Predicting Anxiety Disorders

The key to data driven policy

26-05-2023

M.M.R. (Martijn) Demmink
Master thesis Industrial Engineering & Management
University of Twente

KPMG supervisor:
C.Y.M. (Charlotte) van Leeuwen, MSc

First university supervisor:
Dr. Ir. A.G. (Gréanne) Leeftink

Second university supervisor:
Prof. Dr. C.J.M. (Carine) Doggen

Summary

Introduction

Anxiety disorders are the most prevalent mental disorder in the Netherlands, with more than 1 out of 7 people affected. The goal of this study is to utilize public data from 2015-2020 to predict the prevalence of anxiety disorders for 2021-2026 and identify associated risk and protective factors of anxiety disorders. The study also aims to determine the key factors that should be targeted for interventions by the Dutch Ministry of Health on a national level and by municipalities on a local level.

Methods

We collect data on the number of anxiety disorders for 325 Dutch municipalities from 2015-2020 and social-economic and environmental factors from public data between 2015-2022 for the corresponding municipalities. Data on the number of anxiety disorders for 2021-2022 is not yet available. After cleaning the data, we continue with 130 input factors from public data and one output factor: the proportion of inhabitants with anxiety disorders per municipality. We measure the Pearson correlation of each input factor with the anxiety disorder prevalence. We then perform linear regression, lasso regression, and sensitivity analysis using a neural network to identify national and municipal-level risk and protective factors. By performing these techniques in this order, we gradually increase the complexity of the analysis and gain deeper insights into the relationship between the input factors and the anxiety disorders prevalence. We train and test the neural network on data from 2015-2020 using 5-fold cross-validation to make predictions about 2021-2026.

Results

We find the most positive Pearson correlation with anxiety disorder prevalence to be the proportion of inhabitants receiving social benefit ($r = 0.66$), while we find the most negative correlation with the percentage of inhabitants being married ($r = -0.66$). With lasso regression, we identify the median wealth of couples with no children ($\beta = 0.75$) as the most influential risk factor of anxiety disorder prevalence, and the median wealth of private households ($\beta = -1.06$) as the most influential protective factor of anxiety disorder prevalence. Based on the sensitivity analysis and neural network, we find the most influential risk factor of anxiety disorder prevalence to be the percentage of the population who are divorced (sum of weighted gradients = 0.0101), and the most influential protective factor to be the median wealth of private households (sum of weighted gradients = -0.0062). The neural network is able to predict the prevalence of anxiety disorders for each municipality from 2015-2026 with an MSE of 0.0038 and R^2 of 0.84 for 2015-2020.

Discussion

Our study identifies risk and protective factors related to anxiety disorders using machine learning techniques on public data from 2015-2020. The models produce varying rankings of factors, and we find the neural network to be the most reliable since it is able to capture complex, non-linear relationships between the factors and prevalence of anxiety disorders. Not all of our findings do align with previous literature which could be linked to the temporal scope and use of aggregated public data, which may not fully capture the complexity and prevalence of anxiety disorders.

Conclusions

Risk and protective factors for anxiety disorders are identified with machine learning models. Some factors may be hard to influence with policies, but do give insight into which groups are associated with higher prevalence of anxiety disorders which may guide targeted prevention. Furthermore, the study provides insights into the potential burden of anxiety disorders and the need for intervention, and it guides the development of effective strategies to reduce its prevalence and improve the quality of life for individuals affected by this mental health disorder.

1 Table of contents

1 Introduction

In the last years, there has been increasing awareness around the topic of mental health, as described in the World Health Organization (WHO) Sustainable Development Goals (WHO_1, 2021). Especially the COVID-19 pandemic created momentum to increase the awareness about mental health issues and increased accessibility of mental support (UN, 2021). Factors such as isolation, social distancing, economic instability, uncertainty, fear and increased media attention have increased the awareness of mental health issues (Long, 2020) (Jiang et al., 2020) (Brooks et al., 2020) (Choi et al., 2020).

Mental health related problems cause a lot of harm. For people aged between 15 and 29 years old, suicide is the fourth leading cause of death globally (WHO_2, 2021). For women aged between 15 and 19 years, suicide was even the 2nd leading cause of death in 2019 (Liu et al., 2022).

Data of the WHO show that for every death by suicide there are at least 20 suicidal attempts registered (WHO_1, 2022), not to mention the number of close relatives of victims who have to suffer because loved ones (want to) make an end to their life. This highlights the disease burden of mental health problems and raises the questions with what kinds of psychological problems people are dealing.

Anxiety disorders

The most common mental disorder are anxiety disorders with a prevalence of 1 out of 25 people who suffered from this disorder in 2019 according to the WHO (WHO_2, 2022). A study from the US found that if people ever have made a suicidal attempt in their life, 71% reported to have experience with at least one occurrence of an anxiety disorder (Nepon et al., 2010).

Four longitudinal population-based studies have been conducted to measure and compare the prevalence of current mental disorders, defined using psychiatric diagnostic criteria, before and during the pandemic. Out of these studies, two found no change (Knudesen et al., 2021) (Ayuso-Mateos et al., 2021), one found a decrease (Vloo et al., 2021), and one found an increase in the prevalence of anxiety disorders (Winkler et al., 2021). These studies were often small-scale, limited to high-income countries, conducted in specific locations, and utilized different modes of assessment (such as online versus in-person) to evaluate mental health before and during the pandemic (Penninx et al, 2022). Therefore, on a global scale, it is uncertain whether the temporary increase in mental symptoms related to the pandemic have translated into the development of mental disorders.

Anxiety is normally present in healthy individuals in limited amounts and serves the purpose of allowing the body to be more alert to potential dangers and to improve our focus. In the case of an anxiety disorder, there is an excessive amount of anxiety which causes someone to function improperly, to overreact to certain emotion and to not be in control of his/her responses (Cleveland Clinic, 2020) (APA, 2021). Individuals who have anxiety disorders experience believe that they are out of control of a situation or feel overwhelmed by worry which can manifest into crankiness, sweatiness or shakiness (Watson, 2021).

There are different types of anxiety disorder. The five most prevalent types are (Bandelow, 2015):

- Generalized Anxiety Disorder (GAD)
  GAD involves an ongoing feeling of anxiety in patients for no obvious reason which limits someone his/her functioning in daily life. GAD is different than occasionally worrying and is unrealistic or out of proportion for the situation (WebMD_1, 2021).

- Panic Disorder
  Panic disorder is characterized by unexpected and repeated episodes of immense fear which goes along with physical symptoms such as a pounding heartbeat, chest pain, shortness of breath, breakouts of sweat and dizziness (WebMD_2, 2021).

- Social Anxiety Disorder (SAD)
  SAD is characterized by excessive self-consciousness and overwhelming anxiety in everyday social interactions. There is a constant fear of being judged, embarrassed or ridiculed by others. (WebMD_3, 2021)

- Phobia-related disorders
  These phobias are a lasting and unreasonable fear caused by the presence or thought of a specific object or situation that generally does not cause any danger. In these cases, people experience intense anxiety or try to avoid these stress-full situations or objects. This behavior can interfere with a person its ability to function properly in daily life (WebMD, 2022).

- Agoraphobia
  People with agoraphobia have an intense fear of being in a situation where escape might me difficult or that help would not be available if an emergency happens. Someone with agoraphobia may be afraid of leaving home, being in open or enclosed spaces, travelling on public transport and stranding in line or being in a crowd (NIMH, 2022).

Anxiety disorders can be caused due to genetic predisposition, chemical imbalance in the brain, environmental stress, drug abuse and certain medical conditions (Cleveland Clinic, 2020) (WebMD, 2022). Treatment of anxiety disorders involves psychotherapy and/or medication. Psychotherapy consists of cognitive behavior therapy, which is a ‘talk therapy’ to change someone its thoughts, reactions and behavior to reduce anxiousness or can be exposure therapy where someone is exposed to situations one would have previously avoided. Medication includes specific anti-anxiety medication, antidepressants and beta-blockers and is only used to combat symptoms rather than to cure the anxiety (APA, 2021).

Developments in The Netherlands

In the Netherlands, the prevalence of anxiety disorders between December 2021 and December 2022 was observed to be 15.6%, indicating that more than 1 out of 7 individuals were affected by this condition. (Trimbos_1, 2022). Interestingly, despite the COVID-19 pandemic, there was no significant increase in the occurrence of anxiety disorders in the country. However, there was a temporary rise in the number of individuals reporting feelings of gloominess and anxiety, which did not correspond to an increase in clinical diagnoses of psychological problems during the pandemic. These findings were obtained through a combination of physical and online interviews conducted as part of fieldwork, where Dutch citizens aged between 18 and 75 years were randomly selected from each municipality (Trimbos$_3$, 2023).

Furthermore, it is worth noting that a continuous growth in the number of diagnosed psychological disorders had been observed prior to the onset of the COVID-19 pandemic (Trimbos$_2$, 2022). According to the lead epidemiologist at the Trimbos Institute, potential factors contributing to the initiation of this increasing trend in mental disorders in the Netherlands may include growing performance pressure, a more individual-focused society, and income disparities (Trimbos$_1$, 2022). However, further research is necessary to investigate the relationship between these factors and the rising trend of mental disorders in The Netherlands.

Experts from healthcare and medical science state that the current way in which Dutch healthcare is organized is no longer feasible. According to the Scientific Council for Government Policy (WRR), 1 in 6 Dutch employees work in healthcare (WRR, 2021). If we continue as we are doing now, in the year 2040, 1 in 4 employees would be working in healthcare and between 2050 and 2060 1 in 3 employees would be working in healthcare. This indicates that the healthcare sector is facing significant challenges. Care professionals are under pressure due to the high workload and insufficient cooperation between domains. Access to care and the quality of care are also under pressure. Moreover, healthcare has a task to become sustainable, but simply investing more money is not the solution (IZA, 2022).

Considering the aforementioned developments, it is necessary to invest in prevention, early detection, and a more comprehensive understanding of the required care and its recipients. This requires a shift in focus from hospital-based treatment and care towards promoting health and well-being in people's own environment. (Van de Dungen et al., 2018).

Van de Dungen et al. (2018) propose a transformative approach to mental healthcare, emphasizing the importance of considering individuals' functioning within their environment rather than solely focusing on their limitations. This paradigm shift involves investing in health and well-being through district-oriented prevention and support, aiming to optimize individuals' overall health and vitality while preventing or reducing care needs. Such an approach entails promoting a healthy lifestyle, promoting mental health skills in favorable living environments, and empowering individuals through social support and self-reliance. Notably, support from the social and public sectors, as well as addressing non-care-related issues, holds great value in this framework. Municipalities play a critical role in providing the necessary support (IZA, 2022), while health professionals, recognizing the importance of both medical and social aspects, can gain a comprehensive understanding of individuals' physical, psychological, and social needs.

Problem statement of VWS and VNG

According to KPMG and Phrenos, there is a lack of overview in how mental care is being delivered at a neighborhood level, which kind of mental care is being financed from what act and what the possible future developments of mental care will be (KPMG & Phrenos, p.8, 2022). New acts which have been implemented in the last years, are the following:

- ambulatory care in the curative mental health care (ZVW-WMO);
- entry into force of the long-term care act (WLZ);
- implementation of the Dannenberg vision;
- entry into force of the compulsory care act (WVGGZ) and the care and compulsion act (WZD)
- entry into force of the outline agreement on mental health care (GGZ) in 2019;
- developments in the forensic domain;
- access to the WLZ for people with mental health problems as of 2021

The Ministry of Health, Welfare and Sport (VWS) and the association of Dutch municipalities (VNG) would like more insight into relevant developments regarding the provision of support and care to people with mental health problems both nationally and regionally. They want to gain insight into what care and support is provided and what part of this care and support is financed by which law. VWS and VNG want to be able to follow relevant developments over time and also need policy-neutral interpretation of these developments so that a meaningful discussion, both policy-related and financial, about these developments becomes possible. In a letter to the State Secretary for Health, Welfare and Sport, the chairman of the Trimbos institute emphasizes the importance of taking into account the effects of governmental policy on the mental health within The Netherlands (Van der Hoek, 2020). He proposes a reform of the Dutch Mental Health Care (GGZ) by working more on an evidence-informed policy and evidence-based practice. The GGZ should be considered in a broader social-societal context since a lot of demand for care has an origin within this context. Besides, a lot can be learned from the differences between local regions. Van der Hoek emphasizes the importance of having a short cyclic signaling, response development and effect evaluation by monitoring the developments in the mental care. This should not only be limited to the mental care of the GGZ, but should also include factors from the social domain.

According to IZA, a monitor must both guide the agenda of the administrative consultations and be able to justify to society the improvements in the short and medium term. On the one hand, it provides a tangible understanding what the implementation of agreements means for the provision of care to patients; on the other hand, it provides a focus on the realization of the medium to long term goals for a number of large groups, making the impact of the IZA clear not only to patients but also to health care personnel and finances.

Monitoring data

VWS and the association of Dutch municipalities (VNG) have requested KPMG and Phrenos, a center of expertise for severe mental illness, to create a monitor which visualizes trends and developments of psychological problems within The Netherlands on a local level and national level to create data-driven policy (Dutch Government, 2022).

The most important goals of the monitor are to:

- Gain an improved view of relevant developments in the ambulatory care of mental health, the needs of the patients and clients concerned and the provision of appropriate support and care;
- Provide insight into the regional context and social factors affecting the use and supply of care and support in the region;
- Provide high quality and continuous information that contributes to appropriate policy making at regional and national level in which people with mental vulnerability and quality of care and support are central.

The final product will be a publicly available webpage which everyone can use to gain more insight in the status of mental health problems within The Netherlands, currently measured from 2015 up to and including 2020. The main target audience are policymakers which want to form data-driven policies. To preserve the privacy of individuals, all data has been anonymized and aggregated.

The data sources used for the monitor are:

- Statistics Institute of The Netherlands (CBS) (CBS_1, 2022) (CBS_2, 2022), (CBS_3, 2022) (CBS_4, 2022)
- Medical business intelligence centre 'Vektis' (Confidential data*)
- National Institute for Public Health and the Environment (RIVM)
- A monitor of CBS, RIVM and the joint health service (GGD) called the 'Gezondheidsmonitor' ($Overheid_1$, 2022)
- Municipal Social Domain Monitor (GMSD)
- CBS national police data ($Overheid_2$, 2022)
- National Credit Register (BKR)
- Data platform of the VNG called 'Waarstaatjegemeente'

**The Vektis dataset contains information about the amount of people with a diagnosis for each municipality, the amount of treatment time spent for each diagnosis and its corresponding costs per year over the period of 2015 up to and including 2020. The diagnoses in the dataset are: Attention deficit and behavior, alcohol-related disorder, anxiety disorder, basic GGZ need , bipolar, delirium & dementia (grouped), depressive disorder, eating disorder, other substance-related disorder, childhood disorders, personality disorder, PDD-NOS, residual group diagnosis, schizophrenia and somatoform disorders.*

Predicting anxiety disorders

Next to the monitor which focusses on the visualization of retrospective data, there is a high potential of using the available data (from the sources mentioned above) to make predictions about the development of psychological problems, and specifically the prevalence of anxiety disorders per municipality per year. A prediction model could be used to focus more on prevention in line with the goals of VWS and GGZ while including the social-societal context of the Dutch population and to help on a national and regional level to build data-driven policies. When it becomes more clear within what range each municipality can expect there to be psychological problems, then the policy can be adapted to this estimation.

Various machine learning techniques have been employed for predicting and diagnosing different forms of anxiety disorders, but are mostly based on individual patient data. For instance, Chatterjee et al. (2014) used heart rate variability and a probabilistic machine learning approach to classify individuals with generalized anxiety disorder. Meanwhile, Hilbert et al. (2017) utilized binary support vector machines to analyze multimodal biobehavioral data. Studies have also explored the potential of machine learning for the diagnosis and discrimination of social anxiety disorder using neuroimaging data, such as the use of support vector machines to predict social anxiety disorder (Zhang et al., 2015). Lastly, Sundermann et al. (2017) employed a combined functional magnetic resonance imaging (fMRI) and machine learning approach to distinguish between depressive comorbidity and panic disorder.

Other means of collecting data of a bigger and more varied group is via surveys. Bakkeli et al. (2022) investigates the risk and protective factors for depression during the COVID-19 pandemic in Norway, specifically using machine learning models to predict depression risk and to select models with the best performance for each pandemic phase. The study analyzes survey data collected from 5001 Norwegians and found that decision tree models and regularized regressions had the best performance. Highly ranked predictors of depression that remained stable over time were self-perceived exposure risks, income, compliance with nonpharmaceutical interventions, frequency of being outdoors, contact with family and friends, and work-life conflict. The study concludes that machine learning models consisting of demographic, socioeconomic, behavioral, and epidemiological features can be used for fast "first-hand" screening to diagnose mental health problems and may be helpful for stakeholders and healthcare providers to provide early diagnosis and intervention.

Individual patient data may be more detailed than public data, but it could also lead to bias towards certain demographics or populations (Pannucci & Wilkins, 2010). This could lead to inaccurate predictions for other groups which may be under represented in scientific research. Secondly, patient data is often subject to privacy concerns and regulations (Basil et al., 2022), which may limit the amount and type of data that can be collected and used for analysis. Thirdly, patient data is typically collected after the onset of symptoms, which may limit the ability to make early predictions or detect early warning signs. In contrast, public data may include information about risk factors or early indicators that could help identify individuals or groups at risk for developing mental disorders. Finally, patient data may be difficult to access and may require special permissions or collaborations with healthcare providers or institutions, which can be time-consuming and costly. Public data, on the other hand, is often freely available and can be accessed more easily, allowing for larger sample sizes and more diverse data sets.

Therefore, the use of annually updated publicly available data could add to the prediction and prevention of anxiety disorders on a national and regional level. Therefore, we propose the following main research question:

How can we predict anxiety disorders in the Netherlands between 2021-2026 and identify the socio-economic factors from public data that should be targeted by VWS and municipalities to reduce the prevalence of anxiety disorders at the national and regional level?

We set the years 2021 - 2026 as the scope of prediction to align with the timeline of the IZA and account for the absence of data on the number of anxiety disorders for 2021 and 2022.

We formulate the secondary research questions:

1. What are risk factors and protective factors related to the prevalence of anxiety disorders based on public data between 2015-2020?
2. How can we build an accurate prediction model to predict the prevalence of anxiety disorders between 2021-2026?
3. What factors should VWS and municipalities prioritize for their interventions to decrease the number of anxiety disorders?

2 Methods

The design of the study is twofold. First, we perform linear regression and a neural network to determine which factors are most suitable for predict the prevalence of anxiety disorders. Second, we create a prediction model to determine to what extend we can predict the amount of patients with an anxiety disorder for 2021-2026.

The goal of this study is to:

1. To identify and examine the risk factors and protective factors associated with the prevalence of anxiety disorders, utilizing public data from 2015-2020.
2. To develop a precise prediction model that can accurately predict the prevalence of anxiety disorders for 2021-2026*.
3. To determine the key factors that should be targeted for interventions by VWS and VNG.

**The scope of prediction is set at the years from 2021-2026 since we miss data about the number of anxiety disorders of the years 2021, 2022 and we want to predict the prevalence of anxiety disorders up to and including 2026 in accordance with the scope of IZA (IZA, 2022).*

2.1 Data gathering

We perform a quantitative retrospective study of the proportion of inhabitants with anxiety disorders registered from the 1st of January 2015 to the 31st of December 2020 per Dutch municipality using data from Vektis Medical Business Intelligence Centre. For 'proportion of inhabitants with anxiety disorders', we calculate the proportion of anxiety disorder patients in each municipality relative to the number of inhabitants, in order to account for variations in population size across municipalities. Datasets of the years 2021 and 2022 are not yet completed by Vektis since they are waiting to receive the data from health insurance companies.

We retrieve data from the 1st of January 2015 to the 31st of December 2022 from the public website of the Statistics Netherlands (CBS) and the National Dutch Police for the same Dutch municipalities as the Vektis dataset holds (see Figure 2 for the data flowchart, see Table 11 and Table 12 for further details about the factors from CBS and the Police). We use social-economic and environmental factors from CBS and the Police since these datasets give an extensive overview of each Dutch municipality and may include information about risk factors or early indicators of anxiety disorders. These datasets are freely available and can be accessed easily, allowing for large sample sizes. Furthermore the CBS and Police data are publicly available and yearly updated which contribute to reproducibility of this research for the coming years. For an overview on the availability per database for each year, see Figure 1.

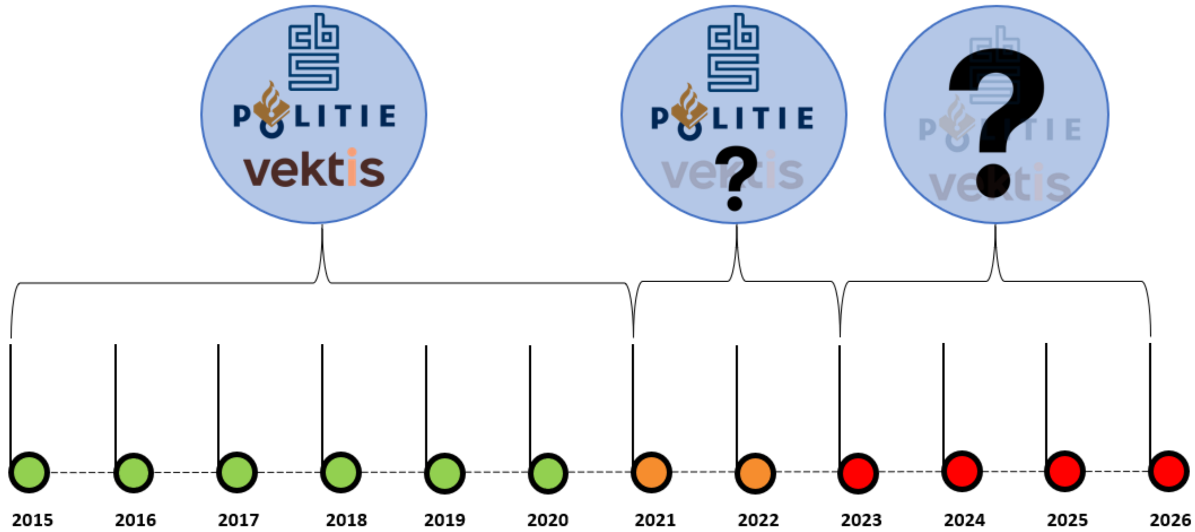

Figure 1 gives an overview of the general database availability for each year between the period of 2015-2026. The green, orange and red dots mean that all databases are available, only the Vektis database is missing and all databases are missing respectively. Vektis has data of the number of anxiety disorders, CBS has data about social-economic and environmental factors, and the Police has data about the number of nuisance reports. All datasets are registered per Dutch municipality per year.

2.1.1 Vektis

Vektis is an information services organization that offers reports on the Dutch healthcare system (Vektis, 2023). Its services include collecting, organizing, and analyzing data on the expenses and usage of medical treatment, devices, and drugs. The aim of Vektis is to enhance the care experience for patients, providers, insurers, and policymakers through data-driven insights.

The data of Vektis contains health figures from the 1st of January 2015 to the 31st of December 2020 about:

- the number of patients with an anxiety disorder per Dutch municipality per year
- the total costs made for anxiety disorders in euro’s per Dutch municipality per year
- the total amount of minutes spent on treatment of anxiety disorders per Dutch municipality per year

We have a dataset with 1,881 rows which we have extracted from the initial dataset received from Vektis by only including the diagnosis ‘anxiety disorder’. The 1,881 rows consist of 325 municipalities with a maximum of 6 years (2015-2020) measured for each municipality. 38 of the municipalities contain 5 years of data (11.6% of the total municipalities), 1 municipality contains 4 years of data, 3 municipalities contain 3 years of data and 5 municipalities contain 2 years of data. Some municipalities exist shorter than the 6 years due to merges and splits in these regions. The dataset consists of 5 columns; year, municipality name, total number of patients, total costs and total treatment minutes. We load the data into R.

There are different types of anxiety disorders, as mentioned in the introduction (see Introduction). The type of anxiety disorder is determined by a psychologist or psychiatrist and included in the Vektis dataset (FMS, 2013). The types of anxiety disorders which are included in the Vektis dataset are:

- Agoraphobia
- Generalized anxiety disorder
- Panic disorder
- Social anxiety disorder
- Specific phobias

The anxiety disorder types which are out of scope of our research are anxiety disorders due to a medical condition, selective mutism, separation anxiety disorder, substance-induced anxiety disorder and other specified anxiety disorders and unspecified anxiety disorders (Mayo Clinic, 2018). These types are out of our scope since these have not been registered in the Vektis dataset.

2.1.2 Statistics Netherlands (CBS)

For the data of CBS, we retrieve data from the 1st of January 2015 to the 31st of December 2022 for all Dutch municipalities (CBS_3, 2022). The CBS dataset holds social-economic and environmental factors per year per municipality containing 2,364 rows and 308 columns of data about different categories (see Table 11 and Table 12). We use CBS data for our research since this could give meaningful insights about the relationship between population, social-economics and environmental factors and the amount of patients with an anxiety disorder per municipality.

2.1.3 National Dutch Police

For the police data, we retrieve data from the 1st of January 2015 to the 31st of December 2022 for all Dutch municipalities ($Overheid_2$, 2022). The police dataset holds figures about the amount of nuisance reports per year per municipality containing 2,082 rows and 11 columns (see Table 11 and Table 12).

2.2 Data preparation

We merge the datasets of Vektis, CBS and the Police based on the year and municipality (See Figure 2). We calculate the number of rows containing only missing values and decide to omit those since they only account for 34 of the 1,983 total rows (1.7%). We also omit the columns containing specific data about patients who have stayed at the treatment location and who have not, since we are interested in the whole population rather than these separate groups. Lastly, we omit columns containing specific localization data about municipalities. We do keep the municipality names to identify the municipalities.

