safe.brussels

Observatory
Paper no. 1

Analysis of recorded crime by district in the Brussels Region

A multivariate and typological approach

MARCH 2026

Introduction

The safe.brussels Observatory plays a key role in the study and understanding of recorded crime in the Brussels-Capital Region. As the regional body responsible for collecting, integrating, and analysing security-related data, it ensures the production of reliable, comparable, and contextualised indicators. Through its methodological expertise and its capacity to cross-reference different sources of information, it provides public decision-makers, field practitioners and the general public with a structured and evidence-based understanding of developments in security in Brussels. This specific analysis forms part of that broader logic of clarification, decision support and transparency.

Recorded crime in the Brussels-Capital Region refers to all criminal offences officially recorded by the police services on the basis of an initial police report. It constitutes an essential indicator for understanding developments in public security, but represents only part of actual crime. Indeed, its volume depends not only on criminal activity itself, but also on institutional and social factors such as police proactivity (intensity of controls, operational priorities and targeted actions) and the willingness of victims or witnesses to report offences. These factors directly influence the number of recorded incidents, independently of the objective evolution of criminal phenomena. Analyses of recorded crime must therefore be interpreted with caution, taking into account of these institutional biases and dynamics.

This analysis aims to explore the links between crime recorded by the police and economic, social, and demographic indicators at territorial level. It seeks to identify territorial patterns of crime recorded in the Brussels-Capital Region, in order to better understand the spatial distribution of phenomena and their local specificities. The approach is based on the use of multivariate statistical methods, in particular Principal Component Analysis (PCA), the cross-referencing of indicators and the construction of a typology, with a view to identifying groups of districts according to their crime profiles. This approach makes it possible to move beyond a purely descriptive reading of the figures and to propose a structured and contextualised interpretation of the spatial dynamics of recorded crime. It ultimately results in a typological approach to the districts based on the variables studied.

1. Residential districts and variables used

The observation units used in the following analysis are the 118 residential districts of the Brussels-Capital Region (Figure 1), as defined by the Monitoring of the Districts1. These districts have an average population of around 10,000 inhabitants, with values ranging from approximately 3,000 to 25,000 inhabitants.

This analysis initially relied on twenty-six quantitative variables, including seven recorded crime variables. The variables were selected on the basis of their availability and the recency of their updates —between 2021 and 2023 for contextual variables, and 2023 for recorded crime variables— as well as their relevance, assessed on the basis of previous analyses of their observed relationships with crime rates.

By grouping highly correlated variables and removing those with very limited explanatory weight at the multivariate level, the analysis was reduced to fourteen quantitative variables: four related to recorded crime, nine socioeconomic and demographic indicators, and one variable relating to SIAMU fire service interventions (Table 1).

FIGURE 1: **The 118 residential districts in the Brussels-Capital Region**

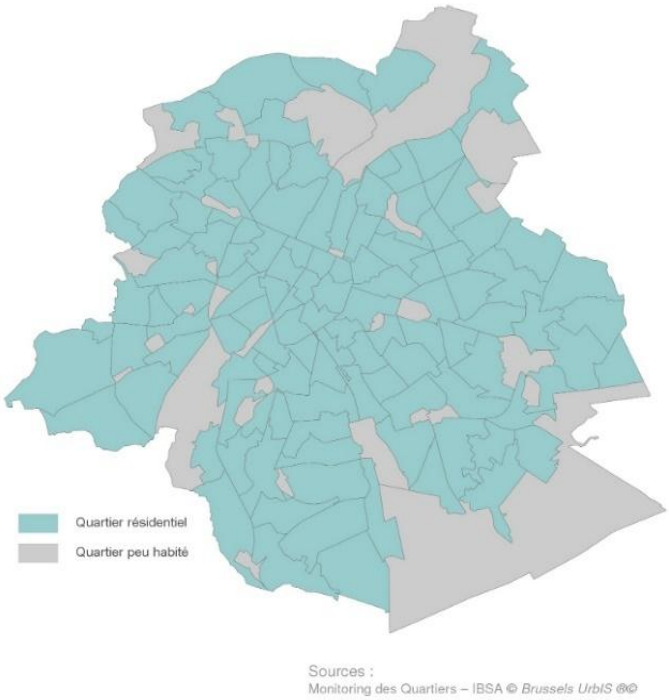


TABLE 1: **The 14 quantitative variables selected for analysis**

4 recorded crime variables and 1 variable for SIAMU interventions	**9 socioeconomic and demographic variables**
• Burglary • Bicycle theft • Drugs (sale and possession) • Violence in public areas (including physical and verbal "threats", "weapons and ammunition", including illegal possession or trafficking, and "sexual violence") • SIAMU intervention (fires)	• Population density • Proportion of households with children • Proportion of single persons under 30 + proportion of ground-floor businesses • Proportion of single persons over 65 + level of greenery • Proportion of males in the population • Proportion of people from OECD countries excluding Turkey and South America • Proportion of people from North Africa and Turkey • Jobseekers • Median income

1. Visit the website of the Region at https://monitoringdesquartiers.brussels/partition-region-de-bruxelles-capitale-quartiers. See also (French and Dutch): Observatory - Image Unit, *Guide méthodologique de l'Observatoire. Analyse et image des phénomènes*, Second Edition, January 2023, pp. 30-32. https://safe.brussels/fr/guide-methodologique-de-lobservatoire-analyse-et-image-des-phenomenes-deuxieme-edition.

We also include three categorical variables. One nominal categorical variable corresponds to the six police zones while two categorical variables are derived from continuous variables: gender and population density. Gender is based on the proportion of men in the population and is defined using four categories (M+, M, F, F+), while population density is defined using three categories (Low, Mid, High). These two latter variables are subject to a bias related to threshold effects resulting from the definition of category boundaries, but they offer a certain level of interpretability in the analysis that follows.

The aim is to identify the correlations between socioeconomic, demographic and urban variables, both among themselves and in relation to recorded crime variables, in order to identify structuring patterns for understanding the territory. To this end, we will first establish the correlation matrix between all variables in the dataset (Section 2). We then conduct a Principal Component Analysis and interpret the results (Section 3), before concluding with a classification of districts into clusters (typology), which is mapped to visualise it within the territory of the Brussels-Capital Region (Section 4).

