A little background on the data

- PROSPR: Population-based Research to Optimize the Screening Process

    - *Project Status: In Progress (2011-)*

- Our research specifically focuses on lung cancer: the leading cause of cancer deaths, with the lowest 5-year relative survival rates of all cancer types.

    — Diagnosis and mortality rates are highest among Black men, though disparities are seen across racial, geographic, and socioeconomic groups

- Right now, over five different community-based health systems alongside Penn are conducting research on ways to improve the various cancer screening processes

Penn Medicine

In this report:

## THE VARIABLES BEING WORKED WITH

- **Comorbidity** is often measured in clinical settings using a Charlson Comorbidity Index (CCI): *A weighted index to predict risk of death within 1 year of hospitalization for patients with specific comorbid conditions.*

- Factors like age, severity of condition, and compound effect of multiple co-morbid diagnoses all contribute to a CCI.

- Any score over 9 or 10 is considered an extremely severe co-morbid condition and a near impossible chance of survival.

Clinical conditions measured for CCI:

- Liver disease
- Diabetes
- Ulcer disease
- Connective tissue disease
- COPD
- Stroke/Paralysis
- Heart failure/heart disease
- Myocardial infarction
- Dementia
- Kidney disease
- Cancer and Leukemia
- AIDS

Penn Medicine

$$\text{Comorbidity Rate} = \frac{n \text{ of patients with comorbidity per tract}}{n \text{ of total patients per tract}}$$

