Portfolio

Spatial Data Analytics / Urban Planning

https://annannzhang.github.io/
annzhangzian00@outlook.com

Ann Zi'an Zhang

Ann Zi'an Zhang

Python, R, SQL, JavaScript, HTML, CSS, Figma
ArcGIS Pro + Online, Excel
Adobe InDesign, AI, PS

SKILLS

Master of Urban Spatial Analytics &
Master of City Planning
University of Pennsylvania
2022-2024

B.A. in Art History, Psychology, and
Science in Society, with High Honors
Wesleyan University, CT
2018-2022

EDUCATION

Economics + Advisory Intern, AECOM,
New York, NY
May 2024 - Current

Summer Geospatial Data Analyst, GeoAdaptive
Boston, MA
Jul 2024 - Aug 2024

Planning Intern, WRT
Philadelphia, PA
Jan 2024 - May 2024

EXPERIENCE

Table of Contents

Girl Power in Transportation Geography

Background & Methodology This capstone project aims to reveal unique contribution of female-identifying researcher-led research in the field of transport geography. The project focuses on Journel of Transport Geography, a well established academic journal in this field. Three approaches were taken for topic analysis on female vs. non-female led research articles - 1) keyword frequency counts, 2) BERTopic, and 3) logistic regression. The results show that female-led research help explore agencies' subjective experience of transit, and gender and social issues, in comparison to the objectivity and physicality of transportation that non-female research tend to focus on.
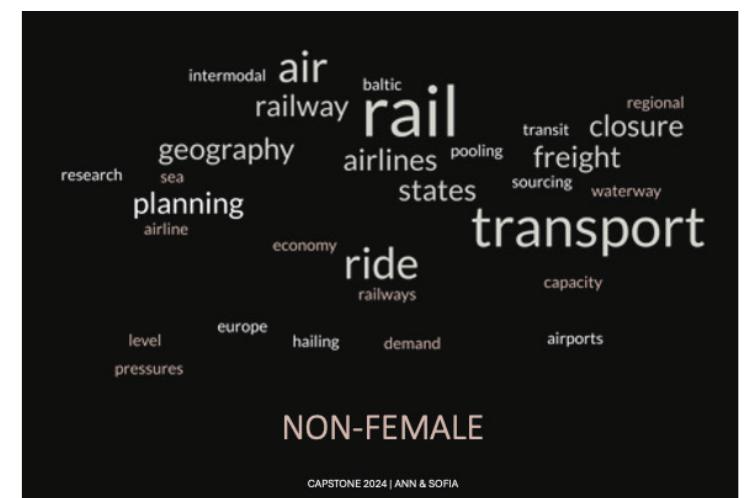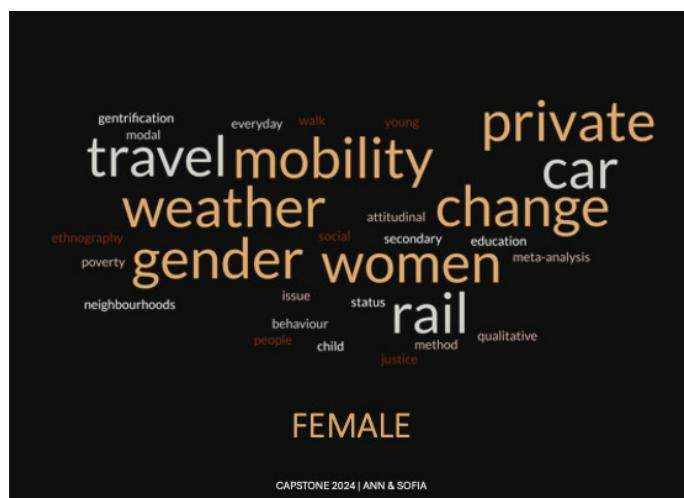
Since the journal's establishment in 1995, we are witnessing an increasing trend of both number of overall articles and number of female-led articles. Among all articles, female-led research articles have an average page count of 9.22 pages (comparing to 9.13 in non-female), an average time cited at 40.99 (comparing to 39.59 in non-female), and an average number of authors at 2.89 (comparing to 2.65 in non-female). Overall, the stats demonstrate an increasing trend in female researchers' involvement and a high quality of their work.

The composite results from 3 approaches are synthesized and summarized in word cloud on the right. In conclusion, female-led research articles are bringing in human-centric perspectives, qualitative research methods, and attention to subjective experiences of agencies traveling into the field of transportation geography, which would otherwise be missing, as non-female researchers tend to focus more on the physical infrastructure and larger network systems as opposed to individual's experience.

4

PART I
Keyword Counts

FEMALE	NON-FEMALE
Mobility	Accessibility
Urban	Urban
Accessibility	Spatial
Public	Model
Analysis	Analysis
Areas	Public
Spatial	Network
Different	Time
Research	Areas
Network	Transportation

The results show an overlap between the two groups regarding frequently occurring keywords, such as urban, accessibility, and spatial. The more unique ones such as mobility, different, and research, are not too useful for distinguishing content that are unique.
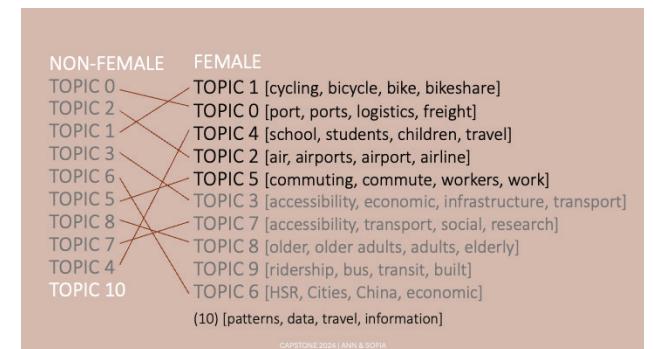
PART II
BERT Topic Identification

The intra-group difference is compared by counting the number of topics in each group. While we are witnessing some overlaps in top ranked topics, we start to see trends such as Topic 4 (school, students, children, travel) ranking higher in the female group and then non-female group.

Topics identified using BERTopic Model

Comparison of Topics for 2 groups

PART III
Logistic Regression

According to results, female-led articles look at subjective experience of transit, specific agencies such as children, students, or young people, and utilizing less-commonly used methodology including attitudinal survey, qualitative analysis, and ethnographic studies. Comparing to non-female groups' focus on the physicality of transit and networks, female-led articles are particularly meaning in bringing in human-centric perspectives and qualitative aspects.

1. Change	1. Railway	1. Walk	1. Governance
2. Everyday	2. Pressures	2. Ethnography	2. Light Rail
3. Mobility Practices	3. Closure	3. Young People	3. Location
4. Behaviour	4. Sea Level	4. Content Analysis	4. Airlines
5. Modal	5. States	5. Quality of Service	5. Air Transport
6. Private	6. Rail Freight	6. Information and Communication Technologies	6. High Speed Rail
7. Practices	7. Rail / Air	7. Travel Diary	7. Networks
8. Neighbourhoods	8. Capacity	8. Motility	8. Globalization
9. Gentrification	9. Railways	9. Mobile methods	9. Developing Countries
10. Weather	10. Demand	10. Inequalities	10. Regulation

Regression on BERT Topics

Regression on Author Keywords

1. Secondary Education	1. Waterway Transport
2. Child	2. Regional Economy
3. Women Status	3. Planning Process
4. Gender Issue	4. Model Test
5. Meta-analysis	5. Infrastructure Planning
6. Qualitative Analysis	6. Sensitivity Analysis
7. Poverty	7. Distribution System
8. Survey Method	8. Uncertainty Analysis
9. Attitudinal Survey	9. Classification
10. Social Justice	10. Service Sector

