The effects of green on mental health between different SES factors in Groningen

How public green and the proximity towards urban green spaces differently impact people with contrasting socio-demographic backgrounds

Date: 18th of August, 2023

Author: Vincent Scholing

Programme: Society, Sustainability & Planning

Supervisor: Samira Ramezani

Colophon

Title:	The effects of green on mental health between different backgrounds in Groningen
Subtitle:	Investigating how public green and the proximity towards urban green spaces differently impact people with contrasting socio-demographic backgrounds
Date:	18th of August, 2023
Word count:	15.618 (including figures & references)

Author:	Vincent Scholing
Student number:	S3539601
Email address (RUG):	V.scholing@student.rug.nl
Email address (private):	Vscholing@gmail.com

Programme:	Master Society, Sustainability & Planning
	Faculty of Spatial Sciences
	University of Groningen
	Landleven 1 9747 AD Groningen
Supervisor:	Dr. Samira Ramzani
Front image:	Felixx Landscape Architects & Planners BV

Acknowledgements

I want to thank my supervisor, Samira Ramezani, for her guidance and help in providing feedback in realising this Master's Thesis. Furthermore, I would like to thank friends and family who were interested enough to discuss certain decision points of the Thesis to give me further insights to develop a well-rounded Thesis.

Abstract

As the city grows, different amenities may need to give space to dwellings. A growing body of literature found that green positively impacts subjective well-being. As a result, investigating whether 'green' is available for everyone is essential in a growing city with limited space.

Literature acknowledges the importance of green, showing outcomes of increased levels of well-being in the short- long term. At the same time, little research has been done on the SES factors. A growing body of research shows that there may be an unequal distribution between different SES groups.

An analysis of the impact of public green and public parks on stress and depression/anxiety risk. A separate analysis is done to see whether public green & distance to public parks differ between SES factors.

Distance to public parks does not impact subjective well-being, while public green does. Outcomes related to the SES factors showed that public green & public parks were relatively equally distributed as there were few differences. But socio-economic differences showed that work participation reduced the distance to a public park. While income decreased the percentage of public green, which contradicts the current literature.

It is recommended that future research focuses on lower geographic levels to grasp differences between households better. This research can be used as a framework for determining what factors to consider in developing new green spaces in the city

Keywords: Subjective mental well-being, public green, socio-economic-status, urban green space

Table of contents

Table of contents

List of figures:

List of tables:

List of images:

1.0 Introduction

1.1 Societal relevance

In the Netherlands there is currently a dwelling shortage, to combat this, the municipality of Groningen shared their "woonvisie" translated to "living vision" on how to achieve the disappearance of a dwelling shortage for all groups in Groningen. Currently in The Netherlands for 89% of its citizens a green place is within 1km of reach (Statistiek, 2012). But with the current need for additional dwellings, the question is whether the relative proximity towards a green space remains doable for the general citizen of Groningen. Additionally, the same problems apply to public urban green, as more need for dwellings may take up spaces that otherwise could have been used for public green. As the municipality of Groningen need to expand available housing to combat the shortage they need to either decide to expand the city or replace different functions of the city. The risk of the latter is that according to literature Public open space (POS) may come into distress and as a consequence lower quality POS result into worse mental health outcomes Francis *et al.*, (2012) .

The EU acknowledges the importance of mental well-being (*Mental health*, 2023) and because of that the European Commission developed a new approach for mental health in 2023. Furthermore, the EU acknowledges & emphasizes the need for urban green spaces (UGS) as **"**Public green space makes room for play, social interaction, creativity, economic activities and entertainment, the very things a city thrives on". Additionally, the EU also addresses future urban challenges such as city heating as a result of global warming (Alberti *et al.*, 2019). Green spaces decrease the negative impact of city heating by absorbing the heat or facilitating shade (US EPA, 2015). The WHO recommends at least 9 m2 of publicly accessible green spaces per person (Didier Vancutsem *et al.*, 2009). The city of Groningen has 34,32 m2 per inhabitant which is quite a bit higher than the recommended number (Joint Research Centre (European Commission) *et al.*, 2019).

The translates into 46 % of public green space. How this compares to the 4 other largest cities of the Netherlands is positive, as Amsterdam has 30%, Rotterdam has 25% Utrecht has 41% and The Hague has 36%. In Berlin urban green is around 49% which is higher than Groningen (*HUGSI for rankings | HUGSI*, no date). But a study done in Berlin by Bertram and Rehdanz (2015) shows that the majority of people in Berlin have too little green located near them.

Thus, even when public urban green may be relatively high it does not immediately imply that every individual or group has equal access to urban green as it may differ from neighbourhood to neighbourhood. Even though, Berlin is different in size and amenities to Groningen. It is important to investigate whether public parks and public green in general are accessible within the city for everyone.

In the municipality of Groningen, there are multiple public parks which include: Het Stadspark, Het Noorderplantsoen, Bessemoerpark, Oosterpark, Pioenpark, Park Selwerd & Het Sterrenbos. These are dispersed through Groningen. These parks are often used for recreational purposes or to exercise. Literature shows that similar parks when subjected to certain qualitative standards which have a positive impact for citizens their well-being in the long term (Bertram and Rehdanz, 2015a; R.F. Hunter *et al.*, 2019; Sharifi *et al.*, 2021a). But also, in the short term, UGS can have a stress-reducing effect on people (Ulrich *et al.*, 1991). Furthermore, green in general such as within the neighbourhood is associated with benefits to mental well-being (MacKerron and Mourato, 2013a; Magdalena van den Berg *et al.*, 2015).

It seems that benefits and the usage of urban green tend to be better/higher for groups with a lower socio-economic status (SES) compared to people with higher SES statuses. (Maas *et al.*, 2006a; McEachan *et al.*, 2016; van den Berg *et al.*, 2016a)

There is no general consensus about the access to public parks as for individuals with a lower SES, access was worse in Australia (Astell-Burt *et al.*, 2014a), while a study in Seoul suggests that individuals with lower SES have better access to public parks (Heo *et al.*, 2021). But it is yet unclear how access to public parks in the northern parts of The Netherlands is.

1.2 Academic relevance

Within the academic world, subjective well-being (SWB) does differ between & within a culture (Suh *et al.*, 1998) as the focus can lie on either the individual or the collective. SWB is often seen as the experience of the individual in the case of the western world (Angner, 2009). Public parks do often have a positive impact on the SWB (Bertram and Rehdanz, 2015a; R.F. Hunter *et al.*, 2019; Sharifi *et al.*, 2021a). This is not only because of leisure activities but also because of the facilitation of multiple activities such as physical activity which boosts the SWB in return (Brown, Schebella and Weber, 2014). Furthermore, green itself seems to have restorative effects as it can reduce stress that is experienced on the short-term, (Ulrich *et al.*, 1991; Grahn and Stigsdotter, 2003)

´Good´ SWB of people is a desired mental state of people (Angner, 2009). ´Good´ experiences can on the short-term cause enjoyment of life & on the long-term contentment. While the other way around is also true as ´bad´ SWB can cause sadness & depression on the long-term respectfully (Diener, 2000; Joshi, 2010). Well-being can in itself also have a upward spiralling effect as people with higher SWB have often closer & more supportive social relationships (Diener and Ryan, 2009a).

What is known about the link between SWB & UGS is quite extensive. MacKerron and Mourato (2013) found a positive association between being in nature & an increased SWB. Also the quantity of green in the environment increases SWB of individuals (Maas *et al.*, 2006a; Magdalena van den Berg *et al.*, 2015). The multifunctional purpose of UGS like the ability to walk through it, have social meetings and the possibility to exercise increases the SWB (Brown, Schebella and Weber, 2014). . General negative emotions such as sadness seem to beneficially change when making use of UGS (Bowler *et al.*, 2010a).

The usage of public parks does seem determined by multiple components such as the quality and the size of the public park (Bowler *et al.*, 2010; Mytton *et al.*, 2012; Koohsari *et al.*, 2018; Sharifi, Nygaard and Stone, 2021). Furthermore, it seems that people with lower level of education make more use of public parks (van den Berg *et al.*, 2016a)

But while a lot is known Sharifi et al., (2021) mentions that there is a knowledge gap in the understanding of how UGS are underused by vulnerable groups which may be caused by a lack of access. Additionally, a better understanding on the usage and general access to green within the neighbourhood would give a better overview of how this affects their mental well-being.

Furthermore, it is unknown to the author that a general study is done about SWB and UGS in the Dutch context, yet alone the northern part of The Netherlands. This general focus and the focus which is more on the SES of people could contribute to literature for better insights into whether an equal distribution of this amenity is needed and whether it positively influences the SWB of people among different SES factors. The general focus on green within a neighbourhood can also be used to determine its effect on the subjective well-being and can make a strong comparison to public parks.

1.3 Research aim & research question

This research aims to try to get a better overview of how people with different socioeconomic- statuses (SES) are negatively affected in their SWB and whether there is a difference between SES factors and public park access. Additionally, a similair comparison is made between public green and SES factors.
These investigations can give an insight into whether access and/or green have a significant impact on people their SWB.

To achieve this aim. The main research question is developed as follows:

"To what extent does access to urban green spaces and public green affect the subjective well-being of different groups of people, and is its access equally distributed among people with different socio-economic statuses in Groningen?

1. To what extent does proximity to urban green spaces affect the subjective well-being of people in the Northern Netherlands?

2. To what extent does public green in neighbourhoods influence the subjective well-being of people?

3. "To what extent is there a difference between people with varying SES factors and their access to a public park?"

4. "To what extent is there a difference between people with different SES factors and the presence of public green

In answering these research questions, a greater understanding can be developed of how public parks and green affect the SWB of citizens in Groningen. Furthermore, the effect on people with different SES factors may give a clearer answer on whether policymakers should focus on accessibility to urban green spaces, having sufficient public green in the neighbourhoods or both.

2.0 Literature review

2.1 Subjective well-being

In literature, SWB can be seen as a person's evaluation of their own live (Diener, 2000; Christina Geng-ging Chi, Ruiying Cai, and Yongfen Li, 2017). More broadly, These evaluations, which can be both positive and negative, include judgments and feelings about life satisfaction (Diener and Ryan, 2009a). The question can be posed "To what degree does a person evaluate her or his life and current situation as good and desirable versus undesirable and negative?" (Diener *et al.*, 2017 P. 134).

Subjective well-being as mentioned by Sharifi, Nygaard and Stone, (2021) can be both eudaimonic or hedonistic the former refers to well-being defined as fulfilment and self-realization and not influenced by recent experiences. While hedonic describes the satisfaction from life, consisting of positive or negative emotions which come from immediate experiences and overall life satisfaction. Additionally on the hedonic view, according to Angner, (2009) SWB is a matter of desired mental state of an individual as individuals attach great value to it and there are associations to SWB related to good social relationships, good health, being productive (Christina Geng-ging Chi, Ruiying Cai, and Yongfen Li, 2017; Joshi, 2010), recreation & feeling of purpose (Diener and Ryan, 2009a). Both concepts should not be seen apart from each other but interrelated.

The definition that will be used in this paper is the hedonic view on SWB, as short-term life experiences are of importance for this paper. Because of this, a deeper understanding of the hedonic view is needed. The hedonic view is how people evaluate their own SWB they do this according to Diener, (2000) & Joshi, (2010) through several separable components. These are: 1. life satisfaction which is the more general judgement of an individual's life, 2. satisfaction with a person's important domain such as work or relationships , 3. positive affect which includes the various feelings people experience when things seem to be going well such as enjoyment of life on the short term & contentment of life on the long-term (Diener and Ryan, 2009a). Finally, 4. negative affect which is the experience that things are not going well which can express itself through feelings of anger or sadness on the short term and longer lasting negative moods on the long term such as depression (Diener and Ryan, 2009a). This underlines the fact that higher general life satisfaction does not immediately imply that either a lot of negative experiences or a lot of positive experiences only shape the SWB of an individual. As Diener and Chan, (2011) mention that evidence shows that both positive and negative feelings have independent effects but can in some instances overlap but this mechanisms is not too clear within literature. All these components play a role which shape the SWB of an individual on both the short & long term.

Proceeding on level of well-being and life satisfaction. A high level can significantly improve life within four areas of life which are health and longevity, work and income, social relations, and societal benefits (Diener and Ryan, 2009a). Additionally this can strengthen the level of well-being again as individuals with higher levels of well-being tend to have closer & supportive social relationships in comparison to people with lower levels of subjective well-being (Diener and Biswas-Diener, 2008). Thus, this shows as mentioned by (Diener and Ryan, 2009 P. 392) that: "people with high subjective well-being tend to have higher levels of self-confidence, warmth, leadership ability, sociability, and more friends to begin with suggests the other side of the causal arrow - people with high subjective well-being actually generate their own social support system".

For this paper the focus lies on two of the four areas which are work and income and social relations. As the other two which are: health & longevity and societal benefits are more

outcomes of either good or bad health. While work & income and social relations have a impact on SWB itself

2.2 Urban green space

Urban green spaces (UGS) within the city can be seen as spaces that are developed for recreational purposes such as public parks & other types of green spaces such as street trees & green roofs (R.F. Hunter *et al.*, 2019) For the purpose of this study, UGS will be split into two parts which is either 'public green' this refers to locations of green within a city that are freely accessible. An example of this would already be a patch of grass alongside water. 'Public parks' are actual physical locations such as The Stadspark & Noorderplantsoen in the city of Groningen. This means that for clarity, public green does take into account public parks in the later analysis while public parks are locations with a number of restrictions are set to define what a public park is. The reasoning including public parks as a separate variable has to do with the recreational possibilities of a public park. As according to Hunter et al., (2019) public parks are used for recreation such as leisure or facilitating exercise. But also for social cohesion according to (van den Berg *et al.*, 2010). Koohsari *et al.*, (2018) mentions that people their willingness to walk in a public park is influenced by the size of the park, where only public parks larger than 1,5ha reported greater odds in their willingness to make use of a public park.

Public parks are public spaces within cities with a high degree of green in them. These urban spaces are according to Hunter et al., (2019) used for recreation such as leisure or facilitating exercise & social cohesion (van den Berg *et al.*, 2010)

2.3 General benefits to mental health associated with green

MacKerron and Mourato, (2013) provides evidence to the link between both nature/green and an increase in SWB while considering a wide range of other factors to control for. Both Magdalena van den Berg *et al.*, (2015) Maas *et al.*, (2006) confirms in their research that there is a strong link between the quantity of green spaces and the subjective well-being. This link shows that a small buffer of green around a person's living environment could already increase the perceived general health in a population. Maas et al., (2006) stresses the outcome of the study that green spaces are not just merely a luxury product. as green space is especially beneficial in general health outcomes for people with a lower SES in comparison to people with a higher SES. When public parks are relatively close an individual could make more use of it. van den Berg *et al.*, (2016) mentions in her cross-sectional study that a higher frequency of using UGS such as public parks resulted in less perceived stress by individuals. Additionally, not only the higher frequency of use but also a higher number of hours within the public park resulted in less perceived stress (van den Bert *et al*., 2016).

Public parks are multifunctional, so the benefits that arise there are not necessarily only because of the green area but also the different recreational options such as exercise and leisure have a positive benefit on the subjective well-being . Brown, Schebella and Weber, (2014) confirms that urban parks do provide a wide range of benefits which are physical, environmental, psychological & social but it does depend on the park type to provide these benefits.

While not only longer term subjective well-being of people are increased, UGS also tends to reduce the level of short term stress of people (Ulrich *et al.*, 1991). In this study the recuperation of stress was faster for students in a natural setting in comparison to the urban environment. To go more into depth, the natural environment had 'restorative influences. These were an increase in positive effects such as calmness and reduced negative effects such as anger and fear. As Ulrich et al., (1991 P.223) mentions: "In general, individuals exposed to the natural settings both reported improved feeling states and evidenced lower stress levels in

physiological indicators." This 'restorative quality' of nature is also confirmed by (Grahn and Stigsdotter, 2003) who found a positive association between self-reported stress levels being lower and increased use of public parks in Sweden.

