RESIDENTIAL BURGLARIES IN GOUDA

A data-driven approach to prediction and prevention

2 February 2018

gemeente gouda

UTRECHT DATA SCHOOL UDS

RESIDENTIAL BURGLARIES IN GOUDA

A data-driven approach to prediction and prevention

By Saïd el Khouani
Yade Rotte
Eva Siderakis
Demi van Weerdenburg

This research report is conducted on behalf of the municipality of Gouda and the Utrecht Data School at Utrecht University

SUMMARY

This research project derived from the Utrecht Data School in collaboration with the municipality of Gouda. The city is affected by a lot of residential burglaries happening in the city. The police already has information about the offenders and the victims, but not about the environmental factors that play a role in this. That is why we decided to focus our research on these factors and asked ourselves "*which factors can be identified and visualised to control or even prevent both attempted an completed burglaries in Gouda*?"

We started by searching for literature about this topic and based on that, we selected variables that seemed to be important. We divided them in four categories: weather and time, neighbourhood conditions, perceived safety and social cohesion. Then we began to collect data from various sources; from the police and municipality, but also the CBS, Kadaster, Belastingdienst and KNMI.

During collecting, combining and cleansing data, we faced some difficulties. It turned out that the CBS and the municipality used different boundaries of the neighbourhoods. This made it difficult to use both CBS and municipality/police data. We decided to use the CBS boundaries, as we could adjust the boundaries from the municipality to the boundaries of the CBS and not the other way around. This way we were able to use statistics from the CBS. We also had some troubles gathering data from the police. This took much longer than expected. Lastly, not all definitions of variables were clear and sometimes different definitions of the same variable were used.

After dealing with these problems the best we could, we started analysing our models. We used different models, because not all variables were comparable. We conducted the analysis three times: once for the total amount, once for the attempted amount and once for the completed amount of burglaries. We found various significant results for:

- *Weather and time*: positive correlation between high wind speed and the amount of burglaries and negative correlations for temperature and hours of sun and burglaries. Moreover, most burglaries happened during autumn and winter and during weekends.
- *Neighbourhood conditions*: positive relationship between population density and amount of burglaries and a negative relationship between household density and burglaries.
- *Perceived safety*: for the overall model, we found positive relations between vegetation, WhatsApp groups and reported violations and the amount of burglaries.
- *Social cohesion*: the items from the 'Stadspeiling' to measure social cohesion were not significant. We did however find significant positive correlations between the amount of children, ethnic heterogeneity and mean income and the amount of burglaries.

It should be noted that these models do not explain everything and should be seen as indicators. We do not know all underlying processes yet, because there are more factors involved. That is why we end our report with recommendations regarding further research about the variables we used, but also some other potential factors are opted.

TABLE OF CONTENTS

SUMMARY 4

TABLE OF CONTENTS 5

1 INTRODUCTION 7

1.1 CURRENT ACTIONS AND CONNECTING INFORMATION 7

1.2 MOTIVATION FOR THE RESEARCH 8

1.3 RESEARCH QUESTIONS 9

1.4 READING GUIDE 9

2 LITERATURE REVIEW 10

2.1 TIME 10

2.2 WEATHER AND SEASONAL CONDITIONS 11

2.3 NEIGHBOURHOOD AND BUILDING CONDITIONS 11

2.4 SURVEILLANCE 13

2.5 SOCIAL COHESION 14

2.6 OFFENDER'S NETWORK 16

2.7 ETHICAL IMPLICATIONS OF DATA-DRIVEN ANALYSIS 17

3 METHODOLOGY 19

3.1 DATA SOURCES 19

3.2 DEFINITIONS AND CONCEPTS 20

3.3 GOUDA DISTRICTS AND NEIGHBOURHOODS 22

3.4 A STATISTIC APPROACH TO NEIGHBOURHOODS 22

3.5 MIDDENWILLENS AND GOUDSE HOUT 24

3.6 PROCEDURE 25

3.7 ETHICAL ASPECTS 25

4 GOUDA & BURGLARY: STORIES, STATS AND FIGURES 27

4.1 RESIDENTIAL BURGLARY TRENDS 27

4.2 REPEATED BURGLARIES 30

4.3 MODUS OPERANDI AND TYPE OF HOUSE 31

4.4 On-Site visit to Goverwelle and Oosterwei 33

5 Results **35**

5.1 Time And Weather 35

5.2 Neighbourhood Conditions 38

5.3 Perceived Safety 40

5.3.1 Street Lighting 41

5.3.2 Dogs 42

5.3.3 WhatsApp Groups 43

5.3.4 Vegetation 46

5.3.5 Reported violations 47

5.3.6 Perceived Safety 49

5.3.7 Perceived Safety Model 49

5.4 Social cohesion 50

5.4.1 Methods 50

5.4.2 Results 55

6 Conclusion **58**

6.1.1 Summary of the findings 58

6.1.2 Differences between Attempted and Completed Burglaries 59

6.1.3 Conclusions 59

6.2 Recommendations 59

6.2.1 Further Investigating of Current Topics 59

6.2.2 Further General Recommendations 62

6.2.3 Additional Options For Further Research 63

6.3 Final Remark 63

References 64

Appendix I: Districts and Neighbourhoods in Gouda **67**

Appendix II: Possible Environmental Predictors of Residential Burglary **69**

Appendix III: Survey questions **70**

1 INTRODUCTION

"*Two people caught during an attempted burglary in a restaurant*", "*Man tries to sell a stolen car*". These tweets are familiar to residents of the city of Gouda. The official police Twitter account keeps the vigilant residents of Gouda on alert about the epidemic of burglaries that have plagued the town for a number of years. The city of Gouda is relatively one of the most crime affected cities within the Netherlands, according to the AD *Misdaadmeter* with a fourth place in the 2017 edition (Gerling, 2016). The most troubling aspect is the increasing number of residential burglaries that has disturbed the municipality. To put this in perspective, the AD *Misdaadmeter* 2017 edition[1] shows that in Gouda 8.91 residential burglaries took place per 10.000 inhabitants. Meanwhile, in the two biggest cities in the Netherlands, respectively Amsterdam and Rotterdam, 5.08 or 4.52 residential burglaries took place per 10.000 inhabitants. Over the past five years, the year 2017 reported 501 residential burglaries, which is finally a decrease of 21 percent. But as still 49.7% percent of all burglaries succeeded and goods were stolen from the house, the problem of preventing residential burglaries is still on top of the agenda of the so-called triangle of the city. The triangle of Gouda, consisting of the municipality, the police and the public prosecution service pointed out that burglary remains a major problem in their city. Cited from a news article of the municipality: "*to gain insights of all numbers and developments, the municipality uses various sources. Connecting information could potentially lead to new understanding. (…) the municipality a variety of information about home ownership, locations of street lighting, alleyways, public vegetation, locations of (security) cameras which all could be potentially combined with the addresses of burglary. The municipality has given a project to students of the Utrecht Data School in order to research different relations and combinations to investigate whether some connections are possible, in order to further improve the policies of the municipality*" (Gouda, 2017; own translation).

1.1 CURRENT ACTIONS AND CONNECTING INFORMATION

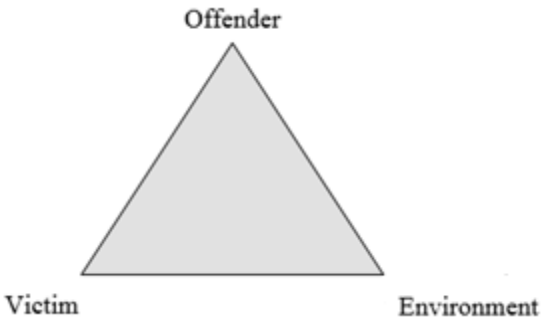
In the past few years, the triangle of Gouda has already implemented various strategies trying to intervene: focussing on both preventive, repressive and undermining tactics. The municipality provides subsidies to improve locks and is weekly active in different districts in the city to inform citizens about burglaries in the area. Furthermore, the municipality encourages citizens' initiatives and informs over a hundred WhatsApp neighbourhood prevention groups. With the Basta 2.0 program, the fight of the local police against residential burglaries continues by using a bigger police force and resources to prevent burglaries and catching burglars. Furthermore, the public prosecution has issued heavier penalties on possessing potential burglary tools such as a crowbar (Gouda, 2017).

[1] The AD Misdaadmeter 2017 uses crime numbers of the year 2016.

1.2 Motivation for the Research

Since the city of Gouda has been plagued by the major issue of residential burglaries, the municipality has been tasked with identifying the causes of these burglaries and how to prevent them. Residential burglary is a sort of crime that personally violates individuals. Erete (2013) defines 'burglary' as "*the unlawful entry into a structure to commit a felony or theft, including forced entry, unlawful entry where no force is used, and attempted forcible entry*". Combatting and preventing burglaries involves different strategies in which alternately is looked at the *offender*, the *victim* or the *environment*, which is identified by criminology literature as the *crime triangle* (figure 1). The idea is based on the fact that any crime is the result of when offenders and targets (the victim) come together in time and space.

Figure 1 – Crime triangle



As starting point for this research was considered that the municipality and the police have already done extensive investigations on both the side of offenders and victims. The environment of Gouda has not been well researched, however it is an important coordinate of the crime triangle which explains outbreaks of crime.

Different authors further stress the importance of the environment in enabling crimes such as residential burglaries. For instance, Moreto (2010) states that any residence can fall victim to an act of burglary, but in reality certain areas and places are more conductive for crime than others. Authors such as Sherman (1995) underline this statement as he found that the place of a crime is six times more predictive of future crime than the offender identity. This further underlines the shift in research from *who done it* into *where done it* as important practice. Moreto (2010) explained the environmental factors of residential burglary in four distinct factors such as surrounding area; household characteristics; immediate design and planning features and; other aspects of lifestyle affecting the location.

1.3 Research Questions

In this research, we try to combine theories from different disciplines into a data-driven approach in predicting and (even) preventing residential burglaries. Within the past decade, the production of big datasets increasingly has influenced governmental analytics. Kitchin (2014) defines big data as *"massive, inter-related datasets that can be connected and utilised in diverse ways"* (p.3). Data could define the way governments aspire to operate as in the case of residential burglaries it is promising to use different (census) analytics and leads to 'game-changing' and new significant questions. In the case of Gouda, there are certain different factors that can be identified from different datasets which can lead to breaking insights in to the criminality issue in Gouda. By focussing on the environment in the crime triangle and a data-driven approach has led to the following research questions which will provide the framework for this study:

Which environmental factors can be identified and visualised to control or even prevent both attempted and completed burglaries in Gouda?

1.4 Reading Guide

This report continues with chapter two in which relevant literature is reviewed in order to build a theoretical framework of potential predictors of the environment that could explain residential burglaries. Chapter three elaborates on the methodology of this project, in which we focus on data sources, definitions and concepts, procedures and ethical aspects that needed to be considered in this data project. The fourth chapter describes the city of Gouda and the challenges it faces with residential burglary by giving different descriptive stories, statistics and figures in order to provide an deeper understanding of residential burglaries in Gouda. Hereby we provide statistics and figures of burglaries that took place between 2013 and 2017. In the next chapter, chapter five, we will elaborate on the findings during the data analysis. The conclusion, which is chapter six, will summarize the foremost significant results and give recommendations for further research and for possible use in policy in order to further prevent residential burglaries in Gouda.

2 LITERATURE REVIEW

Crime has a big impact, which affects the quality of the lives of individuals, households and society itself (Bjørgo, 2016). Residential burglary is a type of crime that leaves individuals personally violated. Rengert and Groff (2011) describe a residential burglary as 'a personal attack' on an individual.

In many cases of residential burglary it seems that even when the home occupant uses prevention measures, such as an alarm, offenders are not deterred (Homel, Macintyre & Wortley, 2013). In many areas affected by residential burglary, the residents do not have a complete control over the environment in which they reside (Rengert & Groff, 2011). Instead of focusing on why residential burglary takes place, different theories increasingly emphasize on where burglaries likely take place. Sherman (1995) describes this as a shift from who done it into where done it. Furthermore, he stated that the place of a crime is six times more predictive of future crime than offender identity. Therefore one is not able to influence the danger of burglary completely.

This theoretical framework describes the potential environmental predictors of residential burglaries, such as time, weather conditions, neighbourhood conditions, surveillance and social cohesion. Several predictors are intertwined and referenced in other chapters, whereas research on the implications of data-driven research is added at the end for thought and to shed light on biases and ethical implications that possible results from this research could bring forth.

2.1 TIME

Patterns between time and burglaries are well known and researched. Time as a predictive variable seems to be tied to other environmental variables. Felson and Poulsen (2003) state that time can be a better predictor when looking at the hourly variations in crime. They researched common patterns in the course of daily life to argue that predictive strategies should be more tuned to the difference between researched areas when looking at the relation between time, space and crime (Felson & Poulsen, 2003). Differences between studies about time patterns might be due to the different measurement of daytime and night-time burglaries, where Felson and Poulsen (2003) stated that the use of true light and darkness conditions by using actual sunrise and sunset information might create better predictive models than simplistic 12-hour timetables.

Montoya, Junger and Ongena (2016) also suggested that burglaries in different parts of the day should have a separate model for burglary prevention, since the variables that burglars consider are based differently on the time of day (Montoya et al., 2016). Results from their study show that daytime burglary seems to be related to territoriality and access control. Building characteristics like having a front garden and having an un-detached house lowers the risk on being burglarized during daytime (0.46 and 0.55 times), while having a corner house is associated with a higher risk of being burglarized during daytime (1.97 times). Night-time burglary seems to be related to access control and the level of difficulty to break in. Having an

un-detached house is also at night associated with a lower risk on being burglarized (0.68 times). Having a front garden and a dog lowers the risk on a burglary (0.72 and 0.36 times), where offender availability was associated with a large increase in burglary risk (11.58 times).

2.2 Weather and seasonal conditions

Weather is an uncontrollable variable, the same counts for seasons and holidays. Literature about the effect of weather on crime does not give a clear answer. As an example, Horrocks and Menclove (2011) conclude that on rainy days there is probably less social control and therefore it would be easier to commit a burglary, whereas Gleaser, Sacerdote and Scheinkman (1996) concluded the opposite. Horrocks and Menclova (2011) state that rain leads to higher levels of burglaries. They base this on the assumption that better weather will increase mobility and social interaction as people are more likely to leave the house. However, good weather can also increase the chance of capable guardians, increasing the chance to be caught in the act (Horrocks & Menclova, 2011). Criminals may also be less motivated to leave the house during bad weather due to the discomfort of being outside in bad weather. In addition, bad weather could make transporting (electronic) goods more difficult (Horrocks & Menclova, 2011).

On the other hand, Brunsdon, Corcoran, Higgs and Ware (2009) and Weisel (2002) found that on sunny days, people spend more time in the outdoors (Weisel, 2002) and often leave their windows open, which makes it easier for the burglar to get in (Brunsdon et al., 2009). This means the weather could influence whether people are at home as well as the ease with which burglars can operate. Ranson (2012) had comparable results. Even though he did not include just burglaries but also other crimes, he found that higher temperatures lead to higher crime rates.

Both Weisel (2002) and Falk (1953) conclude that burglary rates are higher during summer months than during winter months. Falk (1953) adds that especially weekend days are popular crime-days. Schmalleger (1997) also followed this thought with the notion that when the weather is pleasant, people are less likely to be at home and are more likely to leave for a while in the evening or go away for the weekend.

2.3 Neighbourhood and building conditions

Residents often do not have a complete overview of the environment in which they reside, which includes neighbourhood conditions. A few researchers have put emphasis on explaining neighbourhood conditions in relation to burglary. For example, Nasar (1981) has investigated commercial facilities. He found that facilities that are close to major roads experienced more burglaries, especially when there were not a lot of shops located. Weisel (2002) concluded the same for single-family houses. She stated that houses close to big roads are vulnerable, because there is a lot of traffic and burglars do not stand out when they are checking out/return to the neighbourhood. Van Nes and López (2010) on the other hand, found that the higher the amount of space between the house and the main road, the bigger the chance of a burglary. This is partly due to differences in visibility, which can be divided in the factors population density and amount of vegetation.

Results of Wolfe and Mennis (2012) stated that burglary rates were higher in neighbourhoods with a low population density. When controlled for poverty and vegetation, results of Chan (2015) showed the opposite effect: during her geographical research in Jefferson, Kentucky, she found that crime rates increase in neighbourhoods with a high population density. These findings suggest that factors like vegetation and poverty could be of influence in the rate of burglaries. After a lot of observations on-site, Van Nes and López (2010) concluded that the higher the visibility between houses (through windows), the lower the chance of residential burglary. According to Homel et al. (2013), properties with a lack of visibility (because of a fence or vegetation for example) are more vulnerable than properties that are visible.

The amount of vegetation in a neighbourhood might influence visibility of burglars during the burglary or in the following escape, because burglars like the cover from trees (Homel et al., 2013; Weisel, 2002). Fisher and Nasar (1992) also stated that low-lying and dense vegetation is related to more burglaries. However, Wolfe and Mennis (2012) concluded the opposite. They collected data about welfare and the amount of vegetation in Philadelphia. The results show that the crime rates decreased when there was more vegetation. These results were the same for neighbourhoods with a high and a low socioeconomic status. Troy, Grove and O'Neil-Dunne (2012) think this might be explained by the fact that people are more likely to go outside, as those environments are more appealing. This results in more people present to watch out for each other. Nevertheless, effects of vegetation are rather difficult to measure. Whether vegetation decreases inter-visibility does not only depend on the height or width of it, but it also changes throughout the year.

Besides vegetation and population density, academics found other risk factors in the search for burglary prevention. Houses on the corner of a street are more vulnerable, because they offer different escape routes and an easier access (Nasar, 1981; Homel et al., 2013; Weisel, 2002). Donovan and Prestemon (2012) agreed with Homel et al. (2013), but also found that houses on a corner of an intersection experience lower crime rates. In their research, Donovan and Prestemon (2012) studied five types of crime in the municipality of Portland. Houses that were in better conditions were likely to experience crime, because they probably have nicer goods (Donovan & Prestemon, 2012). Weisel (2002) also stated that earlier victims and houses near those victims also tend to be a risk factor, because these neighbourhoods are more researched and thereby easier to break into. Earlier victims also tend to replace the lost goods (Weisel, 2002).

Moreover, several authors concluded that poverty is also related to higher crime rates (Nasar, 1981; Weisel, 2002; Wolfe & Mennis, 2012). They figured that there were more break-ins in dirty areas, areas with a high poverty rate (Nasar, 1981), low social cohesion and ethnic heterogeneity (Hirschfield et al., 2014). Also, the higher the educational attainment, the lower the chance of burglary (Nasar, 1981).

2.4 SURVEILLANCE

There is much literature about the effect of surveillance on the prevention of burglaries, although some literature makes at first glance contradictory claims compared to other literature in this field. The importance of certain variables vary during the day and night, whereas the method and target of burglary also change (Coupe & Blake, 2006). By making better distinctions between the methods and target, better predictions can be made about which houses are more likely to be the target of burglary. As previously stated, different factors are of importance for different day parts when a burglar is looking out for a target (Montoya et al., 2016). Based on possible reasons for burglary in certain houses, it seems appropriate to create separate models for burglary prevention during the day and at night, since the variables that burglars consider are different based on the time of the day.

Different methods can be tried for the prevention of burglaries. Sherman and Weisburd (1995) researched the effect of police patrols. They doubled the amount of patrols in certain areas for ten months and found reduced crime numbers up to 13%. Riccio (1974) had a similar effect, but it found very little direct effect of amount of police patrols and reduction of crime numbers. Lastly, Koper (1995) concluded that police patrols are most effective when they stop for approximately 10 - 15 minutes.

Farrington and Welsh (2002) reviewed the effects of street lightings. Results of thirteen British and American studies showed that earlier research does not give a straight answer about the effect of street lightings. Not all literature found a significant effect, but some authors found a 7% up to 30% decrease of burglary caused by street lighting. In 2008, Welsh and Farrington updated their literature review and concluded the same effects for burglaries. They also found that not only night-time burglaries decreased, but also day-time crimes, which makes it questionable whether there is an interactive effect as well (Welsh & Farrington, 2008). Research of Herbert and Davidson (1995) might help answering this question. They stated that citizens felt more safe to go outside when neighbourhoods were improved with, for example, street lightings. People felt more safe to go outside and were more willing to pay attention to what is happening in their neighbourhood.

