Assesing Demand for Transit Accessesible Apartments in Chicago

Noah Anderson

Project Overview

The methodology laid out in Chapter 1 of Public Policy Analytics: Code & Context for Data Science in Government [3] is utilized in this project to examine transit demand in the City of Chicago. Steif's methodology for indicator variable analysis is applied to open transit data sourced from the City of Chicago, which is subsequently wrangled to ensure compatibility with American Community Survey data. The project employs indicator variables to evaluate transit demand in the City of Chicago.

Introduction

The City of Chicago has been actively promoting Transit Oriented Development (TOD) since 2013, following the adoption of a series of ordinances by the City Council known as the Equitable TOD (ETOD) plan [1]. This plan aims to foster compact and mixed-use development in close proximity to Chicago Transit Authority (CTA) and Metra rail stations, with the objectives of enhancing walkability, boosting transit ridership, reducing greenhouse gas emissions, and creating a more stable tax base. The ETOD plan is geared towards ensuring that all residents, regardless of income, race, ethnicity, age, gender, or immigration status, can benefit from the advantages of pedestrianized areas located near transit stops.

The objective of this project is to leverage indicator variables from the American Community Survey to assess demand for transit access. Furthermore, an informal analysis is conducted on the equity of TOD in Chicago by examining demographic data. The analyzed variables include the percentage of white residents, the percentage of residents holding a bachelor's degree, the percentage below the poverty line, and the average rent. By segmenting ACS tracts into TOD and non-TOD areas, differences between these variables are compared within the two geographic groups. Data is sourced from the 2005-2009 and 2014-2019 Five Year ACS to observe changes in the selected indicator variables before and after the implementation of the ETOD plan by the City of Chicago.

Pulling ACS Data

The chunk presented below utilizes the tidycensus package to extract Five Year ACS data for the 2005-2009 and 2015-2019 surveys, which is focused on Cook County IL. While the 2017-2021 ACS data was also available, it was deemed appropriate to avoid using data that encompasses the pandemic year of 2020 due to its potential impact on transit ridership and popularity. Two distinct data frames are created from the extracted data, each of which are restricted to Cook County IL for further analysis.

```
# Gets data for 2005-2009 Five Year ACS for Cook County and wrangles it
cook09 <-
  get_acs(geography = "tract", variables = c("B25026_001E","B02001_002E","B15001_050E",
                                             "B15001_009E","B19013_001E","B25058_001E",
                                             "B06012_002E"), # Selects target variables
          year=2009, state=17, county=031, geometry=T, output="wide", cache_table = TRU
E) %>%
  st_transform('ESRI:102728') %>% # Converts to a standard coordinate system

  # Renames variables
  rename(TotalPop = B25026_001E, Whites = B02001_002E,
         FemaleBachelors = B15001_050E, MaleBachelors = B15001_009E,
         MedHHInc = B19013_001E, MedRent = B25058_001E,
         TotalPoverty = B06012_002E) %>%
  dplyr::select(-NAME, -starts_with("B")) %>%

  # Converts variables to percentages
  mutate(pctWhite = ifelse(TotalPop > 0, Whites / TotalPop,0),
         pctBachelors = ifelse(TotalPop > 0, ((FemaleBachelors + MaleBachelors) / TotalP
op),0),
         pctPoverty = ifelse(TotalPop > 0, TotalPoverty / TotalPop, 0),
         year = "2009", county = "Cook") %>%
  dplyr::select(-Whites, -FemaleBachelors, -MaleBachelors, -TotalPoverty)

# Repeats steps for 2014-2019 ACS
cook19 <-
  get_acs(geography = "tract", variables = c("B25026_001E","B02001_002E","B15001_050E",
                                             "B15001_009E","B19013_001E","B25058_001E",
                                             "B06012_002E"),
          year=2019, state=17, county=031, geometry=T, output="wide", cache_table = TRU
E) %>%
  st_transform('ESRI:102728') %>%
  rename(TotalPop = B25026_001E, Whites = B02001_002E,
         FemaleBachelors = B15001_050E, MaleBachelors = B15001_009E,
         MedHHInc = B19013_001E, MedRent = B25058_001E,
         TotalPoverty = B06012_002E) %>%
  dplyr::select(-NAME, -starts_with("B")) %>%
  mutate(pctWhite = ifelse(TotalPop > 0, Whites / TotalPop,0),
         pctBachelors = ifelse(TotalPop > 0, ((FemaleBachelors + MaleBachelors) / TotalP
op),0),
         pctPoverty = ifelse(TotalPop > 0, TotalPoverty / TotalPop, 0),
         year = "2019", county = "Cook") %>%
  dplyr::select(-Whites, -FemaleBachelors, -MaleBachelors, -TotalPoverty)
```

