THE ELDERLY LIVEABILITY INDEX

Hassan Bilal

DECEMBER 17, 2015

PROF DANIEL O'BRIEN

SPPUA 5262

Contents

I. Executive Summary

Eco-metrics deals with the measurements of the characteristics of the physical, and social ecology of a neighborhood. In order to learn more about these neighborhoods through eco-metrics, we develop models utilizing latent or conceptual constructs. However, to make inferences about such phenomenon we need to measure observable or manifest variables that drive the development of these conceptual models.

In this project a similar approach was replicated in the development of an elderly livability index. Categorized data pertaining to relevant factors within the Zoning Clearances data set was converted to numerical integers, in order to obtain a zoning score or index by category. The factors under consideration were identified after perusing a research paper published by the American Association of Retired Persons (AARP) titled, "What Is Livable? Community Preferences of Older Adults".

These indexes were later compared, and aggregated to districts and Neighborhood Statistical Area's with the hope of creating a set of values that would indicate desirability of a neighborhood for the elderly. Conversely this index could also highlight the volume, and variability of the amenities that would be required to design an ideal locality for this target demographic. These questions are highly pertinent to the City of Boston, given the growing population of the demographic group over 65 years of age which according to the 2010 US Census.

Regression analysis of the index scores, and the variables available from the 2010 census tracts data showed no positive correlation thereby nullifying any significant or direct relationship between the two. As a consequence, we concluded that the index scores hadn't really shown whether zoning patterns, and public amenities correlated with the desire to live in a neighborhood. There were a number of extraneous factors that had not been included in the analysis which could explain how the elderly self-sorted themselves into the different neighborhoods in Boston, and as a result conclusive results could not be obtained.

II. Introduction

a. Overview of the Dataset

The zoning clearances dataset has been released by the City of Boston as a corollary to its Open Data Initiative after being compiled by the Boston Redevelopment Authority's Planning and Zoning Department, the focal municipal planning and development agency for the city of Boston. The zoning code and the zoning clearances data holds immense importance due the implications it has on the direction of the city's future growth, and the socio economic profile of its denizens.

The dataset consists of a total of 1637 observations and 171 variables. The first 9 variables describe the zoning clearances and categorize them by zone, district, neighborhood, map-no's, articles, sub district identifiers, cartographic identifiers, and regions. The remaining 162 columns represent the zoning land uses which have been classified by usage types. For each specified land use,clearances are either marked 'A' for 'Allowed,' 'C' for 'Conditional,' or 'F' for 'Forbidden or 'A*' in case of an ambiguity.

b. The Concept of Livability

The concept of livability is simple: it assesses which locations in an area provide the best or the worst living conditions. In the contemporary world, assessing livability has a broad range of practical uses, from benchmarking perceptions of development to assigning a hardship allowance as part of an expatriate relocation package. There is no one definition of livability as the concept has been defined differently by various sources.

Livability has been defined as:

"It is the sum of the factors that add up to a community's quality of life—including the built and natural environments, economic prosperity, social stability and equity, educational opportunity, and cultural, entertainment and recreation possibilities" [1]

[1] (Partners for Livable Communities, 2015)

c. AARP's Concept of Livability

The AARP policy paper defines livability as follows:

"A livable community is one that is safe and secure, has affordable and appropriate housing and transportation options, and has supportive community features and services. Once in place, those resources enhance personal independence; allow residents to age in place; and foster residents' engagement in the community's civic, economic, and social life."[2]

d. Development of an Elderly Livability Index

A livability index has therefore been developed for the elderly given the various districts, and zoning clearances provided for the city of Boston by the Boston Zoning Committee. According to the US Census Bureau, one in three Americans is now 50 or above, and in time one in five will be 65+.Therefore, planning for this latent construct is interesting in the sense that it attempts to explain the behavioral and social patterns of an important and ever-growing segment of the US population. It ultimately informs urban planners and policy makers to consider the needs, and preferences of this important demographic group while setting zoning priorities for the various localities in the metropolis.

III. Methodology

While parsing through the dataset, numerous observable variables were identified that would be of great priority to this particular population, and which could be easily accessible given the nature and structure of the zoning clearances dataset. These variables pertained to health care, residential facilities, public services, open spaces, community, and cultural facilities; all of which are of particular value to seniors. These indicators have ultimately been selected on the basis of the AARP report on the community preferences of older adults. This policy paper identified several preferences that are foremost on the minds of the elderly when deciding on localities to reside in post-retirement.

AARP's Public Policy Institute collected data for the study through surveys and focus group discussions. The surveyed population was divided into various segments, which were later classified by household income, gender, metro status, housing tenure, driving status, race, and elderly households with disabilities or a care giver. The data collection process elicited different results from each of these population segments which were then ranked by their inclinations.

---

[2] (AARP Public Policy Insitute, 2014)

According to the above cited report, most members of the population who are over 65 preferred to age in their homes and communities.[3] The importance of proximity to community was a prominent factor yet it varied among the surveyed individuals. Household income was another important factor that influenced their decision making.[4] Increased security through police presence, and the public school systems were other points of discussion that the different sampling groups ranked as being important.

The zoning dataset contained information regarding the approval of city neighborhoods for many of these amenities. These amenities included healthcare facilities such as clinics, clinical laboratories, hospitals, nursing, and convalescent homes etc. Community services included art galleries, auditoriums, cinemas, concert halls etc. The relevant public services comprised of fire stations, penal institutions, police stations etc. whereas residential facilities included elderly housing, multi family dwelling, and one family detached ,and semi attached dwellings etc. Lastly, open spaces were also considered while identifying important manifest variables, and they consisted of recreational camps, stadiums, golf driving ranges etc. All these resources have been recorded as manifest variables in this process, and are meant to be a source of measurement for the latent construct under consideration i.e. the Elderly Livability Index.

a. Calculation of New Measures

The original sub-indexes were created by converting all the factor variables for approvals in our selected zoning categories to the number one, and calculating the means of those values. As a result, we arrived at six sub-index variables which included the healthcare sub-index, the culture sub-index, the public service sub-index, the community sub-index, the residential sub-index. Each of these sub-indexes were individually aggregated at the district level. They were all later aggregated to the district level together, and made a part of a separate data frame. This data frame which was aptly named the ”Elderly_Livability_Index” represented the subindex score by each category, the cumulative total index score ,and the weighted total index scores by preference. In addition, the individual sub-indexes were aggregated to the Neighborhood Statistical Area’s cited in the Zoning NSA Key to provide a further avenue for analysis.

