For office use only
T1________________
T2________________
T3________________
T4________________

For office use only
F1________________
F2________________
F3________________
F4________________

2016

19th Annual High School Mathematical Contest in Modeling (HiMCM) Summary Sheet

Team Control Number: 7231

Problem Chosen: B

Our paper presents a computerized model for optimizing warehouse placement across the continental United States to minimize induced warehouse costs and maximize overnight UPS ground shipping service coverage. Based on geographic travel characteristics such as congestion and infrastructure, our model merges taxation optimizations with warehouse optimizations to curtail charges incurred by both Math Modell's Sporting Goods and its customers.

The model achieved in Part 1 hinges on a contrived variable "shipability" applied to specific locations, a measure of the ease of movement through that location. From a definition of shipability and an explanation for its comprising factors- congestion and infrastructure- we devise an equation for shipability. We then convert the American landscape into a polar coordinate system, forging a metric for r_{ship}, the farthest possible distance travelled from a given location in a given direction with UPS ground-shipping. Alongside this math model, we crafted an algorithm to efficiently choose locations within the United States, working from the outside inwards. The resulting combination of 37 warehouses covered the entirety of inhabited-America.

Part 2 incorporates the effects of taxation by building upon the model generated in Part 1. We first model state borders with Cartesian coordinates. If the farthest point of delivery in a given direction is within state, taxes are levied and vice versa. Consolidating notions of sales tax rates and population density, we further devised a measure of tax liability; the final sections in part 2 endeavor to minimize this total tax liability in light of our Part 1 objectives.

With the addition of clothing taxes, the warehouse game is seemingly profoundly altered. We engineered an adjustment to our model from Part 2, simplistically adjusting the tax incidence to account for clothing inventory. However, considering that sales taxes were weaved into the model in Part 2, and with a justified approximations of clothing sales proportions, we noted that the maps shifted slightly, not as significantly as expected.

Empirical tests of our model yield results that efficiently cover the continental United States and markedly satisfy the additional constraints. For a realistic application of the model, perfection cannot be paramount; to ground our application in reality, we chose to forgo overnight shipping service to some sparsely populated locations within the continental US, such as Death Valley or Yosemite National Park.

Letter to President

Dear President of Math Modell's Sporting Goods,

We hope this letter finds you well, as we have come to a mathematically justified consensus concerning the optimal placement of warehouses for Math Modell's Sporting Goods. Our final recommendation is 42 warehouses placed across the United States, each warehouse with its own unique zip code. The final map of all the warehouses can be found in Appendix C, which accounts for all significant variables.

We justify this recommendation in three different steps. The first part of our analysis stemmed from reverse engineering the UPS Shipping Calculator to determine the distance each UPS driver can drive from any warehouse on any given day. The independent variables that we extrapolated through our mathematical deconstruction were congestion, infrastructure, population density, and placement on a body of water. We accounted for all of the variables and optimized our warehouse placement using the UPS shipping map, coming to the conclusion that we needed 38 warehouses across the United States as is depicted in Appendix A.

The second part of our analysis resulted from incorporating the state specific effects of tax-liability, thus causing us to change our analysis and construct a new graph that aimed to minimize tax-liability for consumers. The Business Insider emphasizes the importance of minimizing tax-liability, as consumers prefer companies to empathize with with their plight of unnecessary increased taxes. With accounting for the aforementioned independent variables and the minimization of tax-liability, we created a new spread of warehouses in the United States, illustrated in Appendix B.

The third part of our analysis resulted from the addition of clothing taxes in the model. We expected a drastic change in the number of warehouses given the severity of state differences in clothing taxes. Rather, the number of warehouses changed from 41 to 42, seeming on face as a minimal change, but there were 5 changes of locations of warehouses. The final map that includes the analysis of clothing taxes is located in Appendix C.

In taking into account the foundational variables from reverse engineering the UPS shipping map, in conjunction with the minimization of tax-liability and optimizing of clothing taxes, we hope that our mathematical model suits the needs of the company's future. The bulk of our model facilitates speedy shipping while concurrently trimming customer tax liabilities and thus the final price of the product; cheaper prices and faster shipping will indeed spark demand for Math Modell's Sporting Goods.

Sincerely,
Team 7231

Shop and Ship

Team 7231

November 20, 2016

Contents

1 Introduction

1.1 Background

With the pervasive digitization of the American lifestyle, consumers can now effortlessly harness the power of the internet at their fingertips. It's a power that entails immediate communication on myriad platforms, raunchy entertainment through mass media, and a swelling online shopping community that can access a breadth of products delivered door-to-door. With entirely-digital vendors such as Amazon and Zappos offering a seamless digital experience coupled with speedy delivery, satisfaction with online shopping is high at 83 percent [4]; considering that online sales are projected to reach 523 billion dollars by 2020 [3], traditional brick and mortar stores are scrambling to establish their web-retail presence. Indeed, the rapid shift in consumer shopping habits have sparked outrage within the commerce community, compelling Forbes' Jeremy Bogaisky to declare "a crisis in retail."[2]

Pressured by online retail and eager to expand across the continental United States, our recreation equipment company, fittingly referred to as Math Modell's Sporting Goods, has recently observed heightened demand for our product. Seeking to establish a permanent and prevalent digital sales presence, Math Modell's has tasked our team with choosing optimal locations for new warehouses. These warehouses must be situated in a manner that minimizes the required number of erected warehouses, thus preventing unnecessary costs, given that all 48 states in the continental United States are serviced with one-day grounding shipping with United Parcel Service (UPS).

1.2 Restatement of Problem

Given that Math Modell's seeks to develop its shipping capacity, we have been tasked with developing a mathematical model encompassing the following constraints and factors:

1. Optimize warehouse placement within the continental United States to minimize the number of warehouses required to ensure that all American residents are offered one-day UPS ground shipping.

2. Incorporate individual state taxation policies to analyze effects to and subsequently minimize customer tax liabilities.

3. Given that Math Modell's seeks to include clothing and apparel- taxed minimally by numerous states- in its online product offerings, discuss the impact of determined warehouse locations and discern changes to optimize warehouse locations.

1.3 Outlining our Variables

Throughout the entire paper, we employ myriad variables to simplify and augment our math model:

- **x:** throughout the paper we use x to represent a location, usually a specific zip-code.
- **S(x)** : shipability, a measure of the ease traveling through location x. S_{max} is the maximum shipability of all locations in America.
- $\mathbf{Infrastructure_x}$: a measure of the strength of infrastructure at location x.
- $\mathbf{Congestion_x}$: a measure of the amount of road congestion at location x.
- $\mathbf{Water_x}$ ($\mathbf{W_x}$): a boolean indicating whether location x is on water.
- $\mathbf{R_{tot}}$: the total mileage of all roads in a specific location.
- $\mathbf{A_x}$: the total area of a location specified by x (either a zip-code or city).
- $\mathbf{D_{pop}}$: the population density in a specific location, Pop is the population of a specific location.
- $\mathbf{r_{ship}}$: the radius of shipment through UPS ground from a specific location at a specific angle. r_{max} is the maximum shipment radius.
- **v** : velocity.
- **Coverage Area** : the total area serviced by a warehouse at a specific location.
- **k** : an adjustment coefficient, equivalent to $\frac{r_{max}}{S_{max}}$.
- $\mathbf{d_{border}}$: the distance from a specific location to a state border in a specific direction.
- $\mathbf{A_{not-taxed}}$: the area served by a specific warehouse that is not subject to taxation.
- $\mathbf{A_{taxed}}$: the area served by a specific warehouse is not subject to taxation.
- $\mathbf{T_w}$: a measure of tax liability of all consumers serviced by a specific warehouse.
- $\mathbf{t_{state}}$: the tax rate of a specific state.
- $\mathbf{pop_{taxed}}$: the population serviced by a specific warehouse that incurs taxes.
- $\mathbf{t_{average}}$: a measure of tax rates incorporating clothing.

2 Optimization Part I

2.1 Assumptions and Simplifications

1. **Assumption:** Weather has no statistically significant effect on the transport distance or time for our goods.

   Justification: We reason that in determining a mathematical model for 1 day shipping of the United States, unique circumstances such as a blizzard or a tornado ought not play into the model because our model is for the large majority of the year. Accounting for weather would be superfluous and unreasonable as weather often cancels shipping altogether. Thus we assume weather has no statistically significant effect on transport time or distance.

2. **Assumption:** Varying levels of traffic and any changes due to time of year are not significant to the problem.

   Justification: We reason that the majority of the shipments will not change based on time of year. The only potential changes might be during major holidays due to traffic, but we argue that the rest of the year is far more important to take into consideration and model rather than the secondary concern of time of year.

3. **Assumption:** We assume shipments occur on a typical workday.

   Justification: We reason that any atypical circumstances that might change a regular work day is not important enough to calculate for two reasons. Firstly, calculating it would not be applicable to the real world on a daily basis as atypical circumstances do not occur often, and thus would be superfluous to calculate. Secondly, adapting our model to account for said circumstances would be near impossible as such circumstances are impossible to predict, leaving us with solely empirical analyses to base our model off of. This would be entirely unproductive and take time and resources away from solving the far more important problem of one-day shipping on a regular work day. Furthermore, we reason that it is impossible to have shipments over the weekend as UPS trucks operate solely during the work week.

4. **Assumption:** We assume that any change in altitude has a minimal effect on the distance for which one day shipments are possible.

   Justification: We reason that one day shipments occur in a relatively local area. Thus, any altitude change would not be drastic enough to cause for a large change in the maximum distance for a one day shipment. Rather, we account for infrastructure on the whole in our optimization, through length of roads and number of highways. We reason that this is sufficient in the goal of accounting for any variability of route between different cities in the country.

2.2 General Strategy

In order to efficiently find an optimal solution, bashing is not permissible. We propose to split the task into 3 parts.

The first step is an approximation of potential warehouse locations. The algorithm is as follows: we place warehouses in certain locations. If our current warehouse system doesn't give us the delivery power we need, we add warehouses in the middle of our existing warehouses. Say we have a warehouse in Nashua, NH and another one in Orlando, FL, we'll put our next warehouse around the middle Atlantic states.

