



Using Spatial Tools to Estimate Freight Costs



John Thompson
Sr Transportation Analyst
Cargill Transportation & Logistics

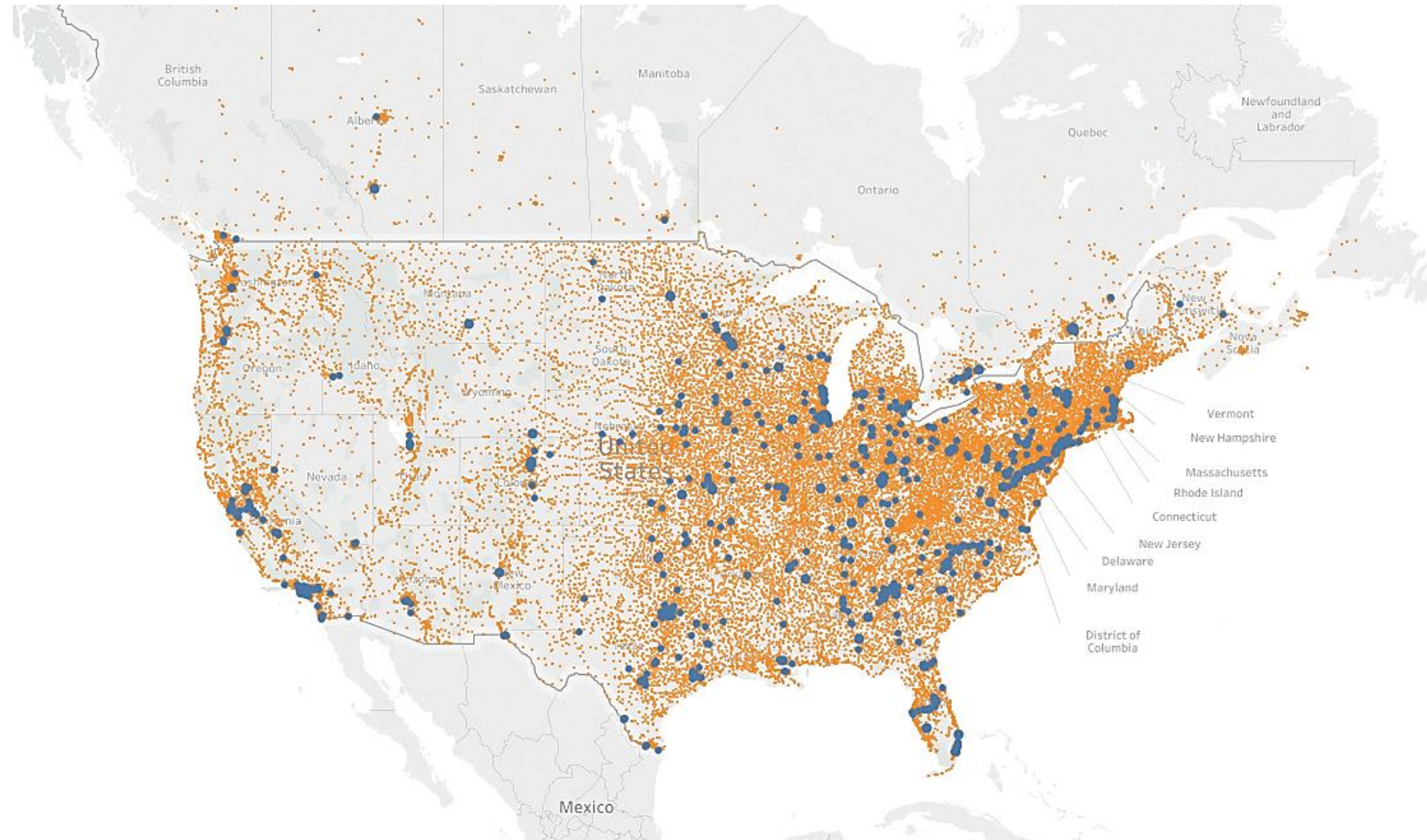
What's The Problem?

IDENTIFY FREIGHT COSTS TO AREAS WE HAVE NO HISTORIC DATA

There are over 32,000 postal codes in the US and Canada

In a typical year, Cargill ships to 220 of them per origin on average.

How can we estimate freight rates to postal codes with no historic data?



How is Full Truck Load Freight Priced?

