American Family Field Mapping

Paulina Leperi

2023-05-12

To begin, I'll acquire CRS data for American Family Field. In this dataset, it's still called Miller Park.

```
stadium_url <- "https://services1.arcgis.com/Hp6G80Pky0om7QvQ/arcgis/rest/services/MajorSportVen
ues/FeatureServer/0/query?outFields=*&where=1%3D1&f=geojson"
stadium <- st_read(stadium_url) %>%
  filter(str_detect(str_to_upper(NAME), "MILLER PARK"))
```

```
## Reading layer `OGRGeoJSON' from data source
##   `https://services1.arcgis.com/Hp6G80Pky0om7QvQ/arcgis/rest/services/MajorSportVenues/Featur
eServer/0/query?outFields=*&where=1%3D1&f=geojson'
##   using driver `GeoJSON'
## Simple feature collection with 842 features and 27 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -157.93 ymin: 17.0697 xmax: -65.7985 ymax: 53.57133
## Geodetic CRS:  WGS 84
```

Once the data is loaded, a recommendation is needed for the best CRS to use. All layers of the maps created in subsequent steps will be transformed to the same CRS.

```
suggest_crs(stadium)
```

```
## # A tibble: 10 × 6
##    crs_code crs_name                        crs_type crs_gcs crs_units crs_proj4
##    <chr>    <chr>                           <chr>      <dbl> <chr>     <chr>
##  1 8160     NAD83(HARN) / WISCRS Kenosha, … project…    4152 us-ft     +proj=tm…
##  2 8159     NAD83(HARN) / WISCRS Kenosha, … project…    4152 m         +proj=tm…
##  3 7613     NAD83(2011) / WISCRS Kenosha, … project…    6318 us-ft     +proj=tm…
##  4 7554     NAD83(2011) / WISCRS Kenosha, … project…    6318 m         +proj=tm…
##  5 6609     NAD83(2011) / Wisconsin South … project…    6318 us-ft     +proj=lc…
##  6 6608     NAD83(2011) / Wisconsin South   project…    6318 m         +proj=lc…
##  7 3700     NAD83(NSRS2007) / Wisconsin So… project…    4759 us-ft     +proj=lc…
##  8 3699     NAD83(NSRS2007) / Wisconsin So… project…    4759 m         +proj=lc…
##  9 32154    NAD83 / Wisconsin South         project…    4269 m         +proj=lc…
## 10 32054    NAD27 / Wisconsin South         project…    4267 us-ft     +proj=lc…
```

I'll use 8159, which measures distance in meters.

```
stadium_t <- st_transform(stadium, 8159)
```

Next, I want to look at the census tracts surrounding the stadium. I'll acquire them using Tidycensus and transform them to the same CRS, then filter to only include the tracts within 50 km of the stadium. (Please note that use of Tidycensus requires an API key.)

```
wi_tracts <- tracts("WI", cb = TRUE, year = 2021) %>%
  st_transform(8159) %>%
  st_filter(stadium_t, .predicate = st_is_within_distance, dist = 50000) %>%
  erase_water()
```

```
## Fetching area water data for your dataset's location...
```

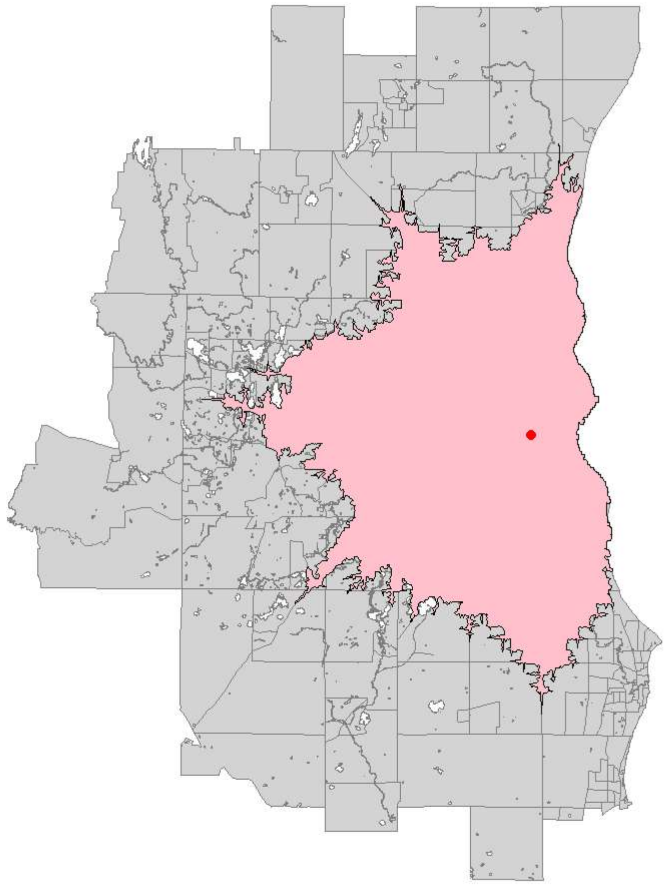
```
## Erasing water area...
## If this is slow, try a larger area threshold value.
```