For the CBS and Police data of the years 2021 and 2022 we impute the missing data for each column and municipality using a linear regression model. For each column, we drop the rows with missing data and check if there are enough data points to fit a linear regression model. If there are two or more data points, we fit a linear regression model to predict the missing values.

If not, we impute the missing data using the mean value of the available data. We then return the dataset with the imputed missing data.
After imputing the missing data we extrapolate the future values for the years 2023, 2024, 2025, and 2026 for each column and municipality using the linear regression model. We loop over each municipality and year, fit a linear regression model to the available data, and extrapolate the future values using the fitted model.

To account for variations in population size among municipalities, we scale each factor to the population size. This correction ensures that differences arising from varying population sizes do not skew the analysis. (see Table 11 and Table 12 for further details about these factors).

The data from Vektis is confidential so to be able to share the results in this report we anonymize the municipalities by setting an ID created by drawing a random integer number of 5 digits. We want to normalize or standardize the data to equalize the scale of each factor and to further anonymize the data corresponding to each municipality. We perform a Shapiro-Wilk test to test if there is a Gaussian distribution in the dataset (Shapiro and Wilk, 1965). We choose to perform normalization instead of standardization since the dataset does not follow a Gaussian distribution ($\alpha<0.05$).

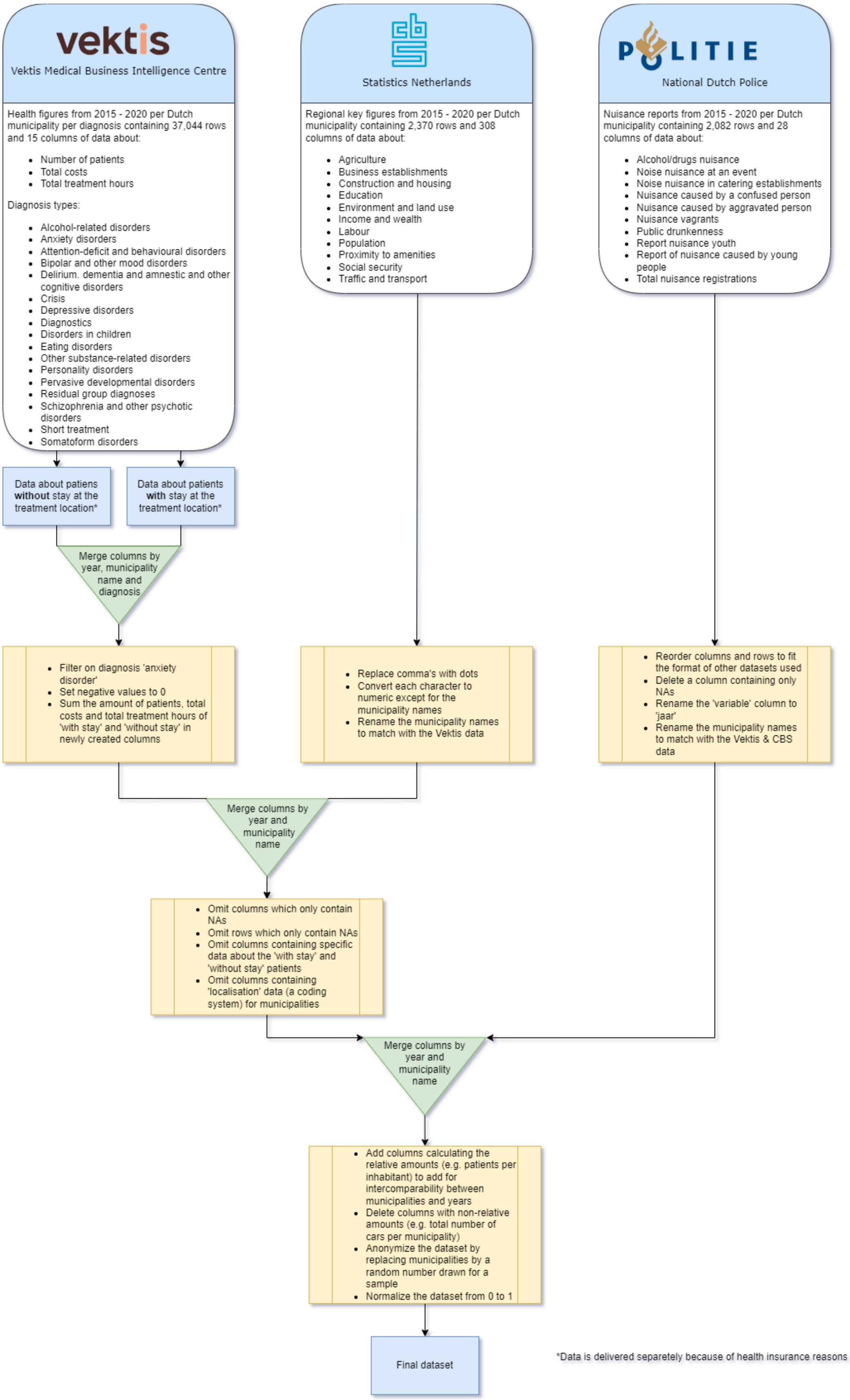

Figure 2 displays a data diagram to visualize how data has been processed in R before performing the data analysis and modelling in Visual Studio Code.

2.3 Data Analysis & Modelling

In the data analysis and modelling phase we want to determine what risk factors and protective factors are related to the prevalence of anxiety disorders based on the dataset used between 2015-2020 and how can we build an accurate prediction model to predict the prevalence of anxiety disorders between 2021-2026. By identifying the risk and protective factors based on public data, we can understand the underlying drivers of the outcome, which can inform targeted and effective interventions. By predicting the future prevalence of anxiety disorders, we gain insights into the potential burden of the condition and the need for intervention.

We start with a correlation heatmap since it is an efficient method to explore the correlation between the input factors and the output factor. By visualizing the correlation heatmap, we can identify which factors have the highest correlation with the target factor to gain a quick overview of the relationships. Next, we use standard linear regression since it is a well-established method that can provide valuable insights into the relationship between the input factors and the output factor. However, we also recognize that linear regression has limitations, such as the assumption of linearity and the potential for overfitting. To address these limitations, we perform lasso regression, which can reduce the number of input factors and prevent overfitting. Finally, we use a neural network, which is a more complex and flexible method that can capture non-linear higher order relationships between the input factors and the output factor. We determine the risk and protective factors for the neural network with a sensitivity analysis. By performing these techniques in this order, we gradually increase the complexity of the analysis and gain deeper insights into the relationship between the input factors and the output factor. We eventually use the neural network to predict the prevalence of anxiety disorders.

2.3.1 Correlation heatmap

A correlation heatmap can be a useful tool in determining the relationships between factors in a dataset. In the case of determining risk and protective factors related to the prevalence of anxiety disorders, a correlation heatmap can help identify which factors are positively or negatively correlated with the proportion of inhabitants with anxiety disorders. By visualizing the correlations between factors, we can quickly identify patterns that may be important in understanding the risk and protective factors for anxiety disorders. The heatmap can also help identify potential multicollinearity among factors, which is important in statistical analyses that aim to identify predictors of anxiety disorder prevalence.

We create a correlation heatmap based on the Pearson correlation of 153 factors (See Table 14) containing normalized data on a scale from 0 to 1. Pearson correlation is a statistical method used to measure the strength and direction of the linear relationship between two factors (Nettleton, 2014). It produces a correlation coefficient that ranges from -1 to 1, with -1 indicating a perfect negative correlation, 0 indicating no correlation, and 1 indicating a perfect positive correlation. The correlation coefficient is calculated by dividing the covariance of the two factors by the product of their standard deviations. A positive correlation coefficient indicates that as the value of one factor increases, the value of the other factor also tends to increase. In contrast, a negative correlation coefficient suggests that as the value of one factor increases, the value of the other factor tends to decrease. The magnitude of the correlation coefficient indicates the strength of the relationship between the factors. We transform the correlation matrix into a heatmap in Excel (See Table 1).

2.3.2 Linear Regression

Linear regression can be a useful tool in identifying potential predictors of anxiety disorder prevalence. By analyzing the relationship between one or more input factors and the output factor of anxiety disorder prevalence, we can determine whether there is a significant linear relationship between the two. Linear regression can also provide information on the direction and strength of the relationship, as well as the statistical significance of the relationship (Schneider et al, 2010).

In the case of identifying risk and protective factors for anxiety disorders, linear regression can help determine which factors are most strongly associated with anxiety disorder prevalence. By expressing the performance of a linear regression model in terms of R^2, we measure how much of the total variability can be explained from the model.

We select the features for our multiple linear regression as follows:

1. Only keep input factors which have a correlation of > 0.4 with the output factor 'the proportion of inhabitants with anxiety disorders which can be interpreted as moderate to strong correlation (Akoglu, 2018) (See Table 2).
2. Select the feature with the highest correlation with the output factor to use in the multiple linear regression formula.
3. Select the next feature which has a correlation of > 0.4 with the input factor but has the lowest correlation with the previously chosen feature(s) to minimize multicollinearity.
4. Repeat step 3 until you cannot find new features which have a correlation < 0.6 with the previously chosen features.

2.3.3 Lasso Regression

In the previous paragraph (See 2.3.2), we performed linear regression based on a limited number of factors selected with the help of a correlation heatmap. By using lasso regression, we can include all factors instead of a limited number. Including all relevant factors in the model can improve the model its ability to accurately predict the outcome factor, which is important to determine the risk factors and protective factors related to the prevalence of anxiety disorders.

Lasso regression is a regularization method that adds a penalty term to the sum of the absolute values of the regression coefficients (Ranstam & Cook, 2018). Lasso regression can effectively reduce the impact of multicollinearity on the regression coefficients and it performs feature selection by shrinking some of the coefficients to exactly zero. This means that Lasso regression can be used to identify which input factors are the most important for predicting the prevalence of anxiety disorders and remove the less important factors from the model.

The input factors are standardized before applying the lasso regression. Standardization ensures that the factors are on the same scale and have the same variance. This is important for Lasso regression because it assumes that all factors are on the same scale and penalizes the absolute size of the coefficients (Hastie et al., 2015). Standardizing the data before applying Lasso regression ensures that the penalty term is applied equally across all factors.

We randomly split the dataset into a training set and a test set, using a proportion of 80% and 20% respectively. We then search over a range of alpha values between 0.00001 and 100 with a step size of 1 to find the best value that minimizes the sum of squared errors between the predicted and actual values. The model is evaluated using R-squared and mean squared error (MSE) for both the training and test data.

2.3.4 Sensitivity Analysis

To determine which factors from public data are risk factors and protective factors related to the prevalence of anxiety disorders we perform a sensitivity analysis on the year 2019 using a neural network (See 2.3.5) to predict the proportion of inhabitants with anxiety disorders. We use the year 2019 since this year has the most recent observations without the influence of the COVID-19 pandemic.

National Sensitivity Analysis

To present a national overview, we perform a sensitivity analysis by altering the percentage for each factor by -5% and +5%. We calculate the gradient for the difference in the effect on the proportion of inhabitants with anxiety disorders. We then sum the gradients of each factor for all the municipalities. Next, we scale these factors for the population size of each municipality and plot the resulting weighted sensitivity scores to identify the factors that have the greatest impact on the output factor. By calculating the gradient for the national sensitivity analysis, we are able to rank the sensitivity of each factor.

Municipality Sensitivity Analysis

For the municipalities, we want to be able to give individual insight in which risk and protective factors there are and what the effect would be if policymakers can alter the value of a risk or protective factor. Therefore, we perform the sensitivity analysis by increasing and decreasing the value of each CBS & Police factor by 5%, 10%, 15%, 20% and 25% while keeping other factors constant. We then input these newly created values into our neural network model (See 2.3.5) to predict the percentual change of the proportion of inhabitants with anxiety disorders per municipality. We can use the most influential factors from CBS & Police to provide a municipality-specific advice on which factors the policymakers should target to decrease the number of anxiety disorders.

2.3.5 Neural Network

To predict the proportion of inhabitants with anxiety disorders over the period of 2021-2026 we build a neural network In the previous paragraphs (see 2.3.2 & 2.3.3), we performed two types of linear regression. Linear regression techniques assume a linear relationship among factors, while a neural network is able to capture complex, non-linear relationships between input factors and the output factor. A neural network is able to capture complex, non-linear relationships between the factors from CBS and Police with the proportion of inhabitants with anxiety disorders.

We construct a neural network model using the TensorFlow library in Visual Studio Code. We create a model where we assign each of the 130 factors of CBS and Police to one of the 17 parallel layers which correspond to the different categories of the CBS and Police data (See Figure 3). By reducing the 130 factors to 17 parallel layers, we decrease the dimensionality of the model, which decreases the computational cost, prevents overfitting and improves generalization (Murphy, 2012).

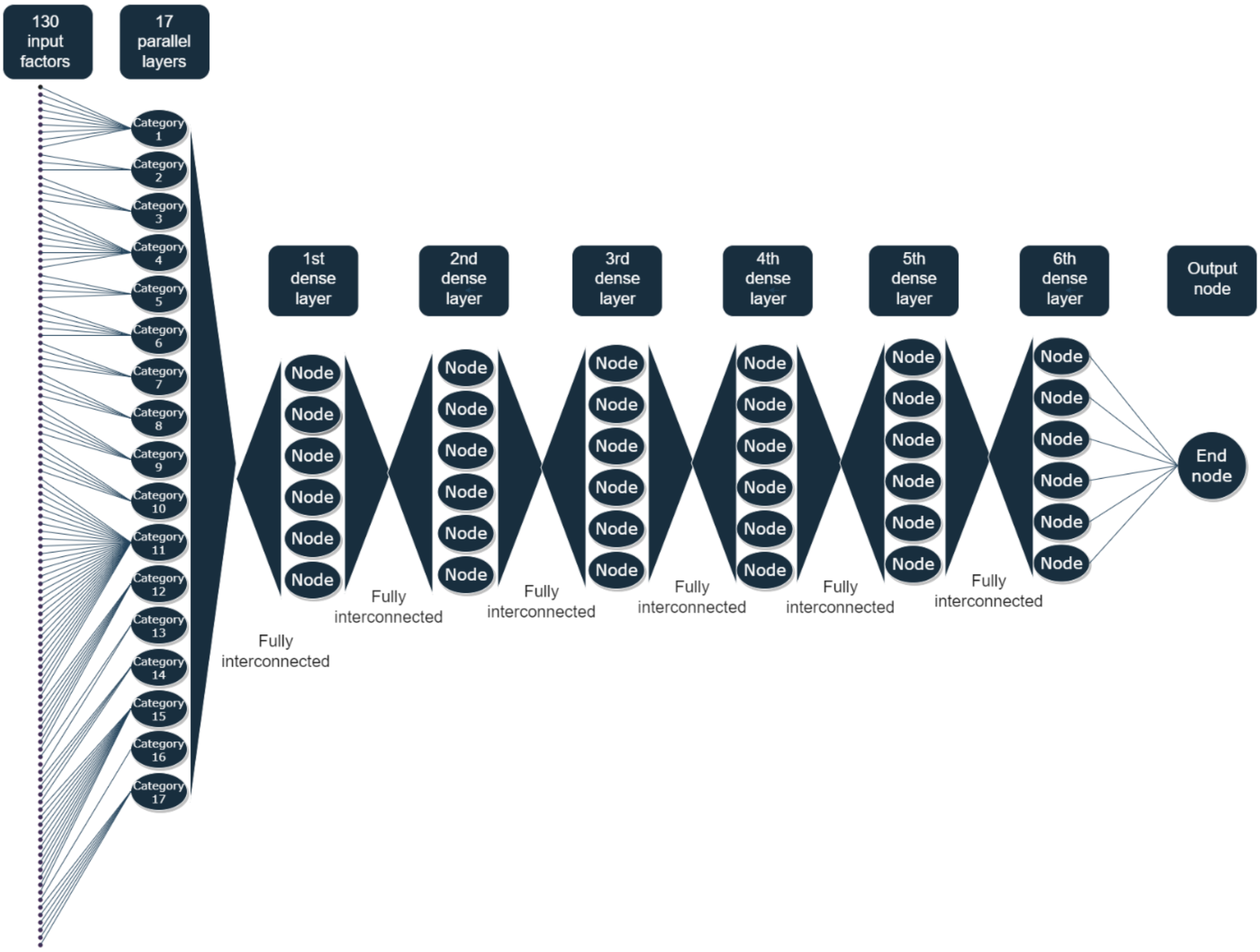

Figure 3 visualizes the final topology of the neural network. 130 factors of CBS and Police are assigned to 17 parallel layers which correspond to the different categories of the CBS and Police data. The ending note consists of the proportion of inhabitants with anxiety disorders.

We assign each of the 130 factors to a category (See Table 11 for more explanation of each factor and its category). In the parallel layers of Figure 3, the following names correspond to the category number given:

1. population age group division (9 factors)
2. population demographic pressure (3 factors)
3. marital status (4 factors)
4. migration background (9 factors)
5. birth- deathrates (4 factors)
6. death cause categories (5 factors)
7. population development (3 factors)
8. household size (4 factors)
9. housing info (5 factors)
10. education and jobs (5 factors)
11. income and assets (30 factors)
12. social security (7 factors)
13. agriculture (9 factors)

14. vehicle ownership (4 factors)
15. distance to facilities (22 factors)
16. environment and land use (1 factor)
17. nuisance Police reports. (6 factors)

After first composing the 17 parallel layers, we concatenate the layers to one dense layer. The concatenated blocks can help to reduce the chance of overfitting by creating a more compact representation of the data that retains the essential information from each block (Murphy, 2012).

We compile the model with the Adam optimizer and an exponential decay learning rate.

Adam (Adaptive Moment Estimation) is an optimizer used for stochastic gradient descent (Kingama & Ba, 2014). It adapts the learning rate for each parameter based on estimates of the first and second moments of the gradients. This can lead to faster convergence and better performance compared to traditional gradient descent algorithms.

Exponential Decay learning rate is a technique used to gradually reduce the learning rate during training (Li & Arora, 2019). This helps the model to converge to a better optimum and improves generalization. It starts with a higher learning rate at the beginning of training and reduces it over time.

We apply L2 regularization to prevent overfitting. L2 regularization, also known as ridge regression, adds a penalty term to the loss function that is proportional to the squared magnitude of the weights in the model (McDonald, 2009). This encourages the model to learn smaller weights, which has the effect of simplifying the model and reducing overfitting. By applying L2 regularization to the concatenated layers instead of every block, the regularization penalty is applied more broadly across the model, rather than targeting specific blocks or layers.

We decide the number of neurons in each layer by experimentation (See Table 5 and Table 6). After setting the desired topology, we perform a run with 10,000 epochs (See Table 7).

We test the model using 5-fold cross-validation, with each fold training and evaluating the model on different subsets of the data. Finally, we evaluate the model on the test data, and we calculate the average validation score (See Table 7).

3 Results

3.1 Linear Regression

We construct a multiple linear regression model with 7 factors based on the threshold of <0.60 collinearity among selected features. The 7 factors selected for the linear regression in unsorted order are:

- Percentage of people who receive social security per municipality*
- Distance from a hospital*
- Households with children*
- Percentage of traffic area*
- Nuisance reports because of a confused person*
- Migration background of the Dutch Antilles and Aruba*
- Nuisance reports because of a homeless person*

**See Table 11 for more details about these factors.*

By selecting these 7 factors as features for our linear regression, we achieve a MSE of 0.027, a multiple R^2 of 0.5941 (See 3.4.1 for interpretation of the results).