[3] (AARP Public Policy Insitute, 2014)
[4] Ibid

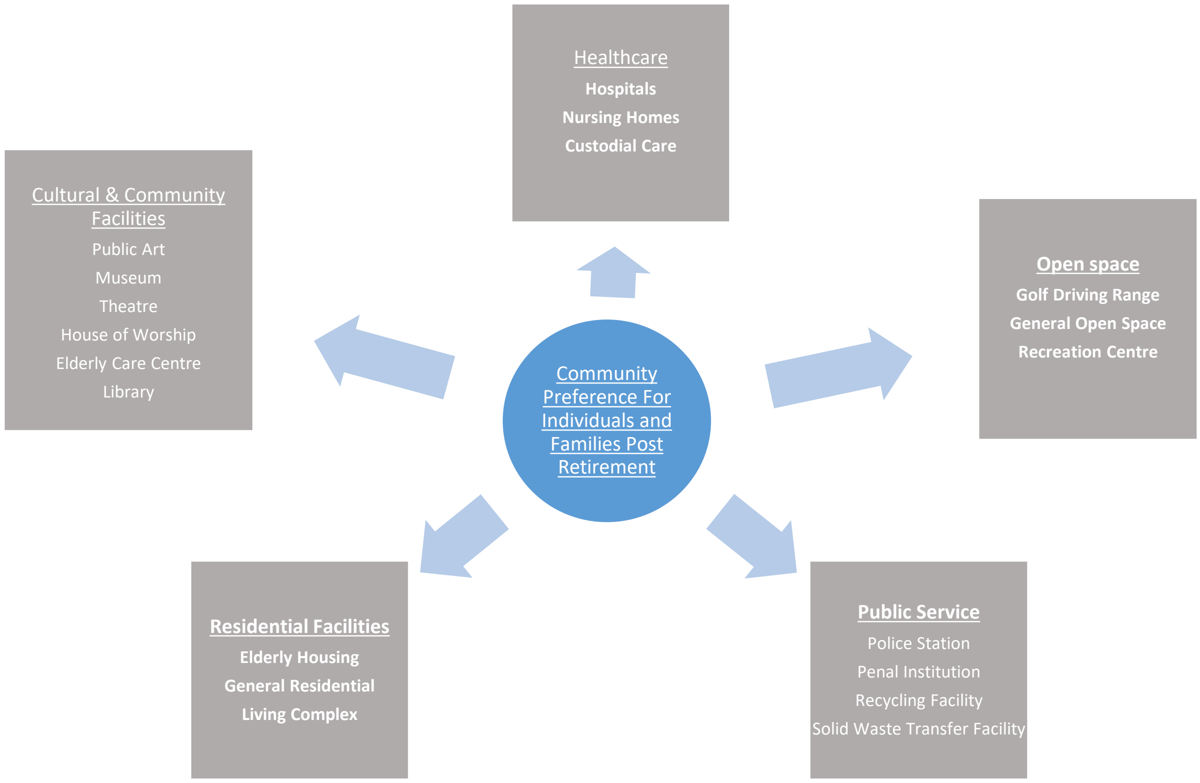

Figure 1-Latent Construct Model

b. Borrowing Measures from other Datasets

The census tracks dataset was the other ancillary source for data utilized in the process of data analysis. While most variables of the census data had a limited application for testing the validity of our index score, we aggregated important demographic information regarding the city of Boston to the NSA level, in order to establish one large main data frame file aptly named,"CensusZoningNSA" by performing a merge function to be able to compare our subindex scores and our census variables at the NSA level. The demographic data that was considered included the proportion of population over 65, the proportion of population that identified itself as White, Black, and Asian. The immigrant population was also considered during this process.

IV. Results

While looking at our main data set we tried to find relationships between the sub-indexes and the census variables by trying to establish a correlation between the two. The most telling variable that could explain the accuracy of our subindex was the proportion of population falling under the age of 65,.We therefore tried to find whether the varying index scores corresponded to the actual size of the over 65 population that resided in the various neighborhoods inside Boston.

Our analysis showed a negative correlation between the numerous sub-index values, and the proportion of population 65, and older. A further regression analysis supplemented our previous findings, and showed no direct or positive relationship between sub-index scores, and present population values for our target demographic within Boston. There was a slight relationship between the cultural amenities subindex, the residential subindex, and the community subindex, and our dependent variable however the values obtained were not significant.

	Estimate Std.	Error	t value	Pr(>|t|)
(Intercept)	0.112995	0.037986	2.975	0.0042
Elderly Cultural Amenities Sub-index	0.075620	0.146693	0.516	0.6081
Elderly Open space Sub-index	-0.033983	0.255030 -	-0.133	0.8944
Elderly Community Sub-index	0.003251	0.147639	0.022	0.9825
Elderly Public Services Sub-index	-0.360341	0.180941	-1.991	0.0509
Elderly Residential Sub-index	0.220382	0.154799	1.424	0.1596
Elderly Healthcare Sub-index	-0.351704	0.243069	-1.447	0.1530

Figure 2-Correlation Analysis

Sub-Indexes	Proportion Of 65 and Older β
Elderly Cultural Amenities Sub-index	**0.188527604**
Elderly Open Space Sub-index	**-0.040959646**
Elderly Community Services Sub-index	**0.009830837**
Elderly Public Services Sub-Index	**-0.408021354**
Elderly Residential Sub-index	**0.418473269**
Elderly Healthcare Sub-index	**-0.357486381**

Figure 3-Standardized β Values

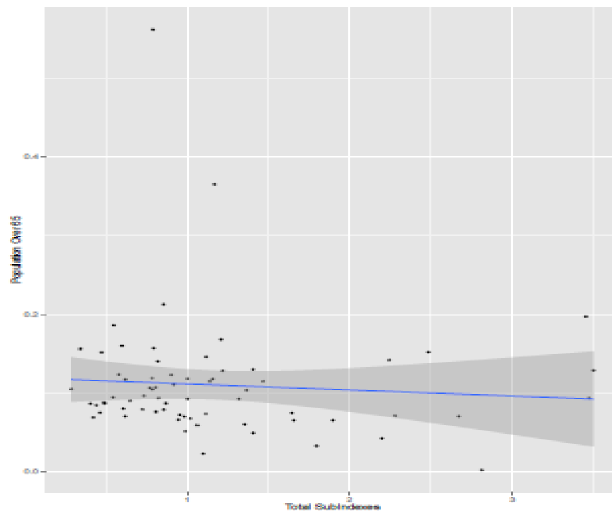

Figure 4-Regression Analysis

A further analysis of these measures through a backwards regression analysis isolating the values of importance such as the elderly community services sub-index, elderly residential sub-index, and the cultural amenities sub-index provided a similar outcome.

