



Master's thesis
Master's Programme in Data Science

Interactive Time Series Visualization of Finnish Housing and Mortgage Markets for Central Banking Analysis

Markku Ylitalo

March 24, 2023

Supervisors: Professor Giulio Jacucci
Dr. Chen He

Examiners: Professor Giulio Jacucci
Dr. Chen He

UNIVERSITY OF HELSINKI

FACULTY OF SCIENCE

P. O. Box 68 (Pietari Kalmin katu 5)
00014 University of Helsinki

Tiedekunta — Fakultet — Faculty		Koulutusohjelma — Utbildningsprogram — Degree programme	
Faculty of Science		Master's Programme in Data Science	
Tekijä — Författare — Author			
Markku Ylitalo			
Työn nimi — Arbetets titel — Title			
Interactive Time Series Visualization of Finnish Housing and Mortgage Markets for Central Banking Analysis			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Master's thesis		March 24, 2023	81
Tiivistelmä — Referat — Abstract			
<p>This Master's thesis covers the visualization process of Finnish housing and mortgage markets by referring Tamara Munzner's nested visualization process model [25]. This work is implemented as an assignment for the Bank of Finland that is the national monetary authority and central bank of Finland. The thesis includes a literature survey in which the different stages of the visualization task are examined by referring them to previous studies, and an experimental part that describes the actual implementation steps of the visualization ensemble, that is an encompassing collection of interactive dashboard sheets regarding Finnish housing and mortgage markets, which was made as a supporting analysis tool for the economists of the Bank of Finland.</p> <p>The domain aspects of this visualization task are validated by arranging an end user survey for the economists of the Bank of Finland. Nearly a hundred open answers were collected and processed from which the fundamental guidelines of the desired end product were formed. By following these guidelines and leaning on the know-how of the previous studies, the concrete visualization task was completed successfully. According to the gathered feedback, the visualization ensemble managed to correspond the expectations of end users comprehensively, and to fulfill its essential purpose as a macroeconomic analysis tool laudably.</p> <p>ACM Computing Classification System (CCS): Human-centered computing → Visualization → Visualization techniques Human-centered computing → Visualization → Empirical studies in visualization</p>			
Avainsanat — Nyckelord — Keywords			
time series visualization, interactive visualization, dashboard, time series data, macroeconomic data			
Säilytyspaikka — Förvaringsställe — Where deposited			
Muita tietoja — Övriga uppgifter — Additional information			

Contents

1	Introduction	2
2	About the Assignment	4
2.1	Motivation	4
2.2	Assignment Description	5
2.3	Selected Time Series Data Sets	6
2.4	Research Questions and Methodologies	7
2.5	Preliminary End User Survey	8
3	Outlining the Visualization Process	13
3.1	Towards Visualization Design	13
3.2	Characteristics of Time Series Data Visualization	16
4	Time Series Visualization Techniques	20
4.1	Potential Techniques	20
4.1.1	Line Charts, Index Charts, and Column Charts	20
4.1.2	Small Multiples	22
4.1.3	Stacked Graphs	23
4.1.4	Ribbon Chart	24
4.1.5	Braided Graph	25
4.1.6	Slick Graph	26
4.1.7	Horizon Graph	27
4.1.8	Heat Map	28
4.1.9	Animations and Interactive Time Line	30
4.2	Interactions in Data Visualization	31
5	Implementing Interactive Time Series Visualization	34
5.1	Data Wrangling	34
5.1.1	Housing Market Data	35
5.1.2	Housing-related Loan Market Data	36

5.1.3	Domestic Indebtedness Data	37
5.1.4	Finnish Economy Barometer Data	37
5.2	Creating and Assembling Visualizations	38
5.2.1	Housing Market Visualization	40
5.2.2	Housing-related Loan Market Visualization	48
5.2.3	Domestic Indebtedness Visualization	52
5.2.4	Finnish Economy Barometer Visualization	52
6	Final Results	56
6.1	Interactive Housing and Mortgage Market Analysis Tool	56
6.2	Final End User Survey	62
7	Conclusions and Future Work	64
7.1	Conclusions	64
7.2	Future Work	67
	Bibliography	68
	Appendix A – Overview of the Final Analysis Tool	72

1. Introduction

Time dependent data analysis constitutes the fundamental basis of the finance and banking industry. Therefore, it is axiomatic that time series, which is a regular and chronologically ordered sequence of data points collected over an interval of time, is the most used data type in the finance and banking domain. Time series data is considered to consist of two main attributes, recorded data values and time stamps referring to the recording points of them [12]. Due to the formidable dependence between money and time, the whole industry tends to have a never-ending necessity to keep track on various economic trends and phenomena both now and over the past. Obviously, time series analysis is a very crucial topic for almost any other industry as well, since it enables to better understand the market and transaction processes. Overall, time series analysis is acknowledged as an effective method to interpret the present and forecast the future in order to make adequate, accurate and effective operational decisions in the right place at the right time.

In the finance and banking context, time series analysis is an essential tool for monitoring the changes of economic-related variables, and the associations of these changes with the other interesting phenomena over the same period. Time series visualizations have an important role supporting this crucial analysis work, since time series themselves are usually just chronological lists of values designated by individual time stamps, and therefore they are quite ponderous to examine without using another data abstraction to illustrate them visually. Data visualizations work as a visual link between a viewer and actual data, and they enhance perceiving, inter alia, changes, patterns, correlation as well as various phenomena relating to source data. In other words, visualizations are visual representations of recorded information, which fundamental purpose is to convert raw data into a form that makes its exploring and browsing more convenient for human eyes [21]. Visualizations also enhance decision making processes by presenting information behind data in a refined and more practical form, which reduces response time significantly [20]. According to the paper [4], to ensure that a visualization fulfills its purpose adequately and performs effectively, it is essential to pay attention to its abilities to maintain all necessary functionalities regardless of contextual

or viewpoint-related changes of use cases. In other words, flexibility is a very fundamental key factor of business, especially finance, related visualizations.

Even though the finance and banking sector utilizes time series visualizations, such as line and bar charts, for daily basis tasks, monitoring, and profound analysis, the range of techniques in use is very narrow; certain visualization techniques have established very deeply in the sector while the others are treated almost as generally acknowledged misfits for any possible task. Since there are already many handy and powerful visualization frameworks available, for example, Power BI and Tableau, it is a pity that many businesses and institutions are still not receptive enough to exploit the benefits of modern visualization techniques. As a counteraction, in this paper, I will introduce a broad business-related visualization project and its implementation process as an example of how to accomplish an extensive but still innovative visualization ensemble step by step.

This thesis considers an extensive collection of time series data visualizations that is implemented as an assignment for the the Bank of Finland. The assignment itself regards supporting Finnish housing market and housing related loan market analysis by building an interactive dashboard collection of data visualizations. In other words, I was requested to compose an interactive-data-visualization-based analysis tool that would support central banking related macroeconomic analysis work of the Bank of Finland's economists. The topic was very challenging, since there were massive amounts of data, the domain was very broad, and it required additional knowledge about central banking and macroeconomics.

The thesis is structured as follows: The Chapter 2 explicates the motivation behind the topic, describes and defines the thesis assignment, introduces available data sources, lists the research questions and methodologies of the thesis assignment, and covers the results of the preliminary end user survey. The Chapter 3 outlines the visualization process, especially from the aspect of time series data, by leaning on previous studies about the topic. The Chapter 4 introduces different visualization techniques for time series data and discusses about interactions in data visualizations. The Chapter 5 reviews the actual implementation stage of the end product that is a collection of interactive dashboard sheets, called also as an analysis tool. The Chapter 6 discusses about the final results of the interactive housing and mortgage market analysis tool and the final end user survey. Finally, the Chapter 7 summarizes the whole assignment, concludes the thesis, and outlines future work regarding both the thesis study and the analysis tool. The overview of the interactive housing and mortgage market analysis tool is displayed in the Appendix A section.

2. About the Assignment

This master's thesis is written as an assignment for the Financial Stability and Statistics department, especially for the Statistical Analysis and Information Services division, for who I have been working during 2021 and 2022. The principal ambition of this thesis is to implement an interactive data visualization of Finnish housing market and housing-related loan market for the economists of the Bank of Finland as an illustrative tool that supports their macroeconomic analysis. Due to the requirement of the Bank of Finland, the final visualization ensemble is created totally by using Microsoft's Power BI software. Power BI is an all-in-one business intelligence application by which different data sources can be queried and combined, data can be modified and parsed, and various visualizations can be build [33]. According to the paper [31], business intelligence consists of techniques, technologies and tools that are designed to enhance decision making processes by providing better access to essential business data by collecting, processing, refining, and visualizing it. According to the paper [6], Power BI has over 30 default visualization frameworks, such as adjustable line and bar chart layouts, while there is also an option to import many other custom visualizations, like heat maps and radius charts (see the Subsections 4.1.8 and 5.2.1), into Power BI. I was allowed to use both default and custom frameworks in this visualization project. Furthermore, Power BI includes many options to automatize data querying and updating processes so data can be retained up-to-date easily [33].

2.1 Motivation

Due to the harsh and business-paralyzing COVID-19 pandemic between 2020-2022, the European Central Bank resorted to reflation actions to compensate huge losses of Euro areal companies, especially the ones who have significant and broad impact to the whole Euro areal economy. Moreover, on account of massive reflation actions, the Euro areal inflation started to increase towards undesirable levels. Therefore, the European Central Bank has planned to increase interest rate noticeably in the near future. However, these interest rate actions are currently, writing this during the spring 2022, frozen, since the disturbing war between Ukraine and Russia started on the 24th of February 2022.

The interest rate policy of the European Central Bank would affect directly on housing markets, because the interest rates of mortgages consist of the margin rates of commercial banks and reference rates which are based on the interest rate of the European Central Bank. At the same time, the housing prices of Turku, Tampere, and especially Helsinki Metropolitan area have continued to increase so high that the housing costs of the biggest Finnish cities are starting to be unbearable for the people having an average income.

There have been a lot of uncertainty in the air during the spring 2022 why housing markets have gathered more attention than usual. Obviously, the increasing costs of housing-related loans have crucial impact for many Finnish citizen, which moreover would scale up over-indebtedness and the housing shortage of affordable dwellings. Due to the current situation, the Bank of Finland had a pressure to implement more comprehensive and illustrative analysis tools about the housing markets in Finland for BOF's internal use.

2.2 Assignment Description

The assignment of my master's thesis is to implement an interactive data visualization of Finnish housing market and housing-related loan market for the Bank of Finland as an illustrative tool that supports central banking related macroeconomic analysis. The end users of this analysis tool are the economists of the Bank of Finland. The visualization is also required to be updatable. Thus it suits either to utilize on a daily basis or for long term monitoring.

The priorities (p) of this assignment, sorted from the most important to less important, are the followings: (p_1) provide an informative and handy analysis tool for the economists of the Bank of Finland, (p_2) compress a lot of time series data concerning housing market and housing-related loan market into a single presentation, and (p_3) introduce new and modern ways to visualize time series data, especially housing market and housing-related loan market data.

There are also a couple of special interests (i) concerning the different subfields of the topic. The Bank of Finland desires to have the following aspects included into the analysis tool: (i_1) the visualization of housing-related loans as a part of housing markets, (i_2) the analysis concerning the interest rates of new drawdowns, (i_3) the dwelling price analysis, and (i_4) the buy-to-let mortgage analysis.

2.3 Selected Time Series Data Sets

The following table itemizes all the separate data sets that were available for me to use in this visualization project of housing and mortgage market related analysis tool.

No.	Data Set	Source	Type
1	Housing-related loan drawdowns	the Bank of Finland	continuous
2	Buy-to-let mortgages	the Bank of Finland	continuous
3	Buy-to-let mortgage drawdowns	the Bank of Finland	continuous
4	Housing-related loans in total	the Bank of Finland	continuous
5	Change of housing-related loans	the Bank of Finland	continuous
6	Consumers' confidence	the Bank of Finland	categorical
7	Consumers' expectations	the Bank of Finland	categorical
8	Consumers' views	the Bank of Finland	categorical
9	Consumers' intentions	the Bank of Finland	categorical
10	Interest rates	the Bank of Finland	continuous
11	Interest rates of new drawdowns	the Bank of Finland	continuous
12	Maturities	the Bank of Finland	continuous
13	Housing corporation loans	the Bank of Finland	continuous
14	Demand of housing-related loans	the Bank of Finland	continuous
15	Change of housing-related loan demand	the Bank of Finland	continuous
16	Amount of new housing-related loans	the Bank of Finland	continuous
17	Household indebtedness	the Bank of Finland	continuous
18	Indebtedness relative to income	the Bank of Finland	continuous
19	Prices of dwellings	the Bank of Finland	continuous
20	Changes of dwelling prices	the Bank of Finland	continuous
21	Dwelling prices relative to rents	the Bank of Finland	continuous
22	Sales of dwellings	the Bank of Finland	continuous
23	Dwelling sales by dwelling types	the Bank of Finland	continuous
24	Value of dwelling sales	Statistics Finland	continuous
25	Housing-related loans by regions	Statistics Finland	continuous

Table 2.1: Selected time series data sets

The data I used in this thesis is collected from the databases of the Bank of Finland and the web tables of Statistics Finland (Tilastokeskus, in Finnish). All this data is open for a public use, and can be found either from the websites of the Bank of Finland or from the websites of Statistics Finland. However, the methods and tool by which data were queried from the databases of the Bank of Finland are confidential and therefore omitted from this thesis. The data sets listed above in the Table 2.1 are the time series, desired by the Bank of Finland, that I was required to include into the final analysis tool. All but consumer barometer time series data sets consist of continuous data. Consumer barometer time series data sets include categorical data.

2.4 Research Questions and Methodologies

The particular research questions I focus on through my master's thesis:

(Q_1) How to provide an informative and handy analysis tool for the economists of the Bank of Finland using interactive time series visualization techniques?

(Q_2) What are the best ways to compress a lot of time series data concerning housing market and housing-related loan market into a single presentation?

(Q_3) What would be new and modern ways to visualize time series data, especially housing market and housing-related loan market data, which the Bank of Finland would consider to exploit broadly in future?

The methodologies I utilize to answer all my research questions:

(M_1) First, I am going to arrange a preliminary end user survey to examine the expectations and desires of the economists of the Bank of Finland. This would help me to implement the housing and mortgage market visualization, that is the final product of my master's thesis, to correspond with its actual purpose as a supportive analysis tool.

(M_2) Second, I will explore and study the previous work concerning my topic, that is interactive time series visualization, and the results of the preliminary end user survey to discover the most suitable visualization techniques for the final analysis tool.

(M_3) Third, I am going to create a couple of prototype visualizations, which I will introduce to the end users of the final analysis tool to gather feedback and opinions about the visualizations and different concepts related to them.

(M_4) After that, I will refine the best prototypes following the collected feedback and the previous studies of my topic.

(M_5) When the final product is fully implemented, I am going to arrange the final end user survey to examine how the analysis tool fulfills its purpose, and whether there any need for future work around it.

2.5 Preliminary End User Survey

According to the paper [19], one of the greatest challenge in the implementation process of a visualization is to understand end users. In an optimal situation, well-build visualization works as an interface between information resources and human capital [19]. Visualizations are meant to function as a tool that enables end users to perceive unintuitive patterns and relations behind data, search explanations for real life phenomena and forecast the future [19]. Based on the study [3] by W. Aigner et al., to successfully match the information contents of a data visualization with the actual needs of the end users, it is necessary to take an event specification step at the very beginning of the visualization process. During the event specification, the end users are requested to formally define their necessities and desires towards the final visualization output. Therefore, also for this visualization project, it is well-reasoned to inquire the end user interests and wishes before starting the actual encoding of the visualizations.

Before proceeding to the practical implementation process, I arranged a preliminary end user survey to examine the initial expectations and desires of the economists of the Bank of Finland regarding the interactive visualization of Finnish housing and mortgage market. These survey results would help me to build the visualization product so that it corresponds to its actual purpose as a supportive macro-economical analysis tool adequately.

The survey consisted of the following open-ended questions:

(1) From your point of view, what would be the most fundamental purpose of the interactive visualizations of the housing and mortgage market? (2) What kind of information you would desire to find from the interactive visualizations of the housing and mortgage market? (3) What additional value these visualizations would offer for you in comparison to plain and tabled statistical data? (4) From your point of view, what are the pros and the cons of using interactive dashboard-based visualizations compared to non-interactive static visualizations? (5) Which interactions, such as

filters, hoverable tool tips and info windows, adjustable time line, selection-based calculations (for example, average, mean, standard deviation), etc. you prefer in dashboards? (6) Which features (such as colors, symbols, curves, line lengths, area sizes, angles, additional text and numbers, etc.) help you to perceive the main points of the visualization most quickly and comprehensively? (7) Which features help you to perceive the details of the visualization most quickly and accurately? (8) How frequently you would use the interactive housing and mortgage market visualizations to support your work? (9) From your point of view, what granularity levels (months, quarters, years, etc.) would be the most relevant to include in the visualizations? (10) From your point of view, are there any other design aspects which should be taken into account?

In this paper, I refer to dashboards as sets of several figures (visualizations), info texts, filtering menus, and other interactive gadgets that collect interrelated information of specific macroeconomic category together, and therefore help the end users to examine and compare data from different sources at the same location without need to browse between separate views [31]. Thus, interactive dashboards are handy data monitoring and analysis tools that should provide all users easy-access to explore, filter and analyze data without almost any previous practice [34]. **The results of the survey:**

The end user survey managed to reach seven economists of the Bank of Finland, who work regularly with housing and mortgage market statistics. The following thoughts, opinions, and desires are gathered and condensed from the answers of the economists of the Bank of Finland, which are the end users of this housing and mortgage market visualization, by the preliminary end user survey introduced above. The results of the survey are itemized by these preliminary end user survey questions from (1) to (10).

(1) According to the end users, the most fundamental purposes of the interactive visualizations of the housing and mortgage market would be to deliver and prepare insightful material related to housing and mortgage markets to support macro-economical analysis, to get a quick overview of the main market trends, and to display multiple data sets, collected from various sources, simultaneously via a single interface. These interactive visualizations may also help to produce a general view of how the housing market and the mortgage market are related over time, in order to determine how they would develop in the future. Moreover, these interactive visualizations support the analysis of the reasons behind the housing loan (mortgage) demand and household finances.

(2) According to the answers, the end users desire to find the following information from the interactive visualizations of the housing and mortgage market: (I) Statistical information and calculations concerning especially mortgages and housing-related loans, such as housing loan stock development, new housing loan drawdowns, volumes, totals, relative changes over time, margins, interest rates, repayment periods, and other credit standards divided, for example, by their geographic location, purpose of the loan, and maturity. (II) Housing market -related information, like developments in housing sales, both new build and existing (old) housing, variables relating to building (permits/starts/finished), the regional development of dwelling prices, both nominal and real, and a division of dwelling type and geographic location. (III) Barometer information concerning the expectations of both households and banks on mortgage market, for example, how much consumers have plans to buy a dwelling or build a house in the next 12 months.