2.4 Shorter term benefits to mental health

Bowler *et al.*, (2010) reviewed studies that focused on the short-term effects on the well-being of respondents after making use of public green. There were beneficial changes on self-reported feelings on energy, anxiety, anger, fatigue and sadness. This underline (Diener, 2000) components of increased positive affect (experiences of people) and reduced negative affect of people with the corresponding positive & negative emotions. (Bowler et *al.*, 2010) looked into before & aftereffects of an activity in nature which were overall positive for the individual.

2.5 Proximity of households to a public park

To make use of an public park, the proximity of should be close enough for an individual to want to make use of it. In this case proximity mostly describes the possibility of walking to a park. Rigolon, (2016) does not find a clear answer in a literature review to the proximity of parks between different SES factors as the literature in general is very contradicting to each other. The author suggests focusing more on the park acreage which is the number or size of parks and recreational facilities and to focus on the park's quality instead. This notion of importance of quality of a park is also shared by (Bowler *et al.*, 2010a; Mytton *et al.*, 2012; Sharifi, Nygaard and Stone, 2021).

The usage of an public park are different among people and is especially dependent on the level of education. A lower level of education would result in a higher use of public parks (van den Berg *et al.*, 2016). Van den Berg *et al.,* (2016) explains this difference between high and low educated and UGS usage as the lower educated group has often fewer possibilities to find stress reducing activities in their direct environment such as public green compared to higher educated people. Furthermore, Mytton *et al.*, (2012) found that people living in the relative greenest areas were most likely to get the recommended amount of physical activity which improves subjective well-being even when restricting it to only urban areas. Additionally (Su *et al.*, 2019) found a familiar outcome related to green vegetation and exercise.

Bertram and Rehdanz, (2015), found an association between life satisfaction and the percentage of public parks to be in close by. This effect was most positive when there was at least 11% of the 1km buffer is green public parks, this article does mention that around 75% of their sample is not able to achieve this 11% UGS within a kilometre of their home. The reason according to Bertram and Rehdanz, (2015) for increased life satisfaction with closer proximity may have to do with a lower threshold to exercise in nature, which increases human health both physically & mentally.

2.6 Demographics & well-being

(Diener and Ryan, 2009a) mention in their general overview the most important demographic variables looked at related to well-being of individuals in literature.
The first important variable is gender, according to (Diener and Ryan, 2009a) while there are differences in experiences of positive & negative emotions. Generally, the differences between men & women and subjective well-being are not substantially different.

Women tended to benefit more than males did from the protective associations between green space on their physical health (Sillman *et al.*, 2022). Furthermore, according to(Braçe, Garrido-Cumbrera and Correa-Fernández, 2021) found that females tended to find qualitative characteristics of public parks more important than males. Finally Bolte, Nanninga and

Dandolo, (2019) did not find any difference between residential green and people their self-rated health.

Secondly, education is an often-used variable, educational attainment and well-being does not reveal any strong trends. While the trend is positive for higher levels of education, it is rather small with an increase between 1% – 3% for well-being (Witter *et al.*, 1984). For all education levels a greener environment was positively associated with subjective well-being according to (Maas *et al.*, 2006a). People with a lower education showed to be most sensitive to green related to their mental health in comparison to higher educated people. Literature shows that the higher the educational level of an household, the more urban green space is available in their area in comparison to groups with lower education (Nesbitt *et al.*, 2019)

Diener and Ryan, (2009) mention that subjective well-being & age while associated with each other does not have a consistent linear relationship with each other but is more dependent on other variables. Xu, Nordin and Aini, (2022) found that for the elderly, green space is still of great importance. Maas et al., (2006) found a positive association for all age groups related to green in the direct environment (1 km buffer around people their home). But there does not seem to be big differences between the different age groups.

Being religious has often a positive correlation with well-being which may be due to the sense of purpose and meaning. But the contradicting part is that countries with the relative highest levels of well-being are not very religious such as Sweden & The Netherlands (Diener and Ryan, 2009).

Social relationship is a variable that is associated with increased positive feelings of subjective well-being. As Diener and Ryan, (2009. P397) mention: "In general, people are simply happier when they are around other people". A lack of social relationships may lead to loneliness. Furthermore, social bonds such as marriage also tend to have a positive effect on the subjective well-being. For adults, perceptions of loneliness was less when there was higher proportion of green including public parks (Chen *et al.*, 2021). Green space also seems to increase social interactions according to (Aram, Solgi and Holden, 2019) and as a result can have beneficial effects on combating loneliness in an neighbourhood.

The employment status of a person, especially unemployment has a negative impact on the subjective well-being of an individual according to (Clark, 2010). As people have a hard time adapting towards unemployment this results in lower levels of subjective well-being. While there is improvement over time, when remaining unemployed. People do not tend to return to their baseline subjective well-being in comparison when being employed or self-employed. Schüle, Gabriel and Bolte, (2017) investigated whether the socioeconomic position of individuals which included the unemployment status as a factor had a relation with public green. The outcome of this research showed that neighbourhoods with a lower socioeconomic position (SEP) had less green space available in comparison to high SEP neighbourhoods

Finally, income has a positive effect on well-being, but this effect does shrink over time dependent on how high the income is (the higher the income the lower the positive effect of subjective well-being is) (Diener and Ryan, 2009). (Astell-Burt *et al.*, 2014a) mention that in Australia the distribution of parks and other green spaces is unequal.

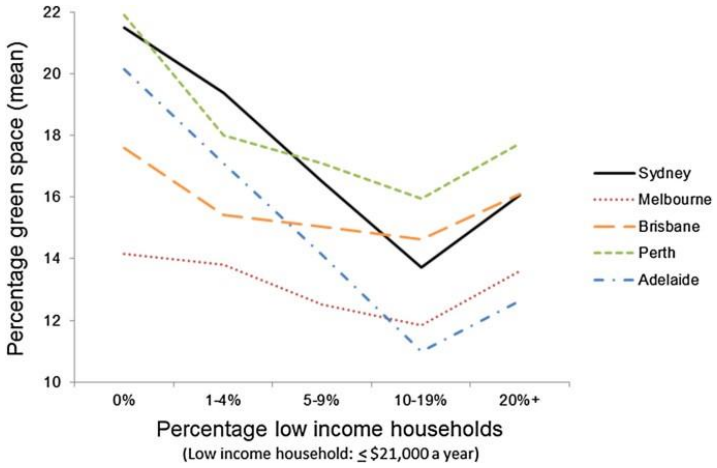

Figure 1 low income households percentage per neighbourhood and the associated green space in the neighbourhood (Astell-Burt et al., 2014)

Being worse for lower income compared to higher income neighbourhoods as seen in figure 1. De Vries, Buijs and Snep, (2020) adds to this as in this paper the percentage of greenspace within 250m was lower for people with a lower income. But this difference was larger for low-urban municipalities compared to high-urban municipalities like Groningen but it was still significant.

These factors above shows the complex relation of subjective well-being as a lot of factors influence mental health including genetic, behavioural & phycological (Diener *et al.*, 2017). Meaning that while a wide collection of social and demographic variables have been included. There are still a number of key factors that are not included, so the complexity of subjective well-being should be kept in mind

2.7 Conceptual model

The conceptual model shown in figure 2 gives a clear overview on how the mechanisms behind public parks and public green. The subjective well-being of individuals is influenced by 3 factors in this model. First of all this is the greenness of a neighbourhood which is expected to have a positive effect on the subjective well-being. The second factor is the proximity / distance to a public park which may encourage public park usage and as a result have an increased level of subjective well-being. Finally, socio-demographic factors including income, work participation, gender, education level, feelings of loneliness and age may have a positive or negative effect people their subjective well-being as this may differ per socio-demographic group. While not only influencing the subjective well-being of an individual. This conceptual model explains that people with different socio-demographic factors have different outcomes when it comes to their greenness of neighbourhood or the proximity that individuals live to a public park.

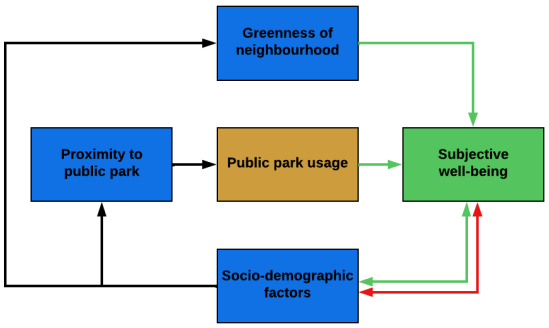

Figure 2 Conceptual model on the proximity & mental health outcomes of UGS

3.0 Methodology

3.1 Research strategy

The aim of this research consists out of 2 main parts, the first one is to investigate whether there is a relationship between the distance people live from a park and their subjective well-being in the city of Groningen. The second aim is to investigate whether public green in the neighbourhood and people their subjective well-being in the city of Groningen. Additionally, this paper also investigates whether people with different socio-economic factors are impacted differently in relation to the distance to a public park and public green. For the research question, it is most logical to consider the use of quantitative data. As the purpose of this study requires more generalizable data which is often collected quantitatively opposed to qualitative (Punch, 2013) As answering this research question it is most efficient to do it with a larger set of numerical respondents and thus a quantitative method is most useful. In the case that the research question focused more on the reasoning behind residents' choices a qualitative research method would be more appropriate (Punch, 2013).

While designing a new survey for this paper has been considered as the benefit of being able to make very specific questions related to RQ makes this paper less reliant on less accurate secondary data. The limitations such as needing a large sample to represent the characteristics of the population as a whole (Burt, Barber and Rigby, 2009) were too substantial as differentiations between SES factors were also needed to be made to answer all the RQ's properly.

3.2 Data characteristics

The secondary data that will be used to perform a statistical analysis on is aggregated data of neighbourhoods in Groningen. The reason why this form of data has been chosen instead of data on a lower geographical level like household level has to do with privacy and the restrictions in accessing the data from governmental bodies. While efforts have been put to acquire this type of data, having neighbourhood aggregated data is the best alternative in answering the RQ's of this paper.

In the next chapters of the methodology, a deeper explanation of the data will be given. Starting with the description of the area, a description of the data itself and the statistical outcomes of the statistical analysis.

Region of the dataset

As mentioned the region that will be analysed is the municipality of Groningen on the neighbourhood level. It is important to mention that both 'wijken' and 'buurten' in Dutch can be expressed as neighbourhoods. 'Buurten' are subdivided wijken as seen in figure 3. So when referring to a neighbourhood a 'buurt' is meant in this paper, in the other case it will be specified such as calling it "larger neighbourhood".

Figure 3 example of the difference between a 'wijk' and a 'buurt' (Basiskaart Groen en Grijs, *2019)*

There are multiple reasons for choosing the municipality of Groningen, the first one has to do with familiarity of the region to make sure that for example less mistakes will be made in determining public parks across the municipality. Another reason has to do with the ambition of the municipality of Groningen to remain green and even increase private and public green availability within the municipality (Gemeente Groningen, 2020).

Groningen is a growing municipality with a generally low mean age (Allecijfers.nl, 2023) this has to do with the fact that there is a school for higher education and a University where a lot of students study which results in a lower mean age. Furthermore, the municipality has a total of 20 larger neighbourhoods (wijken) and approximately 150 neighbourhoods (buurten). The liveability measurerer(leefbaarheidmeter) while generally indicating that living in the larger neighbourhoods tends to be good it does show differences in the general liveability between larger neighbourhoods (Lankhorst, 2020).

To collect the data, multiple sources were used of the year 2020 to make sure no unintentional differences between data sources will cause a type I or type II error. These type of errors mean that either a hypothesis will be rejected while it should be accepted and the other way around (Burt, Barber and Rigby, 2009).

The soil map (Statistiek, 2023a) was used in creating polygons of public parks using the field called 'park & plantsoen' (public park & park) but this 'layer' did need some changes to accommodate the RQ's. Using literature to accommodate these changes, all public parks indicated in the soil map that were smaller than 1,5 hectares (HA) were omitted by using the GIS tool which measures the size of a polygon in HA. The reasoning for this is that Cohen *et al.*, (2010) found that public park usage grew with an increasing size. Koohsari *et al.*, (2018) also mentioned that people with access to a public open space larger than 1,5HA were more likely to actually make use of the public park for different forms of recreation such as walking.

As the qualitative aspects of a public park also plays a role in the usage and enjoyment of the public park (Bowler *et al.*, 2010; Sharifi *et al.*, 2021) such as roads to run on. But not all places had such qualitative characteristics. So because of that, sloping lines across a waterbed such as seen in image 1 were manually removed as these 'green spaces' stretched long enough to be longer than 1,5HA but could not be properly used in comparison to a regular park. As it was hard to be accessed due to the lack of a walking path and cars in front of it such as seen in image 1. The removal of these locations were done using both a 2020 map through Google Earth's terrain map, and also using a GIS 2020 map of terrain. The combination of using both maps tried to reduce human errors in removing green spaces. From the central bureau of statistics (CBS) the "Wijk & buurt statistieken" (Statistiek, 2023c) which are the more general neighbourhoods statistics were used. Secondly, data from the Health monitor for adults and the elderly was used to get data on (mental) health factors (RIVM, 2021). Another source that was used is data from the Klimaateffectatlas (*Basiskaart Groen en Grijs*, 2019) which calculated the total percentage public green in each neighbourhood.

The statistics program that will run the statistical analysis is STATA 17.

3.3 Ethical considerations

All secondary data that has been collected that is about personal information of people has been gathered by governmental bodies whom have a high code regarding privacy (Statistiek, 2023b; *Onze visie op privacy | RIVM*, 2023). Furthermore, due to the data being aggregated it is very hard to link certain information towards an individual. But, this still means that the data gathered should be handled carefully. For this paper, guidelines as mentioned by (Punch, 2013) are important for further ethical deliberation for the duration of writing the paper. That is why the code of conduct will be followed from the University of Groningen to minimize any ethical misjudgement for example that data between different neighbourhoods will be carefully handled to reduce the risk of mismanaging the data while combining all datasets in one Excel sheet (*Ethiek vóór je onderzoek*, 2021). It is important not to just blankly follow a set of rules but as a researcher develop a personal understanding of the ethical principles and contexts (Punch, 2013). Such as carefully handling the data as

mentioned above but also substantiating decisions made related to the data such as what variable to use or not to use in a proper way.

3.4 Data Variables

The variables that were included are subdivided into 3 sections. The main outcome variables (dependent variable), explanatory variables (independent variables) and the control variables.

3.4.1 Outcome variables

To measure the subjective well-being of neighbourhoods two measures were used to make sure the subjective well-being is well captured. The first outcome variable is the percentage of people that mentioned that they had stress or a lot of stress in the past 4 weeks opposed to having no stress or only a little bit of stress in the past 4 weeks (Rijksoverheid, 2023). The second outcome variable was a 10 question sub-survey from the Health monitor 2020 which asks questions related to depression & anxiety these questions can be found in appendix B for further clarification. The variable reports the percentage of people that have mediocre to high risk of depression or anxiety opposed to low or no risk. The combination of both variables will give a better insight into the subjective well-being of people in each neighbourhood for the final discussion in relation to the literature. In the results section both variables will be referred as ‘stress’ and ‘depression / ‘anxiety’ to make the text more comprehensible.

3.4.2 Explanatory variables

The main explanatory variables are the average distance between a neighbourhood & the closest public park and percentage of public green in each neighbourhood.

