**Walkability Index across Minnesota**

**1.0 Introduction**

The United States Environmental Protection Agency (USEPA) created the National Walkability Index to compare and analyze walkability among US communities. In developing the walkability index, a wide number of built environment attributes are considered ranging from demography, diversity, transit, land use, employment etc. The data set used in the study is obtained from the United Nation Census Bureau. According to US EPA, National Walkability Index inputs are:

1. Intersection density (33%) - Higher intersection density is correlated with more walk trips.
2. Proximity to transit stops (33%) - Distance from population center to nearest transit stop in meters. Shorter distances correlate with more walk trips.
3. Land Diversity (33%):
   1. Employment Mix - The mix of employment types in a block group (such as retail, office, or industrial).
   2. Housing and Employment Mix - The mix of employment types and occupied housing. A block group with a diverse set of employment types (such as office, retail, and service) plus many occupied housing units will have a relatively high value.

For the sorting of data, Block Groups (BGs) are statistical divisions of census tracts, are generally defined to contain between 600 and 3,000 people, and are used to present data and control block numbering. The Walkability Index data is divided into four categories (index score range 1−20, as per USEPA) and here referred as walkability buckets:

1.00–5.75 Least walkable

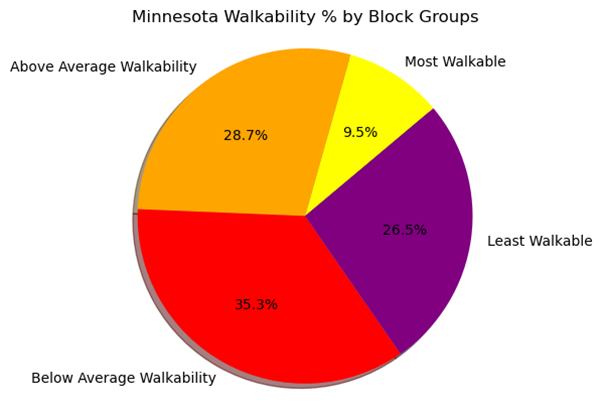
5.76–10.50 Below average walkable

10.51–15.25 Above average walkable

15.26–20.00 Most walkable

**1.1 Minnesota’s National Walkability Score based on block groups**

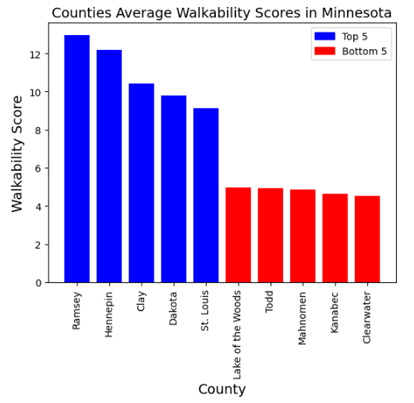
Looking at the National Walkability Index scores, 61.8% of Minnesota falls between 1-10.5 (Plot 1.1). This shows that Minnesota, as a whole, is not a very walkable state. This data was collected by finding all the block groups in Minnesota and sorting them by their National Walkability Score. Once the iteration was done in each category, I found the total number of each category to get the four totals. From there I was able to create a pie chart showing the breakdown of the four categories and how they compare to each other.



Plot 1.1 Minnesota walkability index block wise

**1.2 County wise analysis of walkability index in Minnesota**

Plot 1.2 shows top five and bottom five Minnesota counties walkability index. This information can be helpful when looking at a State and seeing what counties would be easier to get around via walking vs counties that may require some other form of transportation. These scores are gathered by finding the Walkability Index Score for each block group in a County then averaging them. This gave a list of the average scores for all 87 counties in Minnesota.