Regression on Index Words

# All About Arts

Python | (Quarto) https://annannzhang.github.io/Python-Projects/
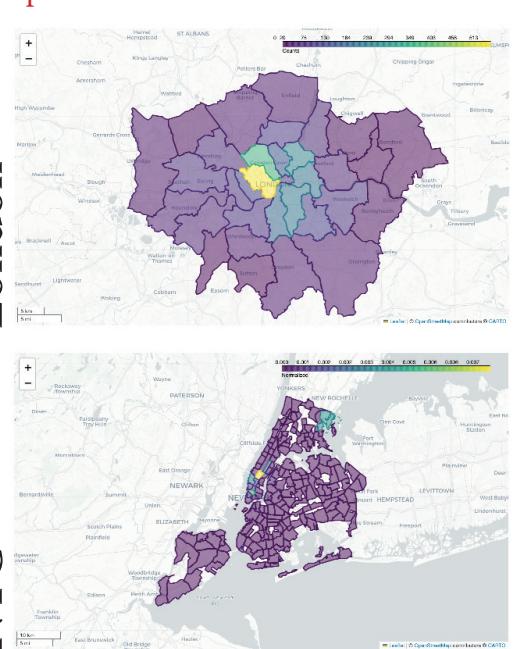
This project is the final project for Geospatial Python class. The project consists of three parts that explore the arts and cultural infrastructure in New York City and London, utilizing different types of analysis in python.
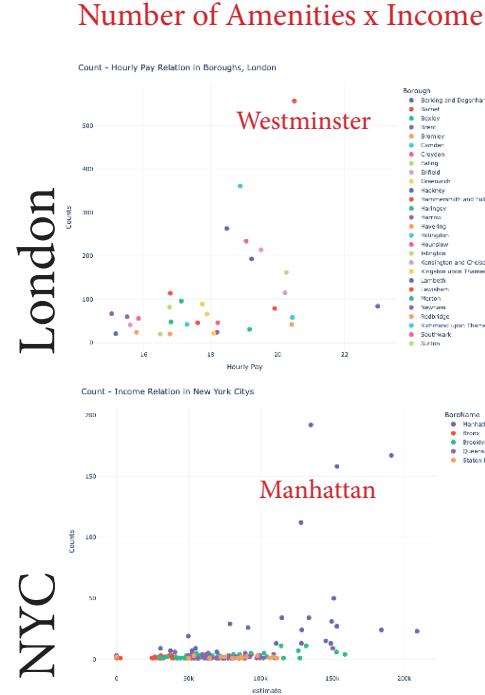
Part I - Arts & Culture in New York & London

Key Findings: Both in NYC and London, we are seeing a disproportionally higher density and more variety of art and cultural infrastructure in city centers. However, interestingly, there isn't a clear positive linear relationship between household income and number of art and cultural amenities – it seems like lower income communities also have access to a decent number of infrastructures, through incomparable to the diversity present in city centers. The most famous and largest museums, including but not limited to the Metropolitan Museum of Art, Modern Museum of Art, and the British Museum, are all located in more of the central area.
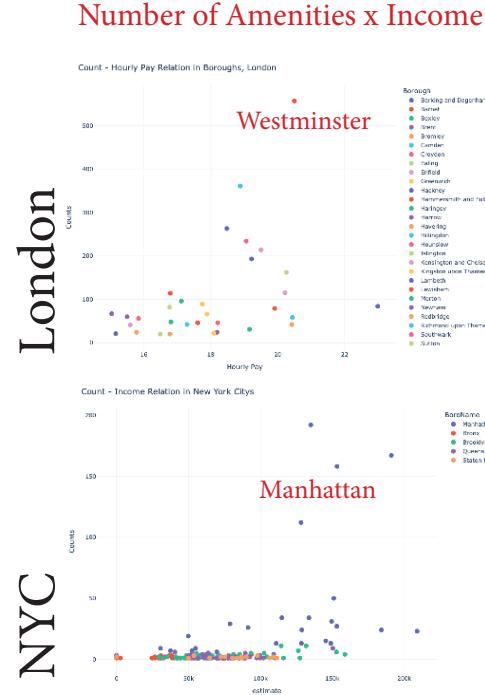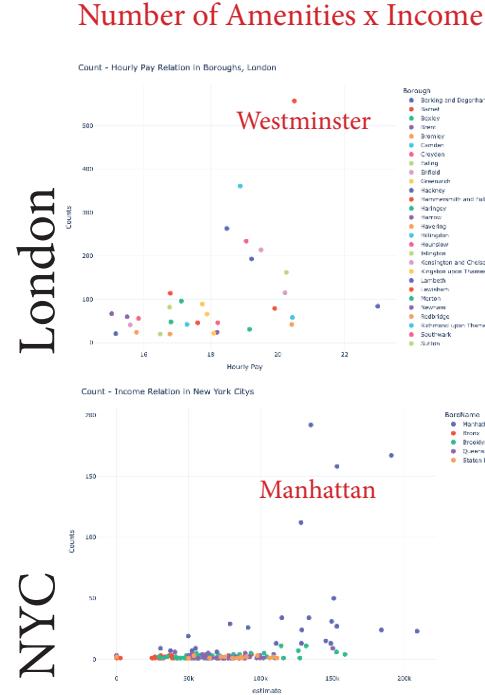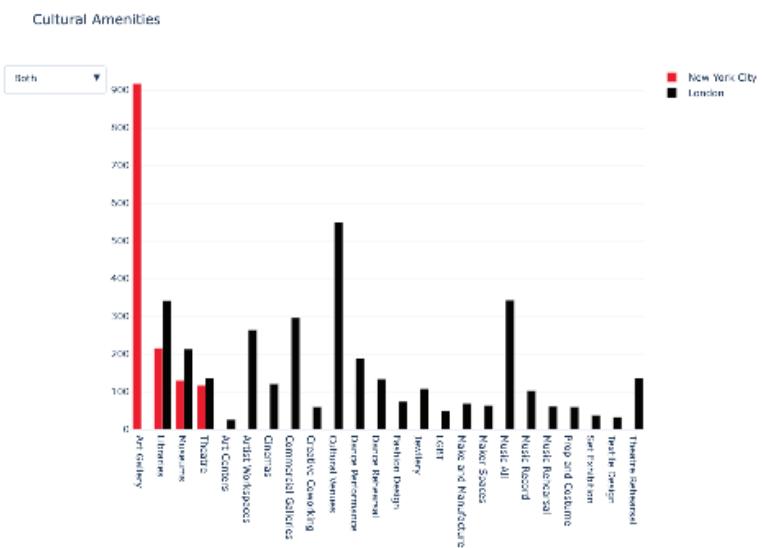

Spatial Distribution of Cultural Amenities


Number of Amenities x Income
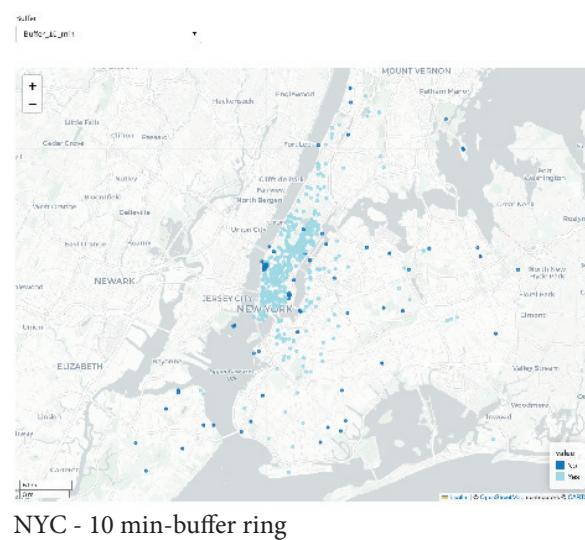
Comparison: Number of each type of arts and cultural amenities

Part II - Public Transit Access and Street Network Analysis

The NYC dashboard shows the amenities that are located within 2-minute, 5 minutes, and 10-minute walkshed buffering, and the London one on the other hand shows 5-minute, 10 minutes, and 15-minute buffers. This decision is made partially due to the fact that the Greater London area is a lot larger than NYC. We can see that most NYC amenities falls within the 10-minute walking buffer and most London ones fall within 15-minute buffer, showing their great accessibility.

NYC - 10 min-buffer ring
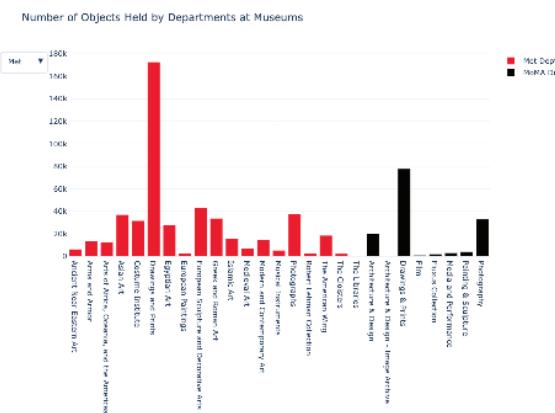
Another potential way to judge their accessibility to outside travelers is through OSM network analysis. Here, I'm using JFK airport and the cultural centers in Manhattan as an example. JFK airport serves as a transsit hub and hence is the the node of origin in this analysis. Sample result is shown in the table below.

|	Name	Address	geometry	shortest_travel_time	shortest_travel_time (in Minute)
0	Alexander Hamilton U.S. Custom House	1 Bowling Grn	POINT (-74.01376 40.70382)	739.8	12.330000
2	American Academy of Arts and Letters	633 W. 155th St.	POINT (-73.94730 40.83385)	982.3	16.371667
3	American Folk Art Museum	45 West 53rd Street	POINT (-73.97810 40.76162)	1043.4	17.390000
5	American Museum of Natural History	Central Park West at 79th Street	POINT (-73.97365 40.78083)	1040.7	17.345000
6	American Numismatic Society	75 Varick St	POINT (-74.00701 40.72353)	703.7	11.728333

