Social interventions connect to the origin of the population

Samuel Schiess

Joschka Geissler

  –> max. 6 -12  pages (doesnt matter)

# Abstract

sdfasf

# Introduction

The world human population is steadily increasing. In developed countries, > 70 % live in cities or urban areas (Bettencourt et al. 2007). With the increase in the population, humankind is facing new challenges and existing problems can get worse. Not only problems of famine, water scarcity or increased energy demand can occur, also socially problematic situations arise when many people live in the same spots (Glaeser and Sacerdote 1999). With the density of humans living in the same area, also psychological problems can occur. The social pressure rises and the concurrence in every domain is increasing, as so many people in the near areas specialize in the same things. On the other hand, people need to choose a lifestyle and find their way of living and are responsible for their own well being. The pressure and responsability can often end in psychological problems, which can lead to domestic violence, self-hatred and even suicide attempts. For actions that do not respect the law, the police is forced to intervention and the reason is being recorded.

On the other hand, not only a rise of the population and the subsequent problems may be observed today. Also the mixing of the cultures and nationalities all over the world, gets more important and brings new possibilities and advantages, but also new problems. In common speaking, the strange cultures have often another treatment of psychological diseases, health, crime and violence than the native culture. This may be right in some cases, but the connection between social troubles and the percent of foreigners in specific areas has never been clearly stated (Entorf and Spengler 2000).

In the case-study of the municipality of Vernier, we have access to data from the interventions due to social reasons and the demographic properties of the area under study. The context of those two variables is often taken as a political argument from right-wing parties, but a more narrow investigation will be provided in this paper. Our study area is formed by the city of Vernier with its 768 hectares and 35 300 inhabitants, situated in the Canton of Geneva and thus a part of the metropolitan area of the city of Geneva (“Vernier Officiel”, n.d.). Mostly the spatial correlation between those two variables will be looked at and also other statistical measurements will be provided, based on the data set and the area given.

# Hypothesis

In the first step we want to investigate the existence and the properties of hotspots in the region of Vernier regarding the two factors social interventions by the police and the origin of the population. Secondly, the relation between those two variables is of interest and their spatial correlation will be investigated. The expected facts are, that both parameters underlie a spatial distribution and hotspots exist and that the positions of the hotspots and thus the parameters are positively correlated to each other.

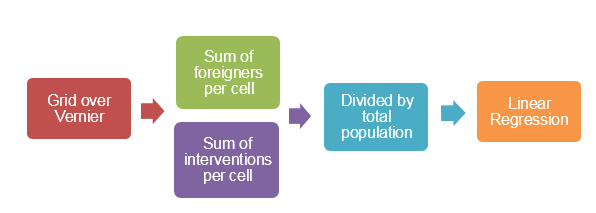
# Data

We use population data from STATPOP of the year 2015 which is annually  surveyed by the Federal Statistical Office of Switzerland for the households of the country (“Canton Geneva”, n.d.). From this data, informations about the total population, the number of swiss, non-swiss and non-european inhabitants have been used. The second dataset used is the data published by the cantonal police department of Geneva. The dataset provideds all police interventions caused by social reasons in the area of study from 2014 to 2017.

# Methods

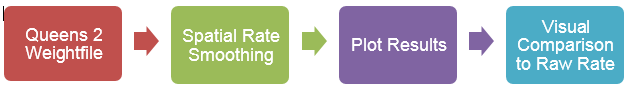
All spatial and statistical analyes have been performed with the software Geoda. QGIS (Version 2.18.) has been used for visualization.

Having the two datasets described, we were able to perform a first statistical analysis. Therefore we performed the processing chain shown in Figure 1. The first step was to create a grid over the study area of vernier with a cell-size of 100 m x 100 m. For each cell, the total number of interventions has been summed up and divided by the total number of the population in this specific cell. Additionally, the number of Swiss, Non-Swiss, and Non-European inhabitants has been computed for each cell. To get the percentage of the different origins, those numbers have also been divided through the total number of the population within the cell. With the obtained variables, we were able to perform a first Linear Regression (Non-Swiss/Interventions, Non-European/Interventions). In addition to the Linear Regression, we also computed the Raw Rate. Therefore, the number of interventions divided by the total population was divided through the percentage of Non-Swiss and Non-European inhabitants within each cell.



