

# Proposal for Semester Project

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

Semester:	FS24
<b>Data:</b>	Walking activity recorded with Strava
<b>Title:</b>	Differentiation of walking patterns
<b>Student 1:</b>	Saskia Gianola
<b>Student 2:</b>	Sarah Wirth

## Abstract

Based on walking data recorded with Strava we want to derive the purpose of the movement. Using the attributes sinuosity, speed and number of stops we want to find out whether an activity belongs to a commuting path, is recreational or belongs to running an errand for example grocery shopping. We want to test if the trajectories differ enough to derive walking patterns. As we recorded the data ourselves, we can add semantic information and test whether the classification is possible.

## Research Questions

Is it possible to derive the type of activity from movement data considering the attributes sinuosity, speed and visited locations? Which type of analytical concept performs best in differentiating between the different activities? Do different concepts perform differently for different people?

## Results / products

We expect to be able to distinguish between different types of activity (commuting path, recreational, running errands). The trajectories are expected to differ regarding sinuosity, speed and number of stops.

As a result we aim to have an estimation of the performance of different analytical concepts at the task of differentiating. This could be presented as misclassification tables for each analytical concept in which we assess which concept succeeded best at discerning the different walking-activities.

## **Data**

We use the activity data from Saskia and maybe another person. If we would be able to start a second data-collection with Saskia or get additional walking data from other students that would be ideal.

## **Analytical concepts**

We intend to use similarity measures to find similar trajectories. Furthermore, we want to extract characteristics of the trajectories to determine sinuosity, speed and number of stops for each trajectory. Then we want to classify the trajectories based on these characteristics. Also, we intend to add information about starting- and endpoints to improve classification. Last, we want to visualize the paths and groups of paths.

## **R concepts**

`library("XML") library("leaflet") library("sf") library("tmap") library("ggplot2") library("tidyverse")` We also intend to write our own functions to facilitate certain steps.

## **Risk analysis**

Our analysis depends on whether our data contains enough trajectories. Therefore we might need more Data. Our plan B is going for travel-mode detection using ca. 2 months of Google-Timeline data, which has rather big temporal intervals, which would be another risk for our plan B. Adding data from another person might also add a challenge as we would need to distinguish between “normal” and “fast” walking-speeds for each person.

## **Questions?**

How many trajectories do we need to have a sufficient amount of data for our project? As the task of differentiating