2. Multivariate approach: the correlation matrix

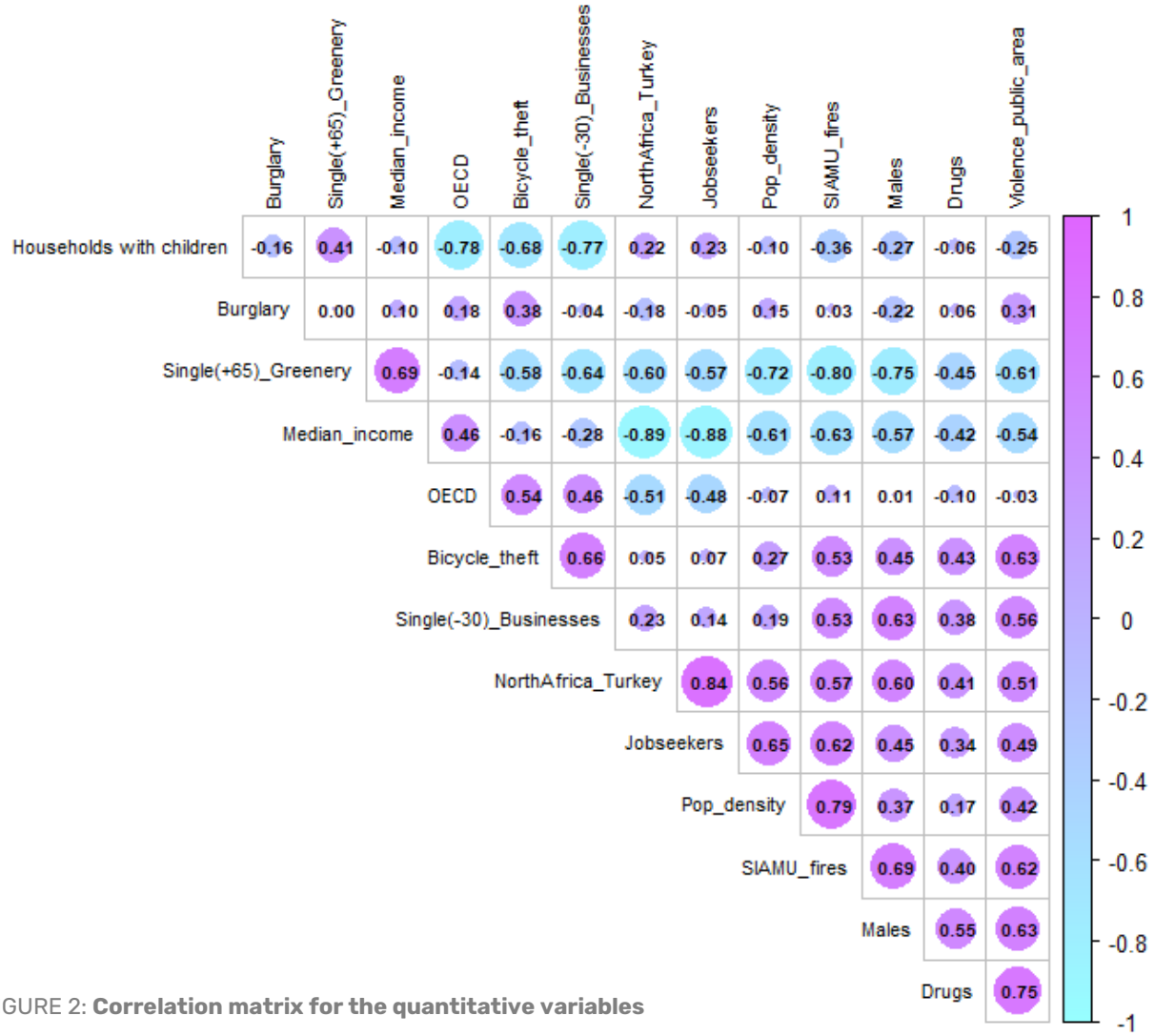


FIGURE 2: **Correlation matrix for the quantitative variables**

The correlation matrix (Figure 2) examines the relationship of each variable and all the others, using Pearson's coefficient2 to measure the strength and direction of the linear relationship between two variables, on a scale ranging from -1 to 1. A coefficient close to 1 indicates a strong positive linear correlation between the two variables: if one increases, so does the other. A coefficient close to 0 indicates that no linear relationship has been identified. Finally, a coefficient close to -1 represents a strong negative linear relationship: if one variable increases, the other decreases.

With fourteen quantitative variables, the correlation matrix shows ninety-one pairwise relationships. By focusing, for illustrative purposes, to a selection of significant correlations, we observe a strong negative correlation between median income and the proportion of jobseekers (-0.88). Median income is positively correlated with the proportion of single persons over 65 and the level of greenery (0.69), as well as with the proportion people originating from OECD countries (0.46), and negatively correlated with violence in public areas, bicycle theft and drug sales and possession.

The proportion of jobseekers is positively associated with the proportion of people from North Africa and Turkey (0.84), population density (0.65), the number of SIAMU fire service interventions (0.62), masculinity (0.45), and recorded crime variables excluding burglary.

We also observe links between SIAMU fire service interventions and population density (0.79), violence in public areas (0.62), bicycle theft (0.53), and the proportion of ground-floor commercial premises and single persons under 30 (0.53). Violence in public areas is strongly correlated with the drug variable (0.75). We reiterate that correlation is not causation, and that two variables with high values for the same district do not necessarily mean that one influences or explains the other.

All these relationships can be visualised as a whole using the correlation circle derived from a Principal Component Analysis (PCA).

3. Dimensionality reduction: principal component analysis (PCA)

Principal Component Analysis (PCA) is a multivariate statistical method used to summarise the information contained in a large number of quantitative variables by transforming them into a reduced number of new variables known as principal components. These components are linear combinations of the standardised initial variables (mean 0 and variance 1), chosen to capture as much of the variability in the data as possible. The first component, or first dimension, explains the largest share of the variance, expressed as a percentage. The second component is independent and orthogonal to the first and explains a smaller share of the variance, and so on. Each component is therefore associated with a dimension and a percentage of variance explained, the sum of all components being equal to 100%.

