

# High Level, Intuitive Look

*Daniel Truver*

*5/19/2018*

## **Problem**

We want to suggest pricing for parking spots in the New York City leased in the short term through a mobile application. It's not rocket surgery here, people, so let's keep it simple. What are the two most important traits of a parking spot? Its location and the time of parking. This is my first idea for a model addressing the spatial aspect of the problem.

## **Preliminary Thoughts (Smoothing Splines)**

Special thanks to Colin Rundel's lecture Spatio-Temporal Modeling (STA-644, Duke University).

It stands to reason that parking spots close together in space are of similar value. If you have just thought of a few anecdotes to disprove that claim, please know that I will ignore you unless you have a larger set of data with which to argue.

We will also assume isotropy in this model. That is, the distance between parking spots is of the same significance in all directions. This is a poor assumption since points in a city are not connected by straight lines, rather streets. We can avoid this problem by changing our metric from Euclidian (distance is the straight line between two points) to Manhattan (distance is the sum of street lengths leading from point A to point B). I will suggest we just accept the problem and stick to Euclidean distance as Manhattan distance is not unique. In further iterations where my pay, and by extension, motivation, are higher we can adjust the model to account for different metrics. For the purpose of this high-level review, just assume we have settled on a metric; it does not change the model significantly.

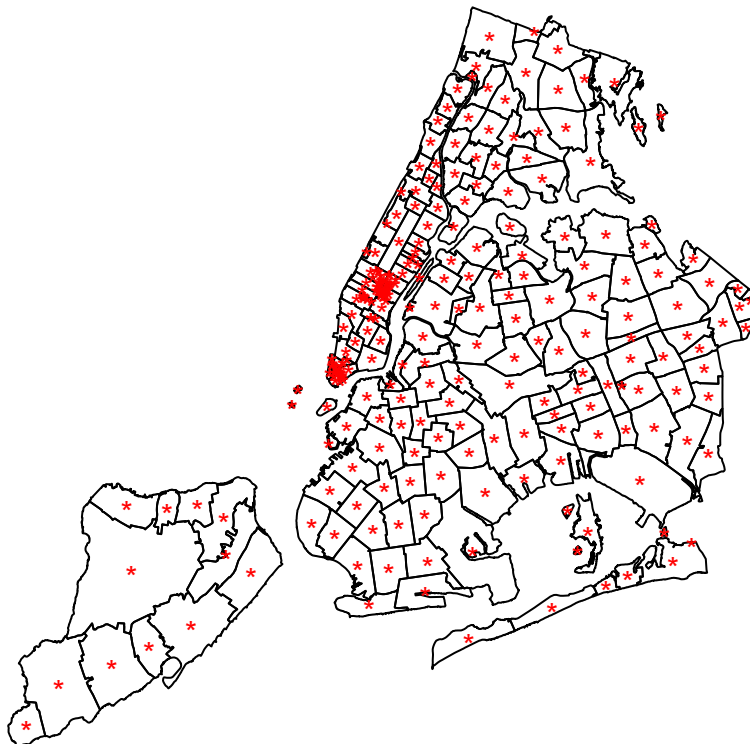
I believe we will have point reference data. If that sentence put you to sleep, don't worry, me too. Essentially, we have a shell that is New York City.

## **New York City Divided by Zip Code**



Within this shell, we will have parking spots with a certain price.

## New York City Divided by Zip Code (centroids as example spots)



As discussed above, we believe the value of parking spots close together should be similar. We will represent this mathematically by saying that the value of a spot at any location in the city is described by a continuous function  $f(x, y)$  where  $x$  and  $y$  are latitude and longitude. Later, once the additional motivation appears, we can add predictors based on expert knowledge and accumulated data. If that last sentence sounded like a bunch of jargon that people really up their own asses would say at a tech conference, it is. Feel free to use it with such people. What it really means is:

Let's say you know that the Whitney is a popular destination on Fridays or that some celebrity who I don't give a shit about is doing an impromptu event at the Supreme store (side thought, Supreme branded parking spots; we can all retire now). Anyway, after contemplating how nice it is that people enjoy art or deepening your contempt for all mankind, we could add a predictor to the model at the Supreme store location or the Whitney that elevates the mean price of parking around that area at the time. This aspect will require observation over time and is near impossible to have at launch. Let's return to the standard model.

We will fit a smoothing spline (will simulate and fit later when motivation is higher) to the data. The smoothing spline for our case, where  $d$  is our metric and  $x_i, y_i, z_i$  are the latitude, longitude, and value of parking spot labeled  $i$  is given by the solution of

$$\arg \min_{f(x,y)} \sum_{i=1}^n (z_i - f(x_i, y_i))^2 + \lambda \int \int \left( \frac{\partial^2 f}{\partial x^2} + 2 \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} \right) dx dy$$

which should have the representation

$$f(x, y) = \sum_{i=1}^n w_i d(x_i, y_i)^2 \log d(x_i, y_i)$$

where  $w_i$  are weights for each spot. You should have software to solve this. If not, you're screwed.

Once the smoothing spline is fit to known spot prices, you will have predictions for all spots on the map. This model is simple but data hungry. Additional predictors will come from people who are experts in the area of parking demand, if we can find such people, or if they existed in the first place.

## **Preliminary Thoughts 2 (Gaussian Process)**

To be continued...