Sub-Indexes	Proportion of 65 and Older β
Elderly Public Services Sub-Index	-0.3762329
Elderly Residential Sub-Index	0.4662358
Elderly Healthcare Sub-index	-0.2786108

The regression line in both cases showed a slightly negative cumulative relationship between our sub-index variables and the proportion of the elderly.

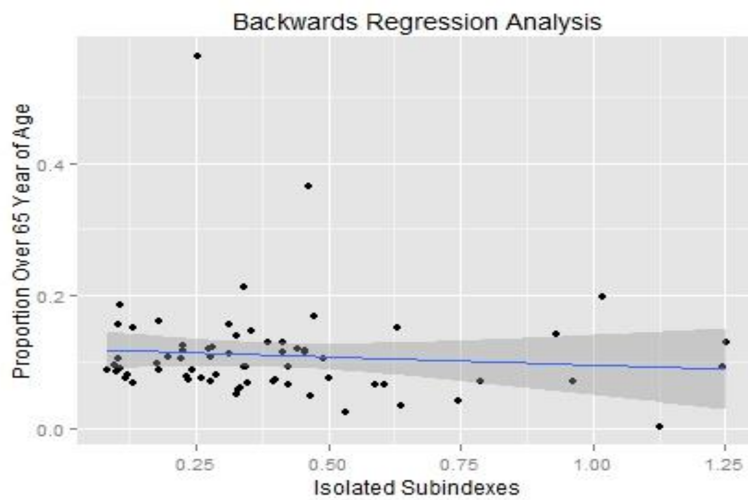

Figure 5-Isolating 3 Indexes for Backwards Regression Analysis

We decided to supplement our statistical findings with some observational data, to see if our findings overlapped with reality, and whether the sub-index scores had any bearing whatsoever on the amenities available at the neighborhood level. This approach would help make better sense of our findings, and isolate any extraneous factors that we might not have considered in our population.

We took a second look at our original zoning clearances data set, to find how the various sub districts were distributed within Boston's neighborhoods, and to isolate the area's where neighborhood audit's would be most useful. Dorchester, South Boston, and Jamaica Plain had the highest number of sub-districts within Boston .Surprisingly the highest number of clearances were also awarded to Dorchester.

On careful observation of our cumulative zoning scores, we found a very different pattern emerging whereby area's showing a lower number of sub-districts such as Bulfinch Triangle, and Boston proper ranked the highest out of all the districts. Unexpectedly none of the major residential areas in Boston scored high within the parameters set of this analysis. Areas with large commercial sub-districts also contained the most clearances for our manifest variables i.e. clearances on healthcare, community services, public services, open spaces etc.

