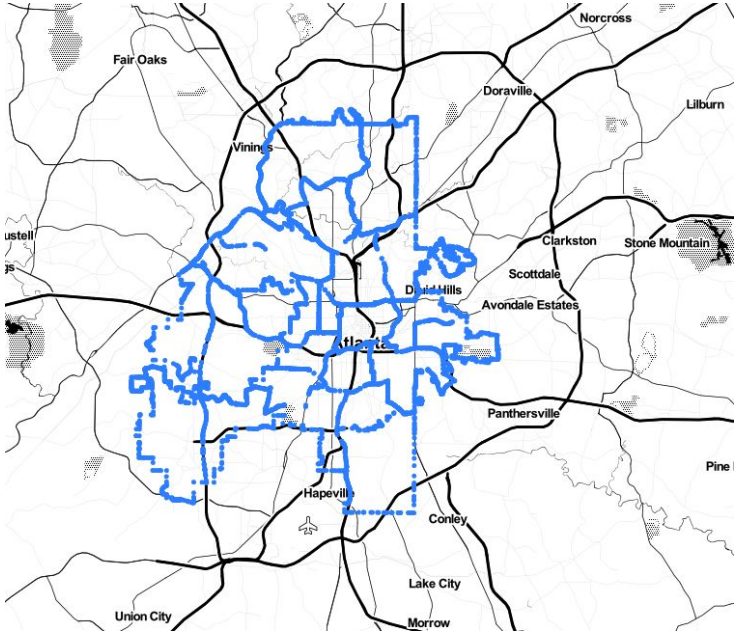
Exploring Atlanta Neighborhoods

Using Machine Learning to Assess the Validity of Administrative Districts

Katie Fullerton

Background

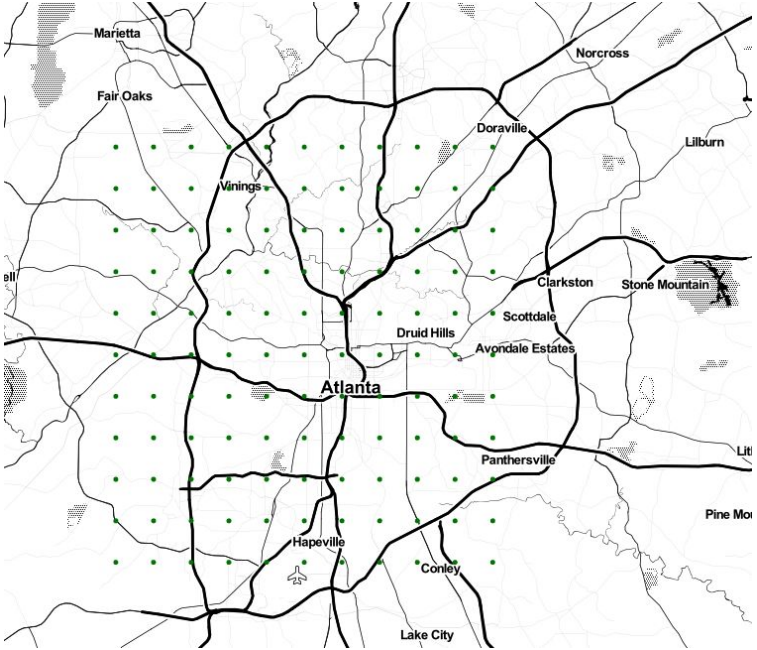
- The City of Atlanta is divided into a number of Neighborhood Planning Units (NPUs) for administrative purposes
- Do these NPUs reflect the way citizens actually use their neighborhoods?
- Realistic administrative boundaries increase civic engagement
- Accurate zoning information is useful to developers



NPU Boundaries in the City of Atlanta

Search Grid

Due to limited number of results for Foursquare API calls, we constructed a search grid that covered the areas mapped in the NPU data



Data Sources

City of Atlanta's GIS System

- Retrieved via 1-time API call
- Contains latitude and longitude of boundary points for each NPU