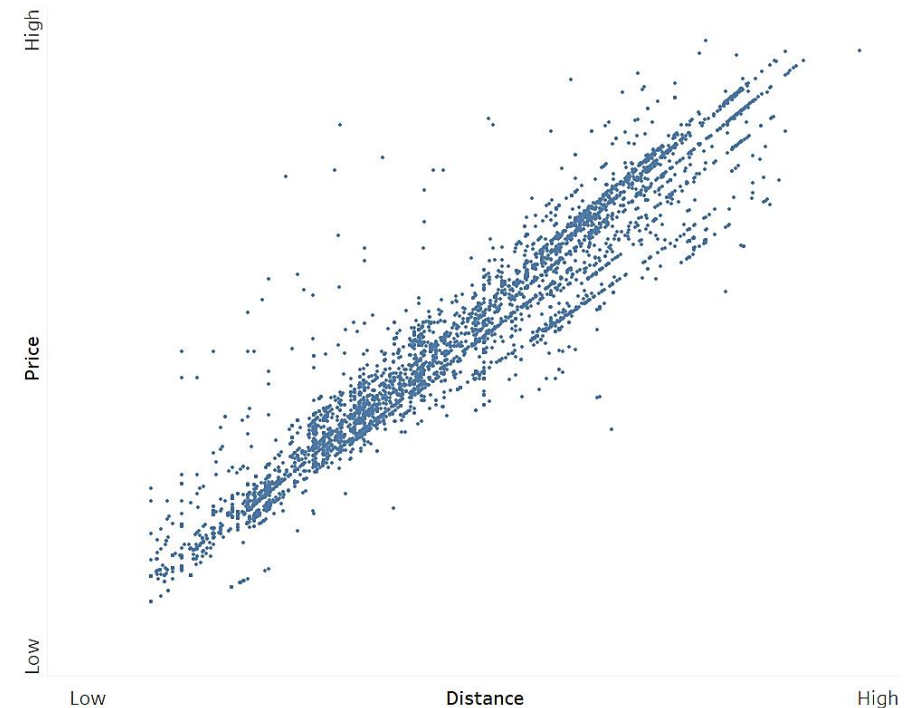
Predominantly a function of distance and geography

According to freight market intelligence firm Chainalytics:

≈ 80% of freight cost is explained by distance

≈ 8% is explained by geographic influence

Other factors combine for an additional 4-5% of the cost.



Possible Methods

ORDINARY LEAST SQUARES REGRESSION

Pros: Multiple variables, can fine tune fit, adapt to changes in market

Cons: Need lots of data, many variables, can get complicated

MILEAGE SCALES

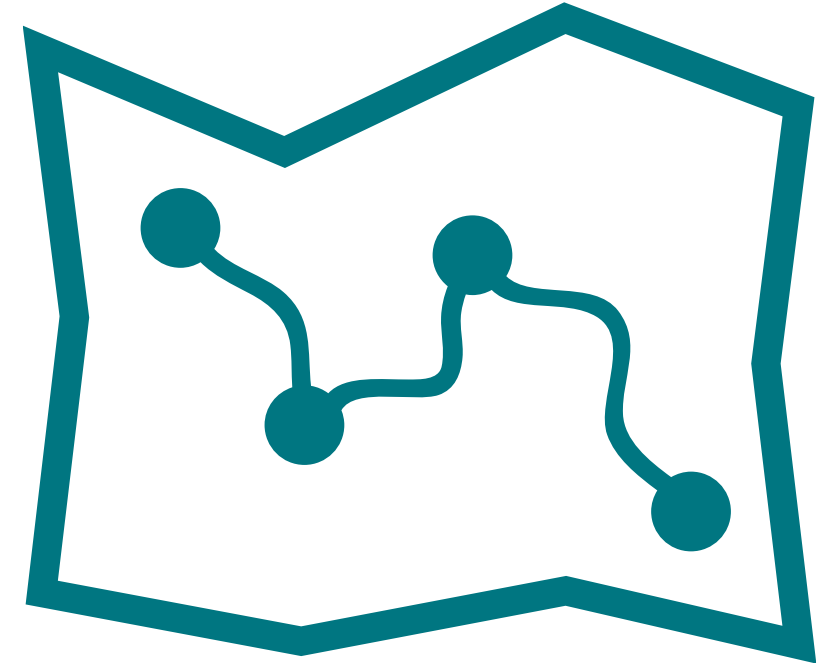
Pros: Estimates costs to a precise destination, easily computed and modeled