Histogram of DV- Comorbidity rate
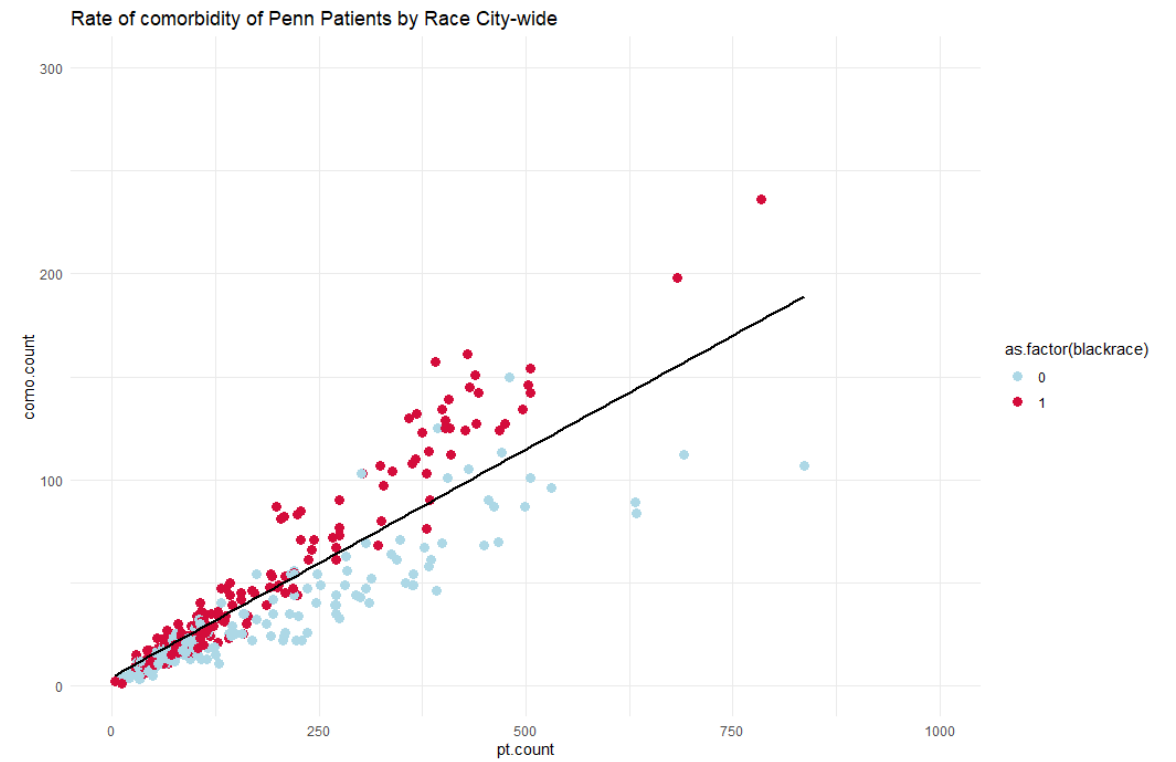
Penn Medicine
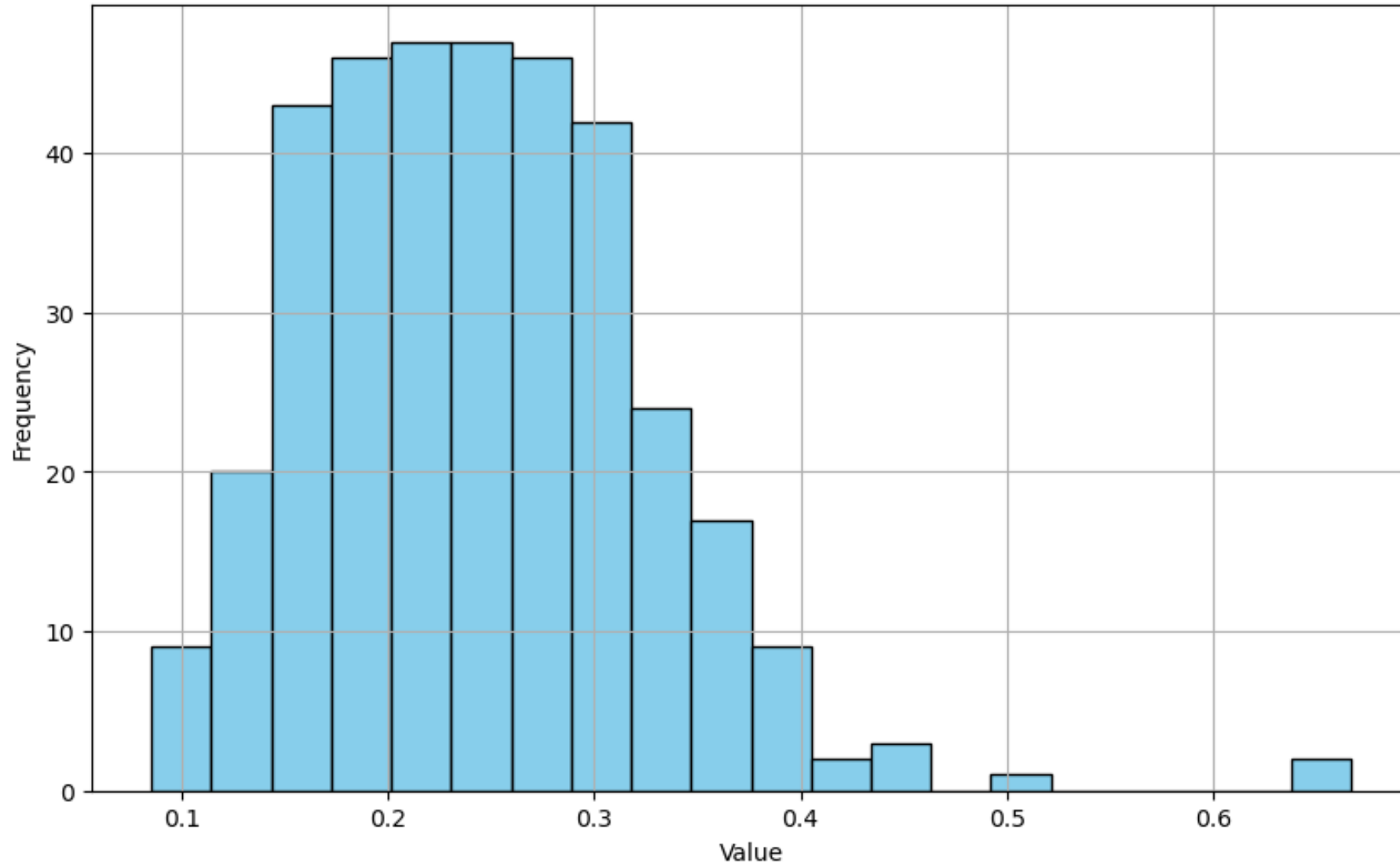
Describing the problem:

OVER-REPRESENTATION OF COMORBIDITY IN BLACK POPULATIONS

Penn Medicine

Key Charts:

Distribution of CoMorbidity by Race
Philadelphia
City Average: 24%

Rate of comorbidity of Penn Patients by Race City-wide
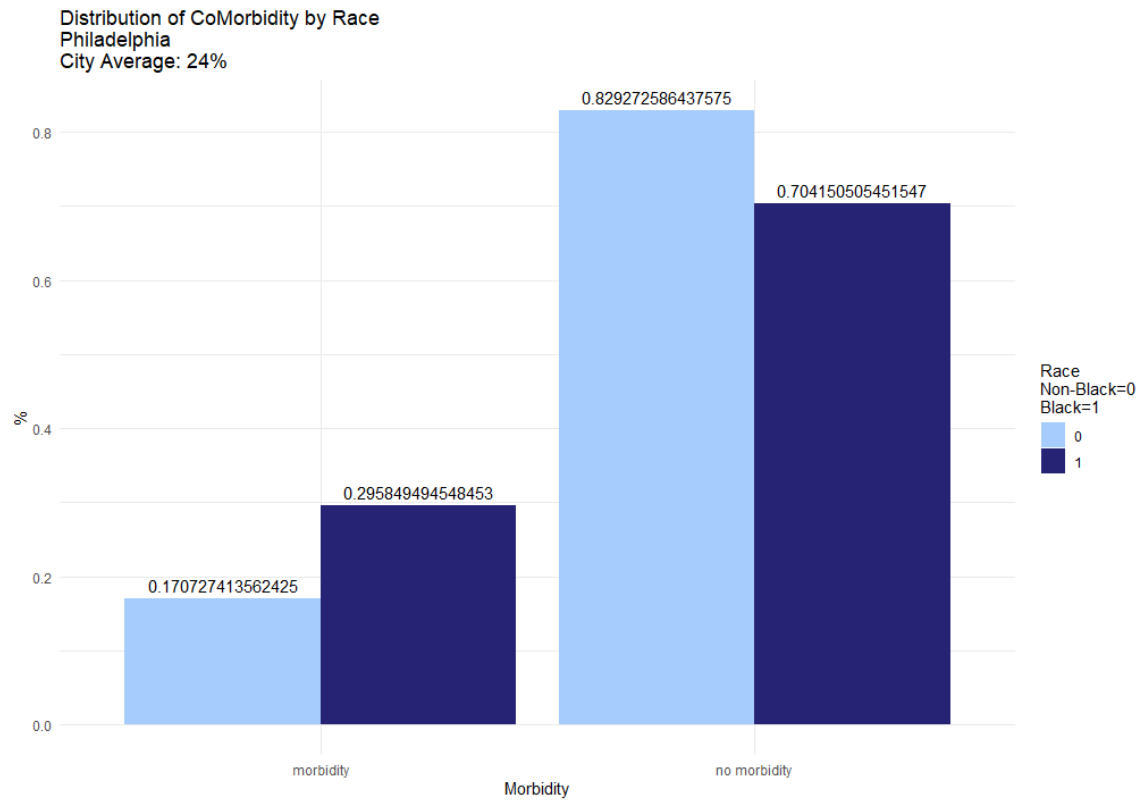
Describing the solution:

USING ENVIRONMENTAL METRICS FROM NAIP IMAGERY TO IDENTIFY WHERE AN FQHC IS MOST NEEDED TO HAVE THE GREATEST PREVENTIVE EFFECT ON COMORBIDITY

Penn Medicine

Why this works:

FEDERALLY QUALIFYING HEALTH CENTERS (FQHC)

- FQHCs are located all throughout the country. Their purpose is to provide subsidized medical services to populations that are the most vulnerable.

- It demands that these health centers be located in places that best serves these populations who are otherwise *neglected by mainstream, private health systems*

- Since we cannot reorganize social structuring to achieve health equity, **we must amend the built environment to best engineer our desired outcome**

Penn Medicine

Why this works:

- Comorbidity is often **more than just a consequence of someone not taking care of their personal health or engaging in unhealthy habits**

- Comorbid conditions are often a product of the built environment and one's social positioning, including *but not limited to:*

  — Housing quality/security

  — Access to reliable transportation

  — Employer-based insurance coverage

  — Labor conditions

  — Access to nutritious foods

  — Proximity to biohazardous environments

  — **Exposure to climate change**

Penn Medicine

"
I QUOTE:

The Role of NAIP:
VEGETATION INDICES

- Land Cover Index: Measure of barren land vs. vegetated land

  - Bare land is associated with urban hot spots, vulnerability to dusty air, vulnerability to drought, and challenges in land development.
    - $(R + G - B) / (R + G + B + 1e\text{-}6)$

- Green Index: Measure of vegetative abundance and plant health

  - Healthy vegetation is associated with cleaner air, drought prevention, heat tolerance, and high property values.
    - $(G - R) / (G + R + 1e\text{-}5)$

Penn Medicine

Ask yourself:

ARE LOW-COST NEIGHBORHOODS SETTLED BY VULNERABLE POPULATIONS BECAUSE ENVIRONMENTAL RISK FROM URBAN DEVELOPMENT REDUCES PROPERTY VALUE FIRST