(3) The end users consider that visualizations would offer the following additional value in comparison to plain and tabled statistical data: Visualizations provide a possibility to inspect and compare multiple data sets from different sources and dimensions quickly and diversely via a single interface, and they enable to perceive correlations and patterns more easily in comparison to plain, tabled, and non-visualized forms of data. In the housing and mortgage market context, visualizations would be helpful for quickly updating one's view on the main developments in the market, instead of having to dig through various statistical files in order to gather the same information. Thus, visualizations bring new and more comprehensive viewpoints towards inspected data and reduce time spend on searching essential information, which support financial analysis significantly. It is also easier to put data into a context, and regard dynamic relationships of different attributes when data is visualized. On the other hand, plain and table-based display format would present statistical data in a more flexible and workable form for excluded additional calculations that may be needed occasionally. However, visualizations would provide a functional overview very efficiently and make the data more interesting also for non-experts, and therefore ease the introduction process of the phenomena being associated with the data.

(4) The pros and the cons of using dashboard-based visualizations compared to non-interactive visualizations according to the end users: (*Pros*) In interactive visualizations, data is updated automatically and it is already preprocessed in a modular format that allows to display instantly diverse visualisations based on different variable combinations selected by users. Interactive visualizations may also include a lot of additional information, like characteristics, calculations, and aggregations, more compactly and user-friendly than stationary data tables. Moreover, interactions can be

utilized to enable quick and effortless alteration between various data sets collected from different sources, and they enable users to manipulate displayed visualization layout, timeline, zooming-level, or perspective. (*Cons*) Metadata and specific values may be more difficult to access directly or they are totally unavailable when they are hidden behind a user interface. Therefore, dashboards may not allow users to produce their own calculations on the underlying data. In addition, the graphs might not be utilized anywhere outside their source dashboard due to some software related incongruity issues or their unsuitable visual appearance. Also, missing data values may cause serious troubles to display data, if an automatic calculation and visualization pipeline does not endure them. Moreover, available data may depend too much on the choices of who made the interactive visualization in the first place, and not necessarily so much on the choices of a final user. In addition, it is not always possible to repeat previous analysis template in the exactly same manner than it was formerly build up if all previous user-based selections of a template are not known, or data is updated automatically over time. In comparison, fixed and stationary cross-section analysis templates offer always and independently the same results and charts despite previous users or inspection times.

(5) According to the answers, the end users prefer the following interactions in dashboards: adjustable time lines, a selection menu for different indicators and calculations, such as the average lengths of the repayment periods of mortgages, diverse filtering options, overall flexibility to mix and match different data sets as user sees fit, a possibility to export the data being associated with the visualizations, user instructions concerning the interactions of the visualizations, hoverable detail windows consisting of exact values and time stamps,

(6) According to the answers of the end users, the following features help them to perceive the main points of the visualization most quickly and comprehensively: distinguishable colors, different color hues in heat maps, and a low component and text density in a single graph. There was also asked for to take color-blind people into account when choosing the color scheme for the graphs.

(7) Indicated by the answers, the following features help the end users to perceive the details of the visualization most quickly and accurately: suitable graph type depending on a data type and phenomenon, constant visualization layout elements, for example, using always the same colors and shapes to represent growth and decline, and hoverable detail windows concerning values and their recording times.

(8) According to the answers, the most of the end users would use the interactive housing and mortgage market visualizations periodically on a monthly basis but some users might utilize it even on a daily basis to support their analysis work.

(9) From the end users' point of view, the following granularity levels would be the most relevant to include in the visualizations: month, quarter, and year frequency, since the most macroeconomic housing and mortgage market data in Finland are available comprehensive enough at the monthly frequency only.

(10) In addition, the end users propounded that the following design aspects should be also taken into account: Fonts should be suitably sized, there should be an ability to change the language of the user interface, graphs should be able to be presented independently without their whole source dashboard, the dashboard environment should enable the access to its metadata, the visualization ensemble should include as many dashboard sheets as necessary to organize information reasonably and logically, that is to say, preference is to avoid to crowd too much content in a single sheet. Moreover, the end users desire that the housing and mortgage market dashboard would be able to be updated, edited, and extended also afterwards.

3. Outlining the Visualization Process

In the book *Information Visualization* [8] written by S. Card, there are provided the comprehensive summary of the principal purposes of the data visualizations. According to the book, visualizations have significant effect for the viewers compared to the situation where there exist raw data or plain tables only. The major cognitive benefits of the visualizations for their users are the followings: visualizations decrease the time spent for searching and perceiving information, expand the capabilities to use the memory capacity more efficiently for analysing information, function as an external process memory, improve the ability of recognizing patterns from data, and provide a possibility to receive diverse and individual information by perceptual examining, reasoning, and reprocessing.

In addition, S. Card states his the book that visualizations enable to explore colossal data sets by sight. Visualizations are also very advantageous for tracking long term events, since they help to monitor multiple aspects simultaneously, and perceive both the details and the big picture of the data alongside. Moreover, visualizations work as abstractions of the real life phenomenon by reducing the information of the chaotic, complex and broad entities.

3.1 Towards Visualization Design

Through this housing and mortgage market visualization project for the Bank of Finland, I exploit the nested visualization process model [25] presented by Tamara Munzner in 2009. The model is established as an advantageous tool for any visualization design and evaluation schemes, since it simplifies the whole implementation process as a stepped fundamental procedure. In other words, the model helps to split visualization design cycle separately into nested layers. Each of these layers include designated evaluation methodologies, which are based on the validation error possibilities appearing at each specific level.

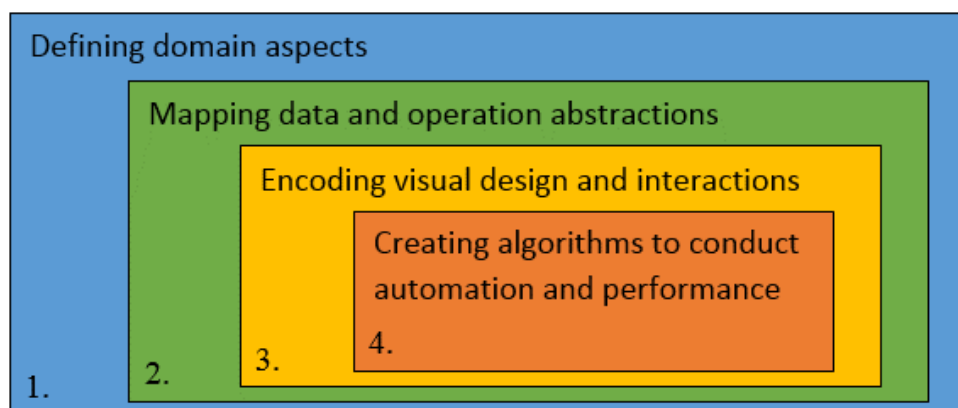


Figure 3.1: The four layers of Munzner’s nested visualization process model, in which proceeding happens step by step towards the inner parts of the model [25].

This nested model of Munzner consists of four layers [25]. The first and the outermost layer concerns characterization of the domain aspects and data. In this paper, Sections 2.1, 2.2, and 2.3 cover the contents of this layer by defining our domain and data aspects regarding the housing and mortgage market visualization project. The second layer, inside the first, covers the mapping processes of the domain aspects of the previous layer. In this layer the domain aspects are transformed into exact and abstract level operations, whereas the data is parsed to correspond proper data types. Section 2.4 covers this abstraction of the domain aspects, whereas Section 5.1 covers the data-related modifications, aggregations, and other fundamental operations that underlie the base visualization functionalities such as filtering and sorting. The third layer, inside the second, consists of all actual design stages of the visualization project that concretizes the abstract level operations defined at the previous layer, including visual encoding and building interactive user interface. Regarding our housing and mortgage market visualization, this process is presented entirely in the Section 5.2. And the fourth, the innermost layer, concerns the automation and performance aspects of the visualization, which in our case are already managed by Microsoft’s Power BI application and its linkages between the databases of the Bank of Finland.

The inner layers inside the first and the outermost layer, which mainly focuses on domain definitions and data sources, consist of layer-specific instances of various design issues which would be reasonable to take into account during any visualization project. The nested structure of Munzner’s process model indicates the relations between these different layers [25]. Therefore, it is important to consider possible validation flaws of each layer carefully before stepping into the next layer. If some imprudent and illogical design choices are made earlier, they will be passed down from the

outer layer to all inner ones. That is to say, the inner layers use every design decision made in the previous layers as an input which defines de facto a base for all further design.

Moreover, any heritable design flaws arised in the outer layers are infeasible to fix or compensate by any following design choices of the inner layers. Thus, Munzner's model considers all visualization tasks as iterative processes, since changing any design aspects of the outer layers forces to retake every step from the modified layer back to the current one [25].

The first validation step of this visualization project is taken in the Section 2.5. This validation step ensures that the domain aspects are defined comprehensively and the actual assignment is clearly and entirely understood. The first validation step is very crucial for the whole visualization project, because all time and effort being spent into the assignment might turn out to be redundant and pointless if the final visualizations provide answers only to the wrong and unnecessary domain questions due to the some kind of misunderstanding.

The second validation step is taken in the Sections 2.5 and 5.1. The main possible flaws of the second layer, that is the mapping data and operation abstractions, are selecting invalid data types and operations in respect of a data model's robustness. To ensure that this validation step is taken adequately into account, nearly a hundred open answers of the preliminary end user survey were carefully processed and considered when selecting suitable data types and operations for the used data model. For example, filtering, sorting and data importing operations were embedded into the end product by the requests of the end users.

The third validation step is taken in the Chapter 4 and in the Section 5.2. The main threat of the third layer, that is the encoding visual design and interactions is to end up using ineffective visualization and interaction techniques that do not transfer information in a clear, unbiased, and user-friendly manner. The validation of this step is ensured by justifying the selected visualization and interaction techniques regarding the actual use cases of the end product. It is important to be sure how well desired information can be received and perceived via used techniques.

Since the automation and performance aspects of the visualization are already managed by Microsoft's Power BI application, the validation step of the fourth layer is omitted from this thesis. The above mentioned validation steps represented upstream validation stage, because these steps were taken before the next inner layers of the nested

model were reached. After the whole implementation process of the end product was completed, the final validation phase was carried out. This validation phase represents the downstream validation. The final validation phase consist of the final end user survey (see the Section 6.2) and its results as well as the feedback collected from the end-user-targeted presentations (see the Chapter 7) of the end product.

3.2 Characteristics of Time Series Data Visualization

According to article [15] by J. Heer et al., time series are an universal type of recorded data, since every real life phenomena tend always to have an chronological time context. Thus, there are an indefinite amount of fields, such as banking and experimental physics, among which time series data is either a relevant or necessary form for collected information. Since real life phenomena are typically very complex and multi-dimensional ensembles, those could be reduced into many time series partitions, each concerning a different aspect of the phenomenon. By utilizing different visualization techniques, such amounts of time series could be represented in order that they are considerably more reasonable to examine by human eye.

In the article *Visual methods for analyzing time-oriented data* [3] W. Aigner et al. state that time series data tend to be really extensive ensembles, since they commonly consist of both numerous data elements and a large assortment of monitored attributes. Therefore, utilizing only simple visualization approaches may drive time series data sets into cluttered and over-stuffed visualization arrangements when these approaches are not able to utilize the scarce space efficiently enough to avoid undesired intersecting or an overall confusion. Instead of using always trivial visualization methods, analytically build data abstractions might provide adequate solutions to display massive data sets as illustrative and uncluttered visualizations.

In keeping with the stance above, the study [17] by Jabbari et al. suggest altogether three general solutions, as an extension to the previous proposition of W. Aigner et al., for reducing overlapping and clutter in the visualizations of time series data. The first solution, the most intuitive one, is to decrease the number of the visualized data points. This decreasing action can be accomplished by distributing either data points, attributes or both into several distinct visualizations instead of a single one. The second is to aggregate data into a form that is more space efficient but still maintains the core information of the data. However, the first solution hampers the comparison between the separated data points and attributes, whereas the second option bears inevitable information loss. Continuing the list, the third, and the final, option is to utilize more

advanced visualization techniques which are designed to reduce the usage of space per a visualized data point, attribute or both. Nevertheless, the third option might easily end up with the awkward, impractical and over-complex solutions which have virtually lost their touch with the essential end users or usage circumstances of the visualization.

Many data visualization related studies have achieved similar conclusions concerning that how the interpretations of the same phenomenon may vary a lot depending used visualization techniques. For example, A. Gogolou et al. managed to provide an evidence, in consequence of their research [13], about how the perceptions of similarities between distinct time series has a significant dependence on the visualization approaches used to present them. The study states that different visualization techniques impact differently people's recognition viewpoints of pattern similarities. Therefore, while choosing a visualization approach for time series data, it is important to consider which aspects between the visualized time series are relevant to perceive having any possible similarities in order to maintain viewers confidence.

The rest of this chapter focuses on the comprehensive study of conceptual visual analytic frameworks for time series and time-oriented data [1] by W. Aigner et al., which describes laudably the most essential principles of the time series data visualization process. According to the study, the all-encompassing analysis framework build on the time series related visualizations includes three main components that define the most of its data characteristics and analytical purposes: an abstract level model of time and the actual data values, a model-integrated computational analysis of the data, and interactions of the user interface construed over the first two components. We will focus on the actual data related components later in the Chapters 4 and 5.

In accordance with the study [1], while modeling the abstraction of time, there are five essences to take into account: scale, scope, structure, viewpoint, and granularity. Each of these essences consists of several subcategories, which altogether define the time abstraction of the visualization. Since all the essences but the granularity of time are pretty straight forward to realise by looking the Table 3.1, we focus only on the concept of granularity. In a time series context, granularity means the partition ratio of the mapped time abstraction. For example, the most commonly used granularities for the mid-term macroeconomic monitoring among the finance industry are days, months, quarters and years. In other words, granularities are the widely-agreed conceptual units to illustrate the different events on a human-invented time line. Due to the nested relations of different time units, by compiling several granularities together, we are able to generate hierarchies, that are also known as calendar structures.

Table 3.1: The alternative composition of a time abstraction

No.	Essence	Types
1	Scale	ordinal, discrete, continuous
2	Scope	points, intervals
3	Structure	linear, cyclic
4	Viewpoint	ordered, branching, diverse
5	Granularity	none, single, multiple (hierarchical)

According to the study [1], The visualization process of time series data is recommended to start by visualizing time itself. To do that, the imaginary dimension of time has to be mapped into a visible form. The two essential mappings of time are (1) converting time to be expressible by some spatial attributes, like length and position, and (2) mapping time into its own physical dimension, and optionally scaling the original time line to better perceptible rate. It is pretty safe to say that, almost every stationary type of visualization ever made exploits some of the spatial attribute mappings considering the abstract nature of time. Instead of utilizing length and position only, time can be also expressed, for example, with angle, area, volume, color, transparency or text. Also the layout of the mapping can vary a lot being almost anything, such as lines, columns, rows, circles and spirals. The second introduced mapping type concerns the physical nature of time is utilized exclusively among animated time series visualizations in which time is represented in a sensate form (see the Section 4.1.9). However, it is also possible to create hybrid mappings by combining these two fundamental ones. In that case time is represented by both the abstract dimension of time using spatial attributes and the physical dimension of time utilizing animations.

Obviously, these essential mappings of time have their own functions and objectives. The stationary mappings of time favour more detail-oriented examination and comparison of the specific parts and points of the visualized time series data, whereas animated visualizations are more suitable to present the changes and the overall development of the bigger picture. Nonetheless, both of these mappings, as well as their hybrid combination, equally support the implementation of the time related interactions, such as zooming time line, switching over different granularity levels and adjusting the frame rate of the animation. The interactions are very important part of the time series visualization, because they enable an individual and user-oriented navigation, adjustments, data mining, and analysis of the time series data. Therefore, it is crucial to design the interactions to directly correspond with the needs and desires of the essential

end users. The same applies with the choice of the encoding techniques of the actual time series data visualizations. We will focus more on both, the interactions and on the comparison of the visualization techniques in the Chapter 4.

4. Time Series Visualization Techniques

According to the paper [16] by J. Heer et al., the graphical perception of visual encodings, which concerns the viewers' general capability to perceive valid information from a visualization, is quite broadly studied subfield among the data visualization research overall. Typically, the graphical perception related studies consider which visual aspects such as different colors, spatial measurements, and arrangements of certain elements are the most effective to illustrate the relevant information of visually encoded data. For example, it is discovered that the position of the visual variables improve viewers' capability to perceive information the most among all types of data: categorical, ordinal, and quantitative. It is also noticed that the coloration of the visual variables performs laudably with categorical data but not so great with the quantitative type of data.

4.1 Potential Techniques

In this section, I will itemize multiple potential time series visualization techniques. Since one of my research questions (see Section 2.4, question Q_4) considers the desire of the Bank of Finland to seek new and effective ways to visualize time series data, especially housing market and housing-related loan market data, I will emphasize more the modern and innovative visualization options than simple line, index and column charts in this visualization project. Along with line, index and column charts, I will introduce nine other potential visualization alternatives for time series data.

4.1.1 Line Charts, Index Charts, and Column Charts

Even though I will try to favour more modern and innovative visualization techniques over the most common ones in this thesis, I begin this itemization process by defining the main features of the line, column and index charts, since they seem to be very popular visualization techniques with time series data. It seems that especially line charts have established their place in this world permanently. For instance, line charts are the most

used visualization techniques in the finance context like they have been ever since first data visualizations were invented [35]. According to the study [24], time series related line charts have been actively utilized in science after they were discovered by Joseph Lambert in 1779. The paper [16] describes laudably the main characteristics of the line charts: Line charts utilize spatial position to illustrate both time and the actual data points in a suitable coordinate system. Usually, horizontal axis represents time and vertical axis represents the values. In this case, time is almost always encoded to proceed from the origin to right, whereas the values increase upwards and decrease downwards. In line charts, the chronologically successive data points are connected with lines, the slopes of which illustrate also the time related changes between distinct data points. Since single lines do not require much areal space, several time series can be presented simultaneously by means of overlapping lines in the same picture and within the same axes. However, if too many time series are tried to fit in the same picture or the overlapping time series resemble too accurately each others, the visualization will end up being cluttered and uneasy to follow [35]. Moreover, the most of the characteristics of the line charts apply also with the index charts. Nonetheless, the index charts are used to illustrate relative changes between the selected time series data points [15]. Index charts have many applications in the finance industry, especially around the stock market.

Column charts, also known as bar charts, are practically like line charts apart from a few differing characteristics of them. Column charts are more reasonable choice to use for distinctly discrete time series data, because there is no slope between adjacent bars to represent their time related changes. Multiple parallel column charts are also more complex to arrange into a single display, since each visible data point requires a colored pillar under it. However, column charts are argued to be more suitable to use while stacked in comparison to line charts, since human eyes struggle to perceive vertical differences correctly if the comparative layers are not both horizontally arranged (see the Subsection 4.1.3). Columns can be also clustered in specific category classes, for example, by using different colors, shapes or textures. According to the paper [10], the main advantage of clustered column charts is that they enable to examine and compare multiple time series of the same category group easily together. Nevertheless, clustered column charts may start to look cluttered as the number of parallel time series increase in a single figure [10].