Wrangling Transit Data and Setting Boundaries

The extracted ACS data pertains to the entirety of Cook County, necessitating an initial step to restrict the tracts solely to the City of Chicago. To accomplish this, the chunk presented below imports a geojson shapefile of Chicago boundary from the official website of the City of Chicago, data.cityofchicago.org. Following several transformations, this shapefile can be utilized to limit the ACS data from the entire Cook County area to solely the City of Chicago.

```
chicago_boundary <- st_read("https://raw.githubusercontent.com/noahba65/Portfolio/main/Chicago-TOD/Data/chicago_city_boundary.geojson") %>%
  st_transform('ESRI:102728') %>% st_sf()
```

```
## Reading layer `chicago_city_boundary' from data source
##   `https://raw.githubusercontent.com/noahba65/Portfolio/main/Chicago-TOD/Data/chicago_city_boundary.geojson'
##   using driver `GeoJSON'
## Simple feature collection with 1 feature and 4 fields
## Geometry type: MULTIPOLYGON
## Dimension:     XY
## Bounding box:  xmin: -87.94011 ymin: 41.64454 xmax: -87.52414 ymax: 42.02304
## Geodetic CRS:  WGS 84
```

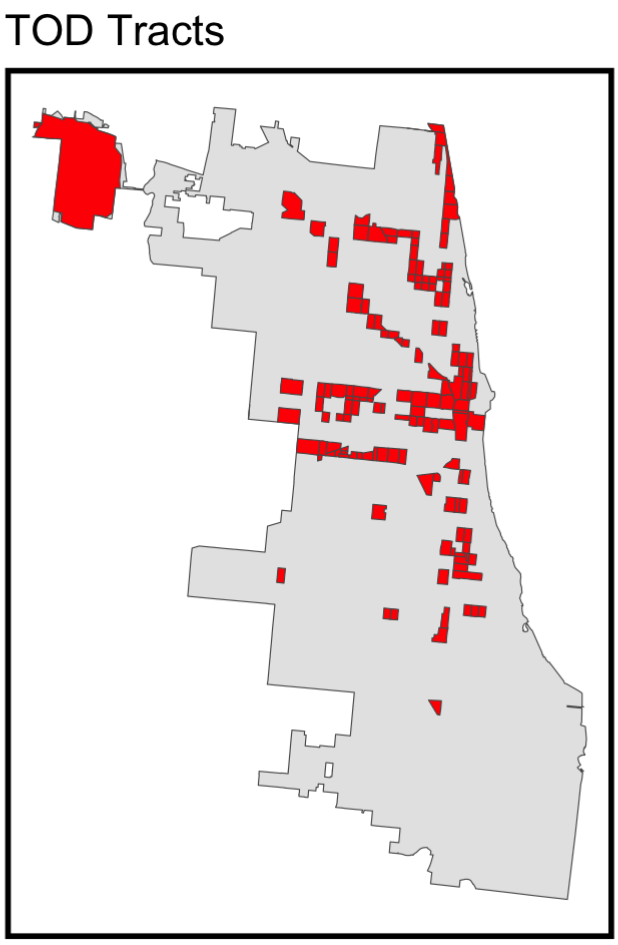
```
# Creating Chicago only ACS tract data for 2009 and 2019 using the st_intersection funciton to only select data and geometries within the Chicago boundary.
chicago09 <- st_intersection(chicago_boundary, cook09)
chicago19 <- st_intersection(chicago_boundary,cook19)
allTracts_chicago <- rbind(chicago09,chicago19)
```