Cons: Need to know driving mileage for every origin/destination pair, doesn't take into account geographic influence

PROXIMITY

Pros: Captures distance and geography, fewer fields required, fast and easy to understand

Cons: May not have nearby data for every destination, less precise than other techniques



Proximity Method

Inverse Distance Weighting Interpolation

Estimate cost using the 3 nearest historic costs

The closer a point is, the more influence it has

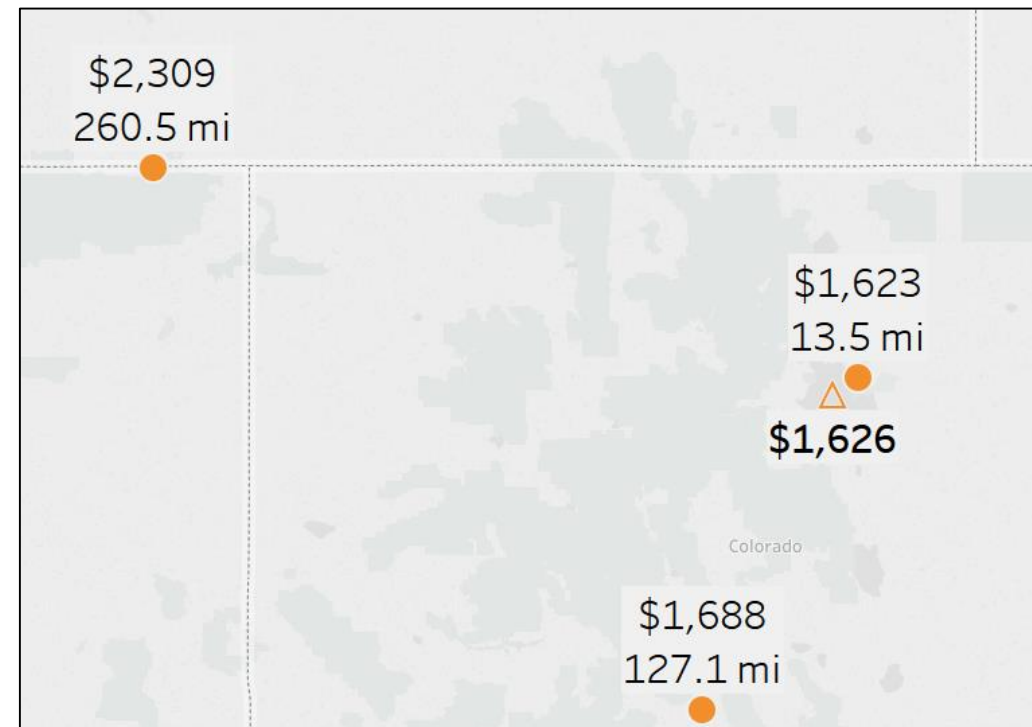
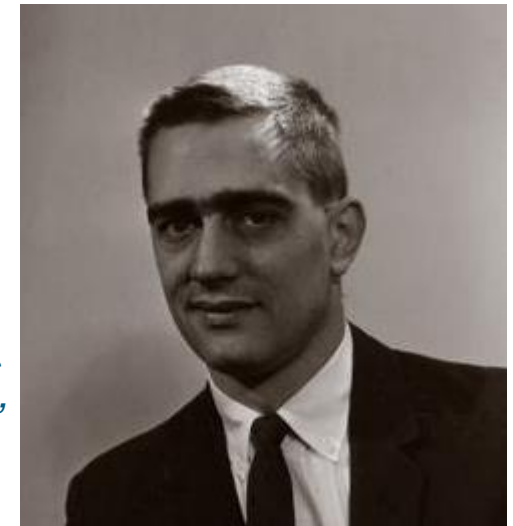
This method eliminates the need to factor in geographical and distance figures since those factors are inherently part of the spatial structure

Example to right:
Average of nearest 3 points: \$1,873
IDWI result: \$1,626

https://en.wikipedia.org/wiki/Inverse_distance_weighting#Shepard's_method

“Everything is related to everything else, but near things are more related than distant things”

*-Waldo R Tobler’s
“First Law of Geography”*



Using the Spatial Tools in Alteryx



Create Points

Take Latitude and Longitude and create a spatial object (centroid)



Find Nearest

Input spatial objects in two groups, Target and Universe.