Part III - Museum Collections: the MET, MoMA, and the British Museum
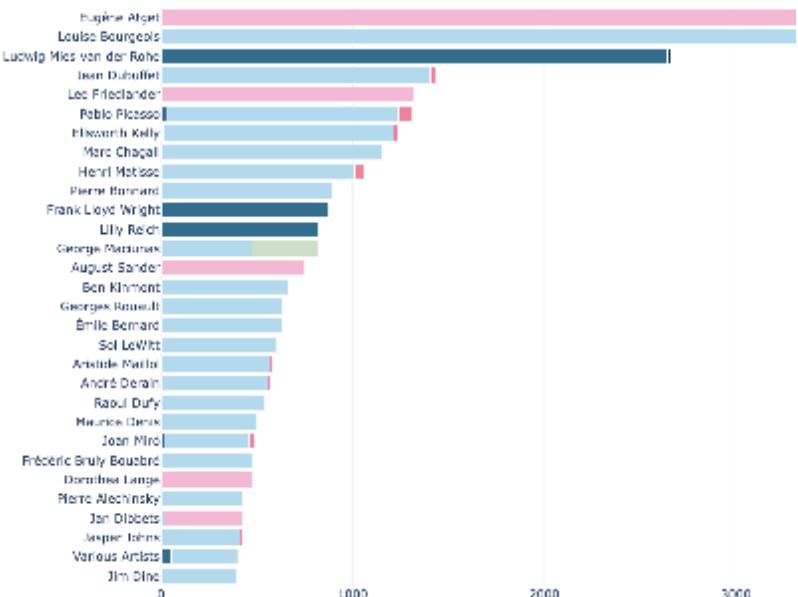
This part of the project zooms into each museum's collections. By utilizing the museums' digital collection such as public dataset by the MET and MoMA on their dedicated github page, or by web scrapping from museum's digital collection result page, I made some interesting analysis of largest collections, top artists, top nationalities, and the most frequent themes of their traditional Chinese artwork collections. This part of the projects aims to explore some possible interesting directions of digital humanities and digitization of museum archives, which has been the trend in this field for the past two decades.

In-Class Projects

Python | (Quarto) https://annannzhang.github.io/Python-Projects/

1 - Street Network Analysis - Crash Data

This project utilizes osmnx (Open Street Maps' networr analysis) to explore spatial patterns of car crash incidence happened in Center City in Philadelphia, PA. The network graph created from OSMnx (Fig. 1) is converted into GeoDataFrame firsts, then used for projecting the crash index geographically onto the an interactive folium map. As the results show in Fig. 2, highways like I-676 Vine Street Expressway has very high crash index in comparison to the rest of the city. This visualization is helpful for advocating for less highways for better travel safety in transportation planning.

Fig 1 - Street Network Graph from OSMnx
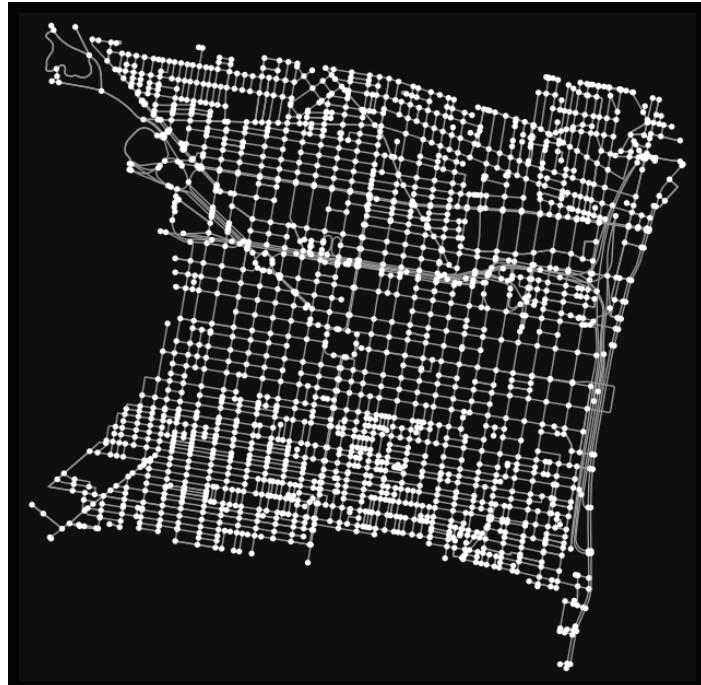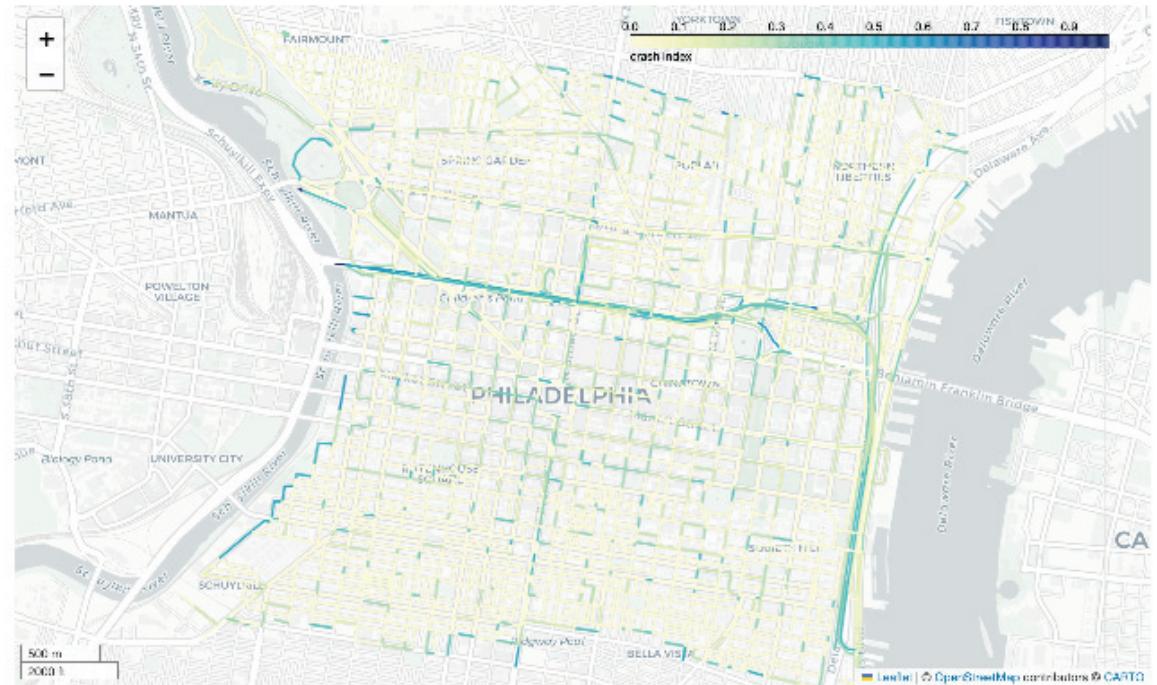
Fig 2 - Crash Index on Street Network
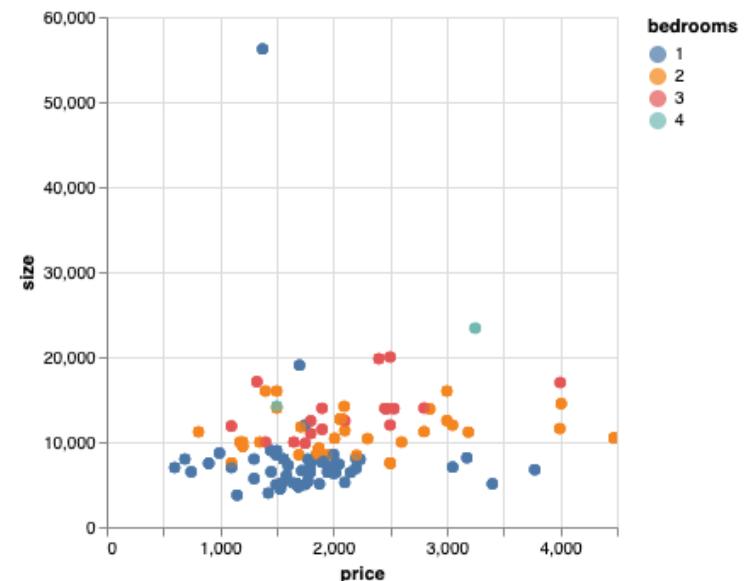
2 - Webscrapping

In this short scrapping project, I used BeautifulSoup to scrap 5 pages of results of Craiglist's Philadelphia region apartments. The raw results are cleaned and organized into the table below. The data was then uesd to make analysis of the apartment market stocks at the time, of which Fig 4 is one of the findings.

