Hadar Neighborhood Dashboard

Interactive Intelligent Systems - 096235 - Spring 2021

Gal Goldstein 204403174, Noa Shmulevich 205737935,
Shoval Zandberg 205791700, Yotam Martin 308044296
Supervised by: Dr. Ofra Amir

Introduction

Cities are human settlements subject to constant evolution. They involve various types of systems, like transportation, housing, health and sanitation, security and protection in continuous interaction and change. In this scenario and based on the deployment of digital technology, a huge amount of city data is produced on a daily basis. This data can be used to build and measure indicators to study and analyze different urban phenomena. Information obtained through such indicators enables us to understand city patterns, develop public policies and implement corrective actions.

Our main goal is to help the decision-makers to understand the current status of Hadar neighborhood, in the fields that interest the Hamal and the municipality of Haifa, and to help them get better decisions, backed by data. The Hamal dataset has many different features that relate to various fields, and decision-makers might miss important insights due to its non-concise structure.

We created a semi-annual brief in the format of a dashboard that summarizes the current state and the various changes that have taken place in Hadar neighborhood, comparing between two consecutive half-years. Our dashboard is focused on various topics and areas such as safety, crime, income, elderly, and population. These different topics are analyzed per statistical area and for the entire neighborhood. The dashboard is publicly available via iis-dashboard.herokuapp.com (allow some time to load if errors appear and refresh after 5 minutes from first access). The code for this dashboard is also publicly available via github.com/yotammarton/iis_dashboard.

As of June 2021 the crime in the Hadar neighborhood is a major and actual problem. The mayor of Haifa, Einat Kalisch-Rotem, said “I have no intention of compromising on the issue of safety in the city of Haifa and in the Hadar neighborhood in particular” [source]. Due to this significant crime problem in the neighborhood, we focus on personal safety and crimes. By having a factual account of all the criminal occurrences that happened in a city during a given time interval, with a geographical location of the infringement and the type of crime, we are able to assign the city’s different regions with safety scores. By that, we can help the decision-makers track this topic closer and make decisions accordingly.

Furthermore, we integrated additional datasets (CCTV locations, 106 call center event log, neighbors conflicts) that help us in the task of crime and personal safety analysis. Some of these datasets were created by us and some received from the Hamal. It should be noted that the initial dataset given to us by the Hamal does not contain enough data samples and does not contain dates and times with variation, and therefore we generated synthetic data for different times. This helps us to show our proof of concept. In the Methods section of this report one can find a detailed explanation regarding how to reproduce the dashboard with new data.

During the project we kept in touch with the Hamal team and got their feedback regarding different parts of the project and changed our work accordingly (focusing more on the crime and safety score, what elements should be part of the safety score, and what kind of data should be presented regarding the other topics).

Our result is an urban interactive dashboard that presents the state of Hadar from different aspects and provides an overview of how Hadar is performing. It provides intelligence to help the Hamal and the relevant decision makers to know and understand the neighborhood and its current state better, and compare its status to previous time frames and previous reports.

Literature review

Our main goal is to get Hadar, and hopefully Haifa, one step forward towards being a smart city. As presented in [1], “Smart city” is defined by IBM as the use of information and communication technology to sense, analyze and integrate the key information of core systems in running cities. The structure of a smart city includes a perception layer (collect information), network layer (transmission and process the information), and application layer (analyze and process massive data). Among the typical applications of smart cities, is smart public service and the construction of social and urban management. This enables them to transfer data, work together, and form the core of the urban management system, achieving seamless management in a command center, similar to the Hamal in Hadar. Some of the main difficulties in building such a city are handling the large-scale space-time information and managing to achieve intelligent analysis and decision support. We address these issues in our work.

For many cities, the smart city has materialized into urban indicator projects. Within these initiatives, the visual representation of data is a fundamental component. One of the visualization options is to use geospatial dashboards. [2] present a critical review of the research and development of geospatial dashboards, including the integration of maps, spatial data analytics, and geographic visualization for decision support and real-time monitoring of smart city performance. The definition of geospatial dashboard according to [2] is a web-based interactive interface that is supported by a platform combining mapping, spatial analysis, and visualization with proven business intelligence tools. An important part of a geospatial dashboard is the indicator which is “a measurement or a set of measures to evaluate a complex social, economic, or physical context”.

Urban dashboards have emerged as a technique to visualize the data in an accessible way. Such an example is the Dublin Dashboard [3], an interactive website that presents data from a variety of sources about Dublin (Ireland) through a series of interactive maps, graphs, and applications. One of the modules described in this paper presents a similar idea to our project, The “Dublin Overview” module. The purpose of this module is to provide a single view of the current values for key indicators. It provides the current value and the difference from the previous values regarding multiple topics. The cognitive load required to understand this module is low. It also enables the user to drill down and get more information by showing an interactive map or a longer time series graph. The temporal phase of the series data is typically monthly, quarterly, or annually.

As mentioned in [4], the public sector usually struggles with adopting dashboards, compared to the profit-oriented private sector. The source for that struggle lies in three main reasons:

1. The differences across the organization’s departments regarding the concept of public value and the complexity to measure it.
2. The large fragmentation of responsibilities across public organizations, as well as the fragmentation of stored government information.
3. The conflict between sharing the organization’s information widely and the commitment to many actors that are involved in public organizations.

When designing a dashboard, we need to consider the cognitive effort our dashboard will require from the users of the system. We act according to the Cognitive Load Theory (CLT) presented in [5] regarding geographic information. CLT assumes that learners have limited working memory, which is connected to unlimited long-term memory.
The author also focused on improving map learning by reducing cognitive load. Total cognitive load must be less than the working-memory capacity of the learner. One may require maps that are as simple as possible in order to manage cognitive load. But, as expertise increases, a simple map may not allow learning to continue for advanced learners. This problem has been called 'the expertise reversal effect'. The visualization's designers should measure both subjective and objective measurements. Subjective measurements are gathered through the mental effort required for a task, typically measured by asking participants to rate the difficulty of a task. The objective measurement represents the performance of participants as they deal with a secondary task that is unrelated to the primary task. They have to also consider the innate differences among the individuals doing the learning- both biological and environmental factors.

[6] Have proposed a theory for multimedia learning which is defined as learning from words and pictures. The theory is based on three assumptions about how the human mind works: (1) The dual-channel assumption- ''humans possess separate information processing channels for verbal and visual material.'' (2) The limited-capacity assumption -''there is only a limited amount of processing capacity available in the verbal and visual channels.'' (3) The active-processing assumption- ''learning require substantial cognitive processing in the verbal and visual channels.'' [6] Identifies five overload scenarios and offers solutions to typical problems. [5] Relates to the scenarios presented in [6] to learning with maps. The main scenarios and solutions that we apply in our work are: (A) Do not overload both verbal and visual channels (for example presenting the text and information on the map in small successive segments). (B) Direct the learner's attention to the critical information (for example, by using bold or italic lettering) and remove extraneous information. (C) Avoid ambiguous representations. (D) Reduce the need to hold information in working memory by presenting the text together with the map.