Street cameras have been proven successful in case studies that found place in Texas (Bridges, 2005) and South Africa (Minnaar, 2007). In both studies, results also showed that, like street lightings, citizens felt safer when there were street cameras present.

Another important factor of surveillance in the prevention of burglary is citizen participation. Erete (2013) investigated the effectiveness of active citizen participation against burglary, and the communication between citizens and police. According to this study, the possible disadvantages of civilian patrols outweigh the expected benefits. Silent participation seems to be more effective as a deterrent to burglars. He concluded that citizen participation in the field of digital tools is much more effective and a minimal risk for citizens (Erete, 2013). An example is the participation in neighbourhood WhatsApp groups and the registration of police initiatives to prevent burglary (e.g. registering a camera with the police, placing lamps

with light sensors on the front and at the back of the house and following burglary prevention courses).

Akkerman (2015) studied the effect of WhatsApp groups in 35 neighbourhoods in Tilburg as a sign of social cohesion within neighbourhoods. After the introduction of the groups, the amount of break-ins was 40% less than before the introduction. This suggests that burglars are being deterrent to neighbourhoods with WhatsApp group signs, which is visible in the amount of break-ins in a neighbourhood before and after the introduction of these groups (Akkerman, 2015).

At another level, dogs can be some sort of surveillance as well. Owning a dog, according to research by Nicolson (1994) has a discouraging effect on burglars, which goes against Erete's (2013) findings. Erete (2013) found that dogs do not scare burglars away, since most dogs are not trained to protect a home and are easily distracted. On the other hand, other authors did find evidence that dogs are effective crime reducers (Homel et al., 2013; Weisel, 2002). The 96 respondents of Homel et al. (2013) stated that big dogs are a physical threat and small dogs are often noisy. Erete (2013) concluded that burglars were more discouraged by a combination of surveillance and social cohesion and relatively less deterred by cameras, alarm systems, dogs and fences (Erete, 2013).

2.5 Social cohesion

Social cohesion has been researched in many forms over the past decades as an influential factor on the degree of crimes in a neighbourhood. It is generally conceptualized as the extent to which people are cooperating to achieve mutual goals in typically a residential environment (Hirschfield & Bowers, 1997). Most literature on social cohesion is based on the social disorganization theory. This theory states that when a group of people has no cohesiveness (or group unification/organization), it is easier for crime to occur (Kawachi, Kennedy & Wilkinson, 1999). The reason behind this phenomenon is that the group cannot confront this danger as one collective, but rather as individuals, reducing the bargaining power of the group and thus their ability to tackle issues collectively. This is one mechanism through which an absence of social cohesion increases the risk for crime.

Erete (2013) organized a panel session with convicted burglars, where they named social cohesion as the most important deterrent against burglary, more than dogs, alarms and cameras. Neighbourhoods where people were talking with each other seemed more cohesive, which discouraged burglars from breaking in to that neighbourhood (Erete, 2013). Erete (2013) stated from his interviews that burglars look at different variables before deciding targets. The amount of people outside or actively monitoring from behind curtains, kids and dog walkers on the street and keeping up the appearance of the neighbourhood may all contribute to the perceived chance of getting caught as assessed by the burglar. The effects of WhatsApp groups in a neighbourhood, as previously stated, is also a deterrent to burglars for break-ins in that neighbourhood (Akkerman, 2015).

Kawachi et al. (1999) linked the social disorganization theory to the concept of social capital, which stands for the extent to which people have access to resources through their social networks. In socially disorganized neighbourhoods resources are not shared among group members, which leads to limited information for members. This makes a group an easier target for criminals, as the chance of repercussions are lower (Kawachi et al., 1999). In these kind of situations, it is the personal benefit a watchful neighbour can earn that is of importance. If there is a chance that a neighbour will reciprocate watchful behaviour in the future, one might consider reporting suspicious activity. The opposite stands also true in neighbourhoods with low social cohesion (and lower trust in strangers), the chance that altruistic behaviour (like reporting a crime) takes place is less likely (Kawachi et al., 1999; Hooghe, 2007). Kawachi et al. (1999) looked at the relationship between social cohesion and crime on a state level, with the focus on the social environment. They distinguish social cohesion in neighbourhoods under two indicators: trust and parental control. Trust is measured as the trust people have in their community, especially in strangers. If trust (and cohesion) is low, the threshold towards reporting wrongdoings is higher and more crimes will be committed.

Bernasco and Luykx (2003) also used the social disorganization theory to explain criminal activity, where they focus on three dimensions: attractiveness, opportunity, and accessibility. They stated that opportunity is the most tied to social cohesion and name residential mobility and ethnic heterogeneity as the main social factors that predict residential burglaries. Residential mobility measures the extent to which people are given time to form bonds with their neighbours (Bernasco and Luykx, 2003). Ethnic heterogeneity is based on the homophily theory: people have a slight preference for people who possess the same characteristics, such as social class, race or education level (McPherson, Smith-Lovin & Cook, 2001). Heterogeneity on an ethnic level in a neighbourhood leads to decreased social interaction and thus to less social cohesion.

Hirschfield and Bowers (1997) looked at how social cohesion affects crime in disadvantaged neighbourhoods. They operationalized social disorganization under a few risk factors, which they used as indirect indicators of social cohesion. They stated that direct measures are virtually impossible to study, since risk factors for social disorganization do not manifest themselves in one clear form. By using the indirect measures of lone-parent households, recent migrants, ethnic heterogeneity and social heterogeneity, they aimed to approach the concept of social cohesion. They opted for the amount of lone-parent households per thousand households, as a measure for parental control (Hirschfield and Bowers, 1997). For the calculation of recent migrants, the permillage (per thousand residents) is used as a measure for residential mobility, because this indicates that a neighbourhood does not have a stable population. Like in the research from Bernasco and Luykx (2003), ethnic and social heterogeneity are used here as a measure for the tendency for homophily. One more direct measure for social cohesion is acquired from looking at the frequency of civilian reports to the police about disturbances.

Martin (2002) took a different approach on operationalizing social cohesion: he used the number of community groups (e.g. community centres) and corrected this for the amount of

houses in a neighbourhood. This approach not only looked at the victims, but also the broader surroundings: which resources are available for people to interact in. This is in line with Hirschfield and Bowers' (1997) argument that more opportunities for the neighbourhood to come together leads to higher social cohesion.

Kawachi et al. (1999) showed that trust in the community and single parents are both significant predictors of criminal activity, specifically burglaries. Bernasco and Luykx (2003) found that residential mobility and ethnic heterogeneity both significantly contribute to the successful selection of targets by burglars. Hirschfield and Bowers' (1997) analysis had similar results for heterogeneity, parental control and mobility. They also showed that in areas with a lower average socioeconomic status, increasing social cohesion, through for example encouraging new residents to participate, or offering support to single parents, can help combat crime levels, especially the burglary and assault rates. Martin (2002) too found a relationship between social cohesion and residential burglaries, with less-cohesive neighbourhoods having a higher burglary rate. Both the victim- and environment-view then provide the same result: social cohesion in a neighbourhood provides protection against crime.

There appears to be a strong relationship between social cohesion and criminal activity, specifically burglaries, from the environment-, offender- and victim-side of the issue. Important recurring operationalization of social cohesion are residential mobility, ethnic heterogeneity, trust, as well as opportunities for the neighbours to get together. These opportunities are limited not only by spatial constraints, such as the availability of community areas, but also time constraints, such as the pressure put on single parents to provide.

2.6 Offender's Network

The social network of the offender is a combined factor of social cohesion and the neighbourhood conditions. According to the 'crime pattern theory' of Brantingham and Brantingham (2008) offenders prefer to commit their burglary in a familiar neighbourhood. Burglars are better able to observe possible victims unnoticed and know which buildings are worth the effort (Brantingham & Brantingham, 2008). Among others, Van Nes (2006) confirmed this theory, and found that most burglars operate within a ratio of 3 km of his/her own house. Menting Lammers, Ruiter and Bernasco (2017) also agreed with this during their study about 19.500 felonies in Den Haag. They added that burglars tend to break-in within their own family. Weisel (2002) stated that if the victim was an easy target in the first place, then this makes it more attractive for a burglar to go back, because they know the property and how they can get in there (Bernasco, 2017; Hearden & Magill, 2004, Weisel, 2002). Moreover, victims often replace their stolen goods (Weisel, 2002). When it seems that the earlier victim took some precautionary measures, the burglar often goes to another house in the same neighbourhood (Bernasco, 2017; Hearden & Magill, 2004; Menting et al., 2016).

2.7 Ethical implications of data-driven analysis

Within the past decade, the production of big datasets increasingly has influenced governmental analytics. Kitchin (2014) defines "big data as massive, inter-related datasets that can be connected and utilized in diverse ways" (p.3). Data defines the ways governments aspire to operate – which leads us entering a new era of Big Data (Boyd & Crawford, 2012; Chan & Bennet-Moses, 2015; Richards & King, 2014). While there is no doubt big data promises new ways of governance of cities by its possibilities (Kitchin, 2014). Chan & Moses (2015) situated the impact of Big Data in the field of criminology, where it is promising to use census analytics and dealing with messy data.

Data is often portrayed as 'just data': objectively collected numbers that describe (urban) phenomena. This implies data as an inherently 'good thing' to make cities more secure, efficient and productive etcetera (Kitchin, 2014). Big Data implies the potential of new insights in city intelligence but on the other hand, are critically questioned by many different scholars. The potential of Big Data is to uncover old issues and new questions through merging and linking data. This provides a dynamic reflection of processes in the city – such predictive tools for crime prediction. The concept of Big Data is controversial itself: the use of data implies a rethinking of the idea of governance. In context, it requires collaboration between authorities and data scientists (Taylor & Richter, 2015). Within criminology, computer modelling and algorithms are frequently more used to predict crimes and to guide policing strategies and justice decisions (Chan & Moses, 2015, p. 25). Predictive policies go further than hotspot analysis; Big Data is used to analyse and predict where, when and why the next crime will likely take place. Hence these predictions are about 'place' and 'time' and involves linking different sets of data.

Kitchin (2014) and Taylor and Richter (2015) both discuss how data is inferred and produced. As Kitchin (2014) puts it: it is not 'just' data. The results of data analytics give the illusion of completeness – while they may cause problems regarding validity and ground truth (Taylor & Richter, 2015). Its potential and richness can cause a Gods Eye View which not only causes taxonomic and ontological problems but foremost ethical problems. When predicting crime, this often involves profiling geographical areas and communities which becomes stigmatized by statistical correlations (Chan & Moses, 2015).

Boyd and Crawford (2012) conclude that working with Big Data is still subjective as it does not necessarily have a closer claim to the objective truth. As they put it: "a model may be mathematically sound (…) but as soon as a researcher seeks to understand what it means, the process of interpretation has begun" (Boyd & Crawford, 2012, p. 178). It does not necessarily mean that what is not seen in the analysis, means it is not there – it is still a sample (Kitchin, 2014). Therefore data analytics may not reflect on every relevant aspect as it may reflect on particular geographies or power structures (Boyd & Crawford, 2012, p. 181).

The discussed research papers all demand for a broad critical perspective that questions how the use of Big Data affect power dynamics of urban governance (Taylor & Richter, 2015). It seems that there is little awareness amongst policy makers regarding the implications of using

citizen data in all four principles. The results of the data-analytics of residential burglaries in Gouda should not be the goal itself – the project should help to learn lessons from those four principles in handling big data in the governmental sphere. To conclude with words of Boyd & Crawford (2012), these new tools participate in shaping the world as we use them. The era of Big Data has just begun and therefore it is important to shed light on assumptions, (ethical) implications and biases of Big Data analytics.

3 METHODOLOGY

3.1 DATA SOURCES

The municipality provided us with a number of data sources from its own collection. Among these are the 'Basisregistratie Adressen en Gebouwen' (BAG), which is a list of all addresses and purposes of buildings, and the 'Basisregistratie Personen' (BRP), which is a list of all citizens of Gouda. Because these are the municipality's own administration files, they are the most accurate and up-to-date lists on these topics. One limitation is that people may have moved, or new buildings may have been opened since the release of the lists, as they only get released once or twice a year.

Furthermore we have access to two years of the 'Stadspeiling', a yearly survey on welfare, filled in by about 2000 citizens each year. For these results, we know in which neighbourhood the respondents live, so we can aggregate the opinions of citizens to a neighbourhood level, allowing us to take these into consideration in our analyses.

Besides the municipality-sourced information, we also incorporate CBS neighbourhood statistics. These include demographic information, such as for example average neighbourhood income, population density, the proportion of immigrants, and the division of age groups across Gouda. Because this data originates from the highest statistical office in the Netherlands, and the municipality uses this same information, we assume this information was the most accurate representation of these topics.

For our analysis on perceived safety, we have numbers on dog population of Gouda. This information is provided by the BSGR, the tax services for the region the municipality is in. The BSGR keeps a dog register, as people in Gouda are required to pay dog taxes. The register relies on people registering their own dog and the BSGT occasionally performs random spot-checks to see if there are unregistered dogs at addresses. But because of this inaccuracy, our results cannot be generalised with certainty.

For our analysis on the influence of the weather, we collected the weather variables from the KNMI database, which we cleaned up in Excel and imported in SPSS. For the vacation days, we used information from the municipality Gouda, where we merged school holidays and national/religious holidays as the vacation variable.

Finally, we have data on all burglaries since 2013 until 2017, which has been provided by the police. This data includes all burglaries at home addresses since 2013, and lists whether a burglary was completed or attempted. We also have information on the modus operandi and the time of the burglary.

3.2 Definitions and Concepts

The municipality makes a distinction between attempted and completed burglaries. Together they are the total amount of burglaries. Attempted burglaries are the ones with signs of forced entry or when there were no goods stolen. Completed burglaries are defined as the burglaries whereby the burglar actually took something and got away. Although the burglar was not caught in the act, it does not mean (s)he was never caught at all.

For this research, we decided to use the tree types of burglary as dependent variables. We look at the total amount of burglaries, and we split it into attempted burglaries and completed (or successful) burglaries.

Moreover, when we talk about 'amounts' in this report, it is important to note that in our analysis we used the *relative* amounts, because neighbourhood sizes differ. The amounts of burglaries are therefore divided by the amount of households per neighbourhood. We use households instead of citizens, because the household as a whole is the potential victim. The absolute and relative amounts of burglaries can be seen in table 1.

Table 1 - Amounts and relative amounts of burglaries per neighbourhood between 2013-2017.

Neighbourhood	Total amount	Attempted amount	Completed amount	Relative total amount	Relative atttempted amount	Relative completed amount
Oosterwei	142	53	89	.254	.095	.159
Voorwillenseweg	49	26	23	.233	.124	.11
Polderbuurt	219	120	99	.183	.1	.083
Componistenbuurt	115	76	39	.18	.116	.06
Vreewijk	87	43	44	.164	.081	.083
Vrijheidsbuurt	47	26	21	.149	.083	.067
Molenbuurt	191	108	83	.145	.082	.063
Mammoet	46	23	23	.133	.067	.067
Slagenbuurt	69	38	31	.131	.072	.059
Kort Haarlem	225	80	145	.13	.046	.084
Sportbuurt	75	40	35	.127	.068	.059
Bloemendaalseweg	97	38	59	.124	.049	.076
Hoevenbuurt	149	62	87	.12	.048	.067
Muziekbuurt	36	18	18	.12	.06	.06
De Gaardenbuurt	28	11	17	.119	.047	.072
Achterwillenseweg	63	26	37	.118	.049	.069
Windrooskwartier en Heesterbuurt	122	56	66	.102	.047	.055
Bodegraafsestraatweg	12	0	12	.1	0	.1
De Korte Akkeren-Oud	309	119	190	.1	.039	.062
De Korte Akkeren-Nieuw	88	42	46	.091	.043	.047
Ouwe Gouwe	188	68	120	.089	.032	.057
Stolwijkersluis-West	3	2	1	.086	.057	.029
Wethouder Venteweg	50	26	24	.079	.041	.038
Wervenbuurt	22	11	11	.075	.037	.037
Grassen- en Waterbuurt	31	15	16	.072	.035	.037
Groenhovenkwartier	49	23	26	.067	.032	.036
Kadenbuurt	83	34	49	.065	.026	.038
Lusten-, Burgen- en Steinenbuurt	45	26	19	.065	.037	.027
Stolwijkersluis-Oost	9	3	6	.06	.02	.04
De Goudse Poort	18	8	10	.05	.022	.028
Hoef-en Veldbuurt	65	28	37	.05	.021	.028
Nieuwe Park-Oost	44	11	33	.049	.012	.037
Weidebloemkwartier	11	3	8	.049	.013	.036
De Baan en omgeving	34	10	24	.047	.014	.033
Statensingel	15	8	7	.043	.023	02
Boerhaavekwartier	67	15	52	.042	.009	.033
Zomenbuurt	38	20	18	.041	.022	.02
Industrieterrein Kromme Gouwe	1	0	1	.04	0	.04
Industrieterrein Langs de Hollandse IJssel	5	1	4	.038	.008	.031
Raam en omgeving	36	15	21	.035	.014	.02
Turfmarkt en omgeving	19	7	12	.03	.011	.019
Nieuwe Markt en omgeving	28	16	12	.029	.017	.013

3.3 GOUDA DISTRICTS AND NEIGHBOURHOODS

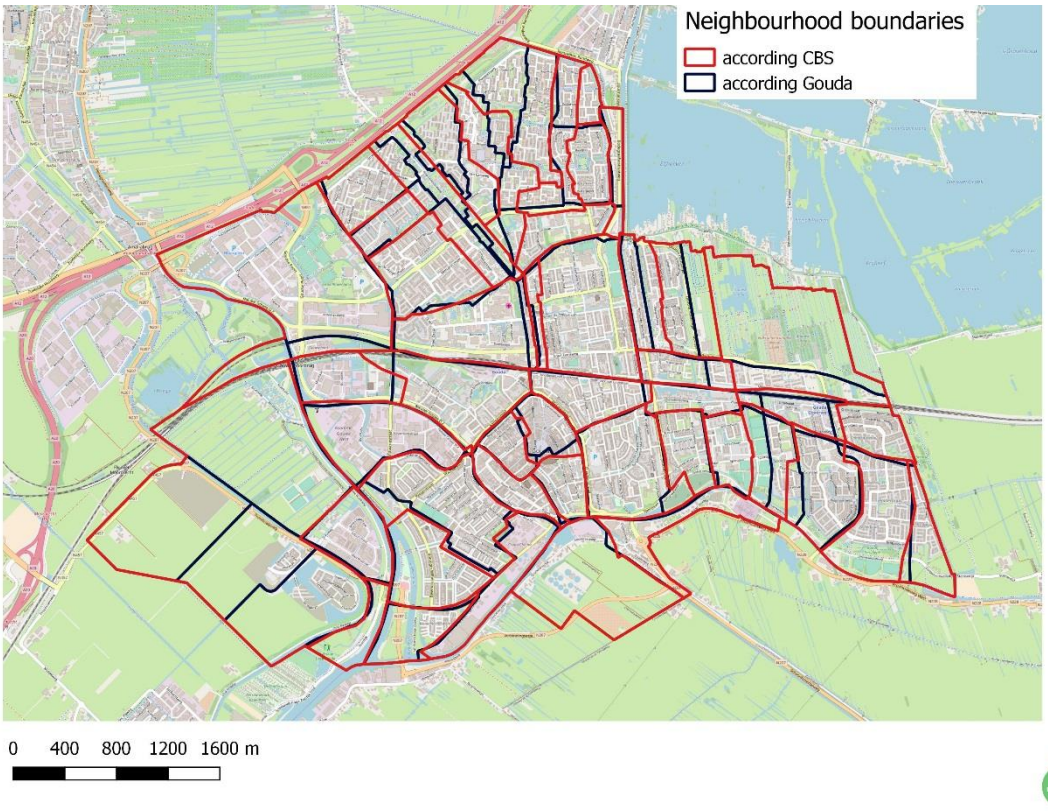
For this project, we chose to analyse burglaries on a neighbourhood level as our characteristics of interest are often shared within a neighbourhood, while they differ a lot between districts. Gouda exists of nine districts and 51 neighbourhoods with different characters (figure 2). Besides the historic inner city, one of the oldest districts of the city is the Korte Akkeren district. Bloemendaal and Plaswijck form the north side of the city, surrounded by the Reeuwijkse Plassen and the Goudse Hout polder district. South of the railroad, Kort Haarlem and Goverwelle form one of the most densely populated districts of the city (respectively 10.112 and 11.288 inhabitants). While the Stolwijkersluis is a very spare populated district with 407 inhabitants. Gouda is currently expanding on the west side of the city and developing new neighbourhoods such as Westergouwe but also (re)developing along the railway such as the Middenwillens neighbourhood.