Fig 3 - Scrapped Results Organized in Table

|	price	size	bedrooms	title	pricepersq
0	2474	13942	3	3bd 2ba, Serene Wooded View, West Chester PA	0.177449
1	2850	13902	2	Apartment by the arts!	0.205006
2	1800	7592	1	A Cozy Living Space In Rittenhouse Square.	0.237092
3	1750	5002	1	Pets welcome in this Beautifully renovated apa...	0.349860
4	1950	8502	2	Enjoy a fantastic living space with tons of na...	0.229358
...	...	...	...	...	...
595	2075	12712	2	Resort-Style Swimming Pool, Extra Storage, 2/BD	0.163232
596	1876	5072	1	On-demand car wash/detailing, Handyman and mai...	0.369874
597	1680	5092	1	Penthouse Hideaway, LVT Flooring, Bike Storage	0.329929
598	2199	6922	1	Fire Pit, On-site Management/Maintenance, Flex...	0.317683
599	1544	5182	1	1 BD, Bike Storage, Washer/Dryer	0.297954

Fig 4 - Apartment Size-Price-Bedroom Relation
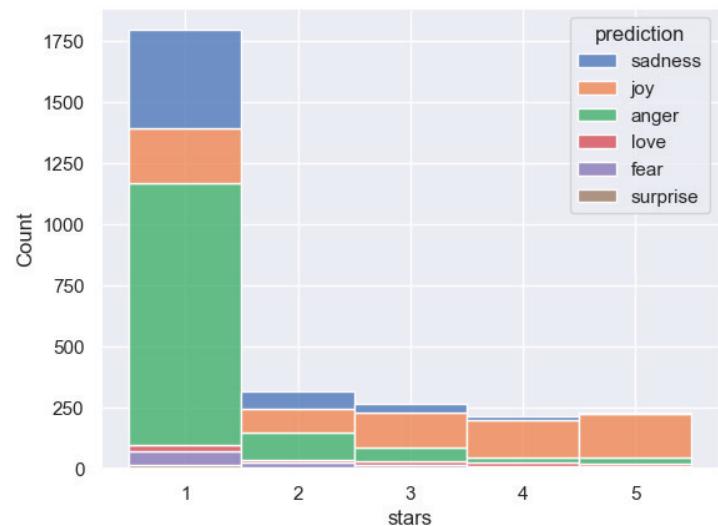
3 - Philly Yelp Reviews Sentiment Analysis

In this in-class project, I applied basic sentiment analysis to some yelp reviews of places in Philly. The sentiment analysis is a form of text analysis that can be used to identify emotions carried through written case. In this case, there are 7 types of basic emotional expression that can be identified, as shown on the right in Fig 6. Fig 5 is a sample of result.

Fig 5 - Sentiment Analysis
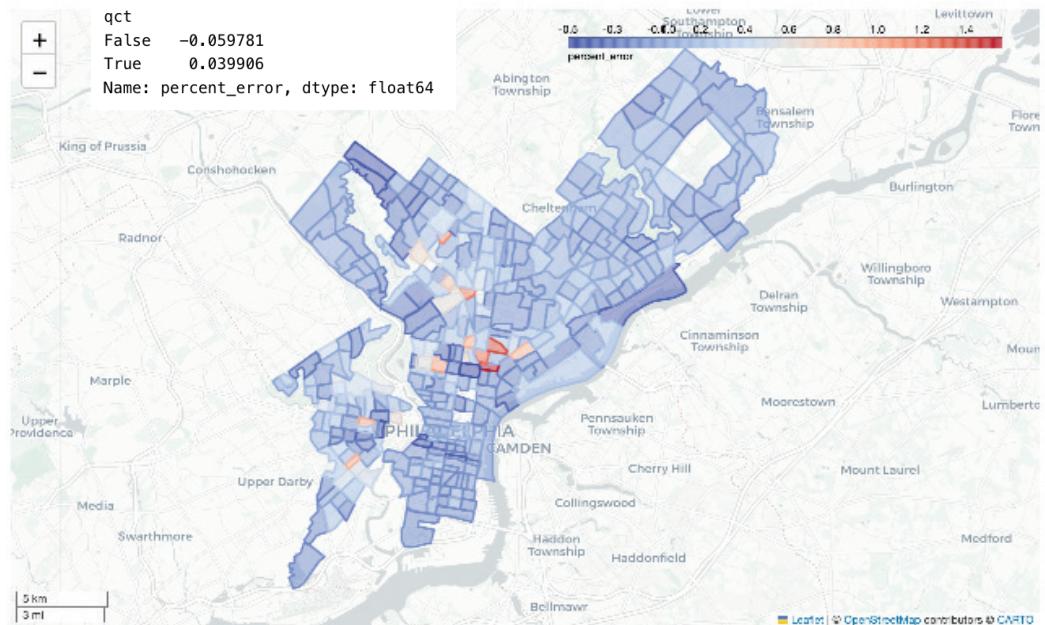
|	sadness	fear	anger	joy	surprise	love	text
0	0.733869	0.250676	0.011039	0.002758	0.001015	0.000643	I know I shouldn't expect much but everything I asked for that was on the drive thru menu was not available. I was actually afraid of what I was going to get once I did get it. I saw the movie "waiting". Word of advice stay clear of this arch. Just so you know I was only trying to order a beverage how pathetic is that.
1	0.000230	0.000126	0.000165	0.998759	0.000246	0.000475	I am only giving 5 stars because the Shamrock Shake is back and delicious!! Too bad it's around only once a year ;(

Fig 6 - Apartment Size-Price-Bedroom Relation
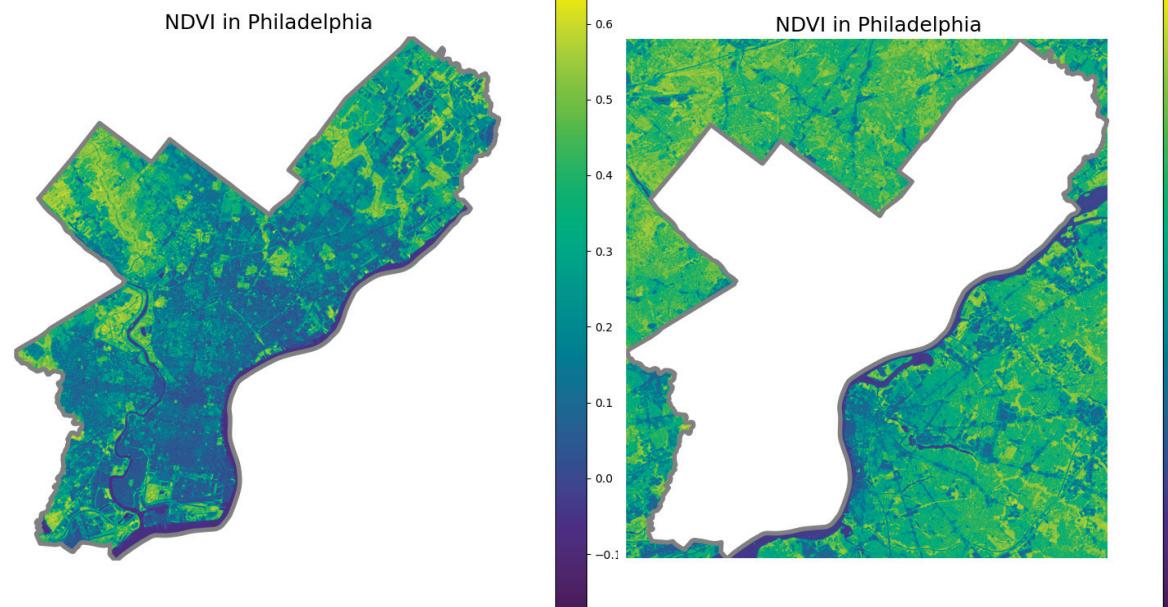
4 - Predictive Modeling of Housing Price in Philadelphia

This project used random forest to train on housing sales data in Philadelphia. The ultimate purpose is to 1) try to train the most accurate predicting model with the given dataset, and 2) explore if there's any predictive biases in different neighborhoods, particularly putting some groups in disadvantage. The percentage error of the trained prediction model is mapped on the right side. We can see several neighorboods are experiencing more errors. And if we compare error of Qualitfying Census Tracts (poverty designation), we will find the QCT experiencing over-valueing, which may cause housing afforddability issue.


```
qct
False    -0.059781
True      0.039906
Name: percent_error, dtype: float64
```