OR

DOES URBAN DEVELOPMENT SUBJECT VULNERABLE POPULATIONS TO UNDUE ENVIRONMENTAL BURDEN FIRST AND THEN THEIR NEIGHBORHOODS LOSE VALUE AFTER?
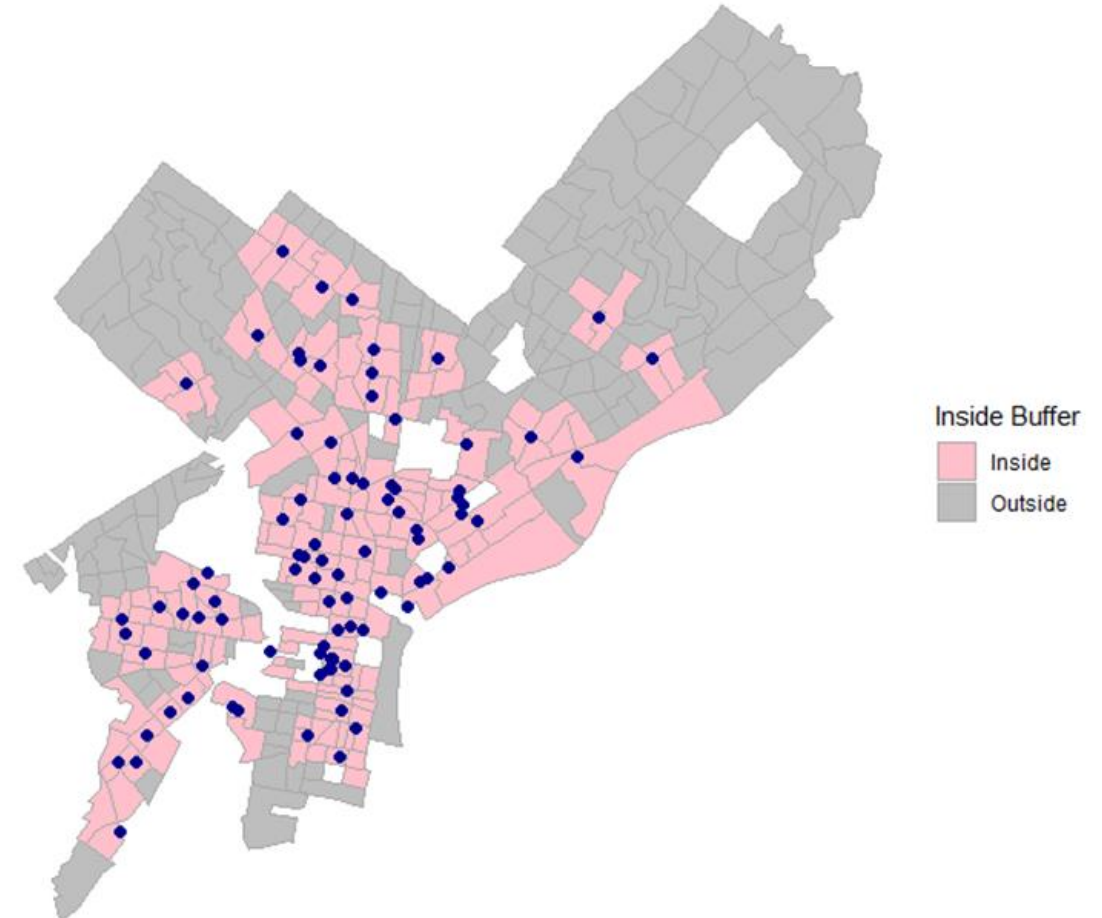
Penn Medicine

The Method:

STEP1

- Creating a ¼ mile buffer around existing FQHC locations in Philadelphia

- There are currently over 60 FQHCs in Philadelphia

- Variable created: Bufferzone

  - We are interested in tracts **outside** of the buffer

Note: Because health data is protected by HIPAA, we do not have point data for patients, just the geocode they live in. When creating the buffer, any tract that even slightly falls within the ¼ buffer is included in its entirety. Therefore, some portions of some tracts may actually be farther than ¼ mile from it's closest FQHC
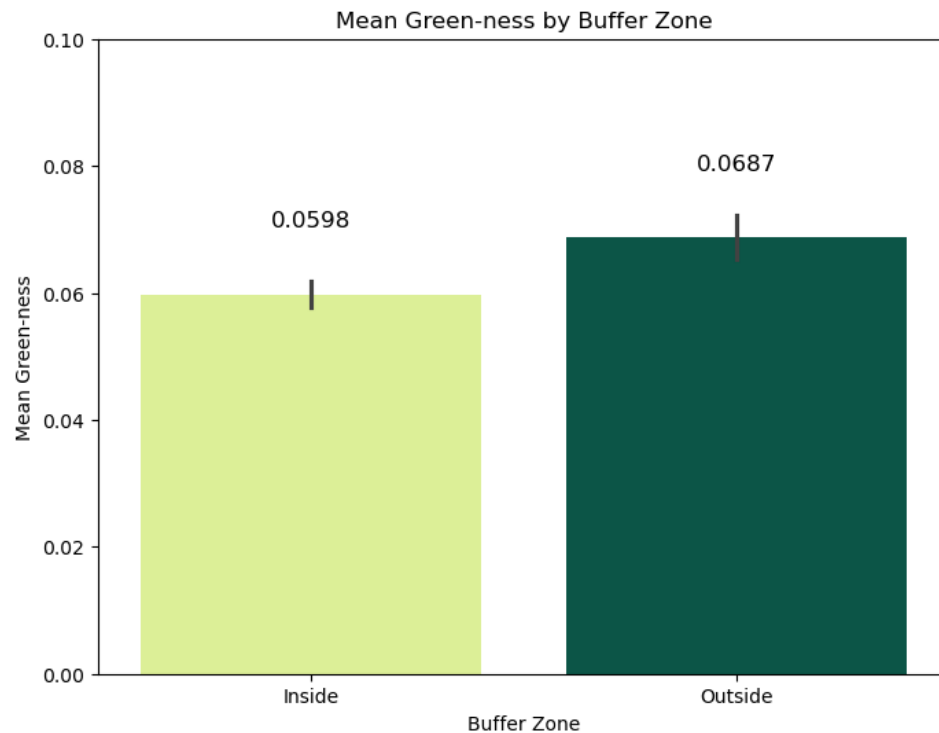
Philadelphia .25 mile Buffer Zone of FQHCs
Inside: n=38,000
Outside: n=23,114

Inside Buffer

Inside
Outside

2021

Penn Medicine

The Method:

- Calculating the indices for each tract, grouping by bufferzone for exploratory purposes

Green-ness	Mean	SD	Min	Max
Inside	0.059	0.02	0.02	0.13
Outside	0.069	0.02	0.02	0.14

Land Cover	Mean	SD	Min	Max
Inside	0.349	0.03	0.29	0.43
Outside	0.354	0.02	0.28	0.42

Mean Green-ness by Buffer Zone

Mean Heat Index by Buffer Zone

ledicine
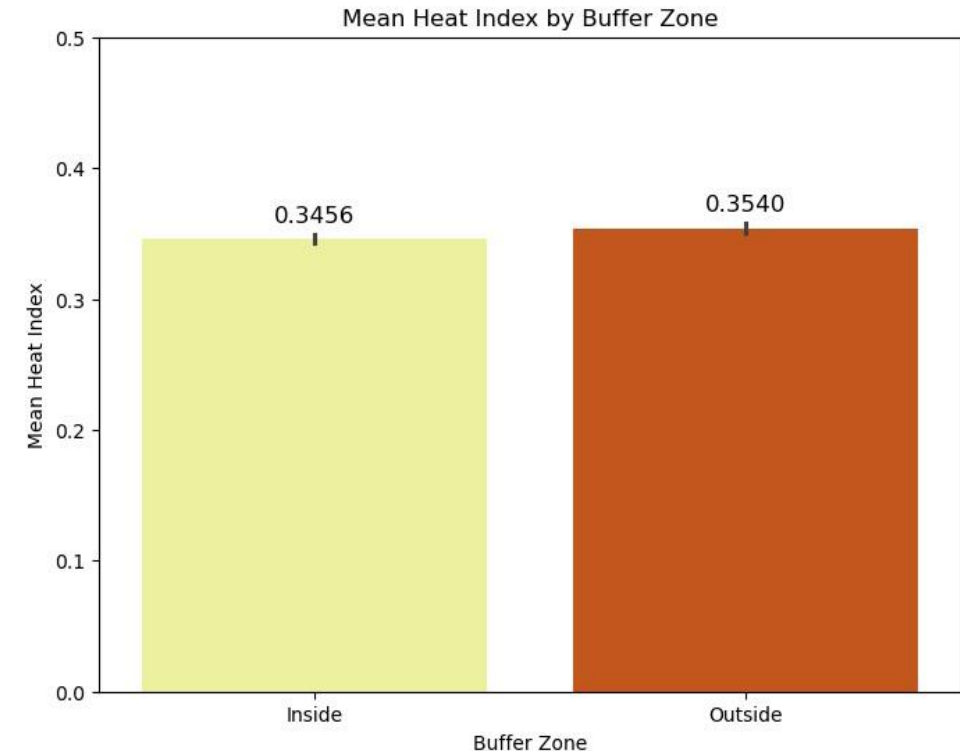
Vegetation Levels with FQHC Locations

