

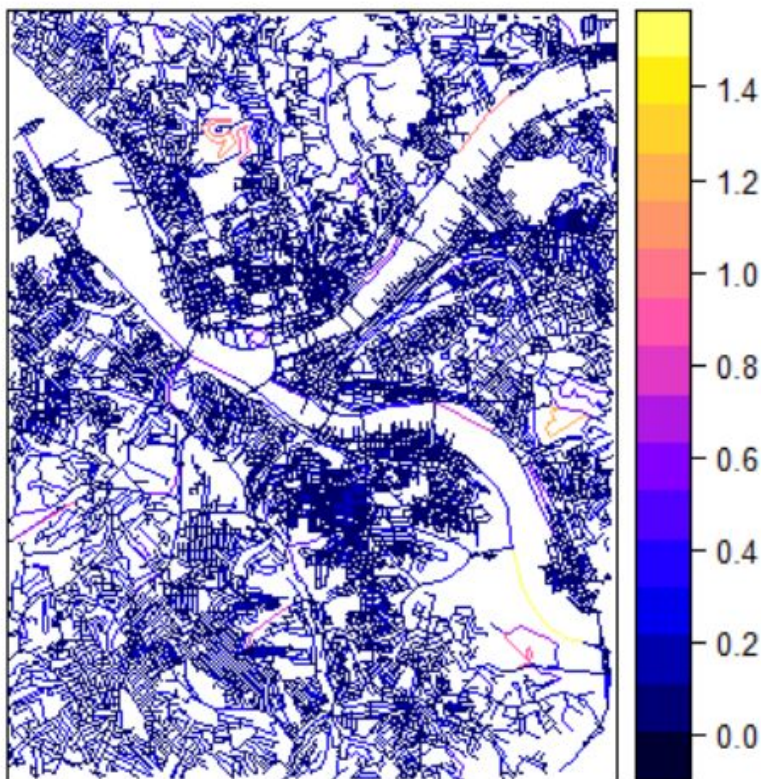
*(also attached is the Rscript)*

### **Spatial Data Manipulation**

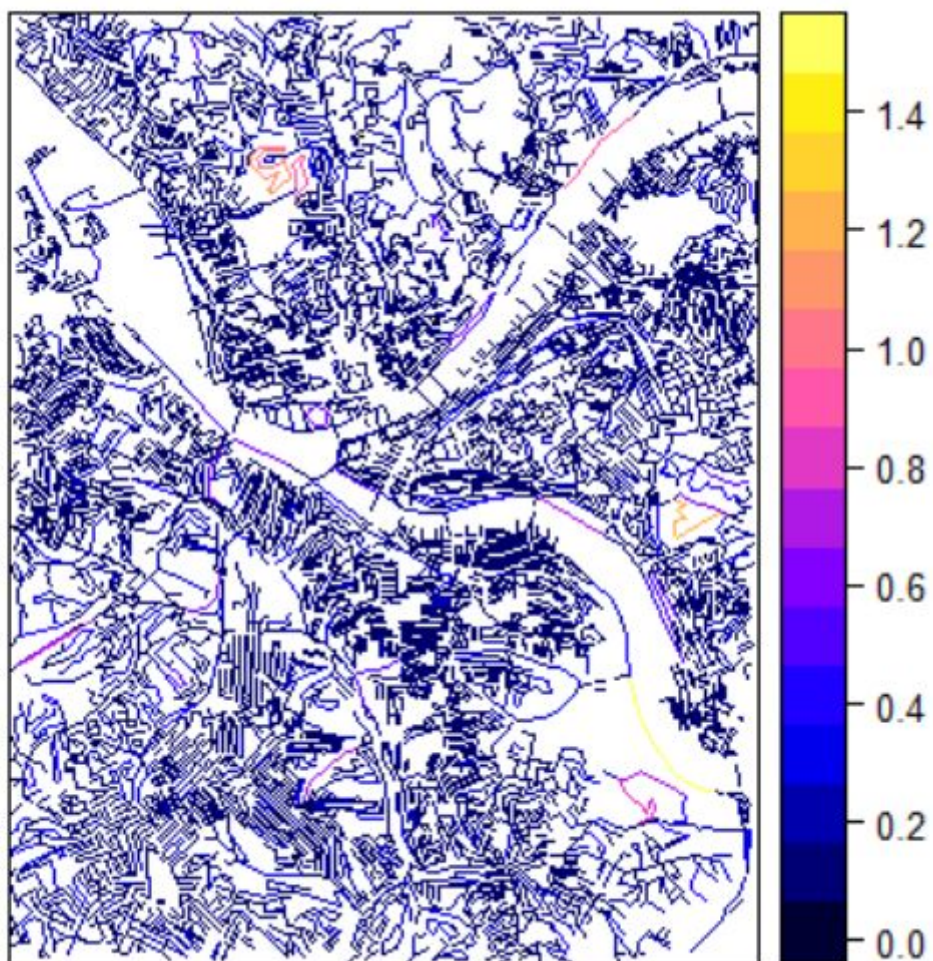
a. There are a total of 22,222 road segments in the shape file