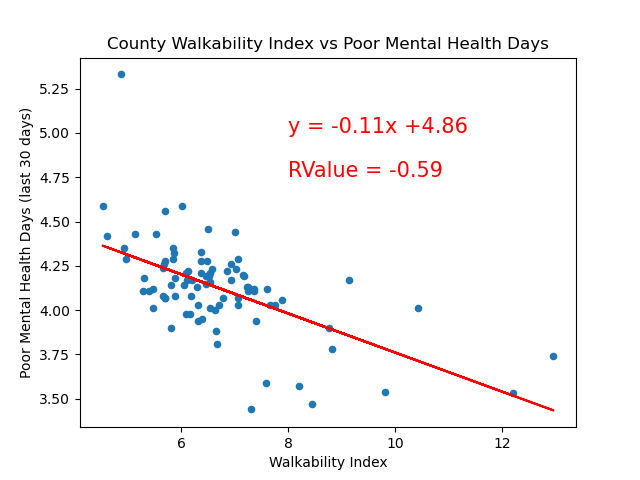


Plot 1.2 Top five and bottom five Minnesota counties walkability index

**2.0 Walkability impact on physical and mental health**

In order to assess how walkability impacts physical and mental health, data was pulled in from United State County Health rankings data source. The US County Health rankings looks at various health measures across every county in the US. The dataset was merged with the walkability data based on counties in Minnesota. In order to view the walkability data by county, a weighted average based on total population per census block group was created.

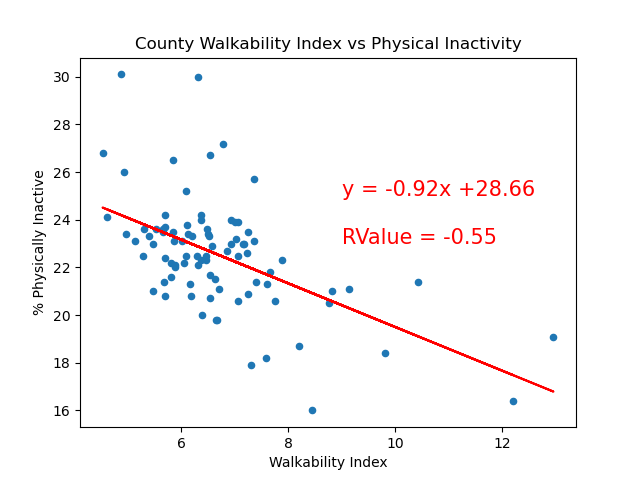
The metric used to evaluate walkability’s impact on mental health was “average number of poor mental health days reported in the last 30 days,” from the county health dataset. We plotted the two metrics in a scatter plot and ran a linear regression in order to determine any relationship. There is a moderate negative correlation (r = -.59) between walkability and poor mental health days. This means that as walkability in a county increases, individuals report having fewer poor mental health days on average.



Plot 2.1 Regression plot between walkability index and mental health

To analyze the relationship between walkability and physical activity, the specific metric ‘percent of adults reporting no physical activity, was taken from the county health dataset. This metric also had a moderate negative correlation (r = -.55) with walkability. As a county’s walkability increases, fewer adults report not having any physical activity or leisure.

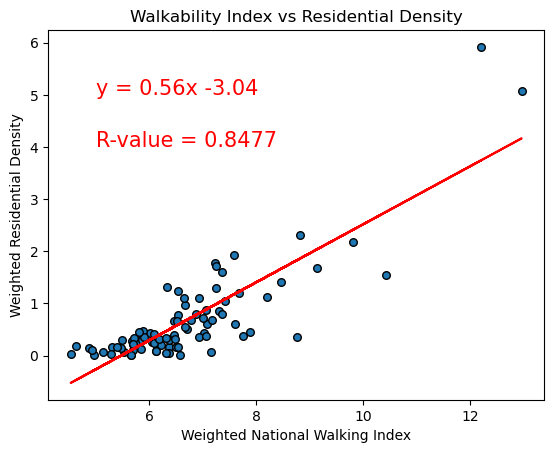
Overall, there is good evidence that having walkable communities leads to better physical and mental health for the population.



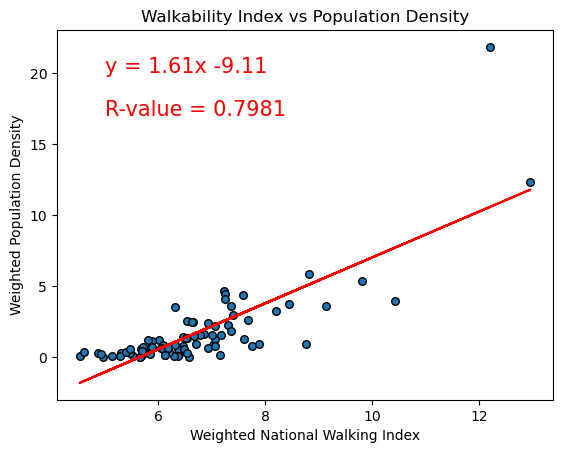
Plot 2.2 Regression plot between walkability index and physical inactivity

**3.1 Impact of population density and housing density on walkability**

Housing and population density has a significant impact on the walkability of a county. Both the residential and population density have a strong positive correlation with walkability with an r-value of 0.8477 and 0.7981, respectively.