5 - Rasterio: NDVI in Philadelphia
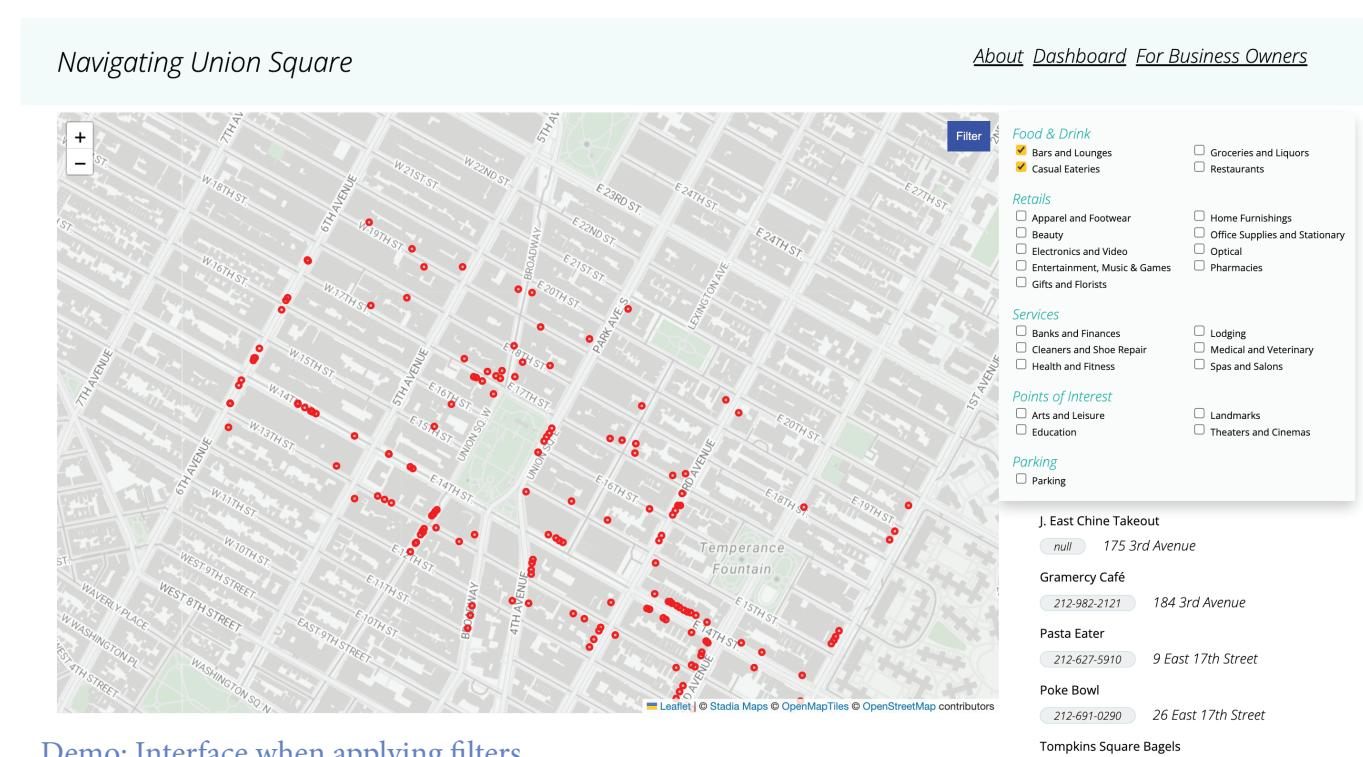
This project aims to explore the green coverage within the boundary of city of Philadelphia versus the immediate outside. Rasterio and matplotlib are used for calculating and visualizing the NDVI, or Normalized Difference Vegetation Index. NDVI is calculated as the ratio between the red and near infrared values. The maps on the right shows that the NDVI in the city tends to be lower than outside of the city. This may be due to the excessive impervious surfaces in the city. Zooming into the map, we can also identify that some neighborhoods in south philly, in particular has lower NDVI comparing to the rest of the city.

NDVI in Philadelphia

NDVI in Philadelphia

Union Sq Biz Dashboard

Dashboard | JavaScript, html, css

This dashboard is created as a part of a JavaScript class for planners and spatial analysts. It displays operating businesses at Union Square in Manhattan, New York City. For long, Union Sq. has an importance in NYC, both in terms of its central geography and vibrant commercial activities. The dashboard opens to both visitors and business owners, with the objective of bridging the two groups for prosperities. The visitors can apply filters to check their desired type of business, or search through the address / businesss entry search bar. Business owners will get a chance to upload introductory or promotional messages, which will show up to visitors when they click on specific business.

Demo; Interface when applying filters.
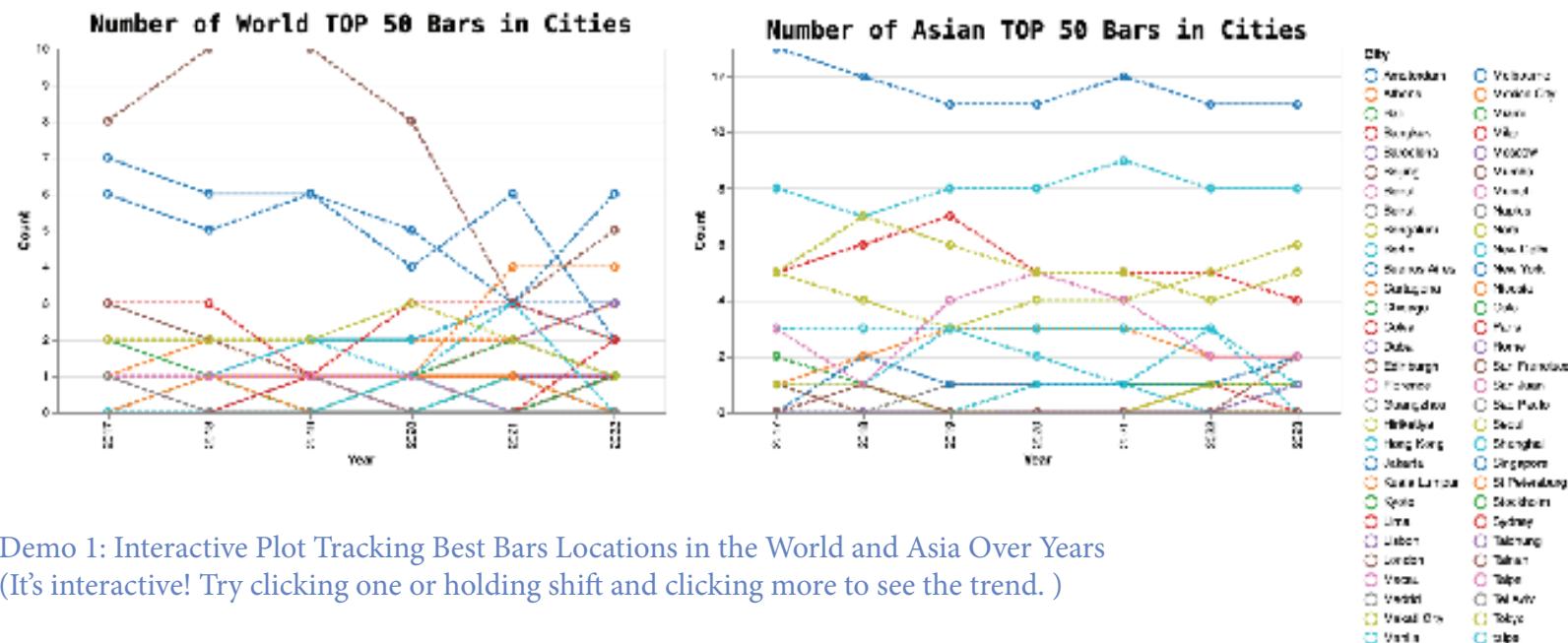
The main motivation behiind this dashboard / navigation app is to prototype a potential channel for promotion and outreach of Business Improvement Districts acorss the country and the world. BID's would have the incentives and resources to promote the local business. And once the template is setup, they only need to have regular maintanence and update the operating businesses.

Looking into the future: As mentioned above, the operational cost may be a burden for BID, especially when they are just starting. Yet in long term, I believe it is meaningful and helpful for local businesses who doesn't have enough resources for online marketing to have a platform to communicate directly to the visitors and potential customers for free.
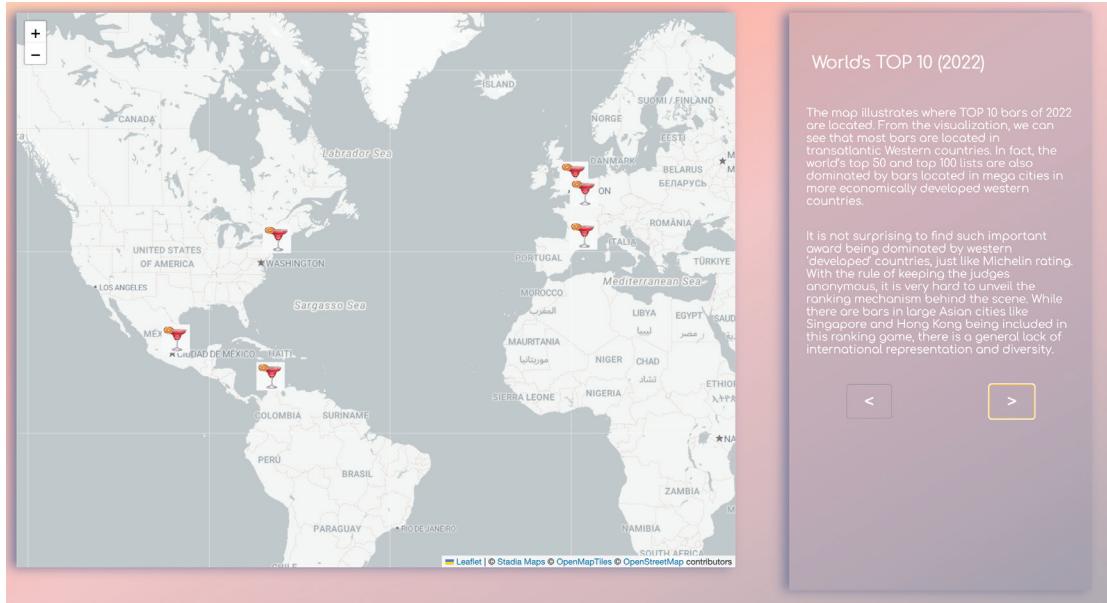
T(ipsy)_T(onight)

