# Assessing the Level of Walkability for Women Using GIS and Location-based Open Data: The Case of New York City

## SUPPLEMENTAL INFORMATION - Methodological Approach

The methodological approach which sets the current research is based on an extended GIS analysis aimed at assessing the level of walkability for women in NYC. In particular, the objective was to identify and characterize a short list of census block groups (a geographical unit used by the United States Census Bureau, which is between the census tracts and the census blocks) where to prioritize interventions focused on enhancing the level of walkability for female pedestrians.

To do so, a series of geolocated datasets were retrieved, sorted, and filtered from several open-data repositories, geoportals and census databases (see Table 2). Data was selected considering the results of the proposed thematic literature review focused on the needs of women while walking. The indicators were analyzed to design a multi-layer map of NYC, based on the spatial distribution of each dataset. In particular, the selected datasets were analyzed considering the spatial distribution of punctual and areal vectors on census block groups. For comparing the various indicators among them, each one has been normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

The indicators were then included in the calculation of the Level of Usefulness Index (LUI), Level Comfort Index (LCI), Level Safety and Security Index (LSSI), and Level of Attractiveness Index (LAI) of NYC for female pedestrians. This was essentially based on a weighted summation of the Z-scores of the variables proposed in this study as well as the calculation of the Walkability for Women Index. Percentile frequency distribution of results made possible the identification of the census block groups (cbg) and Neighborhood Tabulation Areas (NTA) characterized by the lowest Level of Usefulness Index (LUI $\leqslant$ .205 – 10$^{th}$ perc.), Level of Comfort Index (LCI $\leqslant$ .258 – 10$^{th}$ perc.), Level of Safety and Security Index (LSSI $\leqslant$ .258 – 10$^{th}$ perc.), Level of Attractiveness Index (LAI $\leqslant$ .291 – 10$^{th}$ perc.), and Walkability of Women Index (WWI $\leqslant$ .291 – 10$^{th}$ perc.).

Finally, the results of the analysis were further filtered taking into account additional information related to census data and land use data. In particular, the Neighborhood Tabulation Areas (NTA) characterized by the absence of female population, and those characterized by the presences of parks and cemeteries, were not considered for presenting the results of the analysis.

**Level of Usefulness Index**

The calculation of the Level of Usefulness Index (LUI) was based on the City Planning Facilities Database, which aggregates information about facilities that impact NYC neighborhood quality and are owned, operated, funded, licensed, and certified by a City, State, or Federal agency. The considered dataset spans several domains: Health and Human Services (He); Education, Child Welfare, and Youth Facilities (Ed); Public Safety, Emergency Services, and Administration of Justice (Pu); Core Infrastructure and Transportation (In).

Data analysis was based on a series of isochrone maps showing lines of travel time by walking to reach each facility on a 1-10 scale (from 0 to 1 minutes = 10; from 14 to 15 minutes = 1), which were combined to the calculation of the distribution of each dataset on census block groups (cbg). This was aimed at estimating the level of accessibility of relevant services for women by walking. The calculation of the LUI was based on the weighted summation of the Z-scores of each indicator (see Equation 1). According to the results of proposed literature review and thanks to a preliminary sensitivity analysis on weights definition, the constant parameters KHe, KEd, KPu, and KIn were equally balanced ($\sum$ constant parameters = 1). Results were normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

$$\sum_{k=0}^{1} LUI = {}_{K_{He}} Z\_He_{cbg} + {}_{K_{Ed}} Z\_Ed_{cbg} + {}_{K_{Pu}} Z\_Pu_{cbg} + {}_{K_{In}} Z\_In_{cbg} \quad (1)$$

**Level of Comfort Index**

The calculation of the Level of Comfort Index (LCI) was based on a series of datasets related to comfortable pedestrian infrastructures and public spaces: Sidewalks Width (Sw); Sidewalks Area (Sa); Park Zones (Pa); Open Spaces (Op). This allowed us to consider the planimetric base map polygon layer containing sidewalks width and area, open space features, and the localization of beach zones, flagship park zones, large park zones, neighborhood parks, neighborhood playgrounds, and sitting areas.