However, our original algorithm lacks accuracy. Our second step is to thus employ a mathematics model to determine which city in the area is the best. We reverse engineered UPS's model for shipping radius around cities based on their infrastructure and congestion. With this model in mind, we can find potential candidate cities in that area.

The final step is checking whether our implementation works in an efficient manner. We write a program that checks and draws the area that can be covered by UPS's one day shipping, given the warehouses we have. Then we repeat what we just did. Each sequence of those 3 steps will efficiently improve the shipping system's efficiency, and eventually when we cover the entire US it is the optimal solution.

2.3 The 'Shipability' Function

With shipment speeds varying across the continental US from location to location, and lacking access to UPS algorithms that determine transit times, we attempted to reconstruct a mathematical model that calculates shipment speeds for given warehouse locations across the country. The bulk of our work involved determining an equation modeling the shipment radius for a given location, and subsequently adopting the equation to optimize warehouse placement.

In crafting our transit time calculator, which given a warehouse location returns a locus of points within a one-day ground shipment radius, we incorporated a notion we call 'shipability.' Shipability measures the ease of ground-travel through a particular location- that is, higher shipability is associated with faster road travel while lower shipability indicates slower travel.

Bear in mind that shipability is rather volatile, as the ease of movement through a particular location or highway changes day-to-day, hour-to-hour, minute-to-minute. Indeed, the factors affecting transit time (and thus shipability) are manifold; from weather to location to time-of-year to unpredictable crashes, incorporating all these variables would be overwhelming and ineffective. Rather we have focused our shipability function on three noteworthy variables, simplifying and, in a sense, strengthening our analysis:

- **Travel Congestion:** Naturally, greater crowding in a particular area hinders travel. While urban areas are teeming with traffic lights, pedestrians, and windy roads with low speed limits- thus interfering with expedited travel- rural areas are empty and often grant higher speed limits, enabling faster travel. Our measure of travel congestion involves population density, as it accurately characterizes the number of people within a specific zip code. Furthermore, population density is closely correlated with car crashes[8], the largest single catalyst of traffic. Travel congestion as measured by population density thus wields great influence over the travel time through a particular location (shipability).

  To algebraically model traffic congestion, note that ground shipping and trucks driving across a road is analogous to electrical particles traveling through metal wire. Indeed, each aspect of ground transport finds a parallel quantity with electrical wiring; while current represents the speed of movement or shipability, resistance equates to population density and voltage represents the total population. Note that formally, Ohm's law relating current, resistance, and voltage (or in our case shipability, population density and total population respectively)is

$$V = IR$$

where V =Total Population, R =Population Density, I =Shipability. Since total population remains relatively constant, we can discern that shipability depends on the inverse of population density, denoted as D_{pop}.

- **Infrastructure:**

  The quality of infrastructure in a particular location affects a driver's ability to drive through that location; regardless of congestion, a dearth of roads with create for longer, windy routes while an abundance of roads will enable faster point-to-point travel, and greater shipability. We thus measure the availability of a location's infrastructure by the total road length within that location.

  In calculating the infrastructure coefficient, we incorporate the total mileage of all roads sprawling across a specific zip code. Note that road lengths naturally increase with larger zip codes; the size of the zip code must be integrated. One solution is to simply divide the total road mileage by the entire zip code area. This approach, however, would construct a poor representation, as area is two-dimensional and road lengths are one-dimensional values. Rather, to assure that all units and dimensions are accounted for, we divide the total mileage of all roads by the square root of the zip code area. Formally, given a zip-code at location x,this produces the following equation:

$$Infrastructure_x = \frac{R_{tot}}{\sqrt{A_x}}$$

  where R_{tot} is the total road mileage within zip-code x, and A_x measures the area within zip-code x.

- **Water Boolean:** This variable W_x assumes two values: 1 and 0, indicating whether the location is water or land (water corresponds to 0, since the goods cannot be ground-shipped across water).

With the specific coefficients now defined and justified, we can move forward to constructing our shipability function. Let x be a zip code in the continental United States. $S(x)$ is then defined as the 'shipability' of our location x, characterizing the ease of traveling through x. Then

$$S(x) = Congestion_x * Infrastructure_x * Water_x$$

$$= \frac{1}{D_{pop}} * \frac{R_{tot}}{\sqrt{A_x}} * W_x$$

Note that the area of a zip code or A_x is equivalent to the total population of a zip code divided by the population density. That is, we have that

$$A_x = \frac{Pop}{D_{pop}}$$

Thus our shipability function becomes

$$S(x) = \frac{1}{D_{pop}} * \frac{R_{tot}}{\sqrt{\frac{Pop}{D_{pop}}}} * W_x = \frac{1}{D_{pop}} * \frac{R_{tot}}{\sqrt{Pop * D_{pop}}} * W_x$$

2.4 One-Day Ground Shipping

In section 2.3, we conceived of and crafted a notion for our shipability function, or the ease of traveling through a certain zip code. We observed and quantified the factors that affect traveling ease, constructing our final equation accordingly. This section takes the model a step further, harnessing calculus to determine how far from a warehouse one-day shipping can deliver.

Note that $S(x)$ or the shipability of a certain location (x is a zip code), measures the ease of travelling through that location. The shipability of a location is thus comparable to the velocity observed travelling through that location; greater shipability entails easier travel and faster velocities, while lower shipability equates to harder travel and slower velocities. However, in incorporating shipability into distance calulcations, we must use relative shipabilities (as shipability itself has no useful value). We thus employ the notion of relative shipability as $S(x)/S_{max}$, where S_{max} is the maximum shipability across the United States.

In determining the farthest possible distance for one day ground shipping from point x in an angle theta, we must consider the shipability of every point passed as we move from x in the direction of theta. This thus requires integrating over relative shipabilities of locations covered from day 0 of travel to day 1.

Furthermore, simply summing up relative shipabilities of various points would yield values proportional to the one-day ground shipping radius, but not the radius itself. Our equation thus multiplies the final integral- or the sum of all relative shipabilities (to the maximum shipability) at an angle theta from day 0 to day 1- by the maximum radius to calculate the final radius.

To thus calculate the radius of one-day ground UPS shipping (r_{ship}), from x in direction θ, we have the following formula:

$$At\ angle\ \theta,\ r_{ship} = \int_{0\ days}^{1\ day} v\ dt = r_{max} \int_{0\ days}^{1\ day} \frac{S(x)}{S_{max}} dt$$

For the purpose of future calculations, we will use $k = \frac{r_{max}}{S_{max}}$ because k is a constant that incorporates all of the other constants, making it more useful for comparisons. The above equation becomes the following:

$$At\ angle\ \theta,\ r_{ship} = r_{max} \int_{0\ days}^{1\ day} \frac{S(x)}{S_{max}} dt = k \int_{0\ days}^{1\ day} S(x)$$

Figure 1: The radius of shipment (r_{ship}) from point x in direction θ

2.5 Determining Coverage Areas for Warehouse Locations

From section 2.4, we crafted an equation for r_{ship} in direction theta, a measure of how far a package could be shipped in one day in direction theta starting from a warehouse at location x.

To aid fine-tune our empirical model, we employed our equation of r_{ship} to calculate the area of shipping locations covered by placing a warehouse at location x. Note that section 2.4 offered a polar coordinates approach- we calculated our radius of r_{ship} of maximum shipment distance in one day based on a certain direction of travel, theta. To find the total coverage area given a warehouse location at x, we can thus integrate over the angle θ as follows:

$$Coverage\ Area = \int_0^{2\pi} \frac{(r_{ship})^2}{2} d\theta$$

This measure provides for the area of coverage given a warehouse location at x. By comparing coverage areas of different locations, we can tailor our empirical model; determining warehouse locations by choosing zip codes with the highest coverage areas will minimize the total number of warehouses required, thus reducing Math Modell's costs. That is, given two locations A and B within the Continental United States if $CoverageArea_A < CoverageArea_B$, then we would be more inclined to place a warehouse at B.

2.6 First Steps

Our general strategy, as described in section 2.2, is to add new warehouses around the middle point of existing warehouses. It would thus be logical to place our initial warehouses in the four corners of the United States. We chose to place one of them at the headquarter in New Hampshire, and the other three in Florida, California and Washington, naturally in large cities.

We employ a program written in Python OpenCV libraries, which sends the zip of our warehouses to UPS website to download the shipping map, and then overlays maps on top of each other to calculate the area covered by one-day shipping. The overlaying program we wrote takes in multiple UPS maps, and colors areas yellow if yellow on one map (meaning that the area can be distributed from one of the warehouses withing one day). The yellow parts are areas that can be distributed within one day under current warehouse system. The result is below:

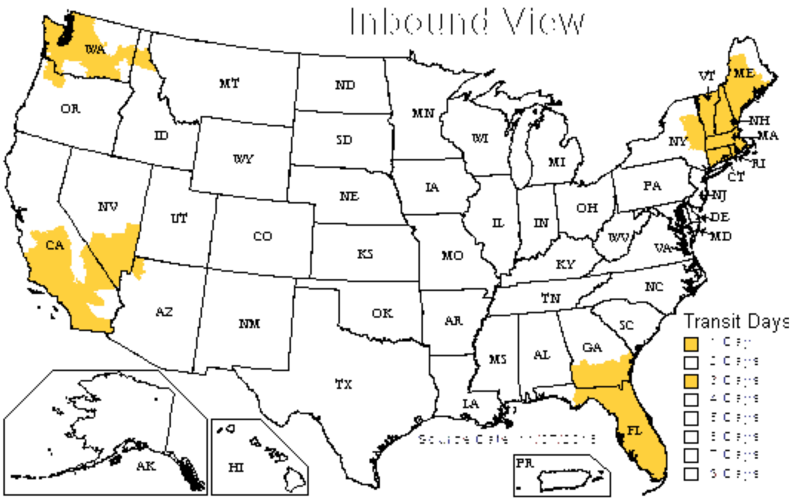

Figure 2: First groups of warehouses.

