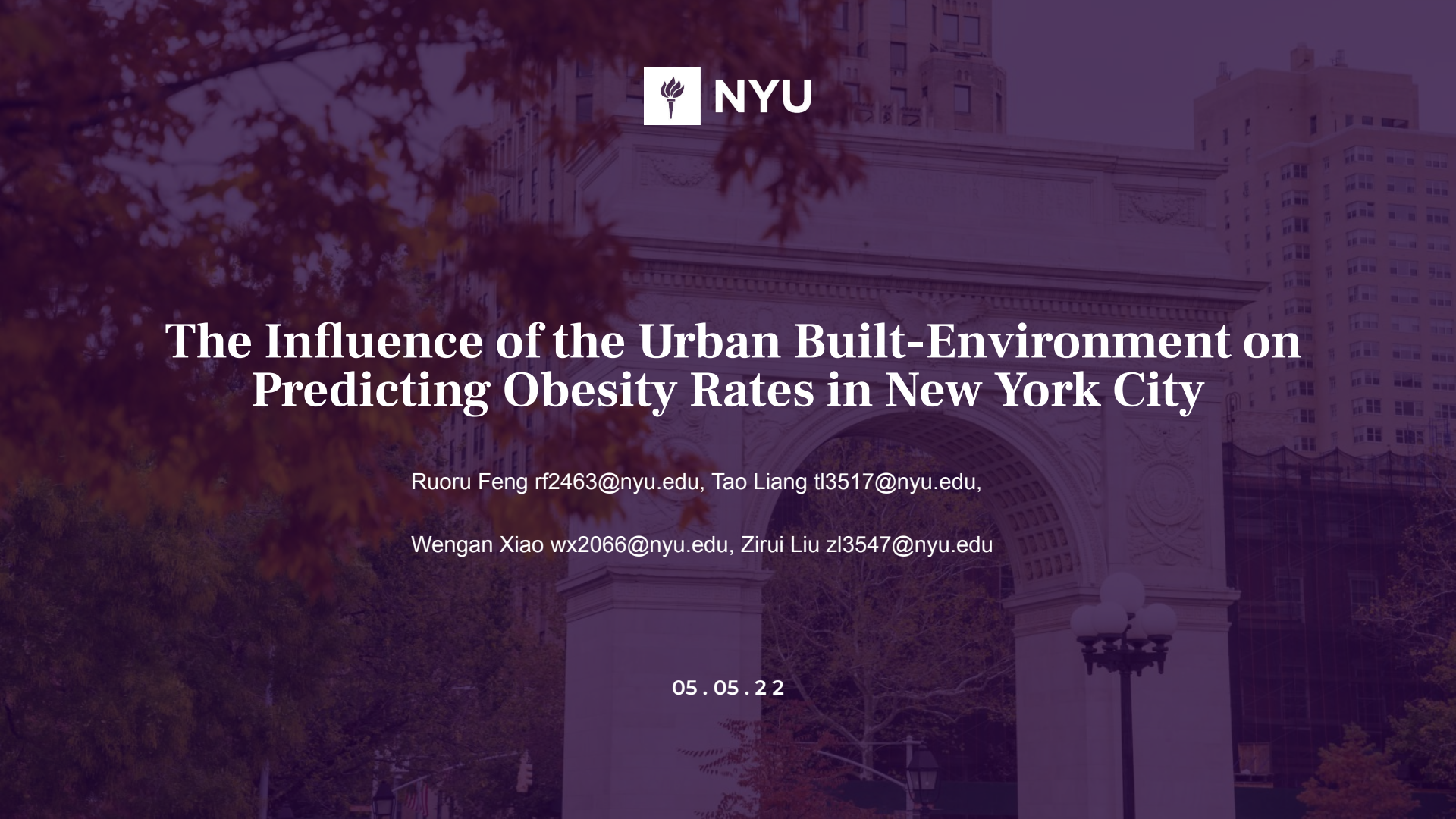


# The Influence of the Urban Built-Environment on Predicting Obesity Rates in New York City

Ruoru Feng rf2463@nyu.edu, Tao Liang tl3517@nyu.edu,

Wengan Xiao wx2066@nyu.edu, Zirui Liu zl3547@nyu.edu

05 . 05 . 2 2

PART 01

# Background

# Built Environment & Obesity

- Built Environment is a conceptually important determinant of obesity in the past decades
- Although most of the study have not yet confirmed the impact of built environment on obesity, many studies had shown that their features have influence on obesity.
- In our project we use historical obesity rate in NYC as our target variable and built environment features as independent variables.
- Our project will gives policymakers and urban planners a measurable idea of how to utilizing urban resources and take action to the vulnerable neighborhoods.