3.4 A STATISTIC APPROACH TO NEIGHBOURHOODS

As previously described in this chapter, we have retrieved data from multiple sources such as the municipality itself and data from the CBS. For spatial purposes, we have found a discrepancy in geographical boundaries according to the data of Gouda itself and CBS as seen in figure 3. The red lines represents the boundaries according CBS and they clearly are off with the official registered boundaries of the municipality.

Figure 2 - Map of neighbourhoods and districts in Gouda.
Source: GIS.GOUDA

Figure 3 – Differences between CBS and Gouda neighbourhoods.



The differences are mainly spatially bound, but CBS furthermore acknowledges less neighbourhoods than the municipality itself does. The differences are in the Goudse Hout/Middenwillens area and in the Oostpolder in Schieland area. In consideration with the municipality, we chose to retain the neighbourhood boundaries according to the CBS. Therefore we re-projected all spatial data to the boundaries of the CBS in QGIS, as we were not able to transform the data the other way around. In this project, we distinguish 47 neighbourhoods in total. For the Oostpolder in Schieland, we have combined the neighbourhoods De Eilanden, Tuinenbuurt, Westergouwe and Oostpolder in Schieland. This is area is still developing and primarily consists of rural area now. Therefore, it did not have a lot of burglaries the past five years. Attached in Appendix I is a table with overview of all neighbourhoods and distinctions between CBS and Gouda.

3.5 MIDDENWILLENS AND GOUDSE HOUT

A special consideration needs to be made for the combined Middenwillens and Goudse Hout area. By combining those two neighbourhoods, the area becomes too skewed for statistical analysis. As seen in figure 4, De Goudse Hout is a recreational nature reserve of 83 hectare while Middenwillens is a small neighbourhood under construction. At this moment, about hundred newly built houses are realized. By combining a large nature reserve and a small area with newly built housing leads to skewed results in the analysis. With nine burglaries in total, we have decided to exclude the Goudse Hout neighbourhood of further analysis.

Figure 2 – Middenwillens and De Goudse Hout.

3.6 Procedure

In order to prepare our data, we had to take different steps, depending on the analysis. This preparation prefaces each result section, per subject. One general preparation that was performed on the data is the aggregation from individual cases to neighbourhood averages. This was necessary for all analysis except for the analysis of the effects of weather and time. We have based our aggregation on the CBS neighbourhoods, and counted all cases within these neighbourhoods since 2013. This results in total, completed and attempted burglaries. Because we divided totals of burglaries by the number of households per neighbourhood, it allows us to compare neighbourhoods with each other, and draw meaningful conclusions on the factors that influence the number of burglaries.

We had to wait for the police data about burglaries for quite some time. Meanwhile, we were cleansing our data and divided the selected variables into four topics (weather and time, neighbourhood conditions, perceived safety and social cohesion). We started building our models based on data the municipality had about burglaries the burglaries of 2016. We then found out about the different boundaries of the municipality and the CBS, as explained above. Therefore, we needed to redo some measurements.

After we got the police data, we could do our actual analysis. We used various kinds of statistical analysis. In Box 1 there is a basic explanation on how to interpret them. We gathered the results and started writing the report and preparing the presentation.

3.7 Ethical aspects

The data used in this report is very sensitive: information on the houses that are prone to burglaries, whether in the past or in the future, would be identifiable if we did not aggregate our results appropriately. On the one hand, individual cases would provide us more specific information, like household composition and income. On the other hand, if results could be traced back to the victims, these could be easily targeted by malevolent readers. In agreement with the municipality and the police, we aggregated our analysis to CBS neighbourhood levels.

Because of this decision, we had a sample size of 47 (neighbourhoods), which is a relatively small sample to measure a lot of predictors. This means we had less variance to analyse, and therefore less statistical power to find significant results. Even though our method comes with its limitations, it does not mean the analyses are meaningless.

Our analysis should be considered an exploratory view into the factors that affect residential burglaries in Gouda. Under no circumstance should policy changes be based on this report, as there are too many possibilities for spurious effects. However, the topics that stand out in this report may very well yield interesting new leads for further research.

BOX 1. EXPLANATION OF SOME STATISTICAL ANALYSIS

Linear regression analysis

Linear regression tests whether the independent variable (X) has an effect on the dependent variable (Y). A straight line in fitted into one formula with estimated values. When we calculate this in SPSS, it will give several parameters. Most important are:

- *F*: *F* checks whether the created model contributes to explain the dependent variable (Y).
- *b*: shows the gradient of the line. When *b* is positive, we have a positive relationship between X and Y (e.g. when *b* = .4, it means that when X goes right by 1, Y increases by .4). A negative *b* means a negative relationship (e.g. *b* = -.4 tells us that when X goes right by 1, Y will decrease by .4).
- B: is the standardised version of *b*. It allows us to compare different types of measurement (like kilos and height).
- *SE* (Standard Error): as stated above, the regression line calculates *estimated* values. The SE tells us how much variability there is between samples. If SE is big, this could indicate that *b* does not give an accurate reflection.
- R^2: gives the amount of variance explained by the model. For example, $R^2 = .35$ means that 35% of Y is explained by X. It also means that 65% of Y is explained by other variables.
- *p*: lastly, *p* shows whether *b* makes a significant contribution. When $p < .05$, it means that there is a 95% chance that the effect of *b* is genuine.

Independent t-test

If we want to know whether there is a significant difference between two variables, we compare their means. SPSS produces the t-statistic. The greater the magnitude of *t*, the bigger the chance that there will be a significant difference. This can be checked by looking at the *p*-value.

One-way ANOVA and post-hoc tests

Sometimes there are more than two variables involved. That is where the one-way ANOVA comes in. Like with linear regression, SPSS produces a F-statistic. If it is significant, there are indeed differences, but we are not sure yet which means actually differ from each other. This is why we conduct post-hoc tests: these show all possible combinations and their (lack of) significance. Again, the *p*-value shows the significance. Lastly, ω^2 is the effect size. Roughly said, ω^2 of .01, .04 and .14 are seen as small, medium and large effect sizes respectively.

Factor analysis

Sometimes a certain subject cannot be measured with only one item. A factor analysis is able to search for clusters of items (like questions from a survey) that fit together and measure the same subject. There are some assumptions. For example, two items need overlap, but also have to have some uniqueness. When the assumptions are met, the factor analysis shows whether the chosen items can be merged into one or more variables. When items actually are put together, we will have to test whether the new variable is reliable. This can be done by checking whether the Cronbach's α is higher than .7.

4 GOUDA & BURGLARY: STORIES, STATS AND FIGURES

Before doing the statistical analysis, we gathered general (descriptive) information about the topic. In this chapter, we created an overview of the burglary trends, hot spots in the city, vulnerable houses and a report about an on-site visit to one of these hot spots.

4.1 RESIDENTIAL BURGLARY TRENDS

The past year implies a break-through in the recent trends of residential burglary in Gouda, as the total amount of burglaries (whether attempted or completed) decreased with 21 percent regarding to 2016 (figure 5). The past year marks successes by different prevention tools such as WhatsApp groups, securing better locks, prevention by informing citizens and more ad-hoc police deployment and capacity. The current aim is, of course, to set this trend in the upcoming years.

Figure 3 - Total amount of burglaries per year.

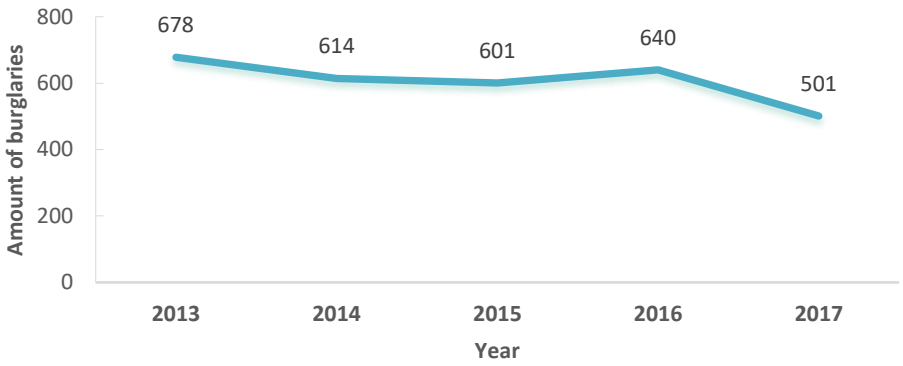
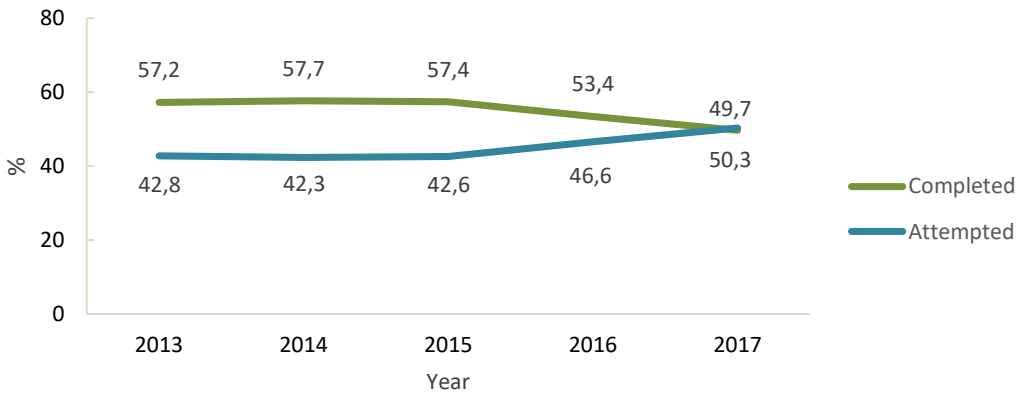


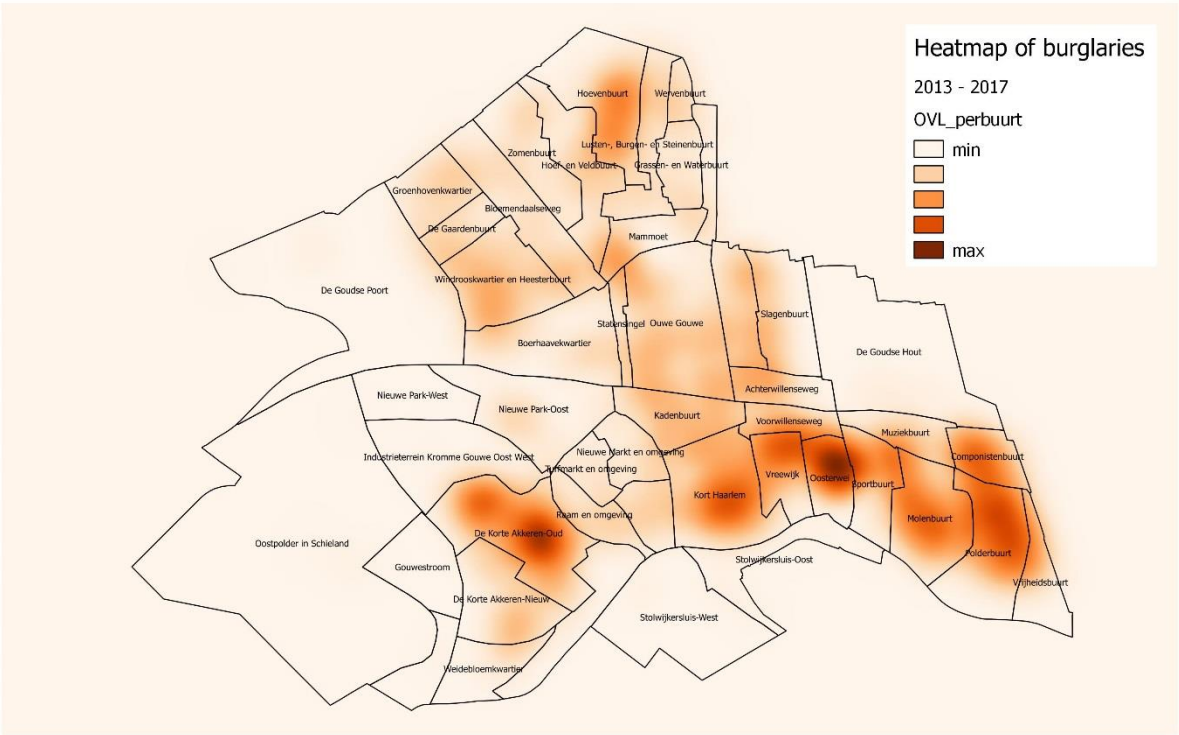
Figure 6 shows the completed and attempted amounts. It can be seen that the difference between the two became smaller over the years. The amount of completed burglaries decreased, which means there are less burglars that actually took some goods.

Figure 4 - Attempted and completed burglaries per year.



As the municipality was uncertain whether any hotspots could be indicated, because residential burglaries seemed to take place everywhere in the city, it is interesting to analyse if any critical areas can be calculated. By plotting all burglaries in QGIS, the following heatmap was created for the period of 2013 to 2017 (figure 7). It indeed confirms residential burglaries take place in almost any neighbourhood. But the frequency over five years clearly shows more problematic areas in the Hoevenbuurt, in the north of the city, and a lot of areas in the south of the city: houses in De Korte Akkeren-Oud, Kort Haarlem, Oosterwei and most neighbourhoods in the Goverwelle district are more often targeted.

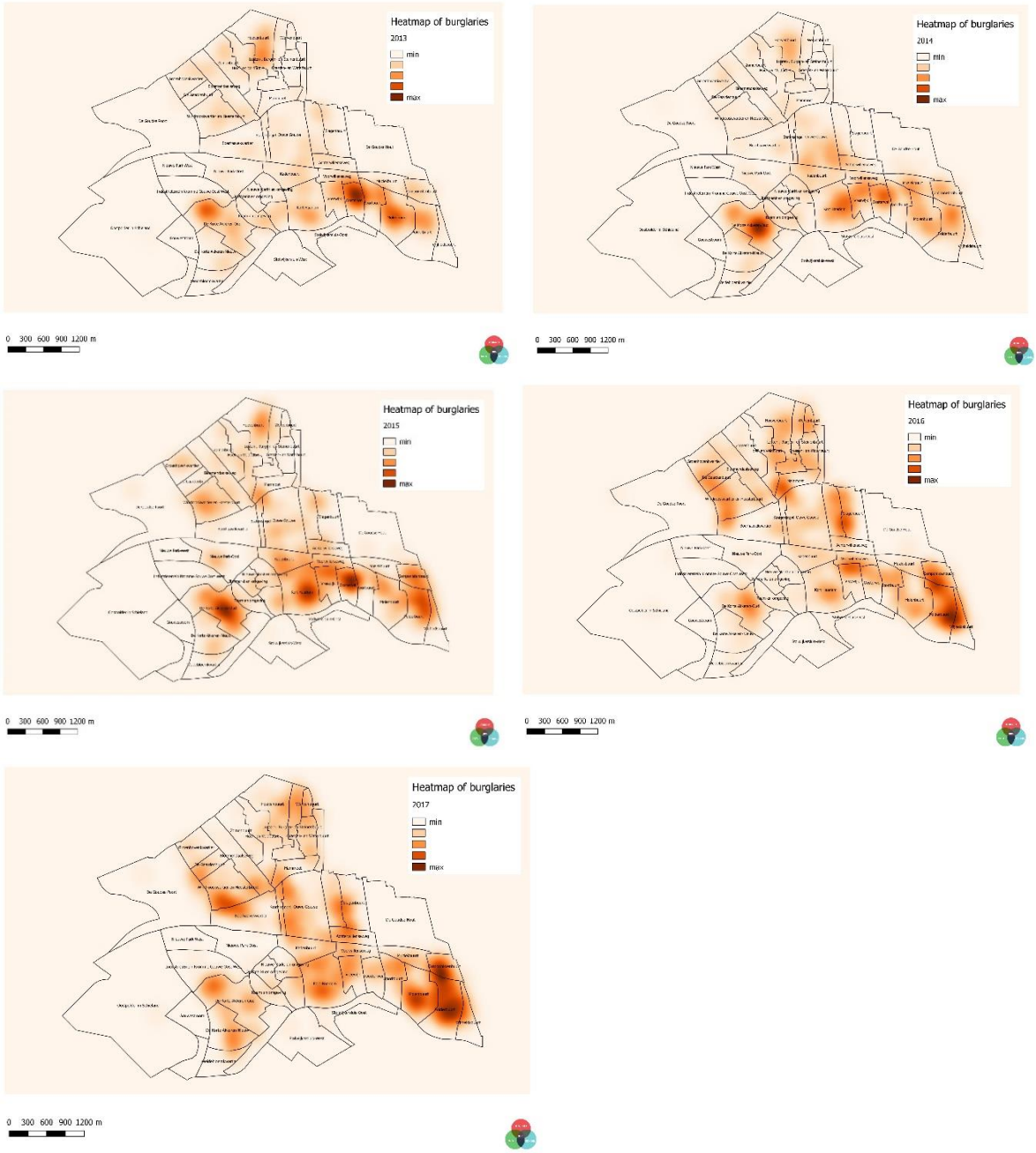
Figure 5 – Heatmap of burglaries in Gouda, 2013-2017.



0 300 600 900 1200 m

To check if these trends vary over the years, the heatmaps were also plotted per year. It shows that burglaries in the Goverwelle district increased over the years, while the burglaries in De Korte Akkeren slightly decreased in frequency. Furthermore, it can be seen that residential burglary became a bigger issue over the years in the northern part of town.

Figure 6 – Heatmaps of burglaries in Gouda, per year.



4.2 Repeated burglaries

Since a lot of neighbourhoods face the problem of burglary, it would be interesting to see whether there are a lot of repeated burglaries. For two reasons, it was difficult to check the exact amount of repeated burglaries per house:

1. The police data did not include a house number suffix. This causes that sometimes one address has in fact two (or more) houses.
2. Sometimes an address is actually a building, like an elderly home. When such a building has had, for example, 5 burglaries, we cannot say whether there were 5 unique victims or whether someone was burgled multiple times.

We did however check the amount of repeated burglaries on street level. When we combined the addresses from the police data and the BAG, we saw that 369 of 835 streets (with home addresses) have had at least one burglary between 2013 and 2017. This is a little over 44%. To see which streets were victimized the most, we used relative amounts, because the length of the streets differ. We divided the amount of burglaries by the amount of home addresses in that street. Other types of buildings were filtered out. Table 2 shows that relatively speaking, a street in the neighbourhood of Ouwe Gouwe has been burgled most often: this street has a relative score of 0.55.

Table 2 – Top 5 of relatively high amount of break-ins.

Street in the neighbourhood of:	District	Relative amount
Ouwe Gouwe	Noord	0.55
Muziekbuurt	Goverwelle	0.50
Oosterwei	Kort Haarlem	0.50
Molenbuurt	Goverwelle	0.50
Oosterwei	Kort Haarlem	0.43

Table 3 shows the streets with (relatively) the lowest amounts of break-ins. Remarkably, the amount of home addresses is much bigger than the amount of home addresses from the top 5.

Table 3 – Top 3 of relatively low amount of break-ins.

Street	Amount of burglaries	Amount of home addresses	Relative amount	Neighbourhood	District
Lekkenburg	1	368	0.003	Hoef- en Veldbuurt	Plaswijck
Raam	1	246	0.004	Raam e.o.	Binnenstad
Nieuwehaven	1	148	0.007	Nieuwe Markt e.o	Binnenstad

4.3 Modus Operandi and Type of House

The big differences between the amount of home addresses described above, might be explained by the type of house. For example, it could be that flats are more difficult to break-in than other types of houses. The BAG dataset provided information about the type of house. In Dutch, these are: 2-*onder-1 kap-, etage/flat-, geschakelde-, hoek-, rij-* and *vrijstaande woningen* and *woonwagens/boten*. Table 4 shows the amount and percentage of burglaries per type of house. Note that the amount of *woonwagens/boten* is very low. Therefore, the calculations for this type of house are not representative.