In [7] the authors describe the term "Design Thinking" and its importance in engineering, specifically for system design. Design thinking is generally defined as an analytic and creative process that engages a person in opportunities to experiment, create and prototype models, gather feedback, and redesign. [7] Specify the different characteristics a design-thinker should have: (1) Human- and environment-centered concern. (2) Ability to visualize. (3) Systemic vision. (4) Ability to use language as a tool. etc. The dashboard design part of our work is highly influenced by these characteristics.

Zooming in on the different topics our semiannual report would hold, we focus on safety and crime analysis in the Hadar neighborhood. In our project, we deal with spatial data. [8] Explains what spatial data is, how to deal with it and the importance of taking it into account. The author distinguishes between two types of spatial analysis methods. The first is often referred to as exploratory spatial data analysis (acronym ESDA) and is concerned with the description and exploration of spatial data. Typically, the result of this analytical method is visualized with the use of geographic information systems (GIS). The second is spatial modeling. This term refers to a set of regression analysis techniques that are adapted to spatial data. We attempt to predict the outcome variable at each location as a function of variables of the focal location and possibly of variables at other locations as well.
The reason why spatial data is important to us when we analyze crime data is because the spatial arrangement of the data might bias the findings of ordinary statistical analysis (because most statistical techniques are based on the assumption of independence between observations, our results may be biased if this assumption does not hold).

[9] Focuses on the city of Bogota as a model for security analysis, defined by official crime records and social strata percentages. With that data the authors classified safer areas in the city. [10] Shows that there is a causal effect of poverty on crime and violence. As such, by having some dataset that can detail the proportion of people living in poverty conditions in each city region, it would be valuable information to determine the safety of a given region. [9] Offers a formula based on social and criminal data to calculate safety scores for each area. We modify this formula for our task. [10] Refers to two measures of poverty. The first is the fraction of the population whose annual income is below the poverty line - this measure captures the incidence of poverty. The second is the poverty gap, defined as the average distance of the poor from the poverty line - this measure captures the depth of poverty. We use a version of the second score because our data contains income information per statistical area.

[11] States that planners and decision makers aim to understand the way the components of the built environment affect unsecured environments. The author suggests a new model to evaluate a qualitative sense of security that is understandable on a human-scale in the built environment.

The Security Rating Index (SRI) establishes a GIS-based, quantifiable system to identify and rate insecure urban spaces to be used by urban planners and city decision makers to evaluate and improve urban resilience. By using a new approach to quantify secured urban features, we can understand the security rating of different areas in the urban fabric and promote the development of better place making. The paper defines the term ‘security’ in the built environment as related to the spatial and morphological design of a location that influences one’s sense of security. The main selected urban elements used for measuring the sense of security inclusively and equally in each urban area were mixed uses, streetlights, building proximity, and intersections (number of intersections and distance between junctions). In conclusion the SRI model can point out urban areas with potential for crime events, hence it has the potential to prevent crime in urban locations prone to low levels of security. The SRI can assist greatly in future development by allowing planners to understand developed urban situations and support crime prevention in existing and new built environments.

From the technical aspect, we examined a few options for data visualization. We decided that an important aspect is reproducibility, which will be best achieved by using the available Python libraries. [12] Presentes an overview and comparison of free Python libraries for data mining and big data analysis. It presents more than 20 libraries and separates them into groups. We focused on the ‘data visualization’ group. Plotly seems to be the most powerful library in the data visualization field, and we found that it is possible to integrate Plotly graphs into web pages with Dash which enables a user to explore the data in an interactive way.

[13] Describes an interactive Visualization for Real-Time Geospatial Data with Plotly Dash. This platform helps the user recognize key observations, view historical context and determine trends, which is exactly our goal. The visualizations can be reproduced with minimal effort. The interface enables presenting details in three primary forms: (1) Hovering (a concise summary of current field values at a given point). (2) Query result (provided via user selection). (3) Trends (time-series plot of the selected attribute). The value of this visual analytics environment to our project lies in its support for time-series data analytics tasks and scalability to large datasets.

Methods

Data preprocessing:

We take the raw data of the Hamal, sub-categorize it to a few categories (population, crime, etc.) and choose the important columns in each category for the analysis. The input to the dashboard creation process should be pairs of tables. Each pair consists of two consecutive half-years tables, both are focused on the same category with identical structure, but different values. The first table data was created at ‘time 0’ which is the time of the previous report, and the second table was created at ‘time 1’ which is the time of the current report. In that way we can compare the same data objects from two consecutive periods of time.

Data generation:

Since the given data by the Hamal is not rich enough and has low variation of dates, to show the proof of concept of our project, we decided to treat the given data as ‘time 0’ tables, and generate new data for the ‘time 1’ tables. We divided the tables to two types: Non-aggregative and aggregative. The non-aggregative tables are tables where each sample represents an individual object / event. For example, table of abandoned houses in the neighberhood, or table of holocaust survivors. The data generation of this type of tables should be based on the real data table (which is the ‘time 0’ table) with some small changes, since in two consecutive reports, we would expect, for example, to see a similar amount of abandoned houses around the city, with some changes. Therefore, the non-aggregative tables data generation include three phases: (1) Duplicate ‘time 0’ table and add x samples, where x is distributed gaussian with mean and deviation which are set in advance. (2) Each attribute value of each added sample will be randomly selected from the existing attribute column values. (3) Remove from the output table of (2) y samples, where y is distributed gaussian as well, with parameters which are set in advance.

Since we decide on the mean and variance of x and y, we can decide what changes to show in the proof of concept, e.g., an increase of ~4% in the number of senior citizens.

The aggregative tables contain row for each statistical area, with aggregated values such as averaged salary. The aggregative tables data generation includes taking the ‘time 0’ table and adding to each cell a value which is distributed gaussian with the mean and the variation of the column of this cell. Another value can be added/subtracted, if an intended increase/decrease for some data cell is desired for demonstration needs. The data extraction from the Hamal raw Excel data file and the data generation process can be found in Google Colab notebook [link].

This notebook should be executed before the rest of the project code (GitHub repository). The reason for separating the notebook from the project code, is because these phases won't be reproduced by the Hamal, since they will use only real data. Yet, we wanted to demonstrate the generation techniques we used.

Needed input for the dashboard creating (reproducibility):

An important aspect of our work is reproducibility. We want to enable the Hamal to easily reproduce the dashboard with any new data that will be collected in the future. Therefore, we specifically described the needed input for the supplied code, in order to output the dashboard [Appendix 1] .We mostly rely on 2 tables for each topic where each represents a different time frame. Most tables derived from the Hamal original data (where our manipulation of the data is in the given GitHub repository code).