# SDOH

Social determinants of health (SDOH) are the conditions in the environments where people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks.

PART 02

# Data

# Data

## Units:

Census Tracts

(2003  out of 2165, population >= 1000)

**Obesity Rate (CDC, 2019)**

**Socio-economic Data (ACS 5, 2019)**

**Education Data (ACS 5, 2019)**

**Natural Environment (EPA, NASA)**

**Built Environment (OSM, LION, PLUTO)**

| | INDICATOR | DESCRIPTION (SOURCE) |
|---|---|---|
| | OBESITY | Obesity rate of each census tract (PLACE) |
| SOCIAL & ECONOMIC | SE_Unemploy | Percentage of people aged 25-64 who are unemployed (ACS) |
| | SE_Income | Median household incomes (dollars)  (ACS) |
| | SE_Food | Percentage of households receiving Food Stamp/SNAP benefits in the past 12 months (ACS) |
| | SE_Cash | Percentage of households receiving Cash assistance in the past 12 months (ACS) |
| | SE_Poverty | Percentage of population living below the Federal Poverty Level (ACS) |
| | SE_English | Percentage of the population speaking a language other than English who speak English less than “very well” (ACS) |
| | SE_Insurance | Percent individuals age 18-64 with health insurance coverage (ACS) |
| EDUCATION | E_Advance | Percentage of people over 25 years old with a bachelor's education or higher (ACS) |
| | E_HighS | Percentage of 15–17-year-old in high school (ACS) |
| | E_PreS | Percent 3- and 4-year-old enrolled in nursery school, preschool, or kindergarten (ACS) |
| NATURAL | NE_PM25 | Mean estimated microparticle (PM2.5) concentration (EPA) |
| | NE_Ozone | Mean estimated 8-hour average ozone concentration (EPA) |
| | NE_Tem | Mean census tract-wide provisional surface temperature on typical summer days (NOAA) |
| BUILD ENVIRONMENT | BE_TreeD | Street trees density in the census tract. (DPR) |
| | BE_LandMix | Land-use mix level of each census tract. (PLUTO) |
| | BE_WalkVSDrive | Sidewalk area divided by road area. (LION) |
| | BE_nodeD | Street intersection density (pedestrian-oriented intersections). (OSM) |
| | BE_Transit | The average number of transits within 0.8 miles (15 minutes) of the census blocks centroid. (OSM) |
| | BE_Park | The average area of parks within 0.8 miles (15 minutes) of the census blocks centroid. (OSM) |
| | BE_HFood | The average number of supermarkets within 0.8 miles (15 minutes) of the census blocks centroid. (OSM) |
| | BE_Restaruant | The total number of restaurants within the census tracts. (OSM) |
| | BE_doctors | The total number of  doctor's practice within the census tracts. (OSM) |
| | BE_pharmacy | The total number of shops where a pharmacist sells medications within the census tracts. (OSM) |
| | BE_gym | The total number of gyms where a pharmacist sells medications within the census tracts. (OSM) |
| | BE_attraction | The total number of attractions where a pharmacist sells medications within the census tracts. (OSM) |
| | BE_JFood_x | The total number of fast-food restaurants within the census tracts. (OSM) |
| | BE_JFood_y | The average number of fast-food restaurants within 0.8 miles (15 minutes) of the census blocks centroid. (OSM) |

