

A Mathematical Analysis of Food Waste Production and Redistribution

Ryan Huang¹ | Michael Vronsky² | Daniel Wang³ | Justin Yu⁴ | Joanne Yuan⁵
Los Altos High School, Los Altos, CA

Abstract. Food insecurity affects millions of individuals in the United States. We develop a model to address food insecurity by repurposing food waste and apply this methodology to Texas. We also develop models to analyze food waste habits of consumers and optimal distribution strategies for food banks. We first consider whether a state wastes enough food to feed its food insecure population, regardless of distribution methods. We determine that there is not enough to feed the food insecure population of Texas assuming each person needs 2,000 kcal per day. To establish a food consumption baseline for different demographics in the United States, we use government data to find the average number of calories needed per day by age, gender, and activity level. Then, to determine how income affects what food types households eat, we use a nonlinear model fit to predict the proportion of income spent on given food types based on annual income. This allows us to calculate how many pounds of food are wasted for any given household. Finally, we analyze three potential food distribution strategies, including fixed and mobile distribution centers. A key feature of our model is its extensibility and the use of computer simulation to model consumers as rational agents. Of the models tested, we find a model with multiple fixed distribution centers to be the most effective in the long-run (after 4.8 years). The paper was originally a submission to the MathWorks Math Modeling Challenge in 2018.

¹ email: ryanxh@uw.edu

² email: michaelvronsky@berkeley.edu

³ email: dwang89@jhu.edu

⁴ email: justinyu@berkeley.edu

⁵ email: joanneyu@mit.edu

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 1.9 million could be served by food waste at 2,000 kcal/day. Alternatively, the entire food-insecure population could be sustained at just 879 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

1. **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]
2. **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	%
\mathbf{V}_s	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, per year	kcal
R_i	The state receipts for the i^{th} food type, per year. This was taken from the TexasFoodData dataset ⁶ . This is the monetary value of the i^{th} food type which is produced in Texas annually.	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 sequentially multiplying the percentage of food left after each step of the process. Essentially,

$$W_{p,i} = 1 - \prod_{s=1}^3 (1 - W_{s,i})$$

Where s stands for each step of the food supply chain, and $W_{s,i}$ is the food wasted at the s^{th} production step for the i^{th} food. $W_{p,i}$ is the total recoverable production waste for the i^{th} food type.

Using the per-production-phase waste percentages provided in the TexasFoodData dataset, we calculated the total waste percentage using the production steps from postharvest to distribution. We put these individual entries into the vector \mathbf{W}_p , where each entry in the vector corresponds to the percentage of production waste for a specific food type.

⁶ All datasets that were provided by M3 can be obtained online at <https://m3challenge.siam.org/node/385>.

Production waste vector W_p

Food type	Production waste
Cereals	0.09
Roots and tubers	0.29
Oilseeds and pulses	0.06
Fruits and vegetables	0.17
Meat	0.10
Fish and seafood	0.15
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 conversion⁷:

$$\frac{\text{kcal}}{\text{USD}} = \frac{\text{lbs}}{\text{USD}} * \frac{\text{kcal}}{\text{lbs}} = \frac{1}{\text{Price}} * \frac{\text{kcal}}{\text{USD}}$$

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]. For example, the price (USD/Lb.) of catfish in Texas is \$7.30, and the kcal/Lb. is 1,038. Thus, for catfish:

$$\frac{\text{kcal}}{\text{USD}} = \frac{1 \text{ lbs}}{7.30 \text{ USD}} * \frac{1038 \text{ kcal}}{1 \text{ lbs}} = 142.19 \frac{\text{kcal}}{\text{USD}}$$

The results of the calculation are as follows:

Caloric values per dollar (V_ζ)

Food type	$\frac{\text{kcal}}{\text{US Dollar}}$
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

⁷ Price is defined as USD per pound.

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 (from the TexasFoodData dataset), $(\frac{kcal}{USDollar})_i$ is the ratio of kilocalories to dollars for the i^{th} food type, and $W_{p,i}$ 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 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 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 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 year (2,000 kcals/day)
Cereals	223.32	305,917
Roots and tubers	132.99	182,175
Oilseeds and pulses	638.79	875,058
Fruits and vegetables	110.91	151,925
Meat	277.15	379,664
Fish and seafood	0.375	514
Milk	3.927	5,380
SUM	1386.13	1,900,633

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 879 kcal/day (calculated by taking the sum of TPW_{tot} , dividing by the total number of people, and multiplying by 365 days).