Distance

The average distance between the public park and neighbourhood were determined using the average Euclidean distance from the centre of each individual neighbourhood which was calculated by ARCGIS PRO towards the closest border of the nearest public park which is the outline of the polygon. Using ARCGIS PRO its computing functions, (the exact steps and calculations done by the program can be found in Appendix B) were used to determine the average distance of each neighbourhood to a public park as precise as possible.

It should be noted though that this way of determining the average distance from a neighbourhood towards a public park is prone to bias. The determination of the central point of the neighbourhood does pose the risk that by chance in some parts of the neighbourhood a public park is really close while in other parts of the neighbourhood no public park is close. This bias is especially the case for larger neighbourhoods

Public green

The second explanatory variable is the percentage of public green per neighbourhood. What exactly is meant by this is to what degree there is green in the neighbourhood that is accessible by the public as seen in figure 4 and 5 The reason for why this is chosen instead of total percentage green per neighbourhood has to do with the fact that green that lays in a private area or on roofs cannot be observed by the public.

*Figure 4 differentiation of 1 (high green) 2 (Gray space) 3 (Low green) (*Basiskaart Groen en Grijs*, 2019)*

*Figure 5 differentiation of pink (public area) orange (private area) (*Basiskaart Groen en Grijs*, 2019) (blue is water, red are houses)*

3.4.3 Control variables

The control variables can be subdivided into 3 groups, socio-economic determinants (income & work participation), social determinants (feelings of loneliness) & socio-demographic determinants (education level, gender & age).. Diener and Ryan, (2009) found these variables to be most important in defining someone their subjective well-being. While being religious or not also has an impact on the subjective well-being of individuals according to the same author, due to data limitations this variable cannot be taken into account

Social determinants

Literature suggests that having fewer contacts and feeling lonely in general has a negative impact on the subjective well-being (Diener and Ryan, 2009). This variable measures the percentage people that feel lonely and takes the average for every neighbourhood.

Socio-economic determinants

These socio-economic determinants are seen by literature as the most influential in explaining people their subjective well-being (Diener and Ryan, 2009). These are the yearly total income per individual living in the neighbourhood, because the data is aggregated the average per neighbourhood is calculated.

Related to this is the percentage work participation per neighbourhood, this variable explains the percentage of people that currently work which is either being in employment or being self-employed. An limitation of this variable is that it does not make a distinction between people who are not-employed by choice (such as taking care of children) and people who did not choose to be unemployed (often still looking for a job). Both options have a different effect on the subjective well-being of the individual the latter in this case having often a worse effect on a person due to the economic hardship of being unemployed without choosing for it (Masarik and Conger, 2017)

Socio-demographic determinants

The education level of people are divided into three groups which are; low education, average education & high education. Table 1 specifies the education levels and the corresponding education in the Dutch system.

Lower education		**Average education**		**Higher education**	
Primary education	Basisschool	Secondary vocational education 2, 3 & 4	MBO 2, 3 & 4	Higher vocational education	HBO
prevocational education	VMBO	Year 4, 5 & 6 of general secondary education	HAVO jaren 4 & 5 / VWO jaren 4, 5 & 6	University education	WO
Year 1 – 3 general secondary education	Eerste 3 jaren HAVO / VWO				
Secondary vocational education	MBO 1				

Table 1 Education level explanation with their Dutch counterparts (Statistiek, 2023b)

Both the average percentage males & females are taken per neighbourhood. Furthermore, the age division is also given per neighbourhood. For this paper age is subdivided into three groups which are; ages 18 – 39, ages 40 to 64 & ages 65 and above to make the age groups relatively the same width, year wise.

3.5 Data preparations

As mentioned, there are 150 neighbourhoods in the original sample of the municipality of Groningen. Haren, Ten Boer & lower part of Oost have been dropped due to a combination of reasons. The first reason has to do with the fact that these areas are less dense in comparison to (most) of the other regions. Secondly, while the omitted areas may have less public parks which increases the distance between public park and neighbourhood. The area is more rich of natural areas such as forests or farmland which are for the purpose of this study omitted as the focus is on public parks as one of the main research questions. But these areas still may have a positive or negative impact on a person his/her subjective well-being. Because of that there is a risk of a type I or II error as not all possibly important factors are taken into account for the remaining samples in the data. For a list of what neighbourhoods are taken into account Appendix A can be consulted.

The next step in the data preparation process was to exclude all neighbourhoods that were either really small with less than 100 inhabitants or did not have data on 1 or multiple variables that will be included in the final model. The reasoning for omitting these neighbourhoods is because the full model is able to factor in all necessary variables that have either a positive or negative impact on the outcome variable. If one neighbourhood is included that does not have one of the key control variables, it would be automatically omitted from the full model and as a result would still not be taken into account.

3.6 Analytical strategy

3.6.1 Overarching statistical technique

A multivariate regression will be used to look for possible relationships between the two outcome variables and the explanatory variables while controlling for: social, socio-economic & socio-demographic factors. The multivariate regression has been chosen for its unique position in being able to run two outcome variables at the same time e.g. a joint regression. Thus it is similar to a regular regression a both have the same individual coefficients & standard errors.

3.6.2 Test for regression assumptions

To legitimately perform a regression, the final model needs to adhere to a number of assumptions as mentioned by Mehmetoglu and Jakobsen, (2016).

Testing whether there is a linear relationship

To test whether there is a linear relationship between the outcome & explanatory variable (Mehmetoglu and Jakobsen, 2016) suggest running a pairwise correlation between the outcome & explanatory variable. These are both insignificant with the p-value between stress and distance being 0,51 & for depression / anxiety and distance a p-value of 0,69. Additionally, a non-parametric test called the spearman rho is also performed as the sample size is relatively small. This shows again an insignificant outcome with a p-value of 0,21 for stress and distance and a p-value of 0,28 for depression / anxiety risk and distance. What these outcomes can entail is that there does not seem to be a linear relationship between the outcome & explanatory variable. A final way of confirming what type the linear relationship is, can be done by graphing both outcome variables which can be seen below.

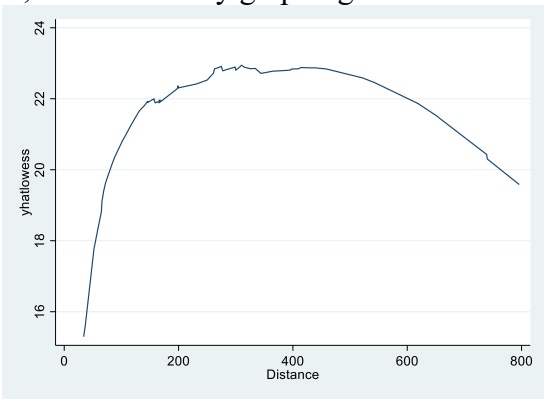

Figure 6 Relation between stress in the past 4 weeks and distance

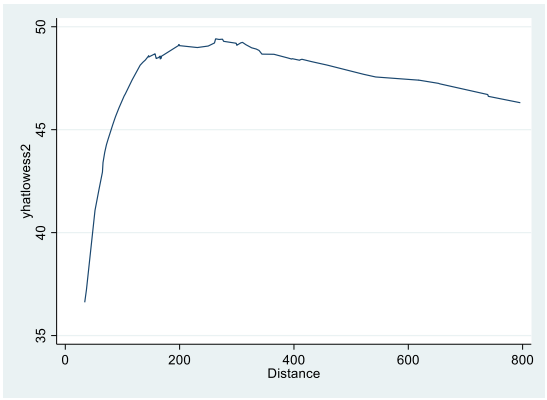

Figure 7 Relation between depression / anxiety risk and distance

Both figure 6 & 7 show that there seems to be relatively lower stress and depression / anxiety risk in the first 200 metres, then it increases but does lower again in when distance increases further. So looking at these figures there is no immediate straight linear relationship but a concave relationship (Mehmetoglu and Jakobsen, 2016).

Running the pairwise correlation for public green and the outcome variables, a significant outcome is observed with a p-value of both stress and depression / anxiety risk < 0,01. To strengthen this expectation the Spearman rho is once more used but also shows a significant outcome for public green with a p-value < 0,01. Meaning that a linear relationship between public green and the outcome variables is expected.

Testing null-model whether it improves with interaction

As the relationship between the outcome variables and the explanatory variable is not linear (Mehmetoglu and Jakobsen, 2016) suggest using a square equation in the model which is going to be used on the distance to a park variable to account for there being a concave relationship when running a regression. To see whether this fits the model, the model should improve the explanatory power when the square equation is added. After running this assumption the interaction shows an improvement of the R-square from the original 0,0068 & 0,0024 towards 0,0514 & 0,0295 for stress and depression / anxiety respectively and thus the square equation can be added to the model.

Testing multicollinearity and homoscedasticity

A model should not have high levels of both multicollinearity and homoscedasticity as this may have a negative effect whether it can explain something about the population after running the regression. As a general rule of thumb multicollinearity should not exceed 10 but it preferably it should be rather as low as possible. When testing for multicollinearity the multicollinearity of distance remains equal to 1,35. But other control variables such as being highly educated have a much higher multicollinearity > 10. Allison, (2012) argues that this can be safely ignored as it is only important for the variables of interest (distance to a public park) and because it equals to 1,35, the regression can be run properly. The same holds true for public green, while a little bit higher than distance at a 2,25 it can still be safely assumed that there is no multicollinearity

Homoscedasticity is the assumption that the error term has constant variance. When running both outcome variables separately using a regression both outcomes are insignificant at a p-value of 0,17 for stress and a p-value of 0,57 for depression / anxiety risk. As both are insignificant it means that the error term does have constant variance and thus there is now problem with homoscedasticity according to (Mehmetoglu and Jakobsen, 2016).
The final test which needs to be taken a look at is whether the distributed errors are normal. This can be done performing a test for skewness and kurtosis. The test outcome is insignificant at a p-value of 0,978 which means that the residuals are not different from a normal distribution and as a result the assumption holds to perform a multivariate regression (Mehmetoglu and Jakobsen, 2016).

To perform a multivariate regression, all assumptions should hold. Which is the case, the only difference from the more common linear regression is that the relation is not straight but a concave relationship which is accounted for by having the square root of distance in the model.

Testing the improvement of the model

The different models are ran using an multivariate regression thus using both the outcome variables at the same time where A is Stress and B is Depression / anxiety which while in the same regression are being calculated separately. The first model which in this case is called the null model is only the outcome variable and the explanatory variable. The explanatory variable is the squared outcome of distance as it is assumed that there is not necessarily a linear relationship as mentioned above In table 2 a summary of the variables in each model can be found. But the general way how each model has been determined is by adding a group of control factors such as socio-demographic factors. What can be seen in is that there is for each model building upon itself a significant improvement per additional model. What is furthermore of interest is that in the null model, stress has a higher explanatory power in the first 2 models while this changes when there are additional control variables added but is equalised again in the final model. This R-squared in the final mode

which has an explanatory power of 98% is really high according to (Mehmetoglu and Jakobsen, 2016). But the same author mentions that the addition of variables always improves the R-squared and thus it can be important to also look at the adjusted R-squared to find out whether the cause is the addition of the variable and not the variable itself.
An multivariate regression is unable to do this. Thus two separate multi-linear regression are performed. In table 4, the outcomes of the adjusted R-squared can be seen which still shows a high explanatory power of 98% for the full model.

Model variable specifications					
	Null-model	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Overarching theme	Base	Physical	Social	Socio-economic	Socio-demographic
Variables included	Depression & anxiety risk	Public green	Feeling lonely	Income	Gender
	Stress in past 4 weeks			Work participation	Age
	Distance-squared				Education

Table 2 The model variable specification where the blue cells = outcome variables, green cell = explanatory variable and orange cells = control variables

R-squared per model									
Null model		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
A	B	A	B	A	B	A	B	A	B
0.051	0.029	0.327	0.261	0.605	0.795	0.780	0.904	0.985	0.984

Table 3 The R-square for each model where A = 'stress in the past 4 weeks' and B = 'depression / anxiety risk'

Adjusted R-squared per model									
Model 1		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
A	B	A	B	A	B	A	B	A	B
0.021	-0.001	0.295	0.226	0.579	0.782	0.758	0.895	0.981	0.979

Table 4 The adjusted R-square for each model where A = 'stress in the past 4 weeks' and B = 'depression / anxiety risk'

3.6.3 Omitting variables

Because the data is aggregated, Statistical computations have the limitation that in the case two variables fully explain each other, one will be dropped due to dependency among the variables (e.g. males and females percentage per neighbourhood) (*FAQ: Estimation commands and omitted variables | Stata*, no date).
So for clarification to the reader a number of variables have been dropped to prevent that the statistical program drops a number of variables itself. The first variables that has been dropped is 'females', while it does not make a big difference as it is almost 50 / 50 distributed with the variable 'males' there are still slightly more males in Groningen thus it has been chosen to keep the biggest group in. Secondly, being averagely educated has been dropped to have a larger gap between being lower and higher educated. Finally, being 65 and over has been omitted as this group also has the lowest association with income and work participation due to the nature of their age period where often people above 65 are retired.
While these groups have been dropped which is a necessity, it does not mean that their influence on each individual neighbourhood is gone as the other variables such as income are still influenced by these groups.

3.7 Descriptives

In figure 8 the average distance to a public park can be seen. There does not seem to be a specific distribution to this. While in the east of the municipality the distance is a bit lower than the other parts of Groningen.

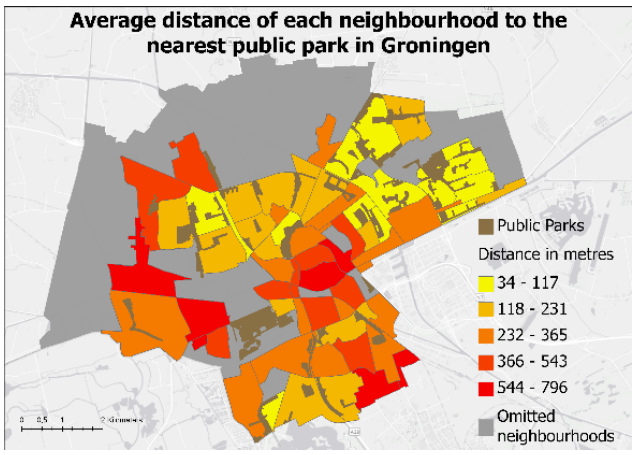

Figure 8 Average distance to a public park per neighbourhoods

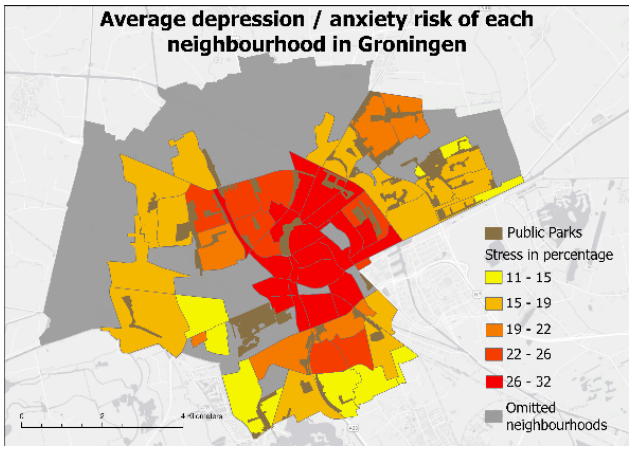

Figure 9 Depression / anxiety risk in Groningen

As data was collected by governmental institutions with their respective sampling strategies the socio-demographic data such as age & education can be seen as representative for each respective neighbourhood. While data on stress & depression have generally a representation of 40 to 100 on neighbourhood level and thus should be handled with more care when making statements about them. In this the variable feeling lonely is an exception as it is sampled through a different survey which is representative on neighbourhood level.