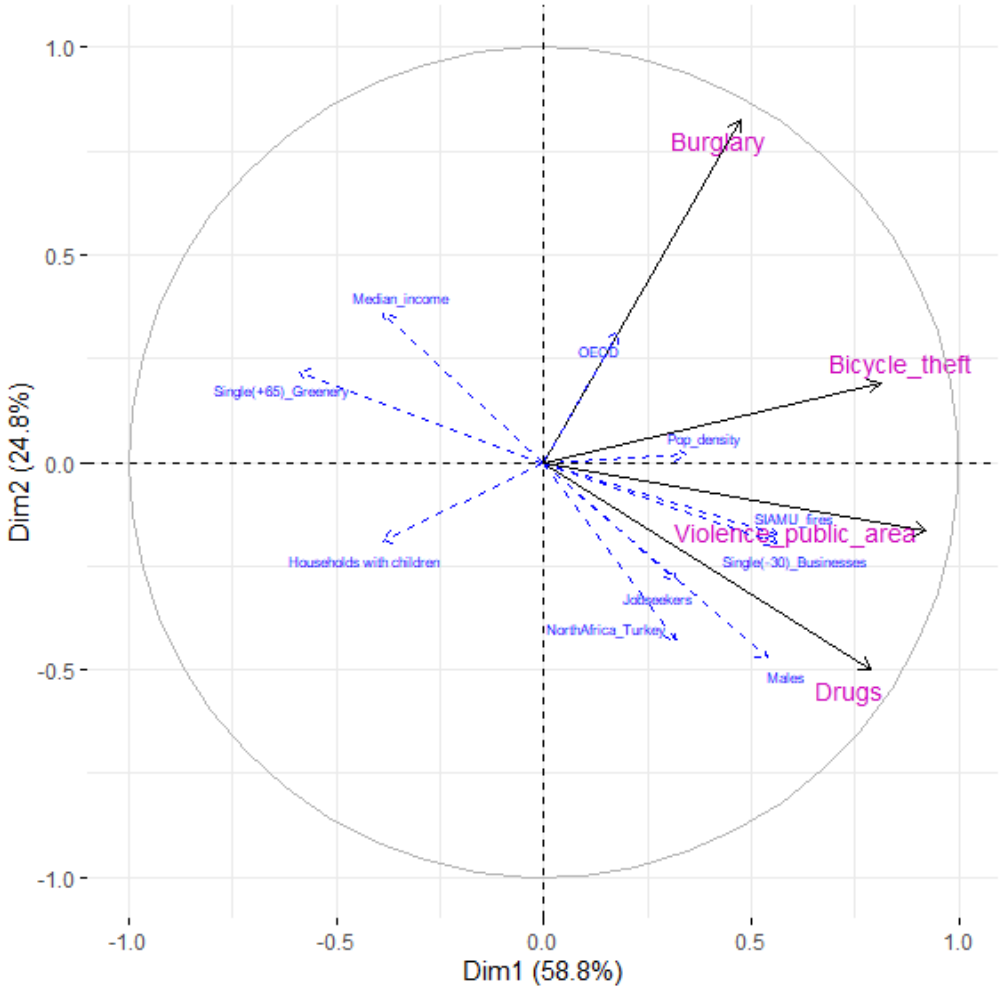
PCA is particularly useful for visualising and interpreting complex data and for identifying hidden structures (groups of individuals or relationships between variables). It makes it possible to simplify the data without losing too much information, thereby facilitating the identification of typical profiles and decision-making. In the context of Observatory's crime analysis conducted by the Image Unit, this applied research tool is used here for an exploratory analysis.

2. The Pearson correlation coefficient is the sum of the products of the deviations from the mean for the two variables under consideration (covariance), divided by the product of their standard deviations (standardisation).

An initial PCA was carried out on the four variables related to recorded crime (Figure 3). This analysis explores the structure of crime. The other variables (dotted arrows) are projected a posteriori onto the axes.

On the first dimension, which explains 58.8% of the variability in recorded crime, all crime variables are positive and point to the right. We can then observe an overall contrast between, on the one hand, low-crime, high-income districts, with a higher level of greenery and a relatively higher proportion of single persons over 65 and households with children, and, on the other hand, high-density districts, with a relatively higher proportion of single persons under 30, a higher proportion of men than women, SIAMU fire service interventions, ground-floor commercial premises and jobseekers, and where the selected crime incidents are concentrated.

FIGURE 3: **PCA - Correlation circle for crime variables**



Conversely, another PCA is carried out on the socioeconomic and demographic variables. This analysis describes the structure of territories and individuals according to their characteristics.

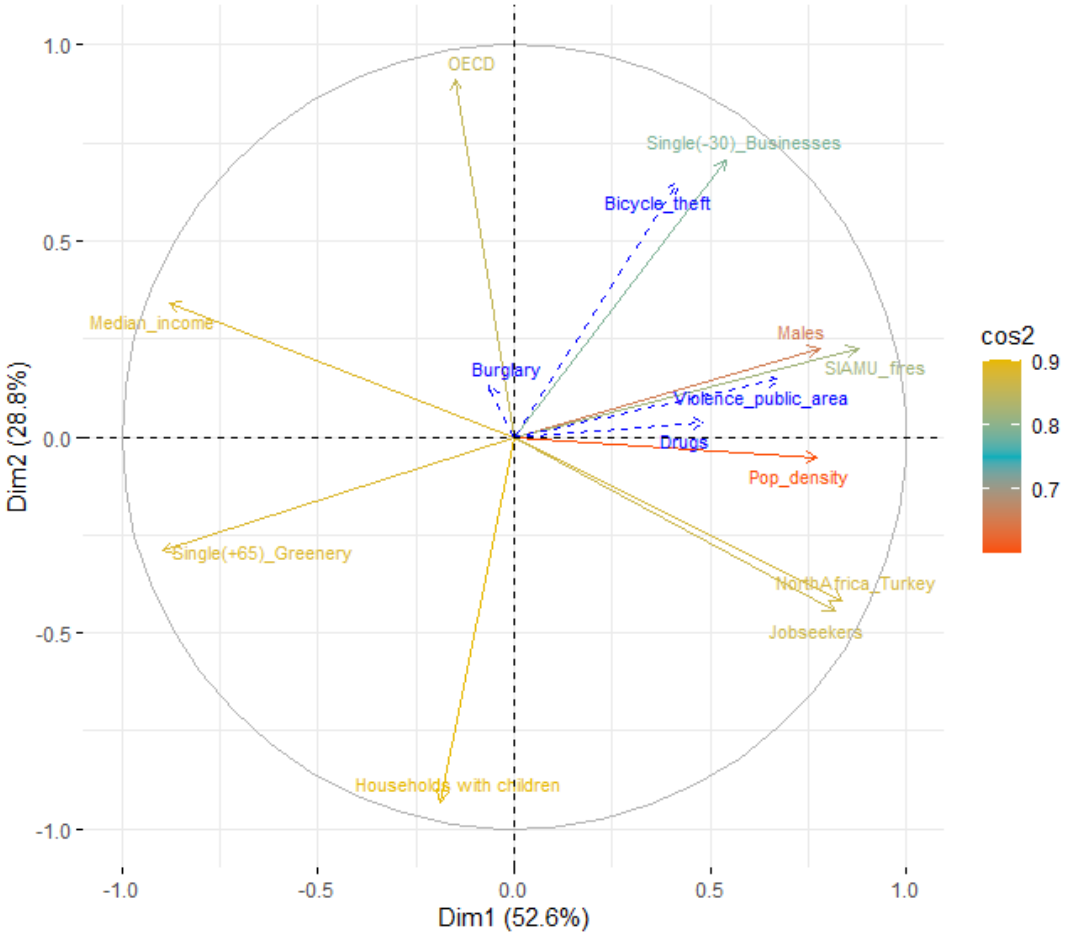
The recorded crime variables are then projected in order to identify the social profiles with which they are associated.