2.7 Algorithm Calculations: Ellensburg vs. Seattle

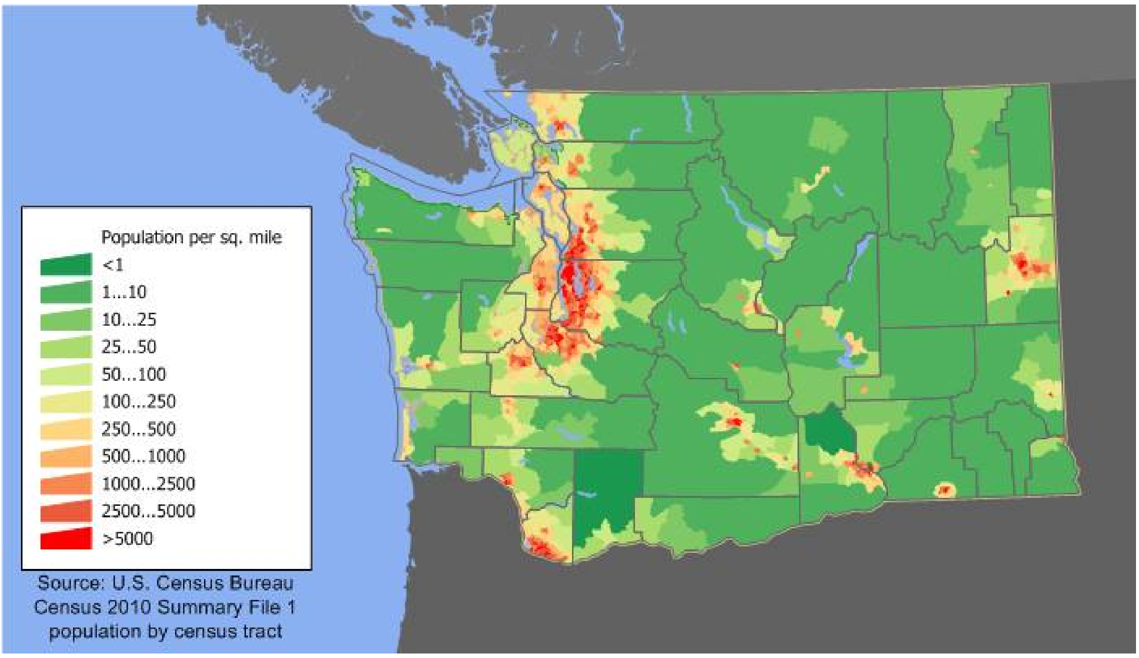

Figure 3: Population density in Washington.

Using the equations generated, we can develop good approximations for the relative difference between how much area the postal service can cover from different zip codes. Take, for example, two relatively close cities in the state of Washington, Seattle (the capital of Washington) and Ellensburg (central Washington).

It is wholly unrealistic to find the shipability at every point on the U.S. map, but calculating the area of one-day ground shipping from a specific location compels an integral. Since integrals require a continuous equation, the shipability must be approximated into a continuous function. To do this, certain locations must model certain areas.

A significant radius beyond Seattle has demographic information similar to Seattle, so information about Seattle will model both its surrounding region, affecting about one-fifth of the radius. Same goes for Ellensburg to a smaller effect, affecting about one-tenth of the radius. The majority of the area of Washington is made up of areas like Greenwater, which is rural and has low population density, so we will model the rest of Washington with information about Greenwater. For estimation purposes, due to the presence of water, the radius for Seattle will decrease by 80 percent.

Here is the information about Seattle, Ellensburg, and Greenwater:

City to variable	Seattle	Ellensburg	Greenwater
Population Density	5032.752	52.40319	51.70167*
Population	806817	30239	20987*
Total Road Length (miles)	1677 (Kubly)	70	50*

* = ("Greenwater, Washington Visitor Information")

Using the shipability, radius, and area functions we formulated in parts 2.3 to 2.5, we can add the table.

Shipability	Seattle	Ellensburg
Approximate radius	$k(.04366)$	$k(.04876)$
Adjusted radius	$k(.03493)$	$k(.04876)$
Approximate area	$k^2(.00383)$	$k^2(.00747)$

Comparing the areas that result from the calculations, we find that the area from Ellensburg is twice the area from Seattle. Noting this, it is more optimal to locate a warehouse in Ellensburg rather than in Seattle. Thus we changed our warehouse from Seattle to Ellensburg. Now our map looks a lot better.

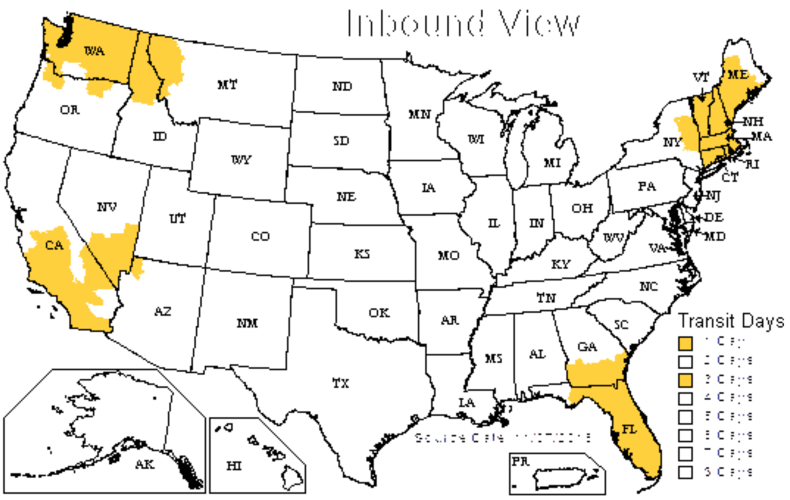

Figure 4: Modified warehouse location.

2.8 Results and Discussion

From our current warehouses, we can discern where to place additional warehouses: one in northern California or Oregon, one around Minnesota or Wisconsin, one around middle Atlantic States and one around Texas. With our modeling of shipability, we can quickly gain a sense of what kind of cities have high shipability, and we put warehouses in those cities: Sacramento, CA, Minneapolis, MN, Richmond, VA and Austin, TX.

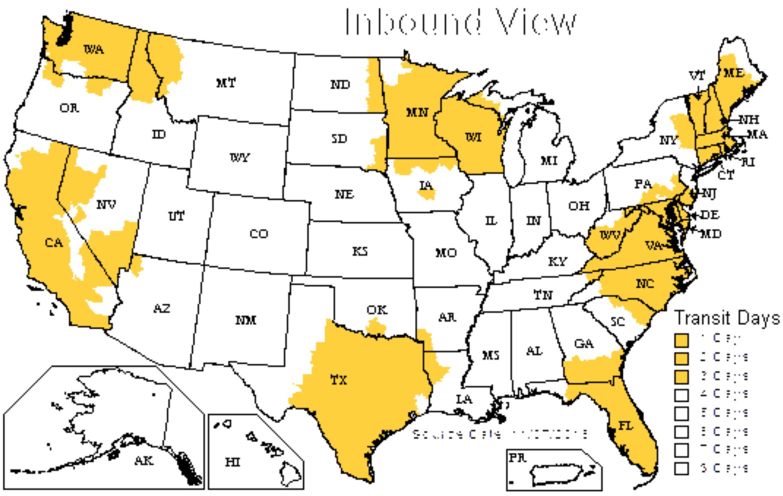

Figure 5: Second groups of warehouses.

Repeating the steps, we put more and more warehouses which cover as much areas with as few overlaps as possible. Our next group of warehouses determined by our general strategy and model are located at the following zip codes: 97470, 59101, 48210, 14850, 16801, 79935, 85029, 35216, 11231.

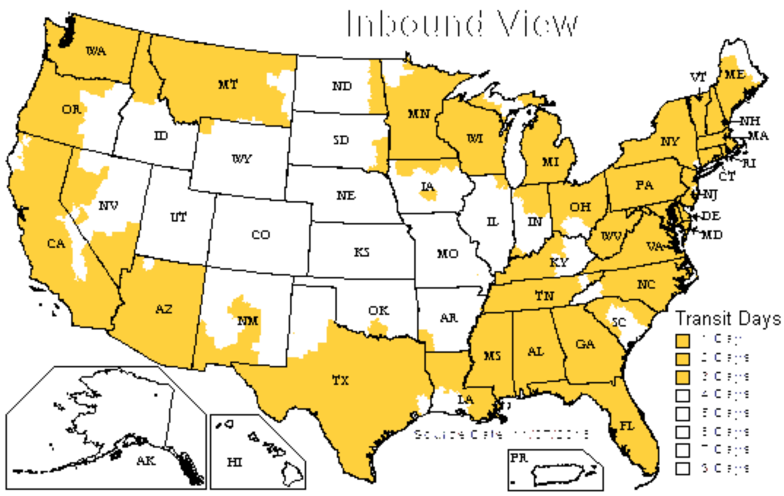

Figure 6: Third groups of warehouses.

Repeating the same steps, we add warehouses to the following zip codes 28211, 40962, 58104, 57543, 83278, 97720, 89104, 70808, 89883.

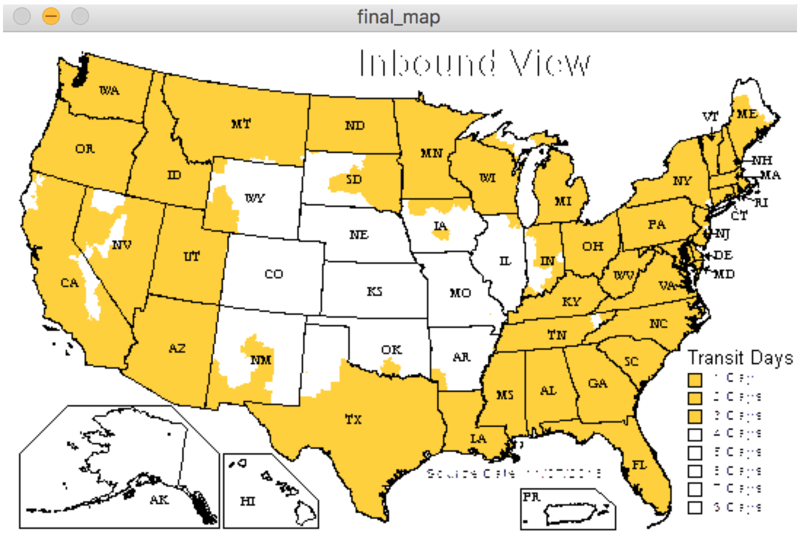

Figure 7: Fourth groups of warehouses.