Subsequently, the next phase involves the preparation of transit data for CTA stops. To accomplish this task, the RMarkdown code leverages the transit data made available on the Chicago City Scapes Github page, where a geojson file of the precise locations of each CTA L entrance is retrieved. A quarter-mile buffer surrounding each L stop is then generated, which serves a crucial role in defining the ACS tracts that will be categorized as TOD and non-TOD. Specifically, all tracts whose centroid is encompassed within the quarter-mile buffer are designated as TOD. The ensuing code chunk generates a visual representation of the selected tracts, which is presented in the accompanying map.

```
cta_stops <- st_read("https://raw.githubusercontent.com/ChicagoCityscape/gis-data/master/stations_cta/cta_entrances.geojson") %>%
  st_transform('ESRI:102728') %>% st_sf()
```

```
## Reading layer `cta_entrances' from data source
##   `https://raw.githubusercontent.com/ChicagoCityscape/gis-data/master/stations_cta/cta_entrances.geojson'
##   using driver `GeoJSON'
## Simple feature collection with 381 features and 3 fields
## Geometry type: POINT
## Dimension:     XY
## Bounding box:  xmin: -87.904 ymin: 41.72156 xmax: -87.6058 ymax: 42.07362
## Geodetic CRS:  WGS 84
```

```
# Creating CTA buffer
cta_buffer <- st_union(st_buffer(cta_stops,5280/4)) %>%
  st_sf()

# Uses the st_centroid function to select which tracts are TOD
tract_buffer_chicago <- st_centroid(chicago09)[cta_buffer,] %>%
  st_drop_geometry() %>%
  left_join(dplyr::select(chicago09,GEOID)) %>%
  st_sf()

ggplot() +
  geom_sf(data = st_union(chicago09)) +
  geom_sf(data = tract_buffer_chicago, fill = "red")  +
  ggtitle("TOD Tracts") + mapTheme()
```