,

A closer examination (Figure 4) shows that the combined variance explained by the first dimension (axis 1 = 52.6%) and the second dimension (axis 2 = 28.8%) reaches 81.4%. These two dimensions represent two new variables which, on their own, explain more than 80% of the variability in the data used here. This graph shows the direction of the crime variables that were subsequently added.

The colour scale indicates the quality of representation of the variables projected onto the constructed axes. A cos² greater than 0.6, combined with an arrow close to the circle, indicates a strong correlation with the axes and therefore plays a more active contribution to their construction. This is the case for all the variables used here.

FIGURE 4: **PCA - Correlation circle for socioeconomic, demographic and urban-planning variables**

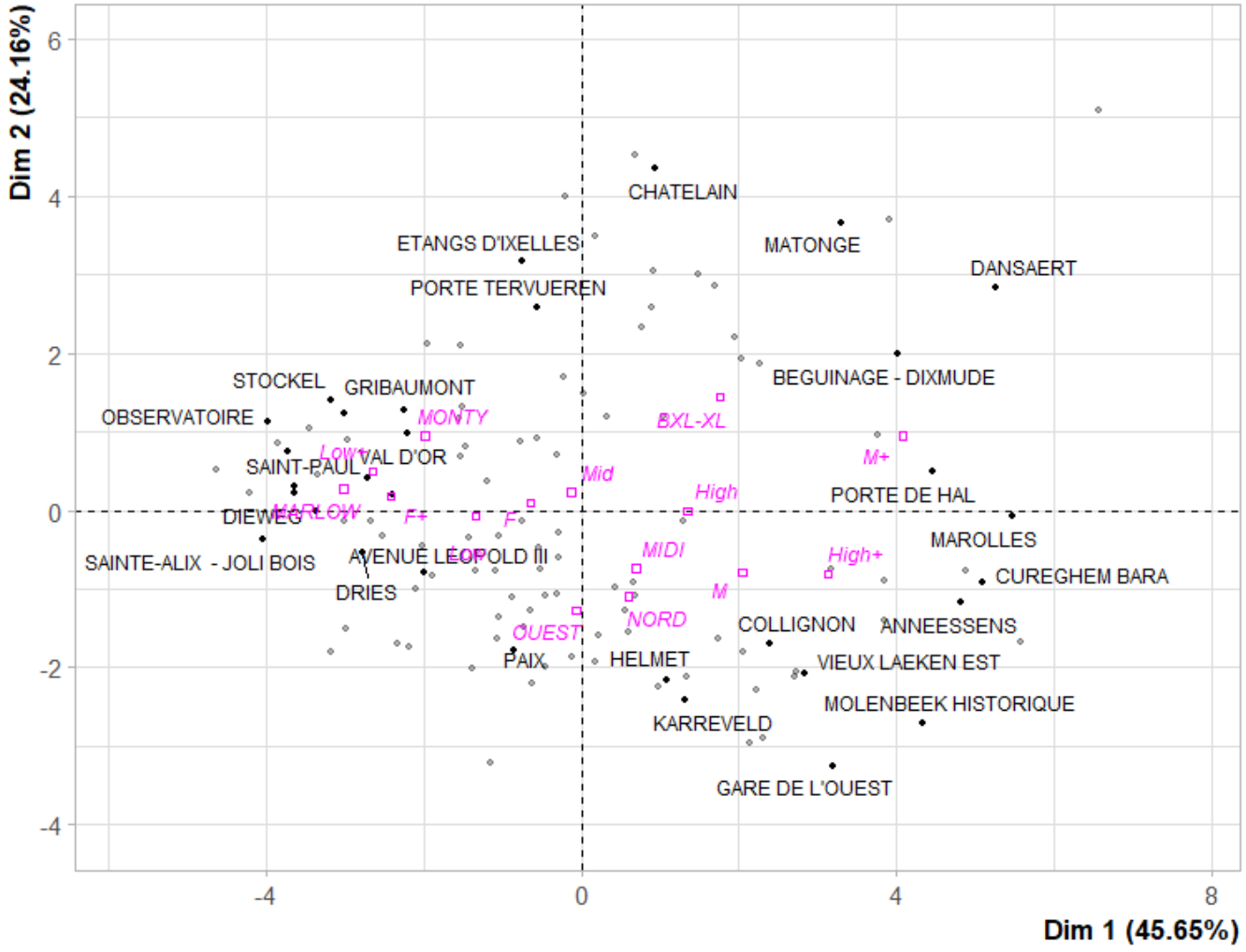


This PCA confirms the earlier observations derived from the correlation matrix: the variables cluster along the axes as anticipated by the Pearson coefficients. However, violence in public areas and the drug variable are more strongly associated with masculinity and population density in districts with the highest number of SIAMU fire service interventions. The arrow-vector for "households with children" is almost at right angles to "median income" and the "proportion of jobseekers" variable.

A right angle indicates an absence of correlation, which may suggest that households with children are as prevalent in higher-income districts as in lower-income ones. The same applies, albeit to a lesser extent, to the variable representing the proportion of people from OECD countries. By contrast, the proportion of people originating from North Africa and Turkey is very strongly correlated with low-income districts.

After visualising the variables, we now present the visualisation of districts on the PCA based on all fourteen quantitative variables (Figure 5). This graph makes it possible to anticipate the classification of districts developed later, as it shows that the districts are distributed across three point clouds: a fairly dense cluster in the centre left, another fairly dense cluster in the bottom right, and a less dense cluster in the top right.

