**Overview:**

In the following project, we are looking at city walkability and how it influences personal behavior. More specifically we analyze if an individual can create a walkable community where no one would require cars. How we approached this topic is by looking for trends in walkable communities like car ownership, income, density, and public transportation.

**Dataset:**

For this project, we found a large dataset on EPA.gov looking at the National Walkability Index in the United States and Puerto Rico. This data provides information at the census block group level and every block is assigned a National Walkability Score. The National Walkability score is measured based on multiple factors such as street density, proximity to transit stops, and diversity of area along with many other criteria.

**Analysis questions:**

Do individuals living in a highly walkable area still choose to have a vehicle?

- Do high-income earners who live in a walkable community own a car?

- Do middle-income earners who live in a walkable community own a car?

- Do low-income earners who live in a walkable community own a car?

If individuals are within proximity to public transit stops and live in a walkable community, are they still likely to own a vehicle?

Is there a correlation between most walkable communities and area based on density?

**Pre-Analysis:**

After looking at the dataset we found on EPA.gov we realized there were a lot of missing values in Census Block Groups located in Puerto Rico. We decided to drop all Census Block Groups in Puerto Rico and just provide an analysis of the rest of the American groups. From here we further cleaned up the dataset by locating columns with stats that pertain to our research and removing any rows with missing values.

**Analysis:**

High-Income Analysis:

In high-income earning, most walkable communities and the percentage of zero car owning households on a census block level there is a weak and negative correlation.

In high-income earning, most walkable communities and the percentage of one car owning households on a census block level there is a weak and positive correlation.

In high-income earning, most walkable communities and percentage of households with two plus cars ownership on a census block level there is a weak and negative correlation.

After analyzing high-income earners living in a walkable city on a census block level, we group the census block levels by area and chart results in a bar graph. We found for high-income earners a strong percentage will at least own one car. For high-income earners, regarding the ownership of two plus cars or no cars, the data fluctuates between the combined Census Block Group area analyzed.

Middle-Income Analysis:

For middle-income earners living in the most walkable communities and comparing these values against the percentage of zero, one, or two-plus car ownership, there seems almost to be no correlation between variables. All three scatter plots have a weak and positive correlation.

After analyzing middle-income earners living in a walkable city on a census block level, we grouped the census block levels by area and chart results in a bar graph. This graph doesn’t display a consistent pattern of car ownership across the census block groups areas.

Low-Income Analysis:

For low-income earners, in the most walkable communities the percentage of zero car ownership households on a census block level, there was a weak and positive correlation.

For low-income earners, in the most walkable communities, the percentage of one or two plus car ownership in households on a census block level both have a weak and negative correlation.

We then group the census block groups by area and chart results in a bar graph. The only consistent behavior charted is most low-income groups do not have a car except for the Columbus, OH, and Montgomery, AL census block group areas.  The New York area presents skewed data due to its high walkability but minimal low-income worker presence.

Transit Stops Analysis:

We analyzed the distance from the population weighted centralized transit stops in meters and car ownership in the most walkable communities. We used a random sample size of 2000 since the dataset is dense and charted results in a scatter plot. This sample size portrays that individuals in walkable communities will own a car whether they are near a transit stop or not.

Density Analysis:

We analyzed the gross density of residential, population, and employment as it relates to high income earners in most walkable areas. In the bar graph, New York, New Jersey, and Pennsylvania had the highest density of all categories.

In the scatter plot graph, there wasn’t any correlation between the most walkable CBG and gross densities.

**Data Bias:**

Some concerns that we noted regarding the data set are that low, middle, and high-income criteria are set to fixed ranges across America. This may provide an inaccurate depiction of low, middle, and high-income earners as the cost of living is not appropriately accounted for.

**Future Possible Analysis:**

Is it possible to influence purchasing behavior in walkable communities to be more eco-friendly?

Does living in a walkable community affect perceived happiness and well-being?

**Conclusion:**

In conclusion, the data analysis does not show a clear indication that living in a walkable community affects car purchasing behavior in a consistent manner across income demographics. Whether you have high, middle, or low income, or are near public transportation in a very walkable area, the average household would still own a car in America. There is no correlation between the most walkable community and areas based on gross population, residential, and employment densities in high income earners.