



WHAT ARE “CONNECTIONS”?

The CONNECTIONS section of OurHome.NYC enables users to do two things: explore the correlation between life expectancy and another variable across the 377 census tracts that contain one or multiple NYCHA developments and examine how individual public housing developments stack up when compared with other developments on the given variables. The CONNECTIONS page asks and answers the question: “Is there a correlation between life expectancy and variable y , and if so, how strong is the relationship?”

The CONNECTIONS page includes a carefully selected set of y variables that inform the x variable, life expectancy at birth (years). By clicking on the dots, you can obtain the (x,y) coordinates for every NYC census tract that contains a NYCHA development. You can also compare two sets of coordinates by clicking on one dot and then passing your cursor over a second dot.

This section includes a subset of the full OurHome.NYC dataset; these indicators were selected based on the social science literature on these topics. Researchers interested in exploring the full range of correlations can access the complete OurHome.NYC datasets by clicking the “DOWNLOAD” tab.

One caveat is that the data come from a variety of years. While all the data were the most recent available in the fall of 2019, the original data sources conduct their studies with different frequency. This should not, however, undermine the value of the analysis for pinpointing need, identifying areas for further research, developing fact-based solutions, and more.

WHAT IS A CORRELATION?

A correlation refers to an array of statistical measurements that describe the relationship between two variables. Correlations can be very useful to understand whether, and to what extent, one variable relates to another. They are usually a first step leading to further research on these topics.

The CONNECTIONS section of OurHome.NYC uses Pearson's correlation coefficient, which measures the linear relationship between two variables. Pearson's correlation coefficient yields a number ranging from -1 to $+1$ and is represented by the symbol r . If r is positive, the linear relationship between the two variables is positive, meaning that as one variable increases across the 377 census tracts, the other also tends to increase. If r is negative, then as one variable increases, the other tends to decrease. The further r is from zero (in either a positive or negative direction), the stronger the relationship between the two variables.

For example, the correlation between “Life Expectancy at Birth (years)” and the indicator “Unemployed (% of ages 16 and older)” has an r value of -0.34 , meaning that, across all NYCHA housing developments, these two variables have a moderate negative correlation. In plain English, we can say that as the share of people who are unemployed decreases, life expectancy tends to increase.

CORRELATION DOES NOT IMPLY CAUSATION

Correlations are statements of statistical relationships. The Pearson's correlation describes the linear relationship between two variables. But it does not prove a change in one variable causes a change in the other variable. For example, ice cream sales and murder rates tend to rise in tandem in places with hot summers. Does that mean that ice cream consumption makes people murderous, or that murder makes people crave ice cream? Of course not. Research shows that they both rise in the summer months because warm weather makes ice cream a particularly appealing treat, and summer is a time when people are more likely to get together and to be outside, where they come into greater contact with one another. Determining causal relationships requires extensive research and subject-matter expertise.

We hope these connections spark discussion, spur understanding, and suggest new solutions for our city's greatest challenges.

Happy exploring!