Table 1 – Amount of burglaries per type of house.

Type of house	Amount of burglaries	Percentage of burglaries
Rij	1080	35.6%
Hoek	740	24.4%
Etage/flat	670	22.1%
2-onder-1 kap	223	7.3%
Vrijstaand	128	4.2%
Geschakeld	50	1.6%
Woonwagen/boot	5	0.2%
Other/unkown	141	4.6%

Table 5 shows correlations between type of house and amount of burglaries. As we can see, the correlations between flats and burglaries is not significant. The correlations between *vrijstaande huizen* and burglaries are average and the correlations between burglaries and *2/1 kap, geschakelde, hoek* and *rijtjeshuizen* are the strongest. This means that relatively seen these four types of houses are most targeted.

Table 5 – Correlations between the amount of burglaries and type of house.

Type of house	Total amount	Attempted amount	Completed amount
2/1 kap	.65**	.60**	.61**
Flat	.17	.25	.06
Geschakeld	.65**	.61**	.61**
Hoek	.65**	.61**	.59**
Rij	.64**	.61**	.59**
Vrijstaand	.34*	.17	.47**
Woonboot/wagen	-.17	-.19	-.13

*$p < .05$, **$p < .01$

The police often reported from which side the burglar entered the house: via the front, the back or the side of the house. It turns out that burglars tend to go in via the back of the house (1010 times), followed by the front (926 times) of the house. The side of the house has been reported 115 times. The low amount could be explained by the fact that a lot of houses do not have a side entrances (linked houses (rijtjeshuizen) for example).

The police did not only report the *side* of entrance, but also the *way* of entrance. These were either via fences, balconies, doors, windows, roofs or others, like basements. Table 6 shows the total amounts. Windows and doors are clearly the most frequently used ways to enter a house. This is not surprising, since a lot of research says that burglars can easily access a house via open windows. Since burglars most often break-in via the back of the house, it might suggest that people also do not lock their back door, or that it is not too hard to open a door quickly.

Table 6 – Amounts of way of entrance.

Way of entrance	Amount
Window	1297
Door	875
Balcony	75
Fence	53
Roof	34
Other	46

Lastly, table 7 shows the correlations between type of house and way of entrance. Among others, we see that *rijtjeshuizen* have significant correlations with all the ways of entrance. The number two (corner houses) and three (flats) also have a lot of significant correlations. This might explain why the burglary rate is high for these houses; apparently burglars have multiple potential options to enter such a house successfully.

Table 7 – Correlations between type of house and way of entrance.

Type of house	Fence	Balcony	Door	Window	Roof	Other
2-onder-1 kap	.23	.30*	.22	.54**	.22	.22
Etage/Flat	.24	.67*	.51*	.43**	.03	.26
Geschakeld	.02	-.12	.03	.44**	.31	-.18
Hoek	63*	.36*	.82*	.88**	.04	.51**
Vrijstaand	.11	.27	.16	.19	.13	.19
Rij	.75**	.47**	.89**	.87**	.37*	.65**
Woonwagen/boot	.34*	-.14	-.02	.31*	.07	-.11

* $p < .05$, ** $p < .01$.

4.4 On-Site visit to Goverwelle and Oosterwei

On November 13, the project team had a tour through the frequently targeted Goverwelle district and Oosterwei neighbourhood with a police officer from Goverwelle[2]. Goverwelle is, especially in the last two years, one of the most targeted districts. The police officer tried to show us why burglars -mostly youth from Goverwelle itself- find it interesting to break into houses. She explains that relatively low punishments, high chance of reoffending and a social criminal network on high schools encourages youth into the criminal path. Some see burglary as the starting point of their criminal career. Characteristic is that the burglars in Goverwelle often break-in in their own nearby environment, but not in their own street.

Figure 7 - Project team on the bike exploring Gouda Goverwelle (own photo).

The police believes in encouraging surveillance both by the police as by citizens. We have cycled through two neighbourhoods in Goverwelle (Polderbuurt and Molenbuurt), which are neighbourhoods with a lot of social housing and lack of social cohesion. The built environment encourages burglary, because the kitchen is in the front while inhabitants are in the back of the house in the living room during the evening. This leads to unawareness of what is happening on the street. Burglars demolish kitchen windows or climb on sheds to get access to windows on the first floor. The police is often trying to create awareness. An example is

[2] Chapter 4.4. is based on stories and statements from the local police officer and is not checked by the Gouda police or municipality.

what they call '*witte voetjes*': police officers mark, for example, open windows with white foot steps to show citizens that it might as well been the footstep of a burglar. Figure 10 until 12 show typical streets in Goverwelle and illustrate how anonymous a burglar can be during night-time. Invisibility plays an important factor, next to building characteristics, lighting and social control. As Goverwelle lacks social cohesion, anyone on the street is quite anonymous. While Oosterwei clearly shows more social cohesion due the existence of a mosque and a community centre, the police officer states.

Figure 8 – (left) A typical house in Goverwelle with the kitchen in the front. In this particular kitchen window, you can see a camera installed. Next to WhatsApp groups, consequently more private security cameras are installed and registered (own photo).

Figure 9 – (right) A typical street in the Molenbuurt, Goverwelle. At night, this street is very dark and again, the kitchens are situated in the front of the house (own photo).

Figure 10 – (left) In the middle of a quiet area in the Polderbuurt, we found a static camera on the corner of the street. When the police and citizens cannot see something, they hope the camera will.

Figure 11 – (right) While green areas benefit health, the high grass on the right side of the picture results in a lack of visibility.

5 RESULTS

After gathering the descriptive information and on-site research, we continued with the statistical analysis. As previously stated, the analysis was conducted in four different topics namely time and weather, neighbourhood conditions, perceived safety and social cohesion (Appendix II). In this chapter, the method per analysis will be further explained and the results will be discussed.

5.1 TIME AND WEATHER

Several time and weather variables could be predictive indicators of when a burglary take place. For example, as mentioned in the literature review, harsh weather could deter burglars to commit a burglary and pleasant weather could mean that people are more likely to leave the house for a while in the evening. For this reason, we collected some historic weather conditions from the KNMI database and other variables from the municipality as school holidays for each day in the period 2013-2017;

- The total amount of rain fallen in a day (in mm);
- The average wind speed (in km/h);
- Hours of rain on a given day (in 0.1 h);
- Hours of sun on a given day (in 0.1 h);
- How many cloud coverage there was on a given day (in okta: a cloud cover estimation with values between 0 and 9);
- The mean temperature (in C°);
- For each day which season it belongs to;
- For each day which day of the week it was;
- And vacation/holidays.

Because the duration of a burglary cannot be pinpointed with a high accuracy in many cases, we decided to only use cases of burglaries that were discovered on the same day. There were 3038 burglaries between the period 2013-2017, but only 1545 of these burglaries had a reliable timestamp and were useable for analysis. Every burglary was categorised by part of day: morning, noon, evening or night. The police calculated the part of day by dividing the alleged duration of the burglary in half and adding it to the earliest time the burglary could have taken place. As can be seen in table 8, the burglaries are also divided by attempt and completion.

Table 8 - The amount of total-, attempted- and completed burglaries per part of day.

	Total amount of burglaries	Attempted amount of burglaries	Completed amount of burglaries
Morning	180	77	103
Noon	402	141	261
Evening	452	179	273
Night	511	269	242
Total	1545	666	879

The six weather variables mentioned above were merged as independent variables into models to check whether they had an influence on attempted and/or completed burglaries. Linear regression (for a basic explanation of the statistical analysis, see Box 1, p. 26) with the total amount of burglaries as the dependent variable, showed that the *lower* the mean temperature[3] and the *higher* the average wind speed[4] is on a day, the *higher* the amount of burglaries.

The second dependent variable was the amount of completed burglaries. Linear regression showed the same effects: the *higher* the average wind speed[5] and the *lower* the mean temperature[6] the *higher* the amount of successful burglaries on that day. Moreover we found that the *lower* the amount of hours of sun[7] on a day, the *higher* the amount of completed burglaries.

The third dependent variable was the amount of attempted burglaries. Linear regression showed that the *lower* the mean temperature[8], the *higher* the amount of attempted burglaries.

Lastly, weekend, day of week and vacation/holiday were independent variables in the linear regression analysis. Results showed a strong correlation between the first two variables, but we could not find a correlation for attempted and completed burglaries respectively.

5.1.2 Part of day

We did an linear regression analysis for burglaries committed at different day parts. We found results in burglaries committed at noon and in the evening. We have not found results in burglaries committed at morning or at night. Linear regression shows an association with the mean temperature[9] and average wind speed[10] in burglaries committed at noon. When we look at the completed burglaries committed at noon, we only find a small correlation with the mean temperature[11].

There are more burglaries in the evening on vacation/holidays than on regular days[12]. Also, more burglaries happen during the evening when the wind speed is higher[13] and the mean temperature is lower[14]. This is different when we look at only the completed burglaries committed in the evening. We could not find a significant correlation with vacation days, but we found a relation with the day of the week[15], the mean temperature[16] and the average wind speed[17]. We could not find these correlations with the attempted burglaries committed at noon.

---

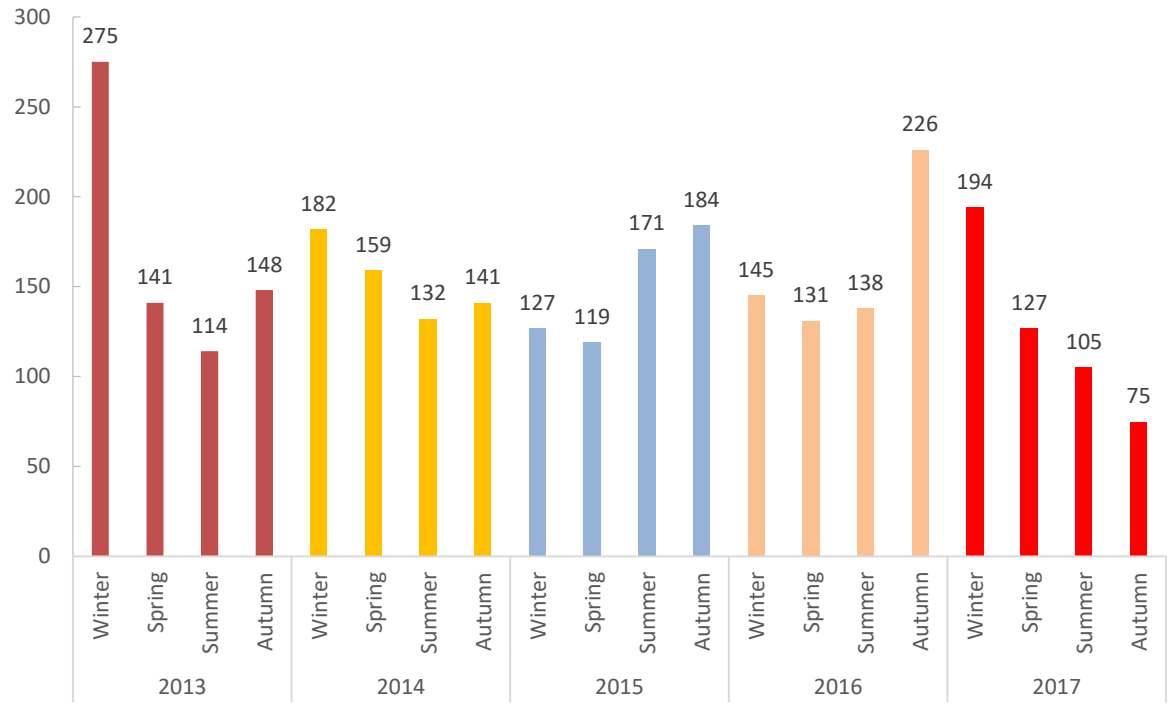
[3] $b = -.019$, $SE = .006$, $\beta = -.121$, $R^2 = .037$, $p < .05$
[4] $b = .019$, $SE = .009$, $\beta = .072$, $R^2 = .037$, $p < .05$
[5] $b = .021$, $SE = .006$, $\beta = .094$, $R^2 = .025$, $p < .05$
[6] $b = -.009$, $SE = .003$, $\beta = -.073$, $R^2 = .025$, $p < .05$
[7] $b = -.002$, $SE = .001$, $\beta = -.091$, $R^2 = .025$, $p < .05$
[8] $b = -.010$, $SE = .003$, $\beta = -.089$, $R^2 = .014$, $p < .05$
[9] $b = -.007$, $SE = .002$, $\beta = -.084$, $R^2 = .017$, $p < .05$
[10] $b = .009$, $SE = .004$, $\beta = .061$, $R^2 = .017$, $p < .05$
[11] $b = -.005$, $SE = .002$, $\beta = -.075$, $R^2 = .013$, $p < .001$
[12] $b = .065$, $SE = .030$, $\beta = .052$, $R^2 = .052$, $p < .05$
[13] $b = -.015$, $SE = .003$, $\beta = -.157$, $R^2 = .052$, $p < .001$
[14] $b = .013$, $SE = .004$, $\beta = .080$, $R^2 = .052$, $p < .001$
[15] $b = .017$, $SE = .008$, $\beta = .084$, $R^2 = .041$, $p < .05$
[16] $b = -.008$, $SE = .002$, $\beta = -.109$, $R^2 = .041$, $p < .001$

[17] $b = .008$, $SE = .003$, $\beta = .063$, $R^2 = .041$, $p < .05$

5.1.3 Seasonal burglaries

We conducted a one-way ANOVA with post-hoc tests (see Box 1, p. 26) to test whether the different seasons had an effect on the amount of burglaries. Because the Levene's test was significant, we used Welch's F. We found that there is a significant effect on the number of burglaries between seasons. There are *more* burglaries in the autumn than in the summer[18], and *more* burglaries in the winter than in the spring[19] and summer[20]. Figure 14 illustrates this result.

Figure 12 - Amount of burglaries per season per year.



We used linear regression to calculate if weather conditions influence seasonal burglaries. We found out that the *higher* the hours of sun[21] on a day in the spring, the *higher* the attempted burglaries. Furthermore, if the mean temperature on a day in the autumn is *lower*[22] and if there is *more* cloud coverage[23], there are *more* attempted burglaries on that day. Finally, we found that there was a small significance in the hours of rain on a summer day, where more hours of rain seems related to the amount of attempted burglaries on that day[24].

---

[18] $F(3, 533.545) = .240$, $p < .05$, est. $\omega^2 = .025$
[19] $F(3, 533.545) = .341$, $p < .001$, est. $\omega^2 = .025$
[20] $F(3, 533.545) = .374$, $p < .001$, est. $\omega^2 = .025$
[21] $b = -.030$, $SE = .003$, $B = -.089$, $R^2 = .049$, $p < .05$
[22] $b = .006$, $SE = .002$, $B = -.123$, $R^2 = .050$, $p < .05$
[23] $b = .106$, $SE = .044$, $B = .286$, $R^2 = .050$, $p < .05$
[24] $b = .007$, $SE = .003$, $B = .207$, $R^2 = .054$, $p < .05$

5.2 Neighbourhood Conditions

As mentioned in the literature review, neighbourhood characteristics can tell a lot about the look and feel of a neighbourhood. This is why we included some general information, namely:

1. Population density: the amount of citizens divided by neighbourhood size.
2. Household density: the amount of households divided by neighbourhood size.
3. Building density: the amount of buildings divided by neighbourhood size.
4. '*WOZ-waarde*': the mean values per neighbourhood.

The population- and household density were retrieved from the CBS (buurt statisiek). The building density was calculated with data from the BAG. Lastly, the dataset with the *WOZ-waarde* was downloaded from QGIS. Table 9 shows the statistics per neighbourhood.

The four variables were merged into a model to check whether they have an effect on burglaries. For the first linear regression, the dependent variable was the total amount of burglaries. The results showed that there are more burglaries when there is a *high* population density[25], but also when there is a *low* household density[26]. This seems conflicting, since one would say that neighbourhoods with low population density also have a fewer amount of households. Earlier research however, was also contradicting about this topic. On the one hand, neighbourhoods with fewer amounts of households are more quiet and this gives the burglar the opportunity to look around and check which houses are easy targets. On the other hand, neighbourhoods with high population density make it easier for the burglar to find his/her victim without being spotted as a suspicious figure.

Linear regression with the attempted amount of burglaries only had the same effect for population density[27]. For completed burglaries, only the household density had a significant effect[28].

[25] $b = 22.90$, $SE = 9.59$, $\beta = -1.12$, $R^2 = .18$, $p < .05$
[26] $b = -63.19$, $SE = 30.34$, $\beta = -1.51$, $p < .05$
[27] $b = 13.32$, $SE = 5.34$, $\beta = 1.17$, $R^2 = .18$, $p < .05$
[28] $b = -35.08$, $SE = 15.78$, $\beta = -1.61$, $R^2 = .18$, $p < .05$

Table 9 - Population-, household- and, building density (per ha) and WOZ-waarde per neighbourhood.

Neighbourhood	District	Population density	Household density	Building density	WOZ-waarde
Raam en omgeving	Binnenstad	118.13	64.69	77.38	167.660
Nieuwe Markt en omgeving	Binnenstad	93.75	59.38	93.81	159.228
De Baan en omgeving	Binnenstad	91.93	48.67	67.13	182.518
Turfmarkt en omgeving	Binnenstad	91.58	52.50	62.75	217.147
Nieuwe Park-Oost	Binnenstad	38.39	18.54	23.29	202.033
Nieuwe Park-West	Binnenstad	-	-	-	-
Groenhovenkwartier	Bloemendaal	66.22	31.74	29.26	177.521
Windrooskwartier en Heesterbuurt	Bloemendaal	64.19	28.57	30.75	198.865
Bloemendaalseweg	Bloemendaal	59.24	20.53	26.37	134.797
De Gaardenbuurt	Bloemendaal	56.45	21.36	22.64	409.595
Boerhaavekwartier	Bloemendaal	54.88	31.80	42.08	147.424
De Goudse Poort	Bloemendaal	4.33	2.71	5.23	135.280
De Korte Akkeren-Oud	De Korte Akkeren	115.14	55.0	57.05	136.812
De Korte Akkeren-Nieuw	De Korte Akkeren	72.58	29.39	32.09	143.520
Weidebloemkwartier	De Korte Akkeren	30.47	13.24	15.75	137.216
Industrieterrein Langs de Hollandse IJssel	De Korte Akkeren	2.20	1.04	3.67	124.871
Componistenbuurt	Goverwelle	105.17	36.49	37.33	209.911
Molenbuurt	Goverwelle	94.28	33.72	39.03	173.060
Polderbuurt	Goverwelle	71.75	27.27	27.95	192.094
Sportbuurt	Goverwelle	56.27	26.82	23.50	120.398
Vrijheidsbuurt	Goverwelle	42.90	15.75	15.75	246.298
Muziekbuurt	Goverwelle	39.39	16.67	16.89	196.055
Kadenbuurt	Kort Haarlem	100.63	47.59	52.81	150.316
Kort Haarlem	Kort Haarlem	82.50	34.70	36.52	190.053
Oosterwei	Kort Haarlem	73.67	31.11	44.28	92.578
Vreewijk	Kort Haarlem	51.33	25.24	29.62	118.241
Voorwillenseweg	Kort Haarlem	44.43	15.0	18.21	296.506
Achterwillenseweg	Noord	77.71	31.47	38.76	98.253
Ouwe Gouwe	Noord	67.09	31.03	35.99	148.902
Wethouder Venteweg	Noord	66.70	31.50	30.75	191.118
Slagenbuurt	Noord	46.48	15.91	17.48	224.782
Statensingel	Noord	44.86	49.29	30.75	177.389
Wervenbuurt	Plaswijck	108.67	49.17	43.83	189.542
Hoef-en Veldbuurt	Plaswijck	73.72	38.38	45.42	158.844
Grassen- en Waterbuurt	Plaswijck	71.88	26.88	29.75	298.556
Hoevenbuurt	Plaswijck	70.43	27.45	27.62	174.138
Lusten-, Burgen- en Steinenbuurt	Plaswijck	69.38	26.73	29.23	205.449
Zomenbuurt	Plaswijck	54.84	29.35	26.71	175,766
Mammoet	Plaswijck	42.44	19.17	20.22	187.973
Bodegraafsestraatweg	Plaswijck	36.33	13.13	22.11	210.621
Stolwijkersluis-Oost	Stolwijkersluis	13.58	6.25	8.63	277.682
Stolwijkersluis-West	Stolwijkersluis	1.49	0	0	376.000
Industrieterrein Kromme Gouwe	Westergouwe	5.30	3.25	7.85	109.771
Oostpolder in Schieland	Westergouwe	1.89	1.24	1.48	154.492
Ringvaartbocht	Westergouwe	1.00	0	0	308.428