FIGURE 5: **PCA - Graph of districts on the first two axes of the PCA for the 14 quantitative variables**

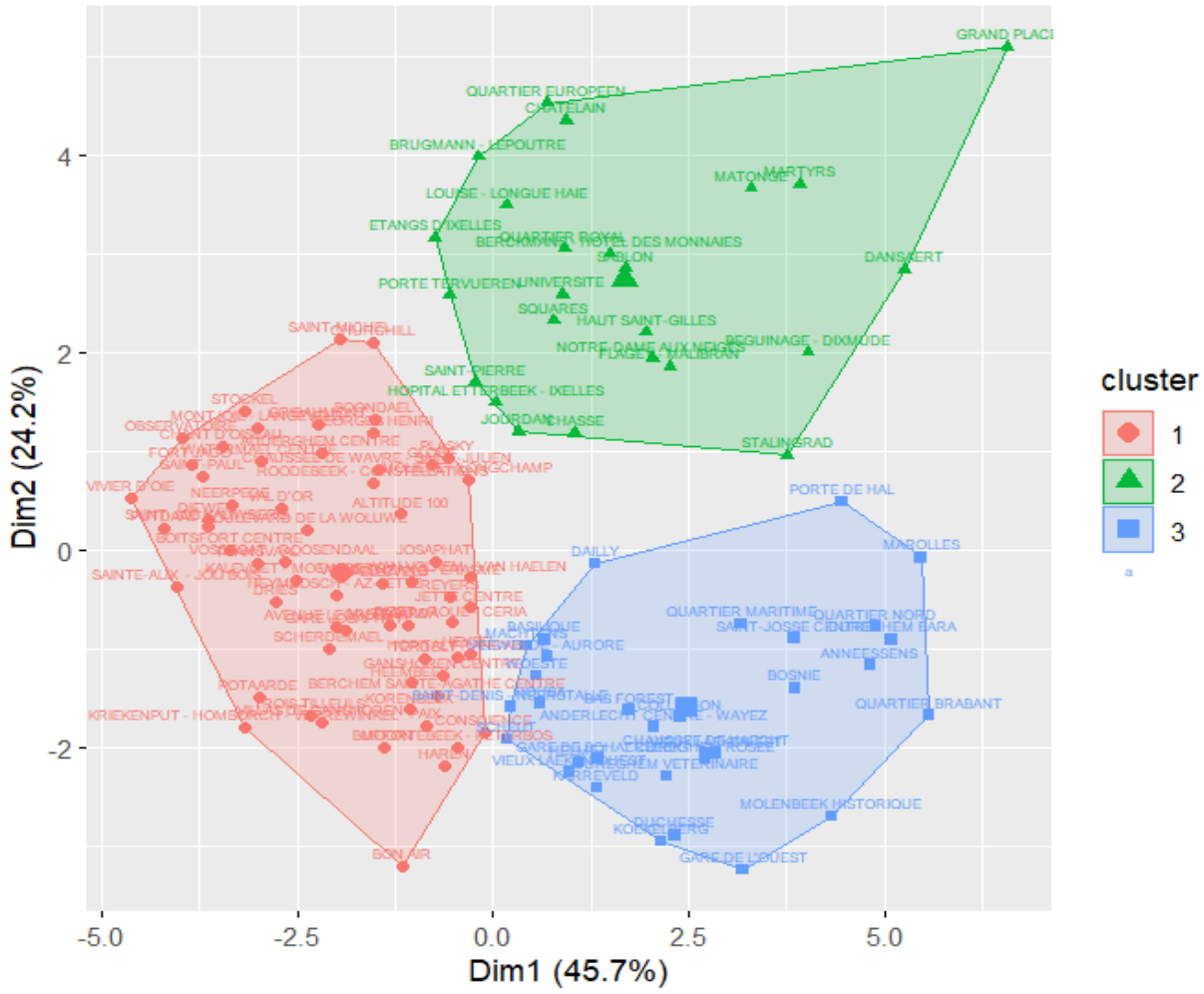


This representation also allows for the a posteriori projection of the categorical variables using colour coding (Density, Police Zone, Gender). We observe a coherent structure for these variables. Population density forms an axis running from the upper left quadrant (Low+) through the center (Mid) to the lower right quadrant (High+). Police zones form a U-shaped structure running from the upper left quadrant (MONTY and MARLOW) to the lower quadrant (OUEST, NORD and MIDI), and rising again towards the upper right quadrant (BXL-XL). Gender combines the two previous variables, following a straight line from F+ (relatively highest proportion of women in the population) to F in the centre (higher proportion of women), then passing through the lower right quadrant to M (higher proportion of men) and raising again towards M+ in the upper right quadrant (highest proportion of men).

4. Hierarchical classification: towards a typology of districts

PCA is also particularly useful for reducing data dimensionality prior to applying other methods such as clustering. To better understand the structure of the data, we applied various hierarchical clustering methods to group districts according to similarities based on the principal components constructed from the fourteen quantitative variables. Several aggregation methods were tested, which are algorithms that group districts into different classes in different ways. We selected Ward's method3, which minimises the internal variance within a class at each successive merging, yielding a fairly balanced final partition. Based on this method, the optimum number of clusters is three (Figure 6).

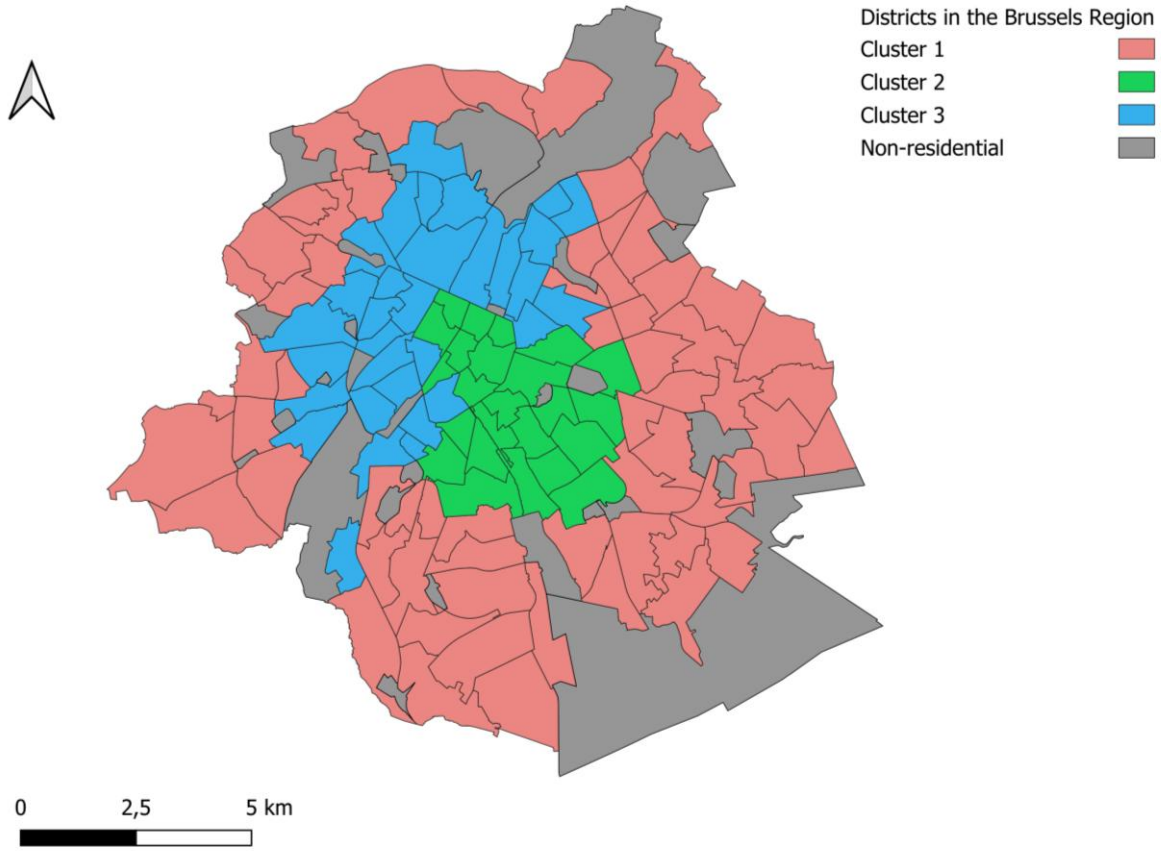
FIGURE 6: **Typology of districts into three groups**