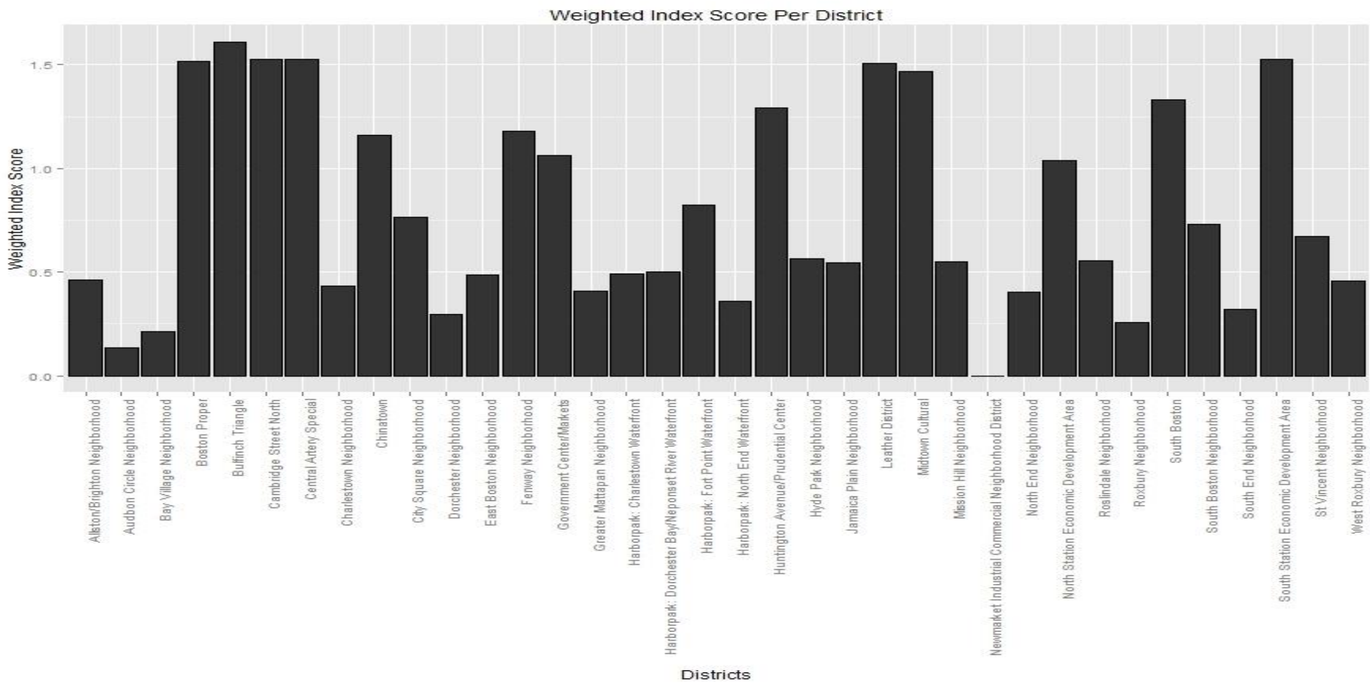

Figure 6-Total Weighted Index Score Per District

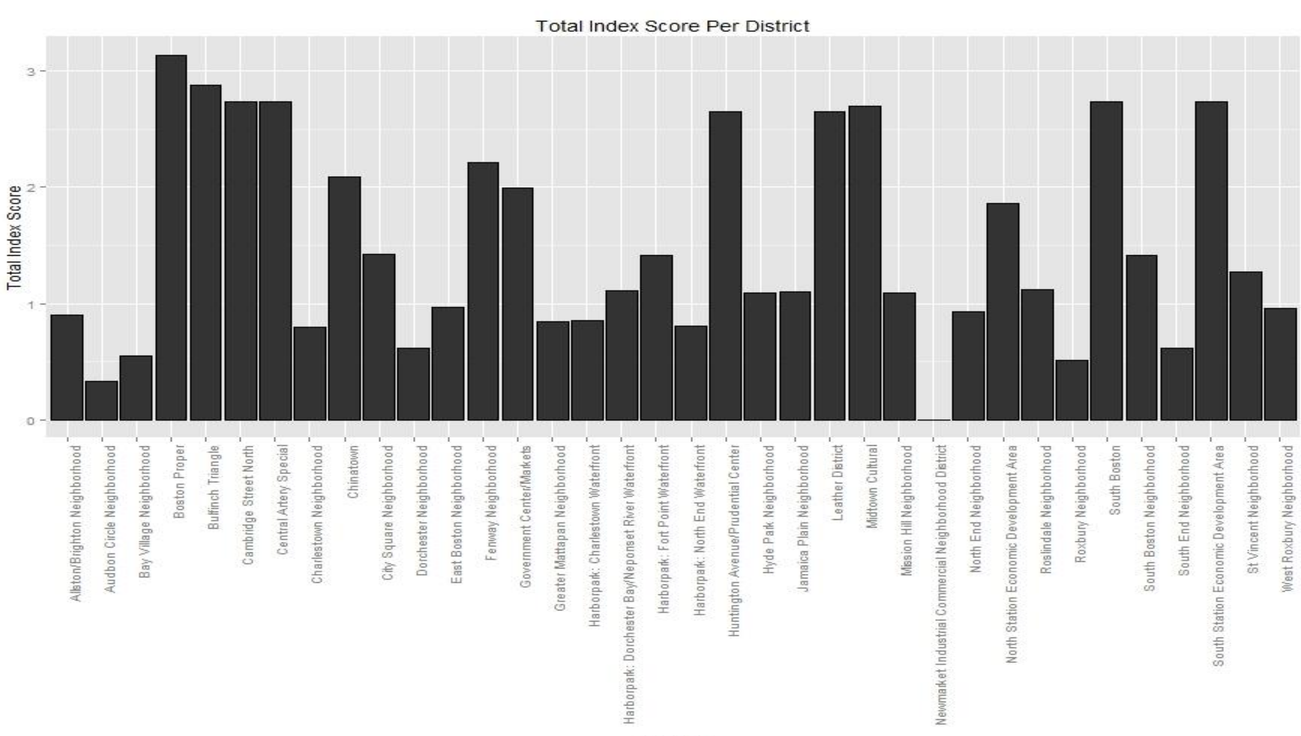

Figure 7-Total Index Score Per District

Out of the total index score of 4, the Bulfinch triangle area which consists of mainly commercial establishments and office space, scored a total of 2.8, and a weighted score of 1.6 out 2. We therefore decided to visit the area to see if the index score indicated had any bearing on the livability of the elderly whatsoever. On visiting the area, we found great disparities between what the index was supposed to indicate, and the actual developments in the area.

Named after the architect Charles Bulfinch, the Bulfinch Triangle is bounded by North Washington, Market, Merrimac, and Causeway Streets. Its vicinity to North Station, the MBTA Haymarket Station, and the AMTRAK terminal, provides great accessibility to commuters. Its vicinity to the numerous historical sites in the city of Boston means that it’s regularly frequented by tourists and visitors alike. Furthermore, the Bulfinch triangle area itself has been designated as a historical site.The fact that the TD Garden which hosts a number of sporting and cultural events each year, is in close proximity to the district can also be cited as a reason for its popularity amongst the residents of Boston, and Massachusetts in general.

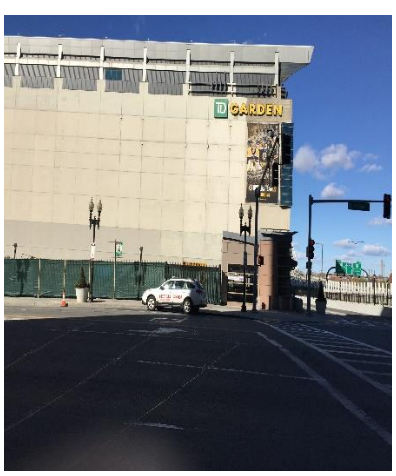

Figure 8-Pictures from the Neighborhood Audit

While all these qualities are noteworthy, they are supposedly not a feature that our demographic group is looking for in a place of residence nor do they figure in our zoning clearances dataset. As we left North station, and scowered the area for observations of interest, we could not help but note the strong commercial nature of the locality. There were a number of business establishments including pubs, restaurants, and other eateries. Construction work could be seen at almost every corner with boards for new housing and commercial developments indicating the large volume of real estate investments going into the area.

Among the categories of the manifest variables that we had considered while formulating the latent construct, were the healthcare facilities in the area which were conspicuous by their absence even though the area had been approved for hospitals, nursing homes, and custodial care facilities. The fact that large hospitals such as the Massachusetts General Hospital, and Tufts Medical that are within close distance of the area may have discouraged the development of similar facilities. Similar extraneous factors such as the close and widespread availability of similar amenities including healthcare, and public services can therefore by cited as a reason as to why this model fails to explain elderly livability.