The final layer in my first visualization will show the area within a 30-minute drive of the stadium.

```
iso30min <- mb_isochrone(
  stadium_t,
  time = 30,
  profile = "driving-traffic"
) %>%
  st_transform(8159)
```

Here’s a map that layers on the three components generated so far.

```
ggplot() +
  geom_sf(data = wi_tracts, color = "grey50", fill = "lightgray") +
  geom_sf(data = iso30min, color = "black", fill = "pink") +
  geom_sf(data = stadium_t, color = "red") +
  theme_void()
```

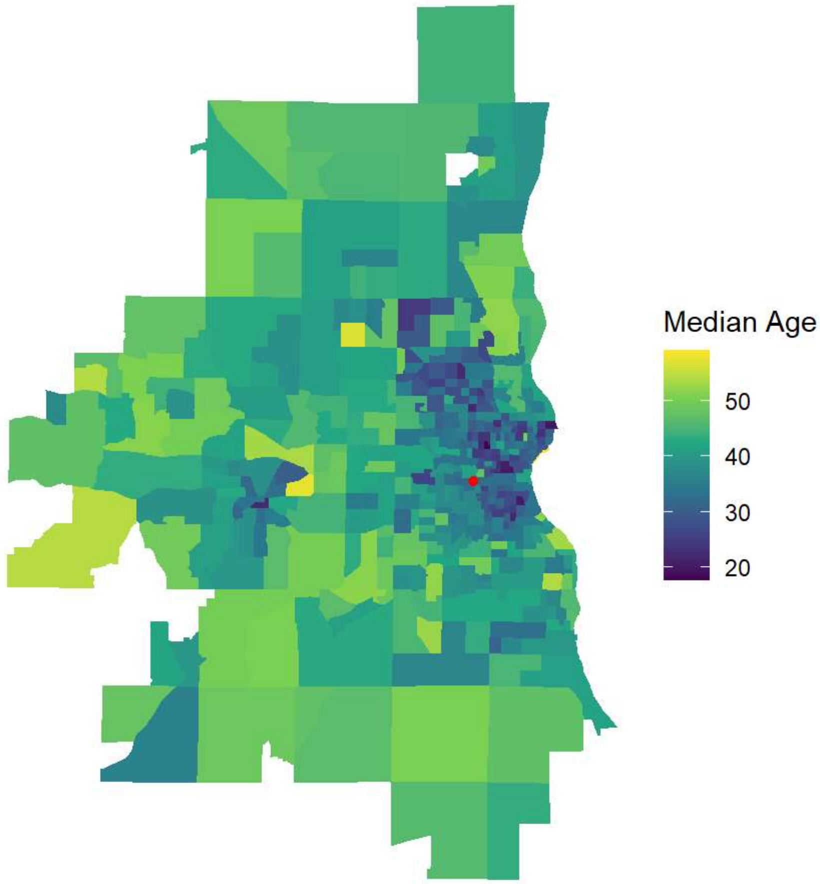


The next thing I want to look at is the age of the population near the stadium. The following code gets median age by census tract for the state of Wisconsin and then filters the dataset to include only those tracts that intersect the 30-minute drive polygon.

```
age_by_tract <- get_acs(
  geography = "tract",
  variables = "B01002_001",
  state = "WI",
  year = 2021,
  geometry = TRUE
) %>%
  st_transform(8159) %>%
  st_filter(iso30min, .predicate = st_intersects) %>%
  na.omit()
```

```
## Getting data from the 2017-2021 5-year ACS
```

```
ggplot(age_by_tract) +
  geom_sf(aes(fill = estimate), color = NA) +
  scale_fill_viridis_c() +
  geom_sf(data = stadium_t, color = "red") +
  theme_void() +
  labs(fill = "Median Age")
```



There appear to be areas to the northeast of the stadium where the median age of the population is low, but this can be quantified in a more precise way.

I will use local spatial autocorrelation analysis to identify “hotspots” where the population differs from total values in the dataset.

Analysis like this can help precisely target advertising for stadium events and promotions.

```
neighbors <- poly2nb(age_by_tract, queen = TRUE)
weights <- nb2listw(include.self(neighbors))
age_by_tract$localG <- localG(age_by_tract$estimate, weights)

age_by_tract <- age_by_tract %>%
  mutate(hotspot = case_when(
    localG >= 2.576 ~ "High cluster",
    localG <= -2.576 ~ "Low cluster",
    TRUE ~ "Not significant"
  ))

ggplot(age_by_tract) +
  geom_sf(aes(fill = hotspot), color = "grey90", size = 0.1) +
  scale_fill_manual(values = c("red", "cyan", "grey")) +
  geom_sf(data = stadium_t, color = "black", size = 5) +
  theme_void() +
  labs(title = "High/Low Median Age \nNear American Family Field",
       fill = "")
```