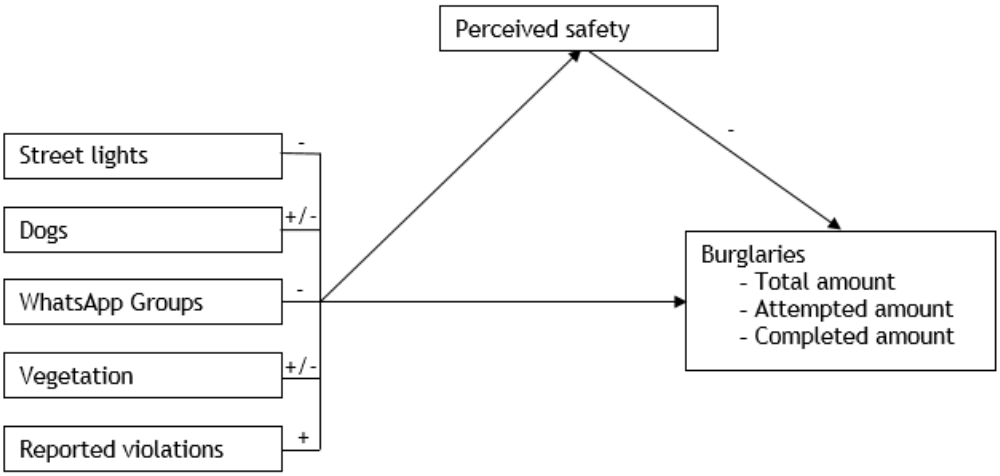
5.3 Perceived Safety

One of the predictors of burglaries could be perceived safety. Literature found several indicators that might reduce burglaries, but is not always unanimous. Figure 15 shows which variables were taken into account during this research, based on the literature review:

- Street lights: as can be seen in Figure 15, we expect a direct negative effect[29] of street lights. Which means that when there are more street lights, the amount of burglaries will decrease. Several authors found that burglars have a high risk of being seen when there are many street lights.
- Dogs: researchers are divergent about the role of dogs. On the one hand, dogs could notify their owners when someone comes near the house. And since dog owners need to walk through the neighbourhood regularly, suspicious behaviour is hard to stay unnoticed. On the other hand, those walks could also create a nice opportunity to break-in while the owner is gone.
- WhatsApp groups: we also expect a negative effect for the amount of WhatsApp groups. These groups keep each other up to date about the whereabouts in the neighbourhood. This way, suspicious behaviour will be noticed quicker.
- Vegetation: vegetation is also a tricky one. On the one hand, vegetation can cause people to go outside which, again, makes it harder for burglars to take a good look around. On the other hand, vegetation can decrease visibility and create nice spots to hide. Moreover, during summer people tend to leave their windows open, which makes it easier for a burglar to come in.
- Reported violations: we expect a positive effect for the amount of reported violations. This means that when the amount of reported violations increase, the amount of burglaries will also increase. Researchers believe that the welfare of a neighbourhood influences the amount of burglaries, since they found that neat neighbourhoods with educated people have less break-ins.
- Perceived fear of citizens: perceived fear might be a mediator for the variables above. A mediation variable (Z) tries to explain the process that underlies the relationship between the independent variable (X) and the dependent variable (Y). In this case, we expect that, for example, street lights (X) will make people feel safe to go outside (Z). When people go outside more often, strange persons will be noticed, which makes it harder for a burglar to check the neighbourhood for a suitable victim and the amount of burglaries will be lower (Y).

[29] Direct effect: the effect of the independent variable (X) on the dependent variable (Y).
Positive direct effect: when X increases, Y also increases (and the other way around).
Negative direct effect: when X increases, Y decreases (and the other way around).

Figure 13 – Perceived safety model.

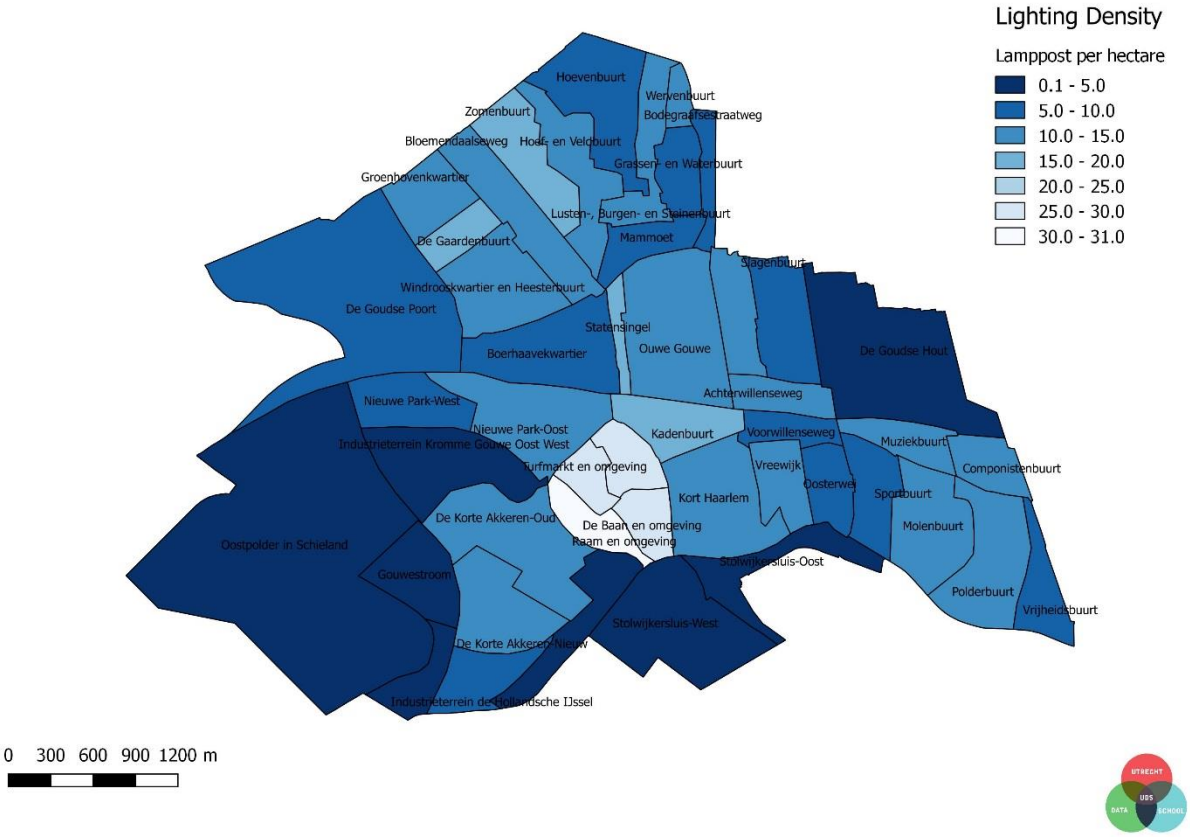


5.3.1 Street Lighting

In order to analyse any lighting density, we have plotted a shapefile with all lampposts managed by the municipality, which we retrieved from GIS.GOUDA and counted the sum of lampposts in each neighbourhood in QGIS. This sum we used to calculate a lighting density (sum of lampposts divided by the surface of the neighbourhood). This value gives us an amount of lampposts per hectare in each neighbourhood. We can hypothesize that the more lamppost per hectare implies the more lit the neighbourhood is. We have visualized the differences on a map in QGIS, to show which neighbourhoods are more dark than others (figure 16). It is obvious that the inner city is the most lit area within the city. The most dark areas are in the polder areas (Oostpolder in Schieland) and sparsely dense area of Stolwijkersluis. It is interesting to see differences in the urban areas. So is the Vrijheidsbuurt darker than other neighbourhoods in the Goverwelle district. To check whether the lighting density of the neighbourhood has a causal relationship, we ran a regression analysis for these variables. Nevertheless, we found that the lighting density has no significant effect on the relative amount of burglaries[30]. This implies we have no causal relationship between the lighting density of the neighbourhood and the amount of burglaries.

[30] $F(1,43) = 3.664$, $p > .05$, $\Delta R^2 = .57$

Figure 14 – Lighting density per neighbourhood.



5.3.2 Dogs

To measure dogs, the *'(honden)belastingdienst'* provided data with the amount of dogs per postal code. Due to the privacy of dog owners, we could not get the amount of dogs per house. The analysis for dogs were therefore performed with postal codes instead of specific addresses. The postal codes with dogs were linked to the corresponding neighbourhoods. The mean amount of dogs per neighbourhood is M = 44.24. Also for dogs we used relative amounts, because neighbourhood sizes differ. Therefore, the amount of dogs was divided by the amount of households per neighbourhood. Next, linear regression analysis with the amount of dogs as the independent variable, and the amounts of respectively total-, attempted- and completed burglaries as the dependent variables were calculated. Results of all linear regressions showed significant effects[31]. The results show a positive relationship between the amount of dogs and the amount of burglaries.

To research the difference between attempted and completed burglaries, a t-test was conducted (see Box 1, p. 26). The independent variable was whether or not there were houses with dogs within the postal code. The results showed that the amount of postal codes that have

[31] Total: b = .016, SE = .004, β = .54, R^2 = .29, p < .001
Attempted: b = .007, SE = .002, β = .41, R^2 = .17, p < .001
Completed: b = .009, SE = .002, β = .60, R^2 = .36, p < .001

a dog was bigger for completed burglaries (*M* = .67, *SE* = .47) than for attempted burglaries (*M* = .58, *SE* = .49). This result was significant[32], but the effect size was low: *r* = .18.

The results implicate a positive effect of having dogs and the amount of burglaries; meaning that the higher the amount of dogs, the higher the amount of burglaries. However, it is not clear what causes this effect, since most likely there are more variables involved. It could be that people that have experienced multiple break-ins decided to get a dog to protect themselves. Another possible explanation is that dog owners have to go for a walk regularly and leave their windows open, since they will only leave for a short amount of time. On the other hand; when people go out often, they might recognize suspicious figures easier.

5.3.3 WhatsApp Groups

WhatsApp prevention groups are one of the most current forms of neighbourhood prevention. The municipality encourages these initiatives by facilitating new registrations, communicating with the administrators and provide budget for street signs (see figure 17).

Figure 15 – WhatsApp street sign in Oosterwei (own photo).

The amount of WhatsApp groups in the city is constantly growing - at the moment of data collection 111 groups were registered by the municipality. It is possible that there are more neighbourhoods, streets or blocks active on WhatsApp, however they are not officially registered at the municipality as this is not obliged.

[32] $t(18225) = -12.438$, $p < .001$

Figure 16 - The distribution of WhatsApp groups per neighbourhood in Gouda (n = 111).

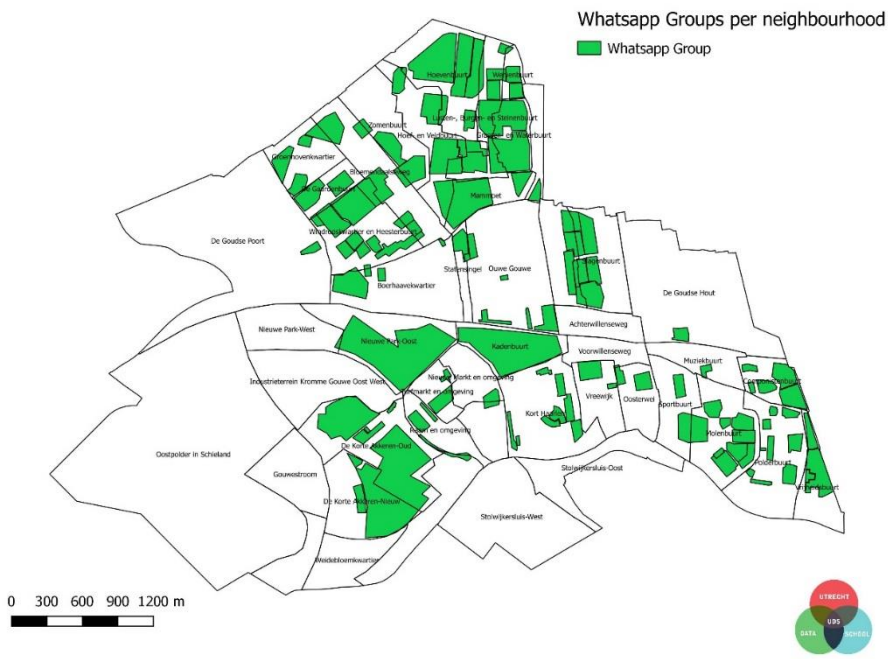


Figure 18 shows the distribution of the 111 WhatsApp groups active in neighbourhood prevention. As some groups overlap the geographical boundaries of neighbourhoods, the total amount of WhatsApp groups was calculated by how many groups were presence within the geographical boundaries. This implicates that some groups are counted more than once as they cross over those boundaries. Windrooskwartier en Heesterbuurt (N=13) had the most groups, followed by two neighbourhoods in the Goverwelle district (Molenbuurt and Polderbuurt respectively, both N=11). Regarding our data, not in every neighbourhood a WhatsApp group is active: for example not in Oostpolder in Schieland or in Stolwijkersluis. When active, those WhatsApp groups form the eyes and ears of the neighbourhood as the police cannot always be present. But how effective are the prevention groups in the case of residential burglaries?

In a regression analysis, we found that the amount of WhatsApp groups within the neighbourhood[33] has a significant correlation with the amount of burglaries. The regression model has an estimated variance of 49.7% for the total amount of burglaries within the neighbourhood[34]. When looked at the relative amount of burglaries in the neighbourhood, no significant effect has been found[35].

Next, by trying to get more insights in the effect of WhatsApp groups, we collaborated with the municipality and made a small survey which was conducted to the 111 registered group administrators (see survey at Appendix III). With a response rate of 33% (N=37), the response is quite low. The first question, questioning in which neighbourhood the WhatsApp group is active, has led to a variation of (loose) interpretation such as that respondents did not claim to be an

[33] $b = 13.95$, $p < .001$
[34] $F(1, 44) = 36.63$, $p < 0.001$, $\Delta R^2 = .442$
[35] $F = .412$, $p < 0.01$, $\Delta R^2 = -.14$

administrator or stated in which district instead of which neighbourhood (such as 'Goverwelle'). Hence why only 27 cases were traced back to one of the Gouda districts. This makes the sample too small for statistical analysis. Therefore we only can state some descriptive results which are not representative for all WhatsApp groups in Gouda.

According the survey and showed in table 10, the most groups are active in the Bloemendaal district, which is relatively seen almost comparable to the total amount of groups (22.2 and 23.98 percent respectively). The average amount of people within the WhatsApp group differs a lot. The average in total is 47.78 with a standard deviation of 32.70 which implicates group size differs a lot in between groups. The smallest group according the survey consists of 10 people while the biggest group consists of 124 people.

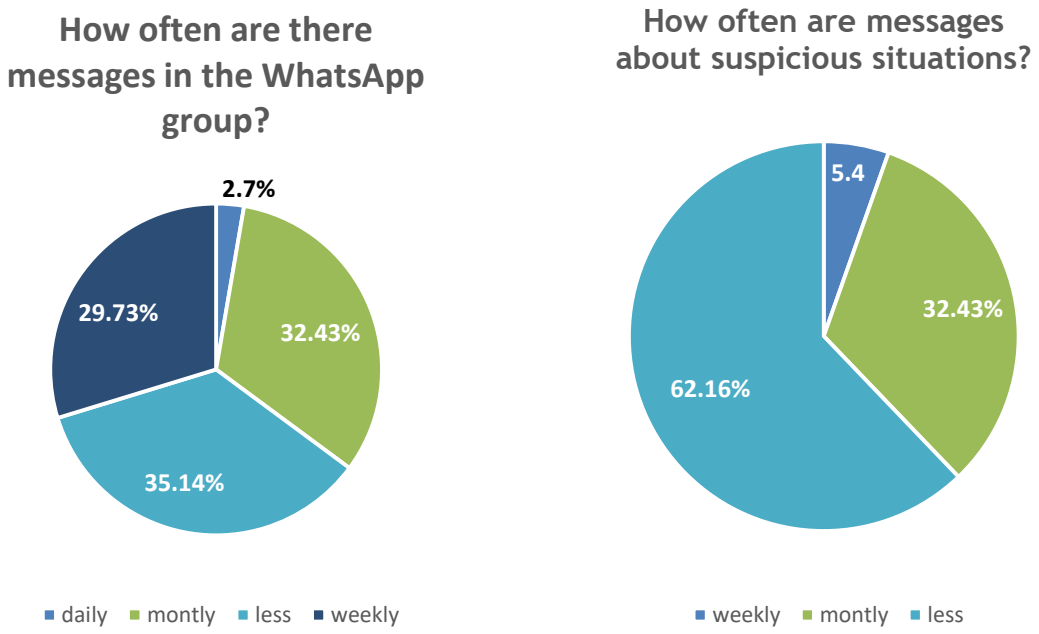
Table 10 – Descriptive results of WhatsApp groups survey (N=27).

Neighbourhood	# of survey	% of survey	# of total	% of total	Avg. People in group	Avg. messages per month
Bloemendaal	6	22.2	41	23.98	43.67	11.33
Plaswijck	5	18.52	36	21.05	76.4	26
Goverwelle	4	14.82	34	19.88	41.25	42.4
Kort Haarlem	4	14.82	17	9.94	33.25	15.5
Binnenstad	3	11.11	10	5.85	72	3.33
Noord	3	11.11	20	11.69	67.67	10.66
Stolwijkersluis	1	3.70	0	0	30	60
De Korte Akkeren	1	3.70	13	7.60	19	5
Westergouwe	0	0	0	0	0	0
Total	**27**	**100%**	**111**	**100%**	**47.78**	**16.89**

Furthermore was asked when the WhatsApp group was established – it is interesting to see that the first groups already formed in 2014 while still quite recently new groups are established. On average, within the groups 16.89 messages were send on a monthly basis. On average, it seems that in Stolwijkersluis and in Goverwelle more messages are send. But it has to be noted that this number is based on perception, Stolwijkersluis only consist of one group and there are a lot of outliers as some have indicated to share more than 100 messages per month.

Looking at figure 19, it seems that the groups are not used on a daily basis a lot. The respondents of the survey almost evenly stated to use their group weekly (29.73%), monthly (32.43%) or less (35.14%). When looking to the contents of the messages, it seems that most messages are non-related to crimes and suspicious situations as the over 62% of respondents states that suspicious things are mentioned less than per month.

Figure 17 – Amount and contents of messages in WhatsApp surveillance groups.



It would be very informative to connect questions from the survey such as activity, group size and date of establishment to further investigate whether the group apps have an effect on burglaries. As previously stated, as the first question could not trace back the groups to neighbourhoods or more even more specific, the registered groups, the information could not be match to data of residential burglaries in the neighbourhood.

5.3.4 Vegetation

The data about vegetation in the municipality of Gouda was retrieved with QGIS from the GIS.GOUDA database. The data consisted of the amount of trees and the size of green areas. The green areas, or vegetation, include grass, hedges and different types of plants. We deliberately kept the amount of trees as a variable on its own, as it was expected that they might affect the sight of the houses. To control for the different sizes of neighbourhoods, the vegetation and amount of trees were divided by the size of the neighbourhoods in hectare to get relative amounts.

Linear regressions have been calculated for the amount of vegetation and trees as the independent variables and the respectively total-, attempted- and completed amount of burglaries. Results show significant relationships for both vegetation[36] and trees[37] on the total- and completed amount of burglaries, but not for attempted burglaries.

