# Introduction and Background

During the spring of 2021, I finished my engineering studies and got a job opportunity in Stockholm. During my studies, I have moved a couple of times, and I have enjoyed some areas more than others. My current living area is by far my favourite. For this project, I was wondering if data science could help inform me about similar areas in Stockholm.

A programme that could use information about a person’s current, or previous, living area to identify suitable areas for relocation could prove useful. Consumers could use the tool to understand what areas they would thrive in. Realtors could also use the programme to provide recommendations for similar areas to their clients and city planners could even use it to understand what characterizes popular living areas. Previously, consumers would rely on stories from friends, or the general reputation of an area to evaluate whether it could be a nice place for them. Using a clustering algorithm and data from foursquare, the guess work is decreased, and consumers could find areas that more suited their needs and preferences. While realtors need to be socially skilled to determine what the needs of a potential client are, and a programme like this could help them to identify customers for objects and provide other suitable proposals if there are many customers interested in the same object. Used on a larger scale, cities may seem an increase in popularity in areas that match people’s needs and preferences, but no not have the best reputation. City planners could also use the information to figure out what combination of services and venues in a neighbourhood attracts certain people. A city with an aging population may for example look at cities where young people live to figure out how to attract them to their city.

# Data

In Sweden, every address has a postal code. The postal code is unique to a neighbourhood and mainly used for mailmen and other logistics actors to deliver mail and packages. The postal code is also accompanied by a city name. Large cities and municipalities will have many postal codes and city names, in that case the city name can refer to a part of a city. Stockholm municipality has 27 such cities, with each city containing a varying number of postal codes. The postal codes in Stockholm municipality will be compared to my current postal code. Data was collected from three main sources:

* The website Postnummerservice (postal code service) was scraped for the cities within Stockholm municipality.
* The database PAP/Lite was used to collect the postal codes for each city and their coordinates. An API was used to connect to the data base, which was free for personal use.
* The Foursquare database provided the nearby venues for each postal codes, using their coordinates.