

# A comparison of Rural Towns in the Netherlands

IBM DATA SCIENCE CERTIFICATE - CAPSTONE PROJECT

Laura Bongaardt  
Delft, Netherlands

July 18, 2021

## Abstract

*Here will be a summary/abstract of the work.*

## I. INTRODUCTION

It is often said that Rural Towns in Western societies are struggling to survive in modern times. Usually it concerns historical places that relied on manual labour jobs or agriculture for income. With automation and globalisation many of those jobs are gone and while cities found new roles in trade and innovation, Rural Towns are still lacking. In this text I want to take stock of the Rural Towns in the Netherlands, from a Data Science perspective, using publicly available data from the CBS, Postcode.site and the Foursquare API.

The objectives are:

- 1) Identify and locate Rural Towns of interest, based on Population Size and Distance to a city.
- 2) Make a comparison of Rural Towns, relative to cities and other towns, based on Population Growth and Average Income.
- 3) Determine which Rural Towns are successful and which are less successful.
- 4) Identify industries that are commonly present in successful Rural Towns.

If we can identify industries that are successful in Rural Towns, this might help to bring more opportunities to towns that are less successful. This text is intended for (local) gov-

ernments to create policies that stimulate local economies, for business men that want to invest in rural towns, people that are looking to move to the countryside or for residents that want to help their communities.

## II. DATA

### i. Central Bureau of Statistics (CBS)

- *Table 84992NED*: List of registered places, cities and towns, in each municipality in the Netherlands. Dated Jan 1, 2021.
- *Table 84799NED*: Population size and area size for each 4-digit postcode in the Netherlands. Data collected for year 2020.
- *Table 83502NED*: Population size for the period 2000-2020, for each 4-digit postcode. Last update Sep 4, 2020.
- *Table 84286NED*: Average income for each 4-digit postcode. Dated year 2018.

### ii. Postcode.site

List of all 4-digit postalcodes belonging with each registered place, city or town. Dated 2021.

### iii. Foursquare API

List venues in the neighborhood of selected Rural Towns.

### III. METHODOLOGY

#### i. Identify Rural Towns

We start by identifying and locating the places of interest. The Central Bureau of Statistics (CBS) collects many different indicators and statistics and makes them publicly available through their data portal: <https://opendata.cbs.nl/>. Data can either be downloaded in csv format or requested with an API. Through Table 84992NED we obtain a list of the registered places, cities and towns, for each municipality. Table 84799NED contains the population and area size for each 4-digit postcode. And by scraping from the tables on <https://postcode.site/> we can get a list of the 4-digit postcodes within each registered place. Latitude and Longitude are obtained through the GeoPy package.

By combining the data, we get a dataframe with registered places, their respective population size, area size, longitude and latitude. Distances and population density can be calculated from these. I am looking for small places located at some distance away from a city center. What constitutes 'small' or 'some distance away' are as of yet vague terms and is to be determined from the available data using distribution plots.

#### ii. Assessment of Rural Towns

After the identification of Rural Towns, I want an assessment of their welfare. Population Growth and Average Income are two basic indicators that can be sourced from tables from the CBS. Population growth will be listed for the period 2010-2020. Average Income, on a 4-digit postcode basis, was last measured in 2018, but we can use this as a relative indication.

We will compare these indicators with cities to get a general assessment of the performance of Rural Towns. Next, we will look at the difference between towns and divide them into best and worst performers.

#### iii. Neighborhoods of Rural Towns

We now have a list of Rural Towns in the Netherlands, whether they are good or bad performers and the Latitude and Longitude that were obtained in step i. So we can use the Foursquare API to list all the venues in each neighborhood. After categorizing the venues, we can build a profile for successful and less successful towns. By using Logistic Regression, and examining feature importance, we can determine whether the presence of certain industries contribute to the success of a town.

This way, I hope to identify which industries, except agriculture, are successful in rural towns. By promoting these industries, local governments might be able to stimulate the development of Rural Towns.