b. Mean length= 0.05980, Max length= 1.46654, Median length=0.04275

c. Map for all the streets:



Map for streets having length greater than mean length:

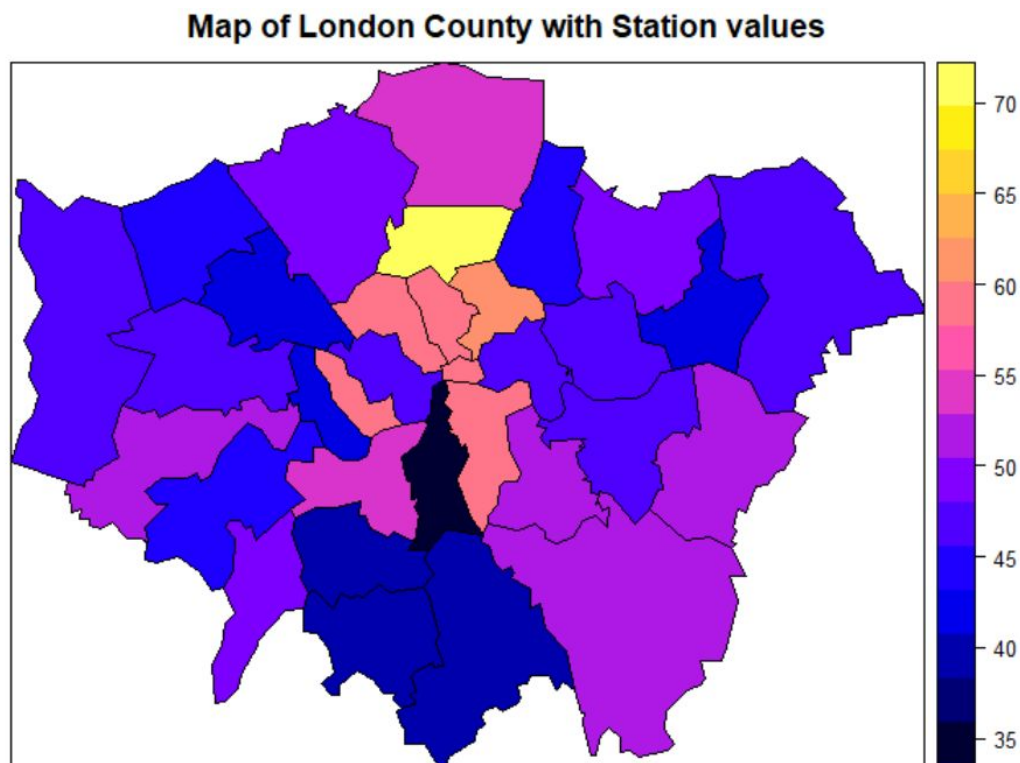


## Spatial Data Aggregation

	stationsagg.Value	lnd.name
1	51.22460	Bromley
2	45.08194	Richmond upon Thames
3	45.70525	Hillingdon
4	46.72895	Havering
5	48.53035	Kingston upon Thames
6	38.35842	Sutton
7	50.59463	Hounslow
8	38.34868	Merton
9	54.84487	Wandsworth
10	40.32173	Croydon
11	35.77615	Lambeth
12	58.68201	Southwark
13	52.50246	Lewisham
14	47.05888	Greenwich
15	45.84224	Ealing
16	41.26989	Hammersmith and Fulham
17	41.90940	Brent
18	44.58654	Harrow
19	48.12921	Barnet
20	58.17340	Islington
21	60.81719	Hackney
22	45.68979	Newham
23	41.49116	Barking and Dagenham