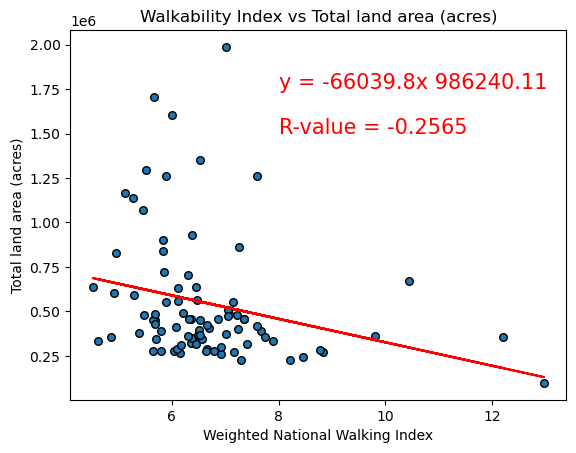


Plot 3.1 Regression plot between weighted walkability index and weighted residential density

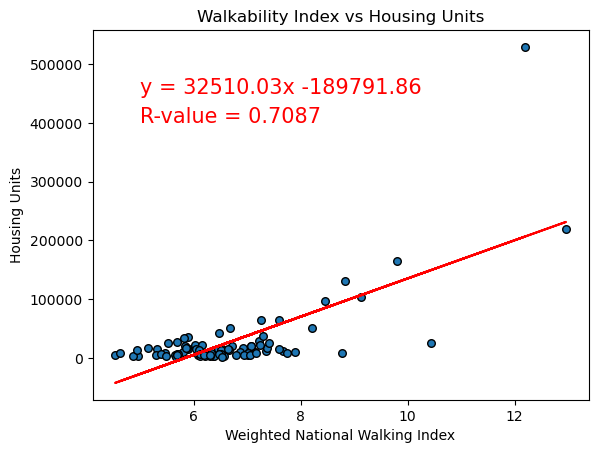


Plot 3.2 Regression plot between weighted walkability index and weighted population density

To investigate whether total housing units and total land area in each county also had an impact on walkability like density has, a correlation analysis was done. Total housing has a strong positive correlation (r = 0.708), total land has a weak negative to no correlation (r = -0.0385 ) with walkability when all the counties were included in the analysis.



Plot 3.3 Linear regression plot between weighted walking index and total land area(acres)



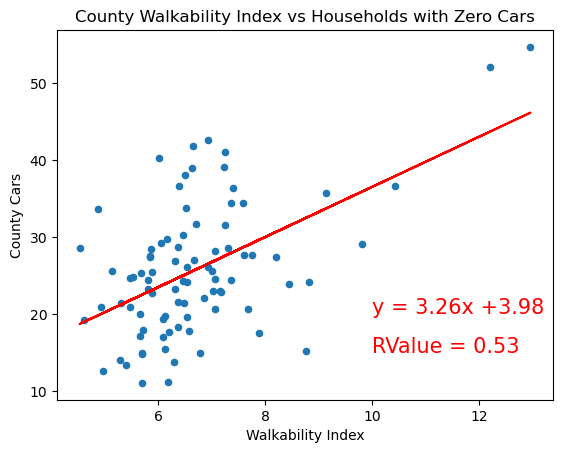
Plot 3.4 Linear regression plot between weighted walking index and housing units

Further analysis of the data revealed that St Louis County, had a large outlier in the dataset with 3.998196e+06 acres compared to Koochiching County which came as the second largest county at 1.986985e+06 acres. When analyzing the data after removing the outlier, there was a slight increase in the correlation (r = -0.256) (Plot 3.3 and 3.4).