Table 1 shows a heatmap correlation matrix of all input factors which have > 0.4 correlation with the proportion of inhabitants with anxiety disorders. For an overview on which numbers belong to which factors, see Table 11.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	1,00	0,66	0,61	0,61	0,60	0,59	0,59	0,59	0,59	0,59	0,57	0,57	0,55	0,55	0,55	0,54	0,54	0,53	0,52	0,51	0,51	0,47	0,48	0,49	0,47	0,46	0,43	-0,40	-0,41	-0,41	-0,41	-0,42	-0,42	-0,47	-0,50	-0,51	-0,54	-0,55	-0,58	-0,66
2	0,66	1,00	0,64	0,56	0,60	0,58	0,76	0,56	0,68	0,68	0,63	0,64	0,58	0,59	0,64	0,56	0,64	0,56	0,58	0,56	0,86	0,48	0,50	0,51	0,44	0,60	0,42	-0,65	-0,47	-0,43	-0,66	-0,51	-0,34	-0,72	-0,71	-0,54	-0,62	-0,64	-0,74	-0,71
3	0,61	0,64	1,00	0,81	0,83	0,85	0,77	0,88	0,68	0,93	0,50	0,85	0,41	0,91	0,80	0,68	0,50	0,92	0,56	0,82	0,43	0,85	0,79	0,87	0,74	0,69	0,64	-0,50	-0,77	-0,71	-0,55	-0,34	-0,50	-0,55	-0,51	-0,77	-0,51	-0,80	-0,75	-0,72
4	0,61	0,56	0,81	1,00	0,74	0,73	0,69	0,77	0,66	0,76	0,51	0,89	0,41	0,73	0,85	0,69	0,43	0,74	0,46	0,64	0,34	0,69	0,64	0,71	0,68	0,73	0,58	-0,37	-0,66	-0,60	-0,40	-0,34	-0,44	-0,44	-0,45	-0,75	-0,52	-0,85	-0,68	-0,73
5	0,60	0,60	0,83	0,74	1,00	0,93	0,64	0,84	0,60	0,83	0,44	0,77	0,42	0,81	0,70	0,62	0,47	0,81	0,55	0,72	0,37	0,73	0,68	0,74	0,67	0,67	0,58	-0,35	-0,66	-0,61	-0,40	-0,30	-0,43	-0,42	-0,43	-0,68	-0,45	-0,70	-0,63	-0,65
6	0,59	0,58	0,85	0,73	0,93	1,00	0,62	0,87	0,62	0,87	0,41	0,75	0,38	0,85	0,71	0,64	0,43	0,83	0,54	0,80	0,37	0,73	0,65	0,75	0,61	0,62	0,54	-0,33	-0,65	-0,60	-0,40	-0,34	-0,43	-0,39	-0,40	-0,67	-0,47	-0,71	-0,62	-0,66
7	0,59	0,76	0,77	0,69	0,64	0,62	1,00	0,62	0,83	0,72	0,61	0,73	0,55	0,66	0,73	0,68	0,64	0,63	0,52	0,63	0,62	0,63	0,61	0,65	0,59	0,59	0,54	-0,71	-0,70	-0,60	-0,67	-0,58	-0,24	-0,74	-0,60	-0,72	-0,70	-0,73	-1,00	-0,79
8	0,59	0,56	0,88	0,77	0,84	0,87	0,62	1,00	0,60	0,90	0,40	0,80	0,40	0,87	0,74	0,67	0,44	0,86	0,55	0,84	0,32	0,75	0,64	0,77	0,63	0,63	0,56	-0,30	-0,65	-0,59	-0,35	-0,30	-0,42	-0,38	-0,40	-0,68	-0,44	-0,74	-0,61	-0,67
9	0,59	0,68	0,68	0,66	0,60	0,62	0,83	0,60	1,00	0,68	0,61	0,60	0,61	0,63	0,69	0,79	0,70	0,54	0,58	0,65	0,51	0,52	0,50	0,56	0,49	0,51	0,39	-0,46	-0,62	-0,51	-0,49	-0,84	-0,11	-0,47	-0,55	-0,69	-0,89	-0,69	-0,84	-0,90
10	0,59	0,68	0,93	0,76	0,83	0,87	0,72	0,90	0,68	1,00	0,46	0,82	0,43	0,93	0,76	0,70	0,51	0,91	0,60	0,88	0,47	0,76	0,69	0,78	0,65	0,63	0,54	-0,44	-0,70	-0,63	-0,49	-0,37	-0,43	-0,50	-0,52	-0,73	-0,51	-0,76	-0,71	-0,73
11	0,57	0,63	0,50	0,51	0,44	0,41	0,61	0,40	0,61	0,46	1,00	0,54	0,54	0,42	0,66	0,29	0,59	0,41	0,42	0,33	0,63	0,40	0,50	0,43	0,49	0,52	0,37	-0,48	-0,23	-0,35	-0,51	-0,56	-0,26	-0,49	-0,64	-0,35	-0,73	-0,66	-0,60	-0,61
12	0,57	0,64	0,85	0,89	0,77	0,75	0,73	0,80	0,60	0,82	0,54	1,00	0,38	0,77	0,89	0,65	0,45	0,81	0,50	0,68	0,44	0,75	0,68	0,76	0,68	0,81	0,65	-0,48	-0,70	-0,63	-0,51	-0,26	-0,44	-0,55	-0,54	-0,74	-0,45	-0,89	-0,71	-0,69
13	0,55	0,58	0,41	0,41	0,42	0,38	0,55	0,40	0,61	0,43	0,54	0,38	1,00	0,36	0,40	0,43	0,81	0,35	0,53	0,36	0,45	0,28	0,27	0,30	0,27	0,34	0,19	-0,36	-0,38	-0,33	-0,35	-0,51	-0,07	-0,39	-0,46	-0,42	-0,55	-0,40	-0,55	-0,56
14	0,55	0,59	0,91	0,73	0,81	0,85	0,66	0,87	0,63	0,93	0,42	0,77	0,36	1,00	0,73	0,65	0,42	0,89	0,49	0,85	0,39	0,77	0,68	0,79	0,66	0,60	0,55	-0,39	-0,68	-0,61	-0,45	-0,33	-0,43	-0,43	-0,44	-0,70	-0,47	-0,73	-0,66	-0,67
15	0,55	0,64	0,80	0,85	0,70	0,71	0,73	0,74	0,69	0,76	0,66	0,89	0,40	0,73	1,00	0,60	0,46	0,70	0,50	0,65	0,48	0,70	0,67	0,72	0,67	0,72	0,59	-0,44	-0,57	-0,62	-0,47	-0,44	-0,43	-0,51	-0,58	-0,65	-0,62	-1,00	-0,72	-0,72
16	0,54	0,56	0,68	0,69	0,62	0,64	0,68	0,67	0,79	0,70	0,29	0,65	0,43	0,65	0,60	1,00	0,54	0,56	0,57	0,71	0,36	0,52	0,43	0,54	0,42	0,50	0,40	-0,33	-0,68	-0,43	-0,33	-0,49	-0,20	-0,39	-0,44	-0,78	-0,59	-0,60	-0,68	-0,92
17	0,54	0,64	0,50	0,43	0,47	0,43	0,64	0,44	0,70	0,51	0,59	0,45	0,81	0,42	0,46	0,54	1,00	0,44	0,66	0,41	0,51	0,35	0,39	0,37	0,32	0,42	0,26	-0,43	-0,51	-0,43	-0,45	-0,54	-0,04	-0,45	-0,56	-0,55	-0,61	-0,46	-0,64	-0,66
18	0,53	0,56	0,92	0,74	0,81	0,83	0,63	0,86	0,54	0,91	0,41	0,81	0,35	0,89	0,70	0,56	0,44	1,00	0,49	0,75	0,37	0,77	0,70	0,77	0,66	0,63	0,53	-0,43	-0,74	-0,67	-0,49	-0,18	-0,47	-0,47	-0,47	-0,74	-0,34	-0,70	-0,62	-0,58
19	0,52	0,58	0,56	0,46	0,55	0,54	0,52	0,55	0,58	0,60	0,42	0,50	0,53	0,49	0,50	0,57	0,66	0,49	1,00	0,52	0,45	0,37	0,32	0,39	0,27	0,42	0,27	-0,33	-0,44	-0,40	-0,35	-0,41	-0,29	-0,39	-0,45	-0,51	-0,47	-0,50	-0,52	-0,62
20	0,51	0,56	0,82	0,64	0,72	0,80	0,63	0,84	0,65	0,88	0,33	0,68	0,36	0,85	0,65	0,71	0,41	0,75	0,52	1,00	0,35	0,65	0,51	0,68	0,50	0,46	0,44	-0,31	-0,62	-0,50	-0,35	-0,42	-0,29	-0,35	-0,36	-0,63	-0,50	-0,65	-0,63	-0,70
21	0,51	0,86	0,43	0,34	0,37	0,37	0,62	0,32	0,51	0,47	0,63	0,44	0,45	0,39	0,48	0,36	0,51	0,37	0,45	0,35	1,00	0,30	0,38	0,33	0,30	0,40	0,28	-0,69	-0,25	-0,24	-0,70	-0,46	-0,30	-0,73	-0,73	-0,31	-0,56	-0,48	-0,62	-0,55
22	0,49	0,51	0,87	0,71	0,74	0,75	0,65	0,77	0,56	0,78	0,43	0,76	0,30	0,79	0,72	0,54	0,37	0,77	0,39	0,68	0,33	1,00	0,92	1,00	0,85	0,64	0,66	-0,41	-0,66	-0,63	-0,49	-0,27	-0,41	-0,47	-0,39	-0,65	-0,42	-0,72	-0,64	-0,59
23	0,48	0,50	0,79	0,64	0,68	0,65	0,61	0,64	0,50	0,69	0,50	0,68	0,27	0,68	0,67	0,43	0,39	0,70	0,32	0,51	0,38	0,93	1,00	0,92	0,82	0,59	0,70	-0,48	-0,58	-0,65	-0,58	-0,24	-0,47	-0,50	-0,42	-0,60	-0,40	-0,67	-0,60	-0,52
24	0,47	0,48	0,85	0,69	0,73	0,73	0,63	0,75	0,52	0,76	0,40	0,75	0,28	0,77	0,70	0,52	0,35	0,77	0,37	0,65	0,30	1,00	0,93	1,00	0,85	0,62	0,66	-0,40	-0,67	-0,63	-0,48	-0,22	-0,41	-0,46	-0,38	-0,65	-0,38	-0,70	-0,62	-0,56
25	0,47	0,44	0,74	0,68	0,67	0,61	0,59	0,63	0,49	0,65	0,49	0,68	0,27	0,66	0,67	0,42	0,32	0,66	0,27	0,50	0,30	0,85	0,82	0,85	1,00	0,63	0,64	-0,43	-0,55	-0,54	-0,53	-0,26	-0,37	-0,45	-0,36	-0,55	-0,42	-0,67	-0,58	-0,51
26	0,46	0,60	0,69	0,73	0,67	0,62	0,59	0,63	0,51	0,63	0,52	0,81	0,34	0,60	0,72	0,50	0,42	0,63	0,42	0,46	0,40	0,62	0,59	0,64	0,63	1,00	0,53	-0,42	-0,55	-0,52	-0,49	-0,26	-0,33	-0,47	-0,48	-0,58	-0,43	-0,72	-0,58	-0,57
27	0,43	0,42	0,64	0,58	0,58	0,54	0,54	0,56	0,39	0,54	0,37	0,65	0,19	0,55	0,59	0,40	0,26	0,53	0,27	0,44	0,28	0,66	0,70	0,66	0,64	0,53	1,00	-0,47	-0,48	-0,52	-0,50	-0,15	-0,45	-0,49	-0,31	-0,53	-0,30	-0,59	-0,53	-0,44
28	-0,40	-0,65	-0,50	-0,37	-0,35	-0,33	-0,71	-0,30	-0,46	-0,44	-0,48	-0,48	-0,36	-0,39	-0,44	-0,33	-0,43	-0,43	-0,33	-0,31	-0,69	-0,40	-0,48	-0,41	-0,43	-0,42	-0,47	1,00	0,50	0,36	0,92	0,29	0,17	0,96	0,72	0,45	0,40	0,44	0,70	0,43
29	-0,41	-0,47	-0,77	-0,66	-0,66	-0,65	-0,70	-0,65	-0,62	-0,70	-0,23	-0,70	-0,38	-0,68	-0,57	-0,68	-0,51	-0,74	-0,44	-0,62	-0,25	-0,67	-0,58	-0,66	-0,55	-0,55	-0,48	0,50	1,00	0,71	0,53	0,18	0,19	0,51	0,41	0,88	0,30	0,57	0,71	0,60
30	-0,41	-0,43	-0,71	-0,60	-0,61	-0,60	-0,60	-0,59	-0,51	-0,63	-0,35	-0,63	-0,33	-0,61	-0,62	-0,43	-0,43	-0,67	-0,40	-0,50	-0,24	-0,63	-0,65	-0,63	-0,54	-0,52	-0,52	0,36	0,71	1,00	0,39	0,17	0,38	0,39	0,38	0,69	0,30	0,62	0,60	0,46
31	-0,41	-0,66	-0,55	-0,40	-0,40	-0,40	-0,67	-0,35	-0,49	-0,49	-0,51	-0,51	-0,35	-0,45	-0,47	-0,33	-0,45	-0,49	-0,35	-0,35	-0,70	-0,48	-0,58	-0,49	-0,53	-0,49	-0,50	0,92	0,53	0,39	1,00	0,31	0,26	0,87	0,71	0,46	0,43	0,47	0,67	0,44
32	-0,42	-0,51	-0,34	-0,34	-0,30	-0,34	-0,58	-0,30	-0,84	-0,37	-0,56	-0,26	-0,51	-0,33	-0,44	-0,49	-0,54	-0,18	-0,41	-0,42	-0,46	-0,22	-0,24	-0,27	-0,26	-0,26	-0,15	0,29	0,18	0,17	0,31	1,00	-0,06	0,27	0,39	0,18	0,95	0,44	0,60	0,68
33	-0,42	-0,34	-0,50	-0,44	-0,43	-0,43	-0,24	-0,42	-0,11	-0,43	-0,26	-0,44	-0,07	-0,43	-0,43	-0,20	-0,04	-0,47	-0,29	-0,29	-0,30	-0,41	-0,47	-0,41	-0,37	-0,33	-0,45	0,17	0,19	0,38	0,26	-0,06	1,00	0,26	0,18	0,27	0,09	0,43	0,22	0,24
34	-0,47	-0,72	-0,55	-0,44	-0,42	-0,39	-0,74	-0,38	-0,47	-0,50	-0,49	-0,55	-0,39	-0,43	-0,51	-0,39	-0,45	-0,47	-0,39	-0,35	-0,73	-0,46	-0,50	-0,47	-0,45	-0,47	-0,49	0,96	0,51	0,39	0,87	0,27	0,26	1,00	0,72	0,49	0,40	0,51	0,73	0,48
35	-0,50	-0,71	-0,51	-0,45	-0,43	-0,40	-0,60	-0,40	-0,55	-0,52	-0,64	-0,54	-0,46	-0,44	-0,58	-0,44	-0,56	-0,47	-0,45	-0,36	-0,73	-0,38	-0,42	-0,39	-0,36	-0,48	-0,31	0,72	0,41	0,38	0,71	0,39	0,18	0,72	1,00	0,47	0,53	0,58	0,59	0,58
36	-0,51	-0,54	-0,77	-0,75	-0,68	-0,67	-0,72	-0,68	-0,69	-0,73	-0,35	-0,74	-0,42	-0,70	-0,65	-0,78	-0,55	-0,74	-0,51	-0,63	-0,31	-0,65	-0,60	-0,65	-0,55	-0,58	-0,53	0,45	0,88	0,69	0,46	0,18	0,27	0,49	0,47	1,00	0,34	0,65	0,72	0,73
37	-0,54	-0,62	-0,51	-0,52	-0,45	-0,47	-0,70	-0,44	-0,89	-0,51	-0,73	-0,45	-0,55	-0,47	-0,62	-0,59	-0,61	-0,34	-0,47	-0,50	-0,56	-0,38	-0,40	-0,42	-0,42	-0,43	-0,30	0,40	0,30	0,30	0,43	0,95	0,09	0,40	0,53	0,34	1,00	0,62	0,71	0,81
38	-0,55	-0,64	-0,80	-0,85	-0,70	-0,71	-0,73	-0,74	-0,69	-0,76	-0,66	-0,89	-0,40	-0,73	-1,00	-0,60	-0,46	-0,70	-0,50	-0,65	-0,48	-0,70	-0,67	-0,72	-0,67	-0,72	-0,59	0,44	0,57	0,62	0,47	0,44	0,43	0,51	0,58	0,65	0,62	1,00	0,72	0,72
39	-0,58	-0,74	-0,75	-0,68	-0,63	-0,62	-1,00	-0,61	-0,84	-0,71	-0,60	-0,71	-0,55	-0,66	-0,72	-0,68	-0,64	-0,62	-0,52	-0,63	-0,62	-0,62	-0,60	-0,64	-0,58	-0,58	-0,53	0,70	0,71	0,60	0,67	0,60	0,22	0,73	0,59	0,72	0,71	0,72	1,00	0,78
40	-0,66	-0,71	-0,72	-0,73	-0,65	-0,66	-0,79	-0,67	-0,90	-0,73	-0,61	-0,69	-0,56	-0,67	-0,72	-0,92	-0,66	-0,58	-0,62	-0,70	-0,55	-0,56	-0,52	-0,59	-0,51	-0,57	-0,44	0,43	0,60	0,46	0,44	0,68	0,24	0,48	0,58	0,73	0,81	0,72	0,78	1,00

We can use a heatmap correlation matrix to visualize the correlations between factors that have a correlation coefficient greater than 0.4 with the output factor, which is the proportion of inhabitants with anxiety disorders. This technique involves plotting a matrix where each cell represents the correlation coefficient between a pair of factors, with the color of the cell indicating the strength and direction of the correlation (See 2.3.1 for more explanation of the Pearson correlation). The green and red cells indicate a negative and positive relationship, respectively, between the input factor and the output factor.

Table 2 shows a list of the top 5 factors with most positive and most negative Pearson correlation scores. The correlation is determined for each input factor with the output factor: proportion of inhabitants with anxiety disorders. The numbers correspond to the numbers listed in Table 1.

No. in heatmap	Factor	Correlation
1	the proportion of inhabitants with anxiety disorders	1.00
	Top 5 most positive correlators	
2	Welfare per inhabitant	0.66
3	Address density	0.61
4	Non-Western migration background	0.61
5	Number of VMBO schools within 5 km	0.60
6	Distance to secondary school	0.59
	Top 5 most negative correlators	
40	Married population	-0.66
39	Houses to buy	-0.58
38	Both parents born in The Netherlands	-0.55
37	Average household size	-0.54
36	Households without children	-0.51

A Pearson correlation measures the strength and direction of the linear association between the factors. The correlation scores range from -1 to 1, where -1 indicates a perfect negative correlation, 0 indicates no correlation, and 1 indicates a perfect positive correlation. A correlation score of 0.66, for example, indicates a moderately strong positive correlation between the factor "Welfare per inhabitant" and the proportion of inhabitants with anxiety disorders. So if the proportion of inhabitants who receive welfare is high in a municipality, the proportion of inhabitants with anxiety disorders also tends to be high, assuming a linear relationship. On the other hand, a correlation score of -0.66 suggests a moderately strong negative correlation between the factor "Married population" and the proportion of inhabitants with anxiety disorders. So if the proportion of inhabitants who are married is high in a municipality, the proportion of inhabitants with anxiety disorders tends to be low, assuming a linear relationship.

Therefore, factors with higher positive correlation scores are more strongly associated with a higher proportion of inhabitants with anxiety disorders, while factors with higher negative correlation scores are more strongly associated with a lower proportion of inhabitants with anxiety disorders. We define those factors as risk factors, and protective factors respectively.

The Pearson correlation coefficient quantifies the linear relationship between factors, although it may not be applicable to every relationship. Even if there’s a linear relationship between two factors, it is important to note that the actual relationship between the input and output factors may not be perfectly linear. Furthermore, it is important to consider that other factors, not included in this study, may also influence the relationship between the included factors.

3.2 Lasso Regression

Table 3 displays the top 5 most positive and most negative coefficients based on Lasso regression which describe the strength and direction between each input factor and the output factor: the proportion of inhabitants with anxiety disorders. (For a complete overview of the results, see Table 15, Table 16 and Table 17)

Factor	Coefficient
Top 5 most positive coefficients	
Net income: Couple with no child	0.75
Median wealth: Employee income	0.68
Population density	0.60
Median wealth: own home	0.49
Standardized income: Self-employment income	0.33
Top 5 most negative coefficients	
Median wealth: Private households	-1.06
Average household size	-0.78
Housing density	-0.65
Standardized income: Private households	-0.58
Distance to daycare center	-0.39

We use the Lasso regression to obtain standardized coefficients, which represent the strength and direction of the relationship between each input factor and the output factor, the proportion of inhabitants with anxiety disorders.

A positive coefficient indicates that an increase in the value of the input factor is associated with an increase in the proportion of inhabitants with anxiety disorders, while a negative coefficient suggests that an increase in the value of the input factor is associated with a decrease in the proportion of inhabitants with anxiety disorders.

For instance, the highest positive coefficient of 0.75 is observed for the input factor " Standardized income: Couple with no child ". This coefficient suggests that a unit increase in the median disposable income of households consisting of couples with no children is associated with a 0.75 unit increase in the proportion of inhabitants with anxiety disorders. In contrast, the highest negative coefficient of -1.06 is observed for the factor " Median wealth: Private households ", indicating that an increase in the median household wealth of households by a unit of 1 is associated with a decrease in the proportion of inhabitants with anxiety disorders by a unit of -1.06.

We should note that the standardized coefficients obtained from the Lasso regression are comparable across factors, regardless of their scales, and can be used to determine the most important predictors. However, the Lasso regression assumes a linear relationship between the factors, which may not hold for every relationship. Even if there's a linear relationship between two factors, it is important to note that the actual relationship between the input and output factors may not be perfectly linear, and other factors may also influence the relationship which are out of scope in this study.

3.3 Sensitivity Analysis

We first focus on the sensitivity analysis on the national level of The Netherlands, and then elaborate on the sensitivity analysis we are able to perform for each individual Dutch municipality.

3.3.1 National Sensitivity Analysis

We calculate the gradient of each factor for all municipalities in 2019 based on a 5% decrease and a 5% increase, and then sum the gradients, weighted by the population size of each municipality, to assess their national impact on the proportion of inhabitants with anxiety disorders.

Table 4 shows the sum of the weighted gradients per factor based on sensitivity analysis with a neural network. The top 5 positive and negative summed gradients are based on a 5% decrease and a 5% increase weighted by the population size of each municipality of public data in 2019. The table assesses the national impact of these factors on the proportion inhabitants with anxiety disorders.

Factor	Sum of weighted gradients
Top 5 most positive sensitivity	
Divorced population	0.0101
Median wealth: Employee income	0.0066
Standardized income: rental property	0.0057
Median wealth: Couple with no child	0.0051
Number of childcare centers within 3 km	0.0048
Top 5 most negative sensitivity	
Median wealth: Private households	-0.0062
Widowed population	-0.0040
Married population	-0.0037
Median wealth: Couple with child(ren)	-0.0035
Total age pressure	-0.0031

The factors with the most positive sensitivity have the greatest positive effect on the prevalence of anxiety disorders, indicating that an increase in these factors predicts a higher proportion of inhabitants with anxiety disorders on a national scale. While the factors with the most negative sensitivity have the greatest negative effect on the prevalence of anxiety disorders, indicating that an increase in these factors predicts a lower proportion of people with anxiety disorders on a national scale. For a full overview of the sum of weighted gradients for each factor, see Table 18.

The sum of weighted gradients provides an indication of the impact of each factor on the proportion of inhabitants with anxiety disorders. However, it is important to note that the values of the sum of weighted gradients do not translate to the predicted change in the proportion of inhabitants with anxiety disorders, but it does give a ranking of effect of each factor. Nevertheless, the results suggest that the factors with the highest positive sensitivity correspond to the risk factors for anxiety disorders, while those with the highest negative sensitivity correspond to the protective factors. These conclusions are based on a sensitivity analysis, where we use a neural network to predict the output factor (refer to section 3.3.4).

The factor ‘divorced population’ scores the most positive sensitivity. This factor is defined as the percentage of inhabitants older than 15 years who have been divorced. A divorce is defined as the civil state arising after dissolution of a marriage or a registered partnership other than by the death of the partner. One reason could be that individuals who have experienced divorce may experience significant stressors that contribute to the development of anxiety disorders. These stressors can include the emotional toll of ending a relationship, financial difficulties, and changes in living arrangements. According to Martin et al. (2022), anxiety disorders are more common among divorced individuals which is in accordance with our finding that a higher percentage of a divorced population in a municipality is associated with a higher proportion of inhabitants with anxiety disorders.

The factor ‘median wealth: private households’ scores the most negative sensitivity. Wealth is the balance of assets and liabilities, while private households are defined as one or more persons who jointly occupy living quarters and provide for themselves the daily necessities of life. Student households are excluded from this definition by CBS. A possible explanation could be that higher wealth is associated with greater stability and security, which could help to alleviate some of the anxiety that comes with financial insecurity (McLaughlin et al., 2012). We do notice that the median wealth of couples without children, median wealth based on employee income and standardized income of people living in rental property are positively associated with the proportion of inhabitants with anxiety disorders. A reason for this may be that factors about wealth and income may be more influenced by the social groups to which individuals belong rather than the level of wealth and income. For example wealth for private households, couples with children and wealth based on self-employed income has a negative sensitivity, while wealth based on employee income and wealth of couples without children has a positivity sensitivity with the proportion of inhabitants with anxiety disorders (See Table 18).

The percentage of inhabitants who are married has a negative sensitivity with anxiety disorder prevalence, meaning that a higher percentage of marriage is associated with lower proportion of inhabitants with anxiety disorders. This is in line with the results of three studies assessing the marital status of individuals with social anxiety disorder which revealed that 43.8% of them were never married in studies conducted by Antony et al. (1998), Davidson et al. (1993), and Wittchen et al. (2000). Another set of three studies, which evaluated the same variable but reported the category "not married," showed that 56.8% were not married or partnered in studies conducted by Dahl and Dahl (2010), Hoge et al. (2008), and Sparrevohn and Rapee (2009). In contrast, the normative data from the US population, using similar definitions of marital/partnered status, indicated that only 30.0% of women and 30.9% of men aged 25-44 were not married or partnered (Goodwin et al., 2010).

3.3.2 Municipality Sensitivity Analysis

To be able to provide municipalities insight in what are the risk and protective factors for their region, we perform sensitivity analysis for each of the 285 municipalities. Table 18 shows the sensitivity analysis of municipality 20558 with varying factor levels and their predicted effect on the proportion of inhabitants with anxiety disorders expressed in a percentage based on the neural network model. In contrast to the summed weights as presented in Table 4, the sensitivity levels of Table 19 and Table 20 do present the predicted change in the proportion of inhabitants with anxiety disorders by altering the values of each factor. The difference in interpretability is due to the method used. For the national analysis, we sum the gradients of 320 municipalities per factor and correct for the population size per municipality, to focus on ranking. For the municipality analysis, we determine the predicted value per percentual change of each factor, to focus on interpretability. Thus providing easy insight to policymakers for municipalities.

For example, we notice that if the number of childcare centers within 3 km would be lowered by 5%, the proportion of inhabitants with anxiety disorders is predicted to decrease by 1.30%. If the number of restaurants within 3 km would be increased by 10%, the proportion of inhabitants with anxiety disorders is predicted to decrease by 2.54%. While factors such as the percentage of unmarried and divorced individuals, and the percentage of the population within the 25-45 age range may be harder to influence through policy interventions, they can provide valuable insight into target groups for municipalities and inform the prioritization of factors to address. This information can be used to provide customized recommendations for each municipality.

3.4 Neural Network

Table 5 presents the performance per topology of the neural network with epochs=100, 6 neurons per dense layer.

Topology	**Number of parameters**	**Performance (in MSE)**
1 neuron per category, 1 dense layer	262	0.0068
1 neuron per category, 2 dense layers	304	0.0059
1 neuron per category, 3 dense layers	346	0.0062
1 neuron per category, 4 dense layers	388	0.0069
1 neuron per category, 5 dense layers	430	0.0055
1 neuron per category, 6 dense layers	472	0.0054
1 neuron per category, 7 dense layers	514	0.0065
1 neuron per category, 8 dense layers	556	0.0064

Table 5 presents the number of parameters and the performance in mean squared error (MSE) for each topology. The different topologies represent the number of dense layers, with each layer having one neuron per category. The results show that the best performance is achieved by the topology with 1 neuron per category and 6 dense layers, which has an MSE of 0.0054. Conversely, the worst performance is achieved by the topology with 1 neuron per category and 1 dense layer, which has an MSE of 0.0068. The number of parameters in the neural network increases as the number of dense layers increases, but the performance does not necessarily improve accordingly. The results suggest that adding more dense layers can improve the neural network's performance, but there is a limit to the number of dense layers that can be added before performance starts to deteriorate.

*Table 6 presents the performance per topology of the neural network with epochs=100, 6 neurons per dense layer, and n = number of factors per category with *minimally 1 neuron per category.*

Topology	Number of parameters	Performance (in MSE)
1 neuron per category, 1 dense layer	262	0.0068
Set 1 or 2 neurons based on the number of factors in each category: < 8 or ≥8, respectively, 1 dense layer	376	0.0063
Set 1, 2, or 3 neurons based on the number of factors in each category: < 8, between 8-16, or >16, respectively, 1 dense layer	442	0.0056
$\sqrt{\boldsymbol{n}}$ neurons per category rounded down to integers*, 1 dense layer	1,300	0.0059
$\frac{\boldsymbol{n}}{\boldsymbol{2}}$ neurons per category rounded down to integers*, 1 dense layer	1,348	0.0054
1 neuron per factor, 1 dense layer	2,783	0.0055

Table 6 shows the performance of different neural network topologies tested with the same number of epochs and 6 neurons per dense layer. We measure the results in MSE and we record the number of parameters for each topology. The first topology with one neuron per category and one dense layer has the fewest parameters but the highest MSE, indicating low complexity. Adding more neurons per category improves the model its performance but adding more neurons not necessarily leads to better performance. We continue with the approach of 1/2/3 neurons per category because of the relatively well performance (MSE=0.0056) with a low number of parameters (n=442). A low number of parameters in a neural network decreases the risk of overfitting (Gupta et al, 2018).

Table 7 presents the performance of the neural network with epochs=10,000, 6 neurons per dense layer.

Topology	Number of parameters	MSE	Avg. validation score 5-fold cross. in MSE
Set 1, 2, or 3 neurons based on the number of factor in each category: < 8, between 8-16, or >16, respectively, 6 dense layers	652	0.0038	0.0069

An MSE of 0.0038 means that the neural network model has a relatively low error rate, indicating that it is accurately predicting the proportion of inhabitants with anxiety disorders. The average validation score for 5-fold cross-validation in MSE of 0.0069 means that the neural network model its performance was consistent across different folds of the data, and that it performs well on unseen data.

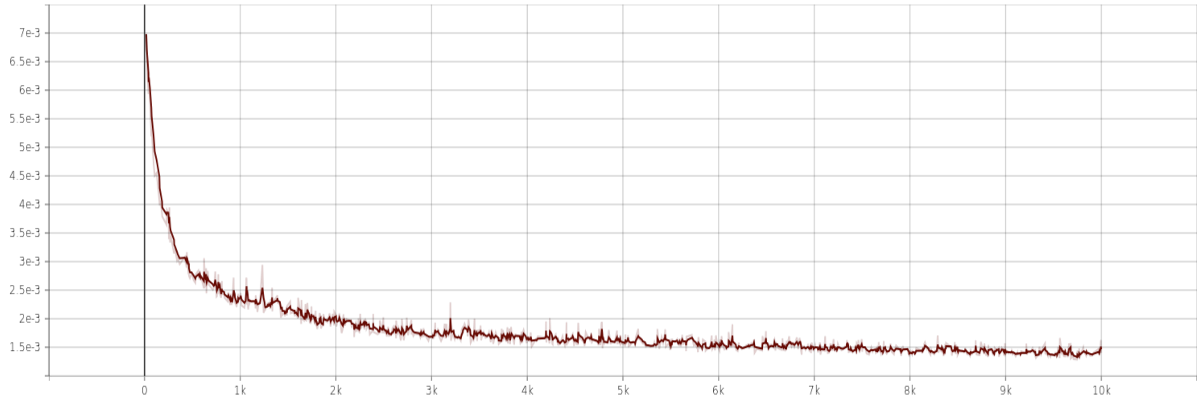

Figure 4 represents the loss function graph by visualizing how the MSE between the predicted output and actual output of the neural network changes during the training process as the number of epochs increases from 0 to 10,000.

