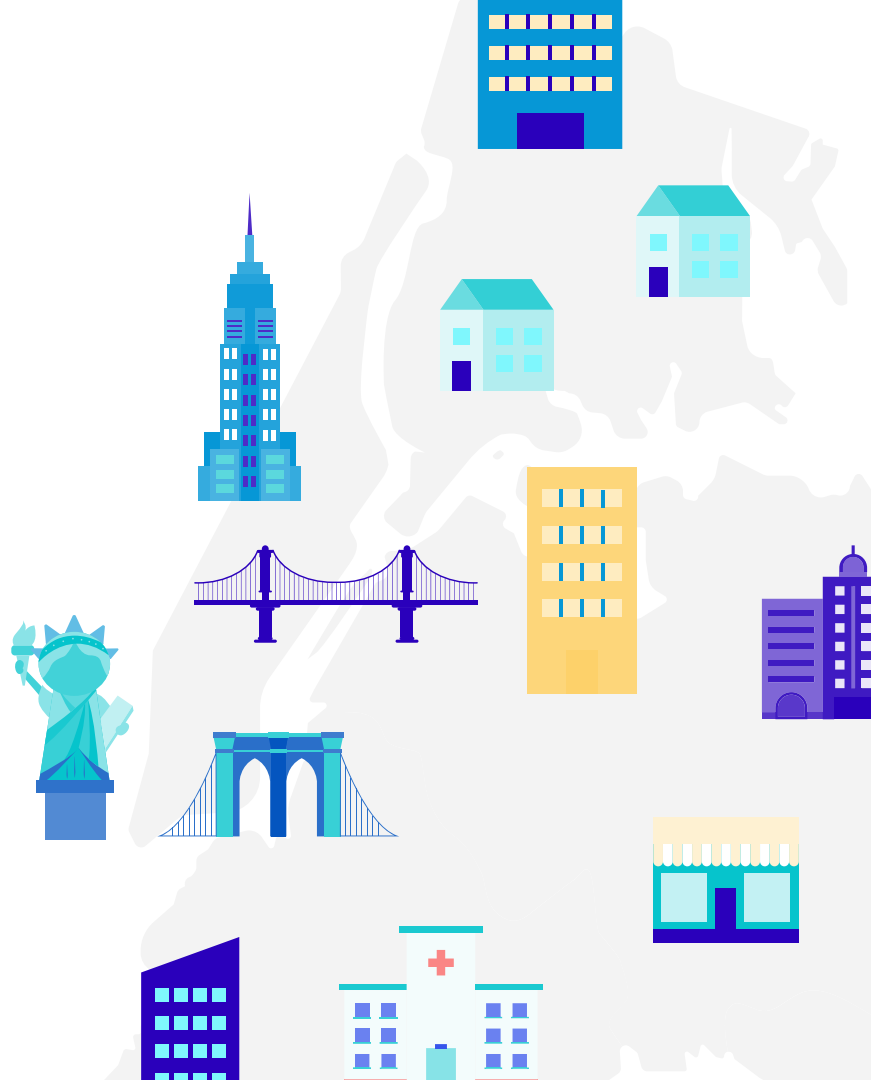


# NYC Limited English Proficient (LEP) Speakers

Quantity of speakers and  
the languages they speak.



# INTRODUCTION

## INSPIRATION

We wanted to better understand the communities that form NYC as it is one of the most diverse cities in the world.

## ISSUE TO RESEARCH

A number of people who live in NYC don't speak English, which can limit their access to different services provided by the City.



## THE DATASET

The dataset contains 136 languages spoken in the 59 community districts of NYC.

## SOURCES

This data comes from the open data portal of the city of New York.

# DATA OVERVIEW



## DATASET 1 (SPEAKERS)

JSON file that contains data about the different languages spoken by Limited English Proficient Speakers in each of NYC's 59 Community Districts and the total number of speakers in each CD.

## DATASET 1 JSON (SPEAKERS)

```
_id: ObjectId('6449df12f030bfe3a19ab90a')
American Community Survey (ACS) Data Time Period: "2015-2019"
Borough: "Manhattan"
Borough Community District Code: "101"
Community District Name: "Battery Park City, Tribeca"
Language: "Gujarati"
LEP Population (Estimate): 12
% of LEP Population: 0.4
CVALEP Population (Estimate): 0
% of CVALEP Population: 0
```

## DATASET 2 (COMMUNITY DISTRICTS)

This dataset contains the coordinates to create the polygons that delimit each of NYC's 59 community districts.