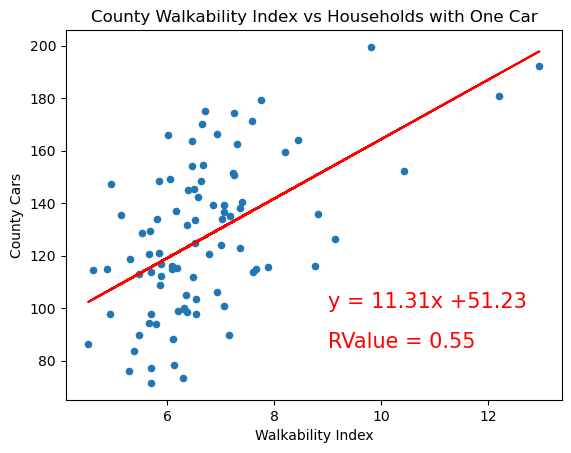
**4.1 Impact of cars per household on the walkability index**

To analyze the impact of numbers of cars per household on the walkability index, analysis was done with data as i) households with zero cars, ii) households with one car, and iii) two or more than two cars per household.

Analyzing the data (Plot 4.1 and Plot 4.2), there was not much difference whether a household has zero or one car. There is a moderate positive correlation associated with the walkability index (r = 0.530 and 0.550).

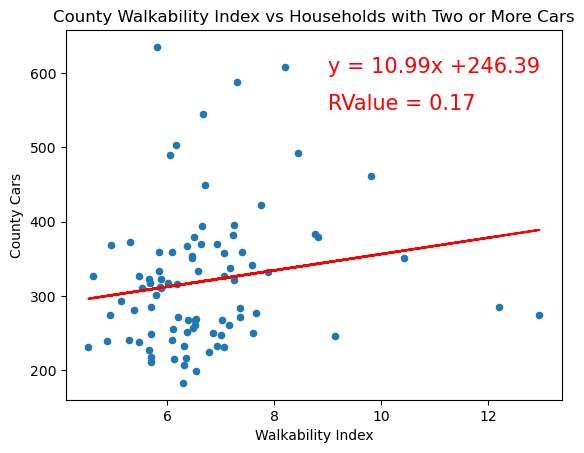
****

Plot 4.1 Regression plot between walkability index and zero cars household

****

Plot 4.2 Regression plot between walkability index and one cars household

A significant decrease in correlation (r = 0.17) was found with two or more cars per household (Plot 4.3). This kind correlation result makes sense since one car only might not be enough for an entire household, so some might not have a choice but to walk. While households with two or more cars have more available cars for them to choose not to walk.

****

Plot 4.3 Regression plot between walkability index and one cars household

**5. 0 Analysis of different wage group employees association with Walkability Index**

One of the attributes, employment does have a significant contribution on the walkability index. Walking is a self support habit. How much time people spend to earn for living and time spent for leisure contributes to the walking habit. In turn, different wage categories like low wage, medium wage and high wage people will have different access to the resources available nearby and that will contribute to developing their walkability.

In this section, an attempt is made to visualize which wage group of employees has higher index value and how the wage group data is associated with the walkability index. The employment data is categorized in three sets.

i) employee wages less than $1250 per month,

ii) employee wages between $1250 and $3333 per month, and

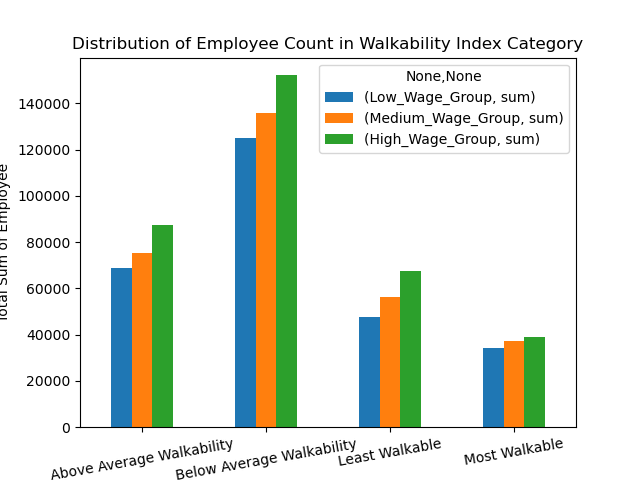
iii) employee wages above $3333 per month.