Creating the dashboards:

As a preliminary step, we planned the structure and content of the dashboard and its division into the various tabs using Excalidraw, which is a virtual collaborative whiteboard tool that enables creating sketch diagrams. Our sketch [link] helped us to begin the code writing process. We use the Dash Plotly python library, since it suits our needs as recommended in [12] and [13]. Dash is a productive Python framework for building web analytic applications. We create several modules - one for each chosen topic. We enable user interaction with the relevant parts of the dashboard, such that the user can choose on which subject and areas he would like to focus and to be able to achieve custom insights. We deployed the dashboard to a public url using Heroku.

Safety tab:

One of the main issues we relate to in our dashboard is the safety in Hadar. As we said before, the personal safety and crime in Hadar is a very central issue that has preoccupied the Haifa municipality recently. We assign a score for each statistical area in Hadar. The higher the score, the safer the area is. This score is an interpolation of different topics that can affect the sense of personal safety in the urban space. Each sub-score was normalized, so that each score of an area would be relative to the other areas in Hadar. This normalization allows the user to have a way to relatively compare the different statistical zones.

The sub-scores are: neighbors conflicts, closed-circuit television cameras, abandoned houses, crimes and types of crimes such as thefts, robberies, body assault and sexual assault, inquiries to local municipality call center (106) regarding security and social issues, average income and population density. Each score is calculated according to the data of each statistical area. We will make sure each topic is represented by a scalar in [0,100] interval so that the scale will be aligned for all the different topics. We will give weight for each topic such that our users will be able to control the weights and interactively determine what topics are more important according to their perception and goals.

More formally, $score_i = \Sigma_{j=1}^{12} w_j \cdot s_j$ is the safety score for the i'th statistical area, where s_j is the score for topic $j \in [1, 12]$ and $w_j \in [0, 1]$is the weight for topic j (which represents the importance of the topic) and $\Sigma_j w_j = 1$.The final score is in [0,100].

We will now elaborate on the score's calculation for each topic:

- For **Neighbors conflicts, abandoned houses** and **inquiries to the local municipality call center** (security and social issues) the score will be calculated according to the amount of population: $1 - \#NumOfEvents / \#population_i$ where $population_i$ is the amount of residents in the i'th statistical area.
- For **closed-circuit television cameras** the score will be calculated per 1,000 people. $\#NumOfEvents * 1000 / \#population_i$. where $population_i$ is the amount of residents in the i'th statistical area.
- **Crime -** we will give two types of scores:
  - Score for total crime in the area will be calculated as 1 minus the percentage of total crime in the area. $1 - \#NumOfCrimesInArea / \#NumOfAllCrimes$.
  - Score for crime type (thefth, robbery, body assault, sexual assault): will be calculted as 1 minus the percentage of specific crimes from all crimes in the area. $1 - \#NumOfCrimeFromSpecificTypeInArea / \#NumOfCrimesInArea$
- **Poverty -** the score will be calculated as the ratio between average income in the area and the average income in Haifa (as found in the official data of the city - link) - $\#AvgIncome_i / \#AvgHaifaIncome$.
- **Population density** in the area. The score will be calculated as the ratio between the density in the area to the density in Haifa. $1 - (\#NumPopulation_i / \#AreaSize) / (\#NumPopulationInHaifa / \#HaifaSize)$.

To be sure we are giving a reliable score for each area we have found articles that suggest how to determine safety score. For instance [10] found that there is a causal effect of poverty on crime and also suggested how to measure poverty. [11] Mentions that a closed-circuit television camera is the most cited element that increases the sense of security. Moreover we consulted The Hamal on issues that are important for them to see in the security aspect.

Overall, we identify 12 elements that create our safety scores. For the ease of usage, as part of the safety tab, we gave the users to set the importance weights for each element. Each such weight is an integer in [1,5], with the option to reset to the default values (3 to all weights), as can be seen in Figure 1.

Figure 1: Importance score’s elements and their values.

Results

We will present our dashboard's tabs. Most of the dashboard's tabs have two modes, as presented in Figure 2. The first mode is the current status in each topic (e.g the amount of crimes, the population size, etc.). The second mode is a comparison between the current report and the previous report. The motivation for focusing on comparison is to see the latest changes in each measure, beyond just looking at the absolute current measurements, as already possible in some of the Hamal's tools. The comparison motif reflects in three ways: (1) Heat map of changes - as appears in Figure 2 ; (2) Annotation of percentages increase/decrease next to every bar in the bar charts - as appears in Figure 3 ; (3) Annotation of the absolute values changes when hovering the bars in the bar charts (e.g. 7→10) - as appears in Figure 3.

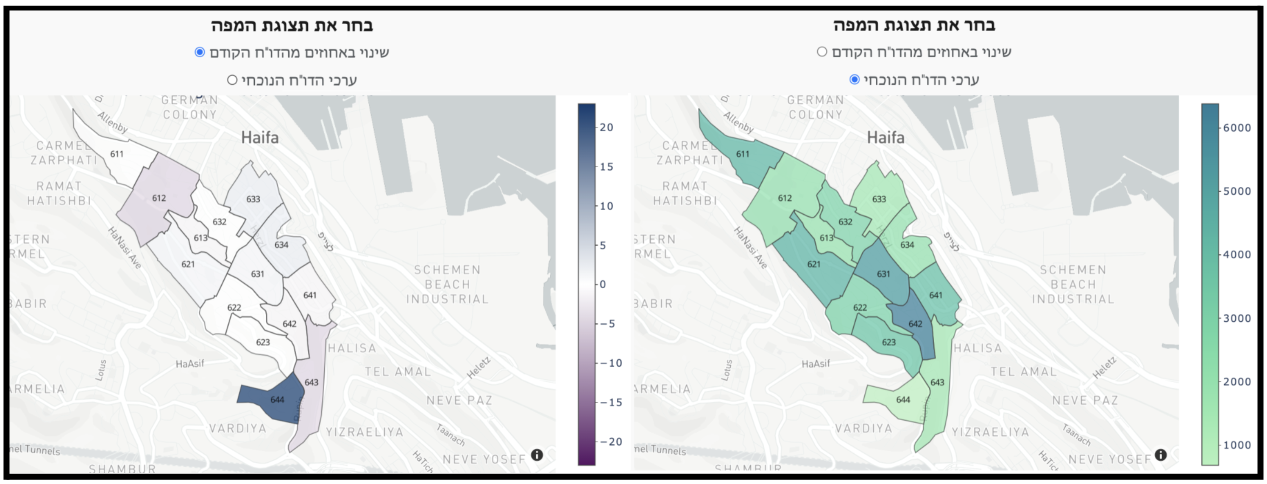


Figure 2: Example of two modes of tab's heatmap: 1) Percentage of change from the previous report (diverging scale was used to emphasize the difference between increase and decrease . 2) Current half of the year values.