## DATASET 2 GEOJSON (COMMUNITY DISTRICTS)

```
▼ 4: Object
  type: "Feature"
  ▼ properties: Object
    boro_cd: "501"
    shape_area: "377232535.649"
    shape_leng: "153751.224394"
  ▼ geometry: Object
    type: "MultiPolygon"
    ▼ coordinates: Array
      ► 0: Array
      ► 1: Array
```

# OBJECTIVE: A DASHBOARD WITH THE FOLLOWING ELEMENTS

## DYNAMIC

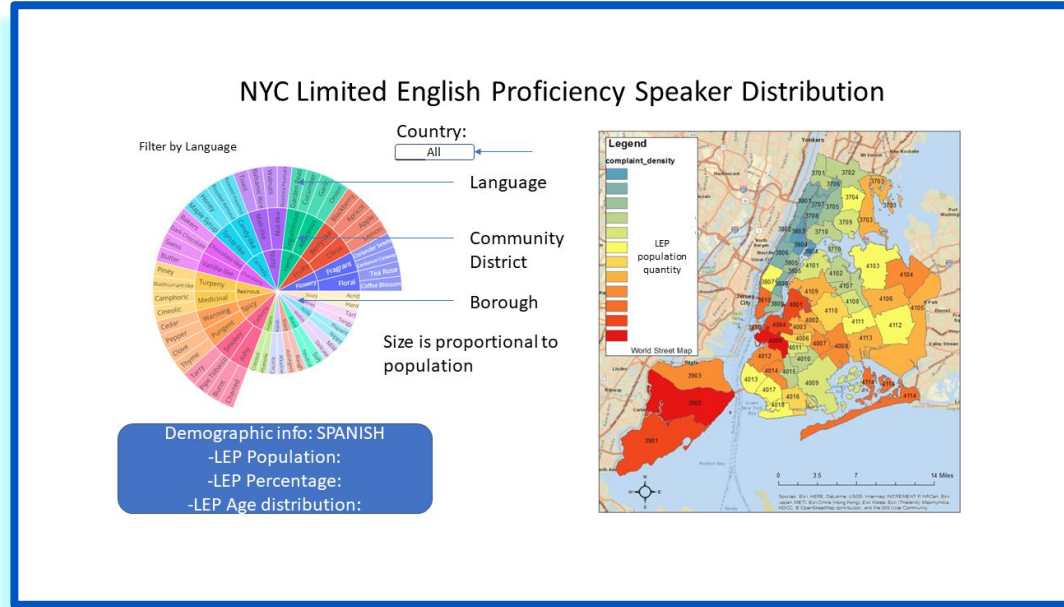
Our user must be able to see and understand the data for any desired language.

## DROPDOWN MENU

For the user to select the language.

## SUMMARY SECTION

A section to summarize the sub-dataset selected by the user.

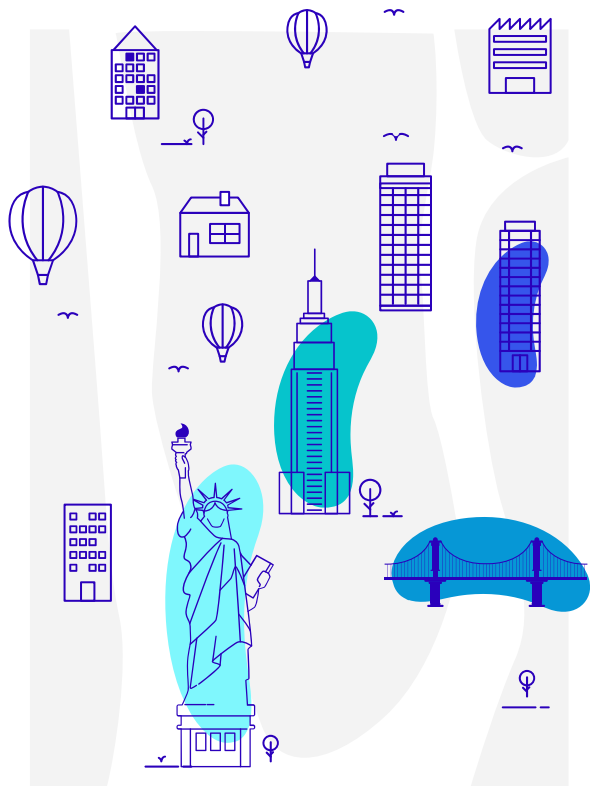


## SUNBURST CHART

A pie-like chart that changes dynamically according to the user interaction.