3. Regarding hierarchical classification, see https://www.shadecoder.com/de/topics/ward-s-method-clustering-a-comprehensive-guide-for-2025. On the 'HCPC' (Hierarchical Classification on Principal Components) function used in the FactoMiner Package with R software, see: https://cran.r-project.org/web/packages/FactoMineR/FactoMineR.pdf p.28.

The typological partition of districts assigns each district to a group. By retrieving this attribute, it can be mapped in order to visualise it concretely within the regional space. We can observe that the spatial representation of the three-cluster typology below (Figure 7) reveals a remarkably coherent territorial structuring into groups of contiguous districts.

FIGURE 7: **Mapping of districts into 3 groups**



Analysis of the mathematical characteristics of the districts derived from the typological construction allows us to make several observations. For cluster 1 (in red), we observe a dominance of districts characterised by high incomes, a high level of greenery and a high proportion of single persons over 65, as well as, to a lesser extent, a high proportion of households with children. Variables related to crime and those associated with jobseekers are much lower than elsewhere. From a spatial perspective, this cluster brings together the "peripheral" districts of the Brussels-Capital Region.

Cluster 2 (in green) contrasts with Cluster 1 on the first dimension, while it contrasts with Cluster 3 on the second dimension (Figure 6). It is characterised by a high proportion of single persons under 30, commercial premises, and a higher proportion of people originating from OECD countries (expat districts). It includes the districts with the lowest proportion of households with children and the highest proportion of bicycle thefts. Masculinity is positive, as are violence in public areas and SIAMU fire service interventions. This cluster includes the central districts as well as those located in the south-east, but close to the city centre.

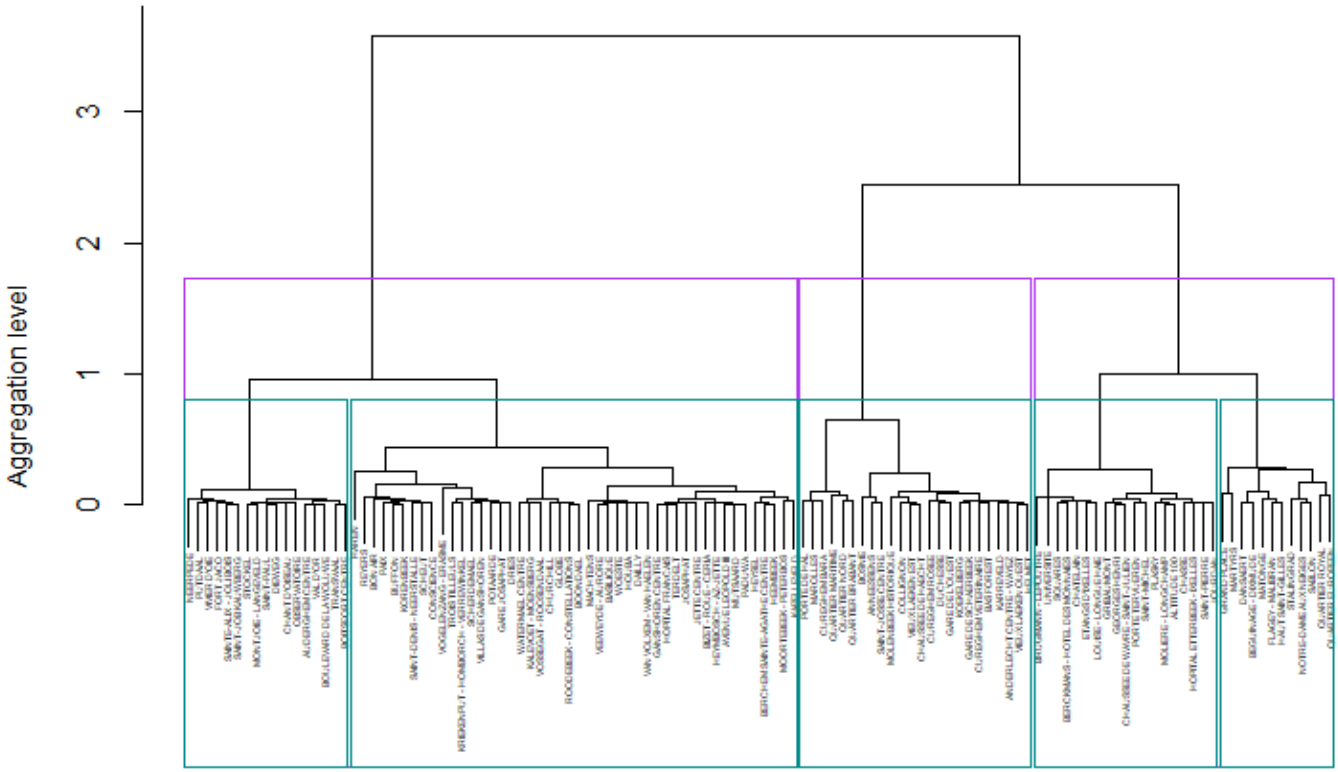
Cluster 3 (in blue) is characterised by districts with the highest rates of jobseekers and people of North African and Turkish origin, high population density and therefore a high level of built-up areas, the highest number of SIAMU fire services interventions, and a higher level of masculinity than elsewhere. By mirror effect, this also corresponds to generally lower incomes and a lower level of greenery. This cluster includes the highest values for violence in public areas, including threats, sexual violence and carrying a weapon, and, to a lesser extent, for the "drug" variable. It mainly comprises districts in the north-west of the Region but close to the city centre. Finally, these districts are characterised by a high proportion of households with children, in contrast to Cluster 2.