1.5 Validation

To validate our model, we converted the values of TPW_i back into pounds using the inverse energy density vector \mathbf{D} from part II §2.4 (which has units of lbs/kcal). Using this calculation, we obtained the total production waste per year in Texas, in units of lbs. This value was approximately 1.72 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 62 lbs. Note that this population does include the food-insecure population, as no distinction is made by the FAO when calculating per-capita waste production.

According to the FAO data [5], 401 pounds 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 110 lbs of annual per capita food waste. While this result is still substantially lower than 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

1. **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 the food waste would never be enough to feed the entire food insecure population.
2. **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

1. **Price estimates:** Our price estimates were based on data from the USDA [6]. Despite this, prices may vary by region, store, distance to store, etc. In general, these factors would increase food prices, and further lower our estimate for the number of people fed by food waste.
2. **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 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 a toddler 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) = 28 years old. 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 teenagers) = 42 years old. [10] We assume that the elderly couple is between 61 and 65 years old, since that range is the average retirement age. [29]
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

Symbol	Definition	Units	Notes
$K(\text{age, gender, activity})$	The daily caloric intake function.	kcal	\mathbb{R}^1
I	Household annual income	USD	
$\text{Dist}(I)$	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 i^{th} term is the number of calories of the i^{th} food type consumed by the household, per annum.	kcal	
D	The vector whose i^{th} element is the inverse energy density of the i^{th} food type	lbs/kcal	
C_{lbs}	The vector whose i^{th} term is the number of pounds of the i^{th} food type consumed by the household, per annum.	lbs	
W_C	The consumption waste vector, whose i^{th} element is the percentage of food wasted at the consumption level for the i^{th} food type	%	
$\text{TCW}_{\text{per annum}}$	The total consumption waste per annum for a household	lbs	

Symbols referenced in Part I §1.3 may be used.

2.4 Solution

In our model, we consider three main factors that significantly affect required calories: age, gender, and activity level. The table below is taken from the Center for Nutrition Policy and Promotion in the USDA [11]:

Age (years)	Activity Level					
	Male			Female		
	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

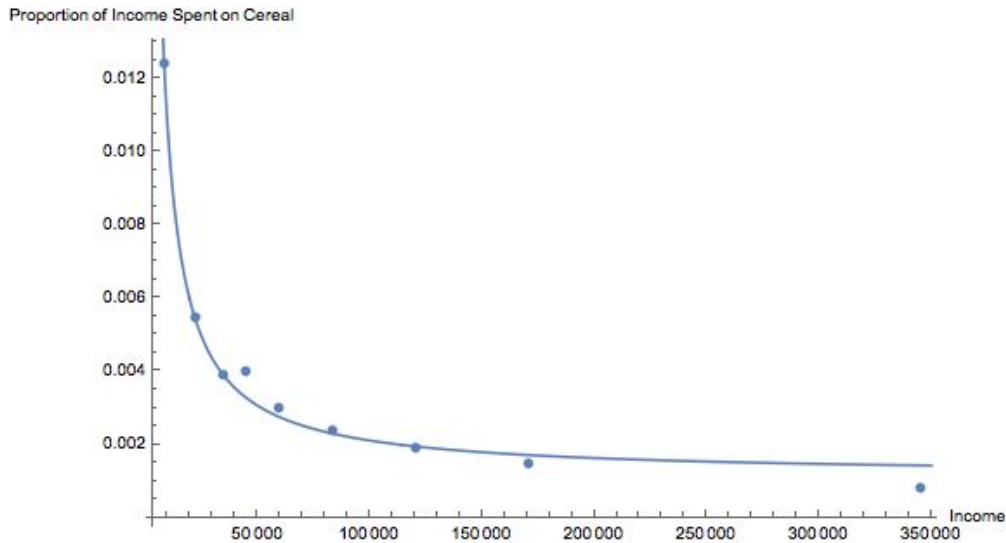
This table defines the daily calorie consumption function we call $K(\text{age}, \text{gender}, \text{activity})$. Using n as the number of people in a household we consider, we can calculate the total calories C_{total} consumed in a household each year as

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

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 (\mathbf{V}_s) to derive the caloric distribution for each household. This operation 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 $\mathbf{Dist}(I)$, where I is the income of a household.