We will summarize the information of each tab, while the appendices contain the full design. Note that the visualizations, texts, interactions, and layout that appears in the appendix are the final result.

While designing the different tabs and the dashboard we were also taking into consideration the cognitive element - presenting the most important graphs on the top of the screen, choosing an intuitive color palette for each scenario, adding explanations for each tab and for each graph if it is necessary and to not rely on the working memory of the users, as suggested in [5] and [6].

First tab [Appendix 2] is the **Main tab**, which serves as a welcome page. This page contains (1) Textual content explaining about the dashboard and the way of use ; (2) A map with division to statistical areas, and (3) A summary table with chosen measurements of current time and the % of change from the last report.

Second tab [Appendix 3] is the **Safety tab**. As mentioned before, the main focus of the dashboard is personal safety and crime. The Safety tab shows the safety score per statistical area. It includes (1) A heatmap with safety score per statistical area ; (2) An interactive widget of topics importance for the score calculation where the user can set its own importance for each component, and (3) Graph that presents all the sub-scores, this data will help to explain the shown scores. There is an option to present results for the whole neighborhood, or per statistical area.

- According to the Hamal needs, we added 2 additional tabs that each contains only part of the safety tab - The **'safety-immersive'** tab contains only the map and scores graph aligned next to each other so it suits the immersive screen which exists in the Hadarion. The **'safety weights'** tab contains only the weights for the safety score. The reason for creating these separate tabs is to enable the Hamal team to present the scores on their touch screen while presenting the graph and map on their immersive screen.

Third tab [Appendix 4] is the **Crime tab**, which includes (1) A heatmap showing the % of change in total crime cases between previous and current report, or the data for the current time ; (2) A trend graph with the number of crimes per month in the last year, and (3) Histograms of crimes by location and type. As in other tabs, there is an option to filter the views for a specific statistical area.

Fourth tab [Appendix 5] is the **Population tab**, which includes (1) A heatmap with % of change in the number of residents between previous and current report, or the data for the current time ; (2) Charts for age distribution, gender and Haredim residents distributions, with notes about the size of change compared to the previous report.

Fifth tab [Appendix 6] is the **Income tab**, which includes (1) A heatmap showing the % of change in average salary between previous and current report, or the data for the current time, and (2) A histogram of average salaries per salary type.

Sixth and last tab [Appendix 7] is the **Elderly tab**, which includes (1) A heatmap showing the % of change in number of seniors, or the data for the current time, (2) Number of holucaust survivors in the chosen statistical area. (3) A chart of distribution of seniors who are alone or receive food, and (4) A chart showing the number of cases per type of help needed for holucaust survivors. There is an option to filter on specific statistical zone.

All the dashboard's tabs contain interactive elements such as additional data while hovering different areas of a given graph as can be seen in Figure 3.

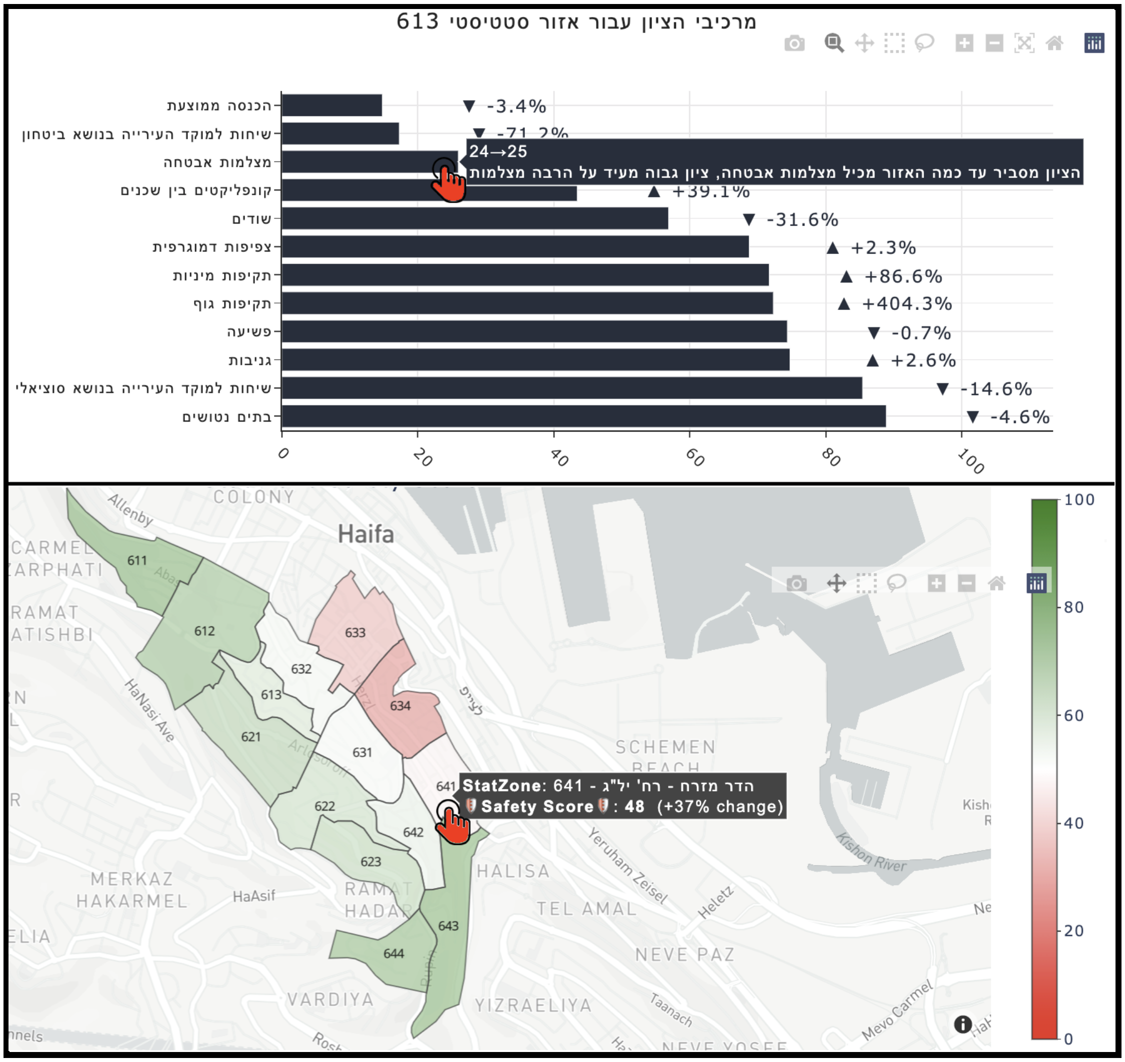


Figure 3: Examples for the additional data presented while hovering the interactive graphs.