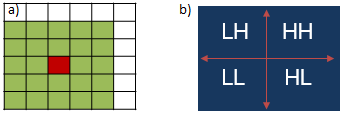
Processing chain for Raw Rate and Linear Regression

In order to find the socalled hotspots of our variables we had to perform a spatial analysis. The workflow for this spatial analysis is shown in Figure 2.



Processing chain for Spatial Rate Smoothening

First, a Queen’s 2 weightfile was computed for the grid of vernier. Figure 3a shows the idea for this weightfile. Each cell with a distance closer or equal to two is taken into account (green) for the computation of the value of the cell of interest (red). The values of the white cells do not have an influence on the calculation of the red value. For instance, if we want to calculate the spatial rate smoothed value for the red cell, we would divide the sum of the first parameter (e.g. Number of Interventions) of all green colored cells through the sum of the second parameter (total population) of all green cells. Several “Spatial Rate Smoothed”-Rates were computed (Percentage of Non-Swiss inhabitants per population, Percentage of Non-European inhabitants per population, Number of Interventions per population). All results have been plotted, visually analysed, discussed and compared to the raw rate which has not been spatially smoothed.



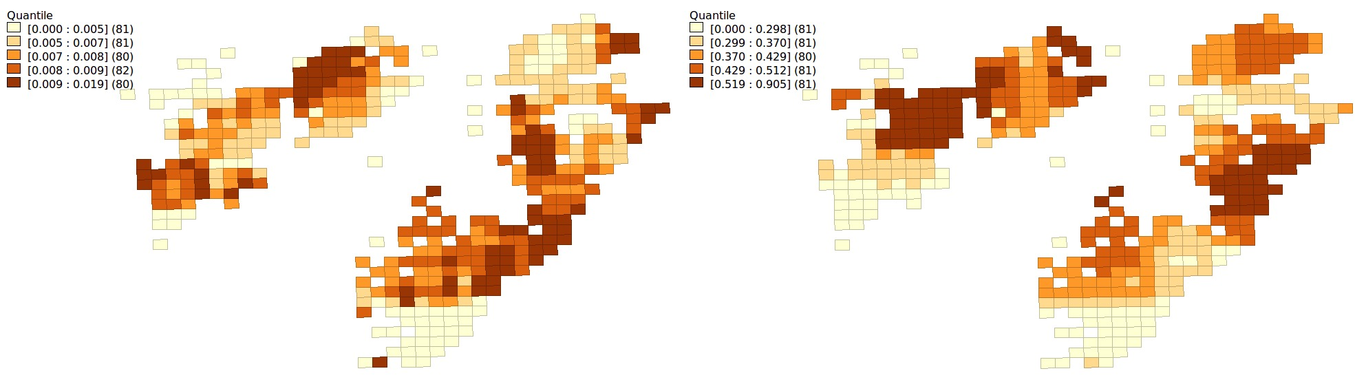
Spatial Analysis; a) Concept of the Queen’s 2 weightfile; b) Concept of the Local Morans I

To investigate dependencies between the variables the Local Morans I-analysis was performed additionally for the raw and the spatially smoothened rates. The Local Morans I provides a possibility to check spatially distributed attributes in its context. As we are interested in the connection between the interventions and the origin of the population, we have chosen the bivariate method. It enables a comparison of the values for the variables with its neighbour and presents the correlation of the two variables at the same time.As Figure 3b shows cells with relatively high values for both variables end up in the top-right quadrant “HH”. Cells with relatively low values for both variables end up in the “LL”-quadrant. Cells with a high and a low  value for the variables are located in the “LH” or “HL”-quadrant respectively. In case the variables of a cell are not significantly high or low (the p value is not small enough to show significance) the cell is categorized as non significant.

The main results have been visualized using QGIS.