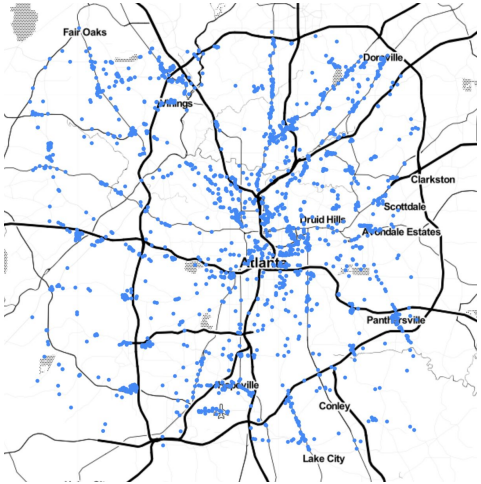

```
{ "attributes": {
        "OBJECTID": 260,
        "LOCALID": null,
        "NAME": "K",
        "GEOTYPE": "NPU",
        "FULLFIPS": null,
        "LEGALAREA": null,
        "ACRES": 1528.29,
        "SQMILES": 2.39,
        "OLDNAME": null,
        "NPU": null
      },
      "geometry": {
        "rings": [
          [
            [
              -84.4173772073577,
              33.772197013770004
            ],
```


Foursquare Venue Data

- Constructed API call for each latitude and longitude point on the grid
- Contains data on location name, category, latitude and longitude
- Data extracted from structured json to dataframe for processing

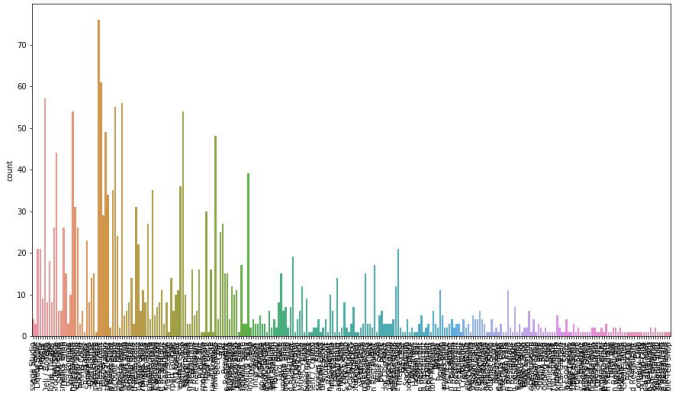
	name	categories	latitude	longitude	neighborhood
0	Zen Massage	Massage Studio	33.666481	-84.549732	NaN
1	Tom Lowe Trap & Skeet Range	Gun Range	33.671201	-84.564226	NaN
2	Camp Creek World of Beverages	Liquor Store	33.657393	-84.511874	NaN
3	Piece of Cake	Bakery	33.656218	-84.513946	NaN
4	Wolf Creek Amphitheater	Theater	33.674711	-84.567392	NaN

Exploratory Data Analysis



Results of Foursquare API calls over whole search grid.

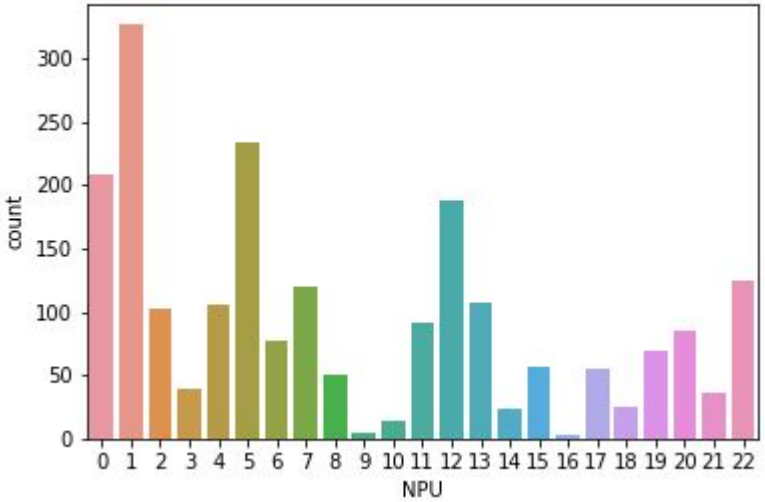
This view confirmed that the search grid method effectively spaced the results.



Plot of the categories listed in Foursquare data vs. number of locations.

From this view, it is clear that there are too many categories with one or two locations for that information to be useful in this analysis.

