Maps of boston

Ari Anisfeld

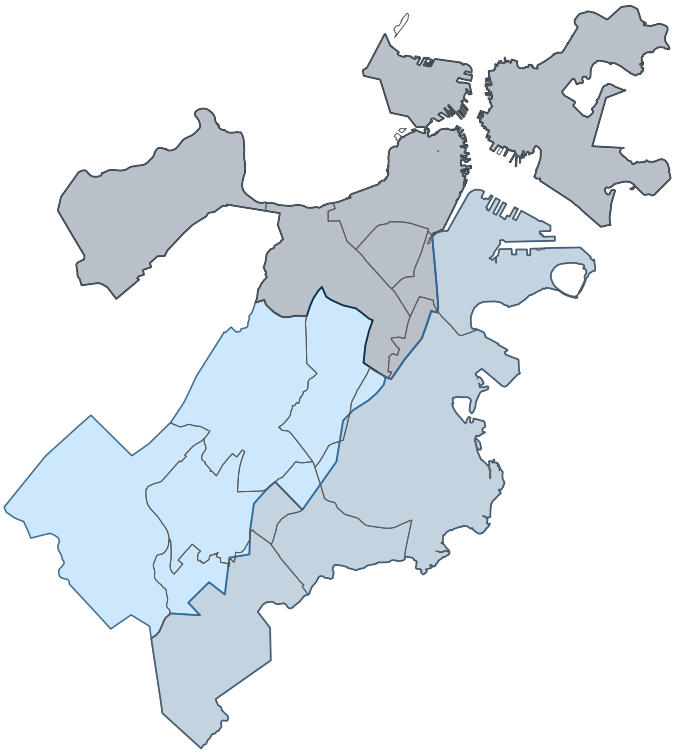
12/7/2021

Making map ignoring zones.

This first pass uses the neighborhood shapes from Boston open data. (https://bostonopendata-boston.opendata.arcgis.com/datasets/3525b0ee6e6b427f9aab5d0a1d0a1a28_0) and using the zones shapefile you previously made.

```
zones <- st_read("data/zones/zones.shp", quiet = TRUE) %>%
          transmute(zone = ifelse(!is.na(north), north, 3),
                 geometry)

neighborhoods_xwalk <-
  read_csv("data/neighborhood_crosswalk.csv",
           skip = 1,
           col_names = c("Name", "neighborhood_code", "x"))  %>%
  select(-x)

raw_neighborhoods <-
    st_read("data/Boston_Neighborhoods/Boston_Neighborhoods.shp", quiet = TRUE)

neighborhoods <-
raw_neighborhoods %>%
  left_join(neighborhoods_xwalk) %>%
  group_by(neighborhood_code) %>%
  combine_polygons() %>%
  filter(!is.na(neighborhood_code))

ggplot() +
  geom_sf(data = zones, aes(fill = zone, color = zone), lwd = .3, alpha = .3, show.legend = FALSE)  +
  geom_sf(data=neighborhoods, lwd = .1, fill=NA) +
  theme_void() +
  labs(title = "Using your zone shape file + Boston open data")
```