24	69.78397	Haringey
25	54.80257	Enfield
26	45.01600	Waltham Forest
27	49.25711	Redbridge
28	51.25654	Bexley
29	58.25932	Kensington and Chelsea
30	47.84540	Westminster
31	57.97391	Camden
32	46.70356	Tower Hamlets
33	58.21482	City of London

**Report:**Above is the mapping of the London county with its respective aggregated values and below is a map describing the same relation where each color corresponds to a different value pointed by the map legend. County of Haringey has the highest aggregated value whereas county of Merton has the lowest value. Majority of the counties have values ranging from 35-45.

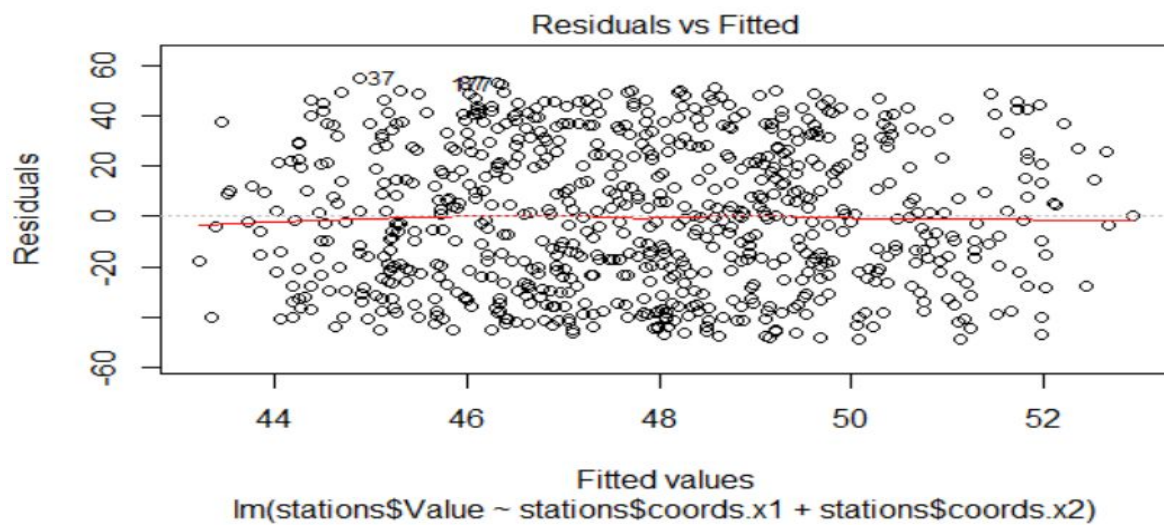




Regression before aggregation:

Coefficients:

(Intercept)	stations\$coords.x1	stations\$coords.x2
-4.419e+01	1.033e-04	2.075e-04