Story Map | JavaScript, html, css
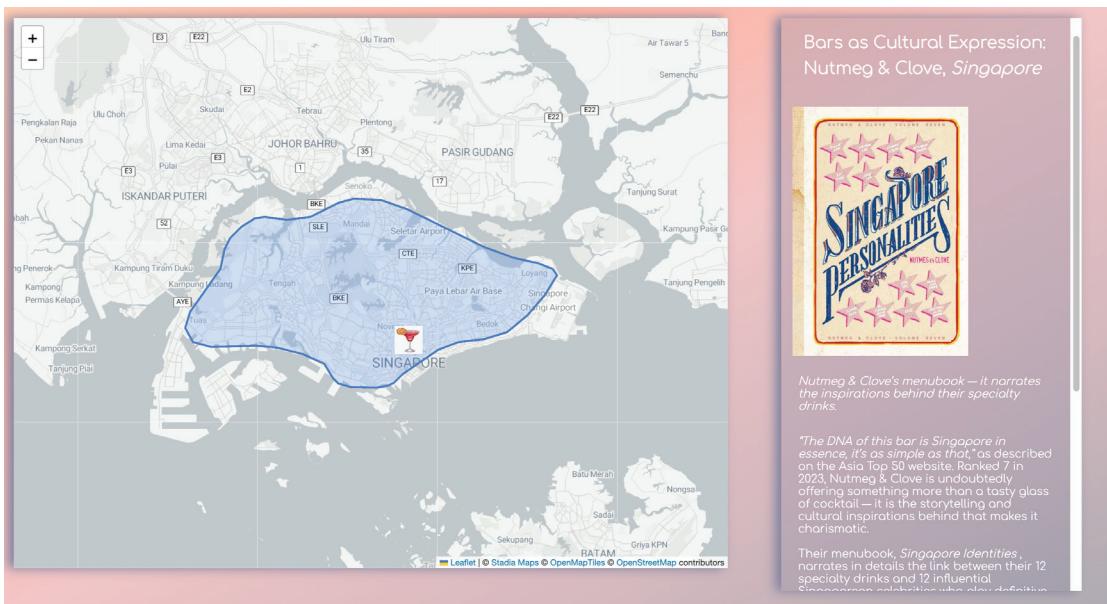
This storymap is born out of my personal interest and passion for drinking and my curiosity in how the bars are judged / marketed around the world. Coming from Asia, I am witnessing the raise of bars in Asia, not only interms of quantity but also quality. Award-wining local bartenders and owners start to integrate the essence of local culture into the drinks, the interior design, and the whole bar experience. On a city level, many metropol-itan cities are home to highly-ranked beloved bars, which turns to be their soft power and attraction to younger generations.


Demo 1: Interactive Plot Tracking Best Bars Locations in the World and Asia Over Years
(It's interactive! Try clicking one or holding shift and clicking more to see the trend.)

Echoing the City's soft power point, I tracked how many Top 50 bars each city is home to from 2017 to 2022 (2023). As you will find when interacting with the graph, cities like London has been home to many top-rated bars over years. Regionally in Asia, Singapore and Hong Kong are undoubtedly the champions in hosting unique and high-quality-bars. During my research, on the other hand, I found Chinese cities like Beijing and Shanghai have been seeing a downfall over the COVID years, potenatily due to the lockdowns. It will be interesting to see whether they will rebound from now.

Demo 2: Storymap Slide Deck Interface

The storymap's slidedecks aim to capture both the general trends around the world (e.g., World's TOP 50 bars are largely located in North America and Europe , showing potential biases of the criteria and ranking systems) and callouts on specific cities or bars that are unique in some ways (e.g., Hong Kong's Lan Kwai Fong and bar industry is very representative.)

Demo 3: Bar Highlight

This is an example of a callout slide for a bar in Singapore. Nutmeg & Clove has been receiving high ranks over years. It is one of the best examples of how bars become windows to introduce local history and culture. Their menu is inspired by some of the most influential Singaporeans, and the drinks and food incorporate local ingredients and recipes. And we are seeing more and more bars in Asia and other places in the World making similar moves towards shaping bars into a cultural entity.

Housing/AirBnB Price Prediction

Regression & Machine Learning | R-Studio
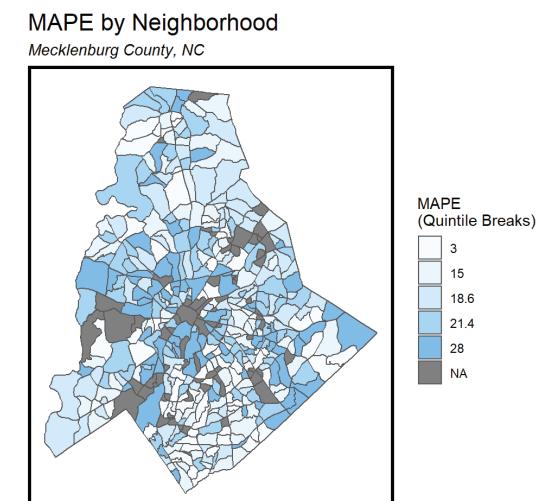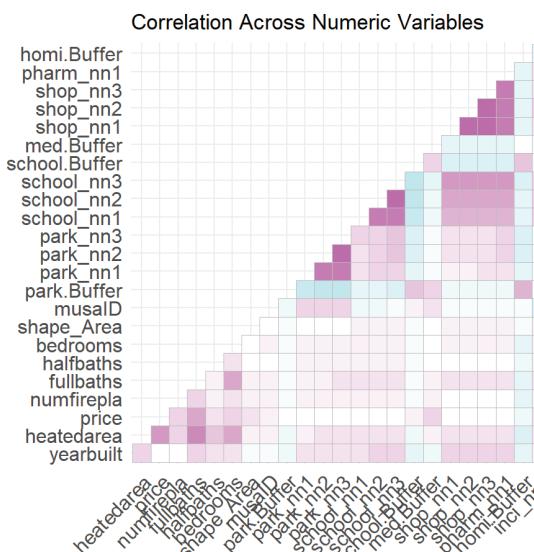
    Two projects in this section both uses basic linear regression and machine learning to conduct price prediction. One is predicting Charlotte's housing price, while the other predicts the airbnb price in Amsterda. Respective geo-social factors and nature of houses for purchase vs. airbnb for short rental are considered when creating the prediction model. All analysis and visualization are done in R-studio.

Charlotte, NC Housing Price Prediction

Project In Collaboration with Ben Keel.

    To train the model, me and my teammate has been exploring critical factors that can contribute to higher predicability. We have identified factors like closest parks, shops, schools, etc. and acquired data from the city's open portal. We then looked at the correlation across all factors to clean up. Using our model, the adjusted R-squared is 0.712, which indicates a decent predicability. The mean absolute percentage error is mapped, and we can identify some clustering near the city center, though not entirely obvious.

```
##
## Regression Results
## ==============================================
##                          Dependent variable:
##                         ----------------------
##                                  price
## ==============================================
## Observations                    37,900
## R2                               0.713
## Adjusted R2                      0.712
## Residual Std. Error    213,352.000 (df = 37789)
## F Statistic          853.159*** (df = 110; 37789)
## ==============================================
## Note:              *p<0.1; **p<0.05; ***p<0.01
```

Correlation Across Numeric Variables

Fig. 2

Home Price & Distance to the (1st/2nd/3rd) Nearest
Mecklenburg County, NC

Fig. 3.3

MAPE by Neighborhood
*Mecklenburg County, NC*

MAPE
(Quintile Breaks)