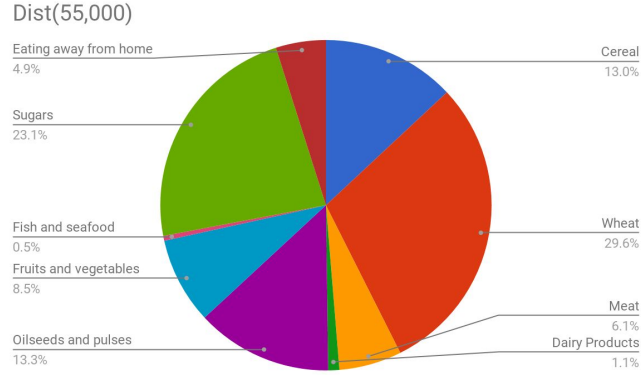
To calculate $\mathbf{Dist}(I)$, we first use the `ConsumerBehaviorBasedonIncome` dataset to find the proportion of a household's annual income that is used to buy foods in the eight different categories: cereal, wheat, meat, dairy products, oilseeds and pulses, fruits and vegetables, fish and seafood, 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 annual 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}_s , which gives us the number of calories consumed annually 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 $\mathbf{Dist}(I)$.

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 \mathbf{C} by \mathbf{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 \mathbf{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,

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

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

In summary,

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

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

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

$$\mathbf{C}_{lbs} \cdot \mathbf{W}_c = TCW_{per \text{ annum}}.$$

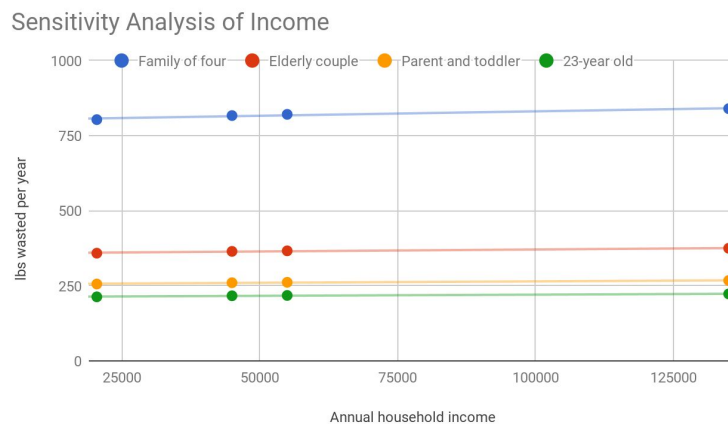
2.5 Results

Description of Household	TWC _{per annum} (lbs)
Single parent with a toddler, annual income of \$20,500	256.4
Family of four (two parents, two teenagers), 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

1. **Simplicity:** 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.
2. **Intuitive:** 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

1. **Assumptions:** 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.
2. **No specificity within income brackets:** 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. Santa Clara County has a constant population density.
4. The price of gasoline per mile for a car is about \$0.19, based on data from 2006, adjusted for inflation ($\kappa_{carTravel} = \$0.19/\text{mile}$) [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 state. This comes from the markup on wholesale groceries, which is assumed to approximately cover the cost of transportation and distribution [23]. Thus, we assume that the cost of transportation is covered by 15% of the value of the food transported.
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]. Distribution centers are paid for upfront, with no mortgage.
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.
19. A trip in the context of mobile distribution centers means trucks travelling from their fixed locations to the central distribution centers or from central distribution centers to their fixed locations.
20. The distribution process takes negligible time.
21. For the mobile distribution system, four roles need to be filled, one for each truck.