Even though line, index, and column charts may seem to be ineffective techniques to visualize more complex and multi-categorical data sets as such, the simplicity of these visualization techniques can also be seen as a benefit while building more sophisticated visualizations. Simple line, index, and column charts are actually very functional methods

to combine with other visualization techniques, like heat maps (see the Subsection 4.1.6) and small multiples (see the Subsection 4.1.2). Line and column charts work also quite well together, if there is a need to include two different time series distinctively in the same figure. For example, in the Figure 4.1 two house price development related lines and a stacked column chart regarding house sale rates are combined in the same figure by utilizing two value axes.

Another approach for these simple visualization methods is to utilize polar coordinate system instead of traditional vertical and horizontal coordinate axes. In a polar coordinate system, first variable's values increase from one center point towards an indefinite perimeter while second variable's values are defined by a rotating angle around this same center point (for examples, see the Figure 5.3 and the Figure 5.6). Albeit polar coordinate charts might support a presented matter well as an eye-catcher, there are a couple of limitations concerning their applicability. Along with most people's inexperience and unfamiliarity towards them, polar coordinate charts are also seen troublesome due to their inability to illustrate negative values or extend rotation-based range without making a chart too dense, and thus cluttered.

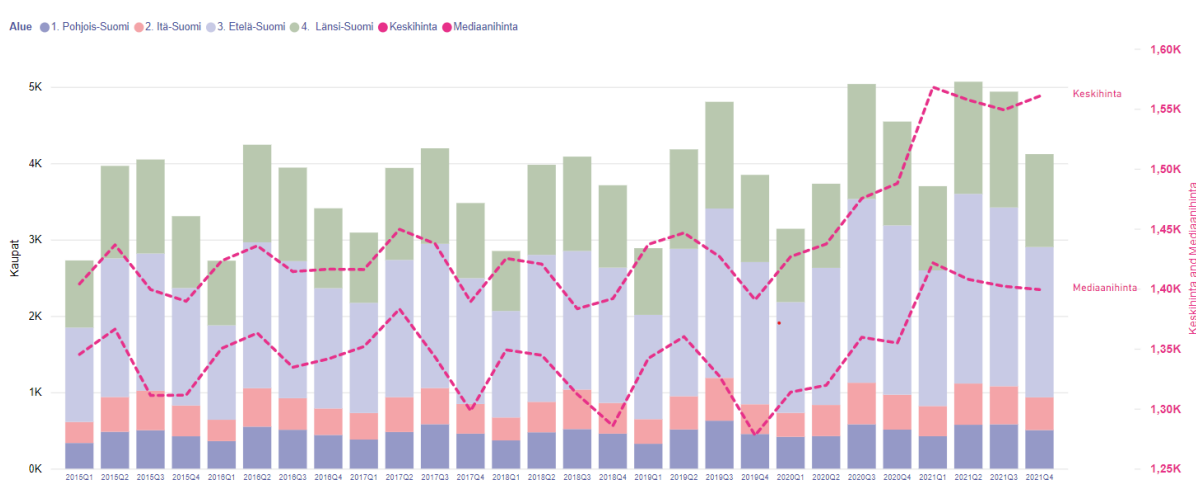


Figure 4.1: The regional housing market transactions in Finland illustrated by the dual-axis combination of stacked column graph and line chart. The stacked bars represent the numbers of housing market transactions (the right y-axis) and the lines express the developments of the average and median prices of detached houses (the left y-axis).

4.1.2 Small Multiples

According to the paper [16], one of the most common substitute for line charts that consist of several time series is to represent each time series individually in a separate chart by means of small multiples (see Figure 4.2). Principally, small multiples are utilized to avoid

overlapping and clutter, instead of trying to fit all information into the same picture. The specific small multiples can be indicated and found from the others by using different identifiers, such as colors or labels. However, small multiples may require a lot of space if there are many time series to display and an intention to maintain the charts on decent size. Moreover, the separation and shrinking of time series visualizations as small multiples trades a lot of resolution, especially vertically. Also the browsing-based comparability of time series declines, since certain charts may end up locating quite far from each others despite the order of the arrangement if there are too many time series included in the small multiples [18]. Besides line charts, small multiples are also well-suited for other time series related visualization categories, such as column charts, horizon graphs, and heat maps [15].

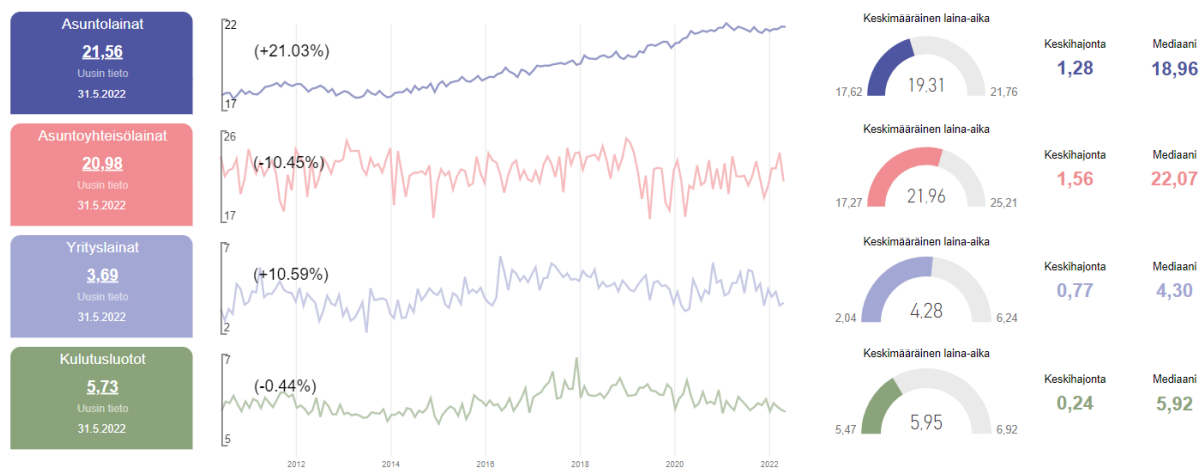


Figure 4.2: The maturities of the categorized domestic loans in Finland illustrated by the small multiples technique. The line charts share the x-axis, even though each of them have an individual stubbed y-axis. The colors are utilized as a reference method to connect both the line charts and the key performance indicators to the right loan categories. The key performance indicators concern respectively the latest published values (colored boxes), minimums, averages, maximums (colored gauges), standard deviations, and medians (colored numbers) of the time series illustrated by the colored line charts.

4.1.3 Stacked Graphs

Since the Bank of Finland already utilizes stacked column graphs a lot in their statistics, regarding the research question Q_4 (see the Section 2.4), it is relevant to focus on stacked line graphs only. The article [15] defines laudably the main characteristics of stacked line graphs, hereinafter shortened as stacked graphs: In stacked graphs, time series data is represent by stacking all data points one on the other starting counting the value-related position of them from on the next lower point that concern the same time point (see the Figure 4.1). The data points of each time series are also connected by colored lines, and the areas beneath of these lines are colored with the same colors than the lines over

them. Like many other visualization techniques, also the stacked graphs have multiple variations. For example, stream graphs are stacked graphs, which align stacked time series visually around the central axis of a canvas. This visualization technique is mainly used to illustrate the multiple cross-section summations of time series over a certain period. Therefore, stacked graphs provide a quick and effective way to perceive the bigger trends and to outline the respective proportions of time series. However, stacked graphs have some significant disadvantages as well.

On the other hand, the paper [16] sums some of the most notable drawbacks concerning stacked graphs: Stacked graphs cannot display negative values, precise comparison between time series is challenging, and it is almost impossible to ascertain the accurate values of stacked time series. Thus, stacked graphs are quite meaningless and awkward choice to visualize time series data if there is no reason to illustrate summed data points. For example, summing temporary velocities or temperatures up together do not usually offer much reasonable information. In addition, it has been noticed that viewers tend to make subconscious misestimations with the stacked graphs by perceiving predominantly areal minimums and maximums being directed arbitrarily in respect of adjacent time series instead of considering the actual vertical distances between the stacked data points. Due to this psychological aspect, stacked graphs should be used cautiously or avoid totally in the contexts which require comparison between different time series. Nonetheless, by using interactive filtering and sorting options, the above mentioned limitations may be resolved quite decently, if these interactions enable examining visualized time series individually, in different order, or in smaller groups.

4.1.4 Ribbon Chart

In 2017, Microsoft's data visualization software Power BI released a new advanced version of traditional stacked column graph called as ribbon chart. The online book [27] explains well the essential aspects of ribbon charts and clarifies their main differences between stacked line graphs and stacked column graphs: Similarly to stacked column graphs, in ribbon charts, the data values of different time series, represented with colored bars or columns, are summed up together by stacking bars one on the other. However, in ribbon charts the order of the bar stack varies over time in respect of the value-related magnitude ranking of the bars unlike in other stacked graphs. In addition, the changes between bars are represented by colored trend line segments. Thus, ribbon charts comprehend the line charts' ability to display trends over time as well as the the stacked graphs' ability to illustrate the total values of different time series over time. Also, the psychological misinterpretation between stacked time series (referring to the Subsection 4.1.3), can be

avoided by using ribbon chart (or stacked column chart). Since there are used bars instead of line-connected points to represent the data values of time series, it is much easier for human eye to concentrate a sight correctly between the comparative interval of adjacent time series segments. In Power BI software, there is also a possibility to include hoverable tooltip windows in both, bars and colored trend line segments, to enable the examination of the exact values of bars and the relative changes between them.

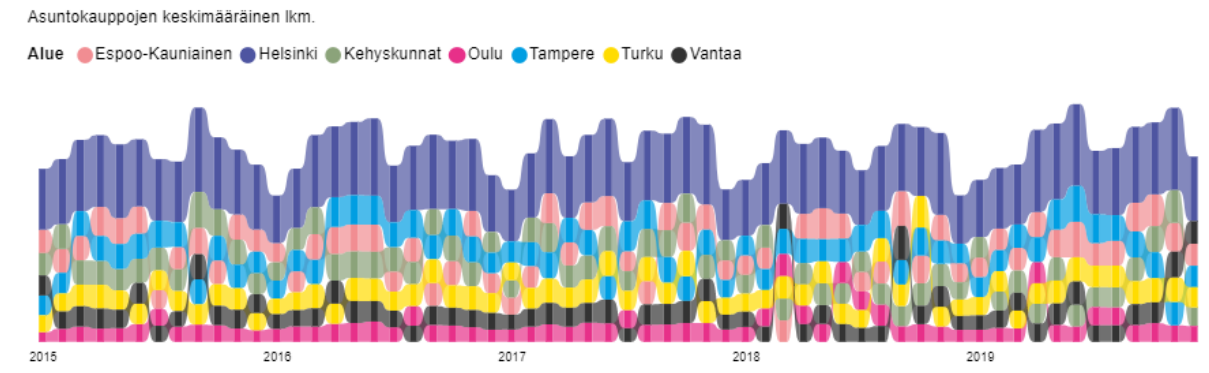


Figure 4.3: The regional housing market transactions in Finland illustrated by the ribbon chart technique. The stacked bars represent the numbers of housing market transactions. In addition, the order of the stacked bars represent the time-dependent magnitude ranking between the different regions of Finland. Therefore, the order of the color-coded bars would not necessarily maintain stationary over the displayed period alike in this figure. The uppermost bar has always the highest ranking, having the number one as the lowest ranking ID, whereas the undermost the lowest ranking, having the highest ranking ID.

4.1.5 Braided Graph

The article [18] introduces a new visualization technique for time series, called as braided graph (see the Figure 4.4). Braided graphs are visually pretty similar to line graphs. Likewise, braided graphs display time series by line-connected data points. However, in the braided graphs, the area beneath each line is colored until the next lower line or the horizontal axis in the same manner as stacked graphs, even though time series are not stacked. Since braided graphs might remind of stacked graphs at first glance, it is important to notice that data points' value-axis-related plumb distances from the time axis represent the actual values of each time series, unlike in stacked graphs, in which data points' plumb distances from the time axis are stacked, and therefore the data point positions of stacked time series represent only the sum of the values of all time series stacked below, including the individual value of an inspected time series as well. Specifically, the most characteristic feature of braided graphs is how they solve the overlapping problem by utilizing an innovative coloring technique. Instead of adjusting the transparencies of the colored areas, braided graphs utilize a procedure to change

the depth-wise layer order between two adjacent areas when their upper boundary lines intersect. After the intersection point, in which time series S_1 and S_2 satisfy the following prerequisites: $S_1(t_i) < S_2(t_i) \wedge S_1(t_i - 1) > S_2(t_i - 1)$, the colored area of the time series S_1 will be lifted depth-wisely from below to the top of the time series S_2 . This sorting procedure ensures that the upper boundary of each time series coloring maintain always visible, and therefore there is no risk that time series might unintentionally disappear under others after their intersection points. However, due to this color segment sorting policy, braided graphs might not be very convenient to use if there are a lot of time series to visualize. Over-dense and thin color segments, which get continuously reorganized, may only bring about visual chaos and clutter.

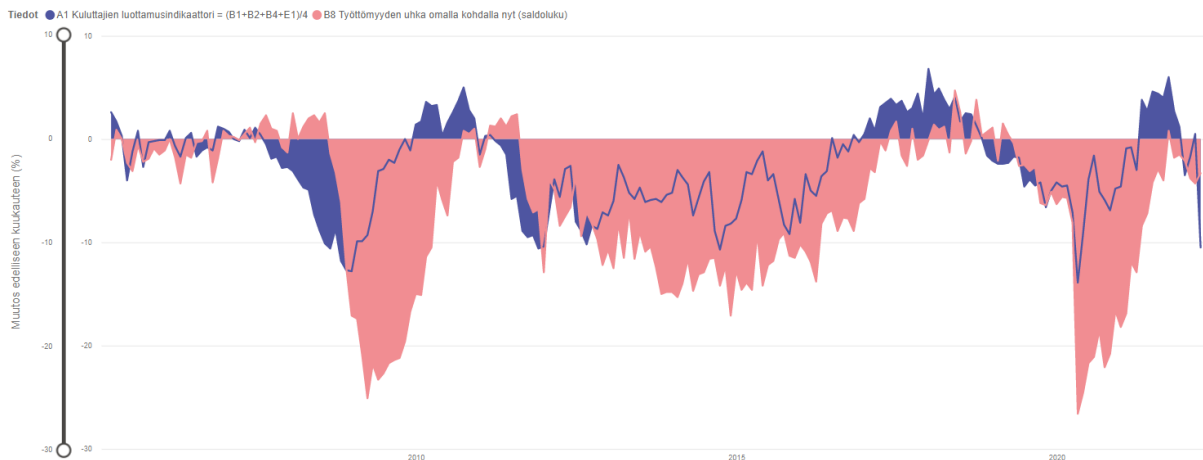


Figure 4.4: The economic confidence of consumers (blue) in comparison to the consumers' current personal threat of unemployment (red) illustrated by line charts. The values represent percentage changes of the poll answers' graded positivity from the previous month. The interjacent areas between the lines of the time series and zero-axis are colored in which way the line charts render the depthwisely stacked impression of themselves for a viewer similarly to braided graphs. However, in this figure the depthwisely stacked order of the line charts remains immutable from beginning to end, whereas in braided graphs the depthwisely stacked orders change whenever the data-related outlines of time series silhouettes intersect (explained in the Subsection 4.1.5).

4.1.6 Slick Graph

Slick graph, introduced in the article [17], is an innovative smoothing-based visualization method for the time series having high variance. Though, there are many other smoothing techniques as well, slick graphs are exceptional in respect of their ability to maintain all information in a visualization even after smoothing, unlike the most of the other techniques which tend to lose always some information while smoothing volatile high frequencies. The key feature of the slick graphs is their frequency-based separation policy, which classifies time series' all data points either into low frequency class or high

frequency class. The main reason for this arrangement is that both of these frequency classes are utilized in slick graphs but their mapping techniques differ. The data points of the low frequency class are mapped based on the value-related vertical positions, and thus visualized by using a common line chart approach, whereas the high frequency data points are mapped based on either color hues or grayscale shades, and visualized under the smooth low frequency line as a heat map.

Slick graphs have performed laudably in many perceiving-related comparison experiments, where the variance is high enough to provide clearly visible changes in the high frequency heat map below the low frequency line. However, users may consider slick graphs pretty confusing if there are no significant changes between the nearby data points, and the heat map segment looks almost mono-colored. Obviously, there is an option to adjust the hue threshold of the heat map segment to being more sensitive for alterations, but this solution may easily lead to users' misinterpretations, since the high frequency changes seem a way more dramatic than they actually should. In addition, slick graphs are suitable to use only with a singular time series at a time, since the heat map segments of slick graphs cannot avoid an awkward clutter if they overlap in the trivial manner, without utilizing, for example, the braided graph technique introduced in the Subsection 4.1.5.

4.1.7 Horizon Graph

The paper [15] defines adequately the main characteristics of horizon graphs: Horizon graph is an ingenious visualization approach to reduce the space requirement of a single graph without trading any vertical resolution of it (see the Figure 4.5). In other words, horizon graph is a very space-efficient technique to advance the information density of a time series still retaining all the original data and details aboard. In horizon graphs, data attributes are mapped based on a vertical position and visualized by line-connected points, like in ordinary line charts. Moreover, in horizon graphs, the vertical dimension, that is related to the domain of time series values, is divided by author-specified threshold steps. These steps are used to classify data attributes in different classes, which constitute a nested layer structure. These layers are utilized to visualize time series data by stacking them depth-orientedly one on the other in order of the nested structure ranking. In addition, the underneath areas of these stacked line subcharts are colored similarly to stacked line charts defined in the Subsection 4.1.3, yet utilizing different hues on different layers depending on their depth-wise position.

Like ordinary line charts, horizon graphs are suitable for displaying negative values as well. In fact, horizon graphs are more reasonable and adequate to utilize with time series having a distinct reference value as a baseline. According to the article [18], this reference value of a horizon graph should not be necessarily a zero. For example, it is pretty justifiable to set 273.15 as a reference value while using Kelvin scale instead of Celsius to display temperatures in a natural science context. Typically, the shifts over the reference value are expressed by changing the whole base colors of the areas beneath lines, for instance from blue to red, instead of usual and more subtle transitions between different hues (the brightness transitions of the same base color).

Even though, horizon graphs have many advantages, like their ability to conserve space while still preserving all the information of an original time series data, and provide a comprehensive trend and pattern overview of data over time by a quick glance, horizon graphs are not so intuitive to interpret for beginners. Also, in many cases, horizon graphs are displayed among the large ensembles of other small horizon graph multiples (see the Subsection 4.1.2), which may be seen as pretty chaotic and not so comparison-friendly technique, if outermost graphs end up being pretty far away from each others, or there is too much variance in individual time series and not so much correlation between them. Similarly to slick graphs (see the Subsection 4.1.6), neither stacking is an option with multiple horizon graphs, since the overlapping color segments would lead to even more chaotic clutter after all. Nevertheless, if the number of the included time series is kept inside reasonable bounds, all included time series have the same reference value as a baseline, and these time series are associated with the same context, horizon graphs are argued to be an ideal technique to compare the trends and patterns of several time series efficiently.