The MSE is plotted against the number of epochs to give insight of how the model its performance is improving with more training. Initially, the MSE is high as the network is making random predictions, but it decreases as the network learns from the training data. The MSE decreases with increasing epochs going from 0 to 10,000 and levels off to a minimum value, indicating that the model converges well.

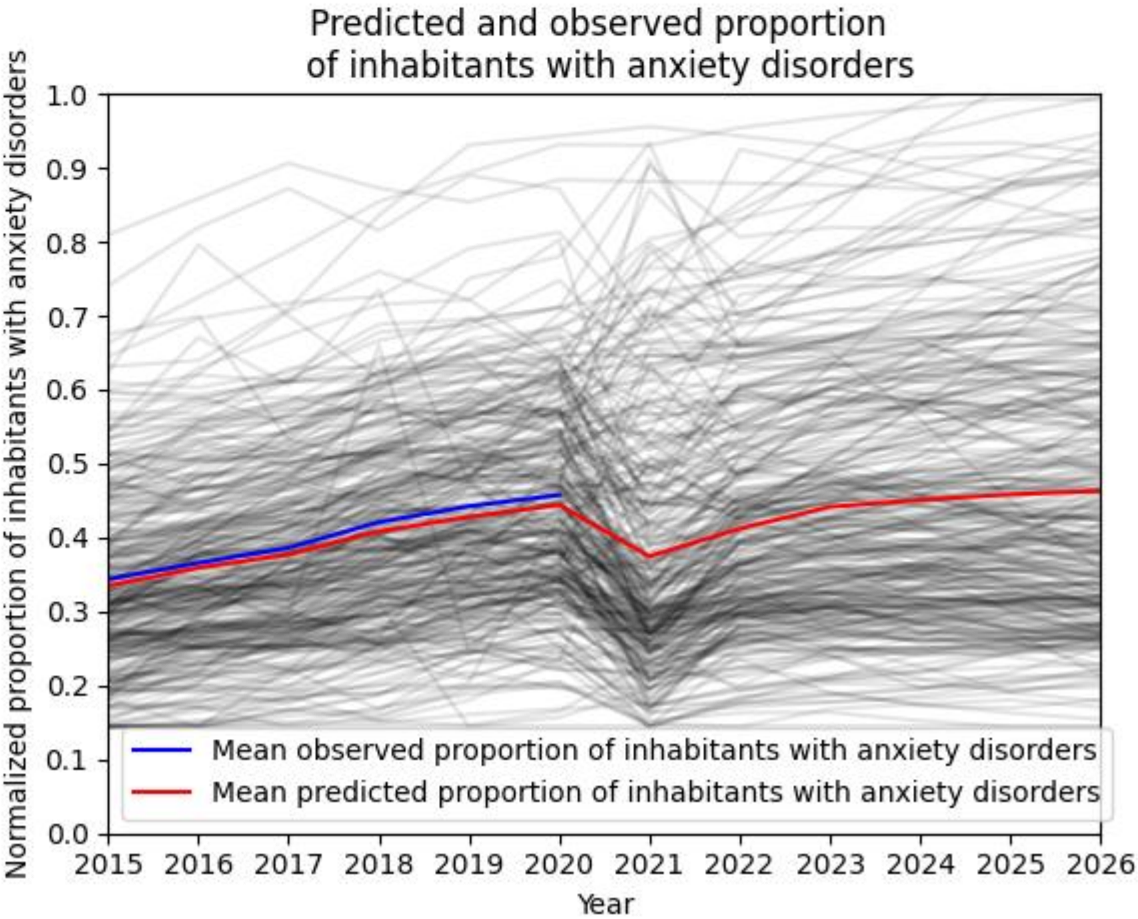

Figure 5 displays a combined graph of all the predictions where each black line resembles one municipality (n=320). The blue and red thick lines resemble the mean observed and mean predicted number of anxiety disorders respectively.

Figure 5 presents all the predictions of the proportion of inhabitants with anxiety disorders for each individual municipality (n=320) represented by a faded black line over the period 2015-2026 and plotted on a normalized Y-axis between 0 and 1. These predictions are based on the neural network model (See Table 7 and Figure 3). The blue and red lines resemble the mean observed and mean predicted proportion of inhabitants with anxiety disorders respectively. The blue line ends after 2020 since the number of anxiety disorders has been observed from 2015 – 2020 (See Figure 1). The mean observed and mean predicted proportion of inhabitants with anxiety disorders is given in Table 8. For observations and predictions of the proportion of inhabitants with anxiety disorders for individual municipalities, see Figure 8.

Table 8 presents the mean observed and mean predicted proportion of inhabitants with anxiety disorders on a national level for the years 2015-2020 and 2015-2026 respectively.

Year	**Mean observed**	**Mean predicted**
2015	0.344	0.335
2016	0.366	0.359
2017	0.386	0.377
2018	0.421	0.408
2019	0.442	0.428
2020	0.458	0.444
2021	Unknown	0.375
2022	Unknown	0.412
2023	Unfinished year	0.441
2024	Future year	0.451
2025	Future year	0.458
2026	Future year	0.463

To evaluate the spread of individual municipalities, we calculate the percentage of predictions for each year that fall within a +/- 0.25 range from the mean predicted proportion of inhabitants with anxiety disorders. On average, 90.0% of the predictions for individual municipalities were within this range. To evaluate the trend among municipalities if we would set the Y-value for t=2015 for all municipalities equal, then we are able to visualize more clearly how the predictions of individual municipalities widen out as years progress (See Figure 6).

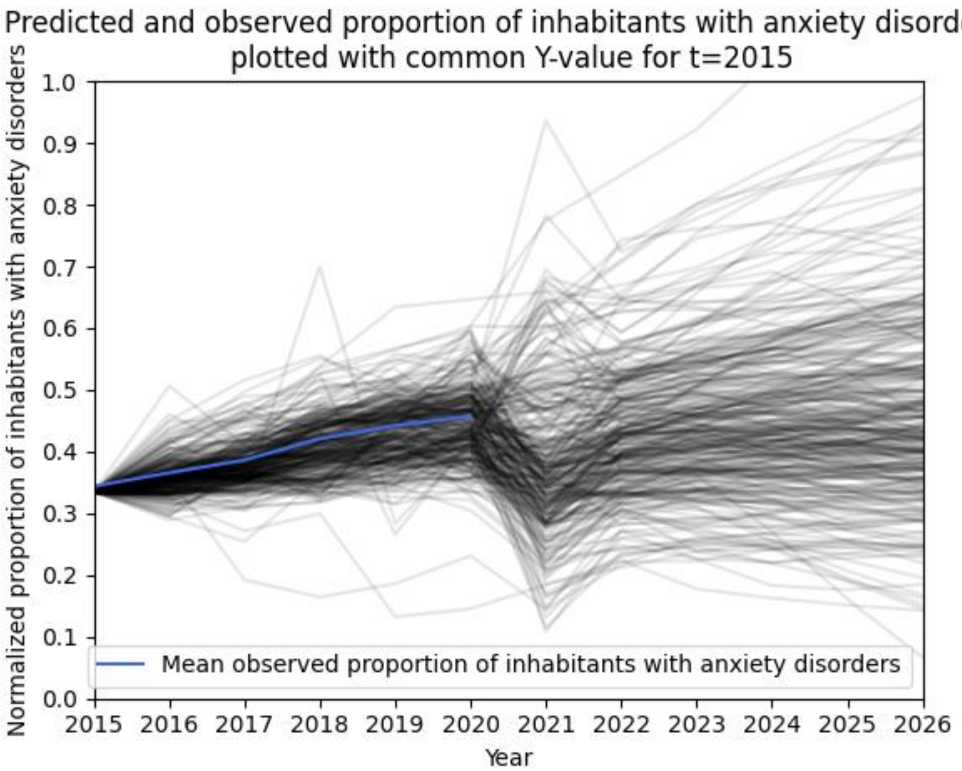

Figure 6 displays a combined graph of all the predictions where each black line resembles one municipality (n=320). The blue line resembles the mean observed number of anxiety disorders. The starting value in 2015 has been set to the mean value for all the municipalities in 2015 (=0.33) to visualize how each prediction spreads over time after a common Y-value for t=2015.

To visualize additional insights into the trends and patterns in the data, we plot the gradient. Specifically, the gradient can help to identify periods of steep increase or decrease, as well as any areas of relative stability or consistency. By plotting the gradient, we can also compare the rate of change between different time periods (Figure 7).

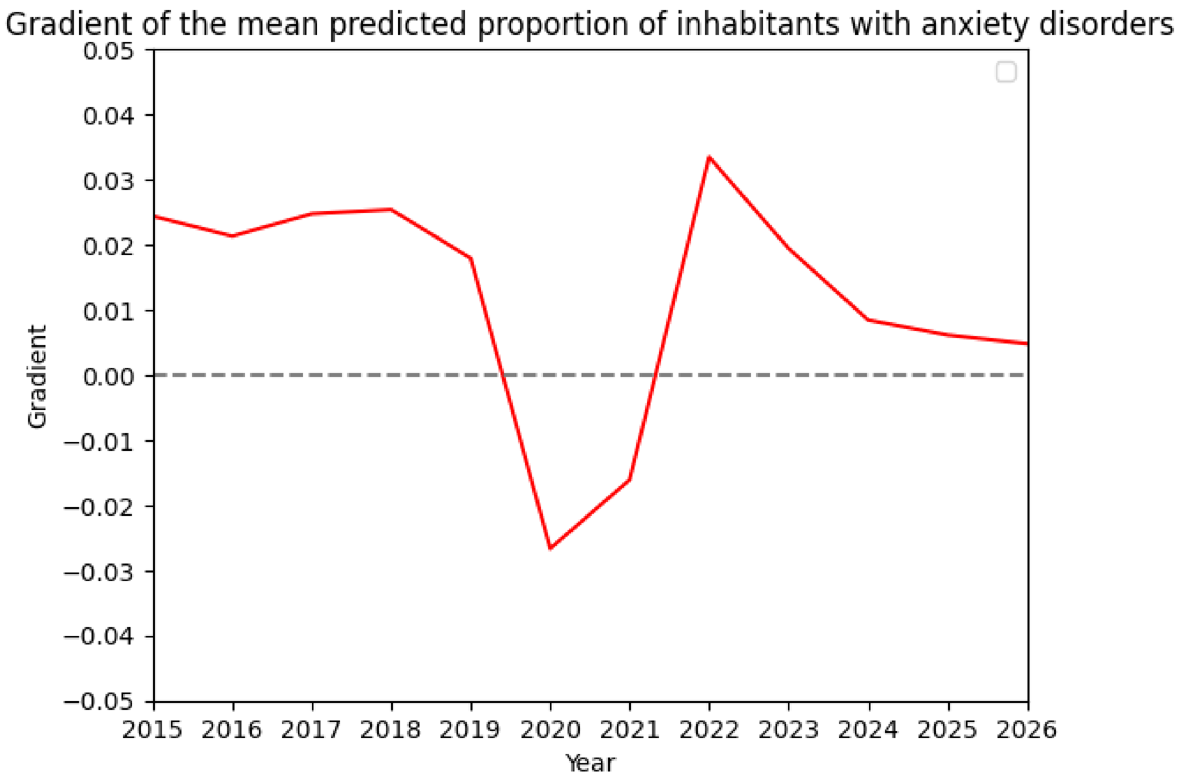

Figure 7 visualizes the gradient of the mean predicted proportion of inhabitants with anxiety disorders based on predictions with the neural network (See Table 7 and Figure 3).

We observe a clear increase in the mean normalized proportion of inhabitants with anxiety disorders between the years 2015 and 2019, with an average gradient of +0.0179 to +0.0254 per year. However, the trend reverses in 2020 and 2021, with negative gradients of -0.0267 and -0.0161, respectively. Following this steep drop in the mean number the proportion of inhabitants with anxiety disorders, we observe positive gradients for the years 2022 to 2026. The year 2022 has the highest gradient at +0.033, which decreases in subsequent years but remains positive, with the lowest gradient observed in 2026 at +0.006.

There could be several factors that could cause the steep drop in the gradient of the mean normalized number of the proportion of inhabitants with anxiety disorders. One possible explanation could be the effects of the COVID-19 pandemic, which may have led to changes in social interactions, lifestyle habits, and overall stress levels of the population. The pandemic may have led to an increased focus on mental health and the availability of mental health services, which could have contributed to a decrease in anxiety disorders among the population.

We observe a consistency between our findings and those reported by Trimbos, indicating a nationwide decrease in the number of anxiety disorders during the COVID-19 pandemic. However, we face a challenge in directly linking the decrease in the mean normalized proportion of inhabitants with anxiety disorders from our neural network to a corresponding decrease of 13.9% to 13.2% of Dutch adults with anxiety disorders before and during the pandemic, respectively, measured by the Trimbos institute (Trimbos$_2$, 2022).

3.4.1 Prediction models summary

Table 9 gives a summary of the performance per methodology used to predict the proportion of inhabitants with anxiety disorders based on public datasets.

Prediction methodology	MSE	R^2
Linear regression based on correlation heatmap selection	0.0273	0.5941
Lasso regression	0.0059	0.7005
Neural network	0.0038	0.8412

We develop three models: linear regression, lasso regression, and a neural network, to analyze the relationship between the input factors and the output factor, which is the proportion of inhabitants with anxiety disorders. Our linear regression model has a moderate performance with an MSE of 0.0273, indicating that the predicted values deviate from the actual values by approximately 0.165 on a normalized scale from 0 to 1. On the other hand, our Lasso regression model has a strong performance with an MSE of 0.0059, which suggests that the predicted values deviate from the actual values by approximately 0.077 on a normalized scale. The neural network model has the best performance with an MSE of 0.0038, implying that the predicted values deviate from the actual values by approximately 0.062 on a normalized scale.

Moreover, our linear regression model has an R^2 of 0.5941, indicating that 59.41% of the variance in the output factor can be explained by the input factors in our model. The Lasso regression model has an R^2 of 0.7005, which is an improvement over the linear regression model. However, our neural network model has the best performance with an R^2 of 0.8412, explaining 84.12% of the variance in the output factor, a higher value than that obtained by the linear regression and Lasso models. Our findings suggest that the neural network model is the most reliable for predicting the output factor, providing valuable insights into the relationship between the input factors and the output factor.

Table 10 shows the top 5 risk and protective factors determined per methodology for anxiety disorder prevalence on a national level. Underlined factors occur twice in the top 5 risk and protective factors for two different methodologies.

	Cross-correlation	**Lasso regression**	**Neural network**
Risk factors			
1	Welfare per inhabitant	Net income: Couple with no child	Divorced population
2	Address density	Median wealth: Employee income	Median wealth: Employee income
3	Non-Western migration background	Population density	Standardized income: rental property
4	Number of VMBO schools within 5 km	Median wealth: own home	Median wealth: Couple with no child
5	Distance to secondary school	Standardized income: Self-employment income	Number of childcare centers within 3 km
Protective factors			
1	Married population	Median wealth: Private households	Median wealth: Private households
2	Houses to buy	Average household size	Widowed population
3	Both parents born in The Netherlands	Housing density	Married population
4	Average household size	Standardized income: Private households	Median wealth: Couple with child(ren)
5	Households without children	Distance to daycare center	Total age pressure

Risk factors with multiple occurrences

In Table 10 we notice that the net income of couples with no child ranks as first and fourth most influential risk factor of anxiety disorder prevalence for lasso regression and the neural network, respectively. Meaning that if the net income of couples with no child increases, the proportion of inhabitants with anxiety disorders is predicted to increase in case of these methods. This is not in line with research done by Helbig et al. (2006) which examined a German population of 2,801 parents and non-parents and found that the prevalence of anxiety disorders tends to decrease as the income level rises. However, this is based on data from Germany in 1999 thus is a different population and timeframe than our scope. It is also possible that couples with no child and higher net income are more likely to have themselves examined by a GP or psychologist, leading to increased reporting and diagnosis of anxiety disorders. This could result in a higher observed prevalence of anxiety disorders among this group.

We notice that the median wealth of inhabitants who earn their income as employee ranks as the second most influential risk factor of anxiety disorder prevalence for both lasso regression and the neural network. Meaning that if the median wealth of inhabitants who earn their income as employee increases, the proportion of inhabitants with anxiety disorders is predicted to increase as well in with these methods. This is not in line with research done by Rutter (2003) which states that a higher socio-economic score is associated with lower risk of anxiety disorders.

Protective factors with multiple occurrences

We notice that the median wealth of inhabitants who belong to private household ranks as the second most influential protective factor of anxiety disorder prevalence for both lasso regression and the neural network. Persons who live alone or together in a housing unit and are able to provide for their daily needs themselves constitute a private household. The results imply that if the median wealth of inhabitants who belong to private household increases, the proportion of inhabitants with anxiety disorders is predicted to increase for the lasso regression and neural network. This could be due to that higher median wealth among individuals in private households may indicate a greater level of financial security and stability. Financial security can contribute to reduced stress levels and provide individuals with a sense of control and confidence, which may act as a protective factor against anxiety disorders (Rutter, 2003).

We notice that the percentage of grown-ups who are married ranks as first and third most influential protective factor of anxiety disorder prevalence for the Pearson correlation and the neural network, respectively. Meaning that if the percentage of grown-ups who are married increases, the proportion of inhabitants with anxiety disorders is predicted to decrease in case of these methods. This is in line with research done by Cariney et al. (2006) with data about 1,346 women which found higher prevalence of anxiety disorders for women with a child and not being married than for women with a child and being married. Research done by Helbig et al. (2006) suggests that partnership especially is associated with lower risk of anxiety disorders if someone has children.

We notice that the average household size ranks as fourth and second most influential protective factor of anxiety disorder prevalence for the Pearson correlation and the neural network, respectively. Meaning that if the average household size increases, the proportion of inhabitants with anxiety disorders is predicted to decrease in case of these methods. This is in in line with research done by Helbig et al. (2006) which examined a German population of 2,801 parents and non-parents and found that the prevalence of anxiety disorders is significantly lower for parents with two children, than for parents with one child. However, this is based on data from Germany in 1999 thus is a different population and timeframe than our scope.

Factors with overlapping definition

Furthermore, we notice that factors related to not having children score as first and third risk factor for lasso regression and the neural network respectively. These factors are the net income of couples with no children (first risk factor, lasso regression), the median wealth of couples with no children (third risk factor, neural network) However, the factor describing the percentage of households without children, does rank as fifth protective factor for the Pearson correlation. Meaning that not having children is related to higher prevalence of anxiety disorders according to the lasso regression model and neural network, but related to lower prevalence of anxiety disorders according to the Pearson correlation. The Pearson correlation measure the linear relationship between variables without correcting for collinearity while the lasso regression and neural network use regularization techniques to penalize for collinearity among factors thus may produce different results than the Pearson correlation.

If we focus on the factors related to inhabitants which do have children, we notice that the median wealth of couples with one (or multiple) child(ren) ranks as fourth protective factor for the neural network, and the average distance to a daycare center ranks as the fifth protective factor for the lasso regression. However, the number of childcare centers within 3 km is ranked as the fifth risk factor for the neural network, and the factors related to children in secondary schools (12-18 years old) are protective factors according to the Pearson correlation scoring fourth and fifth place. These discrepancies among models may arise from the specific characteristics and assumptions of each model, as well as the differences in how they capture and interpret relationships between variables.

4 Discussion

We demonstrate that our study has great potential for data-driven policy making based on public datasets for the prediction of mental disorders and to determine which risk and protective factors to prioritize with policy.

In the past, the lack of comprehensive data made it difficult to assess the effect of certain policies on mental health. However, the development of machine learning techniques has made it easier to analyze large and complex datasets. These tools have become more accessible and easier to use, allowing researchers and policymakers to make data-driven decisions and improve mental health outcomes. With our study we develop models that can provide insight into the effect of factors on anxiety disorders.

Risk and protective factors

The first goal of our study is to determine the risk factors and protective factors related to the prevalence of anxiety disorders based on public data between 2015-2020.

After examining the Pearson correlation score, we observed that the proportion of inhabitants receiving social benefits, the address density, and percentage of inhabitants with a non-western migration background have the highest correlation with anxiety disorders. On the other hand, the percentage of people 15 years and older being married, the percentage of houses available for purchase, and the percentage of inhabitants with a Dutch background have the lowest correlation with anxiety disorders. Nevertheless, we found that there is collinearity among these factors, which is shown in Table 1.

By using lasso regression, we identified that the median wealth of couples with no child, the median wealth of people with an income as employee, and the population density are risk factors for a higher proportion of inhabitants with anxiety disorders. Conversely, the median wealth of private households, the average size of households, and the housing density are protective factors. Furthermore, based on the sensitivity analysis and neural network, we found that the risk factors for a higher proportion of inhabitants with anxiety disorders are the percentage of the population above 15 years who are divorced, the median wealth of people with an income as employee, and the standardized income of people living in rental property. In contrast, the median wealth of private households and the percentage of people above 15 years who are married or widowed were identified as protective factors.

The differences in the ranking of the risk and protective factors between the Pearson correlation, lasso regression, and neural network could be attributed to the characteristics of each technique. While Pearson correlation measures linear relationships between two factors and fails to consider the influence of collinearity, lasso regression selects essential features and adds a penalty term to achieve a better generalization model but assumes a linear relationship between input factors and the output factor. On the other hand, a neural network is powerful in capturing nonlinear higher-order relationships. Therefore, we would prioritize the results of the risk and protective factors obtained from the neural network over those obtained from the lasso regression and Pearson correlation techniques.

Prediction of anxiety disorders

The second goal of our study is to determine how we can build an accurate prediction model to predict the prevalence of anxiety disorders between 2021-2026. We show that we can build a neural network to predict the proportion of inhabitants with an anxiety disorder with a MSE of 0.0038 and a R^2 of 0.8412. Although the predictions for 2021-2016 assume a linear trend for all public data, this assumption may not hold for all factors, indicating a need to improve the trend determination method for each factor.

Prioritization of factors for intervention

The last goal of our study is to determine which factors should VWS and municipalities prioritize for their interventions to decrease the number of anxiety disorders. The national sensitivity analysis based on the neural network does provide meaningful insight on which factors VWS should prioritize their interventions. From Figure 9 we observe that there seem to be a lot of difference in positive and negative sensitivity among the same categories of factors. For example, the income and wealth factors differ a lot based on the group it describes. Income and wealth for employees, people who are self-employed, people who own a home or rent property have a positive sensitivity with the prevalence of anxiety disorders. Meaning that if the value of these factors increase, the proportion of inhabitants with anxiety disorders is predicted to increase as well. The income of private households and couples with children has a negative sensitivity for the prevalence of anxiety disorders, meaning that these factors increase, the proportion of inhabitants with anxiety disorders is predicted to decrease. Having children is associated with lower risk of anxiety disorders (Helbig et al., 2006). Another explanation for these different results could be related to the level of stress associated with different types of employment and housing situations. For example, employees or self-employed individuals with higher income and wealth may experience more job-related stress or work longer hours, which could increase their risk for developing anxiety disorders. Similarly, individuals who own homes or rent property may experience more financial stress or instability, which could also increase their risk for anxiety disorders. However, research about the association between these specific groups and the prevalence of anxiety disorders has not been found in literature.

The age group division shows the percentage of people in age groups younger than 25 years old have a negative sensitivity for the proportion of inhabitants with anxiety disorders. However, the percentage of the population belonging to the age groups 25 up to 80 + have a positive sensitivity (except for 65-80 years) with the proportion of inhabitants with anxiety disorders. A study based on US mental health data reported that the lifetime prevalence of any anxiety disorder was 30.2% in the 18-29 age group, 35.1% in the 30-44 age group and 15.3% among individuals aged 60 years and above (Kessler et al., 2005). The results of our study are in line for the age group of 30-44 years old, but do not show comparable results for the age group of 18-29. This may be due to differences in the population, timeframe and thresholds used.

The sensitivity analysis provides insight for each individual municipality which factors have the most influence on the prevalence of anxiety disorders, thus could give meaningful information about the target group and which risk and protective factors to prioritize.

Limitations

We acknowledge that the data used in our study is limited to public data between 2015-2020, which may not accurately reflect the current situation. This is since there is currently a two-and-a-half-year delay in the data collection process of the number of patients by Vektis. Our study is therefore constrained by the temporal scope. The use of aggregated data means that some relevant factors, such as clinical data on the individual level, is not included in our analysis. Additionally, public data may not fully capture the complexity that may influence mental health outcomes because of confounding factors. Furthermore, our study relies on the assumption that the public data accurately reflects the prevalence of anxiety disorders in each municipality. This may not be entirely accurate as not all cases of anxiety disorders are reported or diagnosed, and there may be differences in how mental health services are accessed and delivered across municipalities.

Within the scope of our research, we focus on 5 different types of anxiety disorders, while not taking into account other types. One type which is not included within the scope of this study is obsessive compulsive disorder (OCD) which had a median 12-month prevalence rate across 21 European studies of 0.7% in 2005 (Wittchen & Jacobi, 2005). In comparison, the included types of disorders such as panic disorders (1.8%), agoraphobia (1.3%), social phobias (2.3%), generalized anxiety disorders (1.7%), and specific phobias (6.4%) demonstrated higher prevalence rates and are included in this study.

Further improvements of the neural network could be made by exploring different topology combinations by varying the number of layers, the number of neurons per layer, the activation function, the learning rate, regularization and optimization algorithms. By retraining the model and adding more years of data when available, the model would be able to improve its accuracy and generalize better to future data. Additionally, including more relevant features and addressing any potential biases in the data could also contribute to better performance. Potential biases could be the timeframe used which may not generalize well to more recent data and the features selected to train the neural network may introduce biases if they are incomplete or exclude relevant variables that could influence the predictions. Biased feature selection can limit the model's ability to capture the full complexity of the problem.