In table 5 the descriptives on the outcome variables show a relative higher average mean for all neighbourhoods compared to the average in The Netherlands which is 17,9% for stress & 45,9% for anxiety / depression (RIVM, 2021). While this average lies above the mean of the Netherlands one could look at both the lower bound & upper bound represented by Min & Max respectively which shows that there is quite a large gap between the different neighbourhoods. Figure 9 & 10 give a clear representation of how these stress levels and depression / anxiety levels are distributed through the municipality. Especially closer to the inner city there is a higher percentage for these two outcome variables.

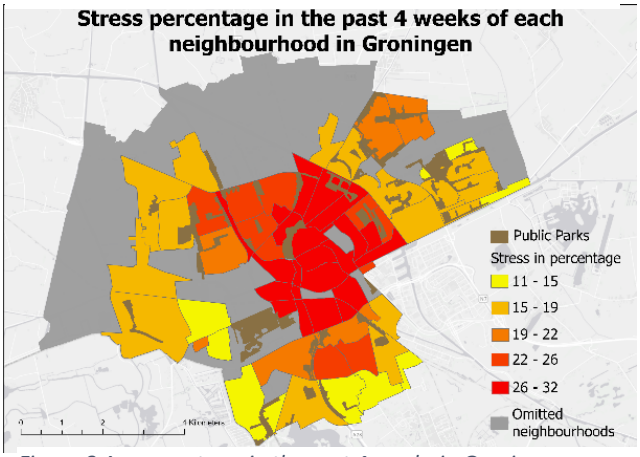

Figure 8 Average stress in the past 4 weeks in Groningen

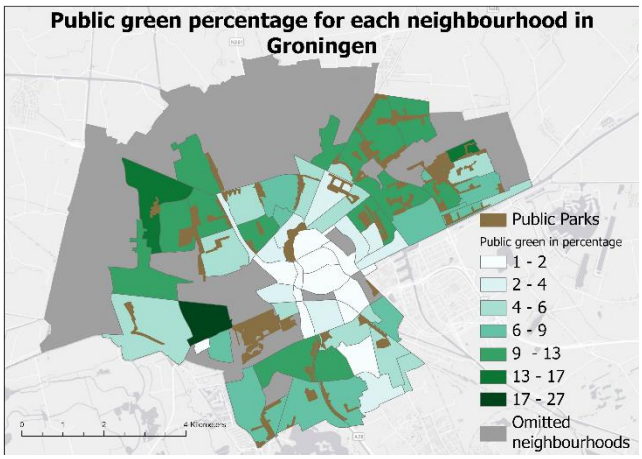

Figure 11 Public green percentage per neighbourhood in Groningen

Some further things that were of interest to look at are the large differences in the percentage of public green as seen in figure 11. The large differences in public green can be attributed by the generally denser inner city of Groningen where the percentages of available public green is lower. While in neighbourhoods that are further away from the inner city and less dense more room is available for green on the sidewalk and smaller urban green spaces. Also the age

structure is of interest as the city of Groningen is a student city with higher education and a university which attributes to the skewed age structure in comparison to other cities.

Comparing the mean scores of the neighbourhoods compared to the Dutch average, the outcome variables show a lower score for the Dutch average. There is also much less public green in the neighbourhoods compared to the Dutch average, but this mainly has to do with the rural area that contributes to that. Two final things that are of interest is as mentioned above about the age distribution that there are relatively a lot of younger people in the neighbourhoods compared to the Dutch average. Furthermore, also the mean of higher education level is quite a bit higher compared to the Dutch average. While income does lack behind compared to the Dutch mean average.

Descriptives	**N = 66**				**Dutch average**
Variable	Mean	Std. Dev.	Min	Max	Mean
Outcome variable					
Stress in past 4 weeks %	21,97	5,93	11,3	31,5	17,9
Depression Anxiety risk %	47,98	8,92	30,6	60,4	45,9
Explanatory variable					
Distance to public park	278	19	34	796	X
Physical determinant					
Public green %	6,71	4,63	0,6	27,19	53,19
Socio determinant					
Feeling lonely %	47,31	5,58	36,7	57,5	47
Socio-economic determinant					
Income X 1000	32,88	9,940	16,8	64,8	34,3
Work participation %	65,49	8,2	38	81	69,6
Socio-demographic determinant					
Male %	50,35	3,6	46,4	75,4	49,7
Female %	49,66	3,6	24,6	53,6	50 ,3
Lowly educated %	17,39	11,81	4	75	27,6
Averagely educated %	37,8	7,94	14	55	41,5
Highly educated %	44,77	14,33	0	75	30,9
Age 18 to 39 %	48,51	20,05	11,95	96,57	32,1
Age 40 to 64 %	34,89	14,46	2,36	64,95	43
Age 65+ %	16,59	9,66	1,07	44,76	24,9

Table 5 Descriptive statistics

3.8 Hypothesis

The expectations of the relations between the outcome, explanatory and control variables are as follow:

1. Larger distances to a park increases stress in the past 4 weeks and depression/anxiety risks
2. Higher percentage of public green decreases stress in the past 4 weeks and depression/anxiety risks
3. People with worse SES factors live further away from a public park
4. People with worse SES factors have less green in their neighbourhood

3.9 Decision-making on the representation of the data

Variables that are included as SES factors

(*Measuring Socio-economic Status and Subjective Social Status*, no date) mentions that individual measures of income, education and occupational social class the resource-based measures. As a consequence, the variables income, work participation, low education and high education were used to interact with the explanatory variables, while age, gender and loneliness were only used as control variables.

Average Marginal effects

The average marginal effects (AME) can be explained as the simple main conditional effects (Mehmetoglu and Jakobsen, 2016); how this is visualised by plotting the slope of a tangent line at y. In this case, this slope can be interpreted as the effects of x on y (*A Beginner's Guide to Marginal Effects | University of Virginia Library Research Data Services + Sciences*, no date).

Making this more concrete, this research investigates the interaction between SES factors: income, work participation, low education, high education and the explanatory variable distance. To make this interaction, the explanatory variable will be multiplied with the SES factor for each model separately to observe changes over the X-axis, which is the SES factor. The reason why this is needed is because distance seems to be non- (concave relationship). As a result, the effects may differ from neighbourhood to neighbourhood with different degrees of SES factors.

The AME will not be considered for public green as an explanatory variable, as previously stated that public green had a p-value < than 0,01 when running the Spearman rho and thus, it can be expected that the relation is linear.

3.10 Results

The results section will start investigating the combination of SES factors and the explanatory variables, next the first full model will be explained but focused on the constant and the control variables. Finally both distance and public green as explanatory variables will be analysed.

3.10.1 Effects of SES factors on the explanatory variables

Before delving into the full models, look at stress and depression/anxiety risk. The effects of the SES factors on the explanatory variables will be investigated. Two separate regressions will be running to examine the effect of the explanatory variables (distance & public green) seen below in Table 6. While only a set of SES factors will be explained. Other control factors are still included, such as age, gender and loneliness, as it is presumed they also influence distance and public green.

Distance

The full model is significant at a p-value of 0,036. The first SES factor, income, is insignificant at a p-value of 0,41, but for the sample, distance seems to increase by 7,5 meters per income increase of 1000 euros as an average per neighbourhood. Work participation is significant at a p-value of 0,04. An increasing work participation lowers the average distance per neighbourhood by 15 metres for every 1 per cent point increase in work participation. This shows that neighbourhoods with higher employment levels enjoy the benefit of living relatively closer to a public park.

Having a lower education also decreases the distance by around 1 metre for every per cent point increase in low education in a neighbourhood for the sample. But because the p-value equals 0,827, the hypothesis is rejected that the difference is not equal to 0 for distance. For high education, the outcome is also insignificant at a p-value of 0,5. But for the sample, the distance to a public park decreases by 4 metres for every increase of 1 per cent point in high education.

Public green

For public green, the p-value of the full model = < 0,01. Ages 18 – 39 contributed to having less access to public green in their neighbourhood. While income also significantly reduced public green, the higher the income in the neighbourhood, the less green there was. But it is important to note that this may be mediated by private green per neighbourhood. Having a higher income often results in more possibilities in owning a garden meaning that the actual green (accounting for green in gardens)

	Distance	**Public green**
Public green	-2.069	X
	(7.157)	
Distance	X	-0.001
		(0.002)
Feelings of loneliness	-24.072***	0.164
	(8.850)	(0.174)
Male	-2.185	0.190
	(8.931)	(0.165)
Low education	-1.102	-0.170*
	(5.019)	(0.091)
High education	-3.930	-0.029
	(5.789)	(0.108)
Income	7.589	-0.345**
	(9.145)	(0.165)
Work participation	-15.961***	0.179*
	(5.353)	(0.105)
Ages 18 - 39	4.195	-0.198***
	(4.116)	(0.073)
Ages 40 - 64	-1.973	0.127
	(4.814)	(0.088)
Constant	2,396.706***	-1.397
	(726.918)	(14.818)

Observations	66	66
R-squared	0.261	0.554
Standard errors in parentheses		
*** p<0.01, ** p<0.05, * p<0.1		

Table 6 Regression between distance to a public park & public green and the control variables

3.10.2 Investigating the effect of the control / SES variables & Constant

While not the focus of the study, there are also a number of variables that, according to the literature, have an impact on the subjective well-being of individuals. Looking at the multivariate regression model is seen in Table 7. Running the full model (Model 1), The most interesting outcome is that loneliness has a relatively high impact in comparison to income or education on both stress and depression/anxiety risk. Feelings of loneliness are significant at a p-value > 0,001 for both outcome variables. With every per cent point increase of reporting feeling lonely, stress increased by 0,3%. For depression/anxiety risk, this was more than double, with an increase of 0,7% on the neighbourhood level. This outcome was expected as the literature mentions that having few social ties and feelings of loneliness generally worsens people's subjective well-being (Diener and Ryan, 2009). But it was interesting that the impact of loneliness is relatively strong compared to other SES factors that are investigated more centrally in this paper.

As mentioned, the variable 'females' and the age of 65+ were omitted. Starting with gender, being a male shows a different sign for the outcome variable as with stress, the coefficient is slightly positive with an outcome of 0,016, while for depression/anxiety risk, the coefficient is negative with 0,02. Nevertheless, both are also insignificant, so for the population, both variables are not proven to be different from 0 as the p-value equals 0,678 for stress and 0,746 for depression/anxiety risk. An explanation for this insignificant outcome is the small differences in male/female percentages present in each neighbourhood, as generally, the division is relatively the same towards 50/50 per neighbourhood. So due to the geographical limitation of the data, differences are harder to observe.

Significant outcomes have been found for 18 – 39-year-olds and 40 – 64-year-olds with a p-value < 0,001 for the different age groups within neighbourhoods. Both coefficients of ages 18 – 39 and 40 – 64 contribute to increasing stress and depression/anxiety risk. But this contribution is more than double for the younger ages in both stress and depression/anxiety risk; this shows that stress and depression/anxiety risk are detrimental the younger an adults age is. As with the gender variable, the ages 65+ were also omitted. Looking back at the averages in the descriptive statistics in Table 5, this could partially explain the higher percentage of stress reporting and depression/anxiety risk as there are more youngsters in the city compared to the average in The Netherlands, which are more prone to these negative mental health indicators.

The way the constant, which is included in all models, can be interpreted is that the constant accounts as a base for the effects of factors that have not been included. While the R2 and adjusted R2 are both relatively high, factors still contribute to explaining stress reports and depression/anxiety risks (Mehmetoglu and Jakobsen, 2016). By including this, the potential errors are reduced.

Model 1	Stress	Depression / Anxiety risk
Distance	-0.000099	0.000985
	(0.002077)	(0.003225)
Distance2	0.000001	-0.000001
	(0.000003)	(0.000004)
Public green	-0.070962**	-0.140309***
	(0.031722)	(0.049245)
Feelings of loneliness	0.303392***	0.705272***
	(0.046209)	(0.071735)
Male	0.016538	-0.019966
	(0.039556)	(0.061406)
Low education	-0.051289**	-0.017726
	(0.022621)	(0.035117)
High education	0.014799	0.038105
	(0.025726)	(0.039937)
Income	-0.007026	-0.162585**
	(0.040743)	(0.063249)
Work participation	-0.067201**	-0.068461
	(0.027259)	(0.042317)
Ages 18 – 39	0.310096***	0.282572***
	(0.018495)	(0.028712)
Ages 40 – 64	0.148423***	0.112162***
	(0.021345)	(0.033135)
Constant	-8.159904**	7.217868
	(3.721708)	(5.777539)
Observations	66	66
R-squared	0.9852	0.9843
Standard errors in parentheses		
***** p<0.01, ** p<0.05, * p<0.1**		

Table 7 Full base model including all the control variables

3.10.3 Models with distance & interaction effects

While all models below in Table 8 are significant with a P-value lower than 0,05. There does not seem to significantly impact either distance or distance² for either outcome variable in model 1. Even for the sample, the outcomes are rather low and do not seem to have a high percentual change in relation to stress in the past four weeks or depression and anxiety risk.

	Model 1	Model 2	Model 3	Model 4	Model 5
Average/highly stressed in the past four weeks					
Distance	-0.000099	-0.000197	0.001177	0.001066	-0.002296
	(0.002077)	(0.002426)	(0.007075)	(0.001954)	(0.002523)
Distance²	0.000001	0.000001	0.000002	0.000001	0.000001
	(0.000003)	(0.000003)	(0.000004)	(0.000003)	(0.000003)
Distance X Income	X	0.000002	X	X	X

		(0.000039)			
Distance X Work participation	x	X	-0.000030	x	X
			(0.000082)		
Distance X Low education	x	x	X	-0.000072***	X
				(0.000022)	
Distance X High Education	x	x	x	X	0.000050
					(0.000034)
Moderate/high risk of depression or anxiety					
Distance	0.000985	0.002489	0.003610	0.001473	-0.001433
	(0.003225)	(0.003955)	(0.010661)	(0.003129)	(0.003978)
Distance²	-0.000001	-0.000001	0.000000	-0.000000	-0.000001
	(0.000004)	(0.000005)	(0.000006)	(0.000004)	(0.000004)
Distance X Income	x	-0.000062	x	x	X
		(0.000064)			
Distance X Work participation	x	X	-0.000047	x	X
			(0.000123)		
Distance X Low education	x	x	x	-0.000046	X
				(0.000035)	
Distance X High Education	x	x	x	X	0.000054
					(0.000054)
Observations	66	66	66	66	66
R-squared	0,98	0,98	0,98	0,98	0,98
P-Value: *=0,1 **=0,05 *=0,001**					

Table 8 shows all models related to distance and their respective interactions per different models.

There is not a linear relationship as tested above. The marginal effects of distance may give an insight into calculating the effect of distance on the well-being of the sample (Graubard and Korn, 1999).

Looking at Figure 12, only a small increase or decrease over the distance can be seen for both stress in the past 4 weeks, which increases over distance, and a small decrease for depression/anxiety risk. The outcomes of this analysis are insignificant, which means that the impact on both variables does not differ from 0 in the population; thus, no difference between the distance to a park and either stress or depression/anxiety risk is observed even though a concave relationship was expected.

This outcome is representative of the sample, though, and how to interpret this is that for every unit increase of distance, people on average who live 200 metres away have their stress level increased by 0,000331. After multiplying, people who live 200 metres away have a 0,07 per cent point increase in their stress level.

As can be seen, even if the outcome would have been significant, the explanatory variable's impact on the outcome variables is very low and probably would be unobservable in the real world.