4.1.8 Heat Map

The article [13] clarifies aptly the main aspects of heat maps, also known as color fields or color maps. The core concept of heat maps is to encode the value range over time by using different colors and hues, instead of usual spatial attributes, to represent the magnitudes of values (see the Figure 4.6). Heat map coloring techniques might vary from stepped color block display to smooth and seamless sliding scale -based approach depending how granular used source data is, and how smoothly the changes and transitions between different recording points of data values are desired to express. Singular heat maps can also include several different color scale transitions, like starting a hue transition from red to yellow and then continuing from yellow to green, or even utilizing the whole rainbow spectre of colors. It is also an option to use only gray scales

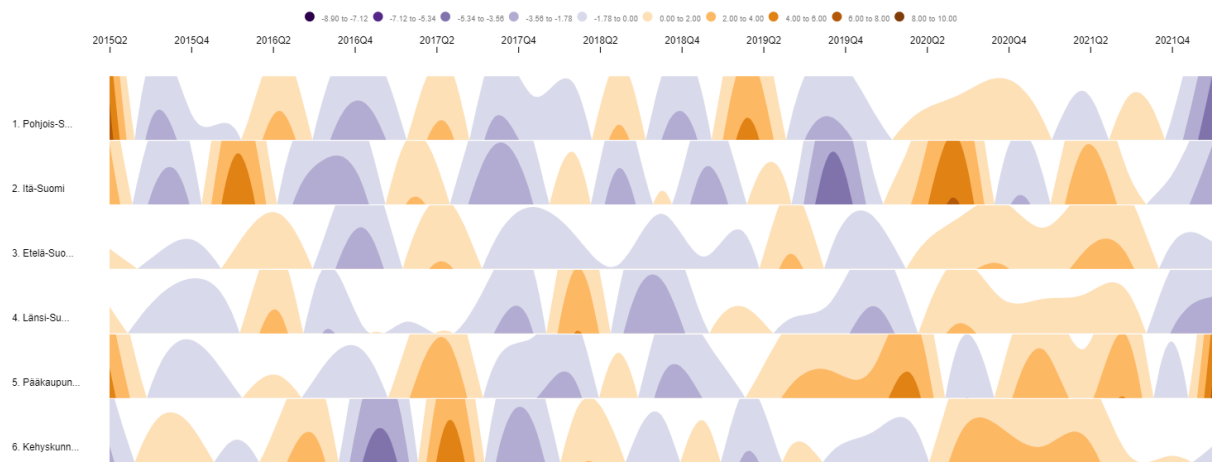


Figure 4.5: The regional housing market index changes in Finland illustrated by the horizon graph technique. Each individual horizon graph in each row concerns a distinct region. The orange-brownish hues represent positive changes which range is divided as five equal-size subranges. Consistently, the purple-bluish hues represent negative changes which range is divided by following the equivalent procedure.

to represent different values. In that case, the brightness encodes the range of values instead of an actual hue alteration. Obviously, the base color of the brightness encoding can be something else than gray or black, even though it is often referred as a gray scale. Moreover, the saturation value of colors can be used to do same trick. Therefore, it is said that there are altogether three color dimensions; hue, brightness, and saturation.

According to the paper [13], the main reasons to exploit heat maps as visualization techniques are their ability to increase a data density in a single graph. Since heat maps are an easy solution to improve significantly the data densities of visualizations, they enable better scalability, and therefore they can be used for large and continuously expanding data sets without becoming shortly cluttered [29]. Thus, heat maps are considered as a very space-efficient visualization technique for time series data [29]. In addition, many studies, like [9], are shown that heat maps are apt to provide a perceivable overview and comprehension over different ranges of time series data at a glance. The paper [9] lean on the human eye -related perception theory by which it reasons how heat maps might advance perception performance significantly in comparison to ordinary line charts. Even though heat maps are brilliant for perceptual averaging and giving a quick overall impression about data, heat maps are not a very adequate visualization technique alone to indicate the specific values of single data points accurately. However, there are already developed a plenty of different visualization applications that utilize the beneficial features of the heat map encoding as a part of a more complex visualization. That is to say, heat map coloring technique is quite common to use as the part of

various visualization fusions to enhance perceptual performance (for example, see the Subsections 4.1.6). There are also presented a more complex heat map approach in the paper [36] that augments traditional heat map technique from one-variable-dependent color points to a two-dimensional plane where each color point is dependent on two variables instead.

Vuosi	tammikuu	helmikuu	maaliskuu	huhtikuu	toukokuu	kesäkuu	heinäkuu	elokuu	syyskuu	lokakuu	marraskuu	joulukuu	
1995													
1996	-5.50	-5.00	-4.50	-5.50	-4.00	-1.90	-1.60	0.10	0.80	0.10	-3.10	-2.20	-3.80
1997	1.10	2.40	2.70	3.60	2.60	2.40	2.00	1.30	3.00	2.80	2.70	2.40	
1998	1.20	0.70	2.60	2.70	2.50	2.50	3.40	2.80	-1.00	-2.20	-0.80	1.30	
1999	0.10	-1.40	0.00	1.10	2.00	3.00	2.40	1.40	1.10	2.20	2.10	2.90	
2000	3.60	3.90	3.40	1.80	2.30	1.80	2.80	1.10	-0.90	-1.60	-0.20	0.70	
2001	1.60	0.50	-1.90	-2.40	-1.00	-2.60	-4.00	-4.40	-5.00	-4.70	-4.90	-2.40	
2002	-0.80	-0.20	1.10	1.00	2.00	-0.10	-0.10	-1.80	-1.00	-1.90	-2.00	-1.30	
2003	-1.30	-3.20	-1.40	-2.20	-1.50	-2.30	-0.60	-1.00	-0.30	-0.60	-0.70	0.00	
2004	0.80	0.20	0.60	0.80	-0.40	-0.60	0.60	0.50	-0.40	-0.80	-0.90	0.60	
2005	0.50	1.20	2.60	1.70	0.30	-4.00	-1.20	0.80	-2.70	-0.30	0.20	-0.10	
2006	-0.10	0.80	-0.70	-1.70	0.10	0.60	-1.70	-1.10	-0.80	-1.10	1.20	1.00	
2007	0.70	0.00	-0.20	0.90	0.10	1.10	0.40	-0.50	-1.90	-1.70	-2.80	-2.70	
2008	-3.10	-3.90	-4.70	-4.90	-7.30	-8.80	-10.10	-10.60	-8.70	-11.80	-12.70	-12.80	
2009	-9.90	-9.90	-9.30	-7.00	-3.10	-2.90	-2.00	-2.30	-1.00	0.00	-1.10	1.40	
2010	1.70	3.60	3.20	3.30	0.30	1.60	2.60	3.70	5.00	2.80	2.00	-1.50	
2011	0.30	0.40	-0.20	-0.60	-1.50	-5.80	-5.50	-8.90	-9.50	-9.30	-10.60	-10.40	
2012	-7.00	-3.90	-5.60	-2.90	-2.60	-8.00	-8.90	-10.20	-8.40	-8.70	-7.10	-7.40	
2013	-6.00	-2.50	-3.70	-5.20	-5.80	-4.70	-6.10	-5.90	-5.80	-6.10	-5.40	-5.20	
2014	-3.00	-3.70	-4.40	-7.40	-5.70	-4.10	-3.20	-8.90	-10.70	-8.40	-8.20	-7.70	
2015	-5.90	-3.20	-3.40	-2.10	-1.20	-4.00	-3.40	-5.90	-8.30	-9.20	-5.80	-8.10	
2016	-3.40	-5.00	-5.50	-3.60	-3.10	0.10	-1.80	-0.50	-1.20	0.40	-0.30	0.50	
2017	2.00	0.80	3.10	3.50	3.90	3.30	3.70	2.60	3.00	4.40	1.90	6.80	
2018	4.30	4.90	3.80	2.90	4.10	1.70	2.50	2.40	1.30	0.10	-1.60	-2.10	
2019	-2.40	-2.40	-2.30	-1.70	-1.80	-4.60	-3.90	-4.50	-4.20	-6.60	-5.00	-4.20	
2020	-4.60	-4.50	-7.10	-13.90	-9.00	-3.90	-1.60	-5.10	-5.90	-6.90	-4.80	-4.60	
2021	-0.90	-0.80	-3.00	3.80	2.70	4.60	4.40	4.00	6.00	2.70	1.20	-3.50	
2022	-1.70	0.50	-10.50										

Figure 4.6: The economic confidence of Finnish consumers illustrated by a calendar heat map, where the grid values represent percentage changes of the poll answers' graded positivity from the previous month. This heat map matrix is defined to having a calendar structure, because its rows represent years, whereas its columns represent months. By utilizing a calendar matrix structure, single time series can be represented more access-time-efficiently in comparison to the common table approach were time is represented by itemizing each individual date point separately forming a lengthy enumeration of values. The green hues represent positive values, the red hues represent negative values, and the saturation represents the distance from the neutral zero point. The higher the saturation the greater the distance, that is to say the greater absolute value.

4.1.9 Animations and Interactive Time Line

Along with the ordinary time series mappings based on spatial attributes, such as position and length, animations are also an adequate option to map time. In contrast to the spatial mappings, animations map time as the duration of displaying a discrete and chronologically ordered set of image frames concerning different time series data point sets. In other words, animation-based visualizations are either interactive or non-interactive videos that play a short movie about the evolution of data values via a graph [23]. Thus, in animations, time is represented by how users experience the passage of it. The paper [16] briefs compactly the main characteristics of the animation mappings of time series: In a time series animation, each frame represents an individual data

recording point under a certain time stamp. Depending the used visualization technique, each frame may or may not include the previous data values of the same time series which have earlier time stamps than the one under inspection. That is to say, each frame includes a visualization concerning either the specific cross-section of the time series at a certain moment only or all the previous values of the same time series as well.

By using animations to map time, there are more extensive range of suitable visualization choices to illustrate the actual data values than there are by using spatial time mappings. In fact, there are not barely any restrictions in respect of the choice of visualization techniques to be compatible with animations, since each animation frame only represent a static display of data points regarding a specific recording moment (time stamp). According to the paper [16], animations have been proven to support the ability of perceiving change-related estimations better than other visualization techniques based on spatial time mappings. However, it is also shown that animations tend to perform significantly weaker with the tasks that regard the actual estimation accuracy compared to the visualization techniques utilizing spatial time mapping methods.

Interactive time lines can be used to map time in the same manner than animations. The main difference between these two is that the animations are procedures that display time slices automatically and chronologically obeying a certain frame rate, and the interactive time lines are more like filters which enable users to navigate through different time stamps and periods manually and arbitrarily. Even though interactive time lines, are equally capable to map time, instead of using them alone, there are usually utilized some spatial time mapping methods as well to facilitate comparison between different data recording points.

4.2 Interactions in Data Visualization

In data visualizations, user interactions are designed to help users to handle large and complex data sets, which cannot be fitted entirely in a single display without losing too much clarity [26]. Interactions are used hand in hand with visual illustrations of data, since interactions modify the display settings of data which affect the appearance of visualizations directly. For example, adjusting a time line and switching between different time aggregations (granularities) are very essential and common user interactions regarding time series data [2]. Moreover, zoomable value axes, hoverable detail/info windows as well as linking and brushing operations are other effective interactions for time series data [22]. In general, interactions enable users to focus on specific viewpoints, themes, details and bigger picture of data alternately [26].

Even though interactions may be mainly seen as great tools for exploring and browsing large data sets, they have some disadvantages as well. The paper [32] discusses diversely about the pros and cons of user interactions in visualizations, and how those interactions may have both a positive and negative impact on a user experience:

(Pros) The most intuitive benefit of using interactions would be their ability to save a display space by showing only definite sub sets of data at a time, and hiding the rest. Also, the whole data set might be too complex or wide to perceive from a single view, even though the display space would be large enough to avoid a visual clutter. In the most of cases, this kind of problems could be easily solved by supporting users' navigation, filtering, sorting, and other display-related inspection choices via an interactive user interface which provides sufficient and clear hints that lead users to principally focus on essential aspects concerning their interests, and conduct them quickly to look at desired direction. For example, the visualization concerning the public transportation network of the city of Paris (see Figure A.1) might be a very successful compaction to capture an enormous public transportation network into a single image, while keeping everything pretty readable and distinguishable despite the size of the data set, but does it conveniently fulfill its purpose as a quick helping hand for a wandering or lost traveller? Even though it is a bit cliché to say, the phrase "less is more" tends to often solve the juxtaposition between practical and overly-complex visualization choices. Therefore, an effectual interaction-based interface should consist only of a few selection options per a view, and avoid confusing viewers. Moreover, the author of the visualization cannot presuppose either any further additional background knowledge about the visualized subject or adequate technical skills to use and adjust the interactive user interface properly and logically from the end users without providing proper instructions for them.

(Cons) Despite the many bright sides of applying interactions to the user interface of the data visualizations, there are naturally some disadvantages to exploit them as well. Above all, if interactions enable users to adjust visualizations overly-free, the overview of the visualized phenomena may become biased due to the subjectively modified mapping settings. That is to say, users might try to either intentionally or unconsciously tune the visualizations in the manner which indicates the desired outcomes emphatically and omits the other aspects which dilute the appearance of the framed phenomena. The second issue concerns the weak comparability between differently formed user-tailored visualizations. Therefore, to ensure that the visualizations interactively formed by different users are still commensurate, the mapping settings used to build up these

visualization have to be the exactly same. Thus, if there are not enough preserved information about the previously used mapping, it might be pretty difficult nor possible to figure out later which settings to choose. The third problem is associated with the limitations of technical performance and time resources. If a user changes the mapping setting, it forces the computer to render the image of the modified visualization again which causes a time delay. Depending how laborious and complex the modified visualization is, as well as, how powerful the computer operating the task is this may take from multiple seconds to minutes. Even though, that kind of singular idle waiting periods would not harm anybody, in the case of trying multiple setting variations in a row, an unproductive waiting time total might increase unbearably high.

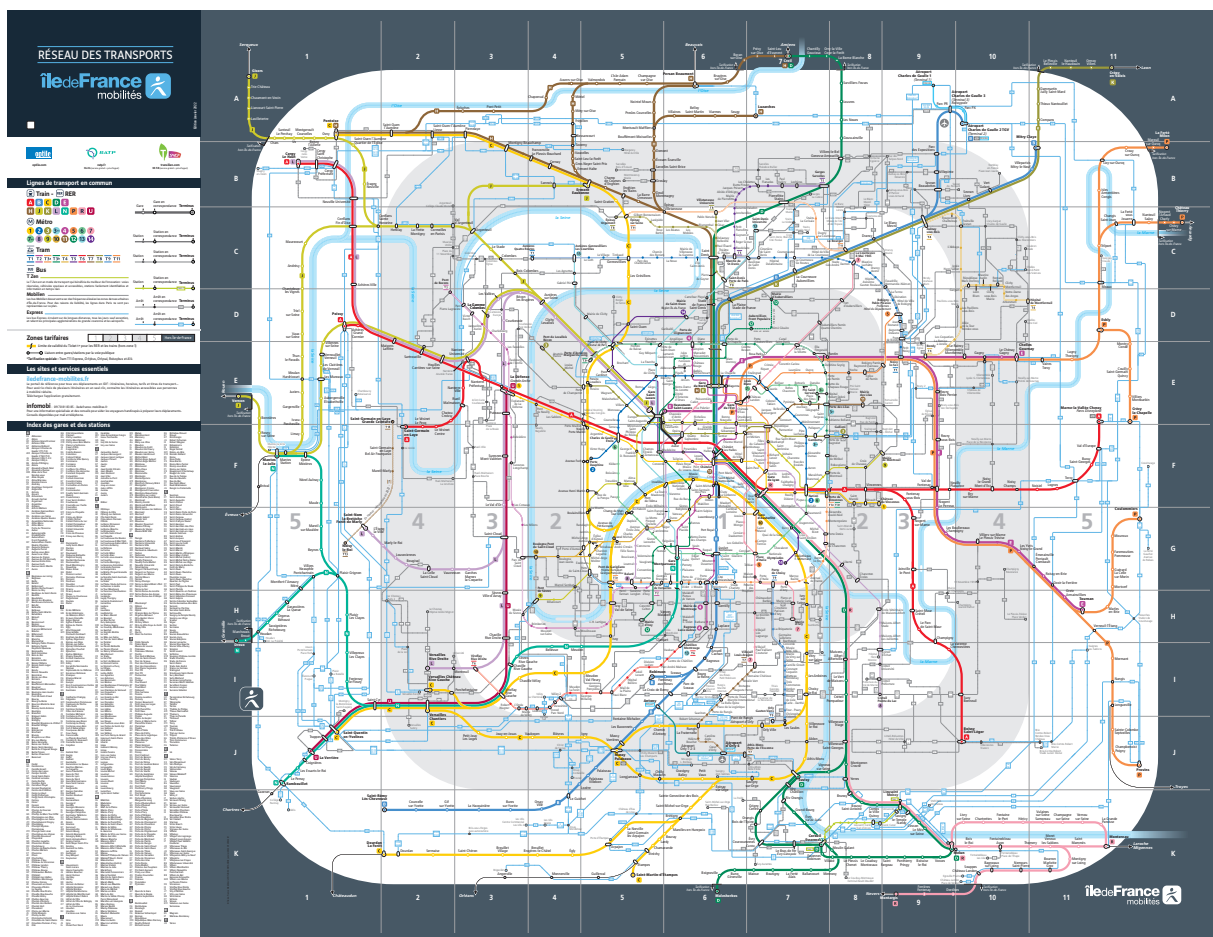


Figure 4.7: The transportation network of the city of Paris. (Source URL: <https://www.ratp.fr/en/planner>, downloaded 20th of May, 2022.)

5. Implementing Interactive Time Series Visualization

5.1 Data Wrangling

The most of the data I utilized for the implementation process of the final visualizations were queried from the databases of the Bank of Finland. In addition, Statistics Finland, that is the national statistical institution in Finland, provided some open data sets concerning Finnish macroeconomic statistics that were crucial for the final outcome of this visualization ensemble. These open data sets, provided by Statistics Finland, were queried by using PxWeb, which is an open source dissemination tool used for sharing tabular data. Since the main focus of this Master's thesis is not to build a visualization pipeline itself, and the Bank of Finland as the national monetary authority and central bank of Finland has strict public documentation restrictions concerning the detailed information of its databases, the data querying phase of this visualization project is omitted purposely in this thesis.

All data sets the Bank of Finland let me to use for the implementation of the visualization ensemble of the Finnish housing and mortgage market were time series organized in the following manner: Each data set consisted of three main dimensions; time stamp, time series category, and actual values. In addition, the majority of the data sets included multiple categorical filtering features, such as location, region, and dwelling type. Thus, every possible combination of the time series categories and their additional filtering categories were itemized under every time stamp having an individual value. Since the data sets covered the decades of recorded financial and economic information, and numerous different filtering combinations had an individual data value point per every time stamp, the data sets were very large and heavy to process. However, the data sets I used were built very punctually and adequately due to the Bank of Finland's precise norms of keeping statistics on.