3.3 Variables

Symbol	Definition	Units	Notes
T	Time in weeks	weeks	For distributor model
t	Time for a food insecure individual to make a one way trip to the nearest station	hours	For consumer model
t_{center}	The time spent at the food distribution center by the consumer	hours	By assumption 3.1.10, 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 hour	\$/hour	$\$18.25/\text{hour}$
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 for the truck to complete a round trip (i.e. from the point where the truck is stationed to the distribution center and back) as calculated by Mathematica	hours	1.6 hrs
C_{kwage}	Weekly wage of drivers working for 19.2 hours a day with a .8 hour daily break	\$/week	$\$13.65 * 20 * 7 = \1911
$C_{\text{transport}}$	Cost of transporting food 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 per week		
$\kappa_{carTravel}$	The cost per mile of private car travel	\$/mile	By assumption 3.1.4, this is \$0.19/mile
κ_{wages}	The minimum wage opportunity cost of travelling to the nearest center	\$/hour	By assumption 3.1.6, this is \$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 per week.	\$	
netFunction	The net benefit to food insecure people, minus operating and fixed costs, as a function of T		

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 by each individual 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:

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

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 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), κ_{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.

On the distributor-side of the model, costs are more specific to each strategy. 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 , the percentage of food insecure people who will come to a center per week to obtain food, by 235,914 to \$24.36, which is the value of food each person will get. The explanation for calculating p is shown later in the problem and the results are shown in Section 3.5.

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

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

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

3.4.0 Simulation Process

In order to test various distribution strategies, we used the principles outlined earlier in a computer simulation done in Mathematica. The main steps in this process were:

1. Generate 500 randomly distributed points within the region of Santa Clara County (using the uniform population distribution from assumption 3.2.3). Each of these represents a food-insecure individual.
2. Find the driving distance and time from each point to the food distribution center.
 - a. If there are multiple centers, pick the closest center by Euclidean distance.
3. Evaluate, using the above method, whether the individual will choose to visit the center (i.e. if their benefits exceed their costs).
4. If the individual would choose to visit the center, record the net benefit to the individual consumer, by taking the monetary value of the food distributed at the center and subtracting the costs, which are given by $(\kappa_{carTravel} * d) + (\kappa_{wages} * (2t + t_{center}))$, otherwise, record the net benefit as zero.

3.4.1 One Distribution Center

The first model we tested was placing a food bank in the geometric centroid of Santa Clara County⁸. 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. Recall that p defined as the percentage of food insecure people predicted to come to a center to obtain food by our simulation (explained further in section 3.4.0)

The following equations describe the distributor-side model. T and $netGain$ are in weeks.

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

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

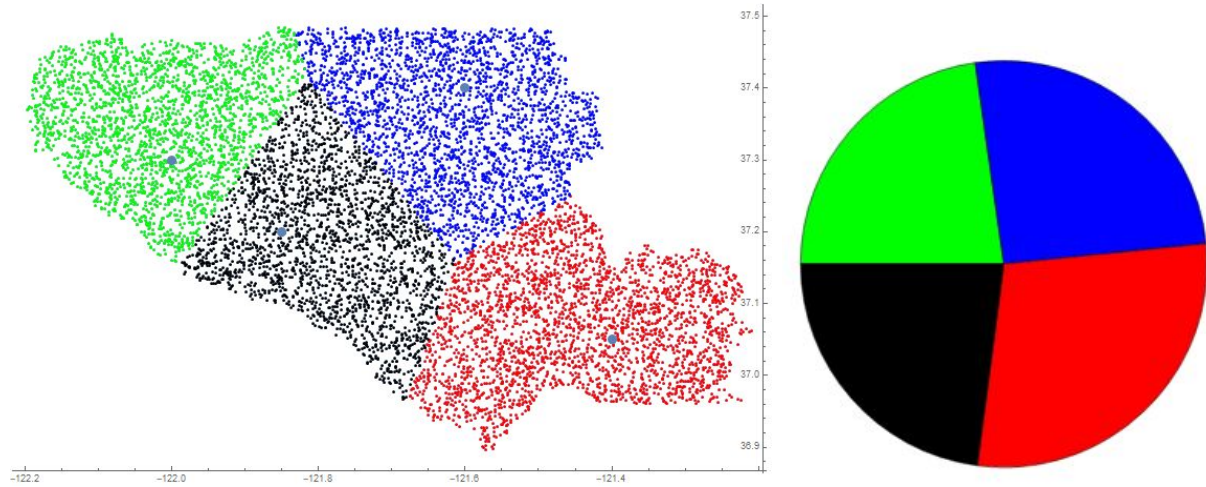
$$netGain = p(N_{fi})(Average\ Net\ Gain) = .6930 * 235,914 * \$7.01647 = \$1147111$$

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

⁸ Under assumption 3.2.3 this is also the mean population center of the county, which is the real underlying reason for its selection

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. By assumption 3.2.3, the population of Santa Clara County would be split approximately equally. 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.