The Bulfinch Triangle area had very few residential buildings to speak of within its district boundaries. The few apartment buildings that were around were either posh condos or small apartments. When we looked up the demographic characteristics of this locality (census tract 203.3, block group 2, Suffolk Country) through the American Fact Finder, we found that there were only 300 individuals over the age of 60 living in the area. This was far less than the number of seniors living in the adjoining parts of Boston (census tract 301,302,203.1). The comparatively increased presence of high rise apartment buildings, and residential quarters in these neighboring areas gave credence to this fact.

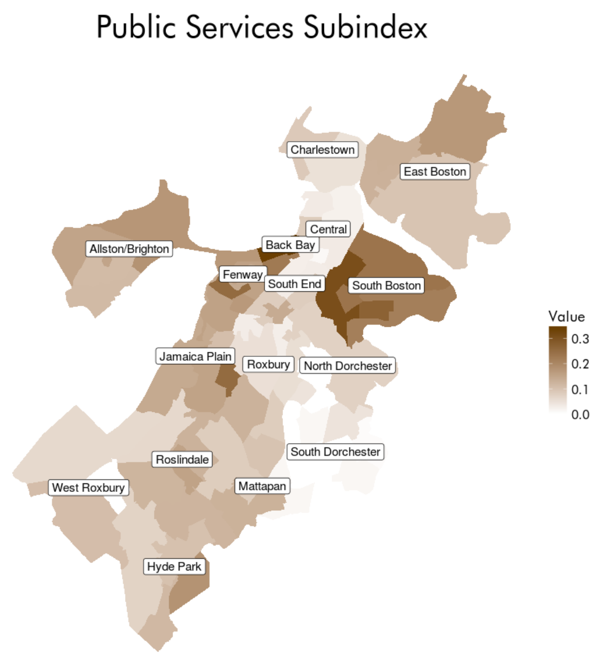

Figure 9-Mapping of the Public Services Amenities Index

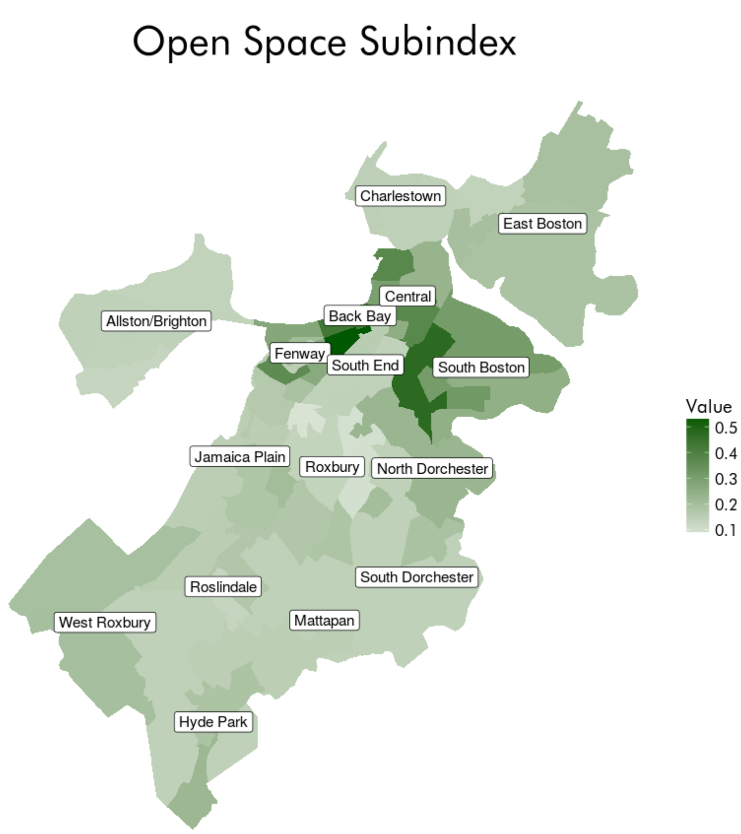

Figure 10-Mapping of the Open space index

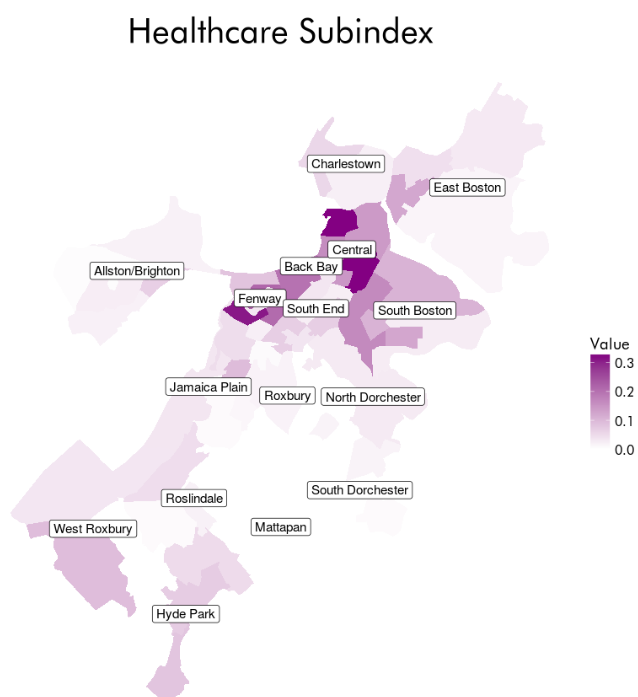

Figure 11-Mapping of the Healthcare subindex

The absence of open space, along with cultural, and community facilities in the area was also noteworthy. This goes to show that while the area's may be approved for the construction of certain facilities by the Boston Reconstruction Authority which in our case included open space areas, and health care services amongst other things, these zoning approvals don't always get translated into reality. Therefore, we feel that my observations during the city exploration assignment do not lend credence to the livability index, and remove any notions of a relationship between residential preferences of the elderly, and the zoning clearances.