Since all the data sets I used were already clean and homogeneous, my focus was principally to combine, manipulate, and aggregate the data as well as calculate new measures from it. For this data wrangling, I used Data Analysis Expressions (hereinafter DAX) that is the native formula and query language for Microsoft PowerPivot, Power BI Desktop, and SQL Server Analysis Services models. Because all the original data sets consisted only of the plain and simple time series recording values without having any further refinement manipulations, I had to process different data sets quite a lot to reveal the desired financial and economic characteristics and calculations. I also needed to create different time hierarchies and other categorical hierarchies in order to build versatile interactive visualizations in which, for example, time granularity, maturity, and regional depth were able to be changed.

5.1.1 Housing Market Data

For the Finnish housing market visualization, I used two different data sources: the first one concerns dwellings and the second one detached houses. The original dwelling data set had five dimensions: time stamp, time series category, value, region, and dwelling type. First, I need to convert all the values from the string data type to the float type. After that, I aggregated all the time stamps from the daily-level order up to the monthly-level order by using either the last available date of each month, the sum of all the values of each month, or the average of all the values of each month as basis for an aggregation depending on the time series categories. For example, it is usually more relevant to use the last available date of each month for the aggregation of economic indices, the sums for market transaction volumes, and the averages for different percentage changes, neither than any other combinations. I also created a date hierarchy in which years, quarters and months form a tree. This date hierarchy enables the end users to navigate between different date granularities in the final interactive visualization. In addition, I coded by DAX a new measure feature that picks only the values having the latest time stamps. The new feature syncs reliably with different filters and date granularities showing always the most recent non-empty values of the user-specified aspects in the user-specified period. By this feature, the end users are also able to get the most up-to-date information separated from the rest of data right away after opening the visualization.

The data set concerning detached houses consisted of four dimensions: time stamp, time series category, value, and region. Similarly to above, I started processing the data by converting the value dimension to the float type. I also needed to fix the format of the time stamps since the Power BI framework was not able to process the time stamps that

included both numbers and letters; quarter-based time stamps each including the letter Q before the quarter index. Since the time stamps had originally the quarter-based granularity, I was only able to create a simple date hierarchy between quarters and years. In addition, I needed to create five new derivative time series by which I was able to separate the specific time series categories, such as market transactions, average market prices, median market prices, quarter-based changes, and year-based changes, from the rest of the categories. These distinct time series enabled me to build interactive visualizations that allowed to display simultaneously the most essential time series either in the same figure or in the same dashboard constantly while shifting the concurrent time series of the other time series categories of the original data set being on display. Furthermore, I added a new measure feature by adapting the same DAX code I created previously; the DAX code that picks only the values having the latest time stamps.

5.1.2 Housing-related Loan Market Data

For the Finnish mortgage and other housing-related loan market visualizations, I used two different data sources: the one that concerns the macroeconomic perspective and the other one that concerns the creditor perspective. All the Finnish mortgage and other housing-related loan market visualizations of the final visualization ensemble, apart from the creditors market share visualization, are from the macroeconomic perspective. They are founded on the single massive data set including altogether 147 different dimensions, including the three main dimensions; time stamp, time series category, and value; similarly to other data sets. Again, I started processing the data by converting the value dimension to the float type. After that, I created a date hierarchy, analogous to the Subsection 5.1.1, in which years, quarters and months form a tree. I build similar date hierarchies for all my data sets in order that data sets would be able to connected together if any essential needs for that intend to do appear afterwards. This date hierarchy also enables the end users to navigate between different date granularities in the final interactive visualization. In addition, I created a new measure feature by adapting the DAX code, being created in the Subsection 5.1.1, again. This essential feature picks only the non-empty values having the latest time stamps taking into account all the user-specified filters and time frame.

The data set concerning the creditor perspective of the Finnish mortgage and other housing-related loan markets had the same amount of dimensions than the first data set, having the most of them equal in fact, again including the three main dimensions; time stamp, time series category, and value. Even though there were nearly one hundred and fifty different dimensions, this data set was much smaller than the one from the macroeconomic perspective, since its recorded time frame was significantly shorter.

Albeit the creditor data set was a lot lighter to use than its macroeconomic giant twin, it had a remarkable size as well. The creditor data set did not require any exceptional refinement actions, and it already included all the essential numbers and percentage shares that I needed to illustrate the market shares of the biggest Finnish creditors. However, I still had to do some fundamental modification, like changing the value dimension to the float type, and creating a date hierarchy that is analogous to the other date hierarchies of the parallel data sets.

Since these data sets were very large, especially the one from the macroeconomic viewpoint, and consisted of an extensive selection of different financial filters, I had to study their structures and contents a while until I had a certainty which filters I need to use together and how to use them properly in order to get any reasonable numbers out of the data. Many of these filters were already built hierarchically structured having different aggregation stages integrated in them. Therefore, it was very precise to either include or exclude the certain filters, for example, regional filters, to avoid over- or underestimating the already aggregated values.

5.1.3 Domestic Indebtedness Data

The domestic indebtedness data set was a relatively small and simple data set consisting only of the three main dimensions: time stamp, time series category, and value. Since the original data was founded on an annual recording granularity, there was neither any relevant reason to aggregate the time dimension of the data more coarse-grained nor construct a date hierarchy, since years was the most infrequent granularity level that the economists of the Bank of Finland regarded to be meaningful for their macroeconomic analysis of Finnish housing and mortgage markets. Nonetheless, I needed to do two modifications to the data; I converted the whole value dimension to the float type, and I created a derivative time series to separate the average income of the Finnish households from the other domestic indebtedness characteristics. By the separated average income time series, I was able to build interactive visualizations which enable to display the Finnish average income constantly while shifting the other concurrent time series of the other time series categories in the same figure.

5.1.4 Finnish Economy Barometer Data

The last data set concerned the polls of the Finnish economy barometers including the three main dimensions: time stamp, time series category, and value. Similarly to the other data sets, I started processing the data by converting the value dimension to the float type. The original time stamps were based on a monthly granularity state, and they

were formatted in the way that combines both numbers and letters, including a year, the letter M, and a month index. I needed to fix the format of the time stamps since the Power BI framework was not able to process the time stamps consisting of both numbers and letters. After that, I created a date hierarchy, analogous to the Subsection 5.1.1, in which years, quarters and months form a tree.

5.2 Creating and Assembling Visualizations

When all data were wrangled, parsed, and collected together, I was able to start the final implementation stage; creating the actual visualizations and assembling them into a dashboard ensemble. Since this visualization collection was meant to be an actively used analysis tool for the economists of the Bank of Finland, it was important to make its visualizations practical and easy to perceive. To ensure about the adequate functionality of the analysis tool, I studied various time series visualization techniques in order to use only those techniques which will be simultaneously practical to use and enhance perceiving important phenomena behind data. In addition, I determined different clutter reduction techniques, from which I selected the most reasonable and compatible ones to utilize in the final visualizations.

The time series data sets consisted of two essential aspects; hierarchically formed timeline and recorded values about the matters of interest. To successfully visualize time series data, both of these aspects have to be taken simultaneously into account. Obviously, many visualization-related requirements of recorded data points are mostly dependent on their context of use, whereas the main concerns of timelines regard either how to retain their chronological sequences in complex visualizations or how to capitalize the analytical benefits of timelines' hierarchical nature. Further in this section, I will explicate, which visualization techniques I ended up using in each dashboard sheet of the final analysis tool, and how I reasoned them.

It is not a secret that very trivial and commonly used line charts might fulfill most requirements expected from simple time series visualization tasks just fine. In fact, according to the paper [2], plain and simple visualization methods, such as line charts, usually tend to outperform more sophisticated techniques in simple data analytical tasks, since simple methods are obviously faster to use and understand. However, line charts have still their limitations, which quickly begin to stand out when the complexity level of data, the information density of data, or the degree of difficulty of analytical tasks increase. Along with traditional line charts, there are also many other effective alternatives to visualize time series data, for example, heat maps and horizon graphs.

According to the paper [13], especially horizon graphs (for horizon graphs, see the Subsection 4.1.7) are an applicable substitute for traditional line charts. However, they may have been found more complicated and slow to perceive for inexperienced viewers.

According to [11], the clutter reduction techniques can be categorized into appearance, spatial distortion, and temporal technique groups. Appearance techniques modify the way how data points are presented in a visualization. For example, color, shape, size, texture, and transparency are trivial properties of data points that can be utilized to enhance distinctiveness by visually separating two or more data elements, such as data points, groups, and even entire time series. Also, filtering techniques like category, value, hierarchy, and time sampling are considered as an appearance-relating techniques, because the presence of data points can be controlled with them. In addition, grouping data points by using visual methods, for example, additional lines, borders, or text, is defined as appearance technique, since it affects how individual data points are perceived and interpreted.

The paper [11] describes spatial distortion techniques as methods that displace data points spatially, for example, by changing each data points' initial position, or stretching the background of data points topologically (either entirely or partially) utilizing fish-eye effect or zooming. However, many spatial distortion techniques seem to fit quite poorly with financial time series data since it is crucial for meaningful economic analysis to keep all data points in chronological order, and therefore we omit this category in this thesis. Instead, temporal techniques are very suitable for time series data. Temporal techniques mainly consist of different animation techniques, which utilize humans' perception of time to arrange and synchronise data points display in relation to actual time. Animations were praised to be very illustrative techniques to display time series data, since the way how time is mapped as actual time that viewer can truly sense seems to be a very logical and intuitive composition. Nevertheless, the economists of the Bank of Finland did not consider animations as practical tools for macroeconomic analysis, because animations make it unnecessary complicated and slow to get access to the latest values, which are, in fact, the most interesting ones normally. Animations also make users to rely heavily on their memory, since all data points are not visible at the same time, and the visual display of data keeps altering regardless of users' state of attentiveness [23]. Obviously, in demanding and precise analysis tasks, this kind of additional and risky human factors are tried to be avoided. Thus, I decided to focus mainly on appearance-related clutter reduction techniques in this project.

From now on, the rest of this section covers the actual design choices of the analysis tool. The selected visualization techniques are introduced and reasoned in respect of their analytical context and dashboard sheet. Each dashboard sheet included into the final visualization ensemble is indicated with an individual identifier from *D1* to *D15*.

5.2.1 Housing Market Visualization

Finnish housing market data included the extensive and diverse set of different time series categories. Since the one of the research questions (see the Section 2.4, Q_2) concerned the best ways to compress large data masses into a moderate number of distinct visualizations, I included various interactive elements, such as filters, time sliders, dropdown menus, and shortcut keys into the dashboards in order to fit as much information into a single dashboard sheet as it was possible without trading too much clarity or simplicity of the end user experience.

***D1*: Regional Dwelling Prices per Square Meter and Dwelling Transactions**

Since prices and transaction volumes are a duet which together embodies the essentials of every market analysis, as an established convention, it is widely acknowledged policy to present these fundamental market variables hand in hand if possible. I had a permission to use the dwelling market-specified data, which consisted, inter alia, of both of these core variables, and therefore it was a natural resolution to include at least one orthodox price-volume assembly in this housing market visualization ensemble. The dashboard sheet *D1* illustrates the regional dwelling prices and transactions in Finland (see the Figure A.2 in the Appendix A section). I decided to represent dwelling prices per square meter, instead of total prices or other price variants, because square prices can be argued to be more commensurate unit for nationwide market analysis due to the high diversity of regional housing standards. For the prices, I decided to utilize a calendar heat map matrix (for example, see the Figure 4.6) to illustrate regional dwelling prices per square meter, because the combination of calendar matrix and heat map encoding techniques enables to compress a lot of time series information space-efficiently into a single visualization. Along with the calendar heat map, I eventually settled on to exploit the ribbon chart technique (see the Subsection 4.1.4) to compactly illustrate both regional market transaction rankings and transaction totals over time. Thus, this dashboard sheet ended up consisting of two time series visualization, omitting the other cross-section visuals and key performance indicators.

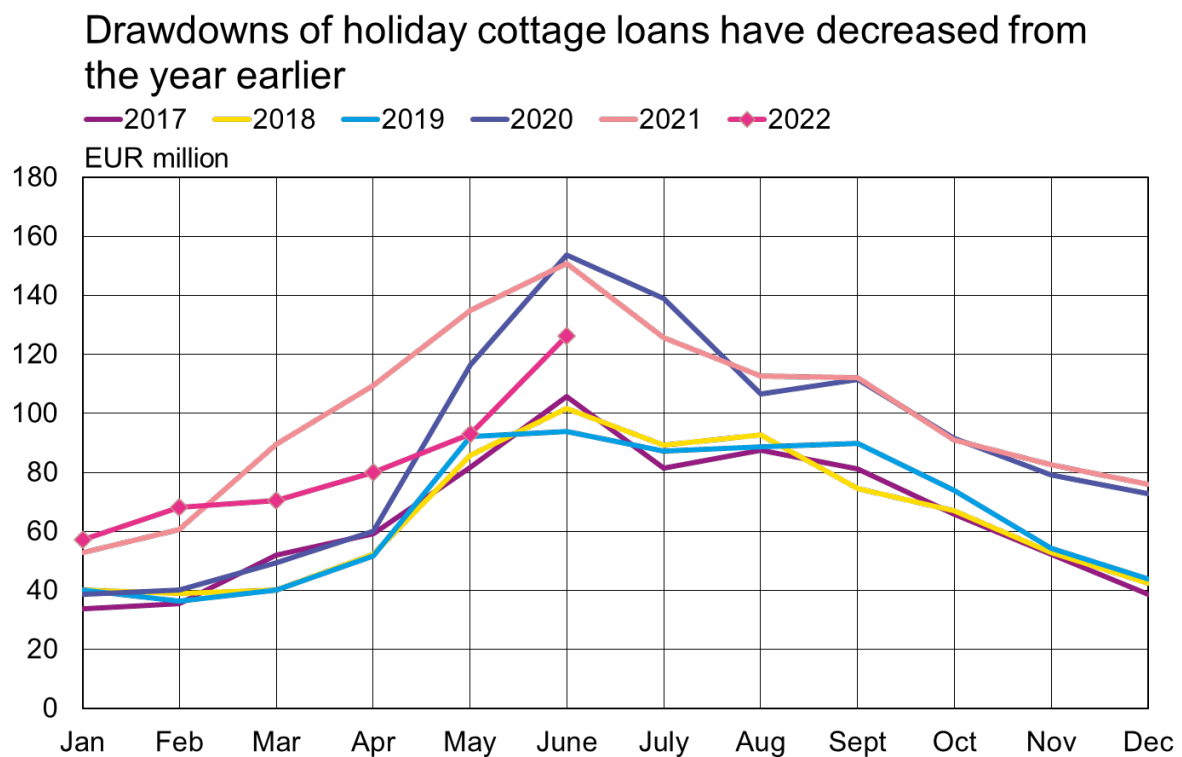
D2: Mortgage Variables in Respect of Dwelling Purchase Decisions and Actual Dwelling Purchases in Finland

The dashboard sheet *D2* collects together the most essential mortgage market variables, such as drawdown-related variables and interest rate variables (see the Figure A.3 in the Appendix A section). Moreover, dashboard concerns consumers' dwelling purchase decisions and actual dwelling purchases. I designed this dashboard to correspond the frequent necessities of the macroeconomic analysis tasks of the end users (see the answers of the preliminary end user survey in the Section 2.5). While being working almost a year for the Bank of Finland, I have learned that the most economic phenomena have a cyclic nature due to either conventional or statutory accounting periods, generally quarters and years. Therefore, one of my aims was to emphasize the cyclic nature of the mortgage and dwelling market variables by partitioning the entire time range distinctly into financial years.

I managed to innovate two exceptional alternatives for traditional line and column charts. My inspirational source for these approaches was the Figure 5.1 from the Bank of Finland's MFI (monetary financial institutions) balance sheet (loans and deposits) and interest rates bulletin [30]. In that figure, months are represented unconnected to specific years from Jan to Dec on a x-axis, whereas financial years are represented by individually colored lines in order to both compress information and ease comparison between the months of different years. Even though this visualization approach certainly has its pros, especially when it is used to illustrate the seasonal peaks of cyclic phenomena, there are some serious cons as well: If some time series are similar enough, the representative lines of these time series might end up being very close to each other, which may cause visual clutter and make it difficult to follow the individual lines effortlessly.

I wanted to fix this drawback, so I encoded two alternative visualization approaches, the one with the original time frequency and the one with the year granularity. For the visualization approach with the more frequent time granularity, I ended up utilizing the full time range on a x-axis to the detriment of the month-related comparison between different years. To enhance the observation of specific months, I added an interactive time slider below this visualization in order to adjust and border the visible time line. I decided to call this visualization technique as a color-partitioned line chart (see the Figure 5.2). The other alternative visualization approach, which I call as a radius chart (see the Figure 5.3), illustrates time series via a polar coordinate system. In this unconventional time series mapping, the full angular circuit around the coordinate center represents a whole time line, equally divided angular sectors represent time units

which are years in our case, and the radiuses of these sectors represent the magnitude of their time-related values. There are also two grid lines included into the radius chart, the outermost being set on the maximum value and the another pointing out the halfway between the zero point and the maximum, in order to enhance the comparison between different years. The main benefits of this polar coordinate presentation are its ability to cut a dash and, in some cases, to reduce a space consumption. According to the consultation of the Bank of Finland's economists, one fundamental function of a successful visualization is its ability to draw attention to the certain subject. Therefore, my secondary aim with these two alternative visualization approaches was to evoke interest, enhance perception, and deepen concentration toward the topic, by utilizing prominent and colorful design and unconventional mapping decisions.



Source: Bank of Finland.

Figure 5.1: The holiday cottage loan drawdowns over the months of the past five and a half years represented by yearly separated and individually colored lines.

D3: Mortgage Variables in Respect of Regional Housing Market Indices and Housing Market Variables

The dashboard sheet *D3* illustrates the essential mortgage market variables together with the housing market indices and other important housing market variables, such as the

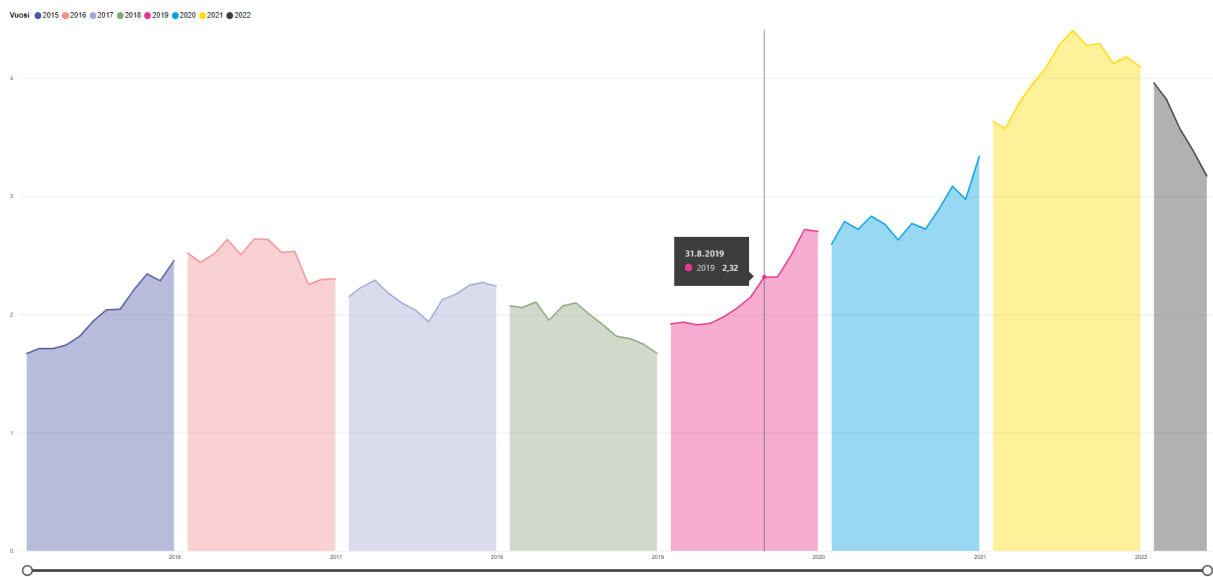


Figure 5.2: The percentage change from the previous month of the housing loan stock of Finland illustrated by a color-partitioned line chart. Here, each colored partition represents an individual calendar year. The granularity of the chart is monthly-based.