PART 02

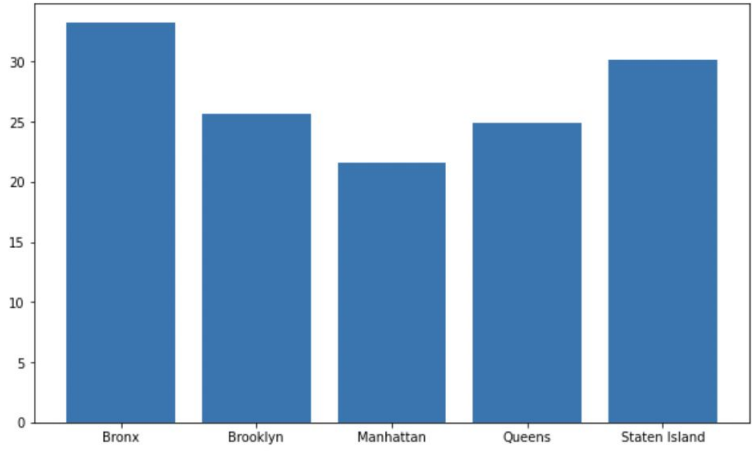
# Obesity Rate

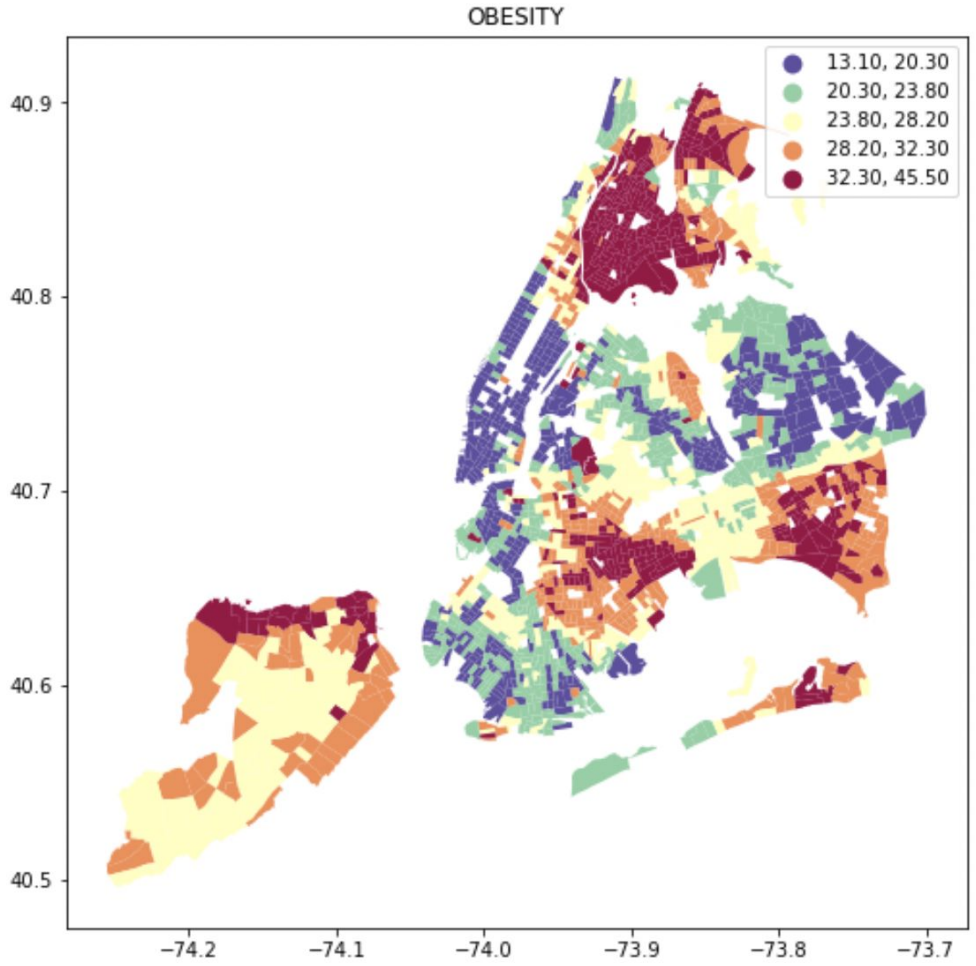
**Mean: 26.38%**

**Standard Deviation: 6.2%**

**Min: 13.10%**

**Max: 45.5%**




NYU

PART 02

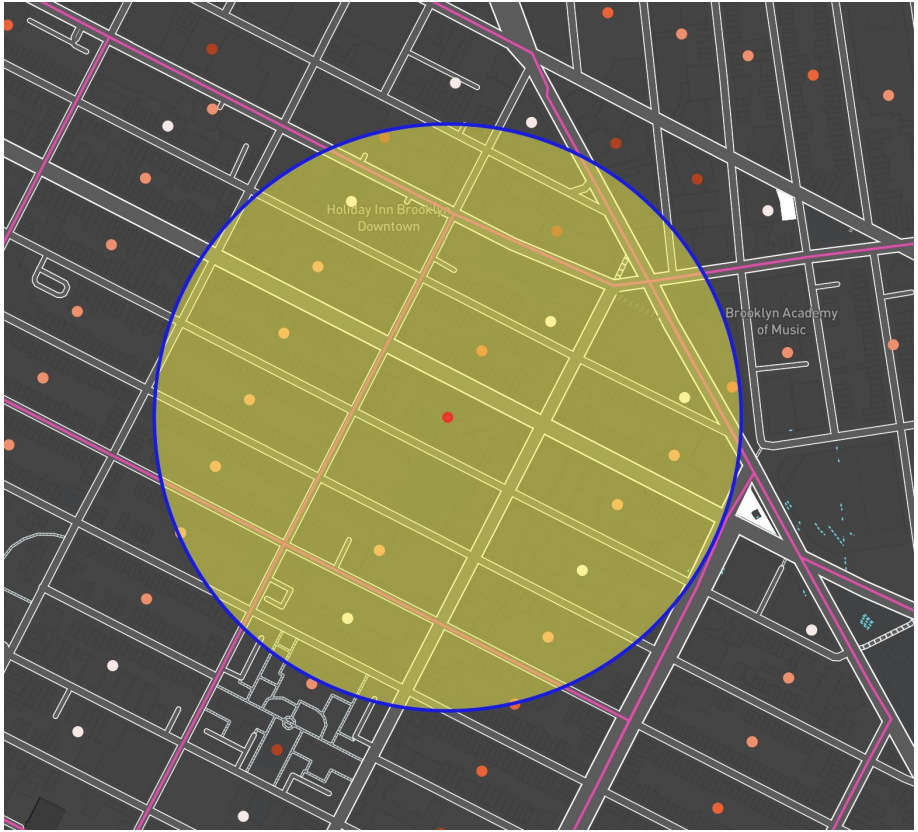
# Feature in 15 min Walking Distance (0.8 mile)

We draw a circle