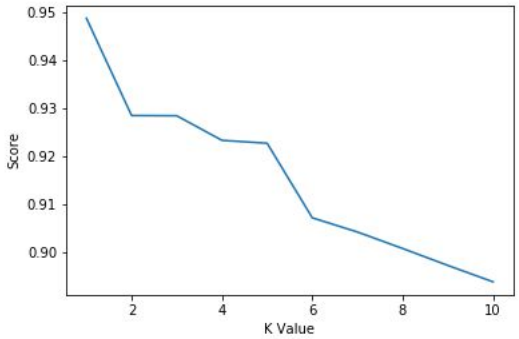
Exploratory Data Analysis cont.



- Each location was assigned to an NPU by assessing which boundary point it was closest too
- The plot shows the number of venues in each NPU
- This information was used to determine the number of cross validation folds used in model training
- 2 NPU’s have only 3 locations, which limits the number of cross validation folds to 3.
- These 2 NPUs could be excluded to increase the model accuracy in future work.

Model

- SKLearn KNeighborsClassifier
- Used cross validation score to determine optimal k value
- Plot does not show characteristic elbow, but does have 2 points where slope sharply decreases
- Selected k = 4 for analysis



Model inputs:

Latitude and longitude of each location

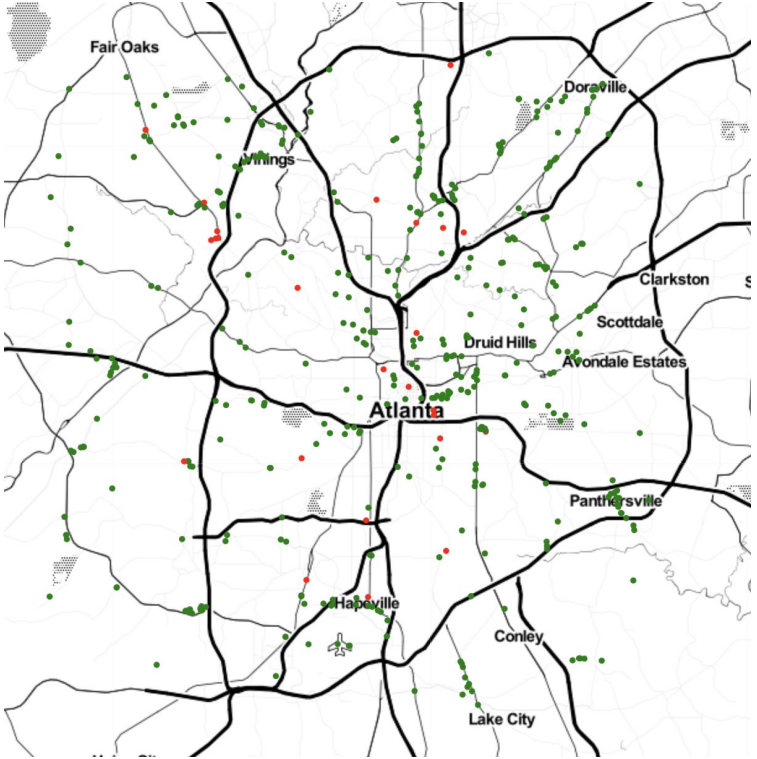
Model labels:

Encoded assigned NPU

Train/test split: 80/20

Results

- Using sklearn's scoring function, score = 0.9414
- To visually analyze, each point in the test set was assigned a color:
  - Green = model predicted the same NPU as the original assignment
  - Red = model prediction did not match NPU assignment
- Majority of test points were correctly assigned
- Red points likely fall along NPU boundaries- further analysis beyond scope.



Discussion

- it appears that the K Neighbors Classifier performs fairly well
- The results of using the model to predict a neighborhood assignment corresponds well with the NPU assignment described above.
- This would suggest that the NPU boundaries do still reflect actual use patterns.
- There are a number of assumptions that were used to simplify this analysis,
  - the method of assigning a venue to an NPU
  - the method of assessing prediction accuracy visually
- The correspondence between the assigned and predicted NPU does suggest that the NPU system is not as outdated as was originally assumed.
- This analysis could be useful for city officials, citizens, and developers.