***

This three-cluster typology can be further refined with a five-cluster typology. The latter is based on the same data and involves splitting Cluster 1 and Cluster 3 into two groups each, resulting in five classes. The dendrogram below (Figure 8) visualises the aggregation of districts, showing the successive mergers of clusters (the "hierarchical tree").

The number of branches crossed, for a given height of the dendrogram (coloured horizontal lines) indicates the number of classes selected. The vertical axis represents the height of the tree, which corresponds to the level of aggregation. The three-class partition is obtained at a higher level, resulting in more global and less homogeneous classes, whereas a lower cut produces five more homogeneous and more detailed classes.

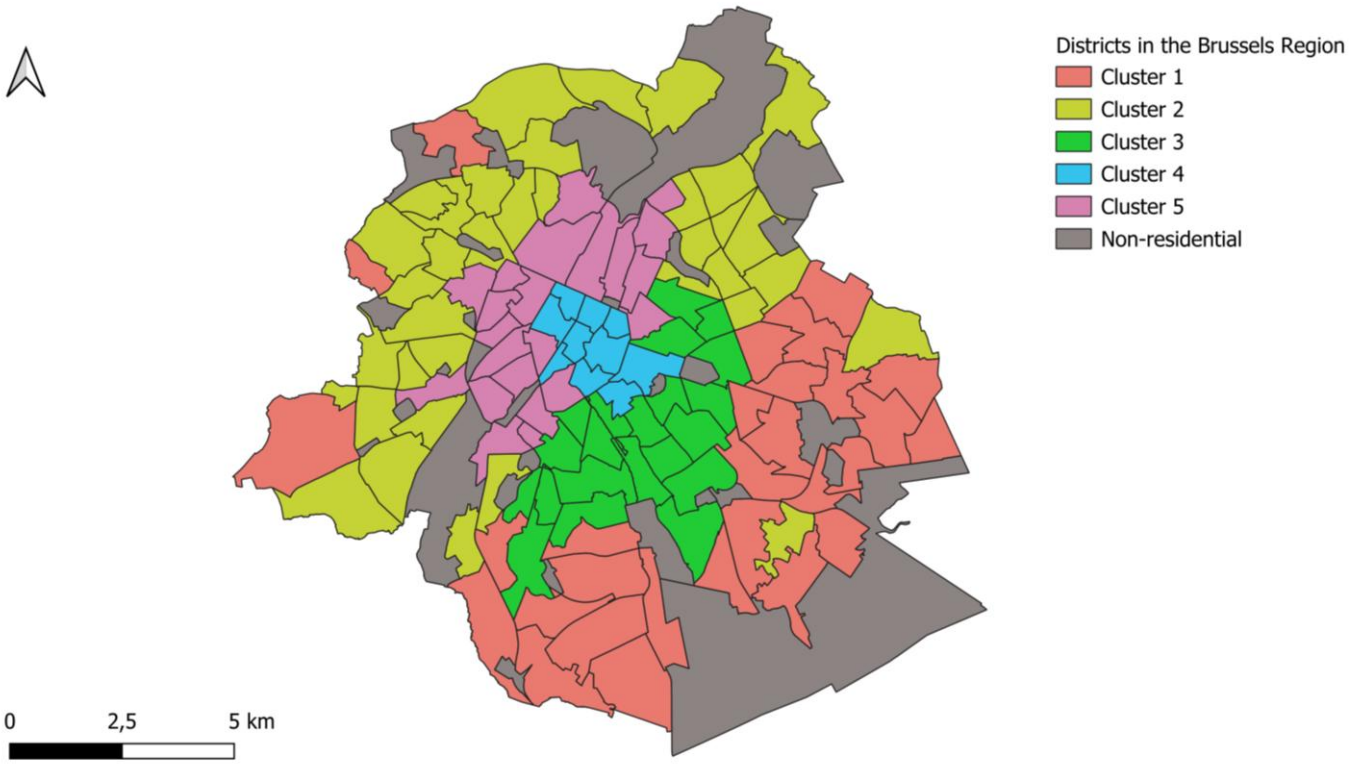
FIGURE 8: **Dendrogram - Grouping into 3 and 5 classes**



After analysing the characteristics of the districts and the variables that influence these divisions, typology can be refined to highlight secondary intra-group differences that are nonetheless clearly present in the three-cluster partition.

The spatial representation of the five-cluster typology (Figure 9) also shows a coherent territorial structuring into groups that are fairly contiguous and tightly grouped.

FIGURE 9: **Mapping of districts into 5 groups**



The peripheral districts are thus divided into two groups which nonetheless remain characterised by negative values on the first dimension, i.e. low levels of recorded crime. Moreover, they share higher incomes, higher levels of greenery and a higher proportion of single persons over 65. However, cluster 1 (in red) mainly covers the south-east of the Region and includes some of the highest values in terms of income, proportion of single persons over 65 and level of greenery, as well as an above-average proportion of households with children. By contrast, this cluster shows very low proportions of jobseekers and people originating from North Africa and Turkey, lower levels of masculinity (these districts therefore have a majority of women in the population), very low population density and a relatively low number of SIAMU fire services interventions. Cluster 2 (in khaki), by contrast, is primarily characterised by a higher proportion of households with children, much lower incomes, few bicycle thefts and a lower proportion of people originating from OECD countries.

The central districts and those located in the south-east but close to the city centre remain characterised by a high proportion of single persons under 30 and bicycle thefts, as well as a very low proportion of households with children. Within this group, a distinction can be made between cluster 3 (in green) which is more strongly characterised by people originating from OECD countries and burglaries, and cluster 4 (in blue), which is more strongly characterised by masculinity, ground-floor commercial premises, violence and drugs.

Finally, cluster 5 (in lilac) includes districts located in the north-west of the Region but close to the city centre[4]. This cluster remains close to the cluster identified in the three-group typology and is characterised by low incomes, a high proportion of jobseekers, people originating from North Africa and Turkey, higher levels of masculinity and population density. It is also strongly positive for violence recorded in public areas.

