

The background is a detailed, isometric illustration of a neighborhood. It features a variety of buildings, including small houses with different colored roofs (red, orange, green), larger apartment complexes, and industrial buildings with smokestacks. There are also cars, buses, and trucks. Landmarks include a large clock, a rainbow, a cow, and several trees. The scene is set against a dark, textured background with faint, repeating patterns of the neighborhood elements.

# Recommending the best neighborhoods

# Recommending neighborhoods

## Relocating...



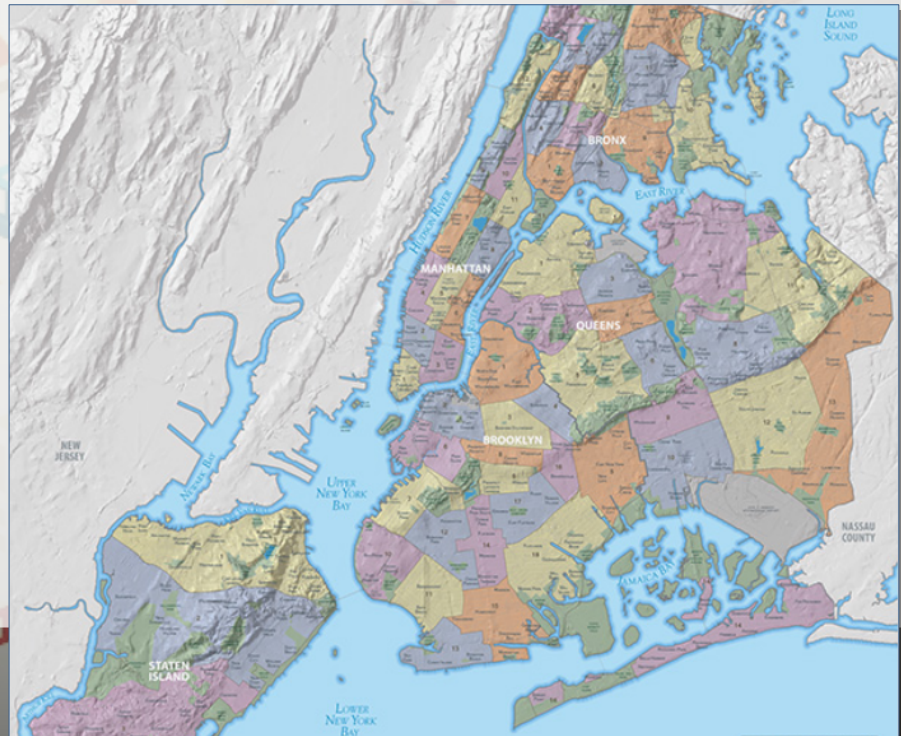
- How to choose the **best neighborhood**

Depends on pricing...  
But also on lifestyles

- Will the customer like the surroundings ?

# Data sources - Geographical

- First need, an **administrative division**
  - Neighborhoods is the usual one
  - From city planning databases (geographical data)
- In the example,  
New York





# Data sources - Categories

- Second, interesting facilities
  - From **Foursquare API**  
(venues search in each area)
    - Coordinates are needed
  - The most important:  
**representative categories**
    - ✓ Describing the neighborhood



# Data sources - Preferences

- **Characteristics**

- Designed features
    - ✓ To define customer lifestyle
  - More can be added, if needed
    - Genre, ethnic group, urban tribe, ...
  - And also granularity
    - Type of music, sports or pets, religion, ...
- Boolean type (yes/no)
    - But 1-0 encoding is useful
      - ✓ For multiplication (zero means no contribution)

YOUNG

MATURE

OLD

SINGLE

MARRIED

CHILDREN

PETS

MUSIC

SPORTS

CULTURE

FASHION

GOING OUT

HEALTH

RELIGIOUS

...

# Data sources - Segmentation

- Preferences in the neighborhood

<b>Categories</b>
Donut Shop
Yoga Studio
Animal shelter
Beach
Community Center
Art Museum
...



<b>Characteristics</b>	
YOUNG	MUSIC
MATURE	SPORTS
OLD	CULTURE
SINGLE	FASHION
MARRIED	GOING OUT
CHILDREN	HEALTH
PETS	RELIGIOUS



## Neighbours segmentation

Which categories are preferred depending on people characteristics?