And finally, we add our last group of warehouses to the following locations: 04747,82601, 47909, 65201, 80926, 69138, 51106, 69201, 87508, 73160, 72002, 67601.

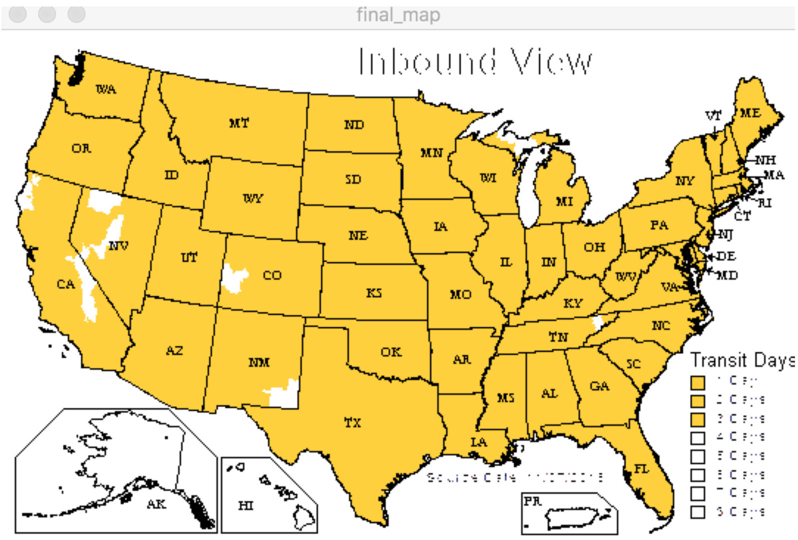

Figure 8: Final groups of warehouses.

We suggest that Math Modell's Sporting Goods builds 38 warehouses in locations specified by the following zip-codes: 03062, 90016, 98926, 32824, 95818, 55407, 23225, 78701, 97470, 59101, 48210, 14850, 16801, 79935, 85029, 35216, 11231, 28211, 40962, 58104, 57543, 83278, 97720, 89104, 70808, 89883, 04747, 82601, 47909, 65201, 80926, 69138, 51106, 69201, 87508, 73160, 72002, 67601 (See Appendix A for map).

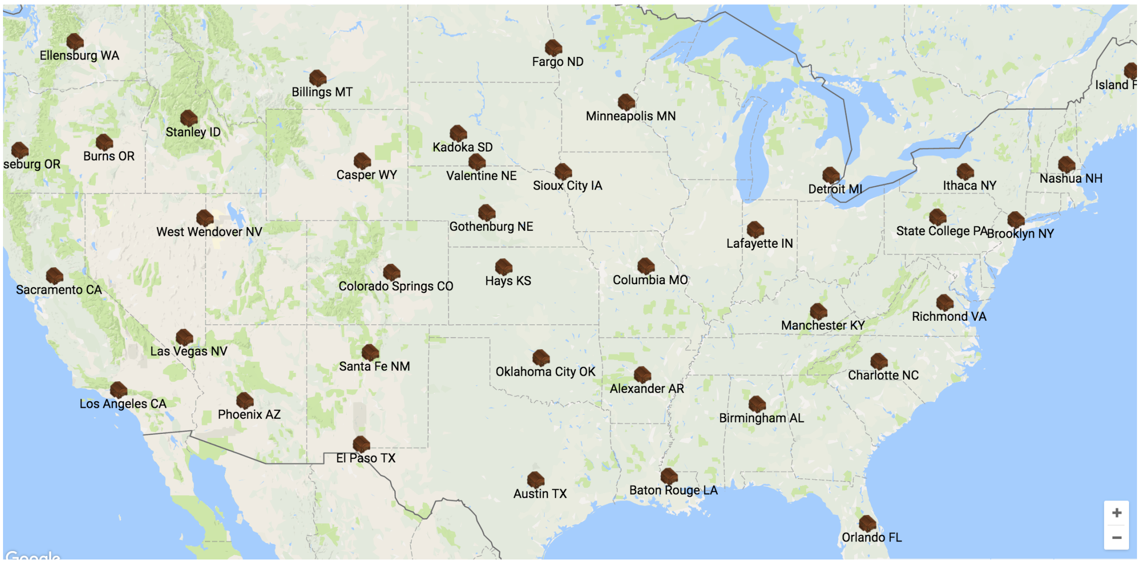

Figure 9: Location of warehouses, part I.

The total coverage area is displayed in Figure 8: Final groups of warehouses. Note that not all of the Continental United States is entirely yellow; there are a few white sections on the map. We understand that Math Modell's hopes to reach the entire American population, but we believe that the white sections shown are so sparsely populated- they include Death Valley National Park, Yosemite National Park, Black Rock Desert, and Dominguez-Escalante National Conservation Area - that establishing warehouses that merely cover those locations would be inefficient and unnecessary. We thus advise that Math Modell's only places warehouses in locations specified by the zip-codes above.

2.9 Sensitivity Analysis

Remember that the shipability of a city is calculated by

$$S(x) = Congestion_x * Infrastructure_x * Water_x$$

$$= \frac{1}{D_{pop}} * \frac{R_{tot}}{\sqrt{A_x}} * W_x$$

Assuming that the city, which we simplify to be a circle, gets larger with population density unchanged, its shipability changes based on $\frac{R_{tot}}{\sqrt{A_x}}$. The road length usually depends on the radius of the city, and the area depends on the square of the radius of the city. Further, $\frac{R_{tot}}{\sqrt{A_x}}$ remains the same, making the city as good a candidate as it was before. But the problem is that as city gets larger, population density changes. We looked at the data on town area versus population density in the entire nation [1]. After omitting outliers, we plotted population density against area. The graph is below:

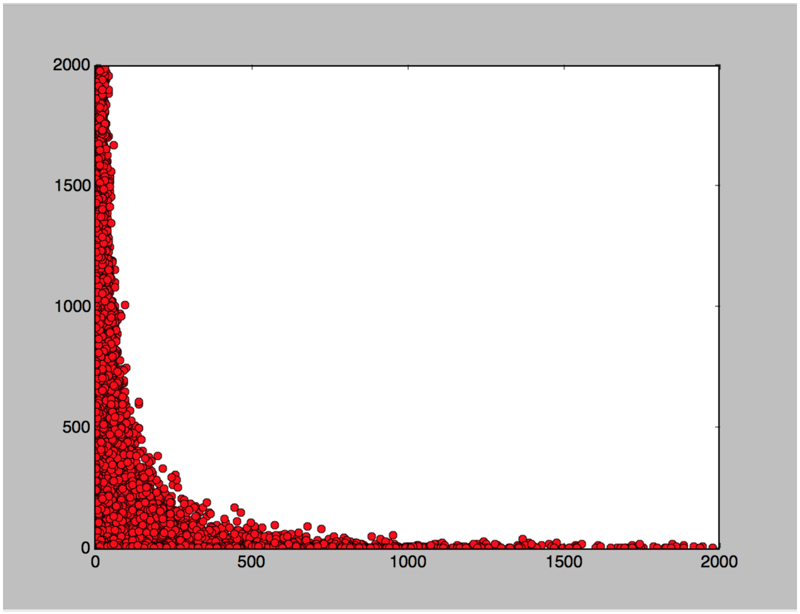

Figure 10: Population density against area.

Although there are too many data to find a function that approximates the shape (the shape looks like an inverse function, but all those points clustered in the bottom left make a clean equation impossible), the graph shows that in general, the bigger a region of zip code is, the less population density is. But in bigger cities the region of zip code is usually smaller, which means that the larger the city, the larger the average population density. And since density is in inverse relationship with size of zip code region and the size of zip code region is in inverse relationship with city's area, population density is directly proportional to the square of the x

$$D_{pop} = \frac{x}{A_{zipcode}}$$

$$A_{zipcode} = \frac{y}{A_{city}}$$

$$D_{pop} = \frac{xA_{city}}{y} = \frac{xRtot^2}{y}$$

Now let's return to our model:

$$S(x) = Congestion_x * Infrastructure_x * Water_x$$

$$= \frac{1}{D_{pop}} * \frac{R_{tot}}{\sqrt{A_x}} * W_x$$

If the city's radius is multiplied by x, then R_{tot} and $\sqrt{A_x}$ are both multiplied by x while D_{pop} is multiplied by x^2. Thus $S(new) = \frac{S(old)}{x^2}$.

3 Sales Tax Part II

3.1 Assumptions and Justifications

1. **Assumption:** We assume the state tax rate for all goods we are selling.

Justification: We argue that this is true for three unique reasons. Firstly, recreational equipment generally follows the state trends for taxes. Thus, we extrapolate that the state taxes would be reasonable to use. Secondly, we reason that the important distinction to be had for the mathematical modeling is the difference in tax rates between states, rather than a difference in taxes between goods. Finally, we reason that

2. **Assumption:** We assume that the model is calculated for days that are not tax-free.

   Justification: We reason that tax-free days happen for a maximum of three days per year, thus making it highly probable that any given day is not a tax-free day. Thus, we argue that the probability of tax-free days is not statistically significant enough for us to account for in our mathematical model.

3. **Simplification** We simplify that state boundaries within the United States are linear.

   Justification: It is nearly impossible for us to calculate the area of shapes that contain curvature that is unique and does not conform to normative shape orientations. Thus, we simplify the boundaries on a minimal scale so as to successfully calculate the areas of the regions for the warehouses. The minimal simplification does not have a statistically significant change to the true area of said regions.

4. **Simplification** The amount of inventory bought at a stor is proportional to the number of people in a region.