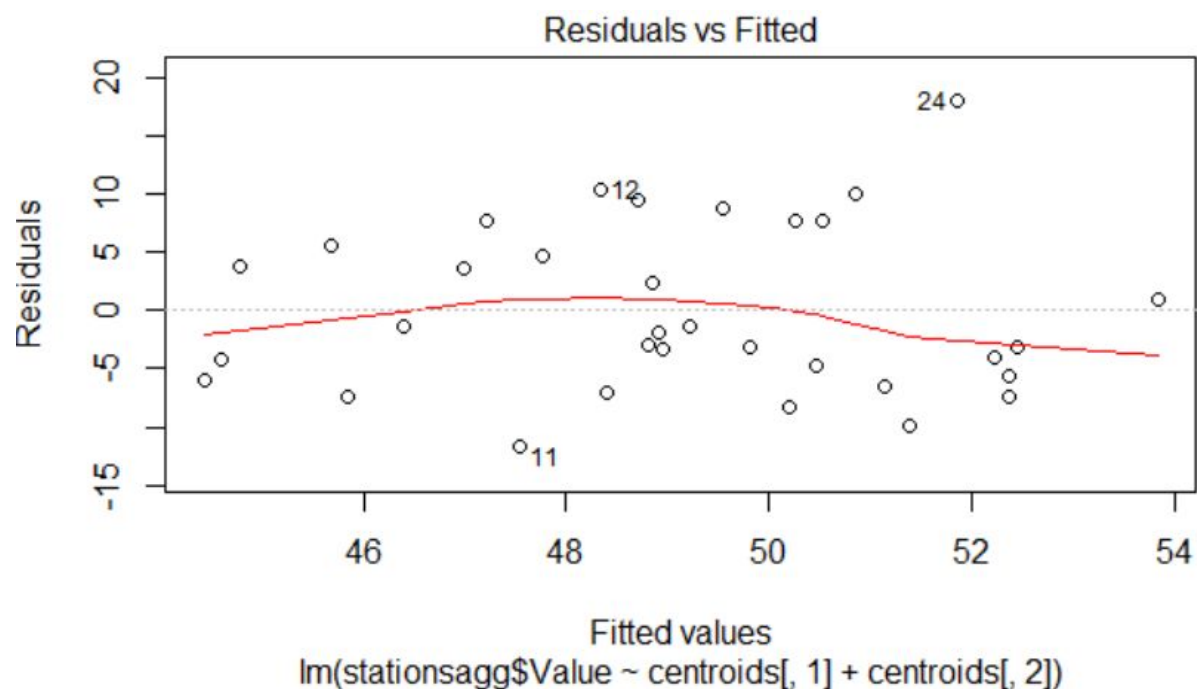


## Value vs X1,X2

Regression after aggregation:

Coefficients:

(Intercept)	stations\$coords.x1	stations\$coords.x2
-4.419e+01	1.033e-04	2.075e-04



## Aggregation Value vs Centroid

Two graphs above are showcasing The Modifiable Areal Unit Problem where result of data aggregation depends on the type of unit we choose. Each results into a different line.

R Script :

```
setwd("C:/Users/Tigmanshu/Documents/pgh_streets/pgh_streets/pgh_streets")
library(rgdal)
s1 <- readOGR(".", "pgh_streets")
summary(s1)
s2 <- s1[s1$LENGTH > 0.05980,]
spplot(s2, z = "LENGTH")

setwd("C:/Users/Tigmanshu/Documents/Spatial Data Analytics")
load("Ind.Rdata")
load("stations.RData")
```

```
summary(lnd)
summary(stations)
data.frame(stations)
stationsagg<-aggregate(stations["Value"],by=lnd, FUN = mean)
print(stationsagg$Value)
spplot(stationsagg,z="Value",main=list(label="Map of London County with Station values"))
relation<-lm(stations$Value~stations$coords.x1+stations$coords.x2)
print(relation)
plot(relation)
centroids <- getSpPPolygonsLabptSlots(lnd)
relation<-lm(stationsagg$Value~centroids[,1]+centroids[,2])
print(relation)
plot(relation)
```