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

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

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

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

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

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 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.

$$.896 * 235914 \text{ people} * \frac{1}{4} * \frac{\$24.36}{\text{person}} * \frac{447.8 \text{ kcal}}{\$} * \frac{001311 \text{ lbs}}{\text{kcal}} \times \frac{1}{9000 \text{ lbs}} = 83.97 \text{ trips} .$$

$$m(t_{\text{driving}}) = 134.4 \text{ hours}$$

In total, four roles need to be filled, one for each truck. Each role requires that the workers at each mobile distribution center work a total of 19.2 hours a day, 7 days a week. While there are labor laws that regulate worker hours and require overtime pay for work in excess of 40 hours per week [27], the work 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. California minimum paid rest [28] is ten minutes every four hours worked, which amounts to about 48 minutes or .8 hours per day.⁹ Thus, the total paid hours for workers in each role will be 20 and the trucks will be operational for 19.2 hours daily.

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

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

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

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

$$\text{netFunction} = (\text{netGain} - C_{\text{variable}} - C_{\text{transport}})T - C_{\text{fixed}} = -1,652,000 + 1,700,210T$$

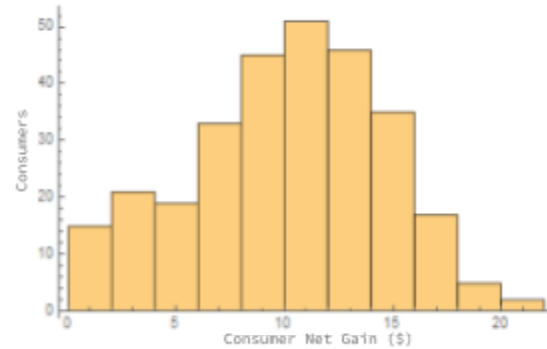
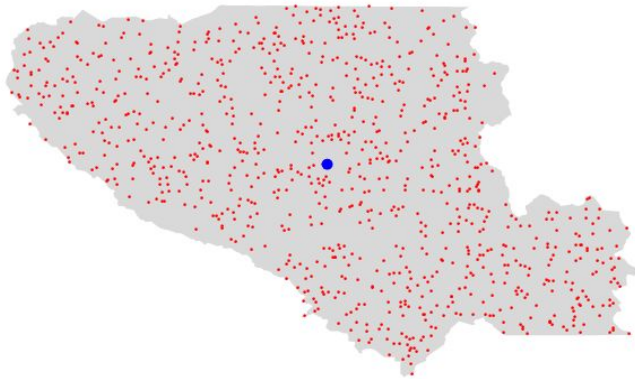
3.5 Results

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

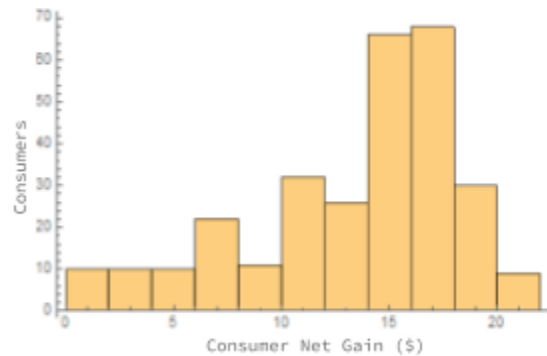
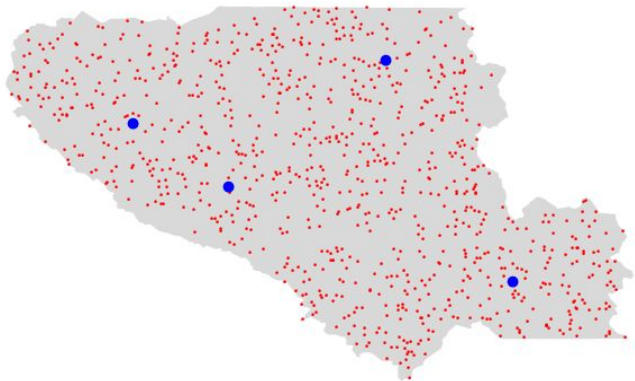
Model	Percentage of people served	Mean net gain per trip
One center	69.3%	\$7.01
Multiple centers	89.6%	\$11.78
One center with trucks	89.6%	\$11.78

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).

