speak["Hallelujah"]
```

```
Graphics[{LightGray, countyshape, Red, Point[test], Blue,
  PointSize[0.017], Point[center]}]
```

```
carTravelCost = 0.19;
wagesLost = 13;
foodDistributedDollars = 24.36;
timeAtCenter = 0.2;
cost = Round[(carTravelCost*
     distances) + (wagesLost*(QuantityMagnitude[traveltimes,
     Quantity[1, "Hours"]] + timeAtCenter)), 0.01];
netGain = Map[#*UnitStep[#] &, (foodDistributedDollars - cost)];
"Percent of residents who access food bank"
N[Count[netGain, x_ /; x > 0]/Length[netGain]]
"Mean net Gain"
Mean[netGain]
Speak["Net gain calculation completed"]
```

```
netGainZeros = Position[netGain, 0.];
Histogram[Delete[netGain, netGainZeros], 10]
```

Model 3.4.2 / 3.4.3 Simulation

```
Clear[distances]
countyshape =
  AdministrativeDivisionData[
   Entity["AdministrativeDivision", {"SantaClaraCounty",
"California",
     "UnitedStates"}], "Polygon"];
countyshape = Polygon[Reverse /@ Flatten[countyshape[[1]][[1]],
1]];
center = {{-122, 37.3}, {-121.6, 37.4}, {-121.4, 37.05}, {-
121.85, 37.2}};
test = Table[RandomPoint[countyshape], {i, 1, 500}];
centerUsed = Flatten[Map[Nearest[center, #] &, test], 1];
distances =
  Quiet[Table[
     TravelDistance[{GeoPosition[Reverse[centerUsed[[i]]]],
     GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}]];
traveltimes =
  Quiet[Table[
     TravelTime[{GeoPosition[Reverse[centerUsed[[i]]]],
     GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}]];
travelcostpermileCar = 0.19;
travelcost = 2*travelcostpermileCar*QuantityMagnitude[distances];
distances = QuantityMagnitude[distances];
```

```
errors = Position[distances, QuantityMagnitude[$Failed]];
test = Delete[test, errors];
distances = Delete[distances, errors];
traveltimes = Delete[traveltimes, errors];
Speak["Hallelujah"]
```

```
Graphics[{LightGray, countyshape, Red, Point[test], Blue,
  PointSize[0.017], Point[center]}]
```

```
carTravelCost = 0.19;
wagesLost = 13;
foodDistributedDollars = 24.36;
timeAtCenter = 0.2;
cost = Round[(carTravelCost*
     distances) + (wagesLost*(QuantityMagnitude[traveltimes,
     Quantity[1, "Hours"]] + timeAtCenter)), 0.01];
netGain = Map[#*UnitStep[#] &, (foodDistributedDollars - cost)];
"Percent of residents who access food bank"
N[Count[netGain, x_ /; x > 0]/Length[netGain]]
"Mean net Gain"
Mean[netGain]
Speak["Net gain calculation completed"]
```

```
netGainZeros = Position[netGain, 0.];
Histogram[Delete[netGain, netGainZeros], 10]
```

Voronoi Partitions

```
countyshape =
  AdministrativeDivisionData[
   Entity["AdministrativeDivision", {"SantaClaraCounty",
"California",
     "UnitedStates"}], "Polygon"];
countyshape = Polygon[Reverse /@ Flatten[countyshape[[1]][[1]],
1]];
colors = {Green, Blue, Red, Black, Purple};
p = 2;
xdist = NormalDistribution[5, 6];
ydist = NormalDistribution[3, 5];
points = Table[RandomPoint[countyshape], {i, 1, 10000}];
hubs = {{-122, 37.3}, {-121.6, 37.4}, {-121.4, 37.05}, {-121.85,
     37.2}};
```

```
minkowski[a_, b_] := (Abs[(b - a)[[1]]^p] + Abs[(b -
a)[[2]]^p])^(1/p);
cindex = Flatten[
   Map[Position[#, Min[#]] &,
     Transpose[
     Table[Map[minkowski[#, hubs[[i]]] &, points], {i, 1,
     Length[hubs]}]]]], 2];
final = Partition[Riffle[points, cindex], 2];
finalg = Partition[Riffle[points, Map[colors[[#]] &, cindex]],
2];
pw = Table[
   Style[finalg[[i]][[1]], finalg[[i]][[2]]], {i, 1,
Length[finalg]}];
Speak["Data calculation completed"]
```

```
pointsize = 0.03;
Show[ListPlot[pw], ListPlot[hubs]]
Speak["Graph generation completed"]
```

```
PieChart[Map[Length, distances], ChartStyle -> colors]
```