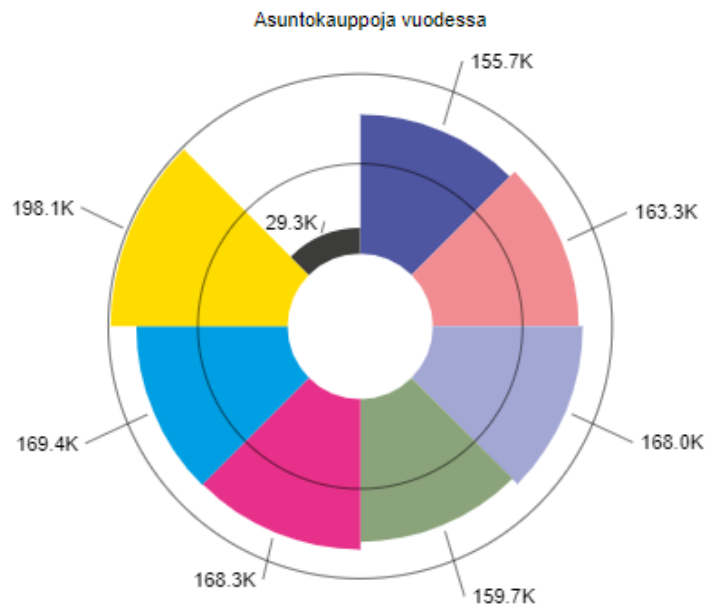


Figure 5.3: The Finnish house market transactions per a year illustrated by a radius chart. Here, the colors represent the same years as the Figure 5.2 displays, and the radiuses represent the numbers of the house market transactions.

number of market transactions, the average length of house marketing disposal periods, and the average house prices per square meter (see the Figure A.4 in the Appendix A section). The principal aim of this dashboard sheet was to include as much information from the development of the regional and otherwise categorized demand of both Finnish

mortgages and owner-occupied houses over time as possible without losing any important information or clarity. Due to many expectations and practical requirements originating from the end users (see the answers of the preliminary end user survey in the Section 2.5), I ended up using traditional line chart presentation in this dashboard. Nonetheless, there are utilized multiple interactive elements in this dashboard sheet to compress information: (1) Housing market -related visualizations include zoomable y-axes to enhance their resolution during a definite analysis. (2) The dashboard sheet has also an interactive time slider in order to adjust the start and end points of the visible time line. (3) Furthermore, the dashboard sheet includes three dropdown menus for economic variables, such as new drawdowns and the proportion of the non-performing loan stock, statistical classes, and Finnish regions.

D4: Distribution of Detached House Market

The dashboard sheet *D4* illustrates the regional distribution of the detached house market in respect of the prices of sold houses and the numbers of market transactions occurred over an inspection period via a scatter plot visualization and a time-dependent dot-size encoding (see the Figure 5.4 below and the Figure A.5 in the Appendix A section). The main purpose of this dashboard sheet was to provide an informative overview of substantial regional differences in the detached house market of Finland. Therefore, I needed to find a reasonable way to visualize multiple regional market distributions close to each other to ensure their convenient and effortless comparison. By utilizing the scatter plot technique and different colors, I was able to display multiple region classes simultaneously in a single plot without getting the view too cluttered. Since I also wanted to retain the time dimension of the data, I decided to encode time by using different dot sizes to indicate different years in the following manner: the smaller the dot the further the year from the present, and vice versa. To be honest, I think that the idea of the connected scatter plot presented in the study [14] would work even better to encode time in a scatter plot, but unfortunately, I did not figure out any reasonable way to connect separate data points chronologically with Power BI's default settings or with any other free Power BI add-ins I was able to use. In addition, this dashboard sheet included multiple interactive elements as well to compress information: (1) The chart includes zoomable x- and y-axis to enhance its resolution during a definite analysis and to reduce the unwanted and hampering clutter caused by the intersecting data point clusters. (2) The dashboard sheet also includes three dropdown menus for the following variables: region, population class, and inspected period as quarters. (3) Furthermore, the dashboard consists of two buttons by which enable to switch between the average price and the median price displays without vanishing other filtering settings off.

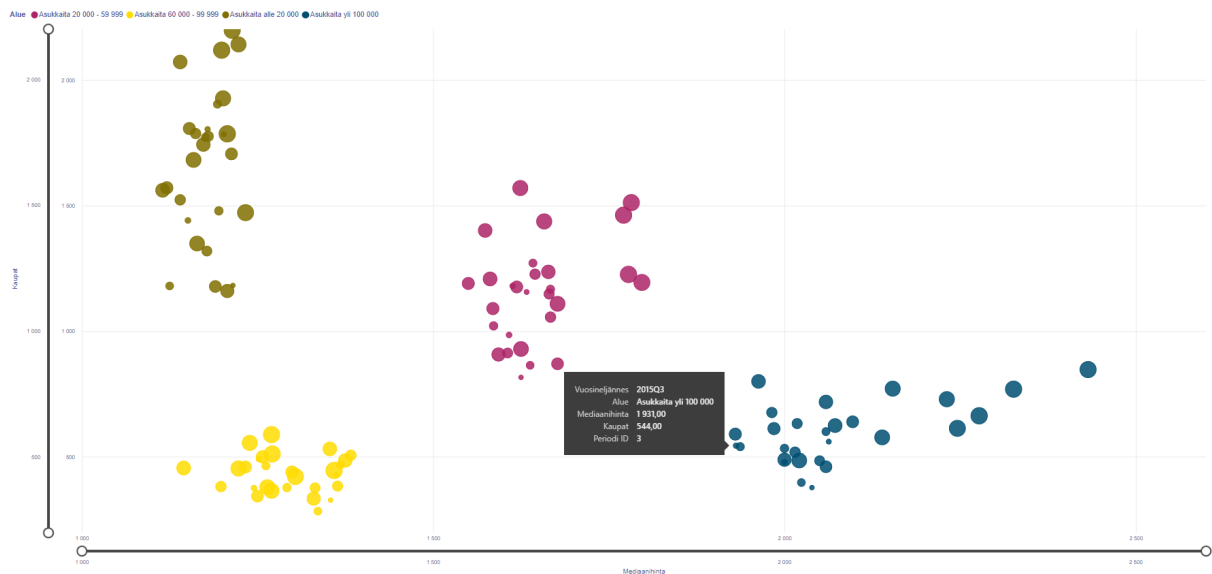


Figure 5.4: The detached house market distribution of four population size class in respect of the prices of sold houses and the numbers of occurred market transactions illustrated by a scatter plot visualization and a time-dependent dot-size encoding.

D5: Regional Outlines of the Most Essential Detached House Market Variables

The dashboard sheet *D5* illustrates the square prices and market transaction numbers of detached houses in Finland by combining a stacked column chart and two line charts into a single figure (see the Figure 4.1 below and the Figure A.6 in the Appendix A section). Since the market transaction numbers and square prices have different units and magnitudes, I had to utilize two value axes (y-axes) to adapt both measures sensibly into the same figure. The left value axis represents the market transaction numbers illustrated by bars, and the right one represents the square prices, which are expressed via two dashed lines; (1) average square prices, (2) median square prices. The dashboard sheet also includes three dropdown menus for the following variables: region, population class, and inspected period as quarters. Moreover, it consist of three interaction dependent key performance indicator boxes displaying the aggregations of the most essential variables over currently viable data points.

D6: Regional Comparison of the Most Essential Detached House Market Variables

The dashboard sheet *D6* displays the most fundamental detached house market variables; market transactions, and average and median prices per a square meter, regionally (see the Figure A.7 in the Appendix A section). The principal aim of this dashboard sheet was to enable an effortless comparison between the aforementioned region-specific detached house market variables over time. Analogously to the dashboard sheet *D3*, I ended up using

traditional line chart presentation also in this dashboard, due to many expectations and practical requirements originating from the end users (see the answers of the preliminary end user survey in the Section 2.5). The dashboard includes only one chart which includes multiple lines of distinct mortgage market variable time series. Nonetheless, there are utilized multiple interactive elements in this dashboard sheet to compress information: (1) The chart includes a zoomable y-axis to enhance its resolution during a definite analysis. (2) The dashboard sheet also includes three dropdown menus for the following variables: region, population class, and inspected period as quarters. (3) In addition, the dashboard consists of three buttons by which a user is able to switch between the three core subjects: market transactions, and average and median prices per a square meter, without losing any other filtering settings.

D7: Regional Changes of Detached House Prices and Market Indices

The dashboard sheet *D7* illustrates regional changes of detached house prices and market indices by utilizing horizon graph and small multiples techniques simultaneously (see the Figure A.8 in the Appendix A section). The main reasons for these visualization encoding choices were (1) to enable more straight forward pattern perceiving process for multiple comparable time series, and (2) fit them all into a single layout without losing any relevant information. Like explicated in the Subsection 4.1.7, horizon graph encoding is an excellent technique to compress information into relatively small space still retaining all original information. In turn, small multiples, covered in the Subsection 4.1.2, is an intuitive technique to collect comparable time series together in order to ease a comparison-based analysis. This dashboard includes also many interactive filter menus, by which a user is able to select different data contents and define the adequate state of the visualization's regional coverage.

D8: National and Regional Development of Detached House Transactions

The dashboard sheet *D8* reveals the regional shares of the detached house selling development in Finland via a waterfall chart (see the Figure 5.5 below and the Figure A.9 in the Appendix A section). The waterfall chart consists of red, green and blue bars, which represent decrease, growth, and the aggregated summation of all regional market transactions in a defined period respectively. Moreover, each one of those bars represents the development of a specific region over a specific period. In our case, periods are denoted either by a quarter, half-annual or annual frequency, because detached house market is tend to function quite calmly keeping its transaction volumes on the moderate level as well as having relatively low volatility over short periods. Thus, from the perspective of the macroeconomic analysis, it is neither very meaningful nor efficient to keep track of the

them. The actual values of the heat map matrix depend on a user-based selection of a detached house market variable. That is to say, all miscellaneous detached house market variables included into this dashboard were connected to the heat map matrix via an interactive filter menu. In addition, along with the heat map matrix technique, I used a bar chart to foreground the most recent values of a selected variable. Moreover this dashboard sheet ended up consisting of three different filtering menus and an interactive min-average-max value gauge.

5.2.2 Housing-related Loan Market Visualization

Housing-related loan market data was queried from the data bases of the Bank of Finland. The data consisted of the large set time series that formed the basis of the whole analysis tool. Therefore, I decided to utilize easily readable, uncluttered, and simplified visualization methods to display multiple time series simultaneously in a single figure. In addition, similarly to housing market data, I included various interactive elements into visualizations in order to compress as much subject-specific information into each dashboard as possible without significant clarity loss.

D10: Outlines of Housing-related Loan Market in Finland

The dashboard sheet *D10* outlines the housing-related loan market variables by utilizing traditional line chart and radius chart (see the Figure 5.6 below and the Figure A.11 in the Appendix A section) concurrently. Analogously to the dashboard *D2*, the radius chart uses a polar coordinate system to illustrate different time series. In the radius chart, the full angular circuit around the coordinate center represents a whole time line, equally divided angular sectors represent time units which are either months, quarters, half-years, or years, and the radiuses of these sectors represent the magnitude of the data point values. The sectors of the radius chart are alternately colored either with bluish or reddish colors in order to separate calendar years from each others. In addition, the sectors of the most recent calendar year are colored with turquoise colors for the same reason. Moreover, by utilizing the dark and light hues in turn, I was able to separate the more fine-grained time units simultaneously. Thus, this radius chart consists of a two-degree time hierarchy based on different colors and color dimensions. This dashboard sheet also included a classic line chart beside the ribbon chart. The purpose of the line chart was only to support the perceiving process of the radius chart, and therefore it provided the same information as the ribbon chart. In addition, this dashboard sheet included multiple interactive elements to compress information: (1) an interactive time slider in order to adjust the start and end points of the visible time line. (2) two dropdown menus for hierarchical economic variables; the first menu for different loan types, and the

second menu for the loan-related variables, such as new drawdowns and the proportion of the non-performing loan stock. (3) Furthermore, the dashboard consists of three buttons which enable to switch between different time series granularities, months, quarters, half-years, and years, without vanishing other filtering settings off.

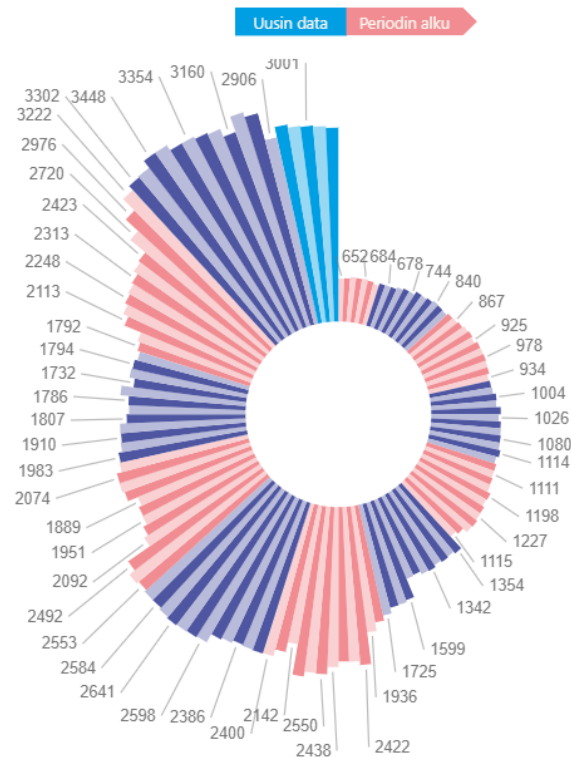


Figure 5.6: Hierarchical housing-related loan market variables with month-granularity illustrated over circa twelve years by utilizing radius chart and color dimensions.

D11: Mortgage Market Shares of Finnish Creditors

The dashboard sheet *D11* illustrates the mortgage market shares of Finnish creditors over time by utilizing a heat map matrix and a clustered column chart (see the Figure 5.7 below and the Figure A.13 in the Appendix A section). Since the heat map matrix of this dashboard sheet is quite analogous to the heat map matrix of the dashboard sheet *D9*, I decided to omit its inspection, because it does not bring any new aspect into the subject. Beside the heat map matrix, I used a clustered column chart to visualize all individual market share developments of the biggest creditors in Finland together. This encoding was able to combine multiple individual column charts together, because all individual column charts were compatible with the same value axis (y-axis), and the time dimension of these individual charts were able to be expressed by using different color attributes. I exploited bluish and reddish colors to visually separate bars years from each others.

In this figure, bars represent different years from 2017 to 2021. In addition, I utilized brightness to illustrate the transition from the past to the present in such a way that the past had less brightness than the present state. The main purpose of this visualization encoding was to provide a quick overview about the market share situation in Finland, and to outline the individual market share development of each creditor separately. This dashboard sheet also included multiple interactive elements to compress information: (1) an interactive time slider in order to adjust the start and end points of the visible time line, (2) a filtering menu concerning Finnish creditors, (3) a pie chart to illustrate the average market shares of the creditors over the inspected period.

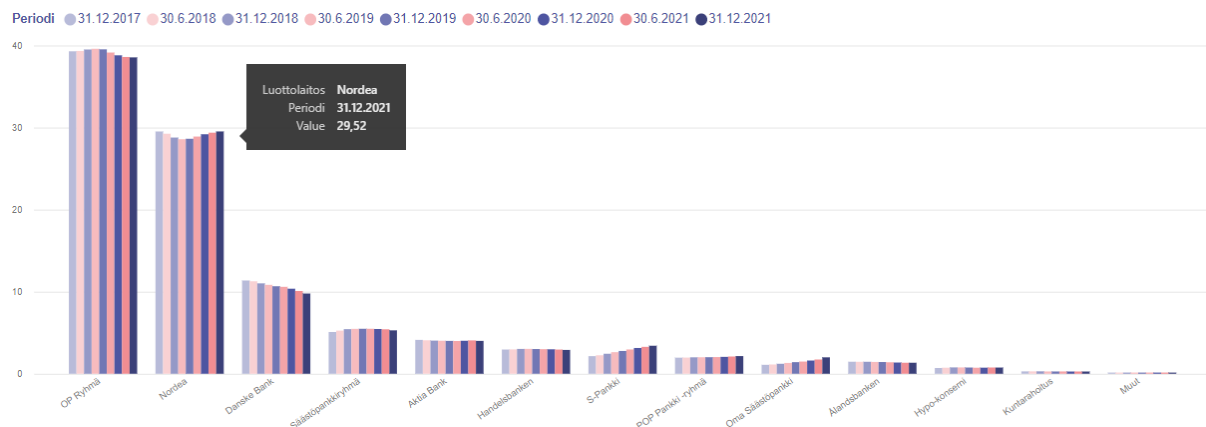


Figure 5.7: The mortgage market shares of Finnish creditors over recent nine years illustrated by a clustered column chart.

D12: Maturities of Housing-related Loans

The dashboard sheet *D12* covers the average maturities of housing-related loans in Finland (see the Figure 5.8). To illustrate the development of these average housing loan maturities over time, I used a small multiples encoding which consisted of cropped individual y-axes for each multiple as well as a shared x-axis (more about small multiples technique in the Subsection 4.1.2). Because different loan types, for example, consumer credits, corporate loans and mortgages, tend to have a high divergence of maturities, the maturity time series of different loan types are not very suitable to display together in a single figure. Since the time series are quite far away from each others, the value range of the shared figure (y-axis) would be forced to be so extensive that the ability to perceive any alteration from an individual time series could almost vanish. Beside clarity, my second ambition was to answer to the research question Q_2 (see the Section 2.4), which calls for space-efficient solutions to visualize time series data. Therefore, I ended up using a small multiples technique by stacking multiple maturity line charts one on the other, and utilizing a shared x-axis for them all. I also cropped the value axes (y-axes) of these line charts to

start above the zero point to both enhance their resolution and to save serious amount of idle space. This dashboard sheet also included multiple interactive elements to compress information: (1) an interactive time slider in order to adjust the start and end points of the visible time line, (2) a filtering menu concerning different moving average calculations, (3) multiple interactive key performance indicators aligned with the related maturity line charts, and referenced to them by using the same colors as the lines.