In this case, the targets are points with unknown freight costs and the universe is all points from our historic data



Heat Map

Generate spatial polygons from individual records grouped by a value.

Render for distribution using the Reporting tools

Configuration – Create Points Tool



Need to have **Latitude** and **Longitude** fields in your data

Identify your fields and the format for your Latitude and Longitude

The screenshot shows the configuration window for the 'Create Points (24)' tool. The title bar is blue with the text 'Create Points (24) - Configuration' and standard window controls. The main area is light gray. On the left, there is a vertical toolbar with icons for settings, navigation, and help. The configuration area contains two dropdown menus: 'X Field (Longitude):' with 'Dest Long' selected, and 'Y Field (Latitude):' with 'Dest Lat' selected. Below these are three radio button options: 'Fields are Lat/Long Floating Point' (selected), 'Fields are Lat/Long Integers (x 1,000,000)', and 'Fields are Projected Floating Points'. At the bottom, there is a text field containing '[WGS84]' and a small gray button with three dots.

Configuration – Find Nearest Tool

Connect Target data set to the “T” input

Connect Universe data set to the “U” Input

Configure your **Targets** by selecting the spatial object you wish to use as your target data

Configure **Universe** by either selecting the spatial object from a table connected to the “U” input, or from a file.

Select the **number of nearest points** to match, **maximum distance**, and if you want to **ignore 0 distance** matches

There is also a “**Select Tool**” dialogue to trim down or rename your field list

Your data will output in two areas, the “**M**” **output** will be your matched targets and corresponding universe values. The “**U**” **output** contains unmatched target records.



Find Nearest (9) - Configuration

Targets (T Input)

Spatial Object Field: Centroid

Universe

Use Records from U Input

Use Records from File or Database:

Spatial Object Field: Centroid

How many nearest points to find? 3

Maximum Distance: 3000 Miles

Dataset: No valid datasets f

Ignore 0 Distance Matches

Configuration – Heat Map Tool

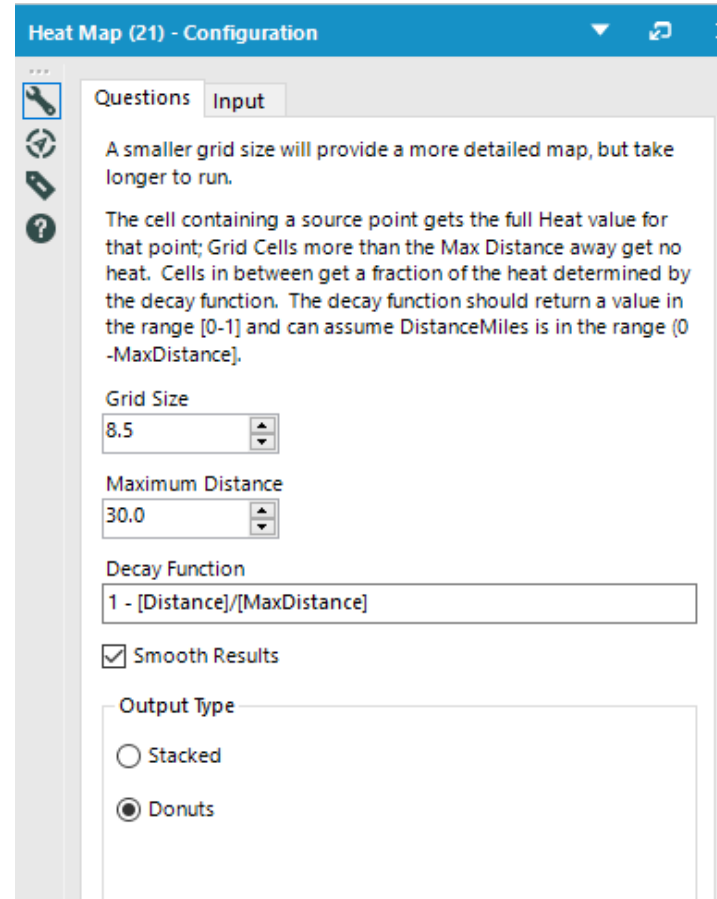
Grid Size sets the smallest size of an area to consider when generating a heat value

Maximum Distance is how far the heat from one grid cell can influence the heat of nearby grid cells. This sets a cutoff on that distance.