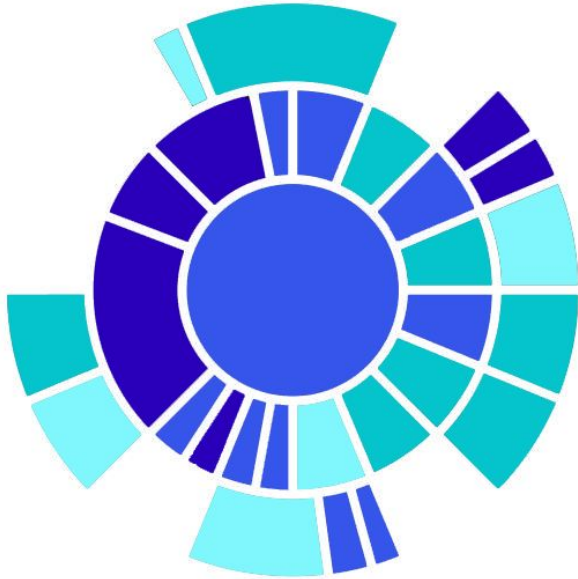
## CHOROPLETH MAP

A map of NYC and its community districts that change color according to the total number of speakers of the selected language in that area.



**DEMO  
TIME**

# MAJOR CHALLENGES FOR GRAPHS



## SUNBURST DATA

The sunburst needed three rows of data. One containing the labels to show on it, one to id the sections and one to specify what the id of the section's parent.

I. e.    L1:    Parent: NYC,  
                  ID: BrooklynID  
                  Label: Brooklyn

          L2:    Parent: BrooklynID  
                  ID: 308  
                  Label: Crown Heights North

          L3:    Parent: 308  
                  ID: Spanish\_308  
                  Label: Spanish

## LAYOUT

We added a `<br>` replacing “,” from the JavaScript file to place sunburst and bar graph info into various rows for better readability

# 4 DIFFERENT API ROUTES FOR OUR GRAPHS:

## [/demographic\\_all](#)

Provides the info for the summary box when “All” is selected.

## [/demographic/<language>](#)

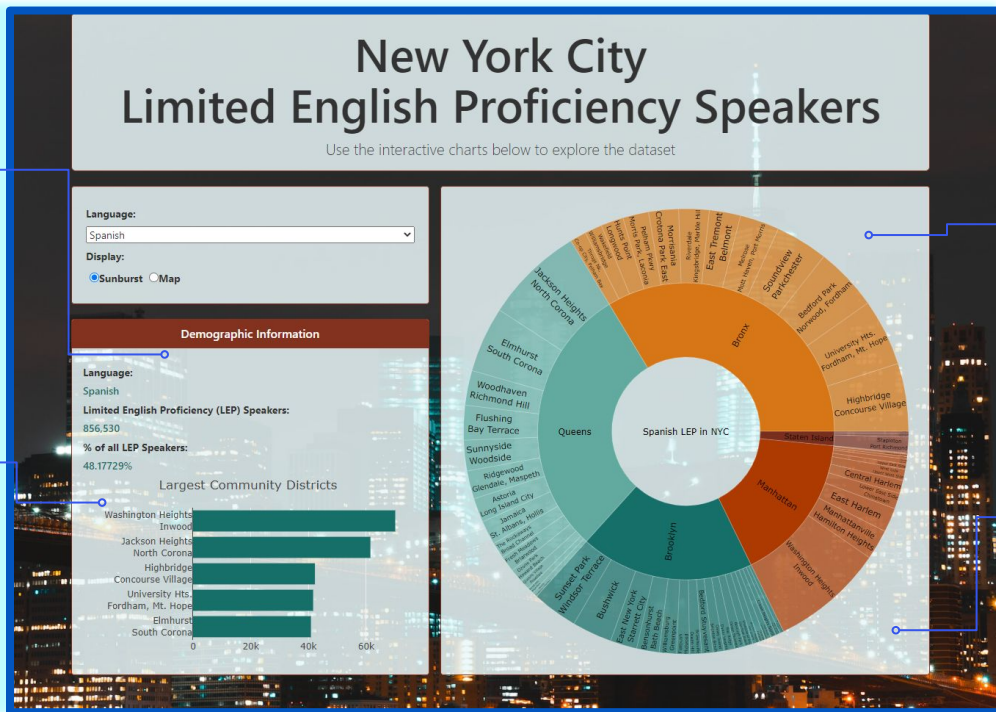
Provides the info for the summary box when a specific language is selected.