Visualizing Indicator Variables

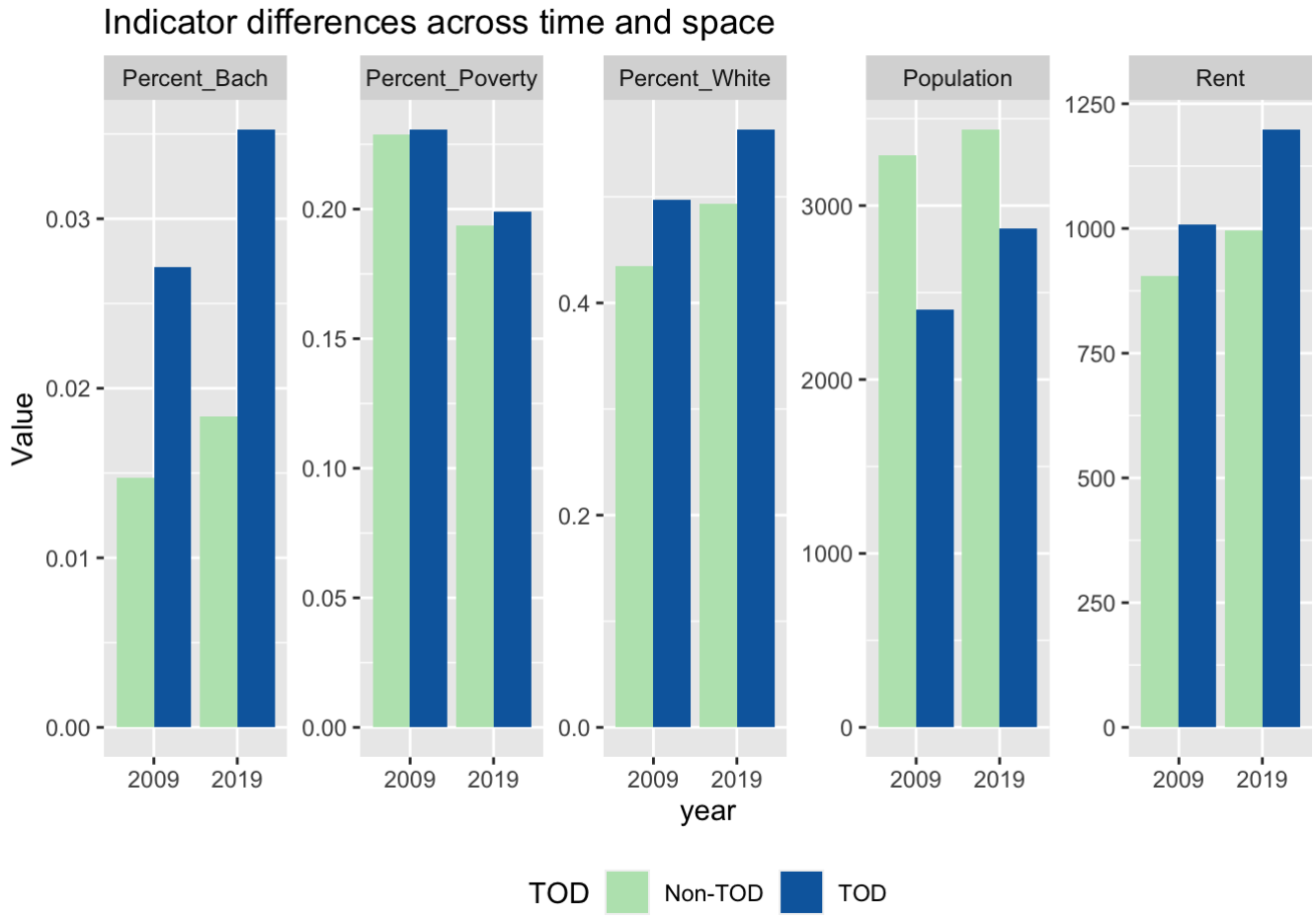
We will proceed by grouping the Chicago data from 2009 and 2019 based on TOD and non-TOD tracts for comparative analysis. Subsequently, the following graphs provide a visual representation of the disparities in the indicators between the two years.

```
# Groups the AllTracts Chicago data frame by TOD and non-TOD
allTracts_chicago.group <-
  rbind(
    st_centroid(allTracts_chicago)[tract_buffer_chicago,] %>%
      st_drop_geometry() %>%
      left_join(allTracts_chicago) %>%
      st_sf() %>%
      mutate(TOD = "TOD"),
    st_centroid(allTracts_chicago)[tract_buffer_chicago, op = st_disjoint] %>%
      st_drop_geometry() %>%
      left_join(allTracts_chicago) %>%
      st_sf() %>%
      mutate(TOD = "Non-TOD")) %>%
  mutate(MedRent.inf = ifelse(year == "2009", MedRent * 1.2, MedRent))
# 2009 rent is adjusted for inflation

# Summarizes the means of the indicator variables
allTracts_chicago.Summary <-
  st_drop_geometry(allTracts_chicago.group) %>%
    group_by(year, TOD) %>%
    summarize(Rent = mean(MedRent.inf, na.rm = T),
              Population = mean(TotalPop, na.rm = T),
              Percent_White = mean(pctWhite, na.rm = T),
              Percent_Bach = mean(pctBachelors, na.rm = T),
              Percent_Poverty = mean(pctPoverty, na.rm = T))

# Creates a cleaner table for summary statistics
kable(allTracts_chicago.Summary) %>%
  kable_styling() %>%
  footnote(general_title = "\n",
           general = "Table 1.2")
```

year	TOD	Rent	Population	Percent_White	Percent_Bach	Percent_Poverty
2009	Non-TOD	905.5855	3288.106	0.4351066	0.0147097	0.2287846
2009	TOD	1008.1401	2402.960	0.4971985	0.0271359	0.2306281
2019	Non-TOD	995.1270	3436.377	0.4933478	0.0183356	0.1936302
2019	TOD	1197.9167	2868.475	0.5632421	0.0352511	0.1989696