with the centroid of each census block and the radius of 0.8.

We calculate total number of the built environment characteristics that falls into the range.

And take the mean value of them.

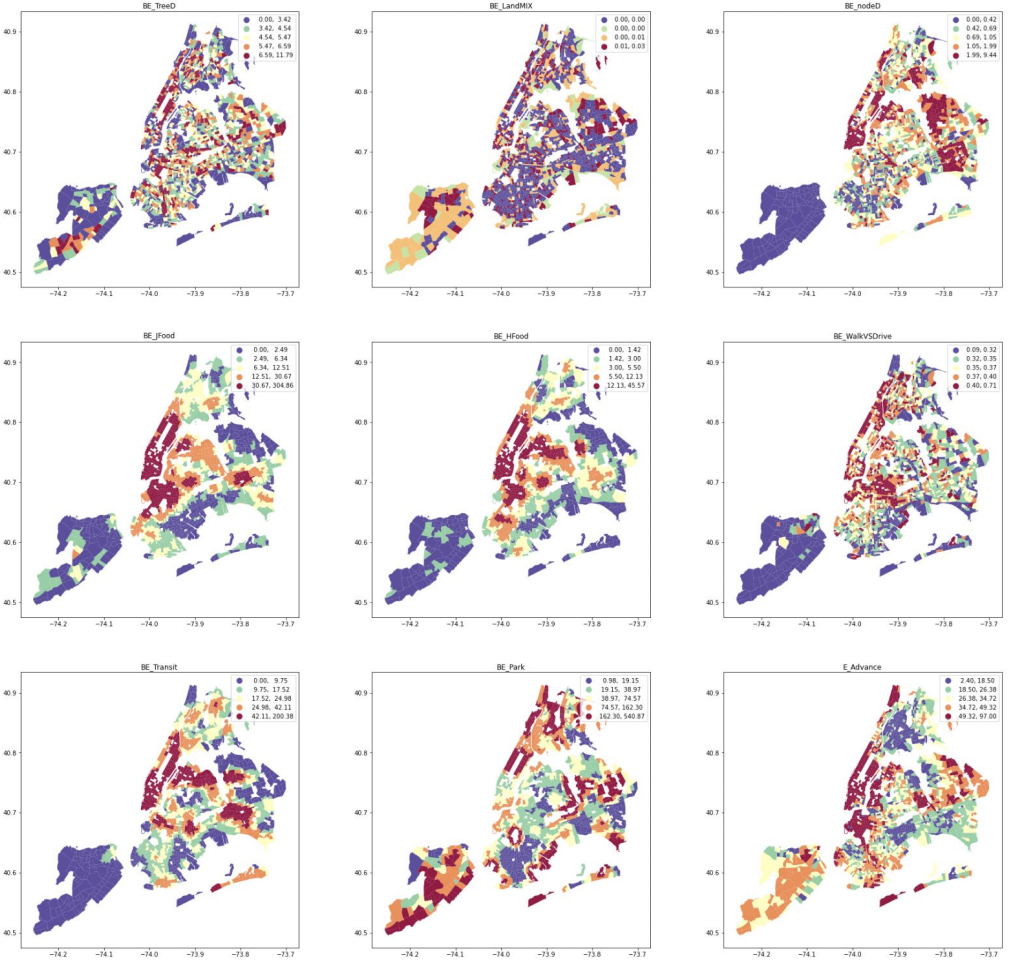
**Supermarkets**
**Fast-food restaurants**
**Park area**
**Transit**