Vegetation Category
Low
Medium
High
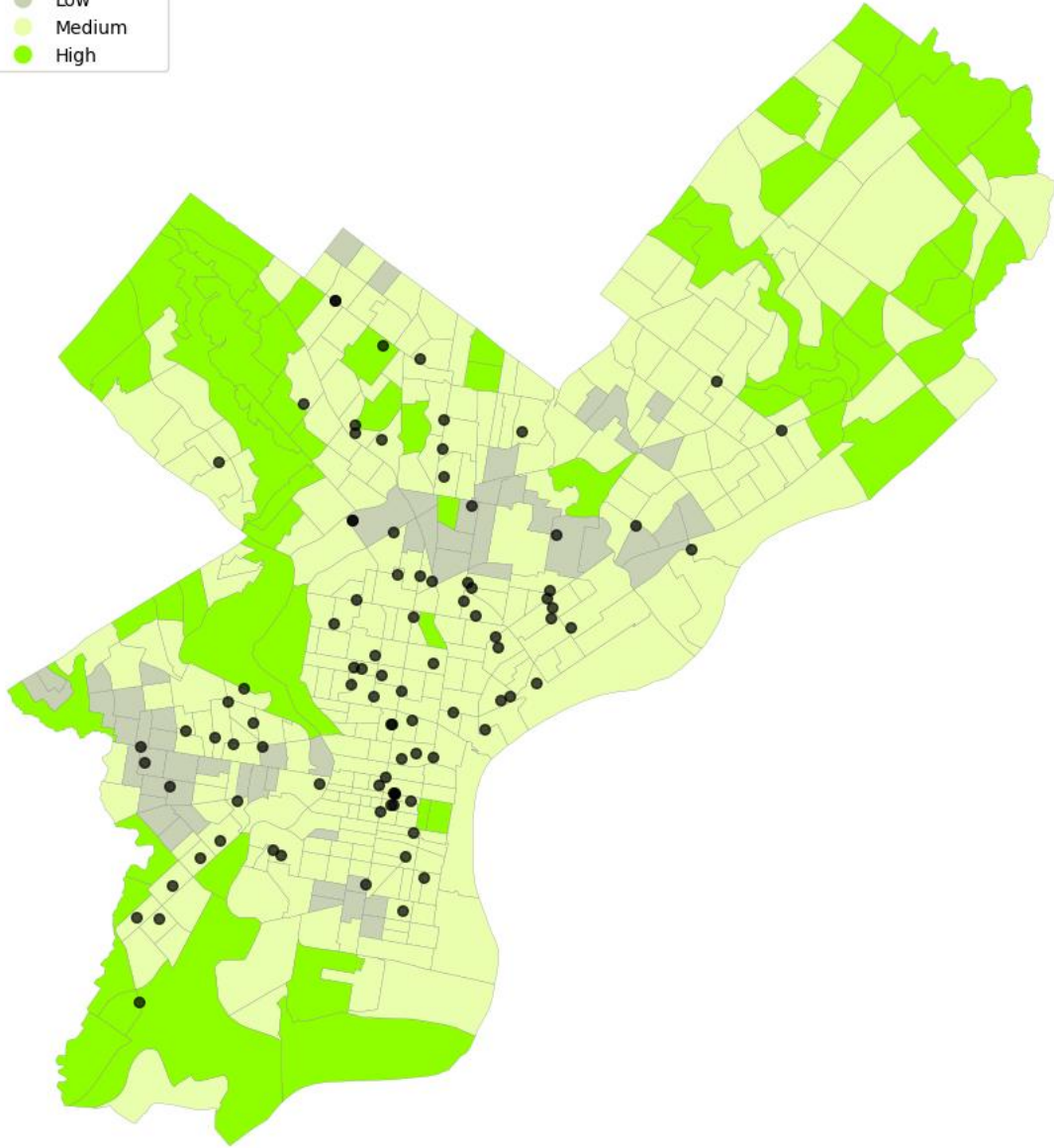
Heat Vulnerability with FQHC Locations