⁹ While this specifically applies to adult employees in the private sector, this should also serve as a reasonable guideline in our case. <https://www.dol.gov/whd/state/rest.htm#California>

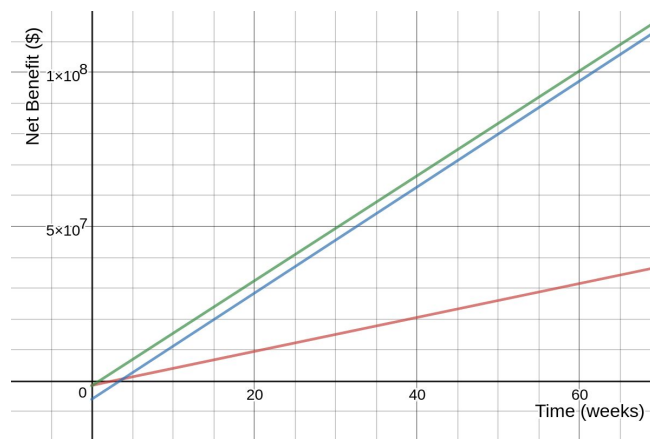


The results for the one-center model are shown above. This consumer model, as expected, generated a lower percentage of consumers served, and was, on average, more expensive due to its higher mean distance from consumers. The histogram shows the distribution of net gains of individuals who *did* go to 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.



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

1. **Realistic driving time:** In our simulations, we used Mathematica's `TravelTime` function, which calculates the driving time between two points, assuming travel at the speed limits and no traffic. These simulations are better than Euclidean or Minkowski estimates because they account for the speed and access inconsistencies of roads.
2. **Extensibility:** The model could be extended to compute results for many different locations of distribution centers. This would allow for better evaluation of where centers should be placed. We could also consider other scenarios, such as what would happen if more stops were allowed for mobile distribution centers.
3. **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

1. **Uniform population density:** For a more accurate model, the population density could be modeled probabilistically using the relative frequency of population in census blocks.
2. **Uniform distribution of food among 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.
3. **Small sample size:** Due to hardware limitations, we could only use simulations of size $n = 500$. Euclidean distance could have been used instead, which would speed up the simulation at the cost of accuracy. The Voronoi partition used a sample size of $n = 10,000$, because it is much faster than the simulation. Since the centers were placed manually, a size of 10,000 was sufficient. If a computerized optimization model for the partition were to be developed, a smaller sample size might be necessary to improve runtime.

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/storageetimes.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>
- [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>
- [27] https://www.dir.ca.gov/dlse/faq_overtime.htm
- [28] <https://www.dol.gov/whd/state/rest.htm#California>
- [29] <https://dqydj.com/average-retirement-age-in-the-united-states/>

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] (*Format imported data*)
fits = Table[
  NonlinearModelFit[data[[i]], a + b/x, {a, b}, x], {i, 1,
  8}]; (*Performs Non-linear Regression*)
Table[fits[[i]][45000], {i, 1, 9}] (*Output graphs*)
```

Model 3.4.1 Simulation

```
Clear[distances]
countyshape =
  AdministrativeDivisionData[
    Entity["AdministrativeDivision", {"SantaClaraCounty",
    "California",
    "UnitedStates"}], "Polygon"];
countyshape =
  Polygon[Reverse /@
    Flatten[countyshape[[1]][[1]],
    1]]; (*Format county shape polygons*)
center = {RegionCentroid[
  countyshape]}; (*Location of distribution center at geometric
\
centroid*)
test = Table[RandomPoint[countyshape], {i, 1, 500}];
centerUsed =
  Flatten[Map[Nearest[center, #] &, test],
  1]; (*If there are multiple centers, pick the closest one*)
distances =
  Quiet[Table[
    TravelDistance[{GeoPosition[Reverse[centerUsed[[i]]],
    GeoPosition[Reverse[test[[i]]]}], {i, 1,
    Length[test]}]]; (*Driving distance calculation*)
traveltimes =
  Quiet[Table[
    TravelTime[{GeoPosition[Reverse[centerUsed[[i]]],
    GeoPosition[Reverse[test[[i]]]}], {i, 1,
    Length[test]}]]; (*Driving time calculation*)
travelcostpermileCar = 0.19; (*From assumption 3.2.4*)
```

