Mapping ecometrics

Hector Antonio Vazquez Brust

November 6, 2015

Mapping ecometrics

Having decided that we want to measure urban segregation in Boston, we are looking for clues that we can actually observe in oour dataset.

To understand the spatial arrangement of different (and maybe unequal) characteristics of the city's built environment, we will select a variable to put on the map.

Working the Tax Assessor's dataset, building value is one the varaibles that usually come to minf when choosing what to compare. We have already created a new variable that measures building value normalized by building size ("AV_BLDG_PER_SF"). We will now create and additional variable, to rank from 1 to 10 the comparative value of a particular building compared to the entirety of Boston's building stock. For the statistically minded, what we'll do is calculate the decile every building on our dataset belongs to.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##     filter, lag
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
#Extract a subset from our columns (Only the variables we are interested in)
parcels <- select(TAdata, AV_LAND, AV_BLDG, AV_TOTAL, GROSS_TAX, LAND_SF, YR_BUILT, YR_REMOD, CT_ID_10,

#Add a column with a 1 to 10 ranking (decile) for the parcel's sq foot value
parcels <- mutate(parcels,  BLDG_RANK = ceiling(rank(AV_BLDG_PER_SF,na.last="keep")/length(which(!is.na
```

So this is how the new decile/rank variable looks:

```
head(select(parcels, AV_LAND_PER_SF, BLDG_RANK ))
```

```
##   AV_LAND_PER_SF BLDG_RANK
## 1       80.95652         2
## 2       72.78261         3
## 3       72.86957         2
## 4       72.95652         3
## 5       46.26866         2
## 6       44.08000         1
```

Let's map our new variable over Boston:

```
require(ggplot2)
require(ggmap)
```

```
## Loading required package: ggmap
```

```
Boston <- get_map(location=c(left = -71.193799, bottom = 42.15, right = -70.985746, top = 42.5))
```

```
## Warning: bounding box given to google - spatial extent only approximate.
```

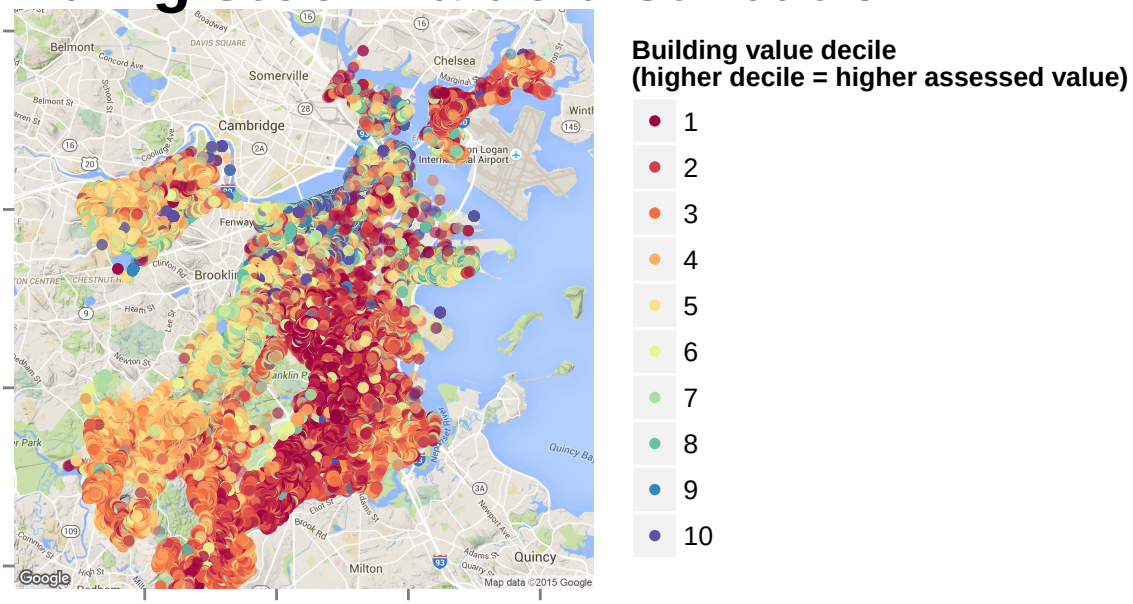
```
## converting bounding box to center/zoom specification. (experimental)
## Map from URL : http://maps.googleapis.com/maps/api/staticmap?center=42.325,-71.089773&zoom=12&size=6
```

```
base<- ggmap(Boston)

base + geom_point(data=parcels, aes(x = X, y = Y, color = as.factor(BLDG_RANK), alpha = .5)) + scale_co
(higher decile = higher assessed value)', alpha = FALSE) + guides(alpha = FALSE) +
  theme(plot.title = element_text(size = 24, face="bold"),
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.x = element_blank(),
        axis.text.y = element_blank())
```

```
## Warning: Removed 2064 rows containing missing values (geom_point).
```

building stock value distribution



This map confirms what the conlusions that we arrived to when we compared neighbourhood mean values: Back Bay/Beacon Hill as the most valued, Dorchester, Mattapan, Roxbury as the least. However, mapping the values does provide further insights: is remarkable how spatially concentrated the high value buildings are, and how close to each other the areas with the higher and lesser values are.

NOte to professor O'Brien:

I'm right now dealing with an implementation of the Shannon index. Once I have it figured out, I'll spatially join the Tax Assessor's dataset with the 2010 census data, and I'll compare built environment diversity with social diversity using maps

```
library(vegan)
```

```
## Loading required package: permute
## Loading required package: lattice
## This is vegan 2.3-1
```