4. We can see that this is roughly the 'Espace de Développement Renforcé du Logement et de la Rénovation' (Enhanced Housing and Renovation Development Area) as defined by the Region in 1993, also sometimes referred to as the 'poverty crescent'.

This typological work on the districts reveals a new and insightful structuring of districts into three or five types. It highlights the importance of urban structure and the built environment in the characteristics of the district clusters identified here. Spatiality emerges as the latent explanatory variable, approached here through the contrast between population and building density on the one hand and the level of greenery on the other, the former being strongly and positively associated with recorded crime. This contrast operates in parallel with another structuring axis, namely the proportion of jobseekers and median income. A third correlated axis contrasts, albeit in a less pronounced manner than the previous two, the proportion of people from North Africa and Turkey with those from OECD countries. With regard to crime, an examination of the other factorial planes shows that recorded incidents of bicycle theft and burglary tend to be more closely correlated with less dense, higher-income districts.

It is therefore possible to summarise, in very general terms, seven contrasts, presented in the table below (Table 2). This table provides a synthesis, derived from PCA and hierarchical classification, of the main structuring axes that differentiate the various types of districts identified (the latent dimensions).

TABLE 2: **General contrasts observed between variables**

Population and building density	Level of greenery
Jobseekers	Median income
North Africa and Turkey	OECD
Businesses and single persons under 30	Single persons over 65
Males	Females
High recorded crime	Low recorded crime
Drugs and urban violence	Burglaries and bicycle thefts

Conclusion

This analysis of recorded crime variables in relation to socioeconomic and demographic variables has highlighted a new and unprecedented understanding of the territory. First, we examined the relationships between heterogeneous variables and the way in which recorded crime variables fit into a broader contextual view of districts.

Second, the typological structuring of districts into three clusters highlighted the spatial dimension and that the centrality of crime the Brussels-Capital Region, with crime levels decreasing as distance from the city centre increases. Based on the typological analysis, the three clusters can be summarised as follows:

TABLE 3: **Characteristics of the districts classified into 3 clusters**

Cluster 1: periphery	**Cluster 2: centre-southeast**	**Cluster 3: centre-north-west**
Income +++	Income -	Jobseekers +++
Greenery +++	Density +++	North Africa and Turkey +++
Single persons over 65 +++	Businesses / persons under 30 +++	Males +++
Household with children +	Household with children -	Density ++
Females +++	OECD +++	Household with children +
Crime -	Crime ++ (especially bicycle theft)	Urban violence +++ Drugs ++
SIAMU -	SIAMU ++	SIAMU ++

We reiterate We reiterate that correlation is not causation, and that two variables with high values for the same district do not necessarily mean that one influences or explains the other. Moreover, recorded crime reflects both individuals' willingness to report offences and police proactivity in terms of investigations and controls.

However, despite these methodological precautions, the present analysis shows that districts with higher density, lower median incomes and higher unemployment rates have significantly higher crime rates. These results suggest (and confirm) that crime is closely correlated with socioeconomic and structural factors, rather than isolated individual behaviour. They therefore argue in favour of long-term, targeted policies aimed at strengthening education, continuing urban regeneration efforts, housing and the development of public spaces, and promoting social and economic integration. Acting on these levers appears essential to reducing social and territorial inequalities and fostering a climate of long-term security and cohesion.

***

It should be noted that the typology adopted here combines recorded crime variables and socioeconomic variables; however, it would also be possible to focus on one or the other, to give greater weight to crime variables, or to construct alternative classifications. It would also be valuable in the future to include temporal data on recorded crime in this type of approach.

The five cluster classification also provides an opportunity to confront current perceptions and raises questions regarding official territorial reference frameworks in administrative matters.

The present analysis is conducted at district level and includes crime variables, which is not the case, for example, in the Belfius Bank typology of municipalities5 or in the KUL standard on the financing of police capacity by municipality, which has been regularly discussed since its introduction in 19986. Beyond territorial scale, it may therefore be useful to consider integrating recorded crime variables in order to refine these typologies.

This exploratory work also highlights the value of having as many variables as possible available at district level. Recorded crime data at district level, as used in this analysis, remain limited to a series of incidents that required a specific request to police services and prior processing by strategic analysts7 to extract relevant and available figures at statistical sector level on a case-by-case basis. These data were then transposed to district level. The list of incidents used in this study would benefit from being expanded.

Finally, it would be advisable for recorded crime data to be further developed, processed and disseminated among research centres and universities in order to encourage debate and improve understanding and analysis of crime in society more broadly.

With regard to socioeconomic and demographic variables, continued census efforts and the regular updating of open-source public data —particularly those from the Monitoring of the Districts— are essential to ensure their timeliness and regular availability for this type of analysis.

***

5. Typology produced in 1997, updated in 2007 and 2017. See Belfius Bank Research Directorate, ''Typologie "socioéconomique" des communes'', Brussels, 2018. Available at: https://research.belfius.be/wp-content/uploads/2018/06/Typologie-des-communes-FR_ tcm_78-150618.pdf.

6. Magali Verdonck (2024), "La norme KUL et ses implications pour les zones de police en Région de Bruxelles-Capitale", *Working Paper* 24/01, Department of Applied Economics, ULB, 34p., available at: https://dulbea.ulb.be/wp-content/uploads/2024/02/Dulbea_WorkingPaper_24.01.pdf.

7. Federal Police, Coordination and Support Directorate Brussels (DCA).

Finally, beyond its empirical results, this study is of major methodological interest: it draws on advanced quantitative methods and cross-referenced, typological approach to provide a renewed perspective on recorded crime in the Brussels-Capital Region. By combining multivariate analysis, dimensional reduction and classification, the study goes beyond a purely descriptive reading of the figures and offers a more structured, nuanced and operational understanding of the phenomena observed.

Moreover, this approach represents a first step that opens up significant perspectives. Its underlying logic, tools and findings are potentially reproducible and could be adapted or shared with other regional observatories or interested institutions. It thus lays the foundations for an evolving analytical framework, facilitating inter-territorial comparisons and the development of a shared data culture in the field of security.

For publications of studies by the Observatory: www.safe.brussels

Author: Nathanaël BAILLY (Observatory - Security Image Unit)

Contact: safe.brussels - Tel: +32 (0) 507 99 11 - contact@safe.brussels - Rue de Ligne, 40 - 1000 Brussels

Responsible publisher: Sophie LAVAUX, Director General of safe.brussels

D/2026/14.168/11