Decay Function sets the rate of change from one grid cell to another within the maximum distance. Controls gradient.

Output type has two options, stacked or donuts. These determine how polygons are stacked on one another.

On the **Input Tab** you choose your spatial object field and the field for generating heat.



Heat Map (21) - Configuration

Questions Input

A smaller grid size will provide a more detailed map, but take longer to run.

The cell containing a source point gets the full Heat value for that point; Grid Cells more than the Max Distance away get no heat. Cells in between get a fraction of the heat determined by the decay function. The decay function should return a value in the range [0-1] and can assume DistanceMiles is in the range (0 -MaxDistance].

Grid Size
8.5

Maximum Distance
30.0

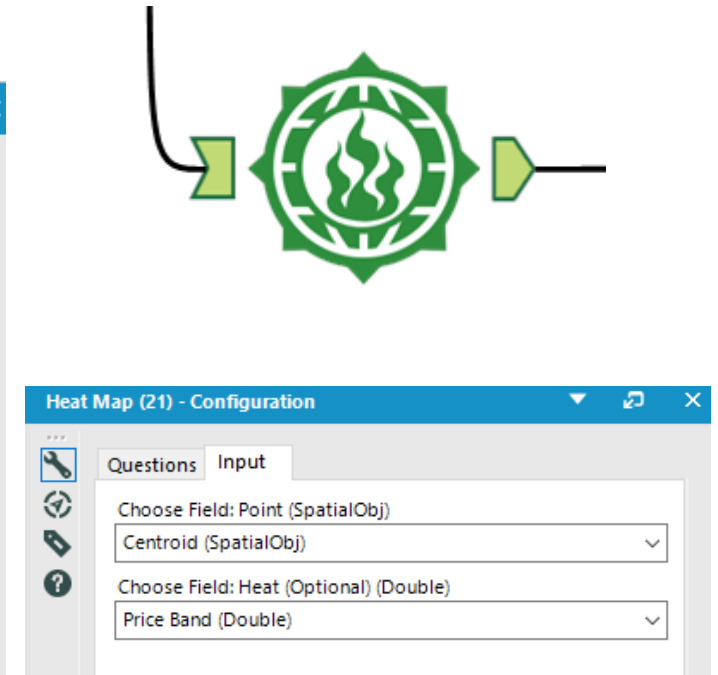
Decay Function
 $1 - [Distance]/[MaxDistance]$

Smooth Results

Output Type

Stacked

Donuts



Workflow Demonstration

Thanks!



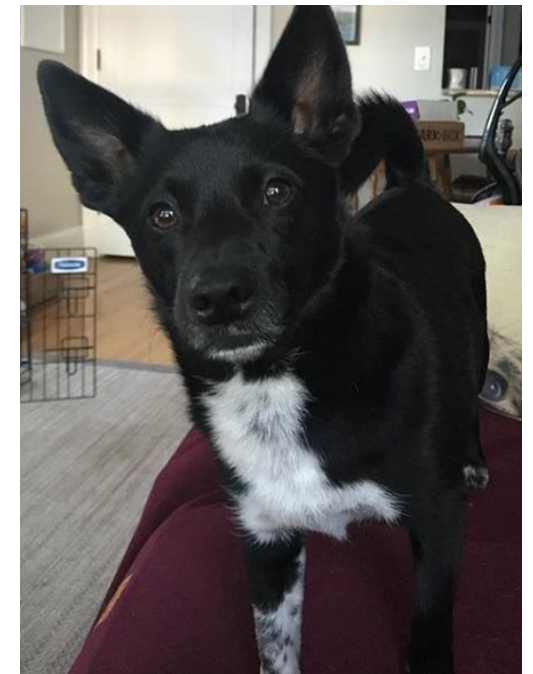
www.linkedin.com/in/johnmthomp/

[+tableau++public public.tableau.com/profile/johnthompson](https://public.tableau.com/profile/johnthompson)

alteryx | COMMUNITY jmt214



@thedognamedfinn



The Cargill logo features a green, curved leaf-like shape positioned above the word "Cargill". The word "Cargill" is written in a bold, italicized, black sans-serif font. A registered trademark symbol (®) is located at the end of the word.

Cargill®