Table 1.2

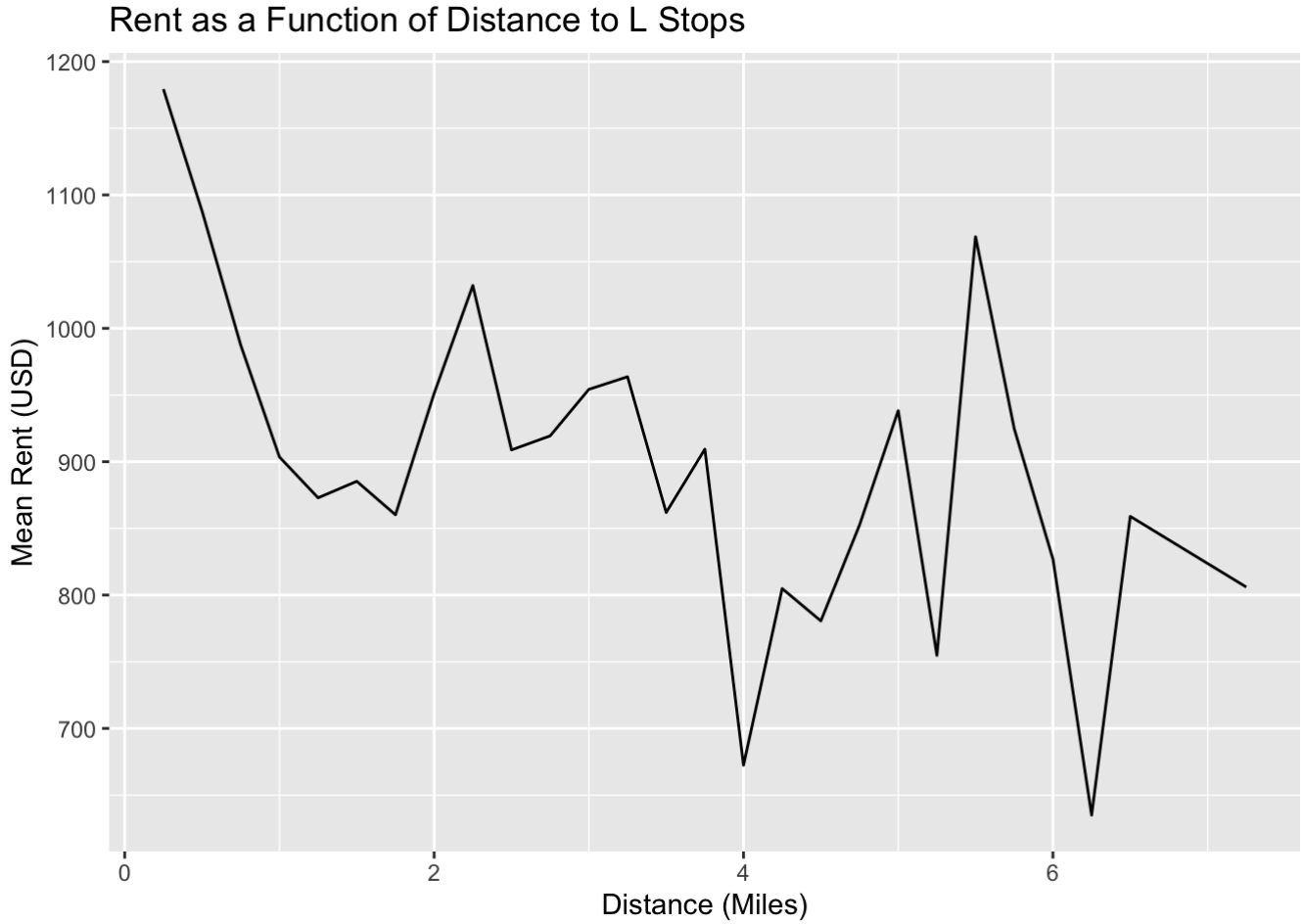
```
# Visualizes the changes in indicators with bar charts
allTracts_chicago.Summary %>%
  gather(Variables, Value, -year, -TOD) %>%
  ggplot(aes(year, Value, fill = TOD)) +
    geom_bar(stat = "identity", position = "dodge") +
    facet_wrap(~Variables, scales = "free", ncol=5) +
    scale_fill_manual(values = c("#bae4bc", "#0868ac")) +
    labs(title = "Indicator differences across time and space")  + theme(legend.position
="bottom")
```