14
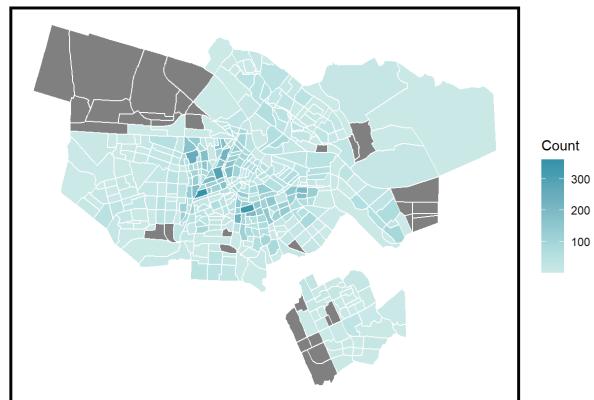
Amsterdam Airbnb Price Prediction

Project In Collaboration with Ben Keel and Tom Sun.
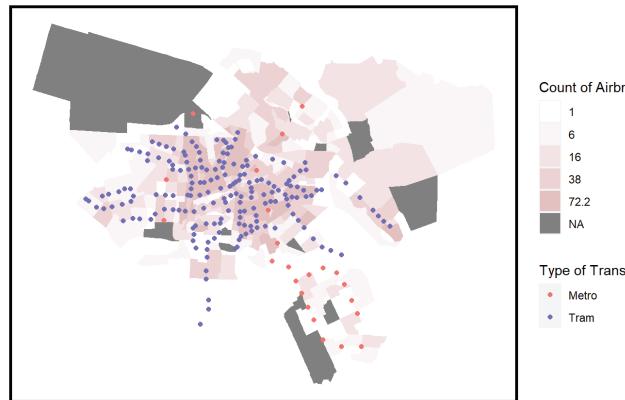
Number of Airbnb per Zipcode
*Amsterdam, NL; 2018*
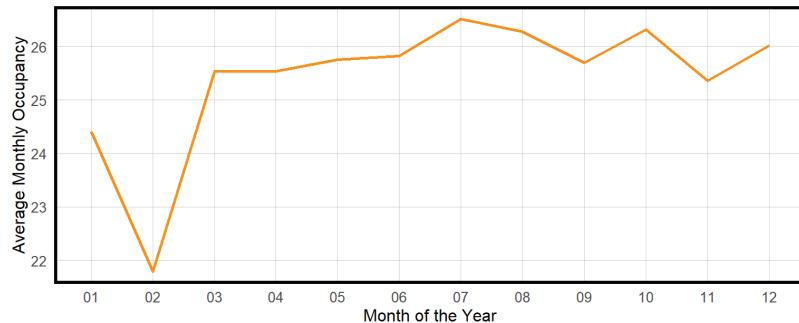
Public Transit & Number of Airbnb
Amsterdam, NL; 2018

The special charateristic and major difference from the Charlotte project is the seasonality and variation of Airbnb prices. Hence, for this project, we not only incorporated factors that are important for tourists, such as public transit and amenities (e.g., pool, parking), but also take seasonality and occupancy data each moth into consideration when we were creating the model.

Airbnb Occupancy Every Month
Amsterdam, NL; 2018
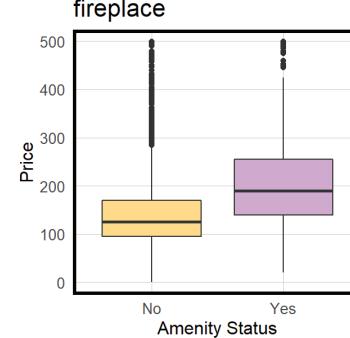
Average Airbnb Price Every Month
Amsterdam, NL; 2018
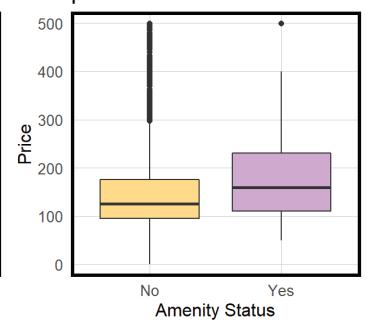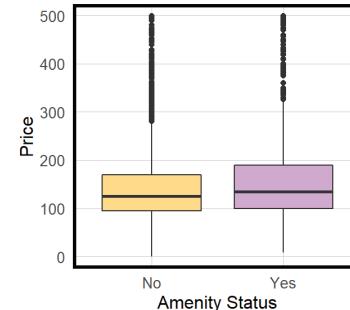
Price w/ vs. w/o Amenities

fireplace

waterfront
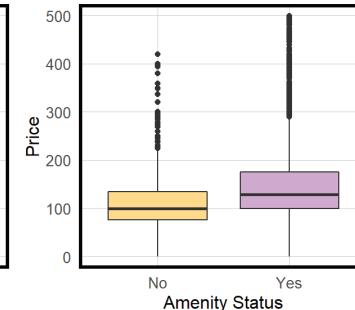
pool
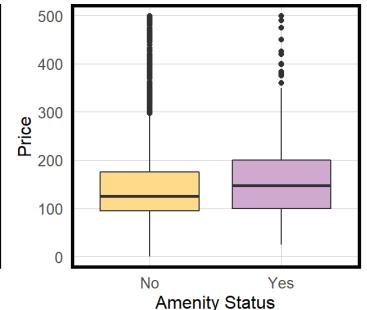
parking

kitchen

AC

15
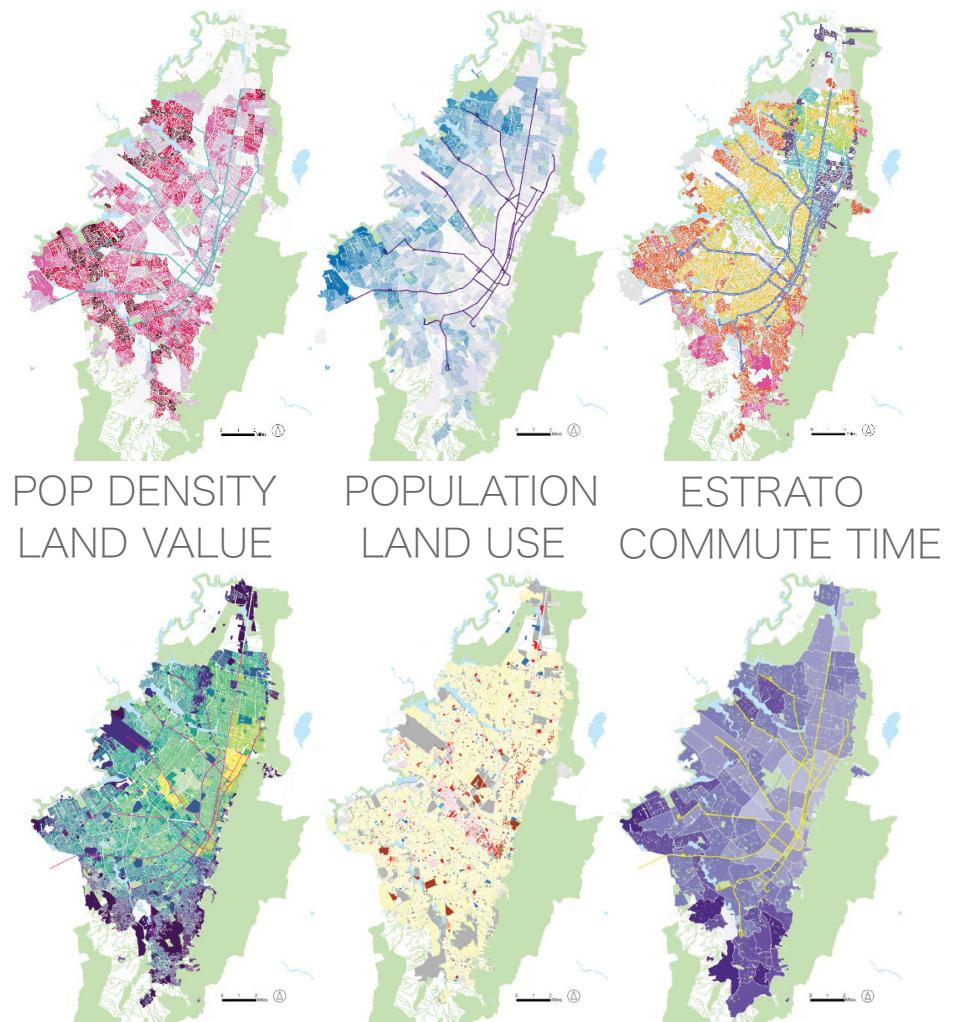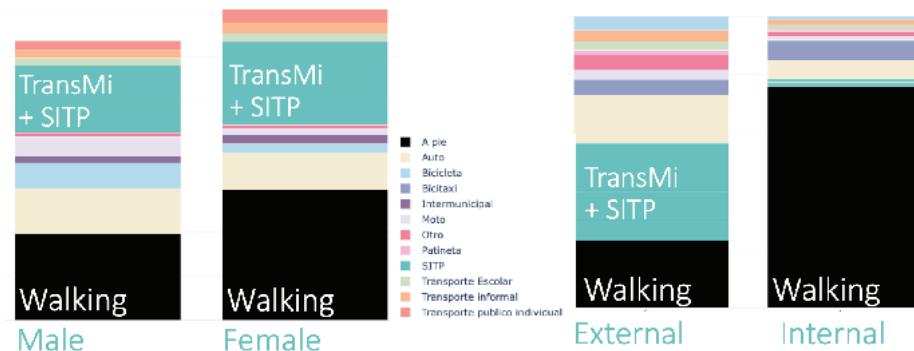
Bogota
Transit Equity & Housing Affordability

Existing Condition Research

Project In Collaboration with Ke Zhou, Evan Zhao, and Teresa Chang;
All work displayed on this page completed by myself.