Heat Category
Low
Medium
High
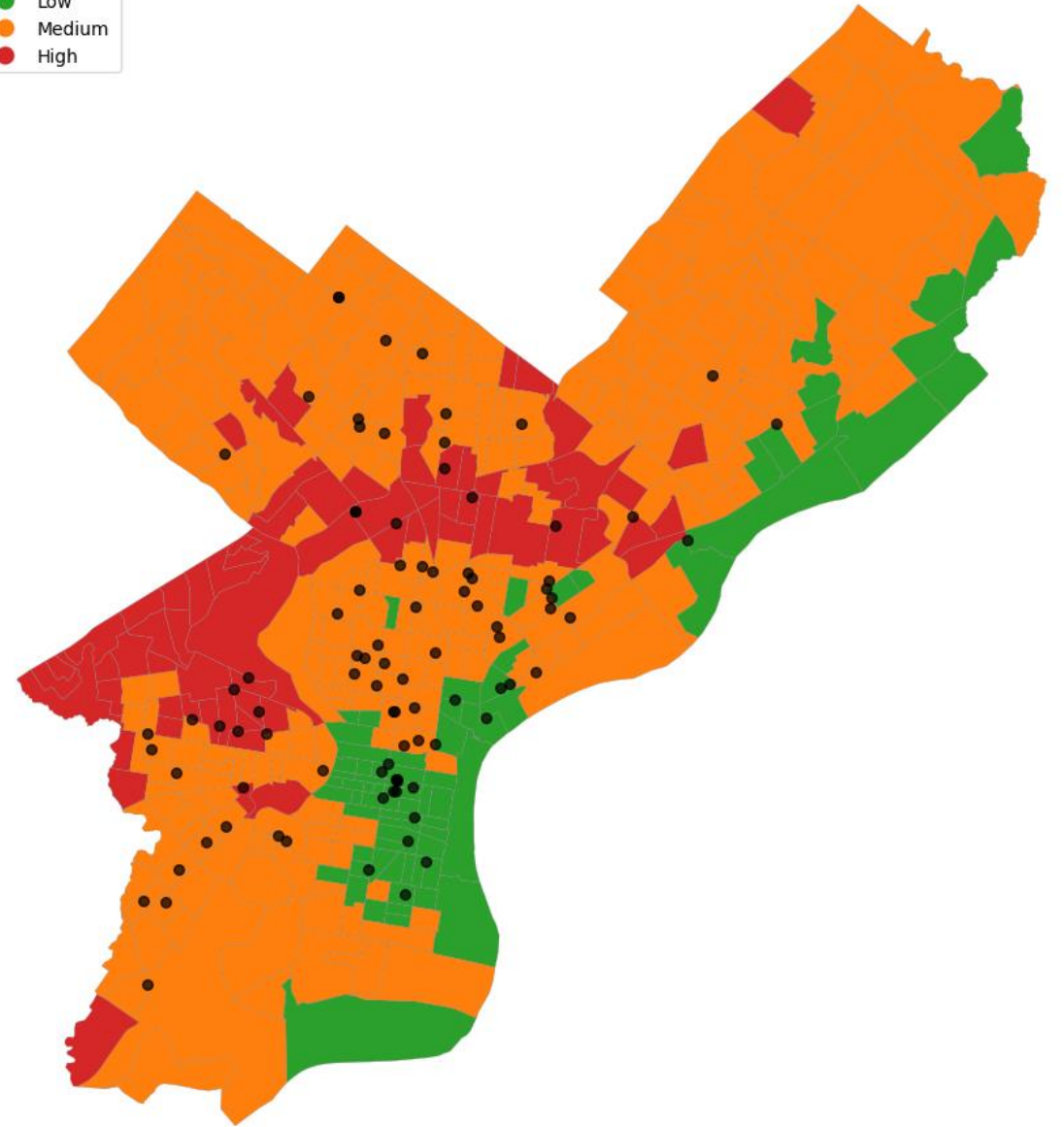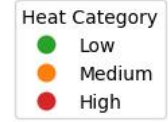
The Method