NYU

PART 02

# Data Heat Map




NYU

PART 02

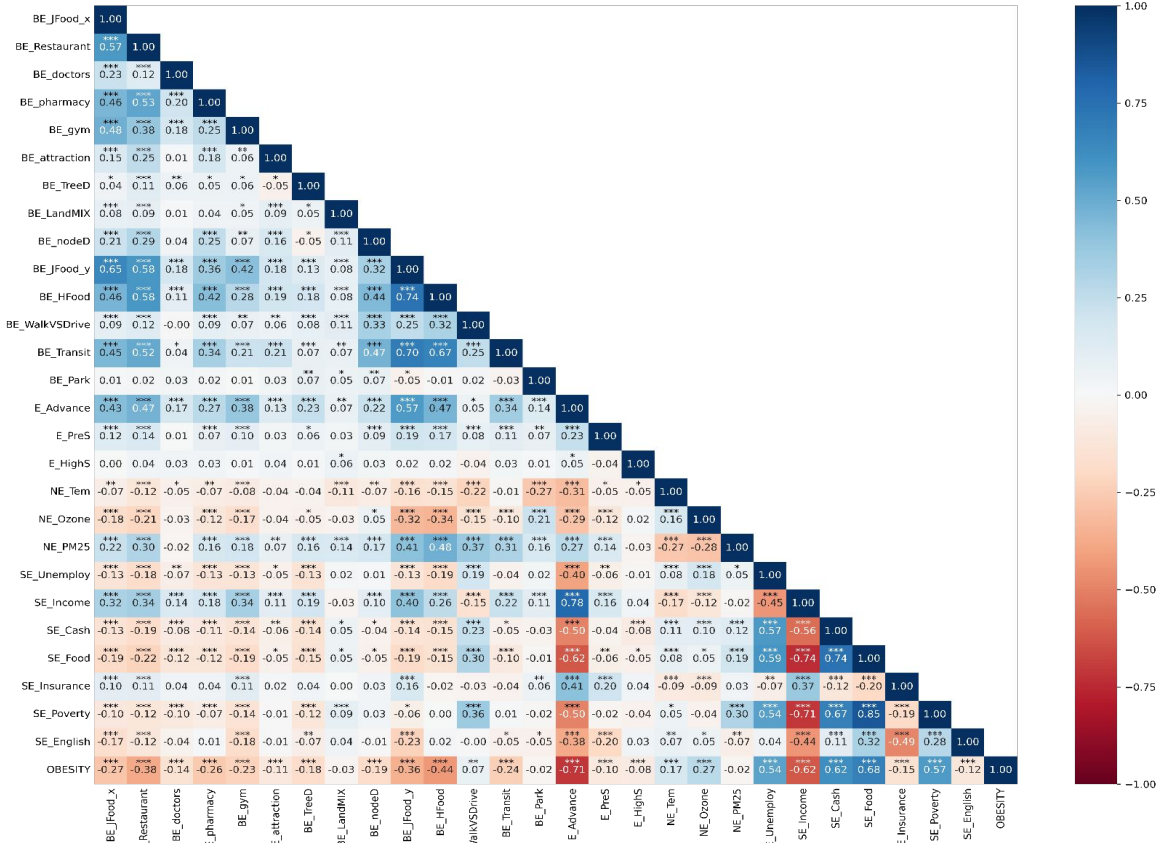
# Pearson correlation coefficient

**P-Value:**

***  < 0.05**

****  < 0.01**

*****  < 0.001**




NYU

PART 03

# Models

# Machine Learning Process

- Stage 1: Non-built environment factors — Use feature correlation
- Stage 2: Built environment factors — combine the feature correlations and feature importance function to filter out all the features to constitute the built environment variables
- Stage 3: Combined dataset of non-built environment factors (Stage 1) + built environment factors (Stage 2), and comparing the results with Stage 1