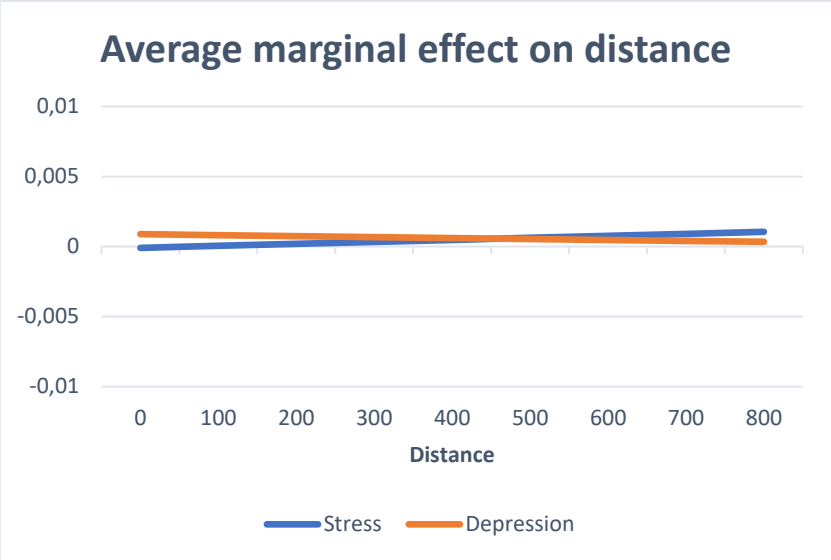

Figure 12: Average marginal effect of distance

So the hypothesis is rejected, and no difference in distance to a park and the impact it has on either stress in the past four weeks or risk of depression/anxiety are observed within the sample.

3.10.3.1 Socio-economic interactions

In the second model, distance and income were given an interaction effect to determine whether income as a socio-economic factor has a different impact in relation to distance to a park e.g. not linear but concave. Starting with the coefficient in the regression table (table 8), there is an insignificant outcome for both outcome variables. The difference is that for depression/anxiety risk, there is a relatively higher downward going slope, while for stress, it remains relatively the same. But as both are insignificant, nothing can be said about the population.

The marginal effects of distance and income on the outcome variables show its impact on the different income groups, as shown on the X-axis in Figure 13. Again, The outcome is insignificant, just like the regression table (table 8) but shows two different patterns. The pattern of depression/anxiety risk shows a rather small but expected outcome related to the second hypothesis. This shows that groups with lower yearly income have an increase in their outcome variable scores which is worse. When income increases, living further away gives a decrease in the depression/anxiety variable scores. But just like the previous margins, this impact is rather small. As with every unit increase of distance, people on average with income of 15.000 have an increase of 0,0015 per cent points. A household within the sample with an income of 15.000 would live 400 metres away. The impact this has on the depression/anxiety risk would be an increase of 0,6 per cent point, which is, again, relatively low. Meaning that the second hypothesis is also rejected as there does not seem that income with distance has a relation as the difference does not seem different to zero for the whole population.

The third model shows the interaction between work participation and distance. This is a negative coefficient in the regression table (table 8), so a downward relationship is expected. The variable is again insignificant for both outcome variables, so nothing can be said about the population. Below, the marginal effects (figure 14) are shown; these are insignificant but do show a downward relation to work participation. Depression/anxiety risk shows a more steep slope. This shows that the sample neighbourhoods with lower work participation are impacted harsher by the distance the neighbourhoods are from a public park compared to neighbourhoods with high levels of work participation. While the slope is steeper for

depression/anxiety risk compared to stress. The impact it has on high work participation levels is almost zero for both.

As the marginal effects are insignificant, nothing can be mentioned for the whole population, and thus, the this hypothesis is also rejected as the relation between distance, and work participation does not seem to differ from zero.

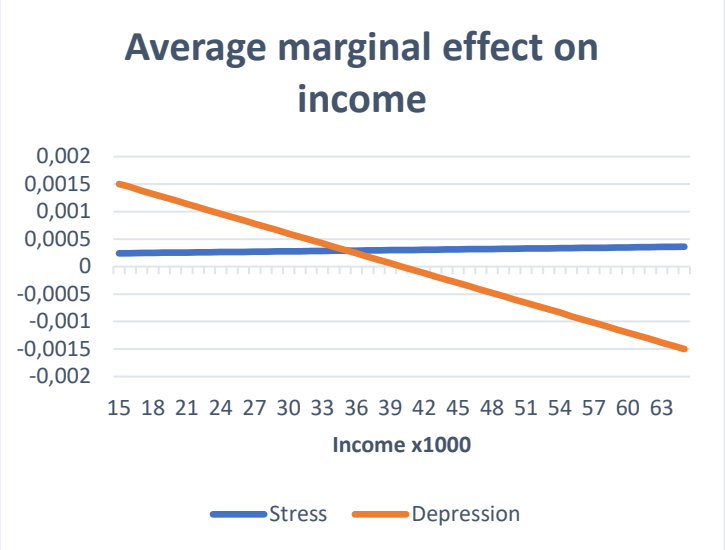

Figure 13: Average marginal effect of an increase in income and distance

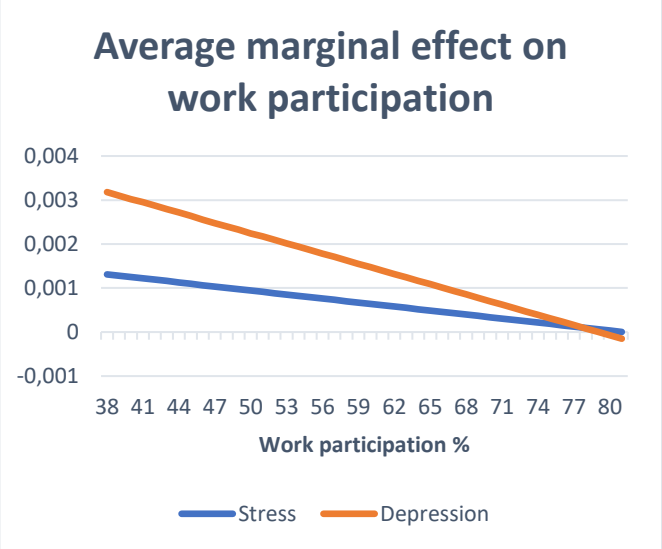

Figure 14: Average marginal effect of an increase in work participation and distance

3.10.3.2 Socio-demographic interactions

Models 4 and 5 are both related to the education levels per neighbourhood. Model 4 shows the interaction effect of low education and distance. This outcome is significant at the p-value of 0,01 (table 8), which shows a decreasing impact between the interaction effect of distance & low education. To further investigate this relation, the average marginal effects can be observed in Figure 15 below. The outcomes are partially significant at the p-value of at least 0,05 on the x-axis for 4% and from 55% onwards to 73%. This shows that as low education in a neighbourhood increases, distance's impact on stress and depression/anxiety becomes lower. But as mentioned, this decrease is relatively low for all the average marginal effects, and thus, the impact, while significant, is rather small.

Model 5 shows the interaction between distance and high education.
The regression (table 8) shows similar to low education, an increase in stress and depression/anxiety, but it is insignificant. So nothing can be said about the population.
Looking into the average marginal effects (figure 16) of this interaction shows an opposite trend in comparison to low education, where neighbourhoods with lower percentages of high education have reduced levels of stress and depression/anxiety. While it increases with higher percentages of education in neighbourhoods.

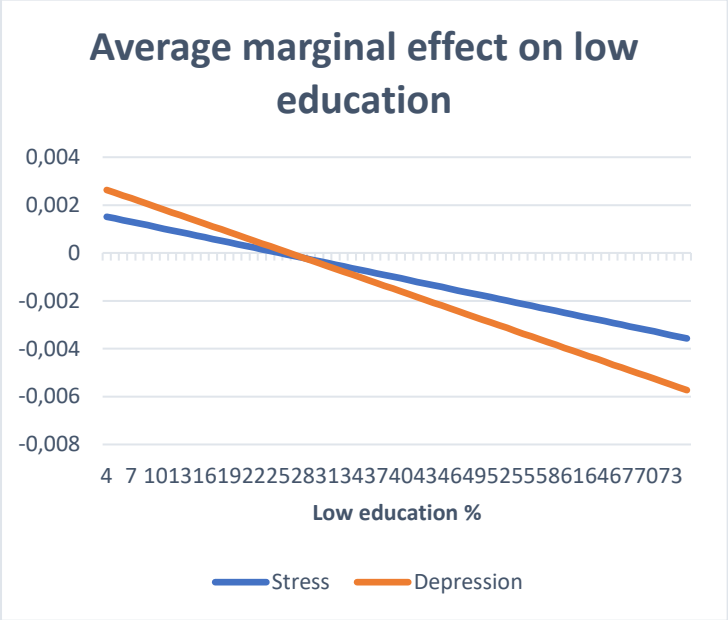

Figure 15: Average marginal effect on low education and distance

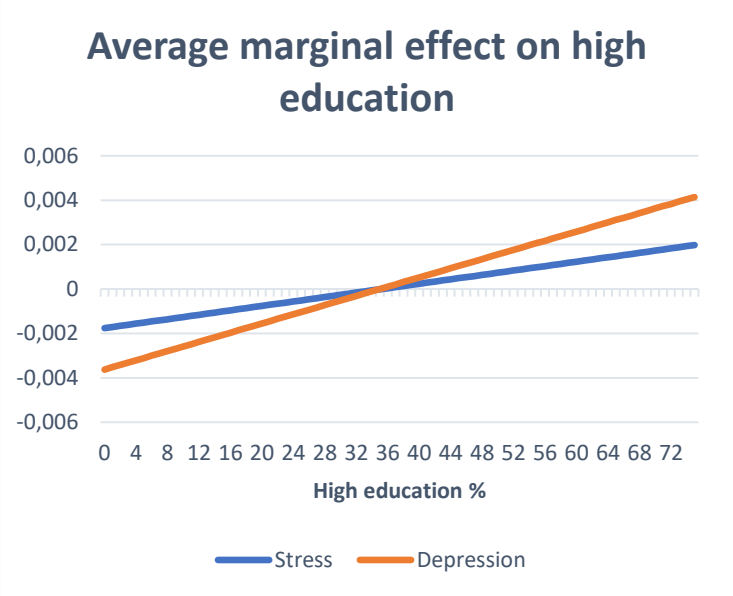

Figure 16: Average marginal effect on high education and distance

But this outcome for the whole model is insignificant.

The hypothesis can be rejected in two parts. The first part is related to the significance, which is partly for low education and fully insignificant for high education. Secondly, it was expected that the outcome variables would decrease in higher-educated neighbourhoods, and for lower-educated neighbourhoods, it would increase. This opposite is apparent for the sample, as higher-educated neighbourhoods have worse outcomes than neighbourhoods with higher percentages of lower education. So because of these two facts, the final hypothesis can be rejected.

3.11 Model with public green and outcome variables

The percentage of public green positively impacts both the stress and depression/anxiety risk with a p-value that is for both lower than a p-value of 0,05 (table 9). For each percentage of green in a neighbourhood, the stress levels are reduced by 0,07 per cent. This is 0,14 for Depression/anxiety risk. This shows that public green has a positive impact in reducing these two mental well-being factors. Putting this in perspective, the mean public green in the neighbourhoods

	Stress	Depression / Anxiety risk
Public green	-0.070690**	-0.140711***
	(0.031432)	(0.048791)
Observations	66	66
R-squared	0.985200	0.984300
Standard errors in parentheses		
***** p<0.01, ** p<0.05, * p<0.1**		

Table 9 Regression of public green with the outcome variables of stress and depression/anxiety risk

of Groningen is rounded up to 7%, so when taking the average of public green, there is a reduction of 0,5% of people reporting stress and around 1% of people that are at risk for depression/anxiety. But this can be higher or lower as this is the average. The full model, including the control variables, can be found in Appendix A

3.12 Concluding results

Looking at the relationship between the SES factors and the explanatory variables, work participation has a significant impact in lowering the distance to a public park with an increase of work participation in the neighbourhood. While public green income has a significant lowering effect, meaning that as a neighbourhood's average income is higher, there seems to be less public green available.
Further, all other SES factors are insignificant. This is in contrast to what was expected. This means that the hypothesis on SES factors and distance / public green must be rejected for the most part. Only for the case of work participation, there is a positive association with distance. Even though income was significant for the case of public green, the outcome was the opposite of what was expected, as it was expected that higher-income neighbourhoods had more public green. But the data shows a contrasting outcome where increasing neighbourhood income had less public green.

When running the first full model using stress and depression/anxiety risk as the outcome variables, it seems that loneliness had a very strong impact on both stress and depression/anxiety risk, with it being a significantly higher impact than the other control variables. Further, the percentage of age groups for each neighbourhood showed that the younger age group (18 to 39) had a worse degree of increase per cent in comparison to 40 to 64-year-olds. Meaning that this outcome may suggest that youngsters have worse stress and depression/anxiety risk as a base in comparison to older adults.

Looking at distance as the explanatory variable and including the AME as the relation between distance and stress and depression/anxiety risk was not linear but concave.
The main results show that there does not seem to be any significant outcomes for distance in relation to either stress or depression/anxiety risk. Furthermore, also the hypothesised relation using the AME of income, work participation, and high education did not have any significant outcome. Only low education showed a decrease in stress when interacting with distance. But this outcome, just like the other outcomes, is relatively close to 0, meaning that the actual relative impact of low education on stress may
have a small impact. This insignificant interaction of the AME was expected as the relation between distance and the SES factors were almost all insignificant. But it was still of importance to measure to investigate as if the outcome would be significant, there was possibly a concave relationship as distance was non-linear.

All hypothesis except for distance interaction with low education in relation to stress can be rejected. When comparing this to the relationship between SES factors and distance as the explanatory variable. But still no concave relation was observed in the AME for low education's interaction with distance. Thus, it is of importance that this result should be carefully considered.

Looking at public green, this variable did show a decrease in stress and depression/anxiety when there was an increase in public green available in a neighbourhood. This means that the hypothesis can be accepted.

Comparing both distance and public green. Public green seems to be the better variable to explain how 'green' can have a stress and depression / anxiety-reducing effect in the neighbourhood. While distance to a public park does not seem to matter too much in the case study of Groningen. The reasoning why this may be the case will be discussed in the next section while comparing it to current literature.

4.0 Discussion and conclusion

4.1 Summary of the findings

The objectives of this research were to look at whether there was a link between the distance people lived from a public park and whether this affects their mental well-being. Secondly, the link between public green and mental well-being was investigated. Finally, the socio-economic factors were considered to see whether there is a relationship with either distance of public green in the municipality of Groningen. To achieve this, a quantitative research method was set up to answer the previously stated questions. In order to answer it, multiple data sources were obtained to find all the right variables. This included aggregated social/demographic/economic neighbourhood data related to the people living there and data on the physical environment, e.g. mean distance to the closest public park and available public green per neighbourhood in the municipality of Groningen.

4.2 Distance as an explanatory variable

The outcomes of the analysis only partially support the conceptual model proposed. The distance/proximity to the closest park seems to have no relationship with stress or depression/anxiety risk (as the subjective well-being measure). This is in line with current literature as (Rigolon, 2016b) did not find any clear outcomes related to the proximity to a public park. Bertram and Rehdanz (2015) mentioned that there was an increase in life satisfaction if enough 'quantity' of a public green is close to a household, showing that proximity is important for their case study. But in their case, the 'buffer' was 1km. The highest distance to a public park in Groningen is approximately 800 metres. The mean distance to a public park is even lower at 278 metres. It may mean that in the case of Groningen, the threshold to access a public park is low enough that there is no difference between the different neighbourhoods. So, answering the first sub-research question, the proximity to urban green spaces does not affect individuals' subjective well-being. But it may impact this if the average distance becomes greater such as in the research of (Bertram and Rehdanz, 2015).