- ✓ A survey is needed...
- ✓ Same Foursquare categories

- Matrix categories - characteristics
  - ✓ Also Boolean, with 1-0 encoding



# Analizing categories

## 1) Calculate representative categories

- From Foursquare venues information
- By number of venues
  - ✓ Describing the neighborhood lifestyle (places and people)

- Methodology:

- Group by category (counting venues)
- Sort by total number (descending)
- Selecting the **top categories** (by default, 10)
  - ✓ The ones with more venues in the area

# Mapping preferences to categories

## 2) Multiplying characteristics

- Preferences vector (characteristics) and
- Segmentation matrix (categories vs characteristics)
  - ✓ For every row, value by value

- Result: 248 rows with unique categories

- Columns with zero, **no contribution to scoring**
- Disabling a characteristic for all categories
  - ✓ Not selected by the customer
- Others are zero from neighbors' preferences
  - ✓ Not selected in segmentation survey



# Calculating neighborhood scores

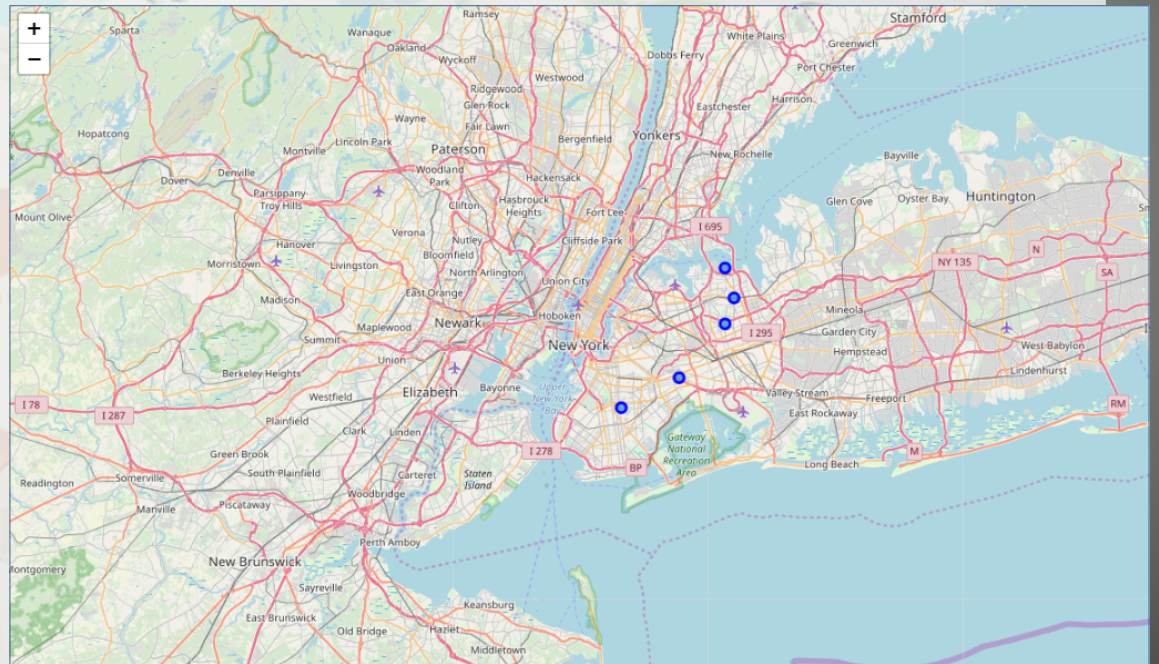
## 3) Calculating scores (for all neighborhoods)

- a) Getting a category multiplying factor
  - Only for the top representative categories
  - Adding all columns for every category
    - ✓ From the previously calculated matrix
- b) Adding a weight by position
  - The more representative, higher contribution
- c) Adding scores for all the top categories

# Presentation

## 4) Drawing on a map

- Selecting top scores
  - Just 5 by default
  - Adding coordinates
- Labels with position
  - And final score



# Results

- Scores are well discriminated
  - For preferences
    - ✓ A zero value disables the contribution
  - For every category
    - ✓ Only flagged characteristics contribute
    - ✓ Weighting pushes the more representative
- Parameters allow full control
  - Foursquare: Radius, number of venues
  - Algorithm: Selected categories and neighborhoods
- Graphical representation enhances experience
  - Map built dynamically
  - But the full list of scores is available



# Conclusions

- Recommendation algorithm
  - Consistent, predictable
- Adaptable for any city and country
  - Administrative division data is needed
    - ✓ May be different (villages, streets, postal codes, ...)
  - Segmentation data (preferences) is needed
    - ✓ Lifestyles may change from place to place
    - ✓ Specific surveys are required
  - A cities repository could be very useful
    - ✓ For geographical and preferences files

# Future directions - I

- More places
  - Adding cities to the repository
  - Other countries
    - ✓ Source data may be inaccurate or difficult to get
- More characteristics
  - Adding new groups (i.e. students, workers, etc.)
  - Increasing granularity (i.e. genre, hobbies, etc.)
- Other geographical divisions
  - Smaller area: Streets
  - But careful with search radius
    - ✓ Small radius may result in very few venues
    - ✓ And bigger, in similar scores for adjacent areas

# Future directions - II

- Recommendations can be prepared in advance
  - Building a library
    - ✓ With pre-calculated scores
  - Based on specific customer profiles







**Recommending**  
the best  
**neighborhoods**