   This is a studio project done in-class. My teammates and I looked closely at the transit equity and housing affordability issue in Bogota, Colombia. We found that despite the previous success of the implementation of the famous BRT system, not all areas in the city receive equal access to such high capacity service. Those majority living on the peripheries are experiencing extremely long commutes and at risk of safety concerns, including crashes and assults. The modal share graph below also shows a difference in gender -- female experience more internal walking trips and external TransMi (BRT) trips while male has more diverse range of choices for travel mode. The common concern on public safety, especially women, inspired my intervention project on enhancing travel safety, which is shown in the next page.

POP DENSITY   POPULATION   ESTRATO
LAND VALUE   LAND USE   COMMUTE TIME

Legend:
- A pie
- Auto
- Bicicleta
- Bicitaxi
- Intermunicioal
- Moto
- Otro
- Patineta
- SITP
- Transporte Escolar
- Transporte Informal
- Transporte publico individual

Male | Female | External | Internal

Intervention Project Proposal: Travel Safety in Bogota

In Collaboration with Teresa Chang; All work displayed on this page completed by myself.

To enhance the safety of traveling during both internal and external trips, *Viaje a Gusto* is designed to be a bipartite program involving two complementary projects. *Viaje sin Peligro* ('Travel without Danger') targets internal walking trips through sidewalk enhancement and streetlight redesign. *Viaje sin Miedo* ('Travel without Fear') aims to tackle safety on public transportation (mainly SITP), particularly the prevalent issue of sexual harassment. Each project engages three key strategies.

Viaje a Gusto

Internal Trips

Viaje sin Peligro

(Travel w/o Danger)
Sidewalk and **Streetlight** Enhancement and Redesign

1 General Sidewalk Enhancement
2 Streetlight Enhancement
**3** Permeability - 'Natural Surveillance'

External Trips

Viaje sin Miedo

(Travel w/o Fear)
Safety on Transit, Less Sexual Harassment Initiative

1 Priority Seating and Waiting Area
2 Social Campaign
3 Text Reporting System

NO AL ACOSO

Para las Mujeres Para los Hombres Para todos

Reportar acoso a la policía / TransMilenio Conductor(a). Viajes a gusto para todos.

NO AL ACOSO, VIAJE SIN MIEDO

Reportar acoso a la policía / TransMilenio Conductor(a). Viajes a gusto para todos.

Poster Design for Social Campaign to raise awareness of sexual harrassment on public transportation.

This is our proposal for increasing permeability and transparency right nearby a major TOD transit hub. By opening up facades, we are encouraging 'natural surveillance' and making pedestrian feel safer when walking.

Scenario

Proposed

Portal de Los Americas
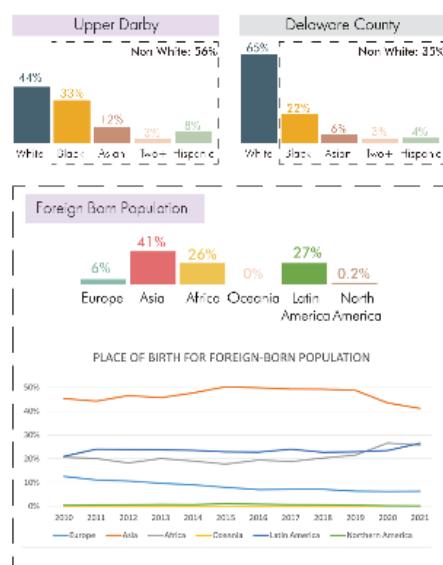
Garrett Road Corridor, Upper Darby, PA
# "The World in One Place"
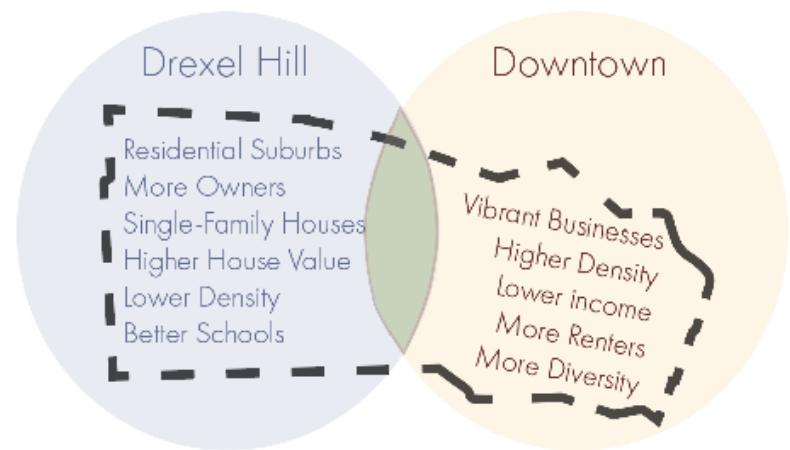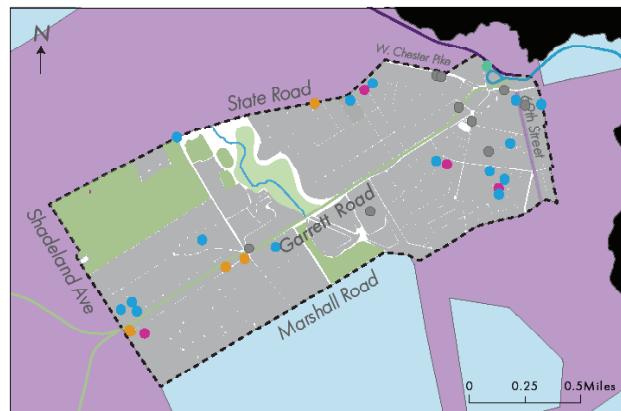
Existing Condition Research

Project In Collaboration with Alexa Ringer, Alex Nelms, Ling Chen, Zhida Ma, Shreya Bansal, and Sidney Kuesters.

For this studio course, my team looked at Garrett Road Corridor in Upper Darby, PA, adjacent to the city of Philadelphia. In addition to the quantitative analysis using census and other public data, we have also focused on qualitative information, especially regarding the socio-cultural aspects. We have identified issues and opportunities for this corridor and Upper Darby township, which includes assets such as largely diverse cultural backgrounds and advantage as a regional transit hub, and issues such as the disparities between the Drexel Hill area versus Downtown Upper Darby.

Left: Large non-white and foreign born population
Bottom: Prevalence of hisotric sites in Upper Darby
Right: Issues and Opportunities summarized by the group

18

Intervention Strategies

Project In Collaboration with Alexa Ringer, Alex Nelms, Ling Chen, Zhida Ma, Shreya Bansal, and Sidney Kuesters.

Upon identifying issues and opportunities, me and my teammates have created a set of intervention strategies and categorized based on their targeted geography. Corridor-wide recommendations are mainly zoning or program-based, which requires mostly top-down effort policy wise and may requires a large amount of investments. Recommendations on Downtown development focuses more on economics and cultural aspects, including setting a BID and creating social campaigns to advertise Upper Darby as a rare cultural enclave in the region. Park and Open Spaces recommendations focus on both equal access to existing green space and creation of new flooding-resilient infrastructure. And Garrett Road recommendations are transportation-oriented, including redesign of right of way, pedestrian walk enhancement for safety, TOD from transforming existing parking lots, etc. Each recommendation / strategy is given a priority and expected timeline. As the table on the right show, most strategies with high priority requires long-term effort and invetsment.

Proposed new ROW, include new greenlane and enhanced pedestrian walk.

Corridor Wide — Downtown — Parks & Open spaces — Garrett Road

	RECOMMENDATIONS		TIME (Years) 1	3	5	7	10
★	Regulation	Zoning Changes			▬		
★	Design	Mixed Use Redevelopments			▬		
	Policy	Community Land Trust				▬	
	Design	Pedestrian Experience Improvements			▬		
	Program	School Programming for Identity & Culture			▬		
★	Policy	Business Improvement District			▬		
	Program	Minority & Disadvantaged Business Enterprise Staff			▬		
	Program	Branding & Marketing Campaign	▬				
	Program	International Restaurant Week	▬				
	Policy	Skill Building Program			▬		
	Policy	Historic Preservation Program			▬		
	Policy	Public Space Improvements			▬		
★	Infrastructure	Park Overhaul			▬		
★	Infrastructure	Flood Infrastructure			▬		
★	Design	Park Connections: Trails				▬	
	Program	Trail Coalition Taskforce	▬				
	Program	Trail Design Competition	▬				
	Program	Tree Inventory	▬				
	Program	Organized Tree Plantings	▬				
	Policy	Commercial Gardening Incentives	▬				
★	Design	ROW Reorganization			▬		
★	Regulation	Parking & Frontage Changes				▬	
★	Policy	Alternative Transportation		▬			

TIME (Years)   1   3   5   7   10