Correlation and linear regression analysis is done to assess the relationship among wage group employees and walkability index.

Table 5.1 Statistical Description of Dataset

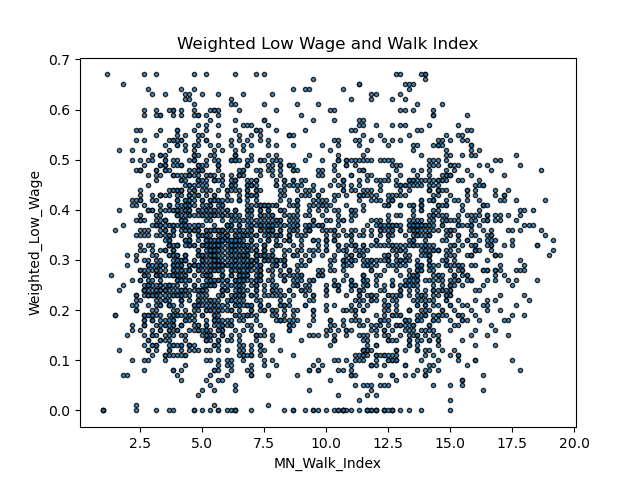
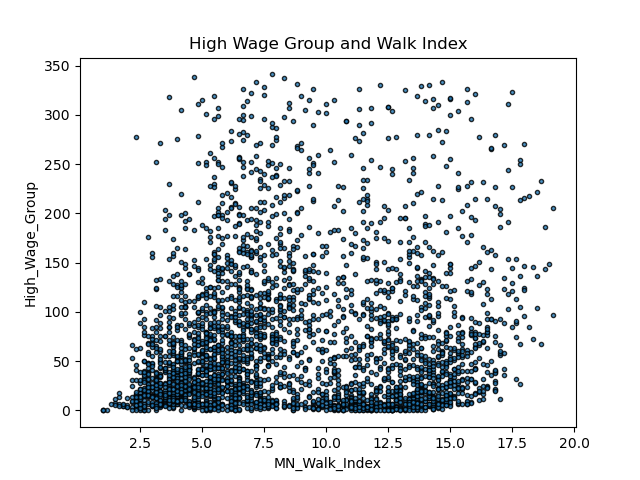
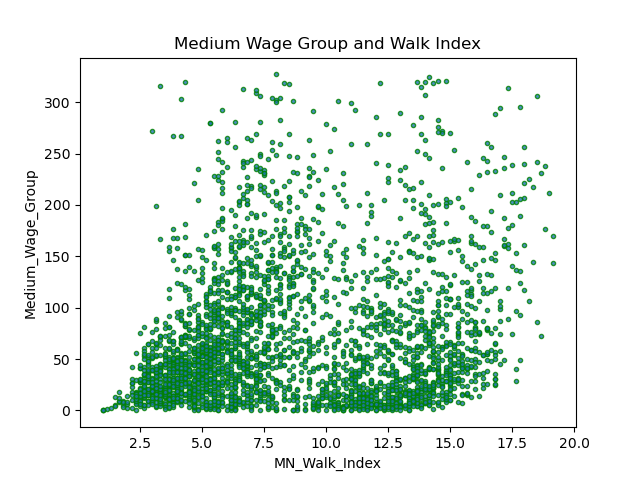
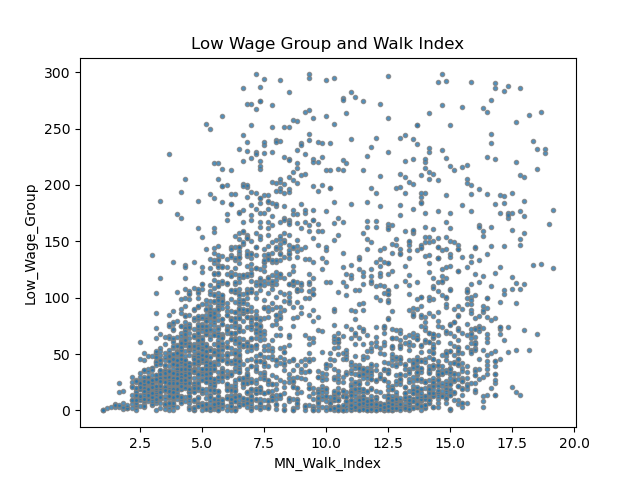
|  | **mean** | **median** | **variance** | **Std. dev** | **Std. error** |
| --- | --- | --- | --- | --- | --- |
| **Walkability index** | 9.18 | 8.5 | 18.09 | 4.25 | 0.06 |
| **Low wage group** | 160.9 | 65.0 | 113433.04 | 336.79 | 5.25 |
| **Medium wage group** | 194.4 | 73.0 | 168241.78 | 410.17 | 6.39 |
| **High wage group** | 336.57 | 76.0 | 1.8 | 1351.49 | 21.07 |
| **Weighted low wage** | 0.304 | 0.3 | 0.02 | 0.15 | 0.002 |