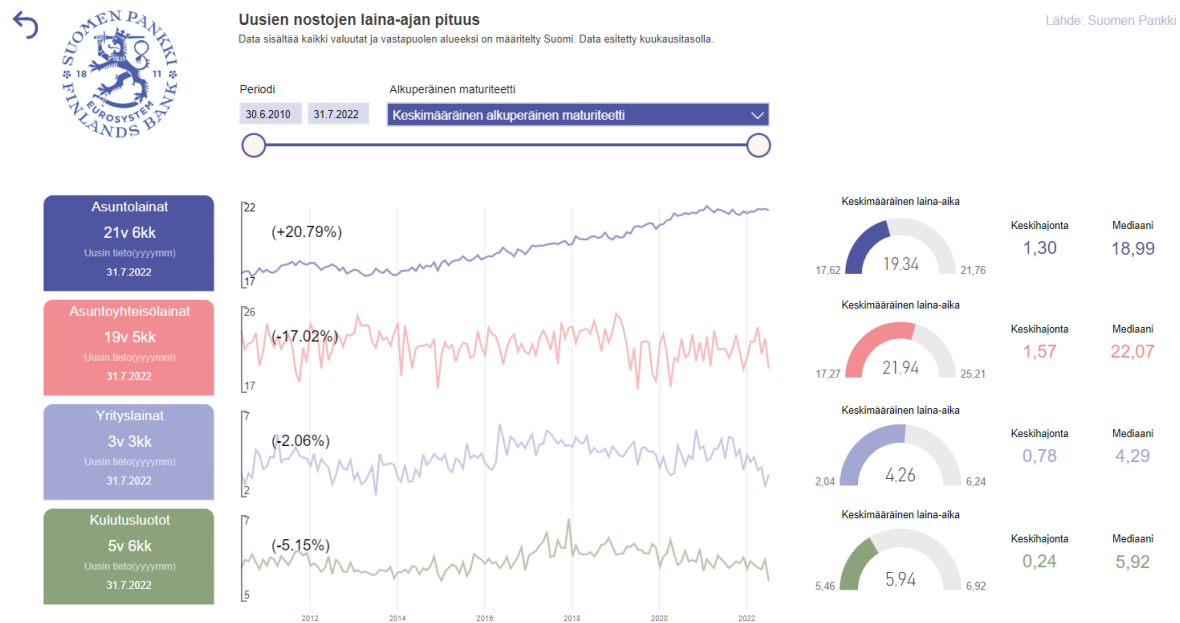


Figure 5.8: The average maturities of housing-related loans in Finland illustrated by using a small multiples encoding which consisted of cropped individual y-axes for each multiple as well as a shared x-axis.

D13: Finnish Domestic Loans in Respect of the Most Essential Loan Variables

Analogously to the dashboard sheets *D3* and *D6*, the dashboard sheet *D13* utilizes traditional line charts due to many strict expectations and practical requirements originating from the end users (see the answers of the preliminary end user survey in the Section 2.5 and the Figure A.15 in the Appendix A section). The principal aim of this dashboard sheet was to enable an effortless comparison between the most essential loan variables of the Finnish domestic loans over time. The dashboard includes three charts each of which includes multiple lines representing individual time series of domestic loan variables. Nonetheless, there are utilized multiple interactive elements in this dashboard sheet to compress information: (1) Each chart includes a zoomable y-axis to enhance its resolution during a definite analysis. (2) The dashboard sheet also includes two dropdown menus for various hierarchical domestic loan variables. (3) In addition, the dashboard consists

of an interactive time slider in order to adjust the start and end points of the visible time line.

5.2.3 Domestic Indebtedness Visualization

Since domestic indebtedness is strongly associated with housing related-loans, it was suggested by the economists of the Bank of Finland that the analysis tool should include also a dashboard sheet about domestic indebtedness. The domestic indebtedness dashboard sheet combines information about the income of Finnish households in comparison to their debts and interest charges. The dashboard sheet also covers the numbers of different debtor types concerning Finnish households.

D14: Domestic Indebtedness and Income

The dashboard sheet *D14* illustrates domestic indebtedness as well as domestic income by three stacked column charts and one clustered column chart (see the Figure 5.9). The stacked column charts are each combined with an income line chart or a trend line chart, which share the same value axes (y-axes) than their host charts. Thus, these three figures represent the same visualization type as the combination of the stacked column chart and line chart in the dashboard sheet *D5*. Likewise, the clustered column chart technique is quite analogous to the clustered column chart exploited in the dashboard sheet *D11*, and therefore I decided to omit its repetitive-like inspection here. In addition, this dashboard sheet includes multiple interactive elements to compress information: (1) Each figure has an interactive time slider in order to adjust the start and end points of the visible time line. (2) The dashboard sheet has also a filtering menu concerning the itemization of the different domestic indebtedness classes.

5.2.4 Finnish Economy Barometer Visualization

The final set of dashboards covers the evolution of Finnish economy barometer poll answers. These answers represent Finnish household's expectations and fears, trust and distrust, towards different economic aspects. In addition, these barometers examine how strong intentions Finnish households have towards different types of purchases or mortgages. Since the economic expectations, trust, and intentions of households reflect strongly to the activity of housing and housing-related loan market, the economists of the Bank of Finland desired that the analysis tool should take economic barometers extensively into account. Therefore, the analysis tool ended up including two dashboards devoted to the dozens of time series about different economic barometers.

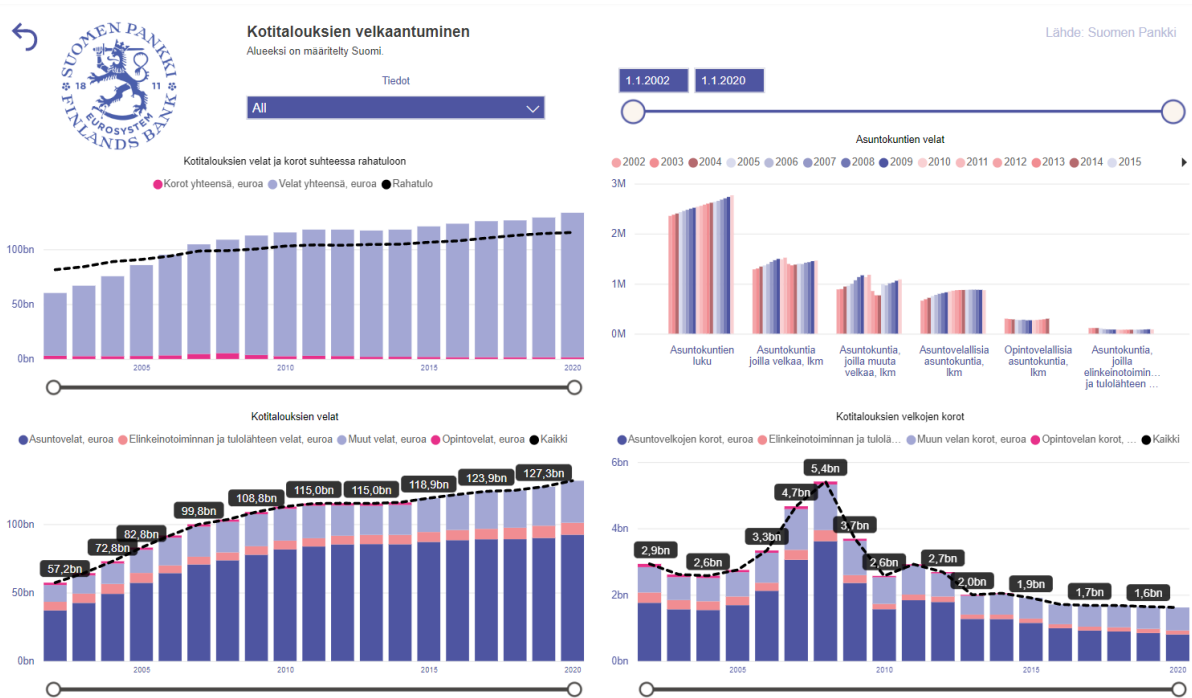


Figure 5.9: The domestic indebtedness and as domestic income in Finland illustrated by three stacked column charts combined with line charts and one clustered column chart.

D15: Atmosphere Changes of the Economic Consumer Barometers

The dashboard sheet *D15* covers the atmosphere changes of the economic consumer barometers over time by using a calendar heat map matrix and two conditionally colored column charts (see the Figure 5.10). Analogously to the calendar heat map matrix used in the dashboard sheet *D1*, the dashboard sheet *D15* utilizes the same visualization encoding to illustrate these atmosphere changes of the economic consumer barometers as a color-coded calendar matrix, in which rows represent years and columns represent months. Also here, the used color scale from the blue hues representing negative values to the red hues representing positive values is reasoned by their suitability to color blind users, and their intuitive and widely acknowledged connotation with negative and positive temperatures. The similar color-coding is utilized in the conditionally colored column charts, in which the columns are attached to the zero axis, and stick out either to negative or positive direction. In addition, this dashboard sheet includes two interactive elements to compress information: (1) an interactive time slider in order to adjust the start and end points of the visible time line. (2) a filtering menu concerning the itemization of the different economic consumer barometers classes and their sub-classes.

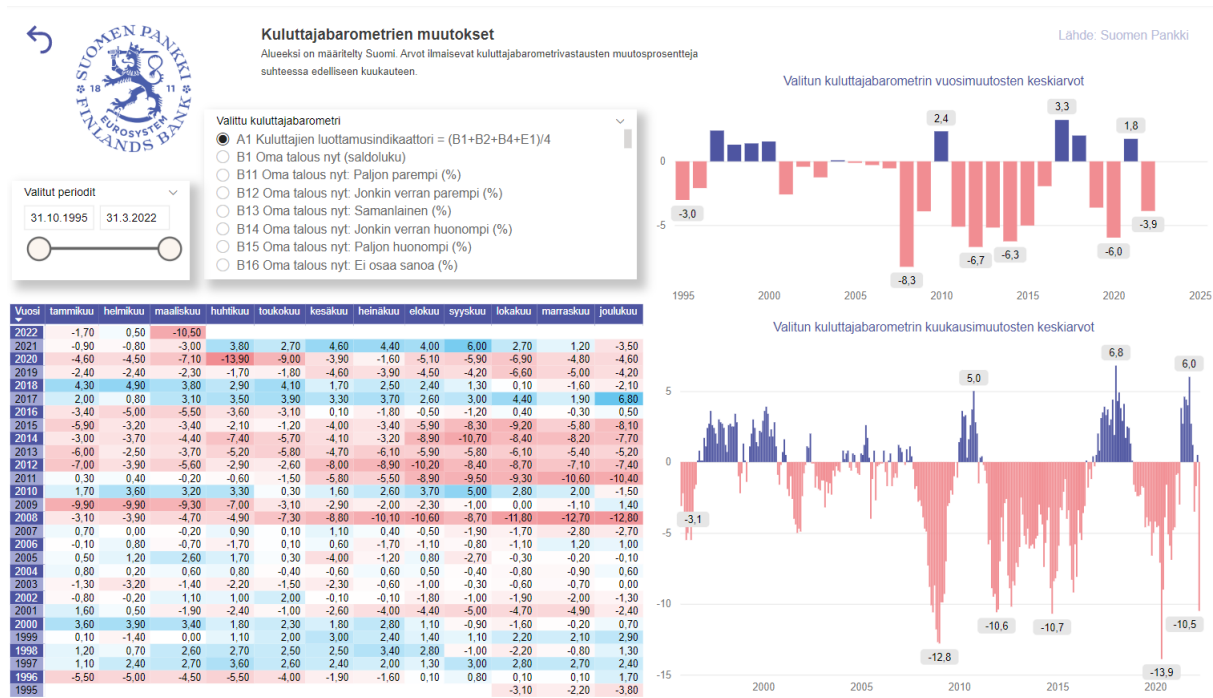


Figure 5.10: The atmosphere changes of the economic consumer barometers over time illustrated by using a calendar heatmap matrix and two conditionally colored column charts.

D16: Comparison between the Atmosphere Development of the Economic Consumer Barometers and the Actual Economic Development

The dashboard sheet *D16* provides a solution to reveal correlations between Finnish mortgage market barometers and the actual mortgage and housing market statistics. This dashboard includes only one figure, which I call as a depth-wisely stacked silhouette chart (see the Figure 5.11 below and the Figure A.18 in the Appendix A section). The silhouette chart resemble the color-partitioned line chart of the dashboard sheet *D2* without its characteristic partition. Giving more thorough definition, the silhouette chart is like a traditional line chart which consists of a colored silhouette between its trend line and the zero axis. Moreover, these silhouettes are designed to be stacked depth-wisely if multiple of them end up intersecting in the same figure. Using multiple silhouette charts together enables easy comparison between the atmosphere development time series of the economic consumer barometers and the actual economic development, and thereby help to reveal correlations between them too, because the colored silhouettes of the charts speed up perceiving the magnitudes of differences as well as make easier to keep track of an individual time series trends. Also, by increasing the transparency ratio of the colored silhouettes, multiple time series were able to display simultaneously without making them undesirably to disappear under another silhouette being stacked depth-wisely higher (for comparison, see the Figure 4.4). However, this visualization encoding endures only a few trend lines

at the same time before it begins to be overly cluttered to examine sensibly.

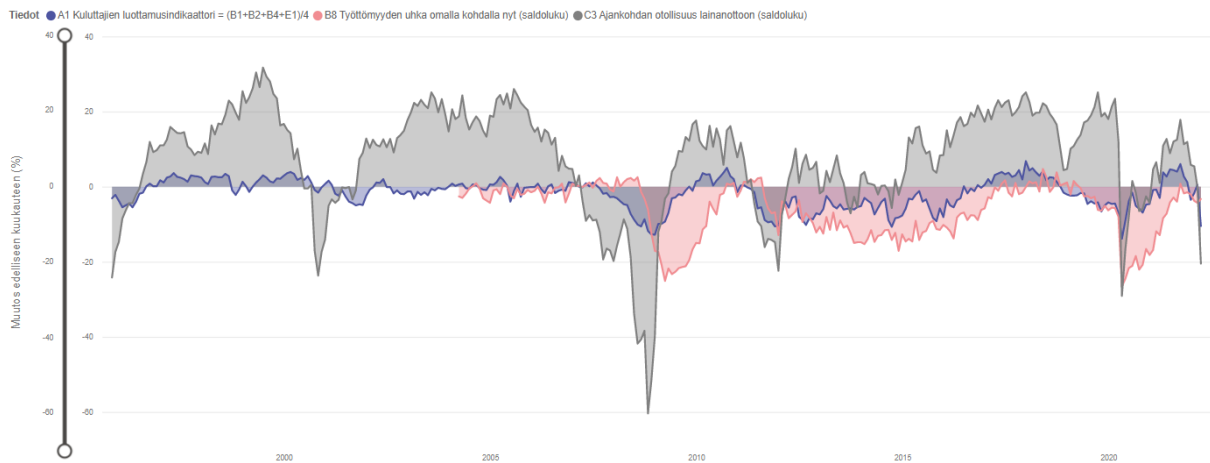


Figure 5.11: Using multiple silhouette charts together enables easy comparison between the atmosphere development time series of the economic consumer barometers and the actual economic development, and thereby help to reveal correlations between them as well.

6. Final Results

The overall ambition of this master’s thesis was to implement interactive time series visualization of Finnish housing and mortgage markets for central banking analysis. Since the subject of this visualization project was very broad and dispersed, I ended up constructing 16 independent dashboard sheets, and altogether 33 time series related visualizations to cover all desired macroeconomic substances around the target topic (see the whole substance list in the Section 2.3 and the dashboard sheet coverage in the Section 5.2). Overall, I utilized 15 different visualization encoding techniques for time series data. The final results achieved a lot of positive feedback from the economists of the Bank of Finland, and the visualization ensemble was put to actual use in-house as a macroeconomic analysis tool. The tool is described to be very comprehensive and adequate support for the analysis tasks regarding Finnish housing and mortgage markets.

6.1 Interactive Housing and Mortgage Market Analysis Tool

For this visualization project, I queried and combined information from six different data sources; the majority of data came from certain data bases of the Bank of Finland, and the rest of it were downloaded from the open source web tables of Statistics Finland. Formulating workable and reliable queries was not the easiest task to do, due to the Bank of Finland’s complex but comprehensive data infrastructures and multiple system security related restrictions. However, great thanks to my colleagues in the Bank of Finland for their help and support, I fortunately managed to figure out how all essential pieces worked together regarding the data querying phase of the project. The most notable accomplishment of the data querying part was to build the whole querying process as an updatable pipeline, which is both simple to maintain and easy to expand when necessary.

As a front end of that data pipeline functions an encompassing dashboard collection, which consists of 16 different dashboard sheets and 33 individual time series

visualizations (covered in the Section 5.2) regarding the housing and mortgage markets in Finland. Obviously, these dashboard sheets also include many other minor visual elements to display data, such as, pie charts, various gauges and other key performance indicators. However, these minor elements have not any direct time relation, and therefore I decided to omit their examination from this time series related thesis. Since this dashboard collection were conceived as a macroeconomic analysis tool for the economists of the Bank of Finland, its main intention was to provide a comprehensive and versatile framework to inspect the Finnish housing and mortgage markets. Therefore, each dashboard sheet is designed to cover an individual viewpoint of the topic. All viewpoints included in the final dashboard collection, represent the preliminary desires of the analysis tool's actual end users; Before the visible implementation process of this visualization project was started, I arranged multiple meetings, where I discussed about the contents, structural elements, functionalities and details of the desired outcome. In addition, I organized a preliminary end user survey by which I managed to collect even more profound needs, thoughts, ideas and viewpoints about the final product.

In this dashboard collection, there are used altogether 15 different visualization encoding techniques that take the time dimension of the time series data into account. These encoding techniques can be classified into four different classes: (1) line and bar chart variations, (2) heat map matrices, (3) scatter plots, and (4) round charts. Next, I will discuss about each of these classes, and describe the end user reactions concerning the interactive visualizations of the analysis tool.

The class of the line and bar chart variations consist of the most extensive and miscellaneous set of different encoding techniques: traditional line charts, small multiples of line charts, line charts having colored silhouettes underneath their trend lines, stacked bar charts combined with traditional line charts, clustered bar charts, conditionally colored bar charts, a waterfall chart, a ribbon chart, and a horizon chart. In respect of time series, all these encoding techniques utilize the conventional coordinate axes in which the horizontal axis represents time and and the vertical axis represents the values of data points. Thus, among line and bar chart variations, coordinate locations are alone a sufficient method to define time and value dimensions of displayed data points. However, the other design choices, such as, colors, shapes, layouts, encoded functions, and interactions, give these visualization types their true characters. Even though traditional line and bar charts are undeniably a valid way to illustrate time series data, especially in the macroeconomic context, the introduced encoding variations of them are very apt either to saturate visualizations with a lot of additional information, or to enhance pattern and trend perceiving, or just to compress information into a more

space-efficient form. For example, horizon graphs are significantly more space-efficient than traditional line charts, and ribbon charts are far more suitable for the volume comparison between individual data points than ordinary stacked bar charts.