Remarkably, results show that the higher the amount of trees, the lower the amount of burglaries. But for vegetation it is the other way around. The more vegetation, the more burglaries. This could be explained by the literature described in the literature review. Like with dogs, Troy et al. (2012) mentioned that trees make it appealing to go outside. This way,

---

[36] Total: $b = .44$, $SE = .22$, $\beta = .35$, $R^2 = .12$, $p < .05$ Completed: $b = .33$, $SE = .11$, $\beta = .50$, $R^2 = .20$, $p < .005$
[37] Total: $b = -31.36$, $SE = 14.26$, $\beta = -.38$, $p < .05$ Completed: $b = -19.07$, $SE = 7.09$ $\beta = -.44$, $p < .01$

suspicious figures might be noticed quickly. On the other hand, Fisher and Nasar (1992) said that especially low dense vegetation, like the vegetation measured during this research, make it easier for a burglar to hide during his/her escape. Also for this variable it is not clear why these effects occurred, since there probably are confounding variables. For example, pruning might play a role. And also, some trees decrease visibility, but this is not always the case.

5.3.5 REPORTED VIOLATIONS

The municipality keeps track of all violations citizens reported, when they were reported and when they were resolved. The data file had all kinds of violations, they varied from a lack of parking lots to garbage laying all over the place. For this research, only violations concerning nuisance by others (from youth or graffiti, discrimination and intimidation), garbage and other kinds of waste/filth (like dead animals, dog poo and broken glass) were taken into account, because these might lead to degradation of the neighbourhood. Table 11 shows the amount of the total- and different types of reported violations.

To measure the effects of reported violations, the amount was first divided by the amount of citizens per neighbourhood, since all citizens can report something. After conducting linear regressions with the relative amounts, only the amount of reported violations and the total amount of burglaries showed significant correlations[38]. This means that the higher amount of reported violations, the higher the amount of burglaries. It is not likely that reported violations directly influence the amount of burglaries, but one explanation could be that violations indicate that citizens in 'clean' and friendly neighbourhoods have higher norms and values and therefore less potential burglars.

[38] $b = .43$, $SE = .08$, $\beta = .61$, $R^2 = .38$, $p < .001$

Table 11 – Various kinds of reported violations per neighbourhood and district.

Neighbourhood	District	Total	Garbage	Waste/filth	Nuisance
De Korte Akkeren-Nieuw	De Korte Akkeren	1041	39	26	4
Kort Haarlem	Kort Haarlem	607	67	31	3
Ouwe Gouwe	Noord	543	62	62	6
De Baan en omgeving	Binnenstad	493	57	37	9
Nieuwe Park-Oost	Binnenstad	490	47	54	9
Windrooskwartier en Heesterbuurt	Bloemendaal	429	43	30	1
Kadenbuurt	Kort Haarlem	414	72	28	9
Lusten-, Burgen- en Steinenbuurt	Plaswijck	409	35	54	3
Nieuwe Markt en omgeving	Binnenstad	406	67	38	17
Polderbuurt	Goverwelle	401	50	47	3
Boerhaavekwartier	Bloemendaal	369	25	28	2
Bloemendaalseweg	Bloemendaal	361	29	42	1
Raam en omgeving	Binnenstad	358	58	25	0
Molenbuurt	Goverwelle	306	27	32	3
Turfmarkt en omgeving	Binnenstad	299	32	27	7
Slagenbuurt	Noord	280	32	51	5
Groenhovenkwartier	Bloemendaal	277	29	46	0
Vreewijk	Kort Haarlem	270	38	21	1
Oosterwei	Kort Haarlem	269	48	17	3
De Korte Akkeren-Oud	De Korte Akkeren	255	154	80	3
Hoevenbuurt	Plaswijck	224	14	39	4
De Goudse Poort	Bloemendaal	223	10	40	3
Componistenbuurt	Goverwelle	222	28	19	2
Achterwillenseweg	Noord	214	18	22	1
Mammoet	Plaswijck	196	9	17	0
Sportbuurt	Goverwelle	190	14	16	0
Grassen- en Waterbuurt	Plaswijck	155	7	17	1
Zomenbuurt	Plaswijck	152	18	27	0
Voorwillenseweg	Kort Haarlem	138	14	11	0
Wervenbuurt	Plaswijck	127	12	20	1
Muziekbuurt	Goverwelle	115	6	7	0
Hoef-en Veldbuurt	Plaswijck	108	55	35	1
Wethouder Venteweg	Noord	76	9	5	1
Vrijheidsbuurt	Goverwelle	71	4	10	1
Statensingel	Noord	70	11	10	1
Industrieterrein Kromme Gouwe	Westergouwe	69	1	2	1
De Gaardenbuurt	Bloemendaal	68	4	8	1
Oostpolder in Schieland	Westergouwe	65	9	7	0
Weidebloemkwartier	De Korte Akkeren	59	3	9	0
Bodegraafsestraatweg	Plaswijck	41	1	4	0
Stolwijkersluis-Oost	Stolwijkersluis	39	1	4	0
Ringvaartbocht	Westergouwe	31	6	1	0
Industrieterrein Langs de Hollandse IJssel	De Korte Akkeren	27	4	3	2
Gouwe Stroom	Westergouwe	27	1	0	2
Stolwijkersluis-West	Stolwijkersluis	12	0	0	0
Nieuwe Park-West	Binnenstad	11	3	8	0

5.3.6 Perceived Safety

Each year the municipality asks a random selection of people to take part in a survey about the city and their wellbeing, the so called Stadspeiling. Five items were selected to measure perceived safety, our mediation variable (see 5.3.). A factor analysis (see Box 1, p. 26), which tests whether items fit together for a certain subject, was conducted with these five items. After checking the assumptions of the factor analysis[39], we were in the clear the create two new variables:

- Feeling unsafe in the neighbourhood. The variable included the questions "*Voelt u zich wel eens onveilig in uw buurt?*" and "*Voelt u zich vaak, soms of zelden onveilig?*" [40]
- Fear of burglary. This scale included the questions "*In mijn buurt wordt regelmatig ingebroken*", "*Ik ben wel eens bang dat er bij mij wordt ingebroken*" and "*Ik heb maatregelen getroffen om te voorkomen dat er bij mij wordt ingebroken*"[41]

5.3.7 Perceived Safety Model

Eventually all the variables above were used for the perceived safety model. The model was tested six times: the total-, attempted- and completed amount of burglaries times two: once with the feeling unsafe as a mediator and once with the fear of burglary.

For all models, the independent variables (X) of the linear regression were the same: street light density, relative amount of dogs, number of WhatsApp groups, relative amount of vegetation, relative amount of trees and relative amount of reported violations, as can be seen in the perceived safety model below. Also step 2 (the effect from the independent variable X on the mediator Z) was the same. For the unsafe feeling, only the amount of dogs was significant[42]. The fear of burglary did not have any significant effects. This means there were no mediations for the models with this mediator.

For the total amount of burglaries, the first step measured direct effects from the independent variables on the total amount of burglaries. Results showed significant effects[43] of vegetation, reported violations and WhatsApp groups. This implicates that the higher the amount of vegetation, the higher the amount of burglaries. And also, the higher the amount of reported violations, the higher the amount of burglaries. These results match the ones we saw earlier in this chapter. This time, however, also WhatsApp groups had a significant effect: the more WhatsApp groups in the neighbourhood, the higher the amount of burglaries.

The linear regression for the attempted amount of burglaries also started with the direct effects of the independent variables on the attempted amount of burglaries. (X --> Y). The

---

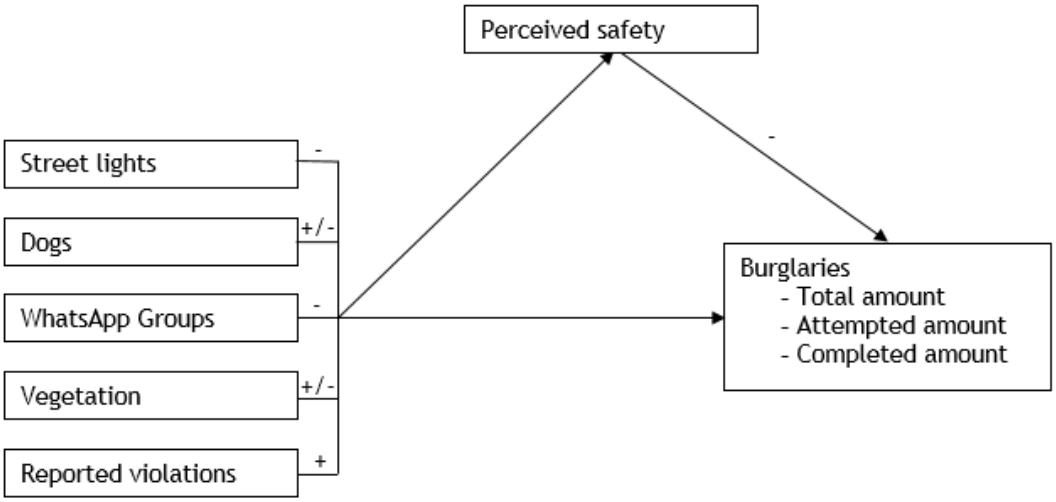
[39] The Kaiser-Meyer-Olkin was .51, two items had an eigenvalue above 1, which was confirmed by the scree plot. Together they explained 88.14% of the variance.
[40] Cronbach's α = .78
[41] Cronbach's α = .76
[42] $b = -.10$, $SE = .03$, $\beta = -.90$, $R^2 = .36$, $p < .001$
[43] Vegetation: $b = .44$, $SE = .19$, $\beta = .35$, $R^2 = .55$, $p < .05$, reported violations: $b = .53$, $SE = .15$, $\beta = .76$, $p < .01$ and WhatsApp groups: $b = .006$, $SE = .003$, $\beta = .28$, $p < .05$

same significant results[44] were found for the amount of reported violations and the amount of WhatsApp groups.

Finally the direct effects of the completed amount of burglaries was tested. Significant results[45] were found for the amount of vegetation, the amount of trees and the amount of reported violations. The results match the results above again, with the addition that the higher the amount of trees, the lower the amount of completed burglaries.

The results are similar to the results we found when we measured them separately. Remarkably, the model shows that WhatsApp groups have a positive relationship with the amounts of total and attempted burglaries. An explanation for this might be that burglars know the neighbours in the WhatsApp group, and earlier research showed that most burglars break-in near their own house and at people they know. It is also possible that the neighbourhood decided to set up a WhatsApp groups because there were a lot of burglaries.

Figure 18 – Perceived safety model.



5.4 Social cohesion

5.4.1 Methods

Social cohesion was initially operationalised as a combination of sub-optimal proxies in our research proposal, such as the number of schools in a neighbourhood, or the number of community centres. The idea was to get an insight into the degree of contact between people in the same neighbourhood, based on the amount of opportunities for them to get together. However, Gouda later provided us with an elaborate survey performed among approximately 2000 residents, asking a wide variety of questions about safety, welfare, and other social issues: the Stadspeiling. One sub-set of question proved to be particularly relevant for our research, asking questions about the degree of contact, but also the atmosphere in the neighbourhood,

[44] Reported violations: $b = .33$, $SE = .09$, $\beta = .86$, $R^2 = .45$, $p < .01$, WhatsApp groups: $b = .004$, $SE = .002$, $\beta = .32$, $p < .05$

[45] Vegetation: $b = .31$, $SE = .10$, $\beta = .47$, $R^2 = .57$, $p < .05$, Trees: $b = -15.91$, $SE = 7.75$, $\beta = -.37$, $p < .05$, Reported violations: $b = .19$, $SE = .08$, $\beta = .53$, $p < .05$

or the acceptance of cultural differences. This became our primary measurement for social cohesion.

Table 12 – Question matrix social cohesion – Stadspeiling 2016.

Question	N	Min	Max	Mean	SD	Factor
6.1 I have a lot of contact with other people in my neighbourhood	1854	1	5	2.40	1.141	1
6.2 I know what happens in my neighbourhood	1823	1	5	2.56	1.038	1
6.3 I feel responsible for the livability in my neighbourhood	1842	1	5	2.03	.906	1
6.4 I feel comfortable in the street I live in	1864	1	5	1.78	.917	2
6.5 In my neighbourhood I'm treated with respect	1827	1	5	1.84	.890	2
6.6 In my neighbourhood cultural differences do not cause problems	1710	1	5	2.33	1.164	2
6.7 I can get help from my neighbours	1811	1	5	1.80	.911	1
6.8 In my neighbourhood people are always ready to help each other	1752	1	5	2.27	1.036	1
Valid N (listwise)	1563					

Eight questions were asked in 2016, while one was dropped in 2017 bringing the total for that year to seven. To give an idea about the differences between the two years: 65% of people reported high levels of contact with their neighbours in 2016, compared to just under 60% in 2017. The questions and their descriptive statistics are listed in table 12 for 2016, and table 13 for 2017. We ran a factor analysis on all questions per year, which resulted in a two-factor solution: one pertaining to having no or negative contact with people in the neighbourhood, factor 1, and one pertaining to respectful behaviour and good atmosphere in a neighbourhood, factor 2. These factors were the same between years, but since 2017 did not ask about one item, it was not a part of the final factor solution either.

Table 13 – Question matrix social cohesion – Stadspeiling 2017.

Question	N	Mi	Max	Mean	SD	Factor
6.1 I have a lot of contact with other people in my neighbourhood	1822	1	5	2.50	1.175	1
6.2 I know what happens in my neighbourhood	1783	1	5	2.67	1.079	1
6.3 I feel comfortable in the street I live in	1831	1	5	1.80	.994	2
6.4 In my neighbourhood I'm treated with respect	1767	1	5	1.86	.948	2
6.5 In my neighbourhood cultural differences do not cause problems	1681	1	5	2.37	1.195	2
6.6 I can get help from my neighbours	1749	1	5	1.85	.978	1
6.7 In my neighbourhood people are always ready to help each other	1680	1	5	2.30	1.095	1
Valid N (listwise)	1506					

We aggregated our factor scores to neighbourhood level, so each neighbourhood got the average value of the factor score of all respondents. Because our analyses span multiple years, we then took the 2016 and 2017 neighbourhood averages together into one averaged construct.

Our measures of social cohesion were plotted on the map of Gouda, so we can see the distribution, and single out problem areas. In figure 21 we see the distribution of respect across Gouda. Here we see that it is mainly the northern neighbourhoods that have an agreeable atmosphere, while the southeast and Korte Akkeren part of the city has more disrespect, and people feel less at ease there. The distribution of the contact measure tends to follow the same distribution. As shown in figure 22, low scores mean positive contact, and high scores mean negative or no contact. The northern neighbourhoods have good neighbourly contact, as does the southern part of the city. In Korte Akkeren and the southeast part of the city there is less contact in most neighbourhoods. We can say there is definitely a correlation between the two measures of social cohesion, but there are enough theoretical and statistical differences in the measures to include both in the analysis.

Figure 19 – Respect per neighbourhood.

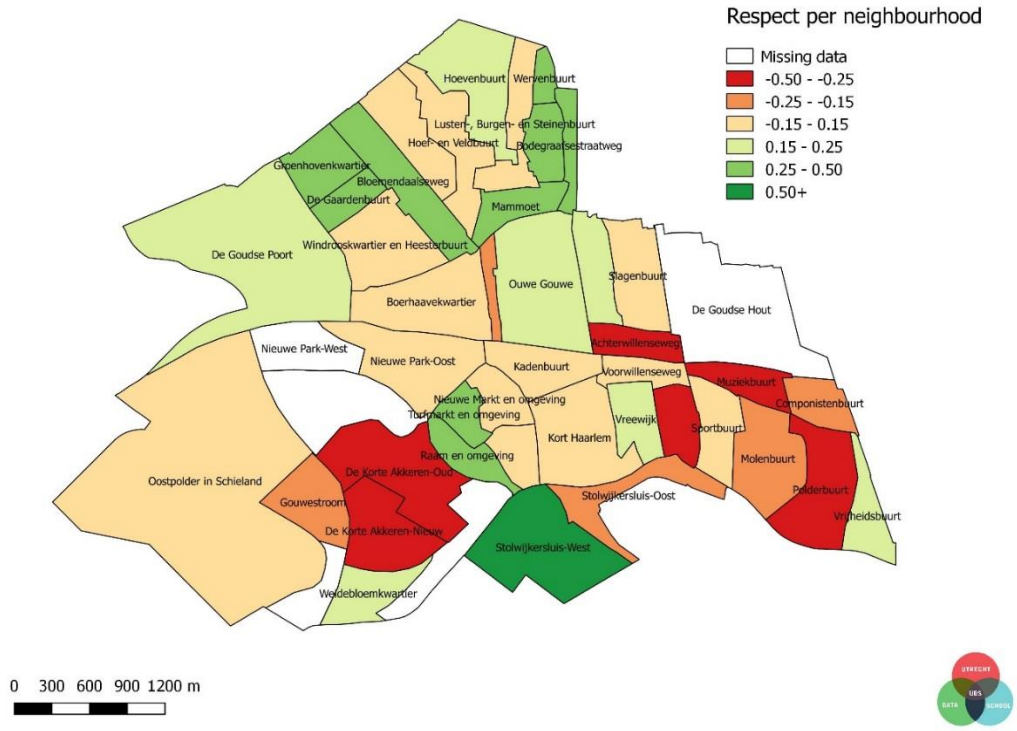
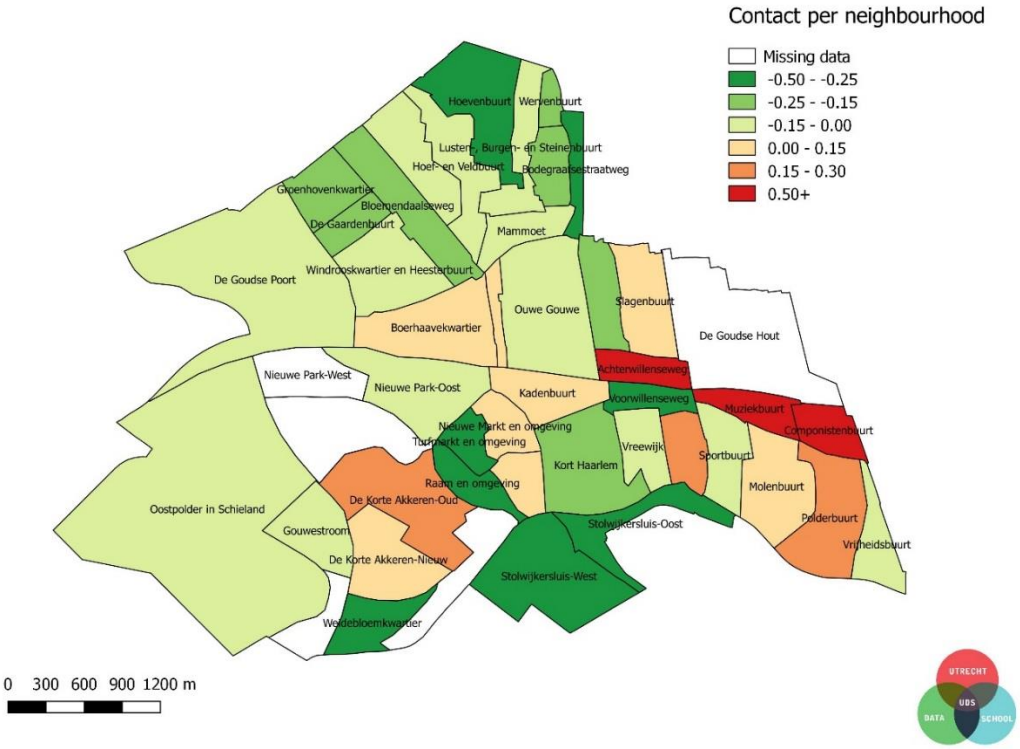


Figure 20 – Contact per neighbourhood.