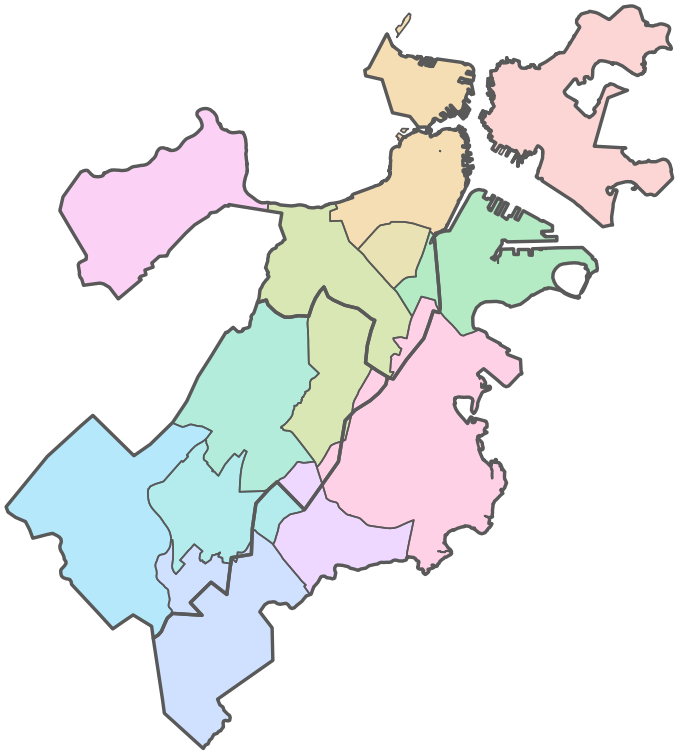
Using your zone shape file + Boston open data



We can also foreground the neighborhood

```
ggplot() +
  geom_sf(data = neighborhoods, aes(fill = as.factor(neighborhood_code)), lwd = .3, alpha = .3, show.leg
  geom_sf(data = zones, lwd = .6, fill=NA) +
  theme_void() +
  labs(title = "Using your zone shape file + Boston open data (alt)")
```

Using your zone shape file + Boston open data (alt)

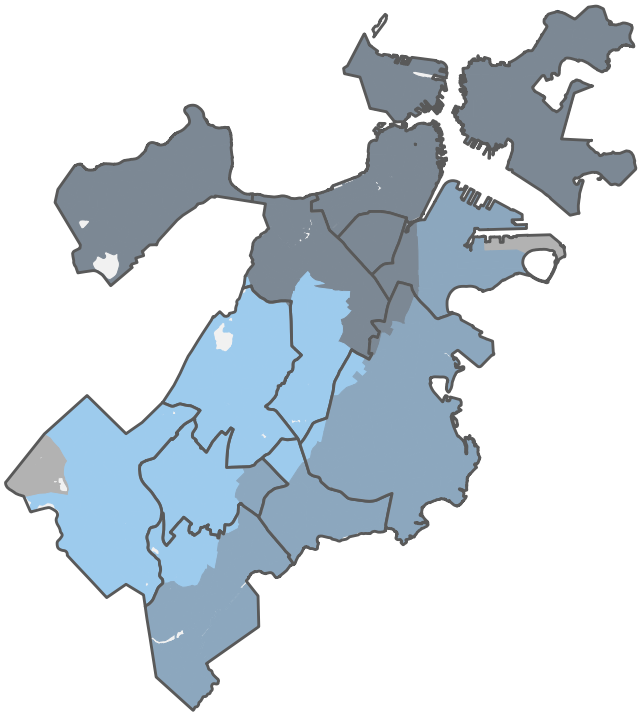


Making maps from tracts

```
geocode_to_neighborhood <- read_csv("data/geocode_to_neighborhoods.csv")
geocode_to_zone  <- haven::read_dta("data/geocodetozone.dta")

geocodes <-
  st_read("data/geocodes/Geocodes.shp", quiet = TRUE) %>%
  # geocodes shapefile has small issues which make certain polygons
  # "invalid" (e.g. st_is_valid(.) returns FALSE).
  # this buffering trick fixes those issues.
  st_simplify() %>%
  st_buffer(dist = 0) %>%
  left_join(geocode_to_zone, "geocode") %>%
  left_join(geocode_to_neighborhood, "geocode")

ggplot() +
  geom_sf(data = geocodes,
          aes(fill = zones, color = zones), lwd = 0, show.legend = FALSE) +
   geom_sf(data = neighborhoods, lwd = .5, alpha = .5,
        show.legend = FALSE) +
  theme_void() +
  labs(title = "Using your geocodes shape file for zone + Boston open data",
       subtitle = "Notice grey and white areas ")
```