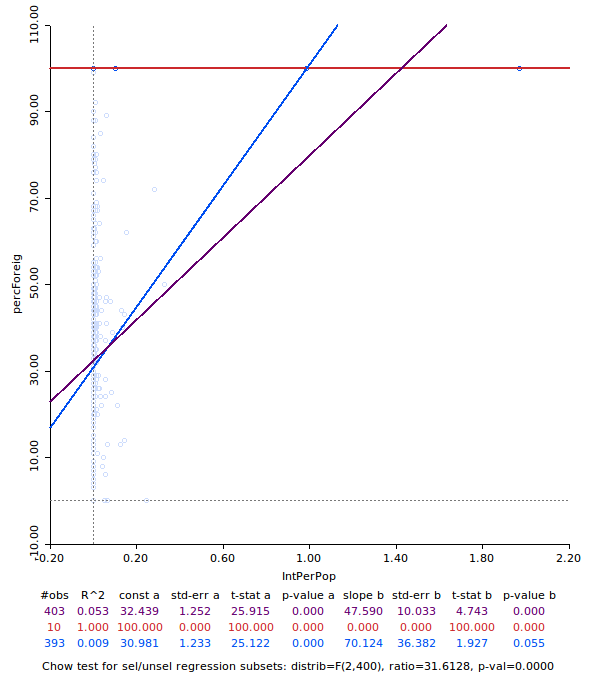
# Results

Regarding the hotspots, a simple representation of the variables over the area has been applied. To observe existing neighbouring effects and to include a smoothening filter, the SRS was applied for both parameters using a Queen 2 weightsfile. The cells are organized in 5 quantiles, with the darkest being the highest and the lightest being the lowest quantile.

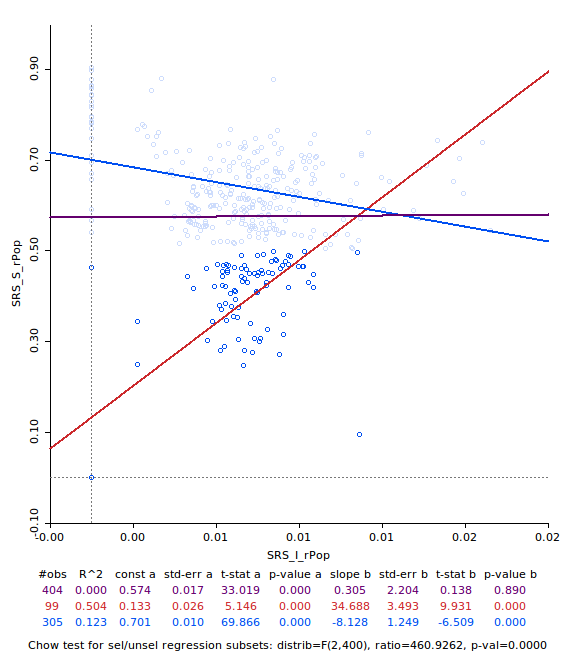


Shows the cells of the municipality of Vernier for the smoothed data using SRS. On the left we see the Interventions divided by the population and on the right the percentage of foreigners.

To research the relation between the two parameters, we first used common statistical methods such as a regression of the data. In the following, the two scatterplots for the two parameters are presented. The first shows the regression and its statistics for the raw data, while the second shows the variable interventions with the percentage of Swiss inhabitants for the smoothed data according to the SRS and the Queens 2 weightfile. in FIGURE the points with a 100% foreign population are selected and the blue line shows the regression without those ten points. In the second scatterplot, the cells with < 50% Swiss citizens are selected and the blue line shows the regression when ignoring those.