Regarding sidewalks features, data analysis was based on calculating the distribution of each dataset on census block groups (cbg). Regarding the other indicators, instead, data analysis was based on a series of isochrone maps showing lines of travel time by walking to reach each park zone and open space on a 1-10 scale (from 0 to 1 minutes = 10; from 14 to 15 minutes = 1). This was aimed at estimating the level of comfort of urban areas for women while walking. The calculation of the LCI was based on a weighted summation of the Z-scores of each indicator (see Equation 2). According to the results of proposed literature review and thanks to a preliminary sensitivity analysis on weights definition, the constant parameters KSw (corresponding to .3), KSa (corresponding to .3), KPa (corresponding to .2), and Kop (corresponding to .2) were weighted to accentuate the impact of pedestrian infrastructures on LCI (∑ constant parameters = 1). Results were normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

$$\sum_{k=0}^{1} LCI = {}_{K_{Sw}} Z\_Sw_{cbg} + {}_{K_{Sa}} Z\_Sa_{cbg} + {}_{K_{Pa}} Z\_Pa_{cbg} + {}_{K_{Op}} Z\_Op_{cbg} \quad (2)$$

**Level of Safety and Security Index**

The calculation of the Level of Safety and Security Index (LSSI) was based on a series of datasets related to pedestrian road safety and security issue: Priority Intersection (Pr); Enhanced Crossings (En); Neighborhood Slow Zones (Ne); NYPD Complaint Map (Co). This allowed us to consider the localization of road intersections with the highest number of pedestrians killed and severely injured, the crossings treated with pedestrian ramps, pedestrian warning signs and high-visibility crosswalk markings, and the neighborhood areas with a 20 mph speed limit. Regarding security issues, the considered dataset includes all violation crimes reported to the New York City Police Department in relation to harassment, obscenity performance rape, sexual crimes and abuses (considering offenses occurring at intersections).

Data analysis was based on calculating the distribution of each dataset on census block groups (cbg), aiming at estimating the level of safety and security of urban areas for women while walking. The calculation of the LSSI was based on a weighted summation of the Z-scores of each indicator (see Equation 3). According to the results of proposed literature review and thanks to a preliminary sensitivity analysis on weights definition, the constant parameters KPr (corresponding to -.2), KEn (corresponding to .1), KNe (corresponding to .1), and KCo (corresponding to -.6) were weighted to accentuate the impact of violation crimes on LSSI (∑ constant parameters = 1). Results were normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

$$\sum_{k=0}^{1} LSSI = {}_{K_{Pr}} Z\_Pr_{cbg} + {}_{K_{En}} Z\_En_{cbg} + {}_{K_{Ne}} Z\_Ne_{cbg} + {}_{K_{Co}} Z\_Co_{cbg} \quad (3)$$

**Level of Attractiveness Index**

The calculation of the Level of Attractiveness Index (LAI) was based on a series of datasets related to vibrant and social public spaces: Sidewalk Cafes (Ca); Sidewalk Restaurants (Re). This allowed us to consider those cafes and restaurants which offer outdoor dining service for eating, reading, working, meeting a friend or taking a rest.

Data analysis was based on calculating the distribution of each dataset on census block groups (cbg), aiming at estimating the level of attractiveness of urban areas for women by walking. The calculation of the LCI was based on a weighted summation of the Z-scores of each indicator (see Equation 4). According to the results of proposed literature review and thanks to a preliminary sensitivity analysis on weights definition, the constant parameters KCa and KRe were equally balanced (∑ constant parameters = 1). Results were normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

$$\sum_{k=0}^{1} LAI = {}_{K_{Ca}} Z\_Ca_{cbg} + {}_{K_{Re}} Z\_Re_{cbg} \tag{4}$$

**Walkability for Women Index**

The results related to the estimated Level of Usefulness Index (LUI), Level of Comfort Index (LCI), Level of Safety and Security Index (LSSI), and Level of Attractiveness Index (LAI) were further analyzed to calculate the proposed Walkability for Women Index (WWI). This was essentially based on a weighted summation of the Z-scores of each index proposed in this study (see Equation 5). According to the results of proposed literature review and thanks to a preliminary sensitivity analysis on weights definition, the constant parameters KLUI (corresponding to .2), KLCI (corresponding to .2), KLSSI (corresponding to .4), and KAI (corresponding to .2) were weighted to accentuate the impact of the level of safety and security on WWI (∑ constant parameters = 1). Results were normalized on a 0-1 scale, creating Z-scores that follow the normal distribution of the values.

$$\sum_{k=0}^{1} WWI = {}_{K_{LUI}} Z\_LUI_{cbg} + {}_{K_{LCI}} Z\_LCI_{cbg} + {}_{K_{LSSI}} Z\_LSSI_{cbg} + {}_{K_{LAI}} Z\_LAI_{cbg} \tag{5}$$