Using your geocodes shape file for zone + Boston open data

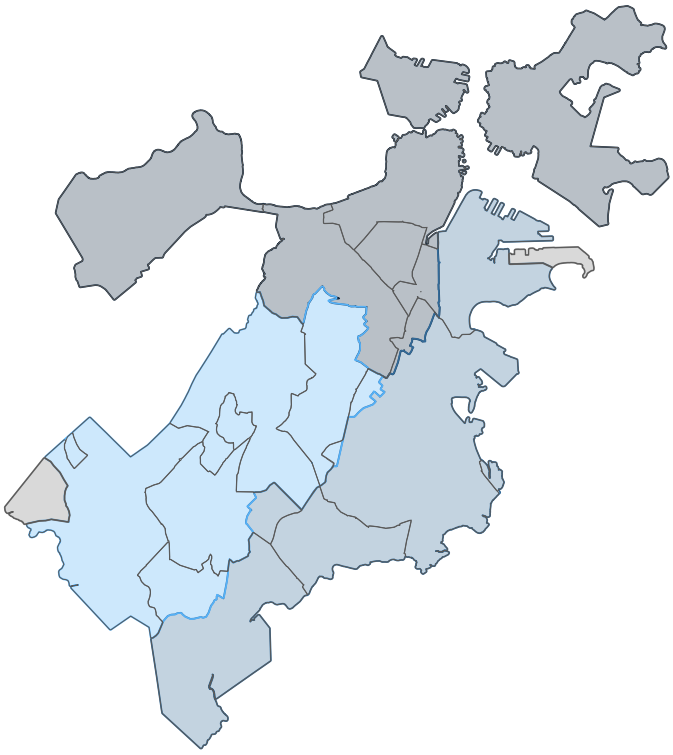
Notice grey and white areas



```
combine_and_clean_polygons <- function(data, grouping_var, TOL = 200){
data  %>%
  group_by({{ grouping_var }}) %>%
  combine_polygons() %>%
  sf_remove_holes() %>%
  st_buffer(-TOL) %>%
  st_buffer(TOL + .1*TOL)
}


these_zones <-
  combine_and_clean_polygons(geocodes, zones, 100)

these_neighborhoods <-
  combine_and_clean_polygons(geocodes, neighborhoodid2, 100)
```

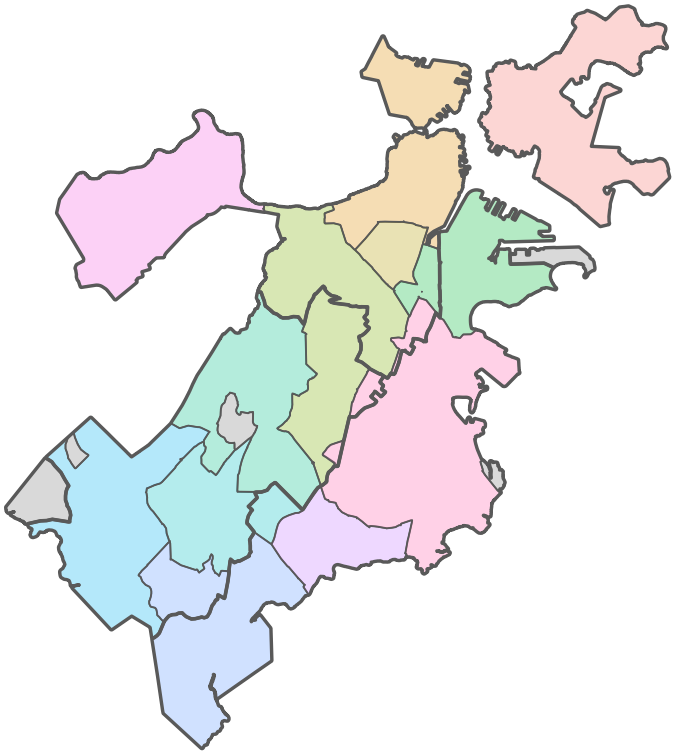
```
ggplot() +
  geom_sf(data = these_zones,
          aes(fill = zones, color = zones),
          lwd = .3, alpha = .3, show.legend = FALSE)  +
  geom_sf(data= these_neighborhoods,
          lwd = .1, fill=NA) +
  theme_void() +
  labs(title = "Using geocodes")
```