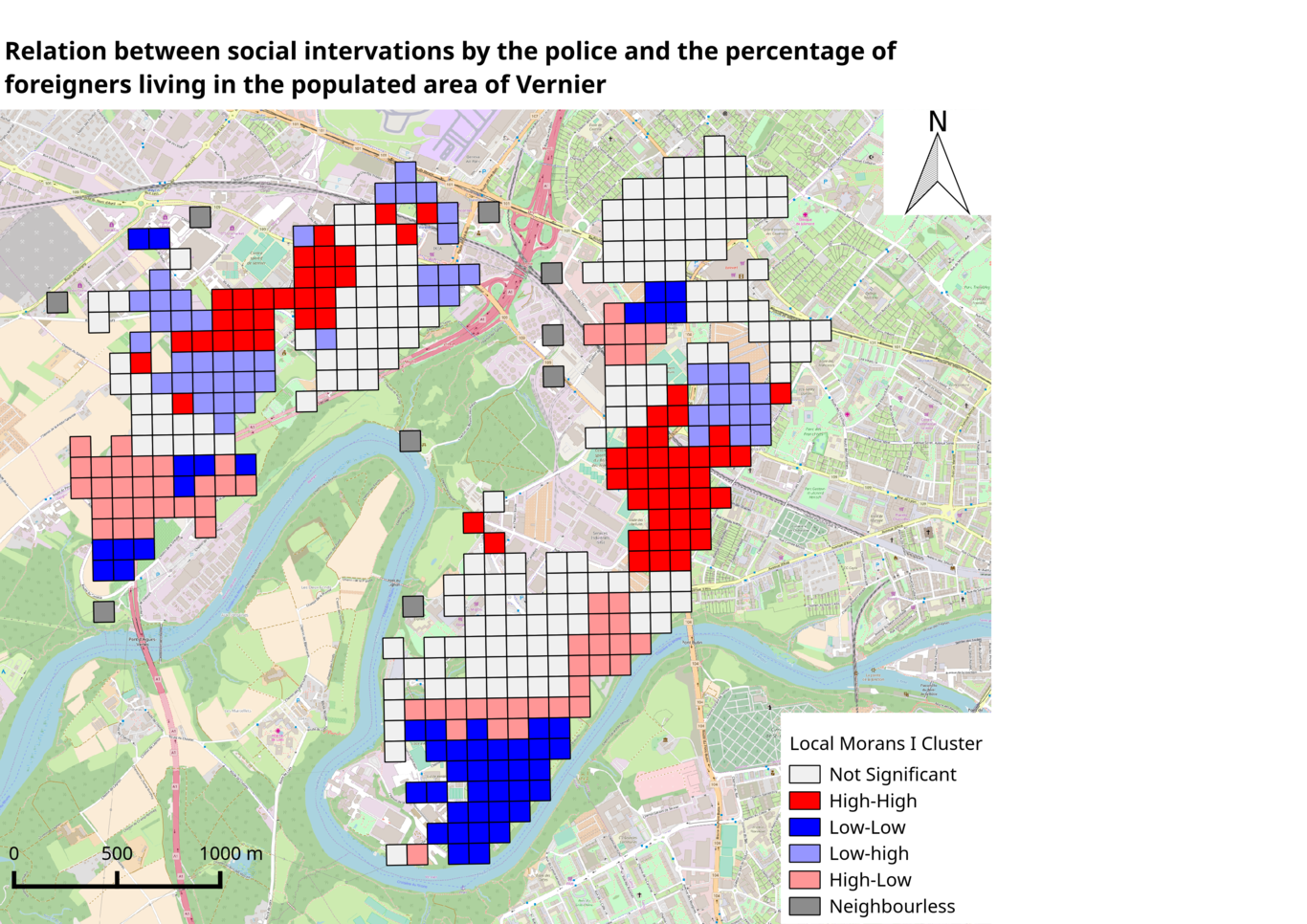


shows the regression line between the interventions due to social reasons per population and the percentage of foreigners in the cell for the raw data. The violet line and data is the regression for the whole dataset, while the blue comprises the same regression when ignoring all the points with 100 % foreigners.