The relationship between distance and the SES factors was investigated. The outcome was that only work participation showed a lowering effect on distance, e.g. as work participation grew in a neighbourhood, the average distance to a public park decreased. Current literature has not looked into this possible relation as it is more often compared to the income of households or SES as whole instead of whether a person is employed / self-employed or not employed. In neighbourhoods with lower work participation, a relation with average income would be expected, but this is not the case. Furthermore, the data limitations may hurt the explanatory power of work participation as the average of a neighbourhood may still leave very large differences within a neighbourhood. As distance had a concave relationship with stress and depression/anxiety risk. The AME may give an insight into the effects of distance over the percentage of work participation. This outcome is partially significant; the difference between the different percentages of work participation does not show large differences even when significant. So it can be concluded that the real-world impact between the different degrees of work participation per neighbourhood is rather small, e.g. no large outliers between work participation percentages in neighbourhoods.

Going to income, the outcome of this paper's statistical analysis, which was that there was no relationship between distance to a public park and income, is in contrast to existing literature. De Vries, Buijs and Snep (2020) suggest that lower-income people had less access to a public park than higher-income groups. The interaction effect of income and distance in relation to stress and depression/anxiety did not show any significant outcomes. While in literature, this interaction has not been researched yet, this outcome was expected because income already had no significant relation with distance. But as de Vries, Buijs and Snep

(2020) suggest a relation for income, an assumption will be made that this outcome of literature can be related to the significant outcome for work participation. Higher degrees of work participation showed that neighbourhoods were closer to a public park. The reason why there is a difference in the income variable and the research of de Vries, Buijs and Snep (2020) may have to do that in their paper; research was done on a lower geographical level being more accurate while this research was done on a neighbourhood level with averages which are more prone to differences within a neighbourhood.

Both low and high education do not have a linear relationship with distance to a public park. and as a result there is no difference in how close a park is and whether people have low or high education. Literature has not looked into this association, meaning that it is difficult to assess whether there is no relationship between distance and education level. The AME of the interaction between distance and low education levels in relation to stress and depression/anxiety risk seemed to be partially significant. However, this percentage point change that the AME show is low, meaning that the impact of this interaction would not be noteworthy. The AME for high education was fully insignificant. Existing literature has not focused on this interaction related to the distance between a public park and education either. So while the expectation is that this paper's outcome is accurate, even with the limitations of the data, it is hard to compare.

While the SES factors related to education have not been looked into separately by current literature, the assumption that education is unrelated to distance to a public park should be assumed carefully, especially because average education level has also been removed due to data limitations. But this paper does show the first signs that there is no difference in distance between a public park and education level.

The result for income, as it showed no difference, should be carefully examined and is somewhat presumed wrong as the literature suggests a difference between income levels and the data limitations of this research. This is the same for work participation, in the sense that the outcome should be carefully interpreted. But as income in literature is significant, an assumption can be made that the outcome for work participation is the same as, to a certain extent, it is related to income.

Relating this to the research question, education attainment may not influence the proximity someone lives from a public park, while income and work participation are more difficult to interpret. It is assumed that, to a certain extent, these SES factors influence how far someone lives from a public park. But this assumption should be further investigated.

4.3 Public green as an explanatory variable

Public green showed a significant outcome: a higher percentage of public green resulted in lower stress and depression/anxiety risk. This outcome is in line with the literature; first of all, this outcome confirms the restorative effects green can have, as mentioned by (Ulrich *et al.*, 1991; Grahn and Stigsdotter, 2003), which have a stress-reducing effect on an individual. And when there is a higher degree of this, there is a higher chance of coming into contact with the public green. Also, the increase in subjective well-being of the individual, which was tried to be captured using both stress and depression/anxiety risk, showed a similar outcome in relation to the literature. This positive effect of public green on a person's subjective well-being was confirmed by a large number of research (Maas *et al.*, 2006b; MacKerron and Mourato, 2013b; van den Berg *et al.*, 2016b; Sharifi *et al.*, 2021b).
This shows that public green is a significant predictor of subjective well-being, as found in this research by looking at both stress and depression/anxiety risk. Current literature also tends to agree that public green improves subjective well-being.

The relation between public green and income was significant but negative; as income increased, there would be less public green in the neighbourhood. This is in contrast to

existing literature that found the opposite to be true that there was more public green, the higher the relative income of households was (Astell-Burt *et al.*, 2014b; de Vries, Buijs and Snep, 2020). One partial reason for this difference may have to do that this paper focuses on public green, while (Astell-Burt *et al.,* 2014) investigated green in general as higher income groups do have easier access to green in, for example, their front- and backyard. But this does not explain the different results (de Vries, Buijs and Snep, 2020).

Looking at the public green and the relation towards work participation was insignificant in this paper. Literature only partially looked at work participation through a more overarching socio-economic position of neighbourhoods, which found that there was less green space available in worse-off neighbourhoods compared to better neighbourhoods (Schüle, Gabriel and Bolte, 2017). That is in contrast with the finding in this paper that did not find a relation, but this difference may appear as the socio-economic position in the paper of Schüle, Gabriel and Bolte (2017) took a wider look than only work participation, which was a factor in this which did also include income as a factor. Other factors such as data limitations may also influence work participation's insignificant effect on public green.

Literature suggests that higher education levels often have more urban green space than groups with lower education (Nesbitt *et al.*, 2019). This is in contrast with this paper, as there does not seem to be any significant relation between the different education levels. One reason for that may be that the geographical data (on the neighbourhood level) has a hard time capturing the individual differences between households and education levels. As the mean for each neighbourhood is taken. While another difference may have to do with the fact that not private green is taken into account, which may show different outcomes.
Both low and higher education did also not seem to have any significant relation to the percentage of public green in this paper, meaning that there does not seem to be a difference between these two educational groups. Research has not touched too much on this topic; thus, while the assumptions can be made that there is no difference. This outcome should be interpreted carefully.

Relating this to the research question, as the outcomes are very mixed, it is difficult to assume whether the socio-economic factors have a significant effect, even though the literature mentions that they probably do. Another study in Groningen may need to be done on a lower scale to make this difference between this research and the current literature. For education levels, it can be somewhat assumed that they do not have any effect, but as with distance, a better look at research on a lower geographical scale should be done.

4.4 Data limitations

4.4.1 Strengths

The data that was used to make the statistical analysis had a number of strengths and weaknesses. Starting with the strength of the data, the data that was used is publicly available, meaning that if another similar comparison wants to be made on the neighbourhood level, this can be rather easily reproduced for another city. Furthermore, the data can give a better overview of what general neighbourhood may benefit from more public green. Giving policy advisors an overview of what neighbourhood needs the most tending to improve access to green as green, especially in a neighbourhood, has beneficial effects on their inhabitants. Furthermore, the socio-economic, socio-demographic, physical and social factors that were taken into account factor in a lot of variances in how the subjective well-being of a person in that neighbourhood. The R-square & adjusted R-square in the statistical analysis shows a very high value.

4.4.2 Weaknesses

But the data also has several weaknesses that severely limit the statistical analysis and, as a consequence, the interpretation related to Groningen in this case. Starting with the geographical level of the data, it is difficult to make strong assumptions for a case study as there can be large differences between households or even streets. An example would be that the overall neighbourhood may have a large percentage of green or be relatively close to green. But the size may overestimate for one household and underestimate for another. In this study, as only neighbourhood data was available, there was no proper way to account for this difference within a neighbourhood. Furthermore, the use of the mean location of each neighbourhood and putting that in relation to the distance to the nearest neighbourhood also may over and underestimate the distance people may have to travel to reach a public park. Resulting in a possible bias, especially for larger neighbourhoods, as the minimum and maximum distance a person lives from the most central point of a neighbourhood becomes higher in comparison to smaller neighbourhoods. The Euclidian faces a similar problem, as it is not a realistic representation of how people will reach a public park.

Another limitation is the use of aggregated data; firstly, making assumptions using averaged data poses risks of overestimating the impact it has on the actual people living in the neighbourhood. Furthermore, this type of data, in comparison to household data, limits the uses of all variables. As mentioned before that, the data cannot run variables that fully explain each other such as males & females, and as such, one of these variables is dropped, losing possibly valuable data. Data on a household level are able to implement this resulting in a more accurate representation of the sample. Further, average data was used per neighbourhood. The total number of cases is far less in comparison if there was data available of households or individuals per neighbourhood. The risk of this data is that a lower number of cases in the sample may cause a non-normal distribution, while with more cases, this risk can be reduced even though the assumption that the data was normally distributed was upheld. Having a larger dataset is less risk aversion in comparison to having only the mean of each neighbourhood. With the presumption that there would be enough correspondents per neighbourhood.

While the variables that are most important in explaining the subjective well-being of individuals were used (Diener and Ryan, 2009b), there is a wide range of factors that may have a certain effect on a person, such as genes, personality, and life choices made (Ngamaba, 2017). Meaning that there is always a risk that there is over or underestimation per individual in a neighbourhood.

In this study,

Finally, the data is only from one singular year. It becomes more difficult to interpret the subjective well-being (stress and depression/anxiety risk) of individuals that answered the survey where the questions originated from. Longitudinal data, meaning data observed over multiple years, could more accurately represent the effect that public green & distance have on the data. As longitudinal data minimises possible outliers that are undiscovered in the data.

4.5 Implications for the built environment and further research

4.5.1 Further research

Starting with a number of research recommendations that can strengthen the proof or disproof of links found in this paper. First of all, research with data on household or individual level can greatly improve the accuracy of data for the municipality of Groningen, especially with the fact that this can reduce the risk of over and underestimating outcomes that were of risk in this paper. Adding to this recommendation, the benefit will be as well that all 'groups' such as all ages will be implemented instead of omitting some groups.

Furthermore, a second recommendation that can build on top of this research is to link the explanatory variables of green and distance with the SES factors in a more robust way. The explanatory variables, outcome variables and control variables have all been run in pairs. It would also be of interest to delve into the relative importance of the variable groups combined to give more accurate recommendations for policy advisors and planners. While this paper did not show a relationship between distance on both stress and depression/anxiety risk, the literature did suggest that there was a relationship. Implementing relatively larger distances, e.g. using a larger city or, as mentioned before, household data to increase the relative distance (as would have happened if not only the mean was taken). Additionally, for future research, household or individual data will be used. The research would also benefit from using real-life distances to a public park to make the statistical analysis most realistic.

Finally, as the same problem arose in (Sharifi *et al.*, 2021b), the determinants of what a public park entails were quite broad. Future research would benefit from having a clearer framework of what a public park exactly entails and why certain other parks that do not fit in this framework are excluded.

4.5.2 Policy recommendations

While distance to a public park did not seem to be a strong explanator in this analysis. A recommendation which was also emphasised by (Astell-Burt *et al.*, 2014b) is that while distance has no significant impact, this does not entail the same is true for other cities. Thus strategies developed by city planners and policymakers should emphasise the importance of the contextual need for public parks but also in the case of public green, which in Groningen is a better determinant.

For the case of Groningen, it is advised that public green between neighbourhoods should be better distributed to accommodate the positive effects public green has on well-being. Especially so with the need for additional dwellings already, public green should not be undervalued in the decision process by planners and the local government in implementing it.

As there are quite large differences between neighbourhoods, a lot of people can still benefit from urban green in their direct environment.

4.6 Conclusion

It is clear through both literature and this research that green is an important factor in a person's daily surroundings. As there is a growing need for more dwellings for a rising population, the importance of public green, whether it is in a small meadow, tree or a park, should not be underestimated. While it does not seem like the distance to a public park matters within the neighbourhoods of Groningen, this may be attributed to a minimum threshold to access a public park. But public green in itself does seem to be a strong predictor of the subjective well-being of individuals. Future research can improve the accuracy of the case study of Groningen when done on a lower geographical level. But this research does show a number of insights that can be used in policy advice and expanding future research.

5.0 Bibliography

Alberti, V. *et al.* (2019) *The Future of Cities*, *JRC Publications Repository*. Available at: https://doi.org/10.2760/375209.

Allecijfers.nl (2023) *Buurt Schoonvelde-West (gemeente Hoogeveen) in cijfers en grafieken (update 2023!)*, *AlleCijfers.nl*. Available at: https://allecijfers.nl/buurt/schoonvelde-west-hoogeveen/ (Accessed: 11 April 2023).

Allison, P. (2012) 'When Can You Safely Ignore Multicollinearity?', *Statistical Horizons*, 10 September. Available at: https://statisticalhorizons.com/multicollinearity/ (Accessed: 6 July 2023).

Angner, E. (2009) 'Subjective well-being', 39, pp. 361–368.

Aram, F., Solgi, E. and Holden, G. (2019) 'The role of green spaces in increasing social interactions in neighborhoods with periodic markets', *Habitat International*, 84, pp. 24–32. Available at: https://doi.org/10.1016/j.habitatint.2018.12.004.

Astell-Burt, T. *et al.* (2014a) 'Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities', *BMC Public Health*, 14(1), p. 292. Available at: https://doi.org/10.1186/1471-2458-14-292.

Astell-Burt, T. *et al.* (2014b) 'Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities', *BMC Public Health*, 14(1), p. 292. Available at: https://doi.org/10.1186/1471-2458-14-292.

Basiskaart Groen en Grijs (2019) *ArcGIS StoryMaps*. Available at: https://storymaps.arcgis.com/stories/0ad6a7a0558749e387c322e74141597c (Accessed: 21 June 2023).

van den Berg, A.E. *et al.* (2010) 'Green space as a buffer between stressful life events and health', *Social Science & Medicine*, 70(8), pp. 1203–1210. Available at: https://doi.org/10.1016/j.socscimed.2010.01.002.

van den Berg, M. *et al.* (2016a) 'Visiting green space is associated with mental health and vitality: A cross-sectional study in four european cities', *Health & Place*, 38, pp. 8–15. Available at: https://doi.org/10.1016/j.healthplace.2016.01.003.

van den Berg, M. *et al.* (2016b) 'Visiting green space is associated with mental health and vitality: A cross-sectional study in four european cities', *Health & Place*, 38, pp. 8–15. Available at: https://doi.org/10.1016/j.healthplace.2016.01.003.

Bertram, C. and Rehdanz, K. (2015a) 'The role of urban green space for human well-being', *Ecological Economics*, 120, pp. 139–152. Available at: https://doi.org/10.1016/j.ecolecon.2015.10.013.

Bertram, C. and Rehdanz, K. (2015b) 'The role of urban green space for human well-being', *Ecological Economics*, 120, pp. 139–152. Available at: https://doi.org/10.1016/j.ecolecon.2015.10.013.

Bolte, G., Nanninga, S. and Dandolo, L. (2019) 'Sex/Gender Differences in the Association between Residential Green Space and Self-Rated Health—A Sex/Gender-Focused Systematic Review', *International Journal of Environmental Research and Public Health*, 16(23), p. 4818. Available at: https://doi.org/10.3390/ijerph16234818.

Bowler, D.E. *et al.* (2010a) 'A systematic review of evidence for the added benefits to health of exposure to natural environments', *BMC Public Health*, 10(1), p. 456. Available at: https://doi.org/10.1186/1471-2458-10-456.

Bowler, D.E. *et al.* (2010b) 'A systematic review of evidence for the added benefits to health of exposure to natural environments', *BMC Public Health*, 10(1), p. 456. Available at: https://doi.org/10.1186/1471-2458-10-456.

Braçe, O., Garrido-Cumbrera, M. and Correa-Fernández, J. (2021) 'Gender differences in the perceptions of green spaces characteristics', *Social Science Quarterly*, 102(6), pp. 2640–2648. Available at: https://doi.org/10.1111/ssqu.13074.

Brown, G., Schebella, M.F. and Weber, D. (2014) 'Using participatory GIS to measure physical activity and urban park benefits', *Landscape and Urban Planning*, 121, pp. 34–44. Available at: https://doi.org/10.1016/j.landurbplan.2013.09.006.

Burt, J., Barber, G. and Rigby, D. (2009) *Elementary Statistics for Geographers*. 3rd edn.