   Justification: Naturally, there will variance of the amount of inventory shipped; certain regions will have people who generally buy more items, while other regions will have people who buy less. However, in this model, due to the high number of people involves, it is safe assume no individual variation. If any one person buys a fixed amount of inventory, then in general we can say that the greater the number of people, the greater the amount of inventory sold.

3.2 The Effects of Taxation

Now that tax liabilities have entered the playing field, the warehouse locations we selected in Part 1 ensure not only the availability of ground shipping to all Americans, but they also profoundly impact the prices of Math Modell's products; sales taxes change based on warehouse locations, and the final product prices observed by customers will change as a result. It is thus imperative that our company Math Modell incorporates tax liabilities into it's warehouse location model, considering that misplaced warehouses could balloon the prices of certain goods and deter product demand.

Indeed, different states have different taxes; while New Hampshire imposes absolutely no taxes, California imposes a 7.5 percent tax rate, the highest of all 50 states. The locations of the warehouses will affect product prices, and thus customer demand. Our model is currently optimized to service everyone in the continental United States, but it fails minimize tax liability. We thus observe warehouses cluttered in highly taxed locations, resulting in suboptimal solutions.

For example, in Part 1 we placed a warehouse in Minneapolis, MN since it maximized coverage area in the region. In a optimized scenario considering taxation, however, since Minnesota has a whopping 6.88% sales tax, the warehouse would likely disappear, and customers in Minnesota would be serviced by warehouses close to the border in surrounding states.

Naturally, by places warehouses in states like Colorado (2.9%) and Oregon (no sales tax) where sales taxes are low or non-existent and avoiding states like California (7.5%) and Mississippi (7%), we can minimize the tax liability for our customers. Furthermore, as sales taxes does not apply for shipments outside state borders, placing warehouses closer to state borders will ensure tax-free coverage of surrounding states, thus reducing tax liability.

In this section, we flesh out the model produced in Part 1 to account for tax liabilities. We first endeavor to model state borders using Cartesian coordinates to obtain distance to a border. We compare this distance to the radius made using the shipability function, introducing a method to calculate total tax liabilities in certain locations. We end with our empirical analysis and adjusted warehouse locations that optimize both tax liability and warehouse placement.

3.3 Modeling State Borders

Since tax rates for shipping across any state border is zero, it is necessary to determine whether the shipping radius area crosses any state borders. These borders can be represented mathematically. In Cartesian coordinates (xCoor, yCoor) represents a location for a warehouse in a state, and (x, y) are the coordinates for some location that does not contain the warehouse. The difference in location between (x, y) and (xCoor, yCoor) can represented as delta(x) and delta(y), which are:

$$\Delta x = x - xCoor$$

$$\Delta y = y - yCoor$$

From here, the borders can be represented as a set of inequalities, where a_n represents slope and b_n represents the value where the dependent variable equals 0:

$$\Delta(y) > a_1 \Delta x + b_1$$

$$\Delta(y) < a_2 \Delta x + b_2$$

$$\Delta(x) > a_3 \Delta y + b_3$$

$$\Delta(x) < a_4 \Delta y + b_4$$

and perhaps other equations, depending on how "boxy" the state borders are.The second two equations are slightly different in case and equations involve a slope of infinity. These equations will intersect and bound the area of a state. When we start at a specific location for a warehouse, all of these equations will evaluate to be logically true.

That's it for the set-up. From any location, a sufficiently long radius will eventually intersect with a border. This is called d_{border}, or distance to the border. If the location where the sufficient radius and border meet is denoted as (x_d, y_d), then we have that $\Delta x_d = x_d - xCoor$ and $\Delta y_d = y_d - yCoor$. This means that:

$$d_{border} = \sqrt{(\Delta x_d)^2 + (\Delta y_d)^2}$$

Now we apply the radius-shipability function at all angles for the same location. If $r > d_{border}$, the radius crosses a border. Since shipping across a border involves no tax, this means that part of the area created from the radius function is not taxed. It is also possible the radius-shipability function gives a value for the radius that never violates any of the inequalities in our set of borders, in which case the entire radius bears the weight of the tax in the state.

3.4 Adjusting our Model

The problem duly notes that orders delivered to locations within a state where warehouses are located will have that state's tax added to the order cost, whereas orders delivered to locations outside a state where a warehouse is located will not be taxed. This poses an complexity that we must incorporate into the model of part 1- that is, we must be able to discern between locations within and outside the warehouse's state boundaries.

From part 3.2, we crafted a method of, given a warehouse location at x, calculating the distance from the warehouse to the state border (d_border) at an angle theta. Below, we extend our model to calculate tax liabilities to customers within a certain location.

Part 1 of the problem introduced a method to calculate the radius of shipment in a certain angle for one-day UPS ground, referred to as r_{ship}. Note that r_{ship} at angle theta can extend beyond the boundaries of the state; r_{ship} can be greater than d_{border}. If so, then the warehouse services out of state residents, who are not liable for taxes. On the other hand, if r_{ship} is less than d_{border}, than the warehouse doesn't extend far enough beyond state boundaries, and every customer within that boundary is tax liable.

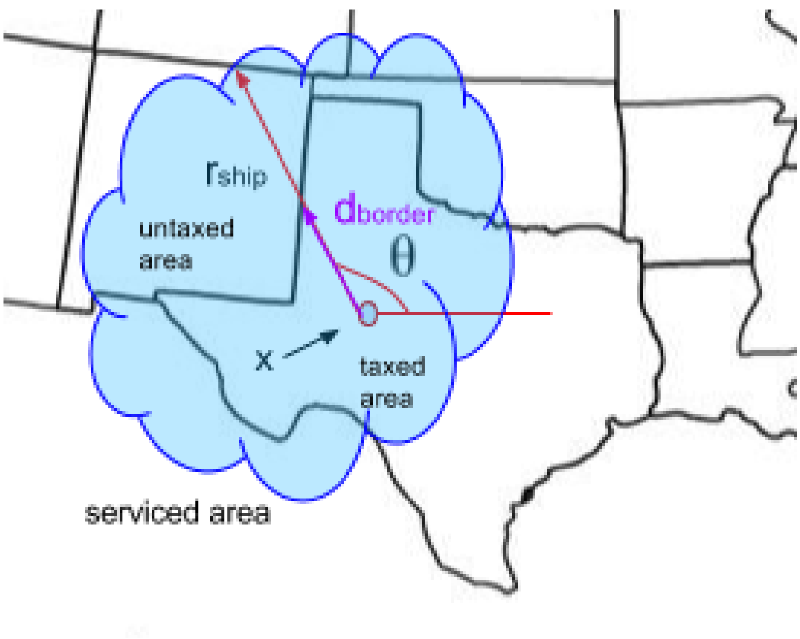

Figure 11: A representation of d_{border} and r_{ship} given a shipment area. Includes taxed and tax-free areas

Note that we cannot directly calculate the area that is taxed within state; we don't know whether for any angle, r_{ship} is less than or greater than d_{border}. However, we can calculate the area that is taxed out of state. If one-day delivery extends out of state, we know that $d_{border} < r_{ship}$ for a given theta, and the distance not taxed (out of state) is characterized by $(r_{ship} - d_{border})$. Summing over all possible angles of theta, we have:

$$A_{not-taxed} = \int_0^{2\pi} \frac{(r_{ship} - d_{border})^2}{2} d\theta$$

where $A_{not-taxed}$ is the tax-free out-of-state area serviced by the stationed warehouse.The above equation provides the total tax free area that a warehouse services; to calculate the tax liabilities leveraged on customers within state, we measure the total land area within state that the warehouse services. This is simply the total out-of-state tax-free area subtracted from the entire area serviced by the warehouse:

$$A_{taxed} = \int_0^{2\pi} \frac{(r_{ship})^2}{2} - \frac{(r_{ship} - d_{border})^2}{2} d\theta$$

In quantifying tax liabilities imposed by specific warehouses, we looked for a measure that would incorporate state tax prices as well as the area within the state serviced by the warehouses. However, the total tax liability of all customers within a specific area depends not on the area of the area taxed (A_{taxed}) but rather the population of the area. We thus multiplied the state tax (t_{state}) by the population of the tax-liable area (Pop_{taxed}), ensuring that when implementing our minimization functionality later, we would minimize the tax-liable area and incorporate individual state taxes into our analysis. Our measure of tax liability is thus as follows:

$$T_w = t_{state} * pop_{taxed} = t_{state} * D_{pop} * A_{taxed}$$

where T_w is a measure of the total tax liability imposed by a single warehouse and D_{pop} is the population density. In observing the effects of taxation on our entire warehouse-placement model, and calculating the total tax liability across the United States with our given model, we simple need to sum the individual tax liabilities of each warehouse across all warehouses:

$$\sum_{warehouses} T_w = \sum_{warehouses} t_{state} * D_{pop} * \int_0^{2\pi} \frac{(r_{ship})^2}{2} - \frac{(r_{ship} - d_{border})^2}{2} d\theta$$

In the next section, we will observe how to minimize this sum, thus minimizing tax liability, without without significantly increasing the number of warehouses erected; we optimize tax liability and warehouse costs alike.

3.5 Model Justification

Just as we justified the previous model comparing Ellensburg and Seattle, we will justify this model comparing two cities, this time Ellensburg and Spokane (both in Washington again). In the previous model, we wanted to maximize the area each of the warehouses covered by one-day ground shipping. However, this did not minimize tax liability. Spokane is located in East Washington, so it is close to the

Idaho border. Ellensburg, in contrast, is located in central Washington, relatively farther from the Oregon border. The analysis here is analogous to the previous problem except for the effect of tax, so we used Greenwater to model the rest of Washington again. The road length for Spokane is estimated from similar cities where data is available. Also, there is no need for an adjusted radius for this case. Below is all the information that does not include tax:

City to variable	Spokane	Ellensburg	Greenwater
Population Density	160.7939	52.40319	51.70167
Population	877485	30239	20987
Total Road Length (miles)	600	70	50

Shipability	Spokane	Ellensburg
Approximate radius	$k(.04161)$	$k(.04876)$
Approximate area	$k^2(.00544)$	$k^2(.00747)$

We find that Spokane covers a smaller area than Ellensburg. This is not a concern because we want to minimize tax liability, not necessarily the number of warehouses. In fact, the area that Spokane does not cover can be covered by another warehouse in North Oregon. This means that any shipping to that area will be tax-free. Comparing the areas, we find that Spokane gives an area about $1 - \frac{area_{Spokane}}{area_{Ellensburg}}$ $= 27$ percent. Because Spokane is in East Washington, we expect the lost area to be in Southwest Washington, where a warehouse in Oregon can cover it.