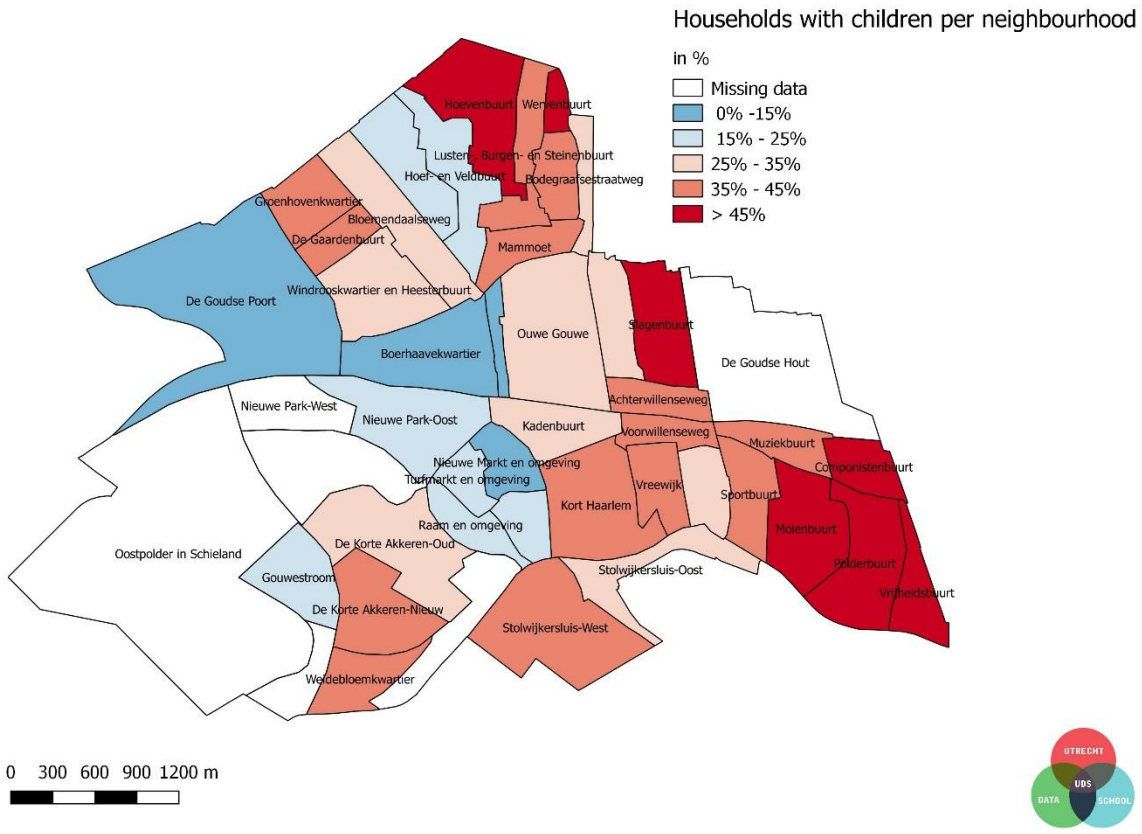


While these measures of social cohesion were our primary interest, we also include a number of demographic controls in our analysis. This data was retrieved from the CBS neighbourhood statistics, with the exception of income, which came from the CBS regional income research. We control for population density, children, immigrants and income.

We expect single houses are easier to target: if many people live close together, disturbances will potentially be seen by more people. Also, we expect that it is harder to break into a residence high up in a flat compared to for example the backdoor of a ground-level house.

We look at the proportion of households with children in a neighbourhood, as an additional measure of social cohesion: children play outside, which increases the number of watchful eyes in a neighbourhood. They also tend to go to school close to home, which brings parents together who otherwise might not talk to each other. However, neighbourhoods with kids may also result in doors being left open more easily, providing more opportunities for burglary. In figure 23 we see the distribution across Gouda, mainly showing us there are a lot of households with children in the southeast and northern part of the city.

Figure 21 – Households with children per neighborhood.



Furthermore we look at the proportion of people with a non-western immigration background in a neighbourhood, According to literature both ethnic heterogeneity and residential mobility lead to an increase of burglaries, because there is less cohesiveness in these groups (Bernasco

& Luykx, 2003; Erete, 2013; Hirschfield & Bowers, 1997). It is therefore expected that higher numbers of immigrants increase the number of burglaries, both successful and total.

Finally we control for the average income in a neighbourhood, as we expect that richer neighbourhoods are more valuable to potential burglars. Income might also explain a large part in variance for the immigration control variable since the two measures are negatively correlated (r = -.711, p <.001). In order to draw meaningful conclusions about the effects of immigrants, who tend to have a lower socioeconomic status, this is the most used control.

Of course, as in any research, many more factors were assumed to be of influence, but because of the privacy concerns, aggregation of our data to neighbourhood level resulted in only 40 cases where enough people filled in the Stadspeiling per neighbourhood, while still retaining access to income information. Including more variables, while possibly of influence, would not result in detectable differences anyway, because of our low sample size.

For our analysis, we run two different multivariate regression models. In the first model, we look at the effects of our predictors on the total amount of burglaries, corrected by the number of households in a neighbourhood. In our second model, we look at the effects of our predictors on the success rate of all burglaries: we take the amount of successful burglaries divided by the total burglaries, corrected for the amount of households, as our dependent variable.

5.4.2 Results

In model one we test the hypothesis that social cohesion has a negative effect on the total amount of burglaries per household in a neighbourhood. The model is significant, meaning that the predictors altogether predict a significant amount of variance in total burglaries[46].

Table 14 – Ordinary least squares regression of total burglaries per household on social cohesion.

Variable name	B	SE
Constant	-83.349	62.889
Limited contact	28.223	44.477
Respectfulness	-4.568	43.440
Population density of households	-.761*	.442
Proportion households with children	2.274***	.544
Proportion immigrants	3.166**	.942
Mean income	.003*	.002
Adjusted R^2	.568	

* p < .05, ** p <.01, *** p <.001

[46] $F(6, 32) = 9.344$, p <.001, $R^2 = .568$

The predictors and their effects are shown in table 14. Neither of our Stadspeiling constructs provide significant results[47]. However, our other proxies and control variables do show a pattern. As expected, neighbourhoods with high population density have less burglaries.

The proportion of households with children also significantly predicts burglaries: in neighbourhoods with many children more burglaries happen per household. Every percent-point increase in the proportion of households with children results in 2.27 more burglaries[48]. This same relationship holds for the proportion of immigrants: for every percent-point increase in non-western immigrants, 3.16 more burglaries happen in that neighbourhood[49]. Finally, there is a significant effect of income: for every thousand euro increase in mean neighbourhood income, 3 more burglaries happen[50].

In model two we test the hypothesis that social cohesion has a negative effect on the success rate of burglaries in a neighbourhood. The model is significant, meaning that the predictors altogether predict a significant amount of variance in the success rate of burglaries[51].

Table 15 - Ordinary least squares regression of success rate of burglaries per household on social cohesion.

Variable name	**B**	**SE**
Constant	.772	.191
Poor contact	-.235	.135
Respectfulness and atmosphere	-.083	.132
Population density of households	-.001	.001
Proportion households with children	-.005*	.002
Proportion immigrants	.000	.003
Mean income	.000	.000
Adjusted R^2	.568	

* $p < .05$, ** $p < .01$, *** $p < .001$

The predictors and their effects are shown in table 15. Just like in our model of total burglaries, neither of our Stadspeiling constructs provide significant results[52]. Population density, the number of immigrants, and income are not significant in this model, while they were significant predictors of the total amount of burglaries. While they do influence how many burglaries happen, they have no effect on how successful those burglaries are.

[47] $p = .530/2$ and $p = .917/2$ for poor contact and respectfulness respectively
[48] $b = 2.274$, $p < .001$
[49] $b = 3.166$, $p = .002/2$
[50] $b = .003$, $p = .069/2$
[51] $F(6, 32) = 2.637$, $p = .034$, $R^2 = .205$
[52] $p = .092$ and $p = .535/2$ for poor contact and respectfulness respectively

The only significant predictor for success rate is the proportion of households with children: in neighbourhoods with many children, burglaries have a lower success rate. Every percent-point increase in the proportion of households with children results in a decrease of .005 in the success rate[53].

To understand why none of our social cohesion measures provide significant results, we look at some correlations between our social cohesion measures and our control variables. As previously stated, our control variables are proxies[54] for other forms of social cohesion that might not be measured by our Stadspeiling constructs, but at the same time there is overlap between the concepts. However, our multicollinearity was not at problematic levels so we decided not to remove any variables.

For example, if ethnic heterogeneity is high, measured by the amount of immigrants, people are also less likely to know their neighbours: a correlation shows that when the proportion of immigrants in a neighbourhood is high, there is less contact between people living in that neighbourhood ($r = .457$, $p = .003$). Similarly, when there are more immigrants in a neighbourhood, there is less contact ($r = -.575$, $p < .001$).

Additionally, when we look at the correlations between income and our social cohesion measure, we see that people in high income neighbourhoods have more contact with each other ($r = -.414$, $p = .009$), and the atmosphere in their neighbourhood is better ($r = .488$, $p = .002$). These are the most important and significant correlations between social cohesion and our control variables. This explains why our measures of social cohesion, when controlling for such similar measures, do not significantly explain additional variance in the amount of burglaries, whether the total or the success rate.

[53] $b = -.001$, $p = .007$

[54] Variable that is not in intself directly relevant, but serves in place of an unobservable or immeasurable variable.

6 CONCLUSION

The initial purpose of this Utrecht Data School project was to gain new insights that were hidden in numbers and various datasets of the municipality of Gouda and the police, which could lead to new understandings of residential burglaries in the city. The focus on environmental data was considered by considering the crime triangle, whereas the environment of Gouda has not been well researched yet in respect of residential burglaries. This is why we asked ourselves: "*Which environmental factors can be identified and visualised to control or even prevent both attempted and completed burglaries in Gouda?*". During the project, we have analysed various topics and variables that might play a role in this. We have divided different variables into four main topics; time and season; neighbourhood conditions; perceived safety and social cohesion as potential predictors of burglary (see Appendix II for an overview of all researched predictors).

6.1.1 SUMMARY OF THE FINDINGS

First, the most important findings of this research project will be summarized here. Considering time and season, we have found a significant correlation between average wind speed and the temperature on a day for burglaries, both completed and attempted, that happens during the afternoon or evening. Therefore can be concluded the more windy or the lower the temperature on that day is, more burglaries take place in the afternoon or evening. Furthermore we found significant correlations for burglaries that took place in the autumn and winter and during weekends.

For neighbourhood conditions, we have created a model that included population-, household- and building density and the mean *WOZ-waarde* per neighbourhood. We found that a high population density and a low household density significantly influenced the amount of burglaries.

We also have created a model for perceived safety with different predictors such as street lights and dog ownership. When we analysed these variables separately, we found a significant results of the amount of dog ownership that implies neighbourhoods with more dogs were often more victimized than neighbourhoods with less dogs. Moreover we found separate relations for vegetation, the amount of trees and reported violations (MOG). When all variables were merged in one model, we found a significant relation for a higher amount of vegetation and an increase in burglaries. In contradictory, a lower amount of trees relates with a higher amount of burglaries. Furthermore we found a relation between when there are more reported violations or WhatsApp prevention groups in the neighbourhood relates with an increase in residential burglaries.

Lastly, we have measured social cohesion with predictors from the Stadspeiling and control variables such as children, income, immigrants and household density. Children resulted to be an important predictor of burglaries; it leads to less completed burglaries however leads to more attempted burglaries. Furthermore there seems to be a link between the amount of immigrants, income and the amount of burglaries.

6.1.2 Differences between Attempted and Completed Burglaries

As mentioned at the beginning of this report, the total amount has been splitted in attempted and completed burglaries to make a further distinction in burglaries. Sometimes variables were significant for either attempted or completed variables, but we did not find big differences that might explain why the completed burglaries decreased, while the attempted burglaries increased.

6.1.3 Conclusions

The findings of this research project has resulted in some interesting significant predictors of environmental factors that influences burglary. As mentioned in chapter three, these analysis should be considered as an exploratory view into predictors that affect residential burglaries in Gouda. As there are many possibilities for spurious effects, our findings will not suggest changes in policy immediately. Nevertheless, it can be concluded that the predictors summarized in our findings are interesting predictors that yield to new leads for further research. As the environment of Gouda was not yet taken into account as explanatory factor of residential burglaries, these predictors suggest that indeed factors as vegetation, WhatsApp groups, population density or children interplays – either positive or negative – with residential burglaries in the neighbourhood. Furthermore, challenges we have faced during the analysis has given us and the municipality insights in data efficiency and working with bigger datasets: resulting from the project of the Utrecht Data School, the municipality has made an agreement with the local police to share more data to work effectively in the near future.

6.2 Recommendations

As we conclude our project with some very interesting predictors, it is recommended for Gouda and the police to further research what these predictors further could imply for chancing policies by acknowledging that further research also implies further investigation of these predictors. This report will be concluded with recommendations for the municipality in two ways; based on contents or based on working data-driven within the organization.

6.2.1 Further Investigating of Current Topics

Even though interesting findings mentioned above were significant, they do not explain all variance in the models (e.g. of neighbourhood conditions, perceived safety and social cohesion) which implicates that there are more factors involved in explaining burglary. This might also explain why some results seem contradicting. That is why it is recommended to research some topics further and more in depth. Hereby we recommend to reconsider the level of research. We have chosen for neighbourhoods as the smallest geographical bounded area. Nevertheless, burglaries and predictors could vary over streets and thus vary between neighbourhoods, or even at house level (see Box 2).

BOX 2: EIGHT STREETS WITH DIFFERENT CHARACTERS

In the descriptive part of the results, we have identified five streets with the highest amount of burglaries (respectively 'high') and three streets with the lowest amount of burglaries (respectively 'low'). With all the researched variables in mind, we compared the two groups. Of course, results cannot be generalised, because the sample size is small and not based on theoretical boundaries whatsoever, but we do see some remarkable differences. For example, three out of five 'highest' streets had one (or even two) schools in the surrounding environment, while the 'lowest' streets did not. Also one of the 'highest' streets had a small playground and one was nearby a children's library. This implies that there are a lot of children in the neighbourhood. Only one of the 'lowest' streets had a small soccer field.

Moreover, the 'lowest' streets were all located in the centre or next to a shopping centre, different kind of religious buildings and other culture-related buildings. Only one of the 'highest' streets had those kinds of buildings in the surrounding area as well.

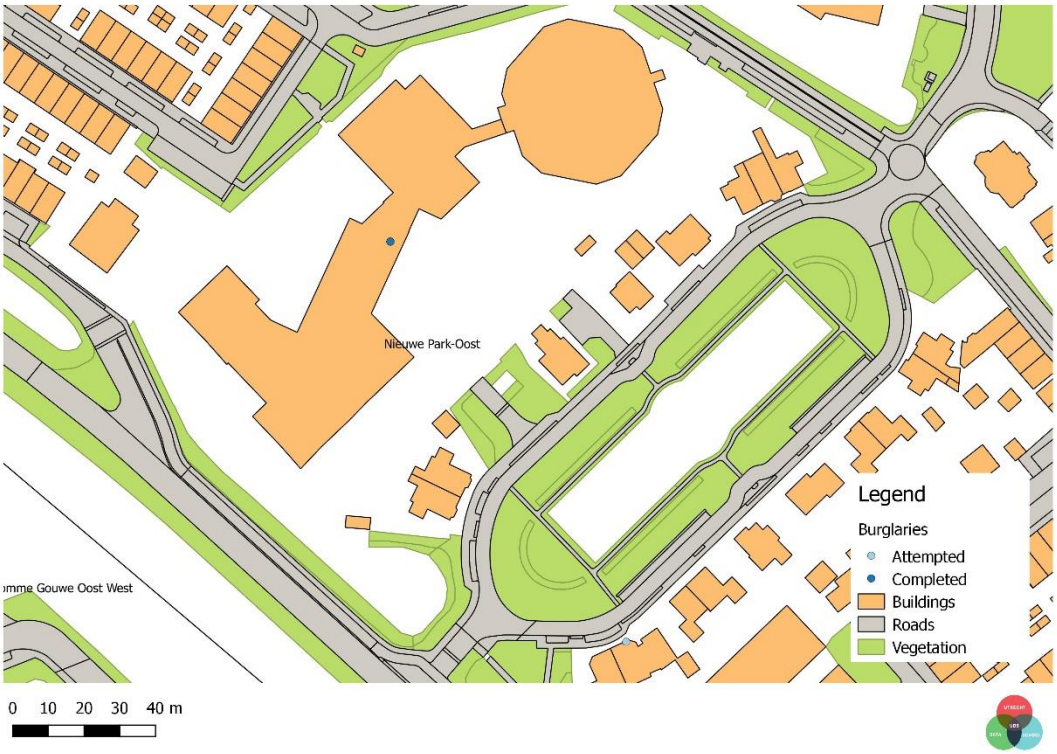
Besides the differences, there were also overlapping street characteristics. For example, there is no distinction between type of house, width of streets, and sport facilities, as both kind of streets had these features. It would be interesting to investigate these aspects for a larger amount of streets.

On top of that, the findings of this research recommends more detailed data. Working with data from different authorities means differences in detail of data. To show its implications, figure 24 shows the case of an elderly home where four times a break-in was reported. Nevertheless, this elderly home consists of over forty homes that share the address. As the police does not register house letters and such in their addresses, this example shows how irregularities occur within the data according to the reality. Even though the good news is that the municipality of Gouda and the police agreed upon a more intensive collaboration with data, it is relevant to further discuss and agree on definitions and details of data.

In line with above, the following recommendations for the predictors within this research would improve the level of detail in the data:

- For analysing the use of WhatsApp groups, it is recommended to conduct a new survey amongst the WhatsApp groups or as item in the Stadspeiling, since not all groups are officially registered. According the regression analysis, WhatsApp groups have a positive relation with the amount of burglaries. Since these groups exists to prevent burglaries, it could be valuable to analyse differences before and after the groups were established in the neighbourhood;

Figure 22 – An elderly home in Gouda with multiple break-ins but a single addresses according the data.



- Social cohesion is something that might differ per street as well. This might also explain why there were no significant results. Some more variables could be taken into account, like events, neighbourhood activities like barbecues and other festivities, community centres, sport clubs, schools and homogeneity within the neighbourhood;
- Vegetation also differs per neighbourhood, as some streets have a lot of hedges, grass or even a pond, and other streets do not have many green areas. Also, the pruning of plants and hedges might differ within the neighbourhood;
- Information about trees would be most usable when they are investigated at house level. As literature suggests, trees could decrease visibility, which makes it easier for the burglar to enter the house unseen. However, not all trees actually influence the sight. And again, pruning might be important as well. More detailed information about trees is necessary to better understand their role. Moreover, people might also have trees in their gardens and those were not available to us;
- Dogs could be investigated at house level. Now we were only able to check the amount of dogs per postal code, but it would be interesting to see whether houses with dogs have had less or more burglaries than houses without dogs;
- The amount of repeated burglaries deserves some more attention. Literature suggest earlier victims are vulnerable, as well as their neighbours;
- Lastly, for weather more factors could be included to get an even more complete description of how weather influences burglaries.

6.2.2 FURTHER GENERAL RECOMMENDATIONS

During the project, it was difficult to gather all data that was needed from different authorities, such as the police, and especially the data about the burglaries was rather important. Furthermore we could not get (full) access to some data we requested. As previously stated, it would be valuable information to gather information about dogs on house level but due privacy reasons we had to work with a dataset on postcode level – which also differs from other data on neighbourhood or street level.

Before we obtained access to the police dataset, we have used the burglary dataset of the municipality to define our models. When we got the police data, we came across some difficulties and differences between or within the data sets:

1. The definitions of certain variables were not always clear, like the definition of '*braak*'. For instance: *'gepoogd slot voordeur te forceren'* was labelled as *'geen braak*', while '*gepoogd keukenraam voorzijde woning open te breken*' did get the '*braak*'-label.
2. One was not sure about all definitions used by the police. For example, the police document included variables labelled *braak (0 = geen braak/1 = braak), omschrijving delict (1 = gekwal. diefstal uit woning/2 = diefstal uit woning, geen braak)* and *voltooid (0 = poging/1 = voltooid)*. The municipality was not sure about the difference between *omschrijving delict* and *braak*.
   For both number 1 and number 2, it is recommended to agree (also with the police) on how to report data to make it more clear and exchangeable within parties.
3. Some variables were in the police data, but not in the data of the municipality. Among other things, the police had information about modus operandi, way of entrance and time of burglary while the municipality did not. On the other hand, the municipality had comments made by the local police officer, which were not in the police dataset.
4. Variables were not always registered the same. For example, the municipality had a variable with all week numbers and a variable called *'dag van de week',* which consisted of a number between 1 and 7 (where 1 is Monday and 7 is Sunday). The police dataset had the complete date, which was easier to work with.
5. In line with number 4, some variables like '*dag of the week*' are difficult to use for data analysis. Another example is 'modus operandi'. It reported both the *way* of entrance as the *side* of entrance, which are actually two different factors. One could reconsider the way of reporting these kinds of variables.
6. It would also be useful to consider to leave uncertain facts out of the reports. To illustrate this, we found that 'part of day' sometimes is the median or mean of the timeframe reported. Of course, it is often hard to tell when a burglary took place exactly, because people are not always at home. But now we did not know when the part of day was clear or not. This caused that we had to leave a lot of data out of the analysis, since we were not sure whether the reported part of day was reliable or not. It might be that this led to fewer significant results.