Discussion

Overall, we learned a lot from this project and mainly gained the knowledge of using a new visualization tool to create dashboards. Furthermore, through this project we were in constant communication with the Hamal team and we modified our dashboard and added details according to their needs, which was a good experience.

A very interesting topic that took a lot of our focus and efforts was the safety in Hadar. This topic was very important to the Hamal team and we invested a lot of time and effort in giving them the most interesting insights. We had many group discussions regarding the elements that we took into consideration when defining the safety score, such as whether more security cameras indicate a safer zone, or rather an unsafe zone that requires more cameras, and whether higher population density or lower income indicates that the zone is less safe. Overall, as Bat-el from the Hamal told us, it seems that our safety scores correspond to their understanding of the safety in the upper and lower areas in Hadar neighborhood. The main contribution of our dashboard to the Hamal's dashboards is the option to compare the state of different topics in different time frames. For each topic we gave both options - to see the absolute values (such as amount of crime, amount of population etc.), or to see the percentage of change from the previous time frame. This is an element that we identify as missing in the Hamal current dashboards and visualizations.

As a further work, we would investigate further the safety aspect and collect more data from the citizens in the neighborhood in order to understand their 'safety feeling' regarding different areas and topics, and combine this into the safety score. We also wanted to add data regarding the street lights in Hadar, and we found out that this data currently does not exist in the municipality, so we would expand this project to a mapping project regarding the street lights locations - this would help us to create different scores for safety during the day and during the night.

We hope that our dashboard will enable the Hamal to get interesting and important insights over Hadar and that they will use it in order to get even more familiar with Hadar.

Figure 4: The Hamal team using our dashboard in the Hadarion (9th August 2021, photo courtesy: Shay Sussman)

References

[1] Su, Kehua & Li, Jie & Fu, Hongbo. (2011). Smart City and the Applications. https://bit.ly/2VEXGoV

[2] Jing, C.; Du, M.; Li, S.; Liu, S. (2019) Geospatial Dashboards for Monitoring Smart City Performance. https://bit.ly/2UZxN2l

[3] McArdle, Gavin & Kitchin, Rob. (2016). THE DUBLIN DASHBOARD: DESIGN AND DEVELOPMENT OF A REAL-TIME ANALYTICAL URBAN DASHBOARD. https://bit.ly/2TiGILZ

[4] Roman A. Vila, Elsa Estevez, and Pablo R. Fillottrani. (2018). The design and use of dashboards for driving decision-making in the public sector. https://bit.ly/3kHpAL0

[5] Rick L. Bunch and Robert Earl Lloyd (2006). The Cognitive Load of Geographic Information https://bit.ly/3rngAwg

[6] Richard E. Mayer and Roxana Moreno (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. https://bit.ly/2UZ7utq

[7] Rim Razzouk, Valerie Shute (2012). What Is Design Thinking and Why Is It Important? https://bit.ly/36JTCpo

[8] Wim Bernasco and Henk Elffers (2009) Handbook of quantitative Criminology pp 699-724, Statistical Analysis of Spatial Crime Data https://bit.ly/2VQ8eBk

[9] André Ferreira, Guillermo Rubiano and Eduardo Mojica-Nava (2018). Urban Security Analysis in the City of Bogotá. https://bit.ly/3qxAbJD

[10] Lakshmi Iyer and Petia Topalova (2014) . Poverty and Crime: Evidence from Rainfall and Trade Shocks in India. https://bit.ly/2VTsEJH

[11] Dalit Shach-Pinsly (2018). Measuring security in the built environment: Evaluating urban vulnerability in a human-scale urban form. https://bit.ly/3wTiMMZ

[12] I. Stančin and A. Jović (2019). An overview and comparison of free Python libraries for data mining and big data analysis. https://bit.ly/2UlDdoK

[13] Sydney Brougham, Jonathan P. Leidig and Greg Wolffe (2019). Interactive Visualization for Real-Time Geospatial Data with Plotly Dash. https://bit.ly/3xT7NV0

