# Funda Data

## House Information

There are 4 files for the Funda data. Two of these contain the houses. One contains the realtors and one contains the review.

### houses.csv (94000 rows, 80 columns)

This file is the ‘raw’ data as it comes from funda.nl.

### houses-preprocessed.csv (85189 rows, 143 columns)

This file contains the preprocessed houses.

### houses-preprocessed-buurten.csv (78807 rows, 212 columns)

This file contains the preprocessed houses, merged with the CBS data on neighborhood level.

## Realtor Information

### agents.csv (4203 rows, 9 columns)

Contains information on the agent level.

## Reviews

### reviews.csv (119080 rows, 11 columns)

This file contains the individual reviews.

# CBS Data

CBS Data has been collected through CBS Statline. The following exports are used.

## Bevolking

Bevolking: Aantal inwoners, Aantal mannen, Leeftijdsgroepen, Burgerlijke staat

Population: Number of Residents, Number of Males, Age Groups, Marital Status

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=4-5,7-15,102&D2=a&HD=170301-1312&HDR=T&STB=G1>

## Bevolking

Bevolking: Allochtonen, Geboorte en Sterfte, Particuliere Huishoudens, Bevolkingsdichtheid

Population: Immigrants, Births and Deaths, Private Households (as opposed to companies), Population Density

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=16-17,23,25,28-32,102&D2=a&HD=170301-1313&HDR=T&STB=G1>

## Living

Wonen: Woningvoorraad, Gemiddelde Woningwaarde, Percentage eengezinswoning, Percentage bewoond, Koopwoningen, In bezit woningcorporatie, In bezit overige verhuurders, Bouwjaar vanaf 2000

Living: Number of houses, Average House Value, Percentage One Household Houses (as opposed to houses for multiple households, Percentage Inhabited, Number of Owned Houses, Number of Houses Owned by Housing Corporations, Number of Houses Owned by Other Types of Renters, Number of Housed Built after 2000

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=33-35,37,39,41-42,45,102&D2=a&HD=170301-1314&HDR=T&STB=G1>

## Income

Inkomen:

Income: Number of Residents with Income, Average Income per Resident with Income, Average Income per Resident, Proportion of Residents with nationally lowest 40% of income, Proportion of Residents with the highest 20% income nation-wide, Actives 15-75 years, Proportion of Households that Belongs to the 40% of households with lowest income, Proportion of Households that Belongs to the 20% of households with highest income, Households with a low income, households below or around social minimum.

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=63-72,102&D2=a&HD=170301-1316&HDR=T&STB=G1>

## Social Security

Sociale zekerheid:

Social Securities: Number of Residents on welfare, Number of Residents on AO, Number of Residents on WW, Number of Residents on AOW

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=73-76,102&D2=a&HD=170301-1316&HDR=T&STB=G1>

## Distance to Facilities

Nabijheid Voorzieningen:

Distance to Facilities: Distance to GP, Distance to Super Market, Distance to Daycare, Distance to Elementary School, Number of Elementary Schools within 3km

<http://statline.cbs.nl/Statweb/publication/?VW=D&DM=SLNL&PA=83220ned&D1=94-98,102&D2=a&HD=170301-1317&HDR=T&STB=G1>

These datasets are combined into ‘postcodes.csv’, a file with [x] rows and [x] columns describing each neighborhood.