Variance in the dataset (Table 5.1) was very high which means the data was not normally distributed. There are greater chances the data was skewed. Data was analyzed for the outliers and was removed accordingly.



**Plot 5.1. Distribution of wage category employee count in walkability index category**

Employee count is highest in the High wage group of employees followed by Medium and Low wage employees in Minnesota (Plot 5.1). The maximum proportion of wage group employees works in the Below and Above average walkability category of workplace. Lowest employee count was found in the most walkable category.

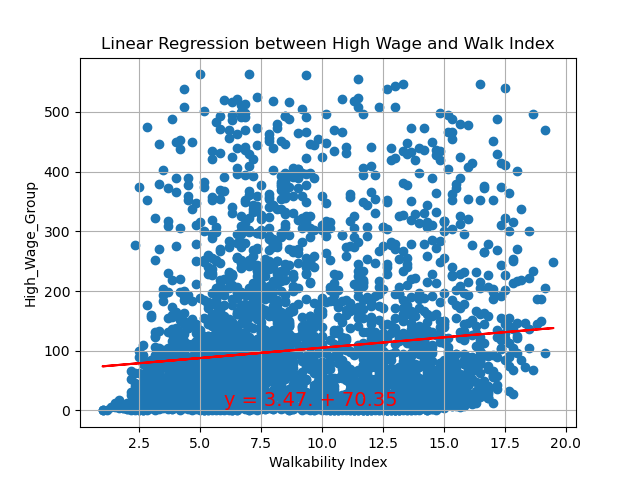
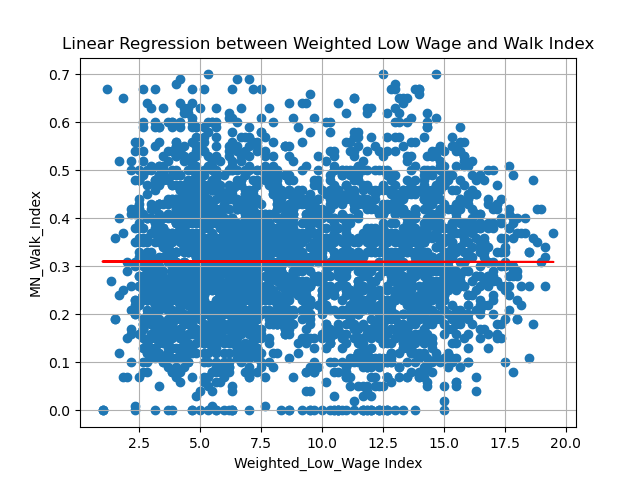
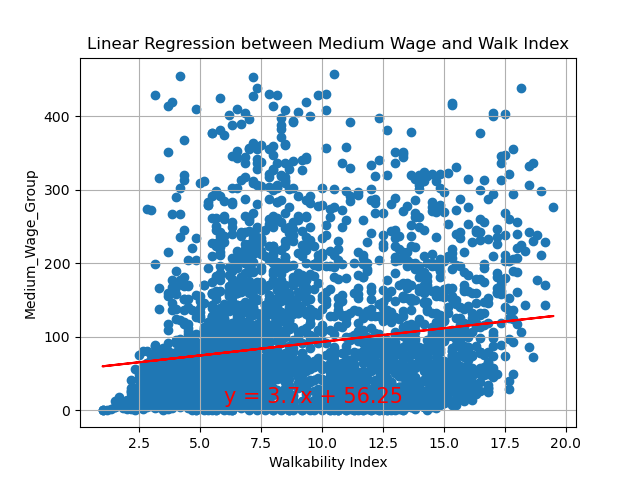
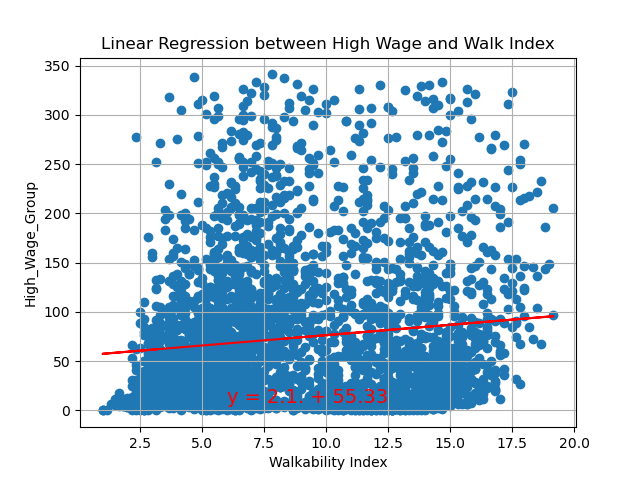
. 