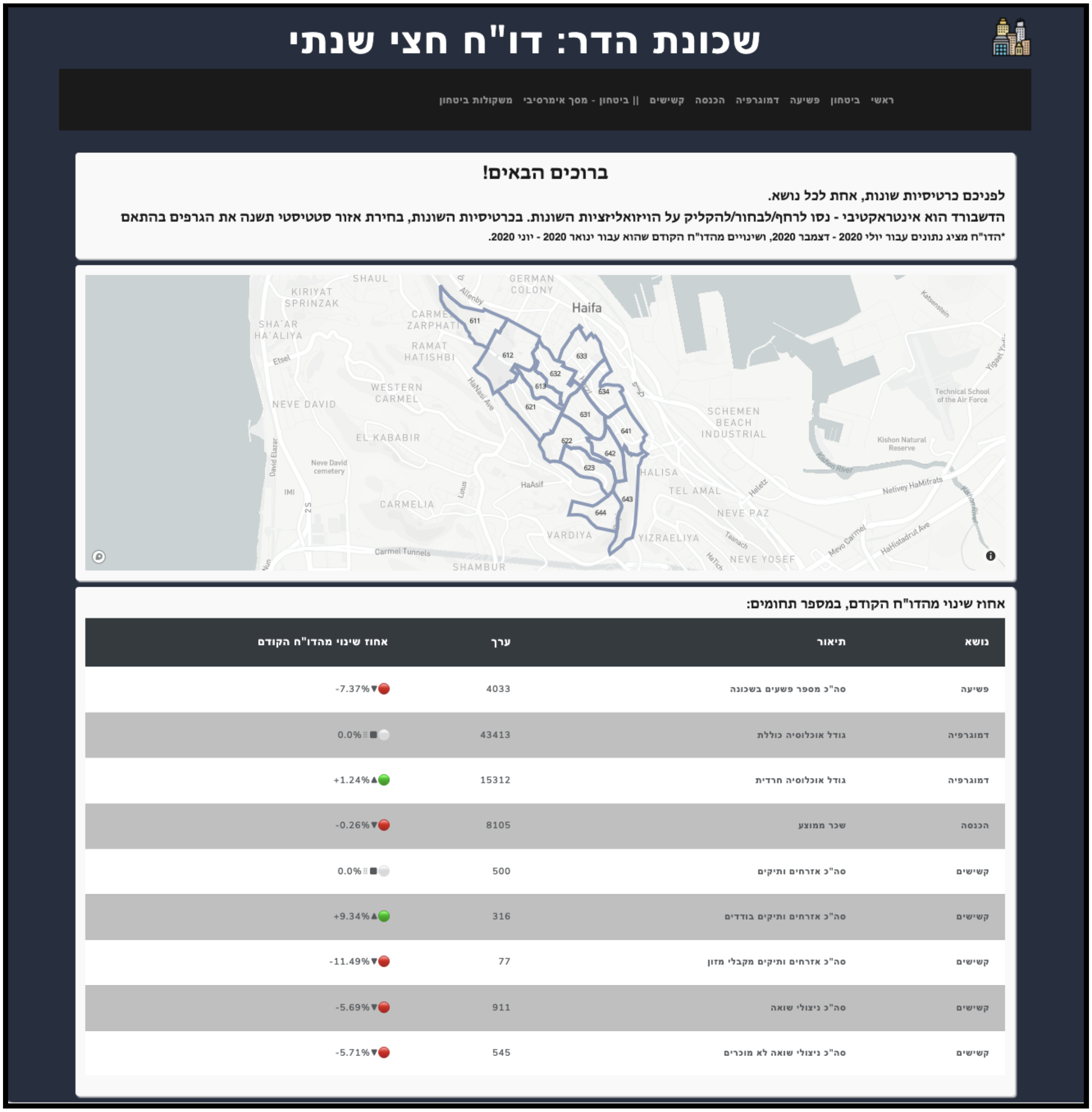
Appendix

Appendix 1 - Required input for the dashboard creation (reproducibility):
The required input to create the dashboard is a directory named "data" file with the following .csv files:

In all files _t0 represents the previous half year and _t1 represents the current half year.

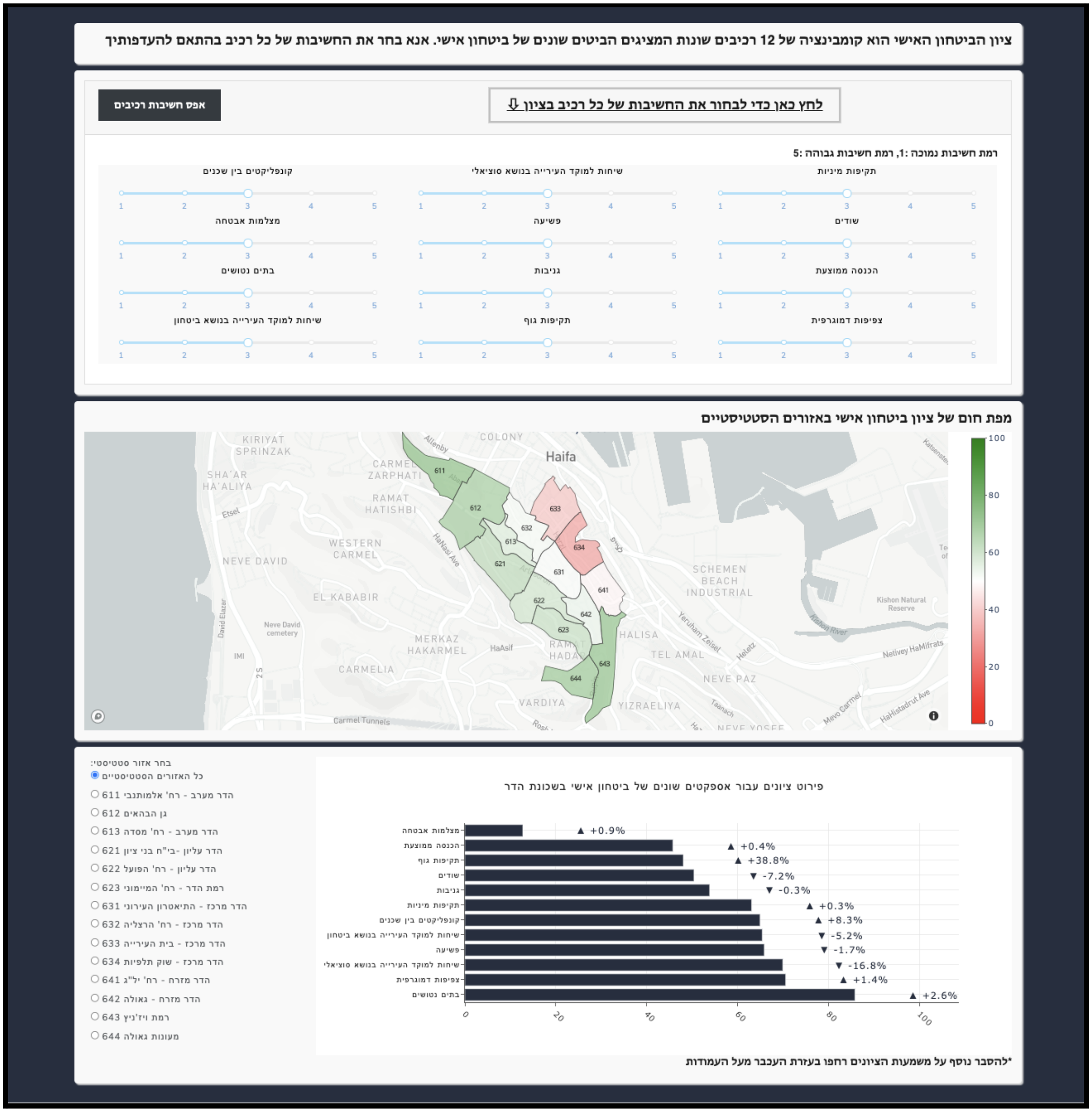
Tables names	**Columns needed (some from Hamal Excel file)**	**Description**
df_holocaust_t0.csv df_holocaust_t1.csv	'StatZone', 'HoloSurvKnwn', 'HoloSurvNdDesc',	Non-aggregative table with details about holocaust survivors.
df_seniors_t0.csv df_seniors_t1.csv	'StatZone', 'SeniorAlone', 'SeniorRecivFood'	Non-aggregative table with details about seniors citizens.
df_haredim_t0.csv df_haredim_t1.csv	'StatZone', 'TotHaredim', 'TotNonHaredim', 'Total', 'PropoHaredim'	Aggregative table with data about Haredi population data, per statistical area.
df_salaries_t0.csv df_salaries_t1.csv	'StatZone', 'ResNum', 'Women', 'Men', 'Children', '18_34', '35_44', '45_54', '55_64', '65_74', '75_84', '85+', 'SalNoHKResNum', 'SalHKResNum', 'SalPenResNum', 'SalSHNoBTLResNum', 'IncSelfResNum', 'SalNoHKAve', 'SalHKAve', 'SalPenAve', 'SalSHNoBTLAve', 'IncSelfAve'	Aggregative table with salaries and demographics data, per statistical area.
df_crimes_cases_t0.csv df_crimes_cases_t1.csv	'StatZone', 'CrimeType', 'CrimeLocType', 'Month'	Non-aggregative table with details about crime events.
df_crime_t0.csv df_crime_t1.csv	'StatZone', 'TotCrimCrime', 'TotBusiBurg', 'TotApartBurg', 'CarTheft', 'TotPropCrime', 'Thefts', 'BodyAssaults', 'SexualAssaults', 'Robbery'	Aggregative table with data about crimes, per statistical area.
df_aband_t0.csv df_aband_t1.csv	'StatZone'	Non-aggregative table with a record for each abandoned house and it's statistical area.
df_conflicts_t0.csv df_conflicts_t1.csv	'StatZone', 'Conflicts' ** This table was created by us using aggregation on the neighbors conflicts data - contains a row per statistical area with the amount of conflicts.	Aggregative table with neighbors conflicts data, per statistical area.
df_106_t0.csv df_106_t1.csv	'StatZone', 'בעל המשימה'	Non aggregative table with data about inquiries to local municipality call center (106).
df_cameras_t0.csv df_cameras_t1.csv	'StatZone', 'lat', 'lon' **External data manually added by us.	Dataframe with cameras locations and it's statistical area.