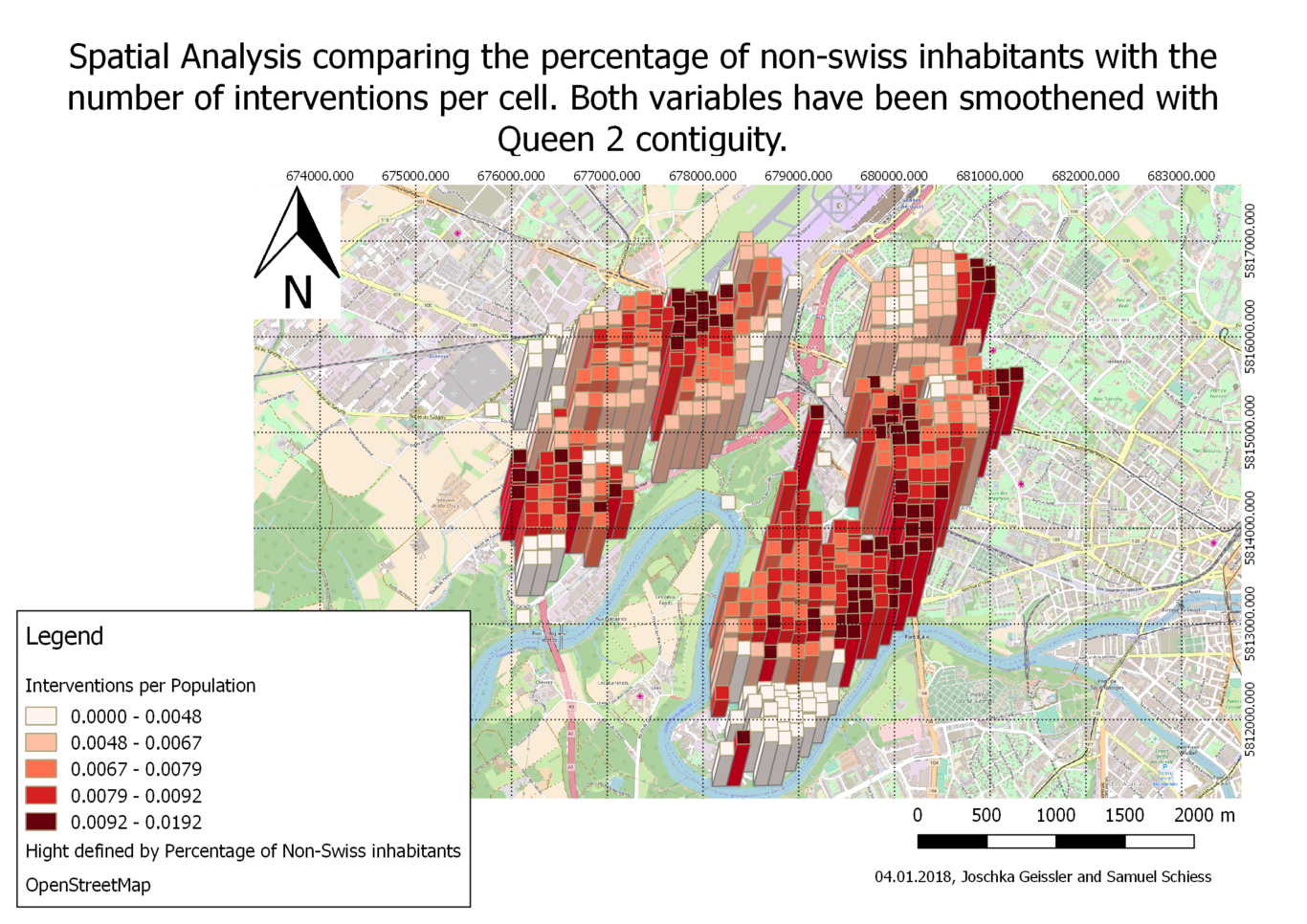


shows the regression line between the interventions due to social reasons per population and the percentage of Swiss inhabitants in the cell for the smoothed data using SRS and the Queens 2 weight file. The violet line is the regression of the whole data and the blue one is the regression when ignoring all points with a Swiss population of < 50 %.

To investigate the spatial correlation between the two variables, we apply the Local Morans I, using the same Queens 2 weightfile as before. The results are plotted on an Open Street Map using the software QGIS. 28% of the cells have a significant relation, which is supporting our hypothesis. 42% of the cells do not have a significant correlation between the two variables and are therefore marked in white.



The map shows the cluster representation of the Local Morans I using the Queens 2 weightfile. A significance of 0.001 was assumed. Two hotspots of High-High relation can be spotted an one large spot of Low-Low relation is situated in the south of the municipality area. The background map has its source in the Open Street Map integrated in QGIS.



This is a caption

# Discussion

Both hypotheses are not thrown considering the smoothed data of the SRS analysis. Nevertheless, it is important to point out, that the smoothing effect of the SRS is quite high and a loss in precision of the data can not be excluded. Therefore, the new information varies from the provided data and the statistical check is less high than with the raw data.

For the analysis of the hotspots, the representation of the raw data and the smoothed data are already a solid base. The percentages of the foreigners living in the cells are already spatially correlated, before smoothing with SRS. Two hotspots with a share of more than 50% foreigners per cell can be detected, both with the raw and the smoothed data, as seen in Figure 4. After applying the SRS, the two hotspots are clearly separated from the other cells. This clustering of the nationalities in urban regions has already been investigated further and this fact also adapts to Swiss cities (Adelman et al. 2016).

Also for the interventions divided by the population some hotspots are established, although not as clear as for the percentage of foreigners. On the other hand, the raw data would not show clusters in any quantile. For this reason, the hotspots of the smoothed data need to be handled with a critical mind. The literature also states a clustering of social interventions. For this purpose, our study area is certainly not large enough. Also a time series of more years would be more meaningful, as the dataset of the interventions is not that high over the studies region.