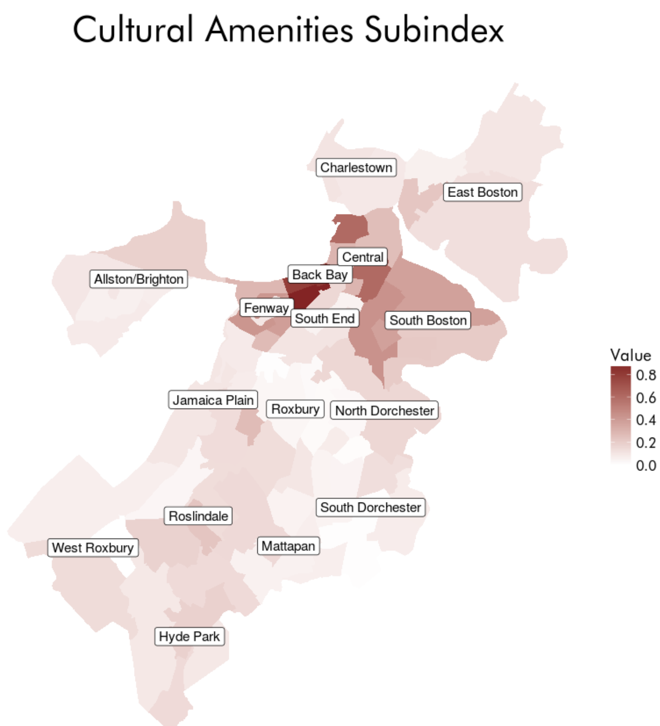

Figure 12-Mapping of the Cultural Amenities Index

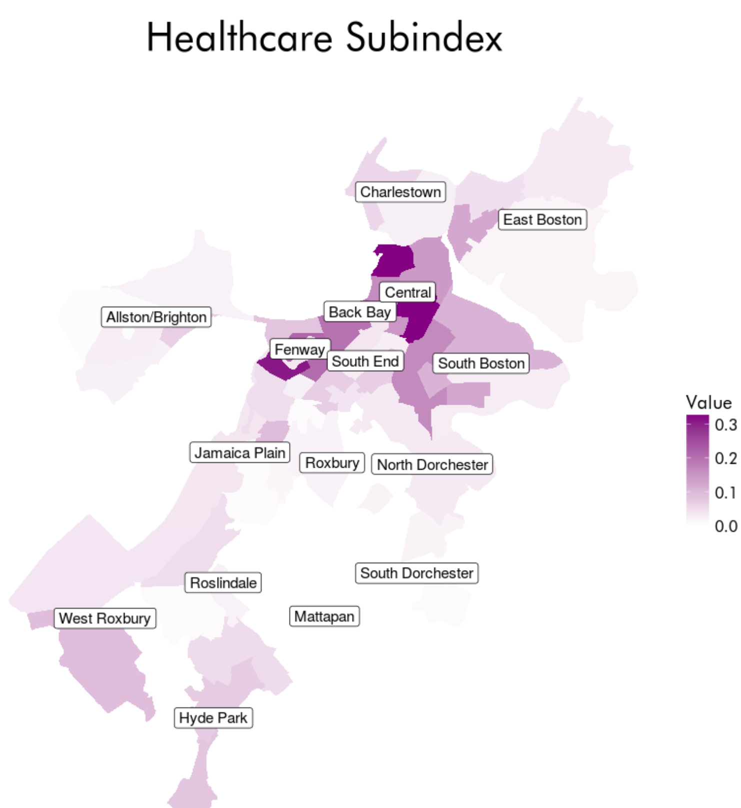

Figure 13-Mapping of the Healthcare Amenities Index

V. Discussion

The index scores have not really shown whether zoning patterns and public amenities correlate with the desire to live in a neighborhood. Therefore we conclude that our target demographic group is self sorting itself into neighborhoods based on various extraneous variables.

The AARP should consequently rehash its approach and look at the optimum mix between the residential amenities, and the other amenities that the elderly would find attractive, and not just focus on individual amenities. In that manner we would be better able to chart out a livability index that would suit the needs and requirements of the elderly population in the city of Boston.

Moreover, the City of Boston can capitalize on the open availability of the zoning clearances dataset, and streamline the process through which individuals and businesses can petition to change zoning clearances. In this way, the data would aid the denizens of the city to partake in the local decision making process.

Works Cited

AARP Public Policy Insitute. (2014). *What Is Livable? Community Preferences of Older Adults.* Washington, DC: AARP.

Partners for Livable Communities. (2015). *Partners for Livable Communities*. Retrieved from http://livable.org/about-us/what-is-livability

Appendix: Annoted R Code

#Uploading all the requisite files

```
Censustracks<-read.csv("F:/R Work/Big Data For Cities/Census Tracts/Census_Tracts.csv")

View(Censustracks)


Ecometrics<-read.csv('F://R Work//Big Data For Cities//Ecometrics//CT_All_Ecometrics_2014.csv')

View(Ecometrics)


ZoningNSA<-read.csv('F:/R Work//Big Data For Cities//Zoning Data//Tracts_to_NSA.csv')

View(ZoningNSA)


Baritracks<-read.csv('F:\\R Work\\Big Data For Cities\\Census Tracts\\Tracts_Boston_BARI.csv')

View(Baritracks)


ZoningNSAID<-read.csv('F:\\R Work\\Big Data For Cities\\Zoning Data\\Zones_with_NSA_R_ID.csv')

View(ZoningNSAID)


ZoningClearancesFinal<-read.csv('F:\\R Work\\Big Data For Cities\\Zoning Data\\Zoning Clearances.csv')

options(scipen = 999)
```

#Checking and identifying required manifest variables

```
options(scipen = 999)

View(ZoningClearancesFinal[90:100])

View(ZoningClearancesFinal[13:18])

View(ZoningClearancesFinal[78:88])

View(ZoningClearancesFinal[71:77])

View(ZoningClearancesFinal[19:29])

View(ZoningClearancesFinal[52:56])
```

#Creating and calculating the subindex variables

```
ZoningClearancesFinal$Elderly_Healthcare_Subindex<-apply(ZoningClearancesFinal[52:56]=='A',1,mean)

ZoningClearancesFinal$Elderly_Culture_Subindex<-apply(ZoningClearancesFinal[19:29]=='A',1,mean)

ZoningClearancesFinal$Elderly_Openspace_Subindex<-apply(ZoningClearancesFinal[71:77]=='A',1,mean)

ZoningClearancesFinal$Elderly_Publicservice_Subindex<-apply(ZoningClearancesFinal[78:88]=='A',1,mean)

ZoningClearancesFinal$Elderly_Community_Subindex<-apply(ZoningClearancesFinal[13:18]=='A',1,mean)

ZoningClearancesFinal$Elderly_Residential_Subindex<-apply(ZoningClearancesFinal[90:100]=='A',1,mean)
```

