New store of IKEA

Working with Spatial Data and the sf package

Background

IKEA is looking to open a new store in Boston, MA. You are in charge of finding suitable locations for the new store. Now you have collected datasets from online GIS resources, and will make sure the locations satisfy the following criteria: (1) The store is located in census tracts with appropriate median income (25,000-80,000) and household value (200,000-500,000); (2) The store will be built in the current "Open Land" areas; (3) The store should be within 2000ft of major roads; (4) The store should be at least 200,000 sq ft.

Data

You are provided with three data shapefiles in the CS_07_spatial_case_study_data folder (which is in both the Lite_SDS and Ess_SDS folders):

1. *Boston_CensusTracts.shp* - Census tracts in Boston; data collected from US Census Bureau (census.gov).
    - Attribute `Med_Income` is the median annual income of a family in this census tract (in dollar).
    - Attribute `Med_HouseV` is the median household value in this census tract (in dollar).
2. *Boston_LandUse.shp* - Land use information of Boston; data collected from MassGIS (mass.gov).
    - Attribute `LU05_DESC` is the land use type, i.e. what this land is used for.
3. *Boston_MajorRoads.shp* - Primary roads in Boston; data collected from US Census Bureau (census.gov).

All datasets have been projected to the State Plane, Massachusetts, Mainland coordinate system with the unit ft.

Tasks

- Determine the candidate locations of a new IKEA store based on the criteria above
- Produce the map of candidate locations

Instructions

1. Load the necessary packages.
2. Read spatial data from the shapefiles provided using `read_sf()`, and check each data to see the geometry and attribute information stored in the data.
3. Filter the census data for the range of `Med_Income` as 25,000-80,000 and the range of `Med_HouseV` as 200,000-500,000.
4. Filter the land use data for just 'Open Land' value for `LU05_DESC`.
5. Create a buffer zone for the major roads in Boston with distance of 2000ft using `st_buffer()`.
6. Check the geometry validity of outputs from step 3-5 using `st_is_valid()`. In case there is any invalid polygon, use `st_make_valid()` to make the geometry valid.
7. Get intersections of all selected areas from steps 3-5 using `st_intersection()` with valid geometries.
    - After intersections, you may get many polygons with attribute information from original data
    - You need to union adjacent polygons to one polygon and also keep polygons with no shared boundaries separate using `st_union()` and `st_cast("POLYGON")`
8. Calculate area of unioned polygons using `st_area()` and convert the output to numeric values.
9. Select the polygons with area larger than 200,000 sq ft.
10. Produce the map with the final candidate polygons, and some other features on the map to provide context (e.g. Boston census tracts with median household value in the census tract).
11. Make sure you also have all essential graph elements, such as title, subtitle, legends, labels, proper colors, font and size and etc. Think about how you can clearly show the candidate sites in the map with some context information. You may want to adjust ranges of x and y axis to focus on area of the final candidate sites.
12. Save your graph into a png file.

What you submit

1. A markdown (.rmd) file, CS07_yourname.rmd (including annotation explaining your code).
2. A png file of your map showing the candidate locations for a new IKEA store (CS07_map_yourname.png).
3. A brief narrative of your results, including what the code does (at beginning of file) and then what you found (at the end of the file) such as how many candidates you found, their general distribution in Boston, etc.