```
travelcost =
2*travelcostpermileCar*QuantityMagnitude[distances];
distances = QuantityMagnitude[distances]; (*Type conversion*)
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]}] (*Generate graphical
representation of simulation*)
```

```
carTravelCost = 0.19; (*Assumption 3.2.4*)
wagesLost = 13.5; (*Assumption 3.2.6*)
foodDistributedDollars = 24.36; (*From section 3.4*)
timeAtCenter = 0.2; (*Assumption 3.2.10*)
cost = Round[(carTravelCost*
  distances) + (wagesLost*(QuantityMagnitude[traveltimes,
  Quantity[1, "Hours"]] + timeAtCenter)), 0.01]; (*From
section 3.4*)
netGain = Map[#*UnitStep[#] &, (foodDistributedDollars - cost)];
"Percent of residents who access food bank"
N[Count[netGain, x_ /; x > 0]/Length[netGain]] (*Count all
netGain>0, get percentage*)
"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]]; (*Format county shape polygons*)
center = {{-122, 37.3}, {-121.6, 37.4}, {-121.4, 37.05},
{-121.85, 37.2}}; (*Center locations from Voronoi*)
test = Table[RandomPoint[countyshape], {i, 1, 500}];
```

```

centerUsed = Flatten[Map[Nearest[center, #] &, test], 1];
(*Nearest center*)
distances =
  Quiet[Table[
    TravelDistance[{GeoPosition[Reverse[centerUsed[[i]]]],
      GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}}];
(*Driving distance calculation*)
traveltimes =
  Quiet[Table[
    TravelTime[{GeoPosition[Reverse[centerUsed[[i]]]],
      GeoPosition[Reverse[test[[i]]]]}], {i, 1, Length[test]}}];
(*Driving time calculation*)
travelcostpermileCar = 0.19; (*From assumption 3.2.4*)
travelcost =
  2*travelcostpermileCar*QuantityMagnitude[distances];
distances = QuantityMagnitude[distances]; (*Type conversion*)
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]}] (*Generate graphical
representation of simulation*)

```

```

carTravelCost = 0.19; (*Assumption 3.2.4*)
wagesLost = 13.5; (*Assumption 3.2.6*)
foodDistributedDollars = 24.36; (*From section 3.4*)
timeAtCenter = 0.2; (*Assumption 3.2.10*)
cost = Round[(carTravelCost*
  distances) + (wagesLost*(QuantityMagnitude[traveltimes,
    Quantity[1, "Hours"]] + timeAtCenter)), 0.01]; (*From
section 3.4*)
netGain = Map[#*UnitStep[#] &, (foodDistributedDollars - cost)];
"Percent of residents who access food bank"
N[Count[netGain, x_ /; x > 0]/Length[netGain]] (*Count all
netGain>0, get percentage*)
"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]]; (*Format county shape*)
colors = {Green, Blue, Red, Black, Purple};
p = 2; (*Minkowski distance degree*)
points = Table[RandomPoint[countyshape], {i, 1, 10000}];
(*n=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); (*Defines Minkowski distance function*)
cindex = Flatten[
  Map[Position[#, Min[#]] &,
    Transpose[
      Table[Map[minkowski[#, hubs[[i]]] &, points], {i, 1,
        Length[hubs]}]]], 2]; (*Voronoi partition*)
final = Partition[Riffle[points, cindex], 2];
finalg = Partition[Riffle[points, Map[colors[[#]] &, cindex]],
2]; (*Apply colors*)
pw = Table[
  Style[finalg[[i]][[1]], finalg[[i]][[2]]], {i, 1,
Length[finalg]};
Speak["Data calculation completed"]

(*Graphical representation*)
pointsize = 0.03;
Show[ListPlot[pw], ListPlot[hubs]]
Speak["Graph generation completed"]

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