Eyeballing the produced graphic below, we see that the area that is not covered by Spokane but covered by Ellensburg approximately one-fith of Washington in its southwest corner, which matches our model very closely. We find with evidence that Spokane is a better location for

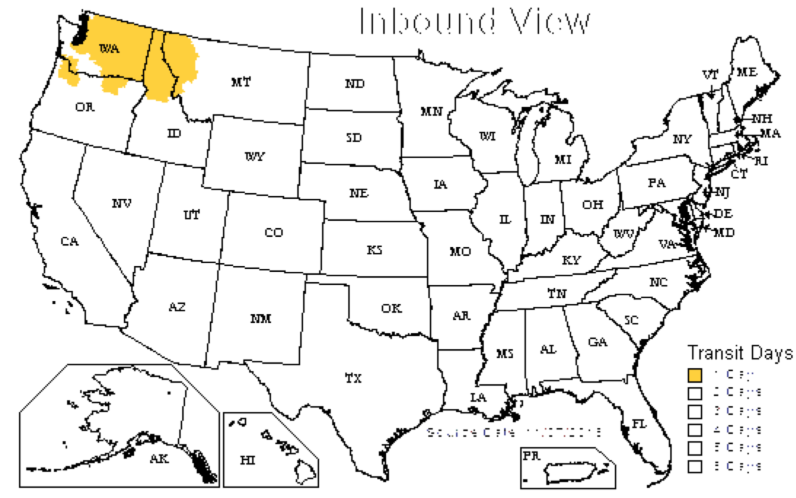

Figure 12: When the warehouse is at Spokane

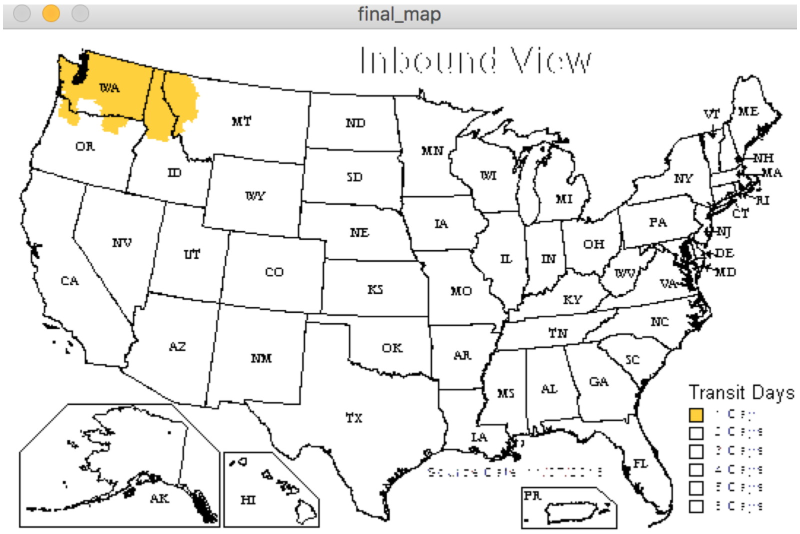

Figure 13: When the warehouse is at Ellensburg

3.6 Relocating the Warehouses

We will change our general strategy a bit. Instead of adding new warehouses between existing warehouses, we add warehouses around the boarder of states, then fill in the middle part. Then we employ our modified model to pinpoint the exact position of the new warehouses and use our program to display what area our warehouse system covers.

We reconsider initial warehouse locations. The one in Nashua stays at the same spot because that's where Math Modell's headquarter is. The one in Florida does not change because if we move it to the center there is no warehouse to cover the southern part of Florida. The one in California remains because the boarder of California are desert and national parks. However, our warehouse in Washington changes from Ellensburg to Spokane and we add an additional warehouse in Joseph. So our initial map looks like this :

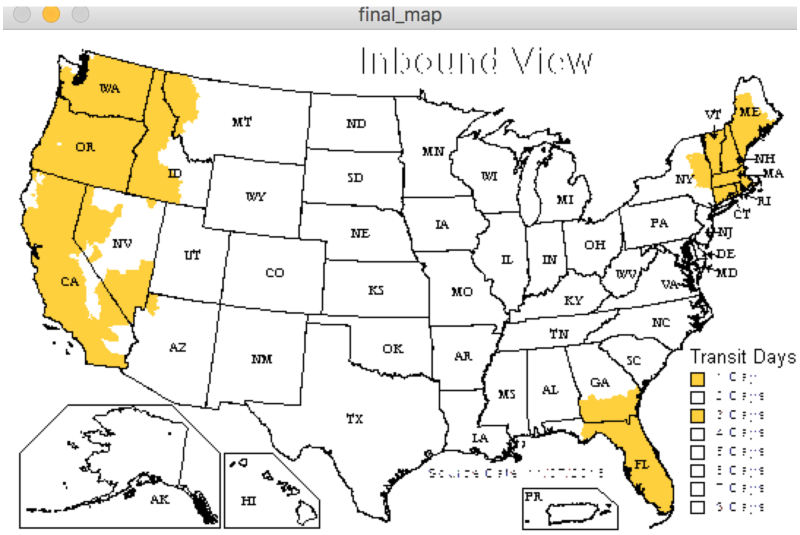

Figure 14: Ininial warehouses.

Using our new general guideline, we try to expand into Montana, Pennsylvania, Georgia, Arizona and Utah. We add warehouses near the boarder of those states and another empty state- for example, on the boarder between Pennsylvania and Ohio or between Georgia and Tennessee. Specifically which warehouse to be added will be estimated by our new model which incorporated tax liability. And our expansion should look like this:

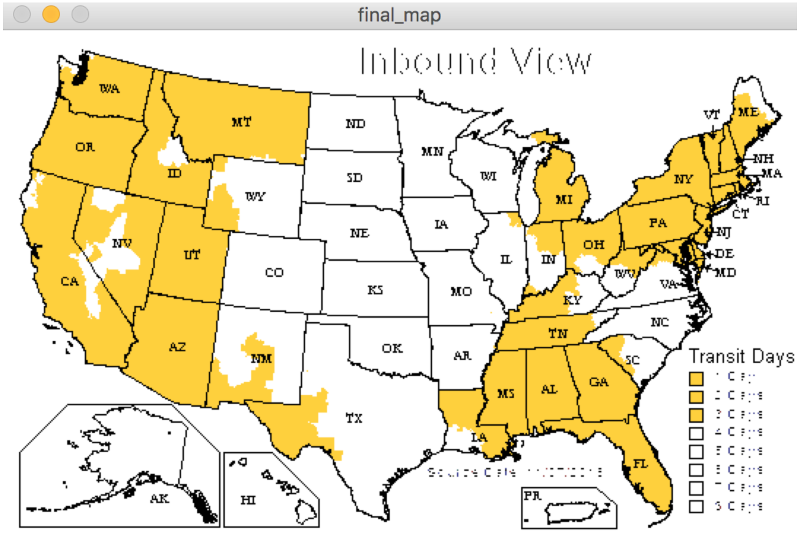

Figure 15: Second group of warehouses.

We use similar strategies in the uncovered states in the middle. And the final map looks pretty much similar to the one we had in Part I of the problem, but there are many relocation of the warehouses.

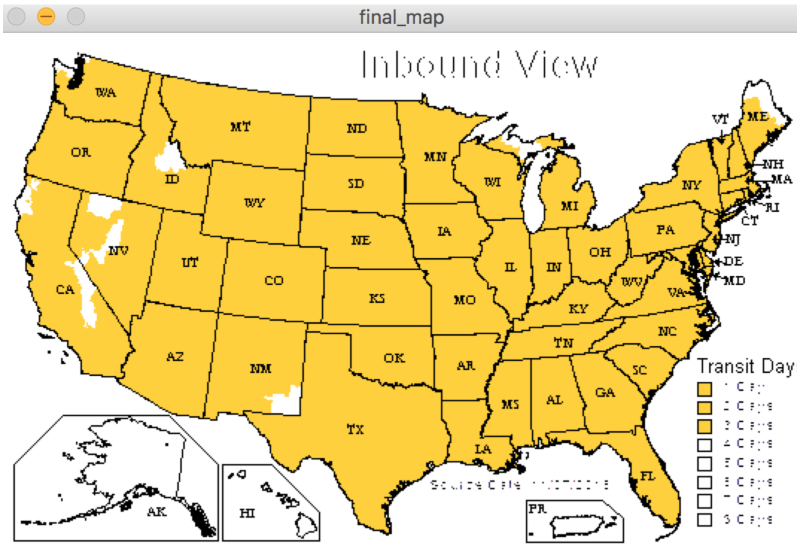

Figure 16: final map of warehouses.

3.7 Results and Discussion

Our empirical analysis as presented above yields a map with 41 zip-codes as follows: 03060, 95822, 92101, 97846, 99204, 97470, 32806, 59301, 59711, 89883, 97470, 59101, 48210, 14850, 16801, 79935, 85029, 35216, 11231, 89169, 23224, 41653, 37660, 29412, 78712, 75501, 70767, 82601, 47909, 65201, 87508, 73160, 72002, 67601, 80219, 55416, 58439, 57469, 51106, 69201, 69301 (see appendix B for a map of these zip codes).

The difference between the zip-codes produced by our Part 1 optimization and the zip-codes produced by this Part 2 optimization are profound. The most obvious shift that emerges when comparing the two is a gravitation towards state borders. Whereas the locations determined in Part 1 were largely in the center of states, Part 2 issued zip-codes that were far closer to state borders, such as Sioux City, Iowa (97846) and West Wendover, Nevada (89883). This is logical considering the harsher tax rates in Iowa (6%) and Nevada (6.85%); to minimize the in-state tax liabilities, our model naturally shifted warehouses towards borders.