Using geocodes



```
ggplot() +
  geom_sf(data = these_neighborhoods,
          aes(fill = as.factor(neighborhoodid2)), lwd = .3, alpha = .3, show.legend = FALSE)  +
  geom_sf(data = these_zones, lwd = .6, fill=NA) +
  theme_void() +
  labs(title = "Using geocodes (alt)")
```

Using geocodes (alt)

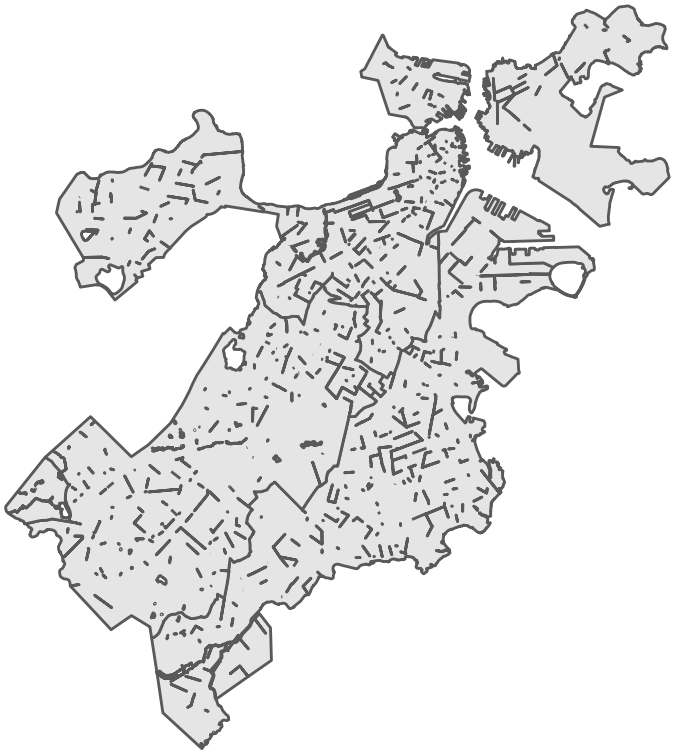


What's the challenge?

The geocodes shapefile has many extraneous lines, which makes it difficult to collapse into polygons.

```
my_plot <- function(data) data %>% ggplot() + geom_sf() + theme_void()

geocodes  %>%
  group_by(zones) %>%
  combine_polygons() %>%
  my_plot() +
  labs(title = "Simply combining leaves a lot of issues")
```

Simply combining leaves a lot of issues



```
geocodes  %>%
  group_by(zones) %>%
  combine_polygons() %>%
  sf_remove_holes() %>%
  my_plot() +
  labs(title = "Removing holes gets rid of many of the small lines",
       subitle = "But we still have lines")
```

Removing holes gets rid of many of the small lines

```
geocodes  %>%
  group_by(zones) %>%
  combine_polygons() %>%
  sf_remove_holes() %>%
  st_buffer(-100) %>%
  st_buffer(110) %>%
  my_plot() +

  labs(title = "Buffering almost does the job")
```

Buffering almost does the job

```
geocodes  %>%
  group_by(zones) %>%
  combine_polygons() %>%
  sf_remove_holes() %>%
  st_buffer(-200) %>%
  st_buffer(220) %>%
  sf_remove_holes() %>%
  my_plot() +
  labs(title = "But buffering too much makes things worse",
       subtitle = "(We lose some corners")
```

But buffering too much makes things worse

(We lose some corners