#Aggregated the Sub-Index Variables by District

```
DistrictHealthAgg<-aggregate(Elderly_Healthcare_Subindex~district,data=ZoningClearancesFinal,mean)
```

View(DistrictHealthAgg)

DistrictCultureAgg<-aggregate(Elderly_Culture_Subindex~district,data=ZoningClearancesFinal,mean)

View(DistrictCultureAgg)

DistrictOpenAgg<-aggregate(Elderly_Openspace_Subindex~district,data=ZoningClearancesFinal,mean)

View(DistrictOpenAgg)

DistrictPubAgg<-aggregate(Elderly_Publicservice_Subindex~district,data=ZoningClearancesFinal,mean)

View(DistrictPubAgg)

DistrictCommAgg<-aggregate(Elderly_Community_Subindex~district,data=ZoningClearancesFinal,mean)

View(DistrictCommAgg)

DistrictResAgg<-aggregate(Elderly_Residential_Subindex~district,data=ZoningClearancesFinal,mean)

View(DistrictResAgg)

#Created a separate dataframe for the Elderly Livability Index

Elderly_Livability_Index<-
data.frame(aggregate(cbind(Elderly_Healthcare_Subindex,Elderly_Community_Subindex,Elderly_Publicservice_Subindex,Elderly_Openspace_Subindex,Elderly_Culture_Subindex,Elderly_Residential_Subindex)~district,ZoningClearancesFinal,mean))

View(Elderly_Livability_Index)

#Created Total ,and Weighted Index Score Variables

Elderly_Livability_Index$Total_Index_Score<-rowSums(Elderly_Livability_Index[,c(2:7)])

Elderly_Livability_Index$Total_Weighted_Score<-
(Elderly_Livability_Index$Elderly_Healthcare_Subindex*1)+(Elderly_Livability_Index$Elderly_Community_Subindex*0.8)+(Elderly_Livability_Index$Elderly_Publicservice_Subindex*0.6)+(Elderly_Livability_Index$Elderly_Openspace_Subindex*0.4)+(Elderly_Livability_Index$Elderly_Culture_Subindex*0.2)+(Elderly_Livability_Index$Elderly_Residential_Subindex*0.1)

```
View(Elderly_Livability_Index)
```

#Merging Zoning Clearances Final & ZoningNSAID

```
ZoningNSAMerge<-merge(ZoningClearancesFinal,ZoningNSAID,by='cartodb_id')
names(ZoningNSAMerge)
```

#Creating Aggregations between Sub-Indexes and NSA Name

```
AggHealthNSA<-aggregate(Elderly_Healthcare_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
AggResNSA<-aggregate(Elderly_Residential_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
AggPubNSA<-aggregate(Elderly_Publicservice_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
AggComNSA<-aggregate(Elderly_Community_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
AggOpenNSA<-aggregate(Elderly_Openspace_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
AggCulNSA<-aggregate(Elderly_Culture_Subindex~NSA_NAME,data=ZoningNSAMerge,mean)
View(AggOpenNSA)
```

#Merging Census Data & ZoningNSA Key

```
CensusNSAMerge<-merge(ZoningNSA,Censustracks,by='CT_ID_10')
names(CensusNSAMerge)

CensusIndex<-merge(CensusNSAMerge,AggHealthNSA,by='NSA_NAME')
CensusIndex<-merge(CensusNSAMerge,AggResNSA,by='NSA_NAME')
CensusIndex<-merge(CensusNSAMerge,AggPubNSA,by='NSA_NAME')
CensusIndex<-merge(CensusNSAMerge,AggComNSA,by="NSA_NAME")
```

```
CensusIndex<-merge(CensusNSAMerge,AggOpenNSA,by='NSA_NAME')
CensusIndex<-merge(CensusNSAMerge,AggCulNSA,by='NSA_NAME')
names(CensusIndex)
```

#Merging Sub-index NSA Aggregations with each other by NSA NAME

```
NSAaggregations<-merge(AggCulNSA,AggOpenNSA,by='NSA_NAME')
NSAaggregations<-merge(NSAaggregations,AggComNSA,by='NSA_NAME')
NSAaggregations<-merge(NSAaggregations,AggPubNSA,by='NSA_NAME')
NSAaggregations<-merge(NSAaggregations,AggResNSA,by='NSA_NAME')
NSAaggregations<-merge(NSAaggregations,AggHealthNSA,by='NSA_NAME')
View(NSAaggregations)

CensusIndex<-merge(CensusNSAMerge,NSAaggregations,by='NSA_NAME')
```

#Aggregating Important Variables with NSA names within the Census dataset

```
Census65Agg<-aggregate(p65older~NSA_NAME,data=CensusIndex,mean)
CensusWhiAgg<-aggregate(propwhite~NSA_NAME,data=CensusIndex,mean)
CensusBlkAgg<-aggregate(propblack~NSA_NAME,data=CensusIndex,mean)
CensusAsAgg<-aggregate(propasian~NSA_NAME,data=CensusIndex,mean)
CensusOtherRaceAgg<-aggregate(propother~NSA_NAME,data=CensusIndex,mean)
CensusImmigAgg<-aggregate(propimmigrant~NSA_NAME,data=CensusIndex,mean)
CensusParkAgg<-aggregate(Park~NSA_NAME,data=CensusIndex,mean)
```