Lastly, as described at the beginning of this report, the boundaries according to Gouda are different to the ones of the CBS. The municipality already contacted the CBS about this, but it would be good to keep monitoring whether the information about Gouda according to other sources is correct.

6.2.3 Additional Options For Further Research

During the project, we have found different other potential predictors of burglary based on theories. The following predictors were not taken into account in this project due data access issues or as they did not fit exactly within the scope of the project. First, we would like to recommend for a data-driven approach of surveillance of the police and cameras to measure whether the presence of surveillance influences burglary. Secondly, it is recommended to take other physical aspects of the environment such as the presence of other types of buildings (schools, stores, offices etcetera); the heights of buildings; household types and escape routes.

6.3 Final Remark

As can been seen, there is still is room for further research and improvements. We were very glad to see the amount of enthusiasm and curiosity within the municipality. Not only about significant results, but also about the whole process. We are very excited that our presence already led to some actions. Because of this, and all the data that is still available, we are convinced that the amount of burglaries will decrease even further in the near future.

REFERENCES

Akkerman, M. (2015). *Effect van het WhatsApp-project in Tilburg op het aantal woninginbraken – een evaluatie*. Master thesis: Tilburg University.

Bernasco, W. (2017). Inbraak bij de buren? Wees op je hoede! Retrieved from: https://www.nemokennislink.nl/publicaties/

Bernasco, W., & Luykx, F. (2003). Effects of attractiveness, opportunity and accessibility to burglars on residential burglary rates of urban neighbourhoods. *Criminology, 41*(3), 981-1002

Bjørgo, T. (2016). Preventing Crime: A Holistic Approach. *Crime Prevention and Security Management*, Hampshire: England.

Boyd, D. & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society, 15*(15), pp. 662 – 679.

Brantingham, P.L., & Brantingham, P.J. (2008). Crime pattern theory. In R. Wortley and L. Mazerolle (Eds.) *Environmental criminology and crime analysis* (pp. 78-94). New York: Willan Publishing.

Bridges, P.C. (2005). *Eye on crime: The applicability of ANPR and Closed Circuit Television within Texas Law Enforcement Agencies*. Arlington: Arlington Police Department.

Brunsdon, C., Corcoran, J., Higgs, G., & Ware, A. (2009). The influence of weather on local geographical patterns of police calls for service. *Environment and Planning B: Planning and Design, 36*, 906-926.

Chan, J., & Moses, L.B. (2015). Is Big Data challenging criminology? *Theoretical Criminology, 20*(1), pp. 21 –39.

Chan, Y.L. (2015). Correlation between crime and vegetation: A case study of Jefferson County, Kentucky. *College of Arts & Sciences Senior Honors Theses*, paper 67.

Coupe, T., & Blake, L. (2006). Daylight and darkness targeting strategies and the risks of being seen at residential burglaries. *Criminology, 44*(2), 431-464.

Donovan, G.H., & Prestemon, J.P. (2012). The effects of trees on crime in Portland, Oregon. *Environment and Behavior, 44*(1), 3-30.

Erete, S.L. (2013). Published in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. *Protecting the home: exploring the roles of technology and citizen activism from a burglar's perspective*. (2507-2516). Paris, France.

Falk, G.J. (1953). The influence of the seasons on the crime rate. *Journal of Criminal Law, Criminology and Police Science, 43*, 199-213.

Farrington, D.P., & Welsh, B.C. (2002). *Effects of improved street lighting on crime: a systematic review*. London: Communication Development Unit.

Felson, M., & Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting, 19*(4), 595-601.

Fisher, B.S., & Nasar, J.L. (1992). Fear of crime in relation to three exterior site features: prospect, refuge and escape. *Environment and Behavior, 133*(24), 35-65.

Gerling, M. (2016 May 21). Misdaadmeter: Goudse regio in de top door inbraken. AD. Retrieved from: https://www.ad.nl/gouda/ad-misdaadmeter-goudse-regio-in-de-top-door- inbraken~a7eb8473/

Gleaser, E., Sacerdote, B., & Scheinkman, J. (1996). Crime and social interactions. *Quarterly Journal of Economics, 111*(2), 507-548.

Gouda (2017; 20 december). Daling in aantal woninginbraken. Retrieved from: https://www.gouda.nl/Bestuur_en_organisatie/Publicaties_en_besluiten/Nieuwsberichten/Daling_in_aantal_woninginbraken

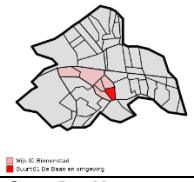
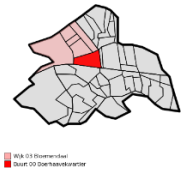
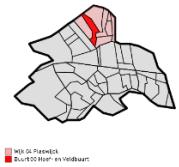
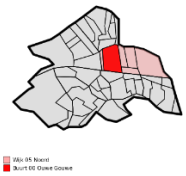
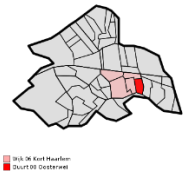
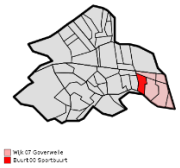
Gouds Dagblad ((017, June 3). Aantal inbraken in Gouda zorgwekkend. Gouds Dagblad. Retrieved from: https://goudsdagblad.nl/aantal-inbraken-gouda-zorgwekkend/

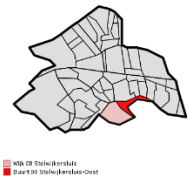
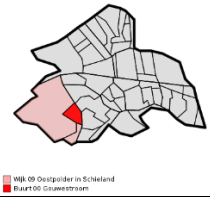
Hearden, I., & Magill, C. (2004). Decision-making by house burglars: offender's' perspective. *Home Office Research Findings* No. 249. London: Home Office.

Herbert, D., & Davidson, N. (1995). Modifying the built environment: the impact of improved street lighting. *Geoforum, 25*(3), 339, 350.
Hirschfield, A., Birkin, M., Brunsdon, C., Malleson, N., & Newton, A., (2014). How places influence crime: The impact of surrounding areas on neighbourhood burglary rates in a British city. *Urban studies, 51*(5), 1057-1072.
Hirschfield, A., & Bowers, K. J. (1997). The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies, 34*(8), 1275-1295.
Homel, R., Macintyre, S., & Wortley, R. (Eds.) (2013). *Cognition and crime: Offender decision-making and script analyses (Crime Science Series, pp. 26-47)*. London: Routledge.
Hooghe, M. (2007). Social capital and diversity generalized trust, social cohesion and regimes of diversity. *Canadian Journal of Political Science/Revue canadienne de science politique, 40*(3), 709-732.
Horrocks, J., & Menclova, A.K. (2011). The effects of weather on crime. *New Zealand Economic Papers, 45(*3), 231-254.
Kawachi, I., Kennedy, B. P., & Wilkinson, R. G. (1999). Crime: social disorganization and relative deprivation. *Social science & medicine, 48*(6), 719-731.
Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *Geo Journal, 79*, 1–14.
Koper, C.S. (1995). Just enough police presence: reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly, 12*(4), 647-673.
López, M., & Van Nes, A. (2007). *Space and Crime in Dutch Built Environments: macro and micro scale spatial conditions for residential burglaries and thefts from cars.* In: Proceedings of the 6th Space Syntax Symposium (6SSS), Istanbul, Turkey.
Martin, D. (2002). Spatial patterns in residential burglary: Assessing the effect of neighborhood social capital. *Journal of Contemporary Criminal Justice, 18*(2), 132-146.
McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology, 27*(1), 415-444.
Menting. B., Lammers, M., Ruiter, S., Bernasco, W. (2016). Family matters: effects of family members' residential areas on crime location choice. *Criminology, 54*(3), 413-433.
Minnaar, A. (2007). The implementation and impact of crime prevention/crime control open street Closed Circuit Television surveillance in South African Central Business Districts. *Surveillance & Society, 4*(3), 174-207.
Moreto, W.D. (2010). Risk Factors of Urban Residential Burglary. *RTM Insights, 4*. Available at: www.riskterrainmodeling.com
Nasar, J.L. (1981). Environmental factors and commercial burglary. J. *Environmental Systems, 11*(1), 49-56.
Nicolson, P. (1994). *The experience of being burgled: A psychological study of the impact of domestic burglary on victims*. Sheffield, UK: Sheffield University
Montoya, L., Junger, M., & Ongena, Y. (2016). The relation between residential property and its surroundings and day-and night-time residential burglary. *Environment and behavior, 48*(4), 515-549.
Politiekeurmerk.nl. (2016). Subsidie maakt woningen inbraak proof. Acquired on 10-11-2017 through https://www.politiekeurmerk.nl/bewoners/nieuws/detail/article/subsidie-maakt-woningen-inbraakproof/
POPCentre (n.d.). The Problem Analysis Triangle [Website]. Obtained from: http://www.popcenter.org/about/?p=triangle
Ranson, M. (2012). *Crime, weather, and climate change*. (M-RCBG Associate Working Paper Series No 8.) Mossavar-Rahmani Center: Harvard Kennedy School.
Rengert, G.F., E. Groff (2011). Residential Burglary: How the Urban Environment and Our Lifestyles Play a Contributing Role. 3rd Edition, ISBN: 978-039806794
Riccio, L.J. (1974). Direct deterrence- an analysis of effectiveness of police patrol and other crime prevention technologies. *Journal of Criminal Justice, 2*(3), 207-217.
Richards, N.M. & King, J.H. (2014). Big Data Ethics. *Wake Forest Law Review, 49*, 389–432.

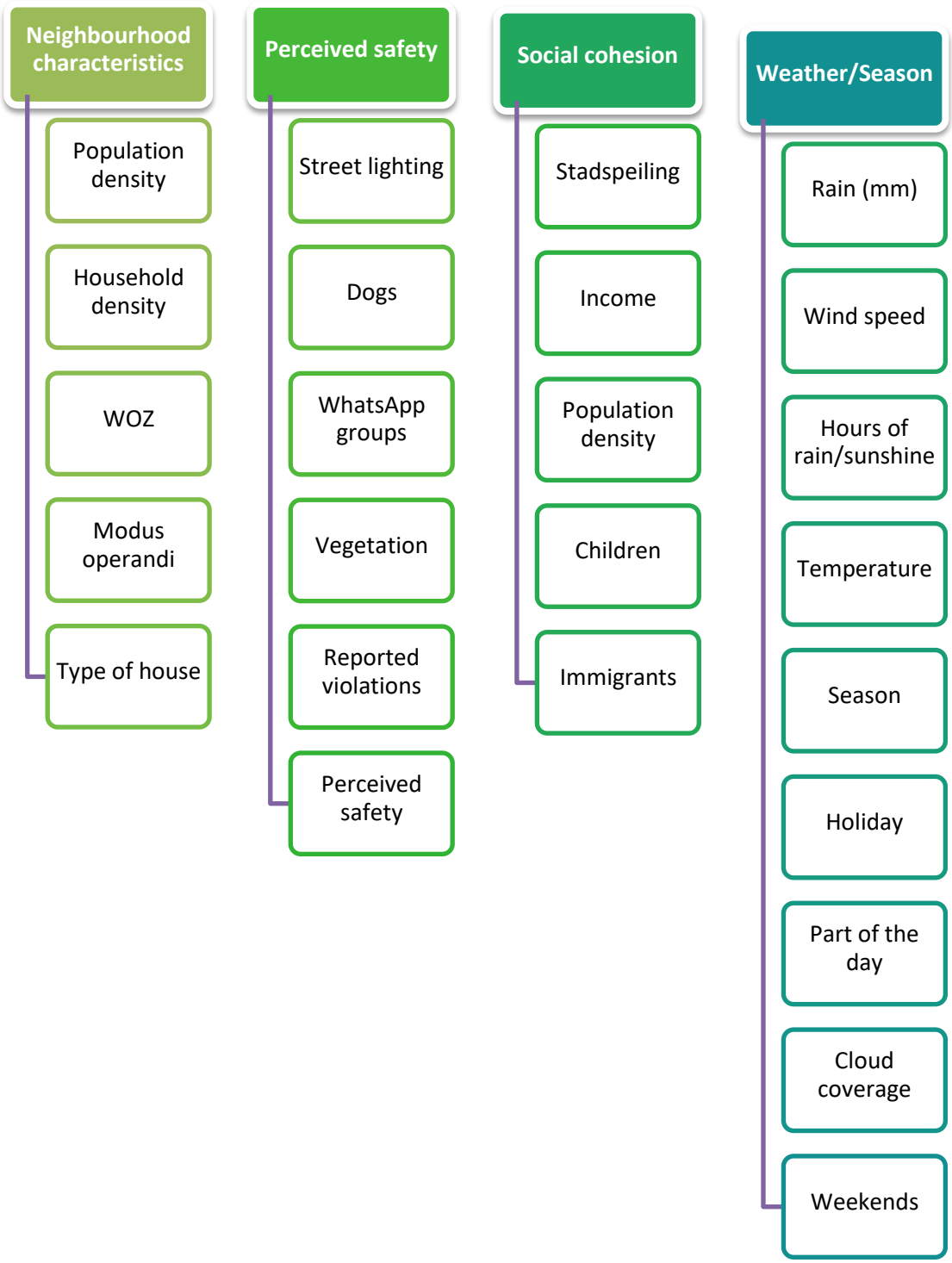
Schmallenger, F., (1997). *Criminal Justice Today*. Englewood Cliffs: New Jersey.
Sherman, L.W. (1995). Hot Spots of Crime and Criminal Careers of Places. In: Eck, J.E., D. Weisburd (Eds). Crime and Place. *Crime Prevention Studies, 1*, 35-52.
Sherman, L.W., & Weisburd, D. (1995). General deterrent effects of police patrol in crime 'hot spots': a randomized controlled trial. *Justice Quarterly, (12)*4, 625-649.
Taylor, L. & Richter, C. (2015). Big Data and Urban Governance. In: Gupta, J., Pfeffer, K., Verrest, H. & Ros-Tonen, M. (Eds). *Geographies of Urban Governance*. Springer: Cham.
Troy, A.J., Grove, M., & O'Neil-Dunne, J. (2012). The relation between tree canopy and crime rates across an urban-rural gradient in greater Baltimore region. *Landscape and Urban Planning, 106*(3), 262-270.
Van Nes, A. (2006). The burglar as a space explorer in his own neighbourhood. *Transactions on Ecology and the Environment, 93*, 671-680.
Van Nes, A., & López, M.J.J. (2011). Macro and micro scale spatial variables and the distribution of residential burglaries and theft from cars. *The Journal of Space Syntax, 1*(2), 296-314.
Weisel, D., (2002). *Burglary of single-family houses*. Washington, DC: UD Department of Justice, Office of Community Oriented Policing Services.
Welsh, B.C., & Farrington, D.P. (2008). *Effects of street lighting on crime*. Lowell: The Campbell Collaboration.
Wolfe, M.K., & Mennis, J. (2012). Does vegetation encourage or suppress urban crime? Evidence from Philadelphia, PA. *Landscape and Urban Planning, 108*, 112-122.

APPENDIX I: DISTRICTS AND NEIGHBOURHOODS IN GOUDA

Districts	Neighbourhoods (Gouda)	Neighbourhoods (CBS)	
1. Binnenstad 	10. Nieuwe Markt e.o. 11. De Baan e.o. 12. Turfmarkt e.o. 13. Raam e.o. 14. Nieuwe Park Oost 15. Nieuwe Park West	BU05130100 BU05130101 BU05130102 BU05130103 BU05130104 BU05130105	Nieuwe Markt e.o. De Baan e.o. Turfmarkt e.o. Raam en omgeving Nieuwe Park-Oost Nieuwe Park-West
2. De Korte Akkeren 	20. De Korte Akkeren Oud 21. De Korte Akkeren Nieuw 22. Industrieterrein Kromme Gouwe 23. Industrieterrein Hollandsche IJssel 24. Weidebloemkwartier	BU05130200 BU05130201 BU05130202 BU05130203 BU05130204	De Korte Akkeren-Oud De Korte Akkeren-Nieuw Industrieterrein Kromme Gouwe Oost West Industrieterrein de Hollandsche IJssel Weidebloemkwartier
3. Bloemendaal 	30. Boerhaavekwartier 31. Windroos en Heesterbuurt 32. Groenhovenkwartier 33. Bloemendaalseweg 34. De Goudse Poort 35. Gaardenbuurt	BU05130300 BU05130301 BU05130302 BU05130303 BU05130304 BU05130305	Boerhaavekwartier Windrooskwartier en Heesterbuurt Groenhovenkwartier Bloemendaalseweg De Goudse Poort De Gaardenbuurt
4. Plaswijck 	40. Hoef- en Veldbuurt 41. Zomenbuurt 42. Hoevenbuurt 43. Lusten-, Burgen- en Steinenbuurt 44. Grassen- Waterbuurt 45. Bodegraafsestraatweg 46. Mammoet 47. Wervenbuurt	BU05130401 BU05130402 BU05130403 BU05130404 BU05130405 BU05130406 BU05130407	Zomenbuurt Hoevenbuurt Lusten-, Burgen- en Steinenbuurt Grassen- en Waterbuurt Bodegraafsestraatweg Mammoet Wervenbuurt
5. Noord 	50. Ouwe Gouwe 51. Statensingel 52. Wethouder Venteweg 53. Achterwillenseweg 54. Slagenbuurt 55. Middenwillens 59. De Goudse Hout	BU05130500 BU05130501 BU05130502 BU05130503 BU05130504 BU05130509 BU05130509	Ouwe Gouwe Statensingel Wethouder Venteweg Achterwillenseweg Slagenbuurt De Goudse Hout De Goudse Hout
6. Kort Haarlem 	60. Oosterwei 61. Vreewijk 62. Voorwillenseweg 63. Kadenbuurt 64. Kort Haarlem	BU05130600 BU05130601 BU05130602 BU05130603 BU05130604	Oosterwei Vreewijk Voorwillenseweg Kadenbuurt Kort Haarlem
7. Goverwelle 	70. Sportbuurt 71. Molenbuurt 72. Polderbuurt 73. Vrijheidsbuurt 74. Componistenbuurt 75. Muziekbuurt	BU05130700 BU05130701 BU05130702 BU05130703 BU05130704 BU05130705	Sportbuurt Molenbuurt Polderbuurt Vrijheidsbuurt Componistenbuurt Muziekbuurt

8. Stolwijkersluis  	80. Stolwijkersluis Oost 89. Stolwijkersluis West	BU05130800 BU05130809	Stolwijkersluis-Oost Stolwijkersluis-West
9. Westergouwe  	90. Gouwestroom 91. Ringvaartbocht 92. Oostpolder in Schieland 93. De Eilanden 94. Tuinenbuurt 99. Westergouwe	BU05130900 BU05130901 BU05130909 BU05130909 BU05130909 BU05130909	Gouwestroom Ringvaartbocht Oostpolder in Schieland Oostpolder in Schieland Oostpolder in Schieland Oostpolder in Schieland

Appendix II: Possible Environmental Predictors of Residential Burglary



APPENDIX III: SURVEY QUESTIONS

1. Ik ben beheerder van de buurtpreventie WhatsApp groep in de buurt … *[invulvraag]*

2. Mijn groep bestaat uit … leden *[invulvraag]*

3. De WhatsApp groep is actief sinds … *[invulvraag]*

4. In de afgelopen twee maanden zijn … berichten geplaatst (schatting) *[invulvraag]*

6. Hoe actief is de WhatsApp groep gemiddeld? Hoe vaak worden er berichten verstuurd? *[multiple choice]*

a. Dagelijks

b. Wekelijks

c. Maandelijks

d. Nog minder

7. Hoe vaak wordt er in de WhatsApp een melding gemaakt over een woninginbraak of een verdachte situatie *[multiple choice]*

a. *Dagelijks*

b. Wekelijks

c. Maandelijks

d. Nog minder