Chen, K. *et al.* (2021) 'How Does Urban Green Space Impact Residents' Mental Health: A Literature Review of Mediators', *International Journal of Environmental Research and Public Health*, 18(22), p. 11746. Available at: https://doi.org/10.3390/ijerph182211746.

Christina Geng-ging Chi, Ruiying Cai, and Yongfen Li (2017) 'Factors influencing residents' subjective well-being at World Heritage Sites', (63), pp. 209–222.

Clark, A.E. (2010) '14 Work, Jobs, and Well-Being Across the Millennium', in E. Diener, D. Kahneman, and J. Helliwell (eds) *International Differences in Well-Being*. Oxford University Press, p. 0. Available at: https://doi.org/10.1093/acprof:oso/9780199732739.003.0014.

Didier Vancutsem *et al.* (2009) *Urban planning, and human health in the European city*. Available at: https://isocarp.org/app/uploads/2014/08/WHO_report_final_version.pdf (Accessed: 8 March 2023).

Diener, E. (2000) 'Subjective well-being: The science of happiness and a proposal for a national index', *American Psychologist*, 55(1), pp. 34–43. Available at: https://doi.org/10.1037/0003-066X.55.1.34.

Diener, E. *et al.* (2017) 'If, Why, and When Subjective Well-Being Influences Health, and Future Needed Research', *Applied Psychology: Health and Well-Being*, 9(2), pp. 133–167. Available at: https://doi.org/10.1111/aphw.12090.

Diener, E. and Biswas-Diener, R. (2008) *Happiness*. Available at: https://onlinelibrary.wiley.com/doi/epub/10.1002/9781444305159 (Accessed: 27 March 2023).

Diener, E. and Chan, M.Y. (2011) 'Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity', *Applied Psychology: Health and Well-Being*, 3(1), pp. 1–43. Available at: https://doi.org/10.1111/j.1758-0854.2010.01045.x.

Diener, E. and Ryan, K. (2009a) 'Subjective well-being : a general overview', *South African Journal of Psychology*, 39(4), pp. 391–406. Available at: https://doi.org/10.10520/EJC98561.

Diener, E. and Ryan, K. (2009b) 'Subjective well-being : a general overview', *South African Journal of Psychology*, 39(4), pp. 391–406. Available at: https://doi.org/10.10520/EJC98561.

Ethiek vóór je onderzoek (2021) *Rijksuniversiteit Groningen*. Available at: https://www.rug.nl/research/research-support-portal/before/ethics/ (Accessed: 23 June 2023).

Francis, J. *et al.* (2012) 'Quality or quantity? Exploring the relationship between Public Open Space attributes and mental health in Perth, Western Australia', *Social Science & Medicine*, 74(10), pp. 1570–1577. Available at: https://doi.org/10.1016/j.socscimed.2012.01.032.

Gemeente Groningen (2020) *Groenplan | Gemeente Groningen*. Available at: https://gemeente.groningen.nl/groenplan (Accessed: 20 June 2023).

Grahn, P. and Stigsdotter, U. (2003) 'Landscape planning and stress'. Available at: http://data0.eklablog.com/sociotopes/perso/documents/landscape%20planning%20and%20stress.pdf (Accessed: 27 March 2023).

Graubard, B.I. and Korn, E.L. (1999) 'Predictive margins with survey data', *Biometrics*, 55(2), pp. 652–659. Available at: https://doi.org/10.1111/j.0006-341x.1999.00652.x.

Heo, S. *et al.* (2021) 'Do persons with low socioeconomic status have less access to greenspace? Application of accessibility index to urban parks in Seoul, South Korea', *Environmental Research Letters*, 16(8), p. 084027. Available at: https://doi.org/10.1088/1748-9326/ac12f1.

HUGSI for rankings | HUGSI (no date). Available at: https://hugsi.green/ranking?mapView=standard&region=Europe&filter=All&tags= (Accessed: 11 August 2023).

Joint Research Centre (European Commission) *et al.* (2019) *Enhancing Resilience Of Urban Ecosystems through Green Infrastructure (EnRoute): final report*. LU: Publications Office of the European Union. Available at: https://data.europa.eu/doi/10.2760/689989 (Accessed: 8 March 2023).

Joshi, U. (2010) 'Subjective Well-Being by Gender', *Journal of Economics and Behavioral Studies*, 1(1), pp. 20–26. Available at: https://doi.org/10.22610/jebs.v1i1.211.

Koohsari, M.J. *et al.* (2018a) 'Are public open space attributes associated with walking and depression?', *Cities*, 74, pp. 119–125. Available at: https://doi.org/10.1016/j.cities.2017.11.011.

Koohsari, M.J. *et al.* (2018b) 'Are public open space attributes associated with walking and depression?', *Cities*, 74, pp. 119–125. Available at: https://doi.org/10.1016/j.cities.2017.11.011.

Lankhorst, R. (2020) 'Enquête leefbaarheid 2020 |'. Available at: https://oisgroningen.nl/enquete-leefbaarheid-2020/ (Accessed: 20 June 2023).

Maas, J. *et al.* (2006a) 'Green space, urbanity, and health: how strong is the relation?', *Journal of Epidemiology & Community Health*, 60(7), pp. 587–592. Available at: https://doi.org/10.1136/jech.2005.043125.

Maas, J. *et al.* (2006b) 'Green space, urbanity, and health: how strong is the relation?', *Journal of Epidemiology & Community Health*, 60(7), pp. 587–592. Available at: https://doi.org/10.1136/jech.2005.043125.

MacKerron, G. and Mourato, S. (2013a) 'Happiness is greater in natural environments', *Global Environmental Change*, 23(5), pp. 992–1000. Available at: https://doi.org/10.1016/j.gloenvcha.2013.03.010.

MacKerron, G. and Mourato, S. (2013b) 'Happiness is greater in natural environments', *Global Environmental Change*, 23(5), pp. 992–1000. Available at: https://doi.org/10.1016/j.gloenvcha.2013.03.010.

Magdalena van den Berg *et al.* (2015) *Health benefits of green spaces in the living environment: A systematic review of epidemiological studies - ScienceDirect*. Available at: https://www.sciencedirect.com/science/article/pii/S1618866715001016?via%3Dihub (Accessed: 1 March 2023).

McEachan, R.R.C. *et al.* (2016) 'The association between green space and depressive symptoms in pregnant women: moderating roles of socioeconomic status and physical activity', *J Epidemiol Community Health*, 70(3), pp. 253–259. Available at: https://doi.org/10.1136/jech-2015-205954.

Mehmetoglu, M. and Jakobsen, T.G. (2016) *Applied Statistics Using Stata: A Guide for the Social Sciences*. SAGE.

Mental health (2023). Available at: https://health.ec.europa.eu/non-communicable-diseases/mental-health_en (Accessed: 27 March 2023).

Mytton, O.T. *et al.* (2012) 'Green space and physical activity: An observational study using Health Survey for England data', *Health & Place*, 18(5), pp. 1034–1041. Available at: https://doi.org/10.1016/j.healthplace.2012.06.003.

Nesbitt, L. *et al.* (2019) 'Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities', *Landscape and Urban Planning*, 181, pp. 51–79. Available at: https://doi.org/10.1016/j.landurbplan.2018.08.007.

Ngamaba, K.H. (2017) 'Determinants of subjective well-being in representative samples of nations', *European Journal of Public Health*, 27(2), pp. 377–382. Available at: https://doi.org/10.1093/eurpub/ckw103.

Onze visie op privacy | RIVM (2023). Available at: https://www.rivm.nl/privacy (Accessed: 21 June 2023).

Punch, K. (2013) *Introduction to social resarch*. London: Sage Publications Ltd.

R.F. Hunter *et al.* (2019) 'Environmental, health, wellbeing, social and equity effects of urban green space interventions: A meta-narrative evidence synthesis', *Environment International*, 130, p. 104923. Available at: https://doi.org/10.1016/j.envint.2019.104923.

Rigolon, A. (2016a) 'A complex landscape of inequity in access to urban parks: A literature review', *Landscape and Urban Planning*, 153, pp. 160–169. Available at: https://doi.org/10.1016/j.landurbplan.2016.05.017.

Rigolon, A. (2016b) 'A complex landscape of inequity in access to urban parks: A literature review', *Landscape and Urban Planning*, 153, pp. 160–169. Available at: https://doi.org/10.1016/j.landurbplan.2016.05.017.

Rijksoverheid (2023) *Gezondheidsmonitor Volwassenen en Ouderen | Gezondheidsmonitor*. Available at: https://www.monitorgezondheid.nl/gezondheidsmonitor-volwassenen-en-ouderen (Accessed: 20 June 2023).

RIVM (2021) *StatLine - Gezondheid per wijk en buurt; 2012/2016/2020 (indeling 2020)*. Available at: https://statline.rivm.nl/#/RIVM/nl/dataset/50090NED/table?ts=1688408360971 (Accessed: 3 July 2023).

Schüle, S.A., Gabriel, K.M.A. and Bolte, G. (2017) 'Relationship between neighbourhood socioeconomic position and neighbourhood public green space availability: An environmental inequality analysis in a large German city applying generalized linear models', *International Journal of Hygiene and Environmental Health*, 220(4), pp. 711–718. Available at: https://doi.org/10.1016/j.ijheh.2017.02.006.

Sharifi, F. *et al.* (2021a) 'Green space and subjective well-being in the Just City: A scoping review', *Environmental Science & Policy*, 120, pp. 118–126. Available at: https://doi.org/10.1016/j.envsci.2021.03.008.

Sharifi, F. *et al.* (2021b) 'Green space and subjective well-being in the Just City: A scoping review', *Environmental Science & Policy*, 120, pp. 118–126. Available at: https://doi.org/10.1016/j.envsci.2021.03.008.

Sharifi, F., Nygaard, A. and Stone, W.M. (2021) 'Heterogeneity in the subjective well-being impact of access to urban green space', *Sustainable Cities and Society*, 74, p. 103244. Available at: https://doi.org/10.1016/j.scs.2021.103244.

Sillman, D. *et al.* (2022) 'Do sex and gender modify the association between green space and physical health? A systematic review', *Environmental Research*, 209, p. 112869. Available at: https://doi.org/10.1016/j.envres.2022.112869.

Statistiek, C.B. voor de (2012) *In Nederland is het groen vaak dichtbij*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/nieuws/2012/23/in-nederland-is-het-groen-vaak-dichtbij (Accessed: 8 March 2023).

Statistiek, C.B. voor de (2023a) *Bestand bodemgebruik*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/bestand-bodemgebruik (Accessed: 21 June 2023).

Statistiek, C.B. voor de (2023b) *Privacy*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/over-ons/organisatie/privacy (Accessed: 21 June 2023).

Statistiek, C.B. voor de (2023c) *Wijk- en buurtkaart 2020*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/geografische-data/wijk-en-buurtkaart-2020 (Accessed: 21 June 2023).

Statistiek, C.B. voor de (2023d) *Wijk- en buurtstatistieken*, *Centraal Bureau voor de Statistiek*. Available at: https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/wijk-en-buurtstatistieken (Accessed: 20 June 2023).

Su, J.G. *et al.* (2019) 'Associations of green space metrics with health and behavior outcomes at different buffer sizes and remote sensing sensor resolutions', *Environment International*, 126, pp. 162–170. Available at: https://doi.org/10.1016/j.envint.2019.02.008.

Suh, E. *et al.* (1998) 'The shifting basis of life satisfaction judgments across cultures: Emotions versus norms', *Journal of Personality and Social Psychology*, 74(2), pp. 482–493. Available at: https://doi.org/10.1037/0022-3514.74.2.482.

Ulrich, R.S. *et al.* (1991) 'Stress recovery during exposure to natural and urban environments', *Journal of Environmental Psychology*, 11(3), pp. 201–230. Available at: https://doi.org/10.1016/S0272-4944(05)80184-7.

US EPA, O. (2015) *Reduce Urban Heat Island Effect*. Available at: https://www.epa.gov/green-infrastructure/reduce-urban-heat-island-effect (Accessed: 11 August 2023).

de Vries, S., Buijs, A.E. and Snep, R.P.H. (2020) 'Environmental Justice in The Netherlands: Presence and Quality of Greenspace Differ by Socioeconomic Status of Neighbourhoods', *Sustainability*, 12(15), p. 5889. Available at: https://doi.org/10.3390/su12155889.

Witter, R.A. *et al.* (1984) 'Education and Subjective Well-Being: A Meta-Analysis', *Educational Evaluation and Policy Analysis*, 6(2), pp. 165–173. Available at: https://doi.org/10.2307/1163911.

Xu, T., Nordin, N.A. and Aini, A.M. (2022) ‘Urban Green Space and Subjective Well-Being of Older People: A Systematic Literature Review’, *International Journal of Environmental Research and Public Health*, 19(21), p. 14227. Available at: https://doi.org/10.3390/ijerph192114227.

6.0 Appendices

6.1 Appendix A

Neighbourhoods included in the analysis			
Badstratenbuurt	De Linie	Indische buurt	Rivierenbuurt
Bangeweer	De Wijert	Klein Martijn	Ruischerbrug
Beijum-Oost	De Wijert-Zuid	Kop van Oost	Ruischerwaard
Beijum-West	Drielanden	Laanhuizen	Schildersbuurt
Binnenstad-Noord	Europapark	Lewenborg-Noord	Selwerd
Binnenstad-Oost	Florabuurt	Lewenborg-West	Sterrebosbuurt
Binnenstad-West	Friesestraatweg	Lewenborg-Zuid	Tuinwijk
Binnenstad-Zuid	Gorechtbuurt	Noorderhoogebrug	Ulgersmaborg
Bloemenbuurt	Gravenburg	Noorderplantsoenbuurt	Van Starkenborgh
Coendersborg	Grunobuurt	Oosterhoogebrug	Villabuurt
Corpus den Hoorn	Helpman	Oosterpoort	Vinkhuizen-Noord
Damsterbuurt	Herewegbuurt	Oranjebuurt	Vinkhuizen-Zuid
De Buitenhof	Hoogkerk Dorp	Paddepoel-Noord	Vogelbuurt
De Held	Hoogkerk-Zuid	Paddepoel-Zuid	Zeeheldenbuurt
De Hoogte	Hoornse Meer	Piccardthof	Zilvermeer
De Hunze	Hoornse Park	Professorenbuurt	
De Kring	Hortusbuurt-Ebbingekwartier	Reitdiep	

