PRE-REGISTRATION / WORKPLAN

Data Mastery Challenge (2024-1B)

Submitted by:

Rakibun Athid

Nawwar Fatima Procheta

Title: Green and Fast Routes for Cyclists and Pedestrians in Utrecht, Netherlands

Description:

In this study, we attempt to replicate and extend the analysis by Ludwig et al. (2021) on the contrasting green and fast routes for cyclists and pedestrians (Dresden and Heidelberg (Germany). It will then replicate the city of Utrecht in Netherlands, with its extensive cycling infrastructure and urban greenery. It provides an analysis of whether these findings are consistent in the new urban context, and whether methods and data sources need to be adapted to these new urban conditions in Utrecht. Most notably, data sources used to calculate the green index are updated, as is its calculation, and terrain steepness is considered in the route weighting function; the local preferences are considered. This study contributes to understanding the broader applicability of green routing methodologies by incorporating Utrecht's unique urban features.

Hypotheses

1. **H1:** While fast routes will be faster, green routes in Utrecht will exhibit higher greenness values than fast routes but will take more time.
2. **H2:** Green routes with fewer detours will be available in urban areas with higher vegetation density in compared to areas with sparse vegetation.
3. **H3:** In neighborhoods with better infrastructure and green spaces, higher population density will lead to greater overlap between green and fast routes.

Design Plan

Study Type: Replication

Study Design:

- **Scope:** For analysis, Utrecht, Netherlands, known for its excellent cycling infrastructure and as the city with the most integration of urban greenery, will be used. The methods are adjusted to the city's unique road network as well as the distribution of greenery in the city.
- **Changes:**

  1. Estimating greenness using OSM data, Sentinel-II imagery and Utrecht municipal datasets.
  2. As neither 3D laser scan data is available, green index calculation relies on OSM and Sentinel-II imagery with adaptations for improved accuracy.
  3. The route weighting function is integrated with terrain steepness to best model cyclist and pedestrian preferences.

Sampling Plan

Existing Data:

- OSM data for road networks, cycle paths, and pedestrian pathways.
- Sentinel-II imagery for vegetation coverage.
- Municipal open data from the City of Utrecht Open Data Portal.

Data Collection Procedures:

1. GIS tools and Python will be used to preprocess the road network and green coverage data.

2. Random routes will be generated within Utrecht's administrative boundaries (start and endpoints).

3. Simulation will include:

Cycling Trips: 1000 routes (originally 5000)

Pedestrian Trips: 1000 routes (originally 5000)

Total Routes: 2000 routes

Variables

Measured Variables:

Measured Variables:

1. Route Duration (minutes): Total Travel time for both fast and green routes.

2. Green Index: Vegetation coverage fraction within a 30-meter buffer to the route.

3.Route Similarity (%): Proportion of overlap between fast and green routes.

4.Deviation (%): Differences in travel time and route distance of fast and green routes.

Analysis Plan

1. Spatial Analysis

Using Python:

- **Data Extraction and Preprocessing:**
    - Extract road networks using OSMnx and clean the data.
    - Use rasterio to process Sentinel-2 imagery for vegetation indices like NDVI.
- **Route Simulation:**
    - Use OSMnx and NetworkX for shortest path routing:
        - Fast route: Minimize travel time.
        - Green route: Optimize for greenery using a weighted approach.

2. Statistical Analysis

Using Python:

- **Descriptive Statistics:**
    - Use pandas to summarize route characteristics (e.g., average travel time, green index).
    - Compare means and distributions of fast vs. green routes.
- **Regression Analysis:**
    - Use Statsmodel or Scikit-learn to examine the relationship between travel time, green index, and other variables.

Use QGIS's built-in tools to calculate area coverage or density of green spaces within route buffers.

Challenges:

1. Data quality variability across urban areas.

2. Low resolution spatial coverage of vegetation data.

Potential for Scaling:

1. Other cities with similar urban features should be able to generalize our findings. With sufficient road and vegetation data available, the methodology shows transferability.

Ethical Considerations:

1. Protect the data privacy in particular for municipal datasets and real-world data validation.

Additional Information:

2. Literature cited: Studies relevant urban planning and GIS, Ludwig et al. (2021)

3. Software: Python, GIS.

4. Data will be released publicly upon publication following data licensing and ethical guidelines.