From the end user perspective, the most interesting and discussed line and bar chart variations were definitely horizon charts, ribbon charts, and those charts that utilize line and bar chart encodings simultaneously. Horizon charts were seen beneficial for macroeconomic analysis, since they are able to compress information considerably and, while being arranged as small multiples, a proper selection of horizon charts could be effectively used for pattern recognition. Pattern recognition benefits were advocated based on horizon charts' special feature to pack data points into vertically equal sized space and still maintain their commensurate scale by exploiting color-coding and depth-wise stacking techniques. This stacking technique enables easier comparison between stacked time series, and therefore it enhances pattern recognition. However, horizon charts were seen quite difficult to use and perceive in the first place. Even though the end users were very receptive to try and utilize horizon graphs from now on both as the part of the analysis tool and likewise elsewhere if proper use cases would emerge.

Ribbon chart technique were seen as compact and useful visualization concept for financial analysis, since, in this technique, the data points of different time series were visually aggregated as total bars over time as in traditional stacked bar charts which are already commonly used visualization technique in the Bank of Finland. Nonetheless, the magnitude-ranking-based stacking order of the data points of different time series was a very appreciated feature, since it enables more effortless comparison between the data points of different time series at a specific time point. This integrated ranking procedure is one of the most characteristic feature of ribbon charts, and therefore it makes a big difference to traditional stacked bar charts. Moreover, the ribbon chart visualization consist of transition lines between the data points of adjacent bars that are related to the same time series. These transition lines reveal more information about time series related changes, such as ranking-based, relative and absolute changes, between two time points after being hovered with a cursor. Also this feature was a remarkable benefit in this visualization technique according the end users. However, ribbon chart technique were also noted to be a little confusing and complicated to use without previous practice, and overly-cluttered if the number of included time series was too high or the variance of the rankings between the data points of different time series.

Line chart and bar chart combinations were also seen as useful method to compress data without losing too much clarity. The analysis tool included two dashboard

sheets which utilized line and bar chart encodings simultaneously in a single figure. In those visualizations, lines were used to illustrate average trend development whereas the bars represented aggregated totals of inspected subjects. In the other dashboard sheet, there were also utilized two separate value axes in the same figure in order to map two incommensurate time series over the same time axis. The axis that associated with the average trend development line had the same color as the line as a reference. The presented combination technique was noted to be quite straightforward and trouble-free to use in both financial and macroeconomic contexts. However, while utilizing two separate value axes in the same figure, there is always a small risk of misunderstanding or oversight an additional value axis and end up with a misconception.

The class of the heat map matrices consist of traditional heat map matrices and calendar heat map matrices. Heat map matrices are used very similarly to traditional matrices. Nonetheless, the biggest difference between these two is the magnitude-based color coding of heat map matrices (see more about heat maps in the Subsection 4.1.8). In many data visualization related studies, colors are proved to improve and speed up perceiving differences, changes, magnitudes and patterns in data. However, that how much colors actually increase distinctiveness depends on a viewer. The first issue concerns that people may have different connotations regarding colors, for example, does red represent growth due to its relation of weather forecasts' higher temperatures, or does it rather represent decrease due to its role as a warning color. The second issue is color blindness and its different types. This is a real issue, since it is commonly known that approximately 8% of men and 0.5% of women worldwide suffer some type of color blindness. Therefore it is recommended to always use color-blind-friendly colors only in data visualizations, if possible. However, on the request of the Bank of Finland, there is only used the color theme of the Bank of Finland in the analysis tool's visualizations. Obviously, this regulation limits coloring options in respect of the issues explicated above. I ended up using reddish and blueish hues in the heat map matrices, because (1) these colors tend to have the biggest mutual contrast regarding the most types of color blindness, and (2) these colors encompass a strong connotation concerning positive and negative values, since they are generally used to express temperatures (on the Celsius scale) based on conceptual color warmth.

Matrices are very applicable and conventional structures to display time series more effectively. Matrices enable to view multiple unattached time series together concurrently, and there are several techniques to express time in them, for example, time can be presented via rows, columns or both. The last option represents calendar matrices, in which two nested layers of a time hierarchy form a chronological grid for the data

point values of time series. Typical examples of calendar matrices in the finance sector are the value grids where rows represent years and columns represent either months or quarters, and vice versa. The calendar heat map matrices of the implemented analysis tool utilize this above mentioned calendar matrix technique. In addition to this calendar structure, the calendar heat map matrices consist of a heat map encoding based on the magnitudes of their grid values. This special feature enhance the intuitive perceiving of data, since the biggest trends of seasonal changes can be distinguished at first glance. The end users of the analysis tool were also very pleased about this feature. The heat map encoding both combined with the calendar structure and without it was seen to ease reading, analysing and searching relevant grid values significantly. Additional colors were also regarded as a feature that makes visualizations more engrossing and maintain viewers' interest longer. However, the matrix-based visualizations were not seen to be the most efficient method to examine the long term changes of simultaneously extensive and fine-grained time series, because, in that case, either the elements of a grid would be unreadable-small, or all the elements of a grid would not fit in a screen at the same time.

The class of the scatter plots consist of two similar scatter plots, in which both axes represent different variables (price and market transactions), colors represent regional sub-classes, and the sizes of scatter dots represent time. The bigger the dot is the more recently it is recorded in comparison to other dots, and vice versa (see the Figure 5.4). The main purpose of these scatter plots were to give an overview about the distribution of the housing market in Finland. The secondary purpose of the scatter plots were to pilot a new visualization technique for revealing regional resemblances and differences based on prices and market volumes. The data set behind this visualization was not completed yet due to the reformed data pipe line of the third party organization, and therefore the scatter plots provided neither factual data values nor adequately structured region classes at the moment. However, the pilot seemed to succeed, since the end users described the new visualization technique as potential and innovative method to inspect the Finnish housing markets. The colored regional clusters were seen illustrative but the end users were quite unfamiliar with the idea of time-related dot sizes, even though the dot size feature was described ingenious on the other hand.

The class of the round charts consist of multiple radius charts utilized in the analysis tool (see the Figure 5.3). The radius charts utilize polar coordinate system like many other round charts that consist of at least two dimensions. In the polar coordinate system, the rotation angle around the central zero point represents horizontal distance (x-axis) in the traditional coordinate system whereas the radiuses represent vertical distance (y-axis) in the traditional coordinate system. For illustrating data points of

time series, the radius charts use ray-like bars, which burst from the central zero point of a chart towards the perimeter of a chart thereby indicating data point values. The radius charts resemble sunburst charts visually a lot, but in the radius charts the rotation angle of bars represents another dimension (in our case time) in contrast to sunburst chart, in which the mutual positions of bars does not imply anything else but the sorting order of bars. The analysis tool includes two kind of radius charts: (1) the one that has individual color references to time units (years), and (2) the one that assort even and odd years by using different colors (red and blue), and even and odd months/quarters/half-years instead by using different levels of brightness (bright and pale).

The end users became interested about the visual appearance of the radius charts, but the visualization technique were seen more challenging to use in the actual analysis work than the other introduced techniques. However, the radius charts were discussed to be very prominent and presentable method to illustrate time series data. The distinctiveness of the radius charts were conceived a beneficial feature regarding especially the in-house and other organizational presentations of the Bank of Finland, which may be occasionally quite information-dense and ponderous to follow nonstop without any stimuli. Originally, the main intention of the radius charts was to draw attention to the core subject by the extraordinary and stimulating time series layout besides the traditional line chart version of the same data designed for actual analysis use. Thus, the radius chart pilot was a successful pilot as a new and innovative visualization technique.

The visualizations were embedded into the 16 interactive dashboard sheets which all were connected to the navigation page of the analysis tool via designated shortcut buttons. The dashboard sheets included either multiple interactive filtering menus, variable menus, granularity menus, time sliders, zoomable value axes, or key performance indicators which each were connected to every visualization (apart from a few exceptions in some dashboard sheets) inside a single dashboard. The end users described interactions as a user friendly and quick method to tune visualizations suitable for definite macroeconomic analysis. Especially, time sliders and different regional filters were seen very useful supporting features for analysis. Interactions were also noted to compress information significantly, since all possible variables and granularity states were hidden under several filtering menus rather than all being visible simultaneously. In addition, the navigation page and the shortcut buttons of each dashboard sheet were seen as a very convenient and agile solution to shift between different dashboard sheets.

6.2 Final End User Survey

After I had managed to get the whole visualization ensemble finished, I arranged the final end user survey to collect feedback, thoughts and criticism about the new macroeconomic analysis tool in action. Since this tool was implemented to support the economists of the Bank of Finland in their housing and mortgage market related analysis tasks, I wanted to gather feedback especially about their user experiences. For this, I decided to use the System Usability Scale (SUS), introduced by John Brooke in 1995 [7], which included ten descriptive statements about a system that can be evaluated by picking responses from a five-step agreement scale; from strongly agree to strongly disagree. The SUS framework was widely argued to be comprehensive and adequate method for examining the usability of a new system, even though it is very time-efficient to arrange [28]. In addition, the SUS surveys can be used on small sample sizes with reliable results. Therefore, the SUS survey was suitable to carry out during the busy working hours of the Bank of Finland's economists. I managed to get six economists to enroll in the SUS survey that was quite analogous with the original SUS framework. Only exception was the optional extra question, in which the respondents was requested to give additional feedback freely.

The SUS statements used in the arranged survey were the followings: (1) I think that I would like to use this system frequently. (2) I found the system unnecessarily complex. (3) I thought the system was easy to use. (4) I think that I would need the support of a technical person to be able to use this system. (5) I found the various functions in this system were well integrated. (6) I thought there was too much inconsistency in this system. (7) I would imagine that most people would learn to use this system very quickly. (8) I found the system very cumbersome to use. (9) I felt very confident using the system. (10) I needed to learn a lot of things before I could get going with this system.

As a summary, the survey's results were excellent and the analysis tool achieved a lot of positive feedback. However, there were also suggested some minor future improvements as well as larger content related extensions in the section of free feedback. These improvements and the plans concerning the future work will be discussed more encompassing in the Section 7.2.

In the arranged SUS survey, respondents were required to choose the response option that matches most to their first reaction to each item, rather than considering any answer for long. The responses of the SUS survey were scored by using the same method introduced in the paper [28]. Analogous to the paper, all scores of the individual response form were summed together by using the following methodology: Each item

had a score range from 0 to 4 points. For the items 1, 3, 5, 7, and 9 the 'strongly agree' responses output 4 points and the 'strongly disagree' responses output 0 points. For the items 2, 4, 6, 8 and 10, the scoring rules are reversed, that is to say, the 'strongly agree' responses output this time 0 points and the 'strongly disagree' responses output 4 points. After all the scores of an individual response form were summed, the sum were furthermore multiplied by 2.5 to obtain the final score of an individual response form. Lastly, the average score of these individual response forms was calculated. This average score represents the actual SUS score of the tested system. The SUS scores have a range from 0 to 100.

The implemented analysis tool concerning the interactively visualized housing and mortgage markets in Finland achieved the SUS score of 90 points. According the article [5], the achieved score is placed in the "best imaginable" SUS score class in which less than 1,7% SUS scores were placed among 959 individual studies. However, in our case, the number of respondents was quite low, which may distort the achieved scores even though the SUS framework were defined as suitable and reliable method to use for small sample sizes as well. At least, it seems that the end users of the finished analysis tool were satisfied with the final outcome of the interactive visualization ensemble and overall results of my work. The analysis tool were described, inter alia, as "...very comprehensive, innovative and versatile tool, which emphasizes the significance of data visualization", and as "...very useful tool for analysing housing markets in Finland".

7. Conclusions and Future Work

The main objective of this thesis was to implement an interactive time series visualization of Finnish housing and mortgage markets for macroeconomic analysis, especially for central banking related analysis. This objective was accomplished during the autumn 2022, and the final results were introduced by me in the Bank of Finland's specialist seminar on the 2nd of November. According to the feedback from the end users, the final outcome of the visualization project corresponded to its principal purpose as a macroeconomic analysis tool laudably. Moreover, the results provided adequate answers to the initial research questions of the thesis. However, the feedback included also a few development suggestions and improvement ideas added by the end users that are discussed in the last section, 7.2 Future Work.

7.1 Conclusions

Before the actual implementing process began, I had defined four research questions regarding this objective: (Q_1) How to provide an informative and handy analysis tool for the economists of the Bank of Finland using interactive time series visualization techniques? (Q_2) What are the best ways to compress a lot of time series data concerning housing market and housing-related loan market into a single presentation? (Q_3) What would be new and modern ways to visualize time series data, especially housing market and housing-related loan market data, which the Bank of Finland would consider to exploit broadly in future? Since these research questions formed the base for the whole thesis, and thus defined the main focuses of the visualization assignment, I will start to review the visualization project based on them.

The answer of the first research question was the hardest one to achieve, since the end users of the analysis tool consisted of economists who valued macroeconomic substance the most in the tool and were not very experienced with modern data visualization frameworks or interactive user interfaces. On the other hand, my specialty focused rather on data processing and visualization techniques than macroeconomics or other advanced finance matters. Therefore, I decided to arrange a quite thorough preliminary

survey for the end users to examine the real interests, expectations, and necessities. By these survey answers, I also tried to figure out the end users' analysis work related customs and other work flow aspects regarding visual data interfaces. Furthermore, the economists of the Bank of Finland were very helpful and cooperative with every finance and macroeconomic issues I had during the entire implementation process of the analysis tool. By the help of the preliminary survey's answers and the collaboration with the economists, I was able to collect together all necessary data and refine it so that the final data visualization ensemble served its purpose as macroeconomic analysis tool adequately. The refining stage consisted of fundamental calculations and categorical aggregations, as well as multiple data wrangling and parsing steps. The user interface including all interactions in it was implemented by following the Bank of Finland's guidelines as well as the desires of the end users. To keep track all intermediate suggestions and opinions, I ended up arranging several meetings with the end users where we negotiated about different options regarding either the user interface or visualization techniques.

To discover the best ways to compress time series data in the housing and mortgage market context, I explored and studied the previous work concerning different variations of time series visualizations and overall visualization process. Moreover, I studied and examined how to integrate different types of interactions into visualizations, since interactions can be used to compress information very effectively. Thus, I ended up using many different interaction techniques diversely in the analysis tool. Furthermore, to ensure that the final visualizations correspond properly the fundamental domain aspects, I followed Tamara Munzner's nested visualization process model [25] during the whole implementation process of the visualization ensemble. I also utilized the results of the preliminary end user survey and additional discussions with the end users to define the most adequate and useful visualization techniques for the final analysis tool.

Overall, I ended up utilizing 15 different visualization encoding techniques for time series data in the final analysis tool. Those 15 visualization techniques were the ones that both I and the end users regarded the most suitable for the analysis tool. To settle upon these visualization choices, I studied and examined many different ways to visualize time series data; new and modern techniques as well as more traditional and conventional ones. After ascertaining several potential visualization techniques for the analysis tool, I discussed with the end users about their preferences in respect of the visualization alternatives, especially from the perspective of the macroeconomic analysis. My second ambition was to propose such visualization techniques that would also enhance analysis work of the other research fields in the Bank of Finland as well.

Since the most of the economists were familiar only with few traditional visualization techniques, such as the most common line and bar chart variations, any more advanced techniques were seen as modern and new. Therefore, many proposed techniques concerned the end users due to their unfamiliarity. However, my proposal also included many visualization techniques that satisfied the end users so thoroughly that I was asked to introduce these new techniques for larger audience in the Bank of Finland's specialist seminar in the autumn 2022. According the end users and the seminar audience, horizon graph, ribbon chart and waterfall chart were very effective visualization techniques. Moreover, radius charts were described to be a prominent and presentable technique to illustrate time series data. However, radius charts were not seen as practical in actual analysis use as traditional line and bar charts, because many analytical aspects were regarded to be harder to perceive from radius charts. After all, the end users and the seminar audience considered that the most adequate and functional visualization techniques of all presented techniques were different line and bar chart variations, and heat map matrices. Nevertheless, all presented visualization approaches drew genuine interest, and many of them were suggested to be exploited actively in the Bank of Finland's other prospective visualization tasks as well.

All in all, the analysis tool fulfilled the expectations and needs of the end users comprehensively. According the final feedback, visualizations corresponded to their actual purpose providing significant help and support for macroeconomic analysis tasks concerning the Finnish mortgage and housing markets. However, the expected outcome of the analysis tool turned out to be a lot more extensive than I imagined in the beginning of the thesis. The actual implementation stage of the tool revealed for me that both the unfamiliar macroeconomic substance and the scattered and enormous data sources slowed down the encoding process of visualizations including their interactive user interfaces. Thus, the extensive size of the tool both engaged a lot of scarce working hours and exposed me into a great many inconveniences during the process. However, the final analysis tool ended up consisting of all wanted information in it by utilizing sixteen linked dashboard sheets, which each included multiple filtering options to compress data. These filters concerned different time series variables such as time, region and data content class. The end users told that they were very satisfied with the interactive user interfaces of the analysis tool, since these consisted of various options to filter visualized time series. Comprehensive filtering options were described to be a crucial feature while doing any serious macroeconomic analysis. Especially the interactions regarding either time or data content settings were seen highly important for analysis tasks. Thus, time sliders and filtering menus were considered equally essential interaction types in the visualizations of the analysis tool.

Even though the whole visualization project was successful since it managed to answer all research questions adequately and the end users were very satisfied with the final outcome as the excellent SUS score of the final end user survey proves, there were also proposed some improvements and expansions by the end users. These aspects are discussed in the Section 7.2, Future work.

7.2 Future Work

Although the final version of the analysis tool was acknowledged as a complete and successful outcome, the end users still pointed out some improvements and expansions concerning it in the final feedback. In fact, the following two development proposals alone came out several times in the feedback: (1) The end users suggested that the analysis tool should include at least one visualization about both investment property loans and summerhouse loans. (2) The end users considered that, in addition to housing related loans, the analysis tool could also cover other household loan types, such as overdrafts, credit card credits, and consumer credits. Thus, the first targets of the future work would concern the above-mentioned expansion proposals 1 and 2.

Moreover, there was a discussion about how modern and up-to-date time series visualization techniques should be regularised in the financial sector of Finland in order to utilize their effectivity, space efficiency, and other context-related benefits. During that discussion, there was also pointed out that it would be crucial to familiarize especially non-technical persons for new visualization techniques and possible interactive elements integrated into those, even though they would need more time, support, and effort to adopt new methods than more technically oriented people, because the majority of target audience tend to usually be non-technical people that might be clueless with the new kind of visualizations. There is a risk of that the whole organization may become stuck with simple, ineffective and inefficient visualization techniques if the majority of the end users keep on being not convinced about other options. Therefore, the second target of the future work could be drawing up encompassing user manual with demonstrative pictures and some clear usage examples.

Bibliography

- [1] W. Aigner, A. Bertone, S. Miksch, C. Tominski, and H. Schumann. Towards a conceptual framework for visual analytics of time and time-oriented data. In *2007 Winter Simulation Conference*, pages 721–729, 2007.
- [2] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski. Visualizing time-oriented data—a systematic view. *Computers & Graphics*, 31(3):401–409, 2007.
- [3] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski. Visual methods for analyzing time-oriented data. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):47–60, 01-02 2008.
- [4] X. Bai, D