VARIABLES	Model 1 Stress	Model 1)epression / Anxiet	Model 2 Stress	Model 2)epression / Anxiet	Model 3 Stress	Model 3)epression / Anxiet	Model 4 Stress	Model 4)epression / Anxiet	Model 5 Stress	Model 5)epression / Anxiety
Distance	-0.000099	0.000985	-0.000197	0.002489	0.001177	0.003610	0.001066	0.001473	-0.002296	-0.001433
	(0.002077)	(0.003225)	(0.002426)	(0.003955)	(0.007075)	(0.010661)	(0.001954)	(0.003129)	(0.002523)	(0.003978)
c.Distance#c.Distance	0.000001	-0.000001	0.000001	-0.000001	0.000002	0.000000	0.000001	-0.000000	0.000001	-0.000001
	(0.000003)	(0.000004)	(0.000003)	(0.000005)	(0.000004)	(0.000006)	(0.000003)	(0.000004)	(0.000003)	(0.000004)
Public_green	-0.070962**	-0.140309***	-0.069442**	-0.107880**	-0.083448**	-0.150501***	-0.072948**	-0.146874***	-0.069178**	-0.139348***
	(0.031722)	(0.049245)	(0.030586)	(0.049857)	(0.035042)	(0.052805)	(0.030020)	(0.048080)	(0.031233)	(0.049245)
Lonely	0.303392***	0.705272***	0.304372***	0.755316***	0.365496***	0.769128***	0.279337***	0.699519***	0.283308***	0.688074***
	(0.046209)	(0.071735)	(0.046264)	(0.075412)	(0.041221)	(0.062117)	(0.042080)	(0.067396)	(0.047849)	(0.075443)
Male	0.016538	-0.019966	0.011925	-0.103631*	0.010844	-0.024429	-0.001600	-0.030802	0.004180	-0.050253
	(0.039556)	(0.061406)	(0.032361)	(0.052750)	(0.042076)	(0.063406)	(0.038095)	(0.061013)	(0.033751)	(0.053215)
Low_education	-0.051289**	-0.017726	-0.052034**	-0.040740	-0.049495*	-0.019295			-0.039165*	-0.016129
	(0.022621)	(0.035117)	(0.022151)	(0.036107)	(0.029234)	(0.044053)			(0.022507)	(0.035488)
High_education	0.014799	0.038105	0.010786	-0.030673	0.011339	0.033282	0.031700	0.038703		
	(0.025726)	(0.039937)	(0.017587)	(0.028667)	(0.027890)	(0.042028)	(0.020585)	(0.032969)		
Income	-0.007026	-0.162585**			-0.008241	-0.162091**	-0.025253	-0.168225***	-0.010982	-0.138995***
	(0.040743)	(0.063249)			(0.043774)	(0.065964)	(0.038006)	(0.060870)	(0.028839)	(0.045471)
Workparticipation	-0.067201**	-0.068461	-0.067469**	-0.071818			-0.064115**	-0.069199	-0.064363**	-0.065027
	(0.027259)	(0.042317)	(0.027255)	(0.044427)			(0.025851)	(0.041402)	(0.026872)	(0.042370)
Age_18_39	0.310096***	0.282572***	0.312699***	0.319936***	0.293815***	0.266541***	0.313458***	0.285325***	0.316081***	0.296168***
	(0.018495)	(0.028712)	(0.015711)	(0.025610)	(0.018242)	(0.027488)	(0.017743)	(0.028417)	(0.016283)	(0.025673)
Age_40_64	0.148423***	0.112162***	0.148519***	0.110118***	0.133093***	0.097963***	0.149669***	0.118512***	0.149392***	0.114060***
	(0.021345)	(0.033135)	(0.021397)	(0.034878)	(0.022327)	(0.033645)	(0.019809)	(0.031726)	(0.020977)	(0.033075)
c.Distance#c.Income			0.000002	-0.000062						
			(0.000039)	(0.000064)						
c.Distance#c.Workparticipation					-0.000030	-0.000047				
					(0.000082)	(0.000123)				
c.Distance#c.Low_education							-0.000072***	-0.000046		
							(0.000022)	(0.000035)		
c.Distance#c.High_education									0.000050	0.000054
									(0.000034)	(0.000054)
Constant	-8.159904**	7.217868	-8.130299**	5.551301	13.634066***	1.486676	-7.604684**	7.621863	-6.570678*	9.468694
	(3.721708)	(5.777539)	(3.803092)	(6.199161)	(3.329315)	(5.016997)	(3.565200)	(5.710008)	(3.706302)	(5.843740)
Observations	66	66	66	66	66	66	66	66	66	66
R-squared	0.985200	0.984300	0.985200	0.982700	0.983600	0.983600	0.986500	0.984700	0.985700	0.984300

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	Stress	Depression / Anxiety risk
Public green	-0.070690**	-0.140711***
	(0.031432)	(0.048791)
Distance	0.000394	0.000256
	(0.000586)	(0.000910)

Feelings of loneliness	0.298456***	0.712575***
	(0.041322)	(0.064145)
Male	0.016412	-0.019778
	(0.039213)	(0.060871)
Low education	-0.050245**	-0.019271
	(0.022033)	(0.034203)
High education	0.014810	0.038089
	(0.025506)	(0.039593)
Income	-0.007318	-0.162153**
	(0.040377)	(0.062677)
Work participation	-0.069581***	-0.064940
	(0.025286)	(0.039252)
Ages 18 to 39	0.310588***	0.281844***
	(0.018230)	(0.028299)
Ages 40 to 64	0.148526***	0.112009***
	(0.021158)	(0.032843)
Constant	-7.857938**	6.771154
	(3.485743)	(5.410948)
Observations	66	66
R-squared	0.985200	0.984300
Standard errors in parentheses		
***** p<0.01, ** p<0.05, * p<0.1**		

6.2 Appendix B

	Always	Often	Sometimes	Occasionally	Never
How often did you feel very fatigued without a proper reason?					
How often did you feel anxious?					
How often did you feel so anxious that you could not become calm?					
How often did you feel hopeless?					
How often did you feel restless?					
How often did you feel so restless that you could not sit still?					
How often did you feel down or depressive?					
How often did you feel that everything costed a lot of energy?					
How often did you feel down that nothing could help you become happy?					
How often did you feel reprehensible, inferior or worthless?					

For each question 1 to 5 points are distributed (5 for always, 1 for never). A score > 15 is a mediocre risk to anxiety disorder or depression (*Risico op depressie | Buurtatlas*, no date)

1. 'Clip' both neighbourhoods and the soil map into the municipality of Groningen
2. 'Euclidian distance' was used in determining the mean distance between public park polygons and neighbourhood polygons
3. 'INT' was performed to transform each cell value of a raster to an integer cell
4. 'Zonal statistics' were used to create raster output combined with the neighbourhood codes combining zone raster (neighbourhoods) and value raster (values of distance to public park)

5. Another 'INT'was performed to transform each cell again from a raster to integer cell
6. 'Raster to Polygon' was used to transform the input raster into a simplified output
7. 'Feature to point' was used to determine the central location of each separate polygon (which is each neighbourhood)
8. 'Spatial join' combining the data in one single table

6.3 Appendix C

The syntax below can be copied to Stata 17. When using the same data as described above in the methods section. The outcomes should be identical/

```
*<Start of Syntax>*
clear all

*Importing excel data
import excel "C:\Thesis SSP\STATA_Data\All_Neighbourhoods.xlsx", sheet("Blad1")

*Step 1 prepare pathway
cd "C:\Thesis SSP\STATA_Data\Data_Files"
save "Neighbourhood_Data", replace

*Step 2 prepare data
rename A Neighbourhoods
rename B Stress
rename C Depression_Anxiety
rename D Lonely
rename E Distance
rename F Public_green
rename G Male
rename H Female
rename I Low_education
rename J Average_education
rename K High_education
rename L Income2
rename M Workparticipation
rename N Age_18_39
rename O Age_40_64
rename P Age_65up

*dropping first row
drop in 1

*Making data readable
destring Neighbourhoods, replace
destring Stress, replace
destring Depression_Anxiety, replace
```

```
destring Lonely, replace
destring Distance, replace
destring Public_green, replace
destring Male, replace
destring Female, replace
destring Low_education, replace
destring Average_education, replace
destring High_education, replace
destring Income2, replace
destring Workparticipation, replace
destring Age_18_39, replace
destring Age_40_64, replace
destring Age_65up, replace

gen Income = Income2/1000
drop Income2
*Step 3 Describe data
ssc install asdoc
asdoc sum, append

pwcorr Stress Distance, sig
pwcorr Depression_Anxiety Distance, sig

spearman Stress Distance, stats(rho p)
spearman Depression_Anxiety Distance, stats(rho p)

*Step 5 suspicion of relation not being linear
lowess Stress Distance, nograph gen(yhatlowess)
line yhatlowess Distance, sort
lowess Depression_Anxiety Distance, nograph gen(yhatlowess2)
line yhatlowess2 Distance, sort


*Step 6 multicollinarity
reg Stress Distance Lonely Public_green Male Female Low_education Average_education
High_education Income Workparticipation Age_18_39 Age_40_64 Age_65up
vif
reg Depression_Anxiety Distance Lonely Public_green Male Female Low_education
Average_education High_education Income Workparticipation Age_18_39 Age_40_64
Age_65up
vif


*step 7 homoscededasticity
mvreg Stress Depression_Anxiety = Distance Public_green Lonely Male Low_education
High_education Income Workparticipation Age_18_39 Age_40_64
```

```
reg  Stress Distance Public_green Lonely Male Low_education High_education Income
Workparticipation Age_18_39 Age_40_64
estat hettest

reg Depression_Anxiety Distance Public_green Lonely Male Low_education High_education
Income Workparticipation Age_18_39 Age_40_64
estat hettest

*normally distributed errorz
mvreg Stress Depression_Anxiety = Distance Public_green Lonely Male Low_education
High_education Income Workparticipation Age_18_39 Age_40_64
predict res3, residual
sktest res3

*Looking whether model improves with interaction
mvreg Stress Depression_Anxiety = Distance
mvreg Stress Depression_Anxiety = c.Distance##c.Distance
margins, at(Distance=(0(100)900))
marginsplot

*Downloading outreg2
ssc install outreg2

**********************************************************************
*****
*Looking at effect of SES factors on explanatory variables
reg Distance Public_green Lonely Male Low_education High_education Income
Workparticipation Age_18_39 Age_40_64

*Looking at income and workparticipation
pwcorr Income Workparticipation


**********************************************************************
*****
*ModelX building 1
mvreg Stress Depression_Anxiety = c.Distance##c.Distance
outreg2 using ModelTest, replace excel dec(3)
*ModelX building 2
mvreg Stress Depression_Anxiety = c.Distance##c.Distance Public_green
outreg2 using ModelTest, append excel dec(3)

*ModelX building 3
mvreg Stress Depression_Anxiety = c.Distance##c.Distance Public_green Lonely
outreg2 using ModelTest, append excel dec(3)
```

```
*ModelX building 4
mvreg Stress Depression_Anxiety = c.Distance##c.Distance Public_green Lonely Income
Workparticipation
outreg2 using ModelTest, append excel dec(3)

*ModelX building 5
mvreg Stress Depression_Anxiety = c.Distance##c.Distance Public_green Lonely Male
Low_education High_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using ModelTest, append excel dec(3)

**********************************************************************
*****
*Distance

*Model 1A
reg Stress c.Distance##c.Distance
outreg2 using ModelTestRegA, replace excel dec(3)
outreg2 using ModelTestRegAx, replace excel dec(3) adjr2

*Model 1B
reg Depression_Anxiety c.Distance##c.Distance
outreg2 using ModelTestRegB, replace excel dec(3)
outreg2 using ModelTestRegBx, replace excel dec(3) adjr2

*Model 2A
reg Stress c.Distance##c.Distance Public_green
outreg2 using ModelTestRegA, append excel dec(3)
outreg2 using ModelTestRegAx, append excel dec(3) adjr2

*Model 2B
reg Depression_Anxiety c.Distance##c.Distance Public_green
outreg2 using ModelTestRegB, append excel dec(3)
outreg2 using ModelTestRegBx, append excel dec(3) adjr2

*Model 3A
reg Stress c.Distance##c.Distance Public_green Lonely
outreg2 using ModelTestRegA, append excel dec(3)
outreg2 using ModelTestRegAx, append excel dec(3) adjr2

*Model 3B
reg Depression_Anxiety c.Distance##c.Distance Public_green Lonely
outreg2 using ModelTestRegB, append excel dec(3)
outreg2 using ModelTestRegBx, append excel dec(3) adjr2

*Model 4A
reg Stress c.Distance##c.Distance Public_green Lonely Income Workparticipation
```

```
outreg2 using ModelTestRegA, append excel dec(3)
outreg2 using ModelTestRegAx, append excel dec(3) adjr2

*Model 4B
reg Depression_Anxiety c.Distance##c.Distance Public_green Lonely Income
Workparticipation
outreg2 using ModelTestRegB, append excel dec(3)
outreg2 using ModelTestRegBx, append excel dec(3) adjr2

*Model 5A
reg Stress c.Distance##c.Distance Lonely Male Low_education High_education Income
Workparticipation Age_18_39 Age_40_64
outreg2 using ModelTestRegA, append excel dec(3)
outreg2 using ModelTestRegAx, append excel dec(3) adjr2

*Model 5B
reg Depression_Anxiety c.Distance##c.Distance Lonely Male Low_education High_education
Income Workparticipation Age_18_39 Age_40_64
outreg2 using ModelTestRegB, append excel dec(3)
outreg2 using ModelTestRegBx, append excel dec(3) adjr2

**********************************************************************
*****
*RQ's related to distance

*Model A1 Using RQ1
mvreg Stress Depression_Anxiety = c.Distance##c.Distance Public_green Lonely Male
Low_education High_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using Regression.xls, replace dec(6)
margins, dydx(Distance) at(Distance=(0(100)800)) post
marginsplot
outreg2 using marginsRQ.xls, replace ctitle(Distance) label

*Hypothesis 2 / 3 using RQ2
mvreg Stress Depression_Anxiety = c.Distance##c.Distance c.Distance#c.Income Public_green
Lonely Male Low_education High_education Workparticipation Age_18_39 Age_40_64
outreg2 using Regression.xls, append dec(6)
margins, dydx(Distance) at(Income=(15(1)65)) post
marginsplot
outreg2 using marginsRQ.xls, append ctitle(Distance) label

mvreg Stress Depression_Anxiety = c.Distance##c.Distance c.Distance#c.Workparticipation
Public_green Lonely Male Low_education High_education Income Age_18_39 Age_40_64
outreg2 using Regression.xls, append dec(6)
margins, dydx(Distance) at(Workparticipation=(38(1)81)) post
marginsplot
outreg2 using marginsRQ.xls, append ctitle(Distance) label
```

```
*Hypothesis 4 / 5 using RQ3
mvreg Stress Depression_Anxiety = c.Distance##c.Distance c.Distance#c.Low_education
Public_green Lonely Male High_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using Regression.xls, append dec(6)
margins, dydx(Distance) at(Low_education=(4(1)75)) post
marginsplot
outreg2 using marginsRQ.xls, append ctitle(Distance) label

mvreg Stress Depression_Anxiety = c.Distance##c.Distance c.Distance#c.High_education
Public_green Lonely Male Low_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using Regression.xls, append dec(6)
margins, dydx(Distance) at(High_education=(0(1)75)) post
outreg2 using marginsRQ.xls, append ctitle(Distance) label
marginsplot

mvreg Stress Depression_Anxiety = c.Distance##c.Distance c.Distance#c.Public_green  Lonely
Male Low_education High_education Income Workparticipation Age_18_39 Age_40_64
margins, dydx(Distance) at(Public_green=(0(3)27)) post
marginsplot

**********************************************************************
*****
*Public park

pwcorr Stress Public_green, sig
pwcorr Depression_Anxiety Public_green, sig

spearman Stress Public_green, stats(rho p)
spearman Depression_Anxiety Public_green, stats(rho p)

*No need to perform Yhatlowess as spearman is significant

*ModelX building 1
mvreg Stress Depression_Anxiety = Public_green
outreg2 using ModelTest_Green, replace excel dec(3)
*ModelX building 2
mvreg Stress Depression_Anxiety = Public_green Distance
outreg2 using ModelTest_Green, append excel dec(3)

*ModelX building 3
mvreg Stress Depression_Anxiety = Public_green Distance Lonely
outreg2 using ModelTest_Green, append excel dec(3)

*ModelX building 4
mvreg Stress Depression_Anxiety = Public_green Distance Lonely Income Workparticipation
```

```
outreg2 using ModelTest_Green, append excel dec(3)

*ModelX building 5
mvreg Stress Depression_Anxiety = Public_green Distance Lonely Male Low_education
High_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using ModelTest_Green, append excel dec(3)

*Model A1 Public green Using RQ1
mvreg Stress Depression_Anxiety = Public_green Distance Lonely Male Low_education
High_education Income Workparticipation Age_18_39 Age_40_64
outreg2 using Regression.xls, replace dec(6)
```