4 Clothing Tax Part III

4.1 Assumptions and Justifications

1. **Assumption:** The proportion of clothing inventory bought from a warehouse to the total inventory is the same at all warehouses.

   Justification: It is impossible to know the variation between different warehouses. In any case, there will be enough people buying from any warehouse that individual variation is unimportant. The effect caused by this variation is negligible when compared to the actual tax effect itself

2. **Assumption:** Each warehouse sells clothing.

   Justification: Indeed, merely adding clothing and apparel to the line of products doesn't ensure that clothing and apparel are stored in each warehouse; some warehouses may be more suited to hold clothing. However, for the sake of simplification, and considering that we want the minimum number of warehouses to service the entire continental United States, we cannot allow two warehouses to service one location. Thus, our assumption that each warehouse sells clothing holds.

4.2 The Effects of Clothing Sales

The addition of clothing to Math Modell's product line has profound impacts on the tax liabilities imposed upon customers. Considering that several states do not tax clothing or have minimal taxes on clothing, the model produced in Parts 1 and 2 would likely not result in optimal solutions necessary to reduce costs for Math Modell's and ensure the least tax liability.

For example, states such as Minnesota and New Jersey, which have 6.88% and 7% sales taxes rates respectively, have limited tax on clothing and shoes, while states like Virginia with a 5.3% sales tax rate taxes clothing rather harshly. The net effect? Our mathematical model from 1 and 2 shifted warehouses towards state borders (See Section 3) and away from states like MN and NJ with higher tax rates; in optimizing merely tax liabilities for recreational goods, the model would be suboptimal when accounting for clothing sales.

Indeed, MN and NJ have lower clothing sales taxes, and depending on the proportion of Math Modell's sales of clothing compared to recreational goods, an optimal solution may involve placing warehouses back in MN and NJ. While our math model from Part 2 also placed warehouses in Virginia, our model from Part 3 may not, considering that Virginia doesn't reduce taxes on clothing and shoes.

In the later parts of this section, we attempt to build upon the models produced in Sections 1 and 2, by slightly altering our total tax liability equation. The incorporation of adjusted tax values for clothing and shoes only strengthens the relevance of our model.

4.3 Incorporating Clothing

Given that certain states have reduced or no taxes on clothing and shoes, Math Modell's expansion into clothing sales will have significant repercussions on our model from Part 2; the model needs to be adjusted for optimal results.

The model here is analogous to the model created in section 3.4; the only difference is the rate at which we tax. Note, however, that since Math Modell's is simply adding apparel to it's product line, we must now differentiate between clothing, which enjoy at time decreased taxes, and recreational goods, which are taxed at the normal sales tax rate.

Since clothing is just a category of item in an inventory, we can break down the inventory into two parts, one which is only clothing, and one which is everything else. Let p be proportion of inventory sold that is clothing. Naturally, this ensures that for every p items in inventory that are clothing, there are $(1 - p)$ items in inventory that are recreational goods (not clothing).

Since inventory that is clothing and is not clothing is taxed differently, we multiply the proportions by their individual tax rates and sum them to obtain the average tax for any item in the inventory. Then if $t_{average}$ is the average tax for any item, we have the equation below:

$$t_{average} = t_{clothing} * p + t_{no-clothing} * (1 - p)$$

While in 3.4, we originally used a simple t_{state} to represent tax rates, we must adjust our original tax liabilities to incorporate this new average taxation. We now plug our average taxation into the original equation from 3.4 involving total tax liability from part 3.4 to obtain:

$$\begin{aligned} T_w &= t_{average} * D_{pop} * A_{taxed} \\ &= t_{average} * D_{pop} * \int_0^{2\pi} \frac{(r_{ship})^2}{2} - \frac{(r_{ship} - d_{border})^2}{2} d\theta \end{aligned}$$

where T_w is the total tax liability imposed by the warehouse. We can now obtain the total tax liability of all warehouse across the continental United States in the same manner as 3.4:

$$\sum_{warehouses} T_w = \sum_{warehouses} t_{average} * D_{pop} * \int_0^{2\pi} \frac{(r_{ship})^2}{2} - \frac{(r_{ship} - d_{border})^2}{2} d\theta$$

This final equation represents the total tax liability, accounting for the reduced clothing tax. The remainder of this section attempts to optimize three major variables, each introduced in part 1, 2 and 3: warehouse count, land area serviced by warehouses, tax liabilities for non-clothing goods, and tax liabilities for clothing goods.

4.4 Results and Discussion

Based on our new model, we must reevaluate the cities where we put our warehouses. The general guideline is that if the total covered area is unchanged and the number of warehouse stays the same, we should try to move warehouses into states that has no taxation on clothes. This doesn't lead to a big change in our warehouse locations. However, after we factored minimizing apparel tax into account, the warehouse that can cover most areas isn't necessarily the most efficient anymore. So we added one more warehouse to make sure that we cover the US. Overall, we have made 5 changes to our warehouse locations in part II: 11231 to 02462, 37660 to 47802, 29412 to 28205, 47909 to 60628 and we add one more warehouse at 28803 (See Appendix C for our final map).

5 Strength and Weakness Analysis

There are numerous strengths that augment our model, and multiple weaknesses that detract from it. Our analysis is below:

5.1 Strengths

1. Our modeling minimizes error in the optimization process. Generally there are two ways to optimize something: Creating a complicated math model that produces a good result right off the bat, or infinite guessing and checking. Neither one of those works out. It's impossible to create a model that optimizes this problem because there are so many variables. Guessing and checking won't work either simply because there are so many cases. But our model first guesses a general region by our mid-point strategy; it then uses a pretty simple math model to figure out a couple candidate cities in that area, and lastly we use the efficient programs we write to check which one of those cities works best. Our model combines the advantages of the two different approaches and thus can do better than any of those.

2. We incorporate a lot of graphs and pictures with our paper. They help emphasize the essentially idea of our model: tackle the problem and approach optimization step by step.

3. We use real-life data to verify our model. We use the case of Seattle versus Ellensburg and Ellensburg versus Spokane to verify the correctness of our model, and the results are very good.

5.2 Weakness

1. We didn't factor in the cost of building warehouses. In reality, warehouse location and warehouse costs can play a major role in warehouse placement. In our model, we choose warehouse locations by zip codes, and zip codes don't

always associate with a cost. We thus failed to factor in the cost of building warehouses.

2. Our paper fails to account for human tastes and preferences. Indeed, Math Modell's Sporting Goods certainly has a specific target market of individuals with similar characteristics, and the store perhaps caters toward more people in a certain part of the country. Perhaps more customers of the store exist in Massachusetts than in Wyoming, but our warehouse placement remains unadjusted. Our model thus fails to account for geographical differences in target market, which would greatly affect warehouse placement.

A Appendix A

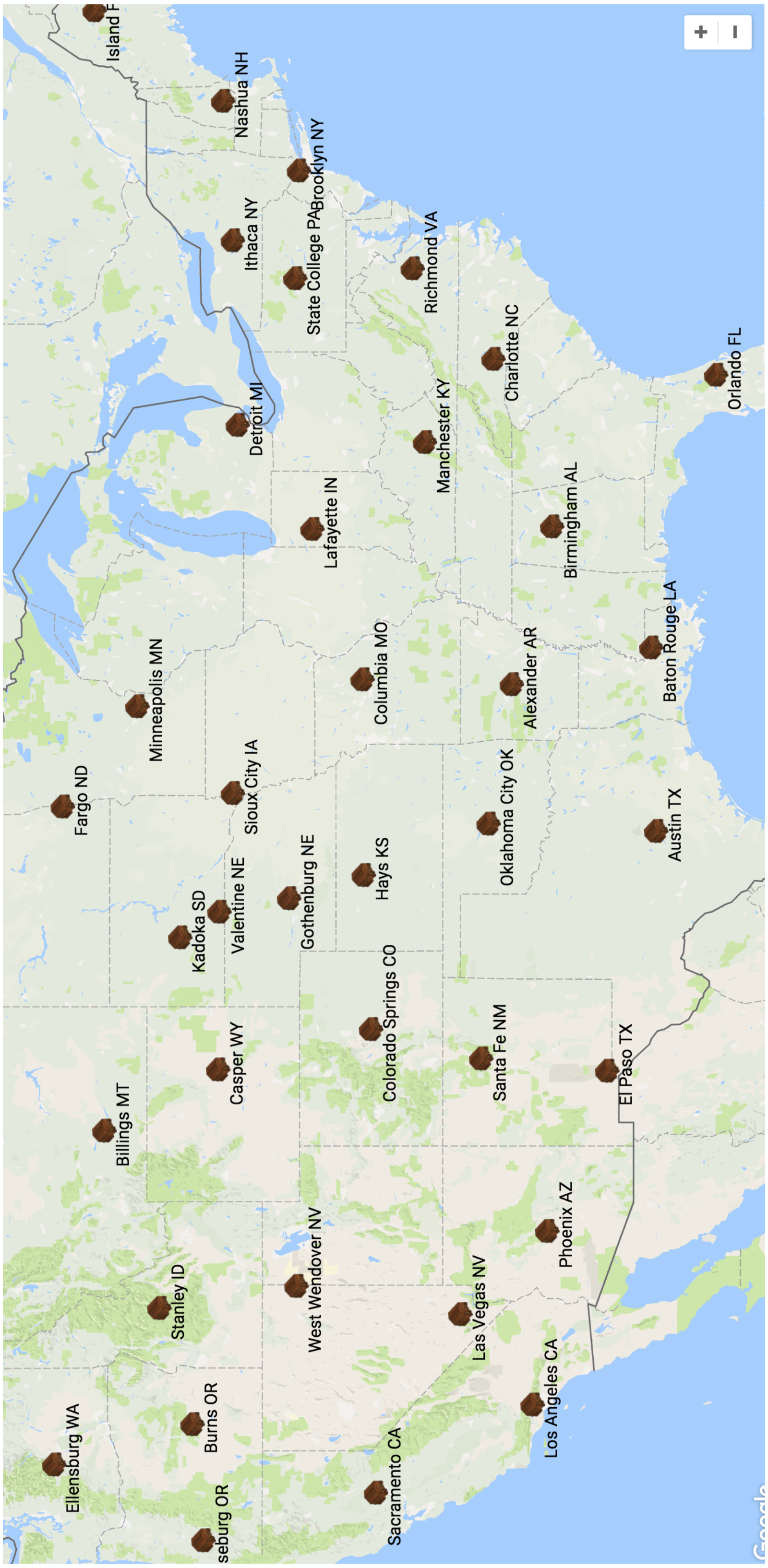

Figure 17: Location of warehouses, part I.