Future implications

When further developed, our model could be used to direct policies and determine the most effective preventive and corrective measures to decrease the number of anxiety disorders. Moreover, we can use cost-benefit analyses to assess the impact of our findings on the costs and number of patients. For example, we can quantify the expected reduction in the number of patients per year and the expected reduction of costs since both are measured in the Vektis data from 2015-2020 for each municipality. Since we can quantify the costs, we can make a cost benefit analysis to check whether policies directly targeting risk factors would be worth implementing. By reducing the number of patients in mental care, we could contribute to a stagnation or reduction in waiting lists for mental health care, especially in the current crisis faced by GGZ. We could predict the risk and protective factors as well on the average treatment time spent per patient, since this is known in the Vektis data for each municipality specified per diagnosis.

Although not all risk and protective factors may be easily adaptable to policy, we can identify which groups are at risk of anxiety disorders in specific regions. This approach is similar to cancer prevention screening programs where certain groups of people are invited for screening based on their age, gender, or other risk factors. By identifying groups at risk of anxiety disorders in specific regions, policymakers can tailor interventions and preventive measures to those who are most likely to benefit from them.

Our study can be used to design new policies by predicting the impact of future developments on the prevalence of anxiety disorders. Policymakers can use our models to assess the effectiveness of potential interventions before they are implemented, thus saving significant time and resources. Moreover, we believe that our methods can be applied to other types of mental disorders, thereby expanding the scope of data-driven policy making. We can change the way VWS, and municipalities make policies and potentially predict other health outcomes beyond anxiety disorders.

5 Conclusion

In conclusion, our study aims to determine the risk and protective factors related to the prevalence of anxiety disorders based on public data from 2015 to 2020 and to build an accurate prediction model to predict the prevalence of anxiety disorders from 2021 to 2026.

We conducted a thorough analysis and modeling phase using various techniques, such as cross-correlation with Pearson correlation, standard linear regression, lasso regression, and a sensitivity analysis based on a neural network. By performing these techniques in order, we gradually increased the complexity of our analysis and gained deeper insights into the relationships between the input factors and the output factor: the proportion of inhabitants with anxiety disorders. The risk and protective factors identified by our models do not necessarily align with previous literature. To enhance the performance of our most promising model, the neural network, potential improvements include exploring various topologies and incorporating additional years of data.

By predicting the future prevalence of anxiety disorders, we gain insights into the potential burden of the condition and the need for intervention. At the same time, by identifying the risk and protective factors based on public data, we can understand the underlying drivers of the outcome, which can inform targeted and effective interventions.

Together, this study provides an understanding of the factors that contribute to anxiety disorders and guide the development of effective strategies to reduce the prevalence of anxiety disorders and improve the quality of life for individuals affected by this mental health disorder.

6 References

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

An Analysis of Variance Test for Normality (Complete Samples) on JSTOR. (n.d.). https://www.jstor.org/stable/2333709

Antony, M. M., Roth, D. A., Swinson, R. P., Huta, V., & Devins, G. M. (1998). Illness Intrusiveness in Individuals with Panic Disorder, Obsessive-Compulsive Disorder, or Social Phobia. *Journal of Nervous and Mental Disease*, *186*(5), 311–315. https://doi.org/10.1097/00005053-199805000-00008

Anxiety Disorders. (n.d.). National Institute of Mental Health (NIMH). https://www.nimh.nih.gov/health/topics/anxiety-disorders

Anxiety Disorders. (2003, February 1). WebMD. https://www.webmd.com/anxiety-panic/guide/anxiety-disorders

Anxiety disorders - Symptoms and causes - Mayo Clinic. (2018, May 4). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/anxiety/symptoms-causes/syc-20350961

Anxiety Disorders: Types, Causes, Symptoms & Treatments. (n.d.-a). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/9536-anxiety-disorders

Anxiety Disorders: Types, Causes, Symptoms & Treatments. (n.d.-b). Cleveland Clinic. https://my.clevelandclinic.org/health/diseases/9536-anxiety-disorders

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, *17*(3), 327–335. https://doi.org/10.31887/dcns.2015.17.3/bbandelow

Basil, N. N., Ambe, S., Ekhator, C., & Fonkem, E. (2022). Health Records Database and Inherent Security Concerns: A Review of the Literature. Cureus. https://doi.org/10.7759/cureus.30168

Cairney, J., Pevalin, D. J., Wade, T. J., Veldhuizen, S., & Arboleda-Flórez, J. (2006). Twelve-Month Psychiatric Disorder among Single and Married Mothers: The Role of Marital History. *The Canadian Journal of Psychiatry*, *51*(10), 671–676. https://doi.org/10.1177/070674370605101007

CBS Statline. (n.d.-a). https://opendata.cbs.nl/

CBS Statline. (n.d.-b). https://opendata.cbs.nl/statline/

Chatterjee, M., Stratou, G., Scherer, S., & Morency, L. (2014). Context-based signal descriptors of heart-rate variability for anxiety assessment. In *Zenodo (CERN European Organization for Nuclear Research)*. CERN European Organization for Nuclear Research. https://doi.org/10.1109/icassp.2014.6854278

Choi, K. R., Heilemann, M. V., Fauer, A., & Mead, M. (2020). A Second Pandemic: Mental Health Spillover From the Novel Coronavirus (COVID-19). *Journal of the American Psychiatric Nurses Association*, *26*(4), 340–343. https://doi.org/10.1177/1078390320919803

Dahl, C. M., & Dahl, A. A. (2010). Lifestyle and social network in individuals with high level of social phobia/anxiety symptoms: a community-based study. *Social Psychiatry and Psychiatric Epidemiology*. https://doi.org/10.1007/s00127-009-0069-6

Davidson, J. R. T., Hughes, D. H., George, L. K., & Blazer, D. G. (1993). The epidemiology of social phobia: findings from the Duke Epidemiological Catchment Area Study. *Psychological Medicine*, *23*(3), 709–718. https://doi.org/10.1017/s0033291700025484

De Gezondheidsmonitors | Gezondheidsmonitor. (n.d.). https://www.monitorgezondheid.nl/

Generalized Anxiety Disorder. (2003, February 6). WebMD. https://www.webmd.com/anxiety-panic/guide/generalized-anxiety-disorder

Geregistreerde overlast; soort overlast, gemeentelijke indeling 2023. (n.d.). Data Overheid. https://data.overheid.nl/en/dataset/5253-geregistreerde-overlast--soort-overlast--gemeente

Gezondheidsmonitor; bevolking 18 jaar of ouder, regio, 2020. (n.d.). Data Overheid. https://data.overheid.nl/en/dataset/16264-gezondheidsmonitor--bevolking-18-jaar-of-ouder--regio--2020

Gibson-Smith, D., Bot, M., Brouwer, I. A., Visser, M., Giltay, E. J., & Penninx, B. W. (2020). Association of food groups with depression and anxiety disorders. *European Journal of Nutrition*, *59*(2), 767–778. https://doi.org/10.1007/s00394-019-01943-4

Goodwin, P. Y., Mosher, W. D., & Chandra, A. (2010). Marriage and cohabitation in the United States: A statistical portrait based on cycle 6 (2002) of the National Survey of Family Growth. *Vital and Health Statistics. Series 23, Data From the National Survey of Family Growth*, *28*, 1–45. https://pubmed.ncbi.nlm.nih.gov/20629347/

Gupta, S., Gupta, R., Ojha, M., & Singh, K. P. (2017). A Comparative Analysis of Various Regularization Techniques to Solve Overfitting Problem in Artificial Neural Network. In *Communications in computer and information science* (pp. 363–371). Springer Science+Business Media. https://doi.org/10.1007/978-981-10-8527-7_30

Helbig, S., Lampert, T., Klose, M., & Jacobi, F. (2006). Is parenthood associated with mental health? *Social Psychiatry and Psychiatric Epidemiology*, *41*(11), 889–896. https://doi.org/10.1007/s00127-006-0113-8

Hilbert, K., Lueken, U., Muehlhan, M., & Beesdo-Baum, K. (2017). Separating generalized anxiety disorder from major depression using clinical, hormonal, and structural MRI data: A multimodal machine learning study. *Brain and Behavior*, *7*(3), e00633. https://doi.org/10.1002/brb3.633

Hoge, E. a., Pollack, M. H., Kaufman, R. E., Zak, P. J., & Simon, N. M. (2008). Oxytocin levels in social anxiety disorder. *CNS Neuroscience & Therapeutics, 14(3), 165–170.*

Home - Waarstaatjegemeente.nl - cijfers en statistieken van gemeenten. (n.d.). https://www.waarstaatjegemeente.nl/

Kingma, D. P. (2014, December 22). *Adam: A Method for Stochastic Optimization*. arXiv.org. https://arxiv.org/abs/1412.6980

Koers, J. (2022a, December 1). Sterke toename psychische aandoeningen bij volwassenen. *Trimbos-instituut*. https://www.trimbos.nl/actueel/nieuws/sterke-toename-psychische-aandoeningen-bij-volwassenen/

Koers, J. (2022b, December 1). Sterke toename psychische aandoeningen bij volwassenen. *Trimbos-instituut*. https://www.trimbos.nl/actueel/nieuws/sterke-toename-psychische-aandoeningen-bij-volwassenen/

Li, Z. (2019, October 16). *An Exponential Learning Rate Schedule for Deep Learning*. arXiv.org. https://arxiv.org/abs/1910.07454

Liu, L., Villavicencio, F., Yeung, D., Perin, J., Lopez, G. R., Strong, K., & Black, R. E. (2022). National, regional, and global causes of mortality in 5–19-year-olds from 2000 to 2019: a systematic analysis. *The Lancet Global Health*, *10*(3), e337–e347. https://doi.org/10.1016/s2214-109x(21)00566-0

Long, N. J. (2020). From social distancing to social containment. *Medicine Anthropology Theory*, *7*(2), 247–260. https://doi.org/10.17157/mat.7.2.791

Martin, P. (2003). The epidemiology of anxiety disorders: a review. *Dialogues in Clinical Neuroscience*, *5*(3), 281–298. https://doi.org/10.31887/dcns.2003.5.3/pmartin

McDonald, G. C. (2009). Ridge regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(1), 93–100. https://doi.org/10.1002/wics.14

McLaughlin, K. A., Costello, E. J., LeBlanc, W. G., Sampson, N. A., & Kessler, R. C. (2012). Socioeconomic Status and Adolescent Mental Disorders. *American Journal of Public Health*, *102*(9), 1742–1750. https://doi.org/10.2105/ajph.2011.300477

Mental Health (MHF). (2018). Mental health: fact sheet on Sustainable Development Goals (SDGs): health targets. *www.who.int*. https://www.who.int/europe/publications/i/item/WHO-EURO-2018-2364-42119-58012

Ministerie van Algemene Zaken. (2022, June 23). *Rapport De juiste zorg op de juiste plek*. Rapport | Rijksoverheid.nl. https://www.rijksoverheid.nl/documenten/rapporten/2018/04/06/rapport-de-juiste-zorg-op-de-juiste-plek

Ministerie van Volksgezondheid, Welzijn en Sport. (2022a, December 7). *Monitor Psychische Problematiek - Samen werken aan zorg, ondersteuning en herstel*. Publicatie | Rijksoverheid.nl. https://www.rijksoverheid.nl/documenten/publicaties/2022/08/29/voortgangsrapportage-monitor-psychische-problematiek

Ministerie van Volksgezondheid, Welzijn en Sport. (2022b, December 7). *Monitor Psychische Problematiek - Samen werken aan zorg, ondersteuning en herstel*. Publicatie | Rijksoverheid.nl. https://www.rijksoverheid.nl/documenten/publicaties/2022/08/29/voortgangsrapportage-monitor-psychische-problematiek

Ministerie van Volksgezondheid, Welzijn en Sport. (2023, January 3). *Integraal Zorgakkoord: "Samen werken aan gezonde zorg."* Rapport | Rijksoverheid.nl. https://www.rijksoverheid.nl/documenten/rapporten/2022/09/16/integraal-zorgakkoord-samen-werken-aan-gezonde-zorg

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. Http:/noiselab.ucsd.edu/ECE228/Murphy_Machine_Learning.pdf. Retrieved May 1, 2023, from https://www.cs.ubc.ca/~murphyk/MLbook/

Nemesis. (2023a). Psychische aandoeningen vóór en tijdens corona. *Nemesis*. https://cijfers.trimbos.nl/nemesis/effect-coronapandemie/vergelijking-corona/

Nemesis. (2023b). Psychische aandoeningen vóór en tijdens corona. *Nemesis*. https://cijfers.trimbos.nl/nemesis/effect-coronapandemie/vergelijking-corona/

Nepon, J., Belik, S., Bolton, J. R., & Sareen, J. (2010). The relationship between anxiety disorders and suicide attempts: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *Depression and Anxiety*, *27*(9), 791–798. https://doi.org/10.1002/da.20674

Nettleton, D. F. (2014). Selection of Variables and Factor Derivation. In *Elsevier eBooks* (pp. 79–104). Elsevier BV. https://doi.org/10.1016/b978-0-12-416602-8.00006-6

O'Brien, C. M. (2016). Statistical Learning with Sparsity: The Lasso and Generalizations. *International Statistical Review*, *84*(1), 156–157. https://doi.org/10.1111/insr.12167

Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and Avoiding Bias in Research. Plastic and Reconstructive Surgery, 126(2), 619–625. https://doi.org/10.1097/prs.0b013e3181de24bc

Penninx, B. W., Benros, M. E., Klein, R. S., & Vinkers, C. H. (2022). How COVID-19 shaped mental health: from infection to pandemic effects. *Nature Medicine*, *28*(10), 2027–2037. https://doi.org/10.1038/s41591-022-02028-2

Richmond, C. (2003, February 6). *Panic Disorder and Panic Attacks*. WebMD. https://www.webmd.com/anxiety-panic/guide/mental-health-panic-disorder

Schneider, A., Hommel, G., & Blettner, M. (2010). Linear Regression Analysis. *Deutsches Arzteblatt International*. https://doi.org/10.3238/arztebl.2010.0776

Sparrevohn, R. M., & Rapee, R. M. (2009). Self-disclosure, emotional expression and intimacy within romantic relationships of people with social phobia. *Behaviour Research and Therapy*, *47*(12), 1074–1078. https://doi.org/10.1016/j.brat.2009.07.016

Specific Phobias. (2003, February 9). WebMD. https://www.webmd.com/anxiety-panic/specific-phobias

Startpagina - Angststoornissen - Richtlijn - Richtlijnendatabase. (n.d.). https://richtlijnendatabase.nl/richtlijn/angststoornissen/angststoornissen_-_startpagina.html

Sukmak, E. V. (2015). Time series forecasting in anxiety disorders of outpatient visits using data mining. *Asia-Pacific Journal of Science and Technology*, *20*(2), 241–253. https://doi.org/10.14456/kkurj.2015.19

United Nations. (n.d.). *The Pandemic Accelerant: How COVID-19 Advanced Our Mental Health Priorities | United Nations*. https://www.un.org/en/un-chronicle/pandemic-accelerant-how-covid-19-advanced-our-mental-health-priorities

Van Der Hoek. (2020, May). *Advies over hervorming van de ggz* [Press release]. https://www.trimbos.nl/docs/69591560-37e2-44b1-b0a1-c19330bd9115.pdf

Vektis - business intelligence centrum voor de zorg | Vektis.nl. (n.d.). Vektis. https://www.vektis.nl/

WebMD Editorial Contributors. (2021, July 13). *Depression vs. Anxiety: Which One Do I Have?* WebMD. https://www.webmd.com/depression/depression-or-anxiety

What are Anxiety Disorders? (n.d.-a). https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders

What are Anxiety Disorders? (n.d.-b). https://www.psychiatry.org/patients-families/anxiety-disorders/what-are-anxiety-disorders

What Is Social Anxiety Disorder or Social Phobia? (2017, March 12). WebMD. https://www.webmd.com/anxiety-panic/guide/mental-health-social-anxiety-disorder

Wittchen, H., & Jacobi, F. (2005). Size and burden of mental disorders in Europe—a critical review and appraisal of 27 studies. European Neuropsychopharmacology, 15(4), 357–376. https://doi.org/10.1016/j.euroneuro.2005.04.012

Wittchen, H., Stein, M. B., & Kessler, R. C. (1999). Social fears and social phobia in a community sample of adolescents and young adults: prevalence, risk factors and co-morbidity. *Psychological Medicine*, *29*(2), 309–323. https://doi.org/10.1017/s0033291798008174

World Health Organization: WHO. (2019). Suicide. *www.who.int*. https://www.who.int/health-topics/suicide

World Health Organization: WHO. (2021). Suicide. *www.who.int*. https://www.who.int/news-room/fact-sheets/detail/suicide

Zhang, W., Yang, X., Lui, S., Meng, Y., Yao, L., Xiao, Y., Deng, W., Zhang, W., & Gong, Q. (2015a). Diagnostic Prediction for Social Anxiety Disorder via Multivariate Pattern Analysis of the Regional Homogeneity. *BioMed Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/763965

Zhang, W., Yang, X., Lui, S., Meng, Y., Yao, L., Xiao, Y., Deng, W., Zhang, W., & Gong, Q. (2015b). Diagnostic Prediction for Social Anxiety Disorder via Multivariate Pattern Analysis of the Regional Homogeneity. *BioMed Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/763965

Zhang, W., Yang, X., Lui, S., Meng, Y., Yao, L., Xiao, Y., Deng, W., Zhang, W., & Gong, Q. (2015c). Diagnostic Prediction for Social Anxiety Disorder via Multivariate Pattern Analysis of the Regional Homogeneity. *BioMed Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/763965

7 Appendices

Table 11 gives explanation of the factor names of the CBS & Police data included for the study (CBS, 2023).

Factor number	Factor name	Explanation
	Age groups	
0	Age group < 5 years	The percentage of people belonging to the age group of younger than 5 years old.
1	Age group 5 - 10 years	The percentage of people belonging to the age group between 5 and 10 years old.
2	Age group 10 - 15 years	The percentage of people belonging to the age group between 10 and 15 years old.
3	Age group 15 - 20 years	The percentage of people belonging to the age group between 15 and 20 years old.
4	Age group 20 - 25 years	The percentage of people belonging to the age group between 20 and 25 years old.
5	Age group 25 - 45 years	The percentage of people belonging to the age group between 25 and 45 years old.
6	Age group 45 - 65 years	The percentage of people belonging to the age group between 45 and 65 years old.
7	Age group 65 - 80 years	The percentage of people belonging to the age group between 65 and 80 years old.
8	Age group > 80 years	The percentage of people belonging to the age group older than 80 years.
9	Total age pressure	Total age pressure: The ratio of the number of people aged 0 to 20 and aged 65 or over to those in the so-called 'productive' age group of 20 to 65. .
10	Green pressure	Green pressure: The ratio of the number of people aged 0 to 20 years to those in the so-called 'productive' age group of 20 to 65 years.
11	Grey pressure	Grey pressure: The ratio of the number of people aged 65 or over to those in the so-called 'productive' age group of 20 to 65.
	Marital status	
12	Unmarried population	The number of people being unmarried (>15 years old) as a percentage of the total population (>15 years old) on January the 1st.
13	Married population	The number of people being married (>15 years old) as a percentage of the total population (>15 years old) on January the 1st.
14	Divorced population	The number of people being divorced (>15 years old) as a percentage of the total population (>15 years old) on January the 1st.
15	Widowed population	The number of people being widowed (>15 years old) as a percentage of the total population (>15 years old) on January the 1st.
	Migration background	
16	Both parents NL-born	Persons whose both parents were born in the Netherlands as a percentage of the total population on 1 January.
17	At least one parent born abroad	Persons whose at least one parent was born abroad as a percentage of the total population as of 1 January.
18	Migration background: Europe, NA, Oceania, Indo, Japan	Persons whose migration background is one of the countries in Europe (excluding Turkey), North America and Oceania, and Indonesia and Japan as a percentage of the total population on 1 January.
19	Migration background: Africa, Latin America, Asia (excl. Indo & Japan), Turkey	Persons whose migration background is one of the countries in Africa, Latin America and Asia (excluding Indonesia and Japan) or Turkey, as a percentage of the total population on 1 January.
20	Migration background: Morocco, Ifni,	Persons whose migration background is Morocco, Ifni, Spanish Sahara or Western Sahara, as a percentage of the total population as of 1 January.

	Spanish Sahara or Western Sahara	
21	Migration background: (Former) Neth. Antilles & Aruba	Persons whose migration background is (former) Netherlands Antilles and Aruba, as a percentage of total population on 1 January.
22	Migration background: Suriname	Persons whose migration background is Suriname, as a percentage of total population as of 1 January.
23	Migration background: Turkey	Persons whose migration background is Turkey, as a percentage of the total population as of 1 January.
24	Non-Western migration background	All persons with a non-Western migration background whose migration background is not equal to Turkey, Morocco, Suriname, the Netherlands Antilles or Aruba, as a percentage of the total population as of 1 January.
	Population development	
25	Population density	Population density: The number of addresses within a circle with a radius of one kilometer around an address, divided by the area of the circle. (expressed in addresses per km²).
26	Live births per inhabitant	The number of live births from 1 January to 31 December of the relevant year per thousand of the average population.
27	Deaths per inhabitant	The number of deaths from 1 January to 31 December of the relevant year in which a competent doctor signed a death certificate, per thousand of the average population.
28	Birth surplus	Birth surplus per thousand of the average population which is the number of live births minus the number of deaths. Also called: natural population growth.
	Reason of death	
29	Deaths due to neoplasm	The number of deaths due to neoplasm as a percentage of the total number of deaths.
30	Deaths due to heart disease	The number of deaths due to heat disease as a percentage of the total number of deaths.
31	Deaths due to respiratory illness	The number of deaths due to respiratory illness as a percentage of the total number of deaths.
32	Deaths due to external causes	The number of deaths due to external causes as a percentage of the total number of deaths.
33	Deaths due to other causes	The number of deaths due to other causes as a percentage of the total number of deaths.
	Immigration and emigration	
34	Domestic migration	Domestic migration rate per thousand of average population. The number of people settled from another municipality within the Netherlands minus the number of people left to another municipality within the Netherlands.
35	Movement mobility	Relocation mobility per thousand of average population. The relocation mobility of a region is calculated as the total of persons moving within municipalities in the region plus the half-sum of persons moving between municipalities persons (settlers plus leavers) in the region.
36	Migration balance	Migration balance per thousand of the average population in the observation year. The migration balance includes persons settling in the Netherlands minus inhabitants leaving the Netherlands to settle outside the Netherlands.
37	Population growth	Population growth per thousand of the initial population on 1 January present in the period in which the relevant changes (births, deaths and the like) occurred.
	Private households	
38	One-person households	One-person households: private households consisting of one person, as a percentage of total private households.
39	Households without children	Multi-person households without children living at home, as a percentage of total private households.
40	Households with child(ren)	Multi-person households with children living at home, as a percentage of total private households. Where a child living at home is defined as a person regardless of age or marital status who has a child-parent relationship with one or two parents belonging to the household. Children living at home include adoptive and stepchildren, but not foster children.
41	Average household size	The number of persons living in private households divided by the number of private households.
	Housing	

42	Balance increase housing	Balance increase dwellings per thousand of dwellings. The balance increase refers to the difference between the number of dwellings added to the stock and the number withdrawn from stock. The calculation of the balance includes administrative corrections.
43	Housing density	Total number of dwellings per km² of land on 1 January.
44	Houses to buy	Percentage of houses owned by the (future) occupant(s) or in use as second homes.
45	Renting houses	Percentage of dwellings that are not occupied by the owner of the dwelling. Properties where no occupant is registered are those where it is likely that the property is destined for the rental market.
46	Ownership unknown	Percentage of houses whose ownership could not be inferred based on various registrations such as the WOZ register, Register of Persons and the housing database 'Kadaster'.
	Education & Employment	
47	Secondary graduates per inhabitant	Ratio of the number of secondary graduates to the population size of a municipality. The number of secondary graduates include VWO, HAVO, VMBO and practical education graduates.
48	MBO graduates per inhabitant	Ratio of the number of MBO graduates to the population size of a municipality.
49	HBO graduates per inhabitant	Ratio of the number of HBO graduates to the population size of a municipality.
50	University graduates per inhabitant	Ratio of the number of university graduates to the population size of a municipality.
51	Jobs per inhabitant	Ratio of the number of jobs to the population size of a municipality. The number of jobs is defined as the average number of jobs in December of employees employed by companies and institutions.
	Net income	
52	Net income: Private households	One or more persons occupying living quarters together and providing for themselves, i.e. non-business, the daily necessities of life (excluding student households).
53	Net income: Single household	Private household consisting of one person. Single-person households, also known as singles, include people who live with others at the same address but run their own household.
54	Net income: Single-parent household	Private household consisting of one parent with one or more children living at home.
55	Net income: Couple with no child	Multi-person household consisting of a couple without child(ren) living at home
56	Net income: Couple with child(ren)	Multiple-person household consisting of a couple with child(ren) living at home
57	Net income: Employee	Households for which worker's pay is the main source of income.
58	Net income: Self-employment	Households for which income as self-employed, director-major shareholder and other self-employed is the main source of income.
59	Net income: Transfer	Households for which benefits, pensions or student loans are the main source of income.
60	Net income: bought home	A household lives in an owner-occupied accommodation.
61	Net income: rental home	A household lives in a rented accommodation.
	Standardized income	**Average standardized income of private households (excluding student households).** Standardized disposable income is disposable income adjusted for differences in household size and composition. This adjustment is made using equivalence factors. The equivalence factor expresses the economies of scale resulting from running a joint household. Using the equivalence factors, all incomes are reduced to the income of a single-person household. In this way, the welfare level of different household types is made comparable. Standardized income is a measure of the welfare of (the members of) a household.
62	Standardized income: Private households	One or more persons occupying living quarters together and providing for themselves, i.e. non-business, the daily necessities of life (excluding student households).