**Plot 5.2. Scatter Plot between walkability index and wage groups**

Scatter plots (Plot 5.2) display the wide spread in the data set and accordingly, the correlation established between the variables is low (r = 0.119 to 0.190) but the association was positive (Table 5.2).

Table 5.2 Correlation matrix between walkability index and wage group

| **CORRELATION MATRIX** | Census Block group | Low Wage Group | Medium Wage Group | High Wage Group | Weighted Low Wage | MN Walk Index |
| --- | --- | --- | --- | --- | --- | --- |
| Census Blockgroup | 1.000 |  |  |  |  |  |
| Low Wage Group | 0.036 | 1.000 |  |  |  |  |
| Medium Wage Group | 0.027 | 0.842 | 1.000 |  |  |  |
| High Wage Group | 0.046 | 0.613 | 0.730 | 1.000 |  |  |
| Weighted Low Wage | 0.015 | 0.299 | -0.020 | -0.281 | 1.000 |  |
| MN Walk ndex | -0.089 | 0.190 | 0.183 | 0.119 | -0.002 | 1.000 |



**Plot 5.2. Regression plot between walkability index and wage groups**

There is a very low association established between the walkability index and the wage groups. None of the wage group employee variables has a very strong correlation with the walkability index. However, the association established was positive. Linear regression fit was able to explain 3% of variability in the Walkability Index (Plot 5.2). One explanation for this association can be that there was a wide variation in the data set. The variance was high that might have caused the skewness in the data. By smoothing the data and reanalyzing can bring some changes in the association.

**6.0 Conclusions**

* As the average walkability score across MN counties increases, the average number of poor mental health days decreases. Walkable neighborhoods boost mental health!
* Both higher Population and Residential density have a strong positive impact on the walkability of a county. However, while the total amount of housing units in a county has a strong positive impact on walkability, total land area (acres) has a weak negative to no correlation with walkability. In conclusion, more people and housing results in better walkability, but surprisingly a larger county does not factor into walkability.
* There is a very weak positive correlation between households with two or more cars and walkability.
* None of the Wage Group Employees had a very strong correlation with the walkability Index. However, the association established was positive.Linear Regression was able to explain 3% of variation in Walkability Index.

**References**

1. www.census.gov/programs-surveys/geography/guidance
2. https://www.countyhealthrankings.org/
3. https://www.epa.gov/sites/default/files/2021-06/documents/epa\_sld\_3.0\_technicaldocumentationuserguide\_may2021.pdf
4. https://www.epa.gov/sites/default/files/2021-06/documents/national\_walkability\_index\_methodology\_and\_user\_guide\_june2021.pdf
5. https://www.census.gov/programs-surveys/geography/about/glossary.html#par\_textimage\_4
6. https://catalog.data.gov/dataset/walkability-index