[Appendix 2](#) - Main dashboard tab.

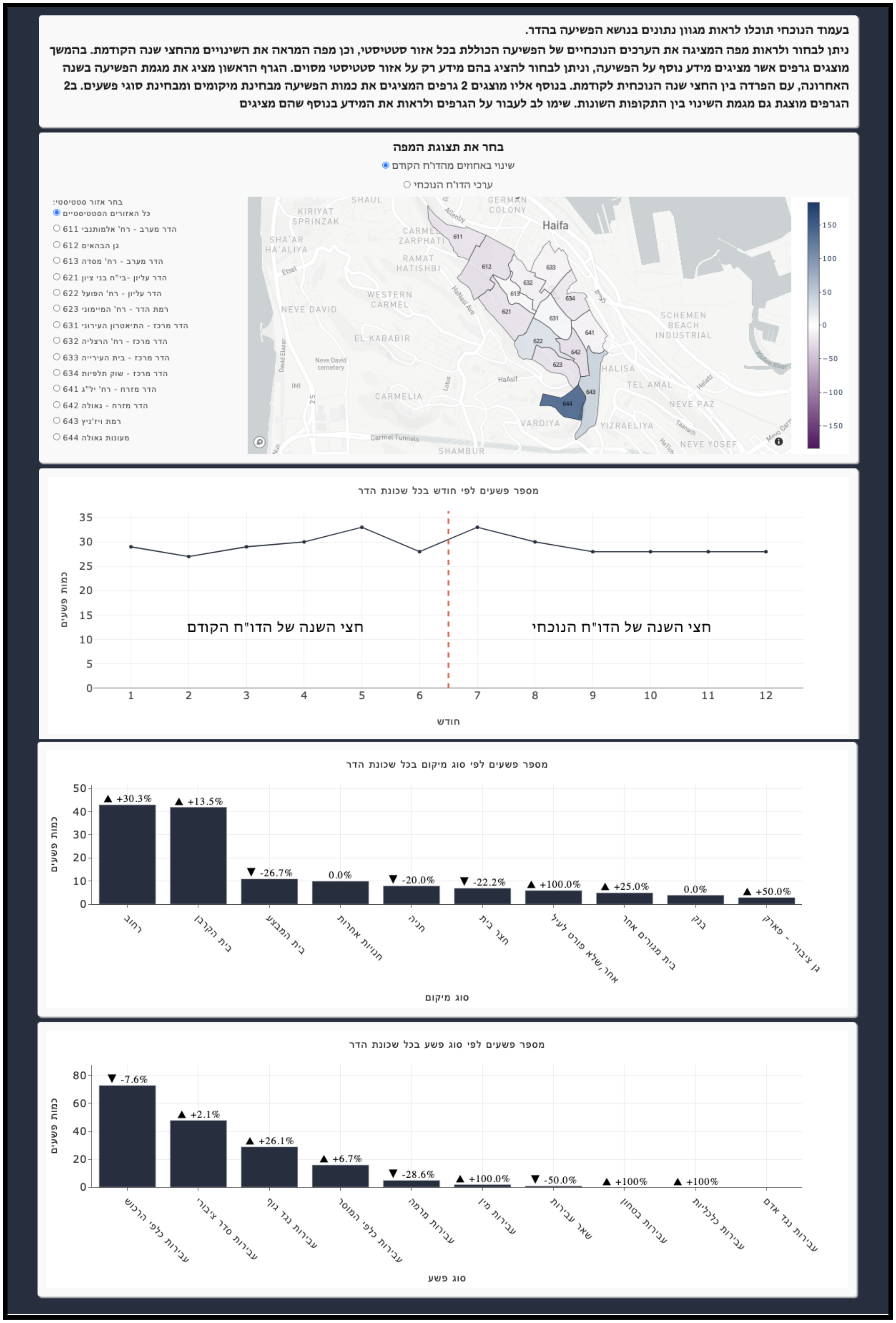


נושא	תיאור	ערך	אחוז שינוי מהדו"ח הקודם
פשיעה	סה"כ מספר פשעים בשכונה	4033	-7.37%
דמוגרפיה	גודל אוכלוסיה כוללת	43413	0.0%
דמוגרפיה	גודל אוכלוסיה חרדית	15312	+1.24%
הכנסה	שכר ממוצע	8105	-0.26%
קשישים	סה"כ אזרחים ותיקים	500	0.0%
קשישים	סה"כ אזרחים ותיקים בודדים	316	+9.34%
קשישים	סה"כ אזרחים ותיקים מקבלי מזון	77	-11.49%
קשישים	סה"כ ניצולי שואה	911	-5.69%
קשישים	סה"כ ניצולי שואה לא מוכרים	545	-5.71%