63	Standardized income: Single household	Private household consisting of one person. Single-person households, also known as singles, include people who live with others at the same address but run their own household.
64	Standardized income: Single-parent household	Private household consisting of one parent with one or more children living at home.
65	Standardized income: Couple with no child	Multi-person household consisting of a couple without child(ren) living at home
66	Standardized income: Couple with child(ren)	Multiple-person household consisting of a couple with child(ren) living at home
67	Standardized income: Employee income	Households for which worker's pay is the main source of income.
68	Standardized income: Self-employment income	Households for which income as self-employed, director-major shareholder and other self-employed is the main source of income.
69	Standardized income: Transfer income	Households for which benefits, pensions or student loans are the main source of income.
70	Standardized income: own home	A household lives in an owner-occupied accommodation.
71	Standardized income: rental property	A household lives in a rented accommodation.
	Median wealth	
72	Median wealth: Private households	One or more persons occupying living quarters together and providing for themselves, i.e. non-business, the daily necessities of life (excluding student households).
73	Median wealth: Single-person household	Private household consisting of one person. Single-person households, also known as singles, include people who live with others at the same address but run their own household.
74	Median wealth: Single-parent household	Private household consisting of one parent with one or more children living at home.
75	Median wealth: Couple with no child	Multi-person household consisting of a couple without child(ren) living at home
76	Median wealth: Couple with child(ren)	Multiple-person household consisting of a couple with child(ren) living at home
77	Median wealth: Employee income	Households for which worker's pay is the main source of income.
78	Median wealth: Self-employed income	Households for which income as self-employed, director-major shareholder and other self-employed is the main source of income.
79	Median wealth: Transfer income	Households for which benefits, pensions or student loans are the main source of income.
80	Median wealth: own home	A household lives in an owner-occupied accommodation.
81	Median wealth: rental property	A household lives in a rented accommodation.
	Social Security	
82	Benefits per inhabitant	Ratio of the total number of people receiving benefits to the population size of a municipality. Benefits include the Unemployment Act (WW), Welfare Act (PW), welfare-related law (IOAW, IOAZ, Bbz), Disability Act (WAO, WIA, WAZ, Wajong, Wajong Act) or the General Old Age Pensions Act (AOW).
83	Benefits without AOW per inhabitant	Ratio of the total number of people receiving benefits up to state pension age to the population size of a municipality.
84	Unemployment per inhabitant	Ratio of the total number of people receiving benefits under the Unemployment Act (WW) to the population size of a municipality.
85	Welfare per inhabitant	Ratio of the total number of people receiving benefits under the Assistance Act or assistance-related law to the population size of a municipality.

86	Disabled per inhabitant	Ratio of the total number of people receiving disability benefits to the population size of a municipality.
87	Wajong per inhabitant	Ratio of the total number of people receiving benefits under the Disability Insurance for Young Disabled Persons Act (Wajong) to the population size of a municipality.
88	AOW per inhabitant	Ratio of the total number of people receiving benefits under the General Old Age Pensions Act (AOW) to the population size of a municipality.
	Agriculture	
89	Arable land area	Percentage of utilized agricultural area used for arable farming as a percentage of the total utilized agricultural area.
90	Horticulture open field	Utilized agricultural area for horticultural open land as a percentage of total utilized agricultural area.
91	Horticulture under glass	Utilized agricultural area under glass as a percentage of the total utilized agricultural area.
92	Perennial grassland	Utilized agricultural area under permanent pasture as a percentage of the total utilized agricultural area.
93	Natural grassland	Utilized agricultural area used for natural grassland as a percentage of total utilized agricultural area.
94	Temporary grassland	Utilized agricultural area used for temporary grassland as a percentage of the total utilized agricultural area.
95	Green fodder	Utilized agricultural area used for green fodder as a percentage of the total utilized agricultural area.
96	Nitrogen excretion	The total amount of nitrogen (N) excreted in kilos per hectare of land.
97	Phosphate excretion	The excreted amount of phosphate in kilos per hectare of land.
98	Kali excretion	Amount of potash present in produced manure expressed as K2O.
	Traffic and transport	
99	Total passenger cars	Passenger cars per thousand population.
100	Passenger cars of private individuals	Private cars per thousand population.
101	Motorcycles	Motorbikes per thousand population.
102	Mopeds	Vehicles with Dutch moped license plates per thousand population.
	Proximity to facilities	
103	Distance to GP surgery	The average distance calculated by road of all residents in an area to the nearest GP practice.
104	Number of GP practices within 3 km	The average number of GP practices within 3 km by road for all residents of an area.
105	Distance to GP surgery	The average distance calculated by road of all residents in an area to the nearest GP surgery.
106	Distance to hospital	The average distance calculated by road of all residents in an area to the nearest hospital.
107	Number of hospitals within 20 km	The average number of hospitals within 20 km by road for all residents of an area.
108	Distance to day-care center	The average distance calculated by road from all residents in an area to the nearest nursery.
109	Number of childcare centers within 3 km	The average number of nurseries within 3 km by road for all residents of an area.
110	Distance to primary school	The average distance calculated by road of all residents in an area to the nearest primary school.
111	Number of primary schools within 3 km	The average number of primary schools (establishments) within 3 km by road for all residents of an area.
112	Distance to school VMBO	The average distance calculated by road from all residents in an area to the nearest school for VMBO education.
113	Number of VMBO schools within 5 km	The average number of pre-vocational secondary schools within 5 km by road for all residents of an area.
114	Distance to secondary school	The average distance calculated by road of all residents in an area to the nearest school for upper general secondary education or preparatory scientific education.
115	Number of high schools within 5 km	The average number of schools for upper general secondary education or preparatory scientific education within 5 km by road for all residents of an area.

116	Distance to large supermarket	The average distance calculated by road from all residents in an area to the nearest large supermarket (minimum area of 150 m2).
117	Number of large supermarkets within 3 km	The average number of large supermarkets within 3 km by road for all residents of an area (minimum area of 150 m2).
118	Distance to restaurant	The average distance calculated by road from all residents in an area to the nearest restaurant.
119	Number of restaurants within 3 km	The average number of restaurants within 3 km by road for all residents of an area.
120	Distance to library	The average distance calculated by road of all residents in an area to the nearest library.
121	Distance to cinema	The average distance calculated by road of all residents in an area to the nearest cinema.
122	Number of cinemas within 10 km	The average number of cinemas within 10 km by road for all residents of an area.
123	Distance to swimming pool	The average distance calculated by road of all residents in an area to the nearest swimming pool.
124	Distance to sports ground	The average distance calculated by road from all residents in an area to the nearest sports ground.
125	Distance to public green spaces	The average distance calculated by road of all residents in an area to the nearest public green space.
126	Distance to main road access	The average distance calculated by road of all residents in an area to the nearest driveway of a national or provincial road.
127	Distance to railway station	The average distance calculated by road of all residents in an area to the nearest railway station.
	Environment and land use	
128	Environmental address density	The environmental address density of an area is determined by first determining, for each residence object, the number of residence objects within a circle with a radius of one kilometer around a residence object, divided by the area of the circle.
129	Traffic area percentage	Percentage of land in use for rail, road and air traffic.
130	Built-up area percentage	Percentage of land in use for living, working, shopping, entertainment, culture and public amenities.
131	Semi-built-up area percentage	Percentage of land with some degree of pavement that is not in use as a traffic area or built-up area.
132	Recreational site percentage	Percentage of land intended for recreational use.
133	Agrarian site percentage	Percentage of land zoned for agricultural use.
134	Forest and open natural terrain percentage	Percentage of land in use as forest or open natural terrain.
135	Traffic area per inhabitant	Hectares of land in use for rail, road and air traffic per 1000 inhabitants.
136	Built-up area per inhabitant	Hectares of land in use for living, working, shopping, entertainment, culture and public amenities per 1000 inhabitants.
137	Semi-built-up area per inhabitant	Hectares of land with some degree of pavement that is not in use as a traffic area or built-up area per 1000 inhabitants.
138	Recreational site per inhabitant	Hectares of land intended for recreational use per 1000 inhabitants.
139	Agrarian site per inhabitant	Hectares of land zoned for agricultural use per 1000 inhabitants.
140	Forest and open natural terrain per inhabitant	Hectares of land in use as forest or open natural terrain per 1000 inhabitants.
	Police nuisance reports	Nuisance is a situation, in which there is a hindrance caused by a condition, person, object or the like and is reported to the police.
141	Nuisance reported per inhabitant	Ratio of the total nuisance reports to the population size of a municipality.
142	Nuisance youth reported per inhabitant	Ratio of youth nuisance reports to the population size of a municipality. Youth nuisance is defined as any report complaining about youth.

143	Public intoxication reported per inhabitant	Ratio of public intoxication nuisance reports to the population size of a municipality. Public intoxication is defined as being in an apparent state of intoxication in a public or publicly accessible place.
144	Nuisance from confused individuals per inhabitant	Ratio of confused individual reports to the population size of a municipality. This is defined as any form of nuisance caused by a confused or overworked person.
145	Nuisance from drug usage per inhabitant	Ratio of drug nuisance reports to the population size of a municipality. Drug nuisance includes any nuisance related to both hard and soft drugs
146	Nuisance from homeless reported per inhabitant	Ratio of homeless nuisance reports to the population size of a municipality.

Table 12 presents the factor names of the public data included for the study. The summary of the statistics is given per factor. For more explanation on the meaning of each factor, see Table 11.

Factor number	**Factor name**	**Min.**	**1st. Quartile**	**Median**	**Mean**	**3rd Quartile**	**Max.**	**Number of NAs**
	Age groups							
0	Age group < 5 years	2.500	4.400	4.800	4.876	5.300	10.000	0
1	Age group 5 - 10 years	2.90	4.90	5.30	5.38	5.80	10.10	0
2	Age group 10 - 15 years	3.200	5.400	5.900	5.912	6.400	9.400	0
3	Age group 15 - 20 years	4.000	5.800	6.200	6.204	6.600	9.000	0
4	Age group 20 - 25 years	2.500	4.800	5.200	5.461	5.700	17.400	0
5	Age group 25 - 45 years	13.00	20.50	21.90	22.25	23.60	36.70	0
6	Age group 45 - 65 years	18.90	28.20	29.70	29.39	31.00	35.00	0
7	Age group 65 - 80 years	6.90	14.00	15.70	15.59	17.10	24.10	0
8	Age group > 80 years	1.800	4.200	4.800	4.935	5.500	10.800	0
9	Total age pressure	44.80	71.70	75.70	75.52	79.70	113.00	0
10	Green pressure	23.9	36.3	39.1	39.3	42.0	71.8	0
11	Grey pressure	15.00	31.80	36.00	36.21	40.20	66.50	0
	Marital status							
12	Unmarried population	21.90	30.80	32.70	33.57	34.80	63.50	0
13	Married population	24.90	49.50	52.40	51.44	54.60	66.00	0
14	Divorced population	2.300	7.200	8.400	8.444	9.700	13.600	0
15	Widowed population	3.100	5.900	6.600	6.544	7.200	10.400	0
	Migration background							
16	Both parents NL-born	44.40	80.50	86.30	84.24	90.30	96.80	0
17	At least one parent born abroad	3.20	9.70	13.70	15.76	19.50	55.60	0
18	Migration background: Europe, NA, Oceania, Indo, Japan	1.400	5.900	7.800	8.549	10.100	46.800	0
19	Migration background: Africa, Latin America, Asia (excl. Indo & Japan), Turkey	1.100	3.200	5.200	7.206	9.000	38.900	0
20	Migration background: Morocco, Ifni, Spanish Sahara or Western Sahara	0.00	0.10	0.50	1.08	1.50	9.70	0

21	Migration background: (Former) Neth. Antilles & Aruba	0.0000	0.2000	0.3000	0.4937	0.6000	41.000	0
22	Migration background: Suriname	0.0000	0.2000	0.5000	0.8781	0.8000	115.000	0
23	Migration background: Turkey	0.000	0.200	0.500	1.227	1.600	10.200	0
24	Non-Western migration background	0.800	2.100	2.900	3.527	4.200	18.800	0
	Population development							
25	Population density	23.0	248.0	480.0	915.8	1184.0	6620.0	0
26	Live births per inhabitant	2.700	8.200	9.100	9.227	10.000	20.300	0
27	Deaths per inhabitant	3.600	8.100	9.300	9.369	10.400	25.500	0
28	Birth surplus	-170.000	-19.000	-0.1000	-0.1424	18.000	160.000	0
	Reason of death							
29	Deaths due to neoplasm	0.0000	0.2881	0.3108	0.3132	0.3352	0.6364	0
30	Deaths due to heart disease	0.09091	0.22876	0.25210	0.25249	0.27505	0.60000	0
31	Deaths due to respiratory illness	0.00000	0.06748	0.08055	0.08111	0.09443	0.25000	0
32	Deaths due to external causes	0.00000	0.04188	0.05195	0.05236	0.06188	0.17647	0
33	Deaths due to other causes	0.1190	0.2678	0.2964	0.3008	0.3262	0.5200	0
	Immigration and emigration							
34	Domestic migration	-541.200	-0.900	2.000	1.341	5.700	57.300	0
35	Movement mobility	52.10	80.00	88.60	92.76	100.20	360.80	0
36	Migration balance	-25.600	0.600	1.500	3.904	3.200	522.000	290
37	Population growth	-33.400	0.700	4.300	5.115	8.500	62.600	0
	Private households							
38	One-person households	18.80	28.20	30.70	32.48	35.10	61.70	0
39	Households without children	19.80	30.20	32.40	31.99	34.30	43.00	0
40	Households with child(ren)	17.60	33.20	35.90	35.53	38.30	58.90	0
41	Average household size	1.640	2.180	2.280	2.272	2.370	3.340	0
	Housing							
42	Balance increase housing	-143.100	3.400	6.900	7.956	11.400	87.800	0
43	Housing density	14.0	106.0	211.0	412.4	532.0	3184.0	0
44	Houses to buy	29.40	59.20	65.20	63.94	70.10	88.20	0
45	Renting houses	11.20	29.20	34.00	35.23	40.00	70.10	0
46	Ownership unknown	0.0000	0.2000	0.6000	0.8342	12.000	70.000	0
	Education & Employment							
47	Secondary graduates per inhabitant	0.004224	0.010338	0.011641	0.011493	0.012719	0.018087	0
48	MBO graduates per inhabitant	0.002602	0.007952	0.009349	0.009276	0.010705	0.019685	0
49	HBO graduates per inhabitant	0.000000	0.002679	0.003104	0.003293	0.003583	0.014818	0
50	University graduates per inhabitant	0.0000000	0.0004734	0.0006996	0.0011124	0.0009369	0.0333490	0
51	Jobs per inhabitant	0.1334	0.2956	0.3737	0.4048	0.4965	13.222	0

	Net income							
52	Net income: Private households	31.60	41.30	44.80	45.76	49.00	109.50	0
53	Net income: Single household	20.10	23.50	25.00	25.44	26.80	53.80	1
54	Net income: Single-parent household	27.20	34.20	36.80	37.32	39.90	83.60	36
55	Net income: Couple with no child	34.90	43.30	46.40	47.51	50.20	107.30	0
56	Net income: Couple with child(ren)	45.8	59.8	64.7	66.7	71.1	181.8	11
57	Net income: Employee	32.90	47.00	50.30	51.03	54.20	94.30	0
58	Net income: Self-employment	45.20	61.95	69.00	70.92	77.70	240.90	14
59	Net income: Transfer	22.50	29.10	31.40	32.32	34.30	98.90	0
60	Net income: bought home	38.60	50.30	54.20	55.47	58.70	141.10	0
61	Net income: rental home	21.50	25.50	26.90	27.06	28.50	37.10	7
	Standardized income							
62	Standardized income: Private households	23.20	28.60	30.60	31.22	33.10	71.50	0
63	Standardized income: Single household	20.20	23.60	25.00	25.47	26.80	53.80	1
64	Standardized income: Single-parent household	18.90	23.30	24.90	25.27	26.90	55.00	36
65	Standardized income: Couple with no child	25.60	31.40	33.60	34.35	36.10	76.80	0
66	Standardized income: Couple with child(ren)	24.80	31.90	34.30	35.28	37.40	93.90	11
67	Standardized income: Employee income	24.60	30.60	32.30	32.87	34.50	56.20	0
68	Standardized income: Self-employment income	28.70	38.70	42.70	44.02	47.60	155.30	15
69	Standardized income: Transfer income	18.60	23.30	25.00	25.71	27.20	70.70	0
70	Standardized income: own home	26.20	33.30	35.60	36.51	38.50	90.90	0
71	Standardized income: rental property	17.20	20.00	21.00	21.13	22.10	29.00	7
	Median wealth							
72	Median wealth: Private households	1.60	40.80	81.00	90.12	125.60	454.10	0
73	Median wealth: Single-person household	1.40	13.90	24.10	33.39	40.60	393.50	0
74	Median wealth: Single-parent household	0.00	5.70	14.30	22.99	29.10	243.80	32
75	Median wealth: Couple with no child	12.9	108.6	158.2	163.2	208.7	686.2	0
76	Median wealth: Couple with child(ren)	0.00	57.23	93.25	101.45	134.18	591.90	11
77	Median wealth: Employee income	0.00	27.10	52.50	58.78	82.10	312.20	0
78	Median wealth: Self-employed income	17.8	144.3	205.9	212.2	269.8	906.7	14
79	Median wealth: Transfer income	1.8	51.0	119.5	125.9	185.0	531.3	0
80	Median wealth: own home	17.0	134.9	177.4	184.7	224.2	719.8	0
81	Median wealth: rental property	0.600	3.000	4.700	4.843	6.400	14.700	6
	Social Security							

82	Benefits per inhabitant	0.1227	0.2520	0.2740	0.2753	0.2978	0.3794	0
83	Benefits without AOW per inhabitant	0.02418	0.06160	0.07291	0.07658	0.08845	0.16537	0
84	Unemployment per inhabitant	0.002407	0.012460	0.015464	0.016180	0.019489	0.041783	0
85	Welfare per inhabitant	0.002723	0.013023	0.017209	0.020782	0.025856	0.078498	0
86	Disabled per inhabitant	0.01209	0.03368	0.03909	0.04077	0.04573	0.08453	0
87	Wajong per inhabitant	0.00000	0.00869	0.01078	0.01200	0.01421	0.03574	0
88	AOW per inhabitant	0.08729	0.17795	0.19895	0.19865	0.21821	0.30764	0
	Agriculture							
89	Arable land area	0.00	2.40	9.10	19.44	31.50	85.50	0
90	Horticulture open field	0.000	0.500	2.300	6.806	7.400	92.800	0
91	Horticulture under glass	0.000	0.000	0.000	1.642	0.400	100.000	0
92	Perennial grassland	0.00	16.60	42.30	42.85	65.50	100.00	0
93	Natural grassland	0.000	1.200	2.900	5.141	6.000	84.400	0
94	Temporary grassland	0.00	4.60	9.60	11.26	16.50	87.90	0
95	Green fodder	0.00	4.10	10.40	11.59	17.90	46.60	0
96	Nitrogen excretion	0.0	132.0	248.0	286.1	344.0	1692.0	0
97	Phosphate excretion	0.00	41.00	77.00	96.78	111.00	887.00	0
98	Kali excretion	0.0	175.2	317.0	315.6	405.8	1693.0	935
	Traffic and transport							
99	Total passenger cars	269.0	476.0	509.0	508.6	540.0	1299.0	0
100	Passenger cars of private individuals	235.0	439.0	478.0	470.8	510.0	585.0	0
101	Motorcycles	14.00	37.00	44.00	44.16	51.00	81.00	0
102	Mopeds	37.00	57.00	67.00	69.55	78.00	169.00	0
	Proximity to facilities							
103	Distance to GP surgery	0.400	0.800	1.000	1.125	1.300	2.800	0
104	Number of GP practices within 3 km	0.600	1.800	3.300	4.773	6.400	38.600	0
105	Distance to GP surgery	1.700	4.400	7.650	8.542	11.100	63.900	317
106	Distance to hospital	1.500	4.600	8.400	9.555	12.800	63.900	0
107	Number of hospitals within 20 km	0.000	1.300	2.100	3.511	4.600	16.800	0
108	Distance to day-care center	0.3000	0.6000	0.8000	0.8809	10.000	227.000	0
109	Number of childcare centers within 3 km	0.000	3.300	5.900	9.194	12.000	91.900	0
110	Distance to primary school	0.4000	0.6000	0.7000	0.7533	0.8000	15.000	0
111	Number of primary schools within 3 km	0.900	3.400	5.900	6.918	9.300	36.100	0
112	Distance to school VMBO	0.400	1.800	2.700	3.338	4.300	14.800	0
113	Number of VMBO schools within 5 km	0.000	0.900	1.900	2.842	3.800	20.400	0
114	Distance to secondary school	1.10	2.20	3.40	4.66	6.00	37.30	0
115	Number of high schools within 5 km	0.000	0.400	1.100	1.844	2.400	17.400	0
116	Distance to large supermarket	0.400	0.800	0.900	1.022	1.200	2.400	0

117	Number of large supermarkets within 3 km	0.800	2.600	4.400	5.722	7.300	40.600	0
118	Distance to restaurant	0.2000	0.7000	0.9000	0.9051	11.000	23.000	0
119	Number of restaurants within 3 km	1.80	7.20	12.70	22.47	25.20	460.60	0
120	Distance to library	0.600	1.400	1.800	2.063	2.300	14.600	0
121	Distance to cinema	0.600	3.700	7.100	8.319	11.300	56.700	0
122	Number of cinemas within 10 km	0.000	0.300	1.000	1.733	2.300	13.600	0
123	Distance to swimming pool	0.900	2.300	3.200	3.875	4.500	34.800	0
124	Distance to sports ground	0.6000	0.8000	0.9000	0.9759	11.000	21.000	1251
125	Distance to public green spaces	0.1000	0.4000	0.5000	0.5138	0.6000	17.000	1251
126	Distance to main road access	0.4	1.2	1.5	1.9	1.9	37.7	0
127	Distance to railway station	1.000	2.500	4.400	6.852	8.400	50.900	0
	Environment and land use							
128	Environmental address density	190	621	996	1183	1552	6074	0
129	Traffic area percentage	0.300	2.600	3.600	4.113	5.100	13.500	1251
130	Built-up area percentage	0.90	6.70	11.90	18.16	25.68	70.20	1251
131	Semi-built-up area percentage	0.100	0.500	1.100	1.786	2.300	12.500	1251
132	Recreational site percentage	0.500	2.000	3.250	4.955	6.600	30.200	1251
133	Agrarian site percentage	0.10	38.85	60.80	55.43	74.78	92.10	1251
134	Forest and open natural terrain percentage	0.20	3.70	10.70	15.56	22.45	96.50	1251
135	Traffic area per inhabitant	1.00	4.00	7.00	8.87	11.00	51.00	1251
136	Built-up area per inhabitant	9.00	19.00	24.00	24.38	28.00	53.00	1251
137	Semi-built-up area per inhabitant	0.000	1.000	2.000	2.846	3.000	23.000	1251
138	Recreational site per inhabitant	2.000	4.000	6.000	8.611	9.000	83.000	1251
139	Agrarian site per inhabitant	0	35	115	190	282	1039	1251
	Forest and open natural terrain per inhabitant	0.00	4.00	17.00	82.03	68.00	4085.00	1251
140	**Police nuisance reports**							
141	Nuisance reported per inhabitant	0.002833	0.011197	0.015521	0.017563	0.021931	0.062673	0
142	Nuisance youth reported per inhabitant	0.0002258	0.0027285	0.0041671	0.0047700	0.0060778	0.0258065	0
143	Public intoxication reported per inhabitant	0.0000000	0.0001081	0.0002048	0.0003018	0.0003830	0.0064516	0
144	Nuisance from confused individuals per inhabitant	0.000000	0.002444	0.003679	0.004200	0.005350	0.026032	0
145	Nuisance from drug usage per inhabitant	0.0000000	0.0006428	0.0010374	0.0014975	0.0017974	0.0118648	0
146	Nuisance from homeless reported per inhabitant	0.0000000	0.0001202	0.0002792	0.0005526	0.0005930	0.0057513	0

Table 13 presents a list of the input factors which have > 0.4 Pearson correlation with the proportion of inhabitants with anxiety disorders. The numbers correspond to the numbers listed in Table 1.