#Merging Aggregated Census Variables together

```
CensusAggMerge<-merge(Census65Agg,CensusWhiAgg,by='NSA_NAME')
```

```
CensusAggMerge<-merge(CensusAggMerge,CensusBlkAgg,by='NSA_NAME')
CensusAggMerge<-merge(CensusAggMerge,CensusOtherRaceAgg,by='NSA_NAME')
CensusAggMerge<-merge(CensusAggMerge,CensusAsAgg,by='NSA_NAME')
CensusAggMerge<-merge(CensusAggMerge,CensusImmigAgg,by='NSA_NAME')
CensusAggMerge<-merge(CensusAggMerge,CensusParkAgg,by='NSA_NAME')
names(CensusAggMerge)



#Merging CensusAggMerge(Aggregated Census Variables) with NSA Aggregations(Sub Index NSA Aggregations)

CensusZoningNSA<-merge(CensusAggMerge,NSAaggregations,by='NSA_NAME')
View(CensusZoningNSA)
names(CensusZoningNSA)



#Now lets create some maps
#Installing and activating required libraries/packages

install.packages('rgeos', type='source')
install.packages('rgdal', type='source')
install.packages("gpclib", type="source")
install.packages('maptools', type='source')
install.packages('dplyr', type='source')
install.packages("extrafont", type="source")
install.packages("ggplot2", type="source")


library(dplyr)
library(extrafont)
library(rgdal)
```

```
library(maptools)

library(ggplot2)

library(rgeos)

library(maptools)


tracts.shp <- readOGR(dsn = 'F://R Work//Big Data For Cities//Census Tracts//Tracts_Boston_2015_BARI//Tracts_Boston BARI.shp', 'Tracts_Boston BARI')


tracts.shp <- merge(tracts.shp, CensusIndex, by = 'CT_ID_10', all.x = TRUE, sort = FALSE)


tracts.df <- fortify(tracts.shp, region = "CT_ID_10")


tracts.df <- merge(tracts.df, tracts.shp@data, by.x = 'id', by.y = 'CT_ID_10', all.x = TRUE)


my.map.theme <- theme(text = element_text(family = "Futura Bk BT"),
                plot.title = element_text(size = 30),
                axis.line = element_blank(),
                axis.title.x = element_blank(),
                axis.title.y = element_blank(),
                axis.text.x = element_blank(),
                axis.text.y = element_blank(),
                #legend.title = element_text(size = 16, face = "bold"),
                legend.title = element_text(size = 14),
                legend.text = element_text(size = 12),
                panel.grid.major = element_blank (), # remove major grid
                panel.grid.minor = element_blank (),  # remove minor grid
                axis.text = element_blank (),
                axis.title = element_blank (),
                axis.ticks = element_blank (),
                strip.text = element_text(size = 20))
```

```
#Get Neighborhood names and locations

NBs.df <- fortify(tracts.shp, region = "BRA_PD.x")
NBcentroids <-aggregate(cbind(long, lat) ~ id, data = NBs.df, FUN = function(x) mean(range(x)))

#Lets try some shorter names

NBcentroids$id <- sub("Fenway/Kenmore", "Fenway", NBcentroids$id)
NBcentroids$id <- sub("Back Bay/Beacon Hill", "Back Bay", NBcentroids$id)

'''

#For improved color scales
library(scales)

#For Public Services
ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Publicservice_Subindex )) +
  geom_polygon() +
  scale_fill_gradient2(high = muted("darkgoldenrod1")) +
  geom_label(data=NBcentroids,
        aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
  labs(title = "Public Services Subindex", fill = "Value") +
  theme_minimal() +  my.map.theme + coord_map()

#For Open Space

ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Openspace_Subindex )) +
```

```
 geom_polygon() +
 scale_fill_gradient2(high = muted("darkgreen")) +
 geom_label(data=NBcentroids,
        aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
 labs(title = "Open Space Subindex", fill = "Value") +
 theme_minimal() +  my.map.theme + coord_map()
```

#For Health Care

```
ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Healthcare_Subindex )) +
 geom_polygon() +
 scale_fill_gradient2(high = muted("darkmagenta")) +
 geom_label(data=NBcentroids,
        aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
 labs(title = "Healthcare Subindex", fill = "Value") +
 theme_minimal() +  my.map.theme + coord_map()
```

#For Cultural Amenities

```
ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Culture_Subindex )) +
 geom_polygon() +
 scale_fill_gradient2(high = muted("darkred")) +
 geom_label(data=NBcentroids,
        aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
 labs(title = "Cultural Amenities Subindex", fill = "Value") +
 theme_minimal() +  my.map.theme + coord_map()
```

#For Community Facilities

```
ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Community_Subindex )) +
```

```
 geom_polygon() +
 scale_fill_gradient2(high = muted("darkblue")) +
 geom_label(data=NBcentroids,
           aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
 labs(title = "Community Facilities Subindex", fill = "Value") +
 theme_minimal() +  my.map.theme + coord_map()
```

#For Residential Amenities

```
ggplot(data=tracts.df, aes(x=long, y=lat, group=group, fill = Elderly_Residential_Subindex )) +
 geom_polygon() +
 scale_fill_gradient2(high = muted("cornsilk4")) +
 geom_label(data=NBcentroids,
           aes(long, lat, label = id, group = NULL, fill=NULL), size=4) +
 labs(title = "Residential Amenities Subindex", fill = "Value") +
 theme_minimal() +  my.map.theme + coord_map()
```

#Merging the relevant files

```
Analysisfile<-merge(merge1,merge2,by='NSA_NAME')
View(Analysisfile)
```

Lets do a Single-sample t-test

#According to the US Census 13% of the population of the US is 65 and over

#Lets try to answer the question:'Is the percentage of the population over 65 years of age in Boston higher than the national average?'

```
t.test(Censustracks$p65older, mu=0.13)
names(Analysisfile)
```

#Lets do a two-sample t-test

#And try to answer the question:'Is the percentage of the population older than 65 higher with greater weighted index scores?'

```
t.test(Analysisfile$p65older,Analysisfile$Weighted_Score, data=Analysisfile)
```

#Lets do a ANOVA test

```
anova<-aov(Analysisfile$p65older~Analysisfile$Weighted_Score, data=Analysisfile)
anova<-aov(Analysisfile$p65older~Analysisfile$district, data=Analysisfile)
```

#Lets compare means

```
TukeyHSD(anova)
```

#Lets visualize group differences

```
require(reshape2)
require(ggplot2)

names(Analysisfile)
melted<-melt(Analysisfile[c(235,183)],id.vars=c("DISTRICT"))

View(melted)
melted$value
means<-aggregate(value~DISTRICT,data=melted,sum)
names(means)[2]<-"mean"
View(means)
```