## [/populations\\_all](#)

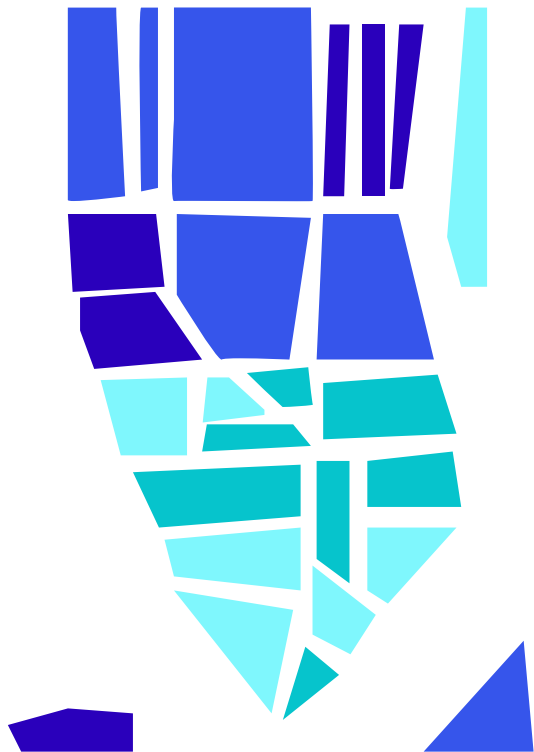
Provides the data for the sunburst chart when “All” is selected.

## [/populations/<language>](#)

Provides the data for the sunburst chart when any specific language is selected.



# MAJOR CHALLENGES FOR THE CHOROPLETH MAP



## MAP DATA

The data we needed for the map wasn't complete (we only had the coordinates, but the data about the speakers was missing).

To solve this we decided to add the properties we needed from the speakers file to the coordinates file using Python.

## HTML

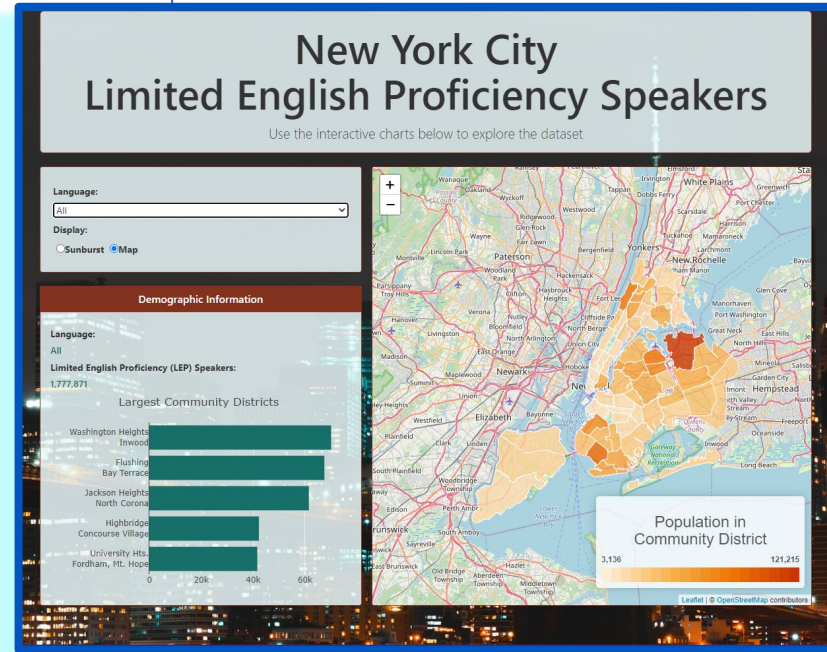
It was challenging to set up the layout to be able to display both the chart and the map in the same page.



## 2 DIFFERENT ROUTES FOR OUR CHOROPLETH MAP:

`/communities/<Language>`

This path filters the speakers JSON for only the specified language and then completes the merge with our GEOJSON coordinates file.



`/communities_all`

This path groups our JSON by community district and sums all of the Limited English Proficiency Speakers for each of the community districts and then completes the merge with our GEOJSON coordinates file.

# FINDINGS



## COMMUNITIES TEND TO CLUSTERIZE

Immigrants from different communities live in different parts of the city and they tend to form clusters according to their place of origin.

## A SHARED LANGUAGE DOESN'T MEAN SOMEONE BELONGS TO THE SAME COMMUNITY

Despite speaking the same language, the clusters can form in different parts of the city.

## QUEENS IS THE BOROUGH WHERE MOST LEPs LIVE

With 35% of them living there – followed by Brooklyn (30%). Therefore, efforts to make City services more accessible should be focused there.