# Without Building Environment

| Model | Score | MAE |
|---|---|---|
| Decision Tree | 0.74 | 2.31 |
| Random Forest | 0.84 | 1.83 |
| Neural Network | | 1.79 |

# With Building Environment

| Model | Score | MAE |
|---|---|---|
| Decision Tree | 0.75 | 2.22 |
| Random Forest | 0.87 | 1.64 |
| Neural Network | | 1.56 |

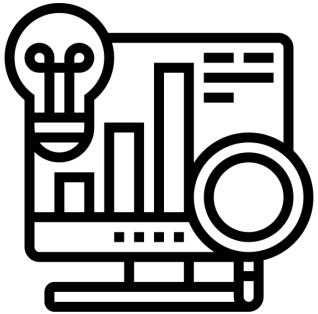
PART 04

# Results

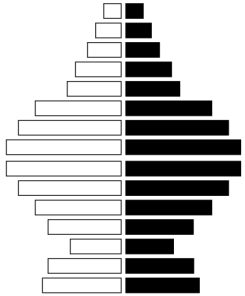
# SHAP

**SHapley Additive exPlanation**

**Game Theory**



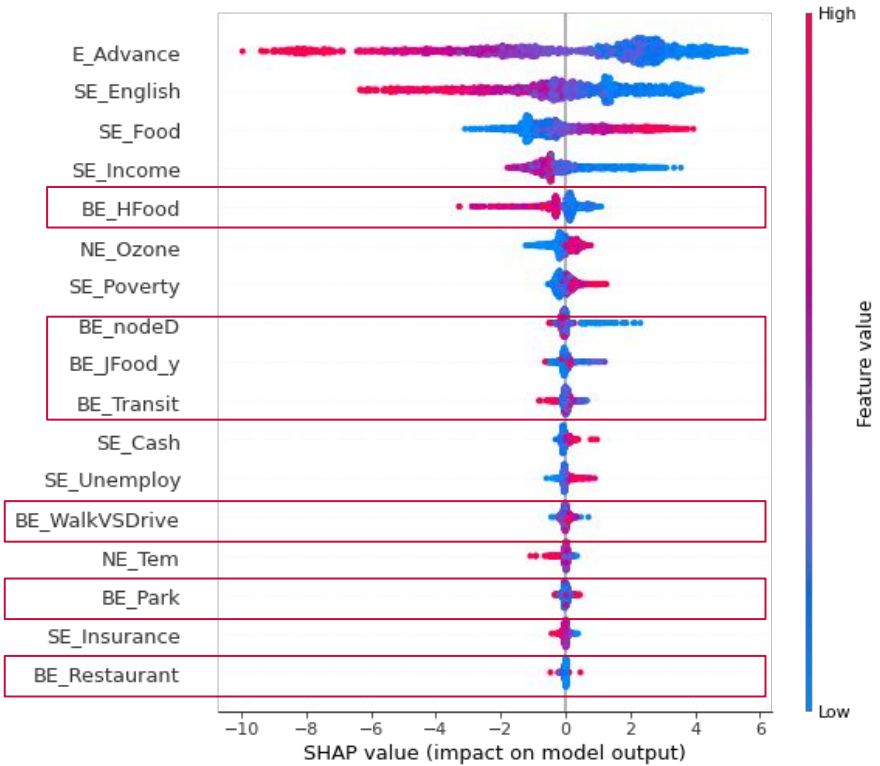
**Model Interpretation Tool**



**Feature - SHAP value**

# Features impacts

## With BE



# Limitation

- The size of the data is relatively small with 2003 rows. The next step might be to consider investigating the obesity rate of entire U.S.
- The improvement in model prediction is not significant might due to the built environment and obesity are rather indirectly related. For example, built environment only encourage physical activities to change obesity rate.

# Reference list

Centers for Disease Control and Prevention. (2019). Places: Local Data for Better Health. https://www.cdc.gov/places/index.html

Levine, J. A. (2011). Poverty and obesity in the US. In Diabetes (Vol. 60, Issue 11, pp. 2667–2668). Am Diabetes Assoc.

Shrider, E. A., Kollar, M., Chen, F., & Semega, J. (2021). Income and poverty in the United States: 2020. Current Population Reports. US Census Bureau, 14.

World Health Organization. (2010). A conceptual framework for action on the social determinants of health. 76. https://apps.who.int/iris/handle/10665/44489

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

# Thank You!