The poverty rates between the two years were comparable, with an overall reduction in 2019. However, the difference in the percentage of residents with a bachelor's degree is striking, with TOD tracts showing a higher concentration of such individuals in 2019. The percentage of white residents in both types of tracts remained largely unchanged, with a difference of approximately 7%. Additionally, the charts demonstrate a significant increase in the disparity of rent prices between TOD and non-TOD tracts from 2009 to 2019. During the ten-year period, the gap in rent prices between the two categories doubled.

To further examine the relationship between rent prices and proximity to train stops, we have created a line plot that records the rent prices at quarter-mile intervals. We used the multiple rings function defined in the source document by Ken Steif and read it into the document with the first chunk, in order to conduct this analysis.

```
allTracts_chicago.rings <-
  st_join(st_centroid(dplyr::select(chicago19, GEOID, year)),
          multipleRingBuffer(st_union(cta_stops), 47520, 5280/4)) %>% # Creates quarter
mile rings around transit stops
  st_drop_geometry() %>%
  left_join(dplyr::select(allTracts_chicago, GEOID, MedRent, year),
            by=c("GEOID"="GEOID", "year"="year")) %>%
  st_sf() %>%
  mutate(distance = distance / 5280) #convert to miles

medRent_Distance_chicago <-  allTracts_chicago.rings %>%
  group_by(distance) %>%
  summarize(mean_rent = mean(MedRent, na.rm = TRUE))

ggplot() +

  geom_line(data = medRent_Distance_chicago, aes(x = distance, y = mean_rent)) +
  xlab("Distance (Miles)") + ylab("Mean Rent (USD)") + ggtitle("Rent as a Function of Di
stance to L Stops")
```



The aforementioned plot depicts a significant increase in rent prices within a quarter mile of L stations, which suggests a high demand for TOD in Chicago. However, beyond the one-mile mark, the relationship becomes less clear and chaotic, indicating that the effect of distance to stops on rent prices diminishes as it is no longer a reasonable walk.