B Appendix B

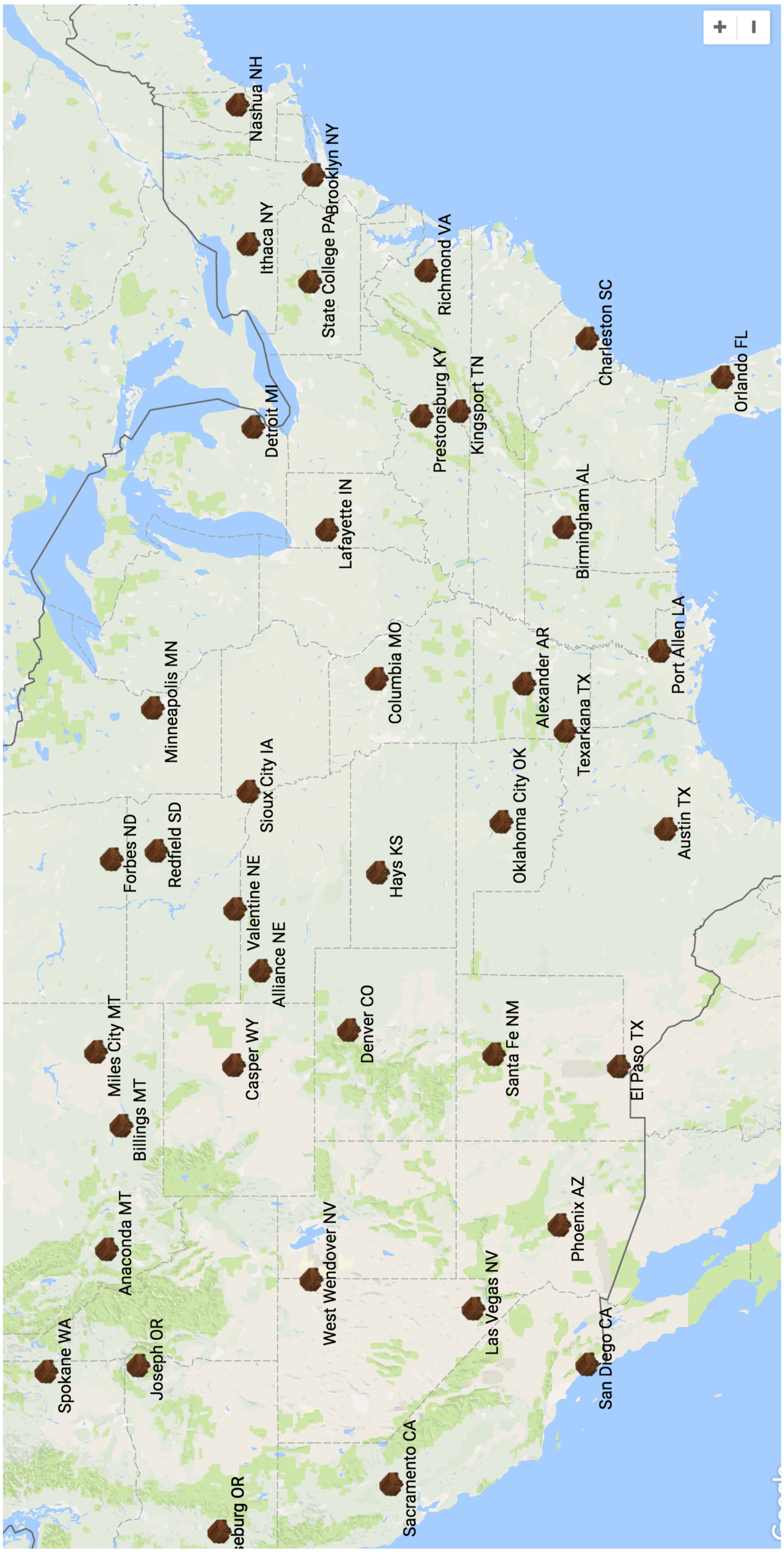

Figure 18: Location of warehouses, part II.

C Appendix C

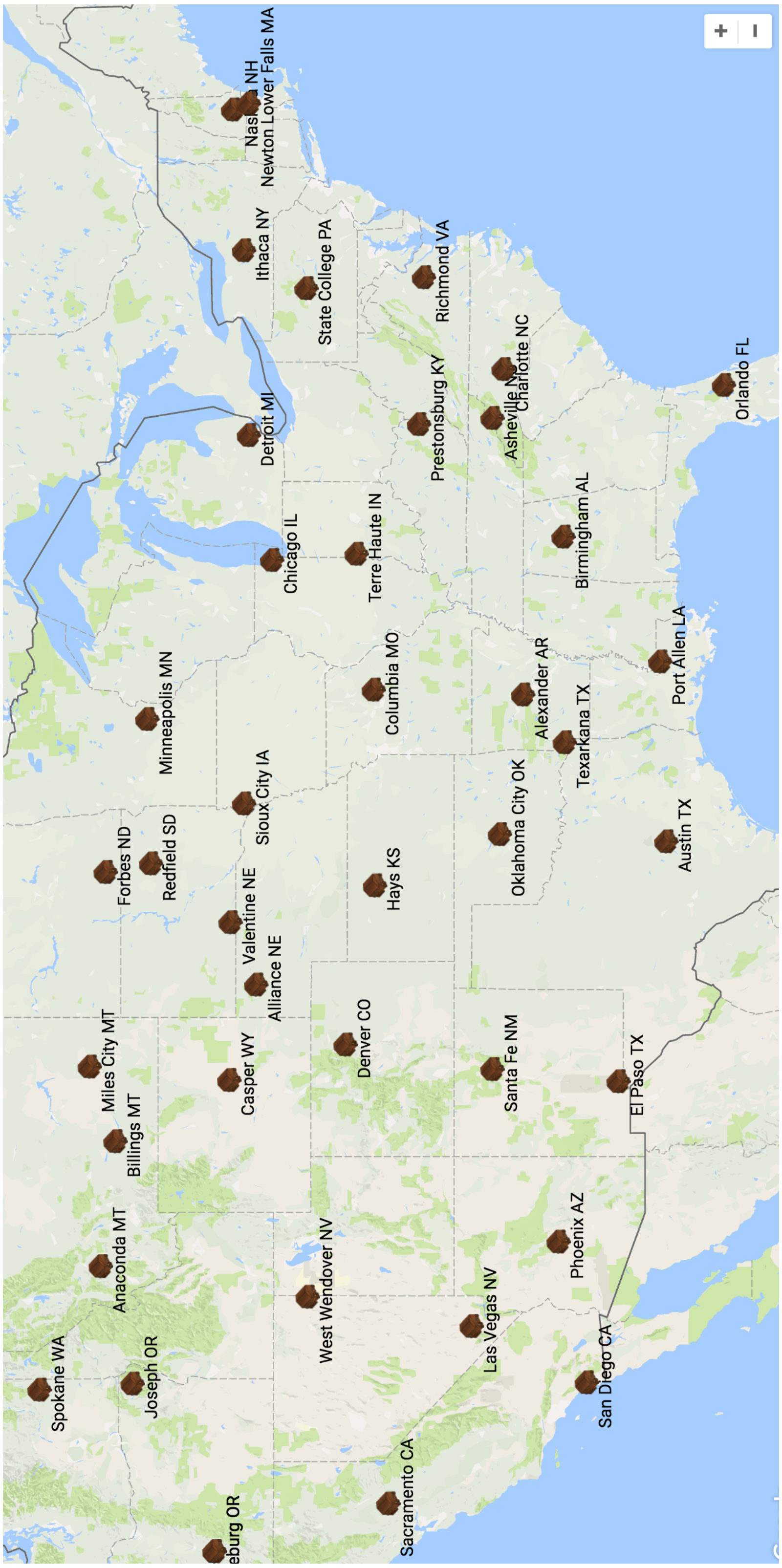

Figure 19: Location of warehouses, part III.

References

[1] Bittner, Jon. "Free US Population Density And Unemployment Rate By Zip Code." The Splitwise Blog. N.p., 06 Jan. 2014. Web. 19 Nov. 2016.

[2] Bogaisky, Jeremy. "Retail In Crisis These Are The Changes Brick and Mortar Stores Must Make." Forbes Magazine. 12 Feb. 2014. Web. 19 Nov. 2016.

[3] Lindner, Matt. "Online sales will reach $523 billion by 2020 in the U.S." Internet Retailer. 29 Jan. 2016. Web. 19 Nov. 2016.

[4] E-Commerce. "83% Of Shoppers Are Satisfied With Their Online Shopping Experiences." Retail TouchPoints. 11 Aug. 2014. Web. 19 Nov. 2016.

[5] Kubly, Scott. "Pavement Management." Department of Transportation. Seattle.gov 2010. Web. 19 Nov. 2016

[6] "Greenwater, Washington Visitor Information." Go Northwest! A Travel Guide. Web. 19 Nov. 2016

[7] Sweeney, Deborah. "5 Ways to Improve Your Company's Tax Liability." Business Insider. 2 Dec. 2010. Web. 19 Nov. 2016

[8] "Population Density and Occurrence of Accidents in Finland." National Center for Biotechnology Information. U.S. National Library of Medicine, n.d. Web. 19 Nov. 2016.