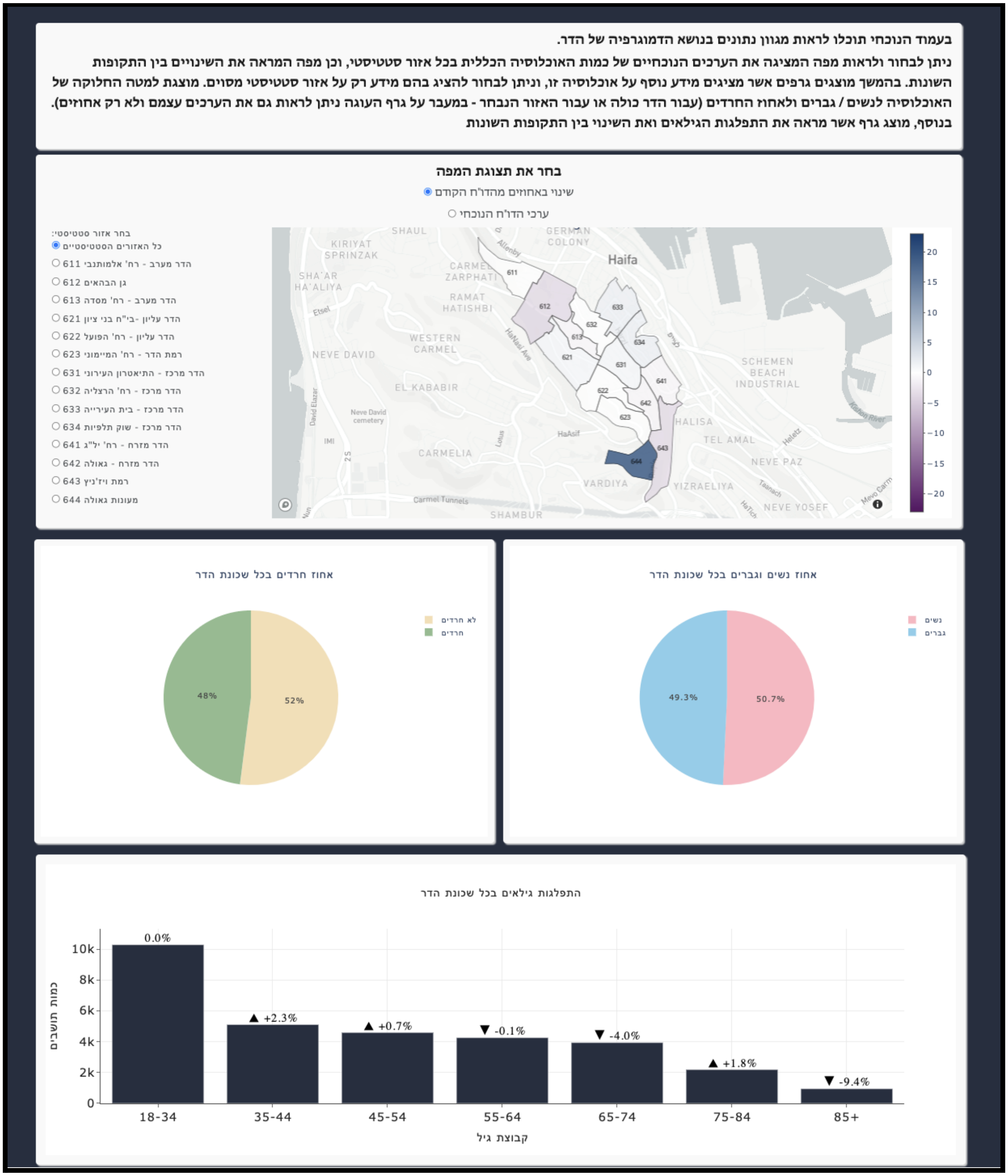
[Appendix 3](#) - Safety dashboard tab.



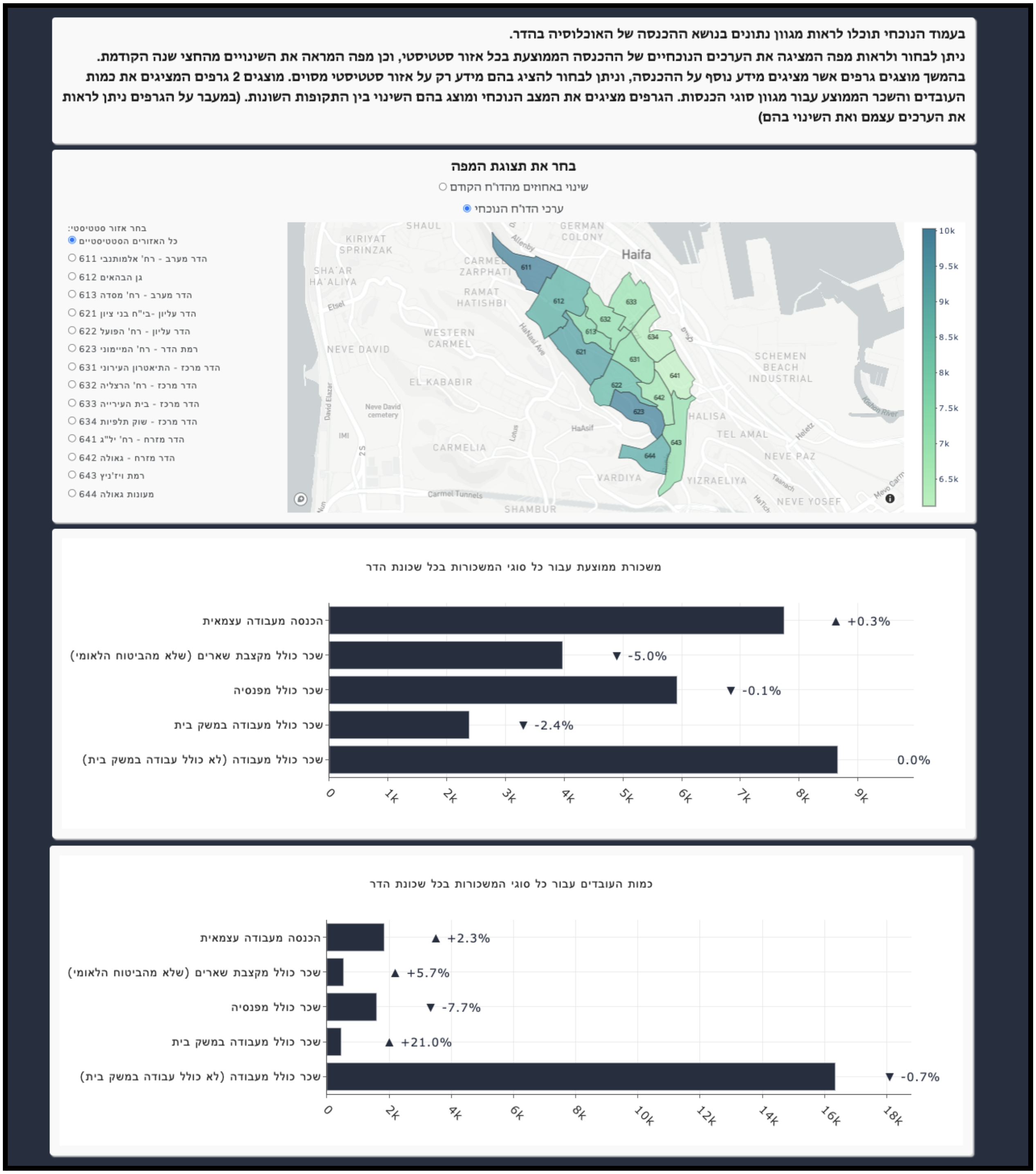
[Appendix 4](#) - Crime dashboard tab.



Appendix 5 - Population dashboard tab.



[Appendix 6](#) - Income dashboard tab.



[Appendix 7](#) - Elderly dashboard tab.