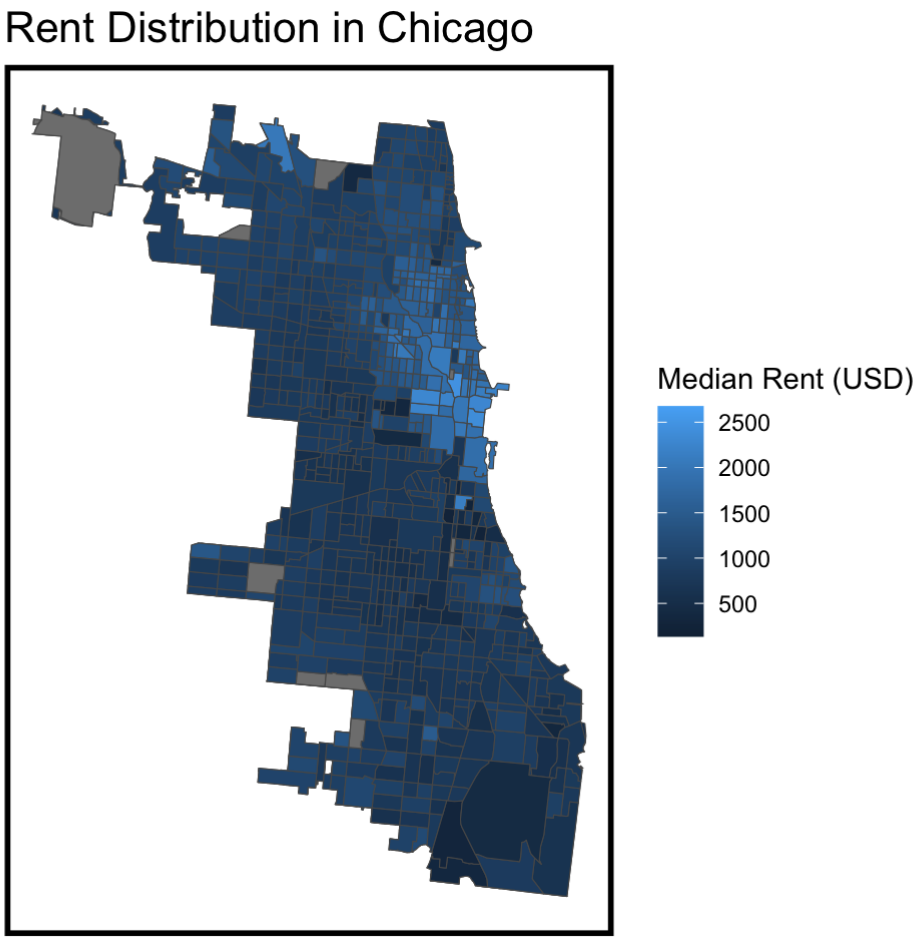
Discussion

While analysis of indicator variables are not conclusive, they would seem to suggest the City of Chicago's ETOD plan has achieved significant success in encouraging the growth of TOD. The population and rent price increasing in TOD tracts in relation to non-TOD tracts indicates increasing demand for TOD in Chicago. While the indicator variables suggest success for TOD in Chicago whether this development is equitable or not could be called into question. White residents continue to be over-represented in TOD tracts suggesting racial inequalities for transit access. While the rising rent prices near transit stations and the increased number highly educated residents indicate success in increasing demand for TOD and perceived attraction to this development style, it presents unique problems in ensuring all socioeconomic groups can benefit from access to public transportation. The City of Chicago should place special emphasis on their efforts to preserve and facilitate the development of affordable housing within TOD zones as laid out in their plan by persevering existing affordable housing and modifying City code to include preference for affordable housing in TOD zones [2].

Limitations

It is important to note some limitations of this analysis. Firstly, the use of indicator variables provides a static view and does not account for statistical significance or variance. Another limitation is the potential for geographical bias within the ACS data. For instance, when examining rent prices, it is clear that the area surrounding Downtown Chicago has substantially higher rents, which could influence the high values associated with TOD areas. It is difficult to discern whether individuals are attracted to TOD areas solely for transit access or for other reasons, such as better job markets or increased amenities. Nonetheless, it is clear that TOD is gaining popularity, and cities nationwide could look to Chicago as a successful example of implementing TOD.

```
ggplot() +
  geom_sf(data = allTracts_chicago, aes(fill = MedRent)) +
  labs(fill = "Median Rent (USD)") + ggtitle("Rent Distribution in Chicago") +
  mapTheme()
```



References

[1] EQUITABLE TRANSIT-ORIENTED DEVELOPMENT. Chicago.gov, City of Chicago, chicago.gov/city/en/sites/equitable-transit-oriented-development/home.htmlchicago.gov/city/en/sites/equitable-transit-oriented-development/home.html.

[2] EQUITABLE TRANSIT-ORIENTED DEVELOPMENT (ETOD) POLICY PLAN. City of Chicago.

[3] Steif, K. (2021). Public Policy Analytics: Code and Context for Data Science in Government. CRC Press.