STEP 3

- Regressing the indices on the comorbidity rate (como.rate~ 0+veg_coverage+heat_coverage+ϵ)

```
                          OLS Regression Results
==============================================================================
Dep. Variable:             como.rate   R-squared:                       0.051
Model:                           OLS   Adj. R-squared:                  0.046
Method:                Least Squares   F-statistic:                     9.590
Date:               Sun, 27 Apr 2025   Prob (F-statistic):           8.79e-05
Time:                       17:08:00   Log-Likelihood:                 398.91
No. Observations:                358   AIC:                            -791.8
Df Residuals:                    355   BIC:                            -780.2
Df Model:                          2
Covariance Type:           nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          -0.0172      0.063     -0.271      0.787      -0.142       0.107
veg_coverage    0.0111      0.211      0.053      0.958      -0.403       0.426
heat_coverage   0.7440      0.171      4.349      0.000       0.408       1.081
==============================================================================
Omnibus:                      95.735   Durbin-Watson:                   1.637
Prob(Omnibus):                 0.000   Jarque-Bera (JB):              361.876
Skew:                          1.125   Prob(JB):                     2.63e-79
Kurtosis:                      7.381   Cond. No.                         54.7
==============================================================================
```
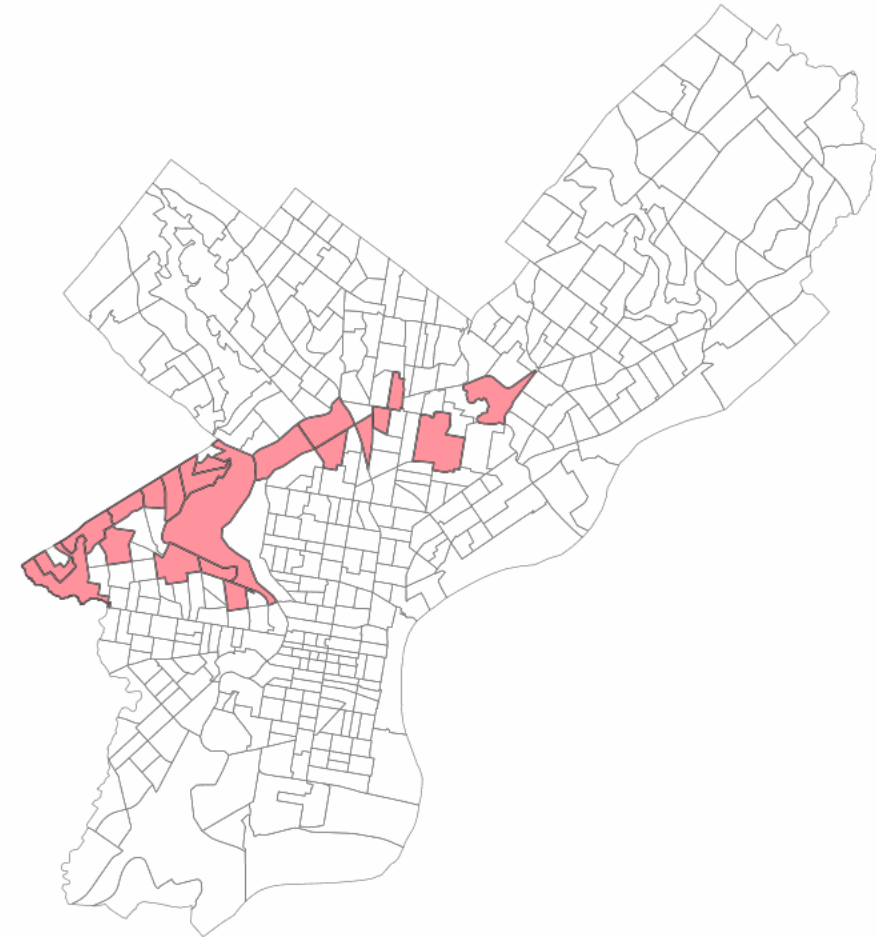
Land Cover Index is statistically significant!!

Penn Medicine

The Method:

STEP 4

- Using AI to identify which tracts are most high-risk:

  - Outside of the buffer zone

  - Land cover index > .375

  - Highest rates of predicted comorbidity

  - Population > 0

  - Rate of poverty > 20%

  - Black population > 40%

All High-Risk Areas

Penn Medicine

Conclusion:

My Top 3 Picks:

But if I had to pick just one….

East Overbrook <3

Click me ;)

Penn Medicine

Limitations

- This method assumes that tracts that fall already inside of the buffer zone would not benefit from a FQHC, when in reality many of those tracts are still extremely vulnerable to environmental risk and disparate health outcomes

- Since the predicted tracts fall on the border of Philadelphia county, it is possible that there are FQHC locations within a "reasonable" distance on the other side of the county line

  — Nonetheless, county divisions are often NOT pedestrian friendly and would only be an option for those with a car

  — Not relevant for 2/3 of the tracts where over 40% do not have a vehicle

Penn Medicine