The boxplots in Figure 5 and 6 do not clearly show a correlation between the variables, neither for the raw data nor for the smoothed. Also when playing with the selection of the cells, no clear correlation can be stated. When normalizing the data, the slopes do not show the expected slope. In addition, the R2do never exceed 0.1, which makes the regression not reliable, with a huge error that can not be explained by the regression curve. Thus, with a statistical analysis of the connection between interventions and percentage of foreigners no clear correlation can be detected.

For the further investigation, a spatial analysis was adapted using the LISA Morans I. When comparing the clustering of the Local Morans I in Figure 7, clear tendencies exist in building two hotspots for the High-High correlation and also one for the Low-Low connection. Also in this case the smoothing effect of the SRS method must be taken into account, which helps to build clusters from the variables. With only 28% of the area, resulting in 113 cells showing a significant relation between the two variables, the effect of the variables on each other are not that clear. The clusters of the High-High connection are in those areas, where it was to be expected according to the hotspots of the smoothed data in Figure 4. The spatial distribution of interventions and foreigner population would be expected as more correlated, according to secondary literature. The same reasons as for the hotspots can be held responsible, as the area is not large enough and the data not that comprising (“Criminality in Switzerland”, n.d.). Also, the interventions due to social reasons do not necessarily correlate with crimes in general.

In general, the two important living areas in Vernier, Lignon and Avanchets, do not specially appear in neither of the analysis. The reason could be, that the interventions are divided by the population living in the cell. But also in the analysis of the percentage of foreigners per cell are the areas not showing special properties regarding the variable. The region of Lignon has become a popular living area and therefore, the prices do not specially attract foreigners or other social classes.

When looking at the results overall, it can clearly be said, that the correlation between the percentage of foreigners living in an area and the social interventions by the police is rather weak. Other properties of the society such as education, wealth, stress etc. must have a more important influence in this region and could be investigated in a further study with more input data regarding those topics.

(“Criminality in Switzerland”, n.d.)

(“Populists in Europe”, n.d.)

We have choosen this Queens2 weightfile because of the following reasons:

* Rook weights will usually have fewer neighbors than queen weights, which leads to the fact that each neighboring cell has more influence. For the analysis performed for this report it was important, that single cells have a low influence.(“Technical Questions | GEODA”, n.d.)
* 10 neighboorless cells for Queen1 and Queen2
* Queens2 was choosen to really get some smoothening

# References

Bettencourt, LM, J Lobo, D Helbing, C Kühnert, and GB West. 2007. “Growth, Innovation, Scaling, and the Pace of Life in Cities.”. *Proc Natl Acad Sci U S A* 104: 7301–6.

Glaeser, Edward L., and Bruce Sacerdote. 1999. “Why Is There More Crime in Cities?”. *Journal of Political Economy* 107 (S6). University of Chicago Press: S225–S258. doi:10.1086/250109.

Entorf, Horst, and Hannes Spengler. 2000. “Socioeconomic and Demographic Factors of Crime in Germany”. *International Review of Law and Economics* 20 (1). Elsevier BV: 75–106. doi:10.1016/s0144-8188(00)00022-3.

n.d. <http://www.vernier.ch/fr/portrait/presentation/.> <http://www.vernier.ch/fr/portrait/presentation/.>

n.d. <https://www.ge.ch/statistique/domaines/01/01_02_2/methodologie.asp.> <https://www.ge.ch/statistique/domaines/01/01_02_2/methodologie.asp.>

Adelman, Robert, Lesley Williams Reid, Gail Markle, Saskia Weiss, and Charles Jaret. 2016. “Urban Crime Rates and the Changing Face of Immigration: Evidence across Four Decades”. *Journal of Ethnicity in Criminal Justice* 15 (1). Informa UK Limited: 52–77. doi:10.1080/15377938.2016.1261057.

n.d. <https://www.vimentis.ch/content/docs/PA_Auslaenderkriminalitaet_Endversion.pdf.> <https://www.vimentis.ch/content/docs/PA_Auslaenderkriminalitaet_Endversion.pdf.>

n.d. <https://interaktiv.tagesanzeiger.ch/2017/populisten-vormarsch/?nosome.> <http://www.tagesanzeiger.ch/19118734.>

n.d. <https://geodacenter.github.io/questions.html#choice.> <https://geodacenter.github.io/questions.html#choice.>