

Proposal for Semester Project

Patterns & Trends in Environmental Data / Computational Movement Analysis Geo 880

Semester:	FS25
Data:	Google Timeline
Title:	Tracking Mobility: Unraveling Movement Patterns and Greenhouse Gas Impact
Student 1:	Chryssolouris Dimitri
Student 2:	Hauser Ramon

Abstract

This study analyzes the movement patterns of two students using Google Timeline tracking and SimaPro for GHG emissions assessment. By segmenting rest and movement phases, identifying key locations, and comparing transportation modes, the project quantifies mobility behaviors and environmental impact. The findings will be visualized through maps and statistical comparisons, ensuring an insightful sustainability analysis.

Research Questions

1. How do the movement patterns of the two students differ?
 - a. What is the ratio between rest and movement? (Segmentation)
 - b. Which is the most frequently visited location?
 - c. How does the use of transportation modes differ between the samples over a specific period?
 - d. What impact do the movement patterns have on each student's CO₂ footprint?

Results / products

The research is expected to reveal notable differences in movement patterns between the two students. For instance, one student rarely uses a car, which could lead to lower greenhouse gas (GHG) emissions, while the other may rely more on motorized transport during the analyzed timewindow, increasing their carbon footprint. The study will quantify the impact of transportation choices on GHG emissions, determine the ratio of rest to movement, and identify the most frequently visited locations. The results will be presented through data visualizations, comparative charts, and sustainability assessments, offering insights into how mobility behaviors influence environmental impact.

Data

The primary data source will be Google Timeline tracking, which will provide detailed movement patterns, including transportation modes, travel distances, and time spent at different locations. The data collection window will be one month, ensuring both students' mobility data is aligned for accurate comparison. Additionally, GHG emissions data will be obtained from SimaPro, which will help quantify the environmental impact of different transportation choices. This will require contextual data on transport emission factors (e.g., CO₂ emissions per km for cars, public transport, cycling, etc.). Further contextual data, such as university schedules, or personal habits, may be useful for interpreting movement patterns. At this stage, not all data has been collected, but Google Timeline tracking and SimaPro will serve as key sources for analysis.

Analytical concepts

Define movement patterns in terms of distance traveled, frequency of trips, and time spent in motion vs. at rest as well as transportation modes and their emission caused. Define “rest” as periods of inactivity (e.g., staying at home, being in one location for extended periods) and “movement” as active transportation or mobility events. Categorize transportation modes (e.g., walking, biking, public transport, car). Collect data on usage frequency, duration, and distance for each mode over a set period (e.g., one week or one month). Compare patterns between the two students. Measure GHG emissions based on transportation modes used (e.g., car, bike, public transport) and distances traveled. Use an established GHG emission calculator (i.e., SimaPro or publicly available emission factors) to quantify the footprint.

R concepts

For this analysis, R will be used for data cleaning, spatial analysis, visualization, and statistical modeling. Key concepts and packages include:

- a. Data manipulation: dplyr, tidyverse for cleaning and structuring Google Timeline data.
- b. Spatial analysis: sf, and raster to analyze GPS locations, travel distances, and movement patterns.
- c. Visualization: ggplot2, tmap, and leaflet for creating maps, route heatmaps, and movement patterns.
- d. GHG emission calculations: EcoInvent dataai for CO₂ footprint calculations.
- e. Time series analysis: lubridate for analyzing movement trends over the one-month period.

Additional spatial analysis methods may include kernel density estimation (KDE) for hotspot detection and network analysis for transportation mode comparisons.

Risk analysis

A potential challenge in this study is missing or incomplete GPS data from Google Timeline. These gaps may arise due to GPS inaccuracies, device settings, or network issues, leading to incomplete movement records. To address this, missing data will be extrapolated using interpolation techniques, such as linear interpolation or time-weighted averages, ensuring continuity in movement trends.

Another issue is the misclassification of transportation modes, where Google Timeline may inaccurately categorize movement (e.g., mistaking walking for biking). To improve data reliability, transportation modes will be cross-verified using movement speed, timestamps, and contextual information, allowing for necessary corrections and more accurate mobility analysis.

Questions?

- a. Given the project timeframe, are we spending our effort in the right places?
- b. Are there any aspects we should simplify or focus on to keep it manageable?
- c. Would a mix of maps, stats, and comparisons be a good way to present our findings?
- d. Do we cover enough methods that we learned in the course? (Fokus on Segmentation, Similarity)
- e. Are there any additional analytical concepts or methods we should consider for a more comprehensive analysis?