No.	Factor name	Correlation
1	patienten_per_inwoner	1.00
	Most positive correlators (decreasing order)	
2	bijstand_per_inwoner	0.66
3	Milieu en bodemgebruik|Bodemgebruik|Omgevingsadressendichtheid	0.61
4	Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond. relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Overig niet-westerse migratieachtergrond	0.61
5	Nabijheid voorzieningen|Onderwijs|Aantal scholen vmbo binnen 5 km	0.60
6	Nabijheid voorzieningen|Onderwijs|Aantal scholen havo/vwo binnen 5 km	0.59
7	Bouwen en wonen|Woningvoorraad|Woningen naar eigendom|Huurwoningen	0.59
8	Nabijheid voorzieningen|Kinderopvang|Aantal kinderdagverblijven binnen 3 km	0.59
9	Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Eenpersoonshuishoudens	0.59
10	Nabijheid voorzieningen|Detailhandel|Aantal grote supermarkten binnen 3 km	0.59
11	Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Gescheiden	0.57
12	Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Totaal niet-westerse migratieachtergrond	0.57
13	overlast_verward_per_inwoner	0.55
14	Nabijheid voorzieningen|Gezondheid|Aantal huisartsenpraktijken binnen 3 km	0.55
15	Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Totaal met migratieachtergrond	0.55
16	Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Ongehuwd	0.54
17	overlast_per_inwoner	0.54
18	Nabijheid voorzieningen|Onderwijs|Aantal basisonderwijsscholen binnen 3 km	0.53
19	overlast_zwervers_per_inwoner	0.52
20	Nabijheid voorzieningen|Horeca|Aantal restaurants binnen 3 km	0.51
21	uitkeringen_zonder_AOW_per_inwoner	0.51
22	Bouwen en wonen|Woningvoorraad|Woningdichtheid	0.49
23	Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Bebouwd terrein	0.48
24	Bevolking|Bevolkingssamenstelling op 1 januari|Bevolkingsdichtheid	0.47
25	Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Recreatieterrein	0.47
26	Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|(voormalige) Nederlandse Antillen, Aruba	0.46
27	Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Verkeersterrein	0.43
	Most negative correlators (increasing negatively)	
28	Inkomen en vermogen|Mediaan vermogen huishoudens|Particuliere huishoudens excl. studenten	-0.40
29	Verkeer en vervoer|Motorvoertuigen|Personenauto's particulieren, relatief	-0.41
30	Verkeer en vervoer|Motorvoertuigen|Motorfietsen, relatief	-0.41
31	Inkomen en vermogen|Mediaan vermogen huishoudens|Bron: Inkomen als zelfstandige	-0.41
32	Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Huishoudens met kinderen	-0.42
33	Nabijheid voorzieningen|Gezondheid|Afstand tot ziekenhuis	-0.42
34	Inkomen en vermogen|Mediaan vermogen huishoudens|Bron: Overdrachtsinkomen	-0.47
35	Inkomen en vermogen|Mediaan vermogen huishoudens|Woningbezit: huurwoning	-0.50
36	Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Huishoudens zonder kinderen	-0.51
37	Bevolking|Particuliere huishoudens|Gemiddelde huishoudensgrootte	-0.54
38	Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Nederlandse achtergrond	-0.55
39	Bouwen en wonen|Woningvoorraad|Woningen naar eigendom|Koopwoningen	-0.58
40	Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Gehuwd	-0.66

Table 14 presents the Pearson correlation between each factor of CBS and Police with the number of anxiety disorders per municipality per year.

Factors	**The proportion of inhabitants with anxiety disorders**
patienten_per_inwoner	1
bijstand_per_inwoner	0.661978478
Milieu en bodemgebruik|Bodemgebruik|Omgevingsadressendichtheid	0.611957681
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Overig niet-westerse migratieachtergrond	0.608283703
Nabijheid voorzieningen|Onderwijs|Aantal scholen vmbo binnen 5 km	0.600197382
Nabijheid voorzieningen|Onderwijs|Aantal scholen havo/vwo binnen 5 km	0.593095508
Bouwen en wonen|Woningvoorraad|Woningen naar eigendom|Huurwoningen	0.592313239
Nabijheid voorzieningen|Kinderopvang|Aantal kinderdagverblijven binnen 3 km	0.591117585
Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Eenpersoonshuishoudens	0.591083844
Nabijheid voorzieningen|Detailhandel|Aantal grote supermarkten binnen 3 km	0.586849222
Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Gescheiden	0.570854276
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Totaal niet-westerse migratieachtergrond	0.567336851
overlast_verward_per_inwoner	0.551498597
Nabijheid voorzieningen|Gezondheid|Aantal huisartsenpraktijken binnen 3 km	0.54880943
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Totaal met migratieachtergrond	0.548692248
Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Ongehuwd	0.543335117
overlast_per_inwoner	0.538373571
Nabijheid voorzieningen|Onderwijs|Aantal basisonderwijsscholen binnen 3 km	0.527225538
overlast_zwervers_per_inwoner	0.517627891
Nabijheid voorzieningen|Horeca|Aantal restaurants binnen 3 km	0.510477871
uitkeringen_zonder_AOW_per_inwoner	0.509083492
Bouwen en wonen|Woningvoorraad|Woningdichtheid	0.48961662
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Bebouwd terrein	0.480396397
Bevolking|Bevolkingssamenstelling op 1 januari|Bevolkingsdichtheid	0.472378348
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Recreatieterrein	0.468813256
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|(voormalige) Nederlandse Antillen, Aruba	0.456320336
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Verkeersterrein	0.42709437
kosten_per_patient	0.410929119
Bevolking|Bevolkingsontwikkeling|Verhuizingen|Verhuismobiliteit, relatief	0.398884078
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Suriname	0.39794658
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|25 tot 45 jaar	0.395564114
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Marokko	0.393558103
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Niet-westerse migratieachtergrond|Turkije	0.386334987
Nabijheid voorzieningen|Vrije tijd en cultuur|Aantal bioscopen binnen 10 km	0.37135502
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Semi-bebouwd terrein	0.367429677

overlast_middelen_per_inwoner	0.356487144
wajong_per_inwoner	0.32955764
arbeidsongeschikt_per_inwoner	0.318696176
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Met migratieachtergrond|Westerse migratieachtergrond	0.313930793
banen_per_inwoner	0.299817188
jaar	0.279093455
overlast_jeugd_per_inwoner	0.271503833
uni_gediplomeerd_per_inwoner	0.267551297
overige_oorzaak_per_sterfte	0.267382061
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|20 tot 25 jaar	0.266513383
behandelminuten_per_patient	0.247451444
behandeluren_per_patient	0.247451444
Nabijheid voorzieningen|Gezondheid|Aantal ziekenhuizen binnen 20 km	0.243442773
openbaar_alcohol_per_inwoner	0.20514733
hbo_gediplomeerd_per_inwoner	0.19181014
Landbouw|Oppervlakte cultuurgrond|Blijvend grasland	0.179330717
Landbouw|Oppervlakte cultuurgrond|Natuurlijk grasland	0.151173252
uitkeringen_per_inwoner	0.142934705
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Woningbezit: eigen woning	0.123249732
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Woningbezit: huurwoning	0.108886945
uitwendige_oorzaak_per_sterfte	0.097824413
Bevolking|Bevolkingsontwikkeling|Geboorte en sterfte|Sterfte, relatief	0.083929586
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Type: Paar, zonder kind	0.083251423
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Type: Paar, zonder kind	0.064475189
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Bron: Inkomen als werknemer	0.060090367
Bevolking|Bevolkingsontwikkeling|Geboorte en sterfte|Geboorte, relatief	0.050569846
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Woningbezit: eigen woning	0.048848059
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Type: Paar, met kind(eren)	0.034891582
Landbouw|Oppervlakte cultuurgrond|Tuinbouw onder glas	0.033489388
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|Jonger dan 5 jaar	0.024071802
Bevolking|Bevolkingsontwikkeling|Immigratie en emigratie|Migratiesaldo, relatief	0.022237142
werkloosheid_per_inwoner	0.018969787
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Type: Paar, met kind(eren)	0.017124869
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|80 jaar of ouder	0.015310775
Landbouw|Oppervlakte cultuurgrond|Akkerbouw	0.004296067
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Woningbezit: huurwoning	-0.002672185
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Bron: Inkomen als zelfstandige	-0.014221188
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Type: Eenpersoonshuishouden	-0.017676895
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Type: Eenpersoonshuishouden	-0.018913505
ademhalingsziekte_per_sterfte	-0.021874569
Bevolking|Bevolkingsontwikkeling|Geboorte en sterfte|Geboorteoverschot, relatief	-0.023067264
Bouwen en wonen|Woningvoorraad|Saldo vermeerdering woningen, relatief	-0.024215633

Bevolking|Bevolkingsontwikkeling|Verhuizingen|Binnenlands migratiesaldo, relatief	-0.026586607
Bevolking|Bevolkingsontwikkeling|Bevolkingsgroei|Bevolkingsgroei, relatief	-0.040957797
Landbouw|Oppervlakte cultuurgrond|Tuinbouw open grond	-0.046450479
gemeente_id	-0.057987187
Nabijheid voorzieningen|Vrije tijd en cultuur|Afstand tot bibliotheek	-0.071589973
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Particuliere huishoudens excl. studenten	-0.077065231
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Bron: Inkomen als werknemer	-0.083746413
Verkeer en vervoer|Motorvoertuigen|Voertuigen met bromfietskenteken (%)	-0.103009827
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Bron: Inkomen als zelfstandige	-0.109837969
Nabijheid voorzieningen|Vrije tijd en cultuur|Afstand tot sportterrein	-0.112247705
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|5 tot 10 jaar	-0.112998502
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Type: Eenoudergezin	-0.135663181
Bouwen en wonen|Woningvoorraad|Woningen naar eigendom|Eigendom onbekend	-0.137286006
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Semi-bebouwd terrein	-0.140631292
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Bos en open natuurlijk terrein	-0.143077054
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld gestandaardiseerd inkomen|Bron: Overdrachtsinkomen	-0.145195513
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Demografische druk|Grijze druk	-0.155073715
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Type: Eenoudergezin	-0.15512438
hartziekte_per_sterfte	-0.159694895
Nabijheid voorzieningen|Verkeer en vervoer|Afstand tot oprit hoofdverkeersweg	-0.160054765
Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Verweduwd	-0.160922893
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|65 tot 80 jaar	-0.16946137
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Particuliere huishoudens excl. studenten	-0.17127908
AOW_per_inwoner	-0.175194041
Inkomen en vermogen|Inkomen van particuliere huishoudens|Gemiddeld besteedbaar inkomen|Bron: Overdrachtsinkomen	-0.175796781
Nabijheid voorzieningen|Horeca|Afstand tot restaurant	-0.187161683
Verkeer en vervoer|Motorvoertuigen|Personenauto's, relatief	-0.200454179
mbo_gediplomeerd_per_inwoner	-0.210482351
neoplasma_per_sterfte	-0.211025857
Landbouw|Mineralenuitscheiding|Kali-uitscheiding	-0.22395736
Inkomen en vermogen|Mediaan vermogen huishoudens|Woningbezit: eigen woning	-0.225129054
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|15 tot 20 jaar	-0.226641598
Landbouw|Mineralenuitscheiding|Stikstofuitscheiding	-0.244553096
Landbouw|Oppervlakte cultuurgrond|Groenvoedergewassen	-0.250889673
Landbouw|Mineralenuitscheiding|Fosfaatuitscheiding	-0.261488606
Nabijheid voorzieningen|Vrije tijd en cultuur|Afstand tot zwembad	-0.264019586
Landbouw|Oppervlakte cultuurgrond|Tijdelijk grasland	-0.26695481
Inkomen en vermogen|Mediaan vermogen huishoudens|Type: Paar, met kind(eren)	-0.267041951
Nabijheid voorzieningen|Kinderopvang|Afstand tot kinderdagverblijf	-0.268897686
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|45 tot 65 jaar	-0.278760068
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Demografische druk|Groene druk	-0.279701124
Inkomen en vermogen|Mediaan vermogen huishoudens|Bron: Inkomen als werknemer	-0.286164687
Nabijheid voorzieningen|Onderwijs|Afstand tot school basisonderwijs	-0.287130384

Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Recreatieterrein	-0.298955245
Nabijheid voorzieningen|Gezondheid|Afstand tot huisartsenpraktijk	-0.305372648
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Bos en open natuurlijk terrein	-0.305549463
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Demografische druk|Totale druk	-0.308259971
Inkomen en vermogen|Mediaan vermogen huishoudens|Type: Eenoudergezin	-0.309963897
Nabijheid voorzieningen|Detailhandel|Afstand tot grote supermarkt	-0.3161478
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Bebouwd terrein	-0.316170165
Nabijheid voorzieningen|Onderwijs|Afstand tot school vmbo	-0.319095743
Nabijheid voorzieningen|Verkeer en vervoer|Afstand tot treinstation	-0.338729212
Inkomen en vermogen|Mediaan vermogen huishoudens|Type: Eenpersoonshuishouden	-0.342588819
Inkomen en vermogen|Mediaan vermogen huishoudens|Type: Paar, zonder kind	-0.343344443
Nabijheid voorzieningen|Vrije tijd en cultuur|Afstand tot bioscoop	-0.345717229
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Verkeersterrein	-0.347315033
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Per inwoner|Agrarisch terrein	-0.364369004
Bevolking|Bevolkingssamenstelling op 1 januari|Leeftijd|Leeftijdsgroepen, relatief|10 tot 15 jaar	-0.366832164
Milieu en bodemgebruik|Bodemgebruik|Naar functie|Percentages|Agrarisch terrein	-0.367537079
middelbaar_gediplomeerd_per_inwoner	-0.380294043
Nabijheid voorzieningen|Gezondheid|Afstand tot huisartsenpost	-0.382196646
Nabijheid voorzieningen|Onderwijs|Afstand tot school havo/vwo	-0.385794885
Nabijheid voorzieningen|Groenvoorzieningen|Afstand tot openbaar groen	-0.38988121
Inkomen en vermogen|Mediaan vermogen huishoudens|Particuliere huishoudens excl. studenten	-0.400016346
Verkeer en vervoer|Motorvoertuigen|Personenauto's particulieren, relatief	-0.405790586
Verkeer en vervoer|Motorvoertuigen|Motorfietsen, relatief	-0.411892697
Inkomen en vermogen|Mediaan vermogen huishoudens|Bron: Inkomen als zelfstandige	-0.414872806
Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Huishoudens met kinderen	-0.417695505
Nabijheid voorzieningen|Gezondheid|Afstand tot ziekenhuis	-0.422990732
Inkomen en vermogen|Mediaan vermogen huishoudens|Bron: Overdrachtsinkomen	-0.465746142
Inkomen en vermogen|Mediaan vermogen huishoudens|Woningbezit: huurwoning	-0.496583472
Bevolking|Particuliere huishoudens|Particuliere huishoudens, relatief|Huishoudens zonder kinderen	-0.51219195
Bevolking|Particuliere huishoudens|Gemiddelde huishoudensgrootte	-0.543438921
Bevolking|Bevolkingssamenstelling op 1 januari|Migratieachtergrond|Migratieachtergrond, relatief|Nederlandse achtergrond	-0.548692248
Bouwen en wonen|Woningvoorraad|Woningen naar eigendom|Koopwoningen	-0.575769125
Bevolking|Bevolkingssamenstelling op 1 januari|Burgerlijke staat|Bevolking 15 jaar of ouder|Gehuwd	-0.656901744

Table 15 presents the positive coefficients in descending order of magnitude based on Lasso regression which describe the strength and direction between each input factor and the output factor: the proportion of inhabitants with anxiety disorders.

Factor	**Coefficient**
Income and wealth|Income of private households|Median disposable income|Type: Couple, without child	0.75
Income and wealth|Median household wealth|Source: Employee income	0.68
Population|Population composition on 1 January|Population density	0.60
Income and wealth|Median household wealth|Home ownership: own home	0.49
Income and wealth|Income of private households|Average standardised income|Source: Self-employment income	0.33
Population|Population composition on 1 January|Migration background|Migration background, relative|Dutch background	0.31
Population|Population composition on 1 January|Age|Demographic pressure|Green pressure	0.29

Population|Population composition on 1 January|Age|Age groups, relative|80 years or older	0.27
Population|Population composition on 1 January|Age|Demographic pressure|Total pressure	0.25
Proximity of facilities|Education|Number of secondary schools within 5 km	0.24
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|Suriname	0.21
Income and wealth|Median household wealth|Source: Transfer Income	0.20
nuisance_confused_per_inhabitant	0.19
Population|Population composition on 1 January|Age|Age groups, relative|45 to 65 years old	0.15
Assistance_per_inhabitant	0.15
Income and wealth|Income of private households|Average disposable income|Home ownership: rental property	0.13
nuisance_youth_per_inhabitant	0.13
benefits_without_AOW_per_inhabitant	0.12
nuisance_resources_per_inhabitant	0.11
Proximity to amenities|Leisure and culture|Distance to cinema	0.10
Income and assets|Median household assets|Type: Single-parent households	0.10
Population|Population trends|Movements|Movement mobility, relative	0.09
mbo_graduate_per_inhabitant	0.09
Proximity of facilities|Horeca|Distance to restaurant	0.09
Environment and land use|Soil use|Environmental address density	0.08
Population|Population growth|Population growth, relative	0.08
Proximity to amenities|Education|Distance to grammar school	0.08
Proximity of facilities|Childcare|Number of childcare centres within 3 km	0.08
Building and housing|Housing stock|Houses by ownership|Houses to buy	0.07
Income and wealth|Income of private households|Average standardised income|Type: Single-parent households	0.06
Population|Private households|Private households, relative|Households without children	0.05
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|Other non-Western migration background	0.05
Proximity to facilities|Education|Distance to secondary school	0.04
Proximity to amenities|Health|Number of hospitals within 20 km	0.04
Agriculture|Area of utilized agricultural area|Perennial grassland	0.04
Proximity of facilities|Education|Number of secondary schools within 5 km	0.04
Agriculture|Area under cultivation|Tillage	0.03
Building and living|housing stock|housing growth, relative	0.03
jobs_per_inhabitant	0.03
Population|Population composition on 1 January|Age|Age groups, relative|25 to 45 years old	0.02
respiratory disease_per_death	0.02
Population|Population composition on 1 January|Civil status|Population 15 years or older|Unmarried	0.02
heart disease_per_mortality	0.02
Population|Population composition on 1 January|Civil status|Population 15 years or older|Unmarried	0.02
Proximity to amenities|Retail|Distance to large supermarket	0.02
Nuisance_vagrants_per_inhabitant	0.01
Agriculture|Area of utilized agricultural area|Natural grassland	0.01

Table 16 presents the negative coefficients in descending order of magnitude based on Lasso regression which describe the strength and direction between each input factor and the output factor: the proportion of inhabitants with anxiety disorders.

Factor	**Coefficient**
Income and wealth|Median household wealth|Private households excluding students	-1.06
Population|Private households|Average household size	-0.78
Building and housing|Housing stock|Housing density	-0.65
Income and wealth|Income of private households|Average standardised income|Private households excluding students	-0.58
Proximity of facilities|Childcare|Distance to day-care centre	-0.39
Income and wealth|Income of private households|Average standardised income|Type: Couple, without child	-0.25
Income and wealth|Median household wealth|Source: Self-employment income	-0.25
Population|Population composition on 1 January|Civil status|Population 15 years or older|Married	-0.24
Income and wealth|Income of private households|Average disposable income|Source: Self-employed income	-0.20
AOW_per_inhabitant	-0.20
Income and wealth|Median household wealth|Home ownership: rental property	-0.16

Nuisance_per_inhabitant	-0.15
Proximity to amenities|Health|Distance to hospital	-0.15
high school graduate_per_inhabitant	-0.14
Population|Population composition on 1 January|Age|Age groups, relative|Younger than 5 years old	-0.13
Population|population evolution|Movements|domestic migration, relative	-0.12
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|(former) Netherlands Antilles, Aruba	-0.11
benefits_per_inhabitant	-0.11
Proximity of facilities|Education|Number of primary schools within 3 km	-0.10
uni_graduate_per_inhabitant	-0.09
Proximity of facilities|Leisure and culture|Number of cinemas within 10 km	-0.08
Proximity of facilities|Education|Distance to primary school	-0.07
Proximity of facilities|Traffic and transport|Distance to railway station	-0.06
Proximity of facilities|Hospitality|Number of restaurants within 3 km	-0.06
Population|Population composition on 1 January|Age|Age groups, relative|15 to 20 years old	-0.06
Income and wealth|Income of private households|Average standardised income|Home ownership: rental property	-0.06
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|Morocco	-0.05
Population|population evolution|Births and deaths|Births, relative	-0.05
public_alcohol_per_inhabitant	-0.05
Agriculture|Area of utilized agricultural area|Temporary pasture	-0.04
neoplasm_per_death	-0.04
Traffic and transport|Motor vehicles|Vehicles with moped license plates (%)	-0.04
Income and wealth|Median household wealth|Type: Single household	-0.04
Traffic and transport|Motor vehicles|Passenger cars, relative	-0.03
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Western migration background	-0.03
Proximity of facilities|Health|Distance to GP surgery	-0.03
Agriculture|Minerals excretion|Nitrogen excretion	-0.02
Traffic and transport|Motor vehicles|Motorcycles, relative	-0.02
Population|Population composition on 1 January|Age|Age groups, relative|10 to 15 years old	-0.02
college-educated_per_inhabitant	-0.02
Building and living|Housing stock|Houses by ownership|Ownership unknown	-0.02
Proximity to facilities|Free time and culture|Distance to library	-0.02
Agriculture|Area of cultivated land|Horticulture open land	-0.01
Proximity to facilities|Health|Number of GP practices within 3 km	-0.01

Table 17 presents the coefficients with 0.00 as magnitude based on Lasso regression which describe the strength and direction between each input factor and the output factor: the proportion of inhabitants with anxiety disorders.

Factor	**Coefficient**
Population|Population composition on 1 January|Age|Age groups, relative|20 to 25 years old	0.00
incapacitated_per_inhabitant	0.00
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Total with migration background	0.00
Unemployment_per_inhabitant	0.00
Proximity to amenities|Leisure and culture|Distance to swimming pool	0.00
Population|Population composition on 1 January|Civil status|Population 15 years or older|Married	0.00
wajong_per_inhabitant	0.00
Population|Population composition on 1 January|Age|Demographic pressure|Grey pressure	0.00
Proximity of facilities|Traffic and transport|Distance to main road access point	0.00
Income and wealth|Median household wealth|Type: Couple, with child(ren)	0.00

Traffic and transport|Motor vehicles|Private cars, relative	0.00
Population|Population composition on 1 January|Age|Age groups, relative|65 to 80 years old	0.00
Agriculture|Area of utilized agricultural area|Horticulture under glass	0.00
Agriculture|Area of utilized agricultural area|Animal forage crops	0.00
Population|Population composition on 1 January|Age|Age groups, relative|5 to 10 years	0.00
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|Total non-Western migration background	0.00
Agriculture|Minerals excretion|Phosphates excretion	0.00
Income and wealth|Median household wealth|Type: Couple, no child	0.00
Income and wealth|Income of private households|Median disposable income|Home ownership: own home	0.00
Building and housing|Housing stock|Houses by ownership|Rentals	0.00
Income and wealth|Income of private households|Average disposable income|Private households excluding students	0.00
Income and wealth|Income of private households|Average disposable income|Type: Single household	0.00
Population|Private households|Private households, relative|Households with children	0.00
Income and wealth|Income of private households|Average disposable income|Type: Couple, with child(ren)	0.00
Population|Private households|Private households, relative|One-person households	0.00
Income and wealth|Income of private households|Average disposable income|Source: Employee income	0.00
Income and wealth|Income of private households|Average disposable income|Source: Transfer income	0.00
other_cause_per_mortality	0.00
Income and wealth|Income of private households|Average standardised income|Home ownership: own home	0.00
external_cause_per_mortality	0.00
Income and wealth|Income of private households|Average standardised income|Type: single household	0.00
Income and wealth|Income of private households|Average standardised income|Type: Couple, with child(ren)	0.00
Population|population trends|Births and deaths|Birth surplus, relative	0.00
Population|population evolution|Birth and mortality|Death of birth, relative	0.00
Income and wealth|Income of private households|Average standardised income|Source: Employee income	0.00
Income and wealth|Income of private households|Average standardised income|Source: Transfer income	0.00
Proximity to amenities|Retail|Number of large supermarkets within 3 km	0.00
Population|Population composition on 1 January|Migration background|Migration background, relative|With migration background|Non-Western migration background|Turkey	0.00
Income and wealth|Income of private households|Average disposable income|Type: Single-parent household	0.00

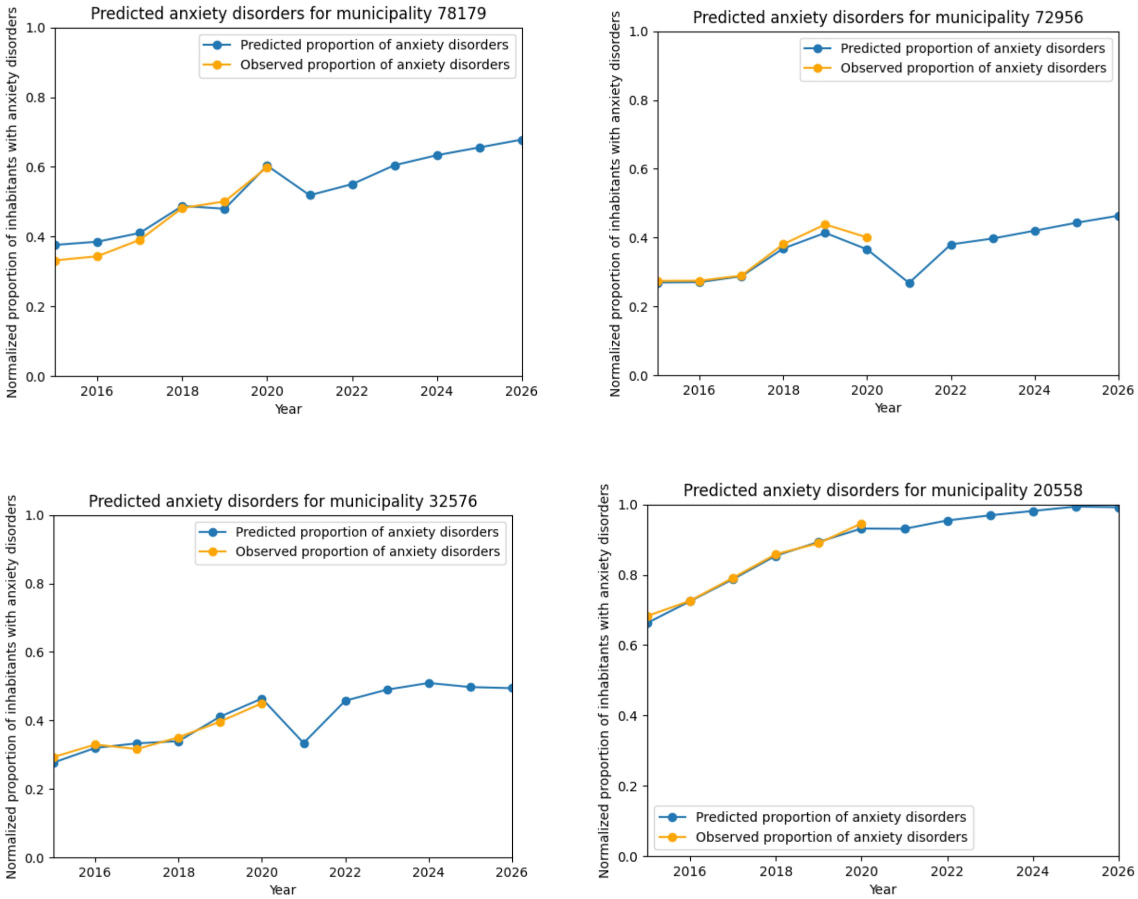

Figure 8 displays 4 of the 325 plots of the proportion of inhabitants with anxiety disorders observed versus predicted by the neural network per year per municipality. The number of anxiety disorders has been measured from 2015 – 2020 as shown in Figure 1.

Table 18 presents the summed gradients of each factor for all municipalities in 2019 based on a 5% decrease and a 5% increase weighted by the population size of each municipality. The table assesses the national impact of these factors on the proportion inhabitants with anxiety disorders.

Factor	**Sum of weighted gradients**
Divorced population	0.0101
Median wealth: Employee income	0.0066
Standardized income: rental property	0.0057
Median wealth: Couple with no child	0.0051
Number of childcare centers within 3 km	0.0048
Unmarried population	0.0047
Number of high schools within 5 km	0.0040
Grey pressure	0.0034
Median wealth: own home	0.0033
Standardized income: Self-employment income	0.0031
Welfare per inhabitant	0.0029
Standardized income: own home	0.0025
Standardized income: Employee income	0.0022
Both parents NL-born	0.0019
Net income: Couple with child(ren)	0.0018
Distance to school VMBO	0.0015
Age group 45 - 65 years	0.0013
Number of VMBO schools within 5 km	0.0013
Benefits per inhabitant	0.0012
Distance to restaurant	0.0012
Age group 25 - 45 years	0.0012
Perennial grassland	0.0012
Households with child(ren)	0.0010
Non-Western migration background	0.0010
Standardized income: Couple with no child	0.0009
Age group > 80 years	0.0009
Distance to cinema	0.0008
Jobs per inhabitant	0.0007
Passenger cars of private individuals	0.0006
Distance to hospital	0.0005
Arable land area	0.0004
Net income: Transfer	0.0004
Distance to large supermarket	0.0004
Net income: Couple with no child	0.0004
Migration background: Africa, Latin America, Asia (excl. Indo & Japan), Turkey	0.0004
Wajong per inhabitant	0.0003
Phosphate excretion	0.0003
Distance to secondary school	0.0003
Birth surplus	0.0003

Median wealth: Single-parent household	0.0003
Natural grassland	0.0003
Standardized income: Transfer income	0.0002
Population growth	0.0002
Net income: Single-parent household	0.0002
University graduates per inhabitant	0.0002
Movement mobility	0.0001
MBO graduates per inhabitant	0.0001
Distance to main road access	0.0001
Distance to library	0.0001
Number of GP practices within 3 km	0.0001
Deaths due to heart disease	0.0000
Deaths due to other causes	0.0000
Total passenger cars	0.0000
Deaths due to neoplasm	0.0000
Horticulture under glass	0.0000
Nuisance from confused individuals per inhabitant	0.0000
Nuisance from drug usage per inhabitant	0.0000
Nuisance from homeless reported per inhabitant	0.0000
Nuisance youth reported per inhabitant	0.0000
Public intoxication reported per inhabitant	0.0000
Balance increase housing	0.0000
COVID-19 pandemic	0.0000
Environmental address density	0.0000
Houses to buy	0.0000
Housing density	0.0000
Nuisance reported per inhabitant	0.0000
Ownership unknown	0.0000
Renting houses	0.0000
Temporary grassland	0.0000
Live births per inhabitant	0.0000
Deaths due to respiratory illness	-0.0001
Average household size	-0.0001
Distance to swimming pool	-0.0001
Migration background: (Former) Neth. Antilles & Aruba	-0.0001
Benefits without AOW per inhabitant	-0.0001
Standardized income: Single household	-0.0001
Disabled per inhabitant	-0.0002
Deaths per inhabitant	-0.0002
Deaths due to external causes	-0.0002
Number of hospitals within 20 km	-0.0002
Net income: Single household	-0.0002
Distance to day-care center	-0.0002
Number of primary schools within 3 km	-0.0002

AOW per inhabitant	-0.0003
Motorcycles	-0.0003
Age group 65 - 80 years	-0.0003
Domestic migration	-0.0003
Migration background: Suriname	-0.0004
Population density	-0.0004
Migration background: Morocco, Ifni, Spanish Sahara or Western Sahara	-0.0005
Unemployment per inhabitant	-0.0005
Horticulture open field	-0.0005
Nitrogen excretion	-0.0005
Green pressure	-0.0006
Households without children	-0.0006
Mopeds	-0.0007
Median wealth: Single-person household	-0.0007
Green fodder	-0.0007
Distance to GP surgery	-0.0007
At least one parent born abroad	-0.0007
Age group 15 - 20 years	-0.0007
Number of cinemas within 10 km	-0.0008
Median wealth: Transfer income	-0.0008
Migration background: Europe, NA, Oceania, Indo, Japan	-0.0008
Number of large supermarkets within 3 km	-0.0009
Distance to railway station	-0.0010
Standardized income: Couple with child(ren)	-0.0011
Migration background: Turkey	-0.0012
Net income: Self-employment	-0.0013
One-person households	-0.0014
Age group < 5 years	-0.0014
Standardized income: Single-parent household	-0.0014
Net income: Employee	-0.0014
Age group 5 - 10 years	-0.0014
Net income: Private households	-0.0015
Net income: bought home	-0.0015
Age group 20 - 25 years	-0.0015
HBO graduates per inhabitant	-0.0017
Secondary graduates per inhabitant	-0.0018
Median wealth: rental property	-0.0019
Net income: rental home	-0.0020
Distance to primary school	-0.0021
Age group 10 - 15 years	-0.0021
Median wealth: Self-employed income	-0.0021
Number of restaurants within 3 km	-0.0025
Standardized income: Private households	-0.0025
Total age pressure	-0.0031

Median wealth: Couple with child(ren)	-0.0035
Married population	-0.0037
Widowed population	-0.0040
Median wealth: Private households	-0.0062

Table 19 presents the sensitivity analysis of municipality 20558 (1 out 285 total municipalities) with varying factor levels from CBS and Police and their predicted effect on the proportion of inhabitants with anxiety disorders expressed in a percentage based on the neural network model.

Factor	Percentage of change for each factor										
	-25%	-20%	-15%	-10%	-5%	0%	5%	10%	15%	20%	25%
Distance to cinema	0,06%	0,05%	0,04%	0,03%	0,01%	0,00%	-0,01%	-0,03%	-0,04%	-0,05%	-0,06%
Distance to day-care center	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Distance to hospital	0,02%	0,01%	0,01%	0,01%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,01%	-0,02%
Distance to large supermarket	-0,14%	-0,11%	-0,09%	-0,06%	-0,03%	0,00%	0,03%	0,06%	0,09%	0,11%	0,14%
Distance to library	0,01%	0,01%	0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,01%
Distance to main road access	0,02%	0,02%	0,01%	0,01%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,02%	-0,02%
Distance to primary school	0,08%	0,06%	0,05%	0,03%	0,02%	0,00%	-0,02%	-0,03%	-0,05%	-0,06%	-0,08%
Distance to railway station	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%	0,01%
Distance to restaurant	0,07%	0,06%	0,04%	0,03%	0,01%	0,00%	-0,01%	-0,03%	-0,04%	-0,06%	-0,07%
Distance to school VMBO	-0,13%	-0,11%	-0,08%	-0,05%	-0,03%	0,00%	0,03%	0,05%	0,08%	0,11%	0,13%
Distance to secondary school	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Distance to swimming pool	0,20%	0,16%	0,12%	0,08%	0,04%	0,00%	-0,04%	-0,08%	-0,12%	-0,16%	-0,20%
Nuisance from confused individuals p	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Nuisance from drug usage per inhabit	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Nuisance from homeless reported per	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Nuisance youth reported per inhabita	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Number of childcare centers within 3	-6,51%	-5,21%	-3,91%	-2,61%	-1,30%	0,00%	1,30%	2,04%	2,01%	1,98%	1,95%
Number of cinemas within 10 km	2,17%	2,01%	1,51%	1,01%	0,50%	0,00%	-0,50%	-1,01%	-1,51%	-2,01%	-2,51%
Number of GP practices within 3 km	0,85%	0,68%	0,51%	0,34%	0,17%	0,00%	-0,17%	-0,34%	-0,51%	-0,68%	-0,85%
Number of high schools within 5 km	-5,78%	-4,62%	-3,47%	-2,31%	-1,16%	0,00%	1,16%	2,05%	2,02%	2,00%	1,97%
Number of hospitals within 20 km	2,10%	1,68%	1,26%	0,84%	0,42%	0,00%	-0,42%	-0,84%	-1,26%	-1,68%	-2,10%
Number of large supermarkets within	1,27%	1,02%	0,76%	0,51%	0,25%	0,00%	-0,25%	-0,51%	-0,76%	-1,02%	-1,27%
Number of primary schools within 3 kı	0,56%	0,44%	0,33%	0,22%	0,11%	0,00%	-0,11%	-0,22%	-0,33%	-0,44%	-0,56%
Number of restaurants within 3 km	2,01%	2,03%	2,05%	2,07%	1,27%	0,00%	-1,27%	-2,54%	-3,81%	-5,07%	-6,34%
Number of VMBO schools within 5 km	-0,78%	-0,62%	-0,47%	-0,31%	-0,16%	0,00%	0,16%	0,31%	0,47%	0,62%	0,78%
Public intoxication reported per inhab	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Age group < 5 years	2,10%	1,68%	1,26%	0,84%	0,42%	0,00%	-0,42%	-0,84%	-1,26%	-1,68%	-2,10%
Age group > 80 years	0,57%	0,46%	0,34%	0,23%	0,11%	0,00%	-0,11%	-0,23%	-0,34%	-0,46%	-0,57%
Age group 10 - 15 years	-0,03%	-0,03%	-0,02%	-0,01%	-0,01%	0,00%	0,01%	0,01%	0,02%	0,03%	0,03%
Age group 15 - 20 years	0,64%	0,52%	0,39%	0,26%	0,13%	0,00%	-0,13%	-0,26%	-0,39%	-0,52%	-0,64%
Age group 20 - 25 years	0,43%	0,34%	0,26%	0,17%	0,09%	0,00%	-0,09%	-0,17%	-0,26%	-0,34%	-0,43%
Age group 25 - 45 years	-6,13%	-4,90%	-3,68%	-2,45%	-1,23%	0,00%	1,23%	2,45%	3,68%	4,90%	6,13%
Age group 45 - 65 years	-0,46%	-0,37%	-0,28%	-0,19%	-0,09%	0,00%	0,09%	0,19%	0,28%	0,37%	0,46%
Age group 5 - 10 years	1,34%	1,07%	0,80%	0,54%	0,27%	0,00%	-0,27%	-0,54%	-0,80%	-1,07%	-1,34%
Age group 65 - 80 years	-0,04%	-0,03%	-0,02%	-0,02%	-0,01%	0,00%	0,01%	0,02%	0,02%	0,03%	0,04%
AOW per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Arable land area	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
At least one parent born abroad	0,76%	0,61%	0,46%	0,31%	0,15%	0,00%	-0,15%	-0,31%	-0,46%	-0,61%	-0,76%
Average household size	0,02%	0,02%	0,01%	0,01%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,02%	-0,02%
Balance increase housing	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Benefits per inhabitant	0,05%	0,04%	0,03%	0,02%	0,01%	0,00%	-0,01%	-0,02%	-0,03%	-0,04%	-0,05%
Benefits without AOW per inhabitant	-0,01%	-0,01%	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,01%	0,01%	0,01%	0,01%
Birth surplus	0,07%	0,05%	0,04%	0,03%	0,01%	0,00%	-0,01%	-0,03%	-0,04%	-0,05%	-0,07%
Both parents NL-born	-0,02%	-0,01%	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,01%	0,01%	0,01%	0,02%
COVID-19 pandemic	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Deaths due to external causes	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	-0,15%	-0,31%
Deaths due to heart disease	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Deaths due to neoplasm	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Deaths due to other causes	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Deaths due to respiratory illness	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Deaths per inhabitant	-0,01%	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%	0,01%	0,01%
Disabled per inhabitant	-0,01%	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%	0,01%	0,01%
Distance to GP surgery	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Divorced population	-6,78%	-5,42%	-4,07%	-2,71%	-1,36%	0,00%	1,36%	2,71%	4,07%	5,42%	6,78%
Domestic migration	2,19%	1,76%	1,32%	0,88%	0,44%	0,00%	-0,13%	-0,13%	-0,13%	-0,13%	-0,13%
Environmental address density	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Green fodder	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Green pressure	0,06%	0,05%	0,04%	0,02%	0,01%	0,00%	-0,01%	-0,02%	-0,04%	-0,05%	-0,06%
Grey pressure	-0,18%	-0,15%	-0,11%	-0,07%	-0,04%	0,00%	0,04%	0,07%	0,11%	0,15%	0,18%
HBO graduates per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Horticulture open field	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Horticulture under glass	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Households with child(ren)	-0,41%	-0,33%	-0,24%	-0,16%	-0,08%	0,00%	0,08%	0,16%	0,24%	0,33%	0,41%
Households without children	0,06%	0,04%	0,03%	0,02%	0,01%	0,00%	-0,01%	-0,02%	-0,03%	-0,04%	-0,06%

Table 20 presents the sensitivity analysis of municipality 20558 (1 out 285 total municipalities) with varying factor levels from CBS and Police and their predicted effect on the proportion of inhabitants with anxiety disorders expressed in a percentage based on the neural network model.

Households without children	0,06%	0,04%	0,03%	0,02%	0,01%	0,00%	-0,01%	-0,02%	-0,03%	-0,04%	-0,06%
Houses to buy	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Housing density	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Jobs per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Live births per inhabitant	-0,01%	-0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,01%	0,01%
Married population	0,56%	0,45%	0,34%	0,23%	0,11%	0,00%	-0,11%	-0,23%	-0,34%	-0,45%	-0,56%
MBO graduates per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Couple with child(ren)	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Couple with no child	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Employee income	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: own home	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Private households	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: rental property	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Self-employed income	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Single-parent househ	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Single-person househ	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Median wealth: Transfer income	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Migration background: (Former) Neth.	-0,03%	-0,03%	-0,02%	-0,01%	-0,01%	0,00%	0,01%	0,01%	0,02%	0,03%	0,03%
Migration background: Africa, Latin Am	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Migration background: Europe, NA, Oc	-0,08%	-0,06%	-0,05%	-0,03%	-0,02%	0,00%	0,02%	0,03%	0,05%	0,06%	0,08%
Migration background: Morocco, Ifni, S	0,26%	0,21%	0,16%	0,11%	0,05%	0,00%	-0,05%	-0,11%	-0,16%	-0,21%	-0,26%
Migration background: Suriname	0,25%	0,20%	0,15%	0,10%	0,05%	0,00%	-0,05%	-0,10%	-0,15%	-0,20%	-0,25%
Migration background: Turkey	0,23%	0,18%	0,14%	0,09%	0,05%	0,00%	-0,05%	-0,09%	-0,14%	-0,18%	-0,23%
Mopeds	-0,67%	-0,54%	-0,40%	-0,27%	-0,13%	0,00%	0,13%	0,27%	0,40%	0,54%	0,67%
Motorcycles	-0,09%	-0,07%	-0,06%	-0,04%	-0,02%	0,00%	0,02%	0,04%	0,06%	0,07%	0,09%
Movement mobility	-0,13%	-0,13%	-0,13%	-0,13%	-0,13%	0,00%	0,14%	0,29%	0,43%	0,58%	0,72%
Natural grassland	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Transfer	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: bought home	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Couple with child(ren)	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Couple with no child	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Employee	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Private households	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: rental home	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Self-employment	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Single household	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Net income: Single-parent household	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Nitrogen excretion	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Non-Western migration background	1,52%	1,21%	0,91%	0,61%	0,30%	0,00%	-0,30%	-0,61%	-0,91%	-1,21%	-1,52%
Nuisance reported per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
One-person households	1,22%	0,97%	0,73%	0,49%	0,24%	0,00%	-0,24%	-0,49%	-0,73%	-0,97%	-1,22%
Ownership unknown	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Passenger cars of private individuals	0,01%	0,01%	0,01%	0,00%	0,00%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,01%
Perennial grassland	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Phosphate excretion	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Population density	-0,17%	-0,13%	-0,10%	-0,07%	-0,03%	0,00%	0,03%	0,07%	0,10%	0,13%	0,17%
Population growth	-0,13%	-0,13%	-0,13%	-0,13%	-0,13%	0,00%	0,26%	0,52%	0,78%	1,04%	1,30%
Renting houses	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Secondary graduates per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Couple with chil	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Couple with no	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Employee incom	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: own home	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Private househo	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: rental property	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Self-employmer	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Single househol	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Single-parent hc	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Standardized income: Transfer income	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Temporary grassland	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Total age pressure	0,07%	0,06%	0,04%	0,03%	0,01%	0,00%	-0,01%	-0,03%	-0,04%	-0,06%	-0,07%
Total passenger cars	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Unemployment per inhabitant	-0,04%	-0,03%	-0,02%	-0,01%	-0,01%	0,00%	0,01%	0,01%	0,02%	0,03%	0,04%
University graduates per inhabitant	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Unmarried population	-5,68%	-4,54%	-3,41%	-2,27%	-1,14%	0,00%	1,14%	2,27%	3,41%	4,54%	5,68%
Wajong per inhabitant	0,01%	0,01%	0,01%	0,01%	0,00%	0,00%	0,00%	-0,01%	-0,01%	-0,01%	-0,01%
Welfare per inhabitant	0,90%	0,35%	0,26%	0,17%	0,09%	0,00%	-0,09%	-0,17%	-0,26%	-0,35%	-0,43%
Widowed population	0,12%	0,10%	0,07%	0,05%	0,02%	0,00%	-0,02%	-0,05%	-0,07%	-0,10%	-0,12%

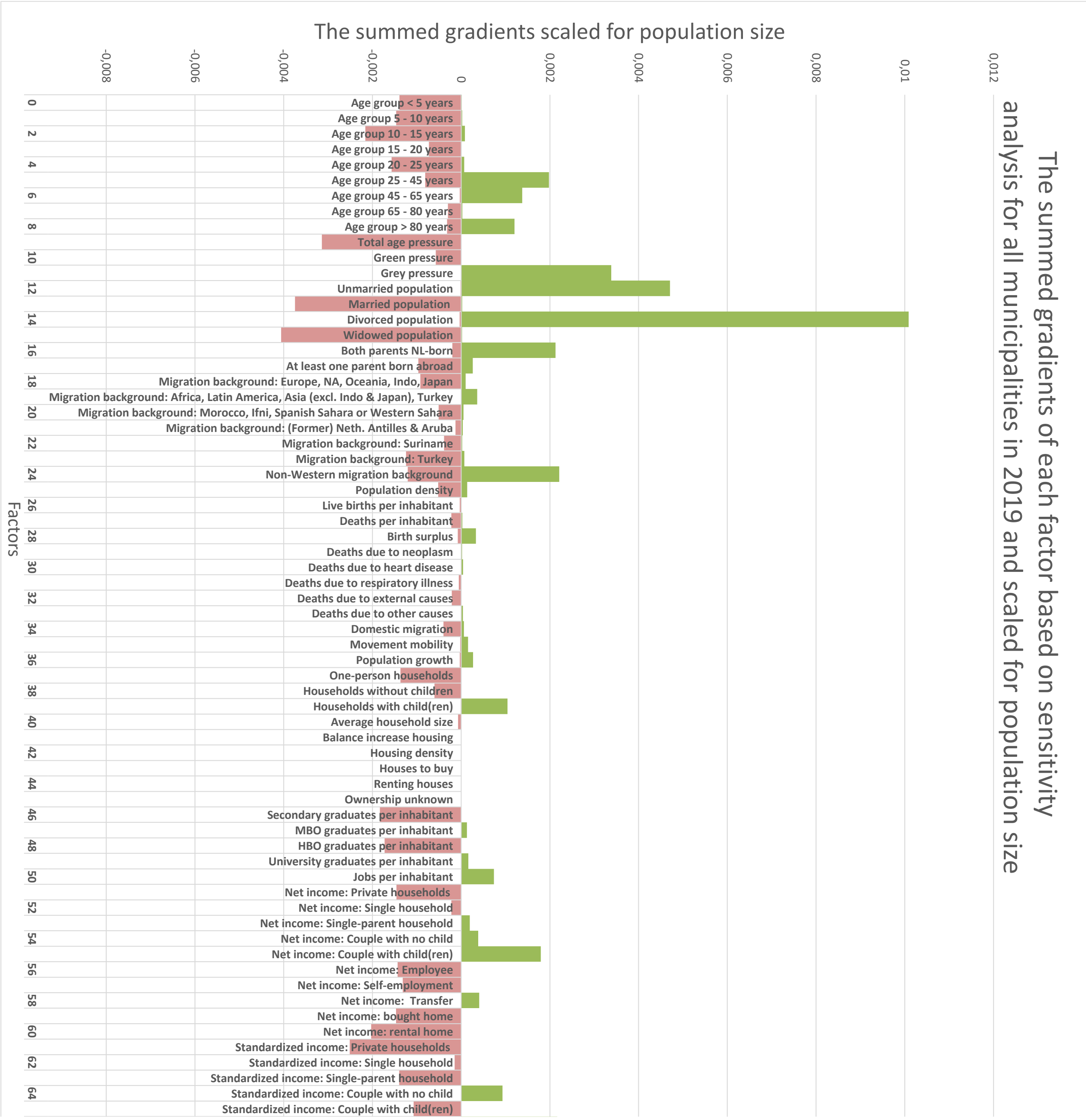

Figure 9 visualizes the summed gradients of each factor for all municipalities in 2019 based on a 5% decrease and a 5% increase weighted by the population size of each municipality. The figure assesses the national impact of these factors on the proportion inhabitants with anxiety disorders (part 1 of 2).

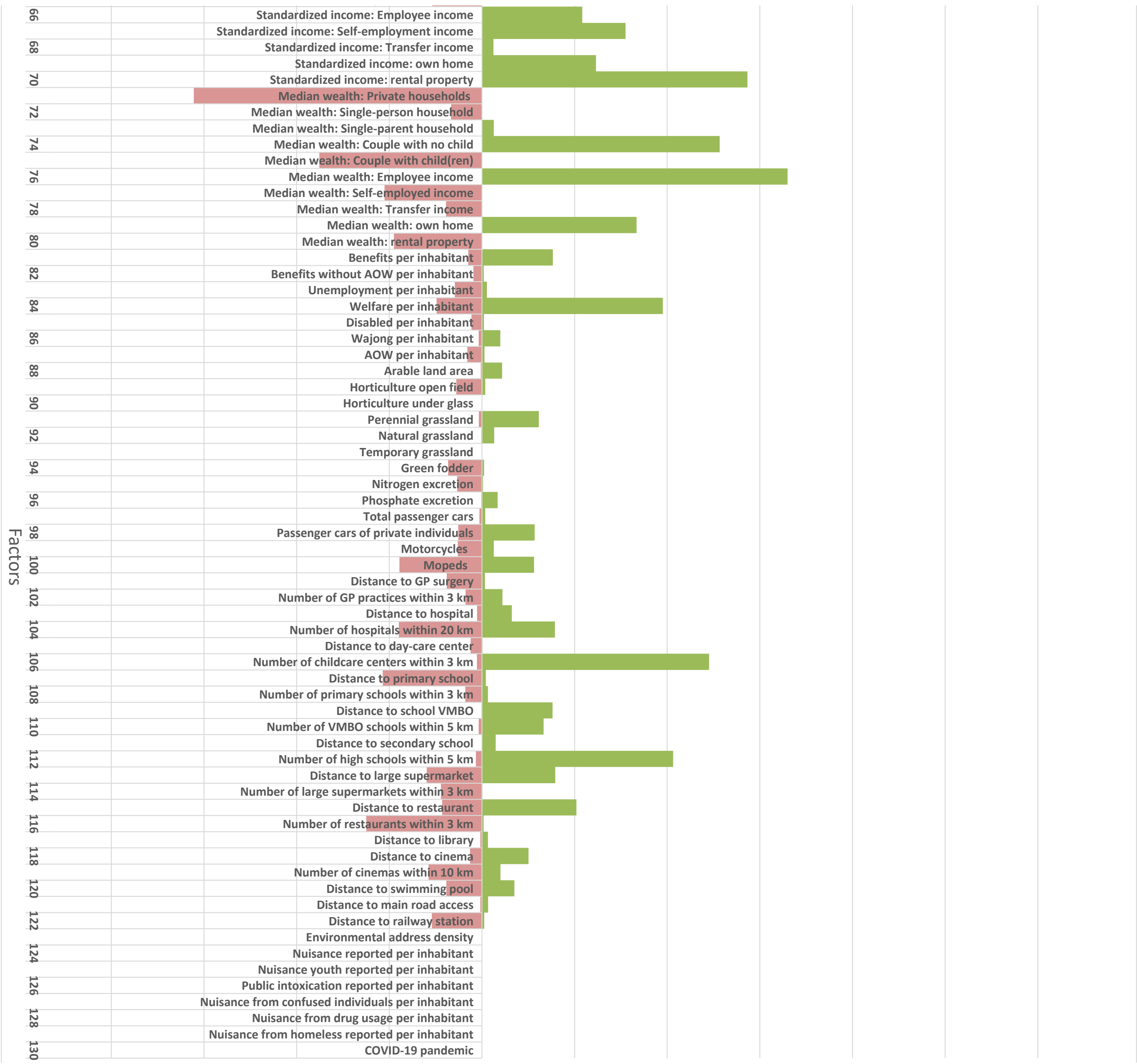

Figure 10 visualizes the summed gradients of each factor for all municipalities in 2019 based on a 5% decrease and a 5% increase weighted by the population size of each municipality. The figure assesses the national impact of these factors on the proportion inhabitants with anxiety disorders (part 2 of 2).