**Priced out: Factors influencing contemporary Dutch house prices**

**Research motivation**

Due to the ongoing Dutch housing shortage and increasing prices, it becomes harder for residents to buy their first house. 80% of adolescents looking to purchase their first house are in severe doubt about whether they will ever be able to buy a house. Especially for single households, this insecurity reaches high levels. It is estimated that single starters with a modal income (€38.000,-) can only get a mortgage worth €150.000,-, whereas the average house in the cheaper segment is priced twice this amount (BBNVARA, 2023). ([Geen 'rijke vriend'? Hoe kom je dan ooit aan een koophuis?! - BNNVARA](https://www.bnnvara.nl/artikelen/hoe-kom-je-als-alleenstaande-starter-ooit-aan-een-huis)).

The purpose of this project is to get an insight in the prices of houses and their characteristics. The research aims to create a profile of the most important factors that influence the house prices. A multiple linear regression model is built to estimate house prices based on their characteristics. In addition, it could help starters to compare the price of a house that they are targeting to the mean price of houses with similar characteristics.

**Data retrieval**

For this project, data from Dutch housing platform Funda was used. The data was retrieved using a programmed web bot by B. Lusse and shows all houses with corresponding information that were for sale through Funda between 08/01/2022 and 08/03/2022. More information about the scraping details can be found by following [this link](https://github.com/bryanlusse/HousePrices__Webscraper)). In addition tot he scraped data, geographic information was retrieved from [OpenDataSoft](https://public.opendatasoft.com/explore/dataset/georef-netherlands-postcode-pc4/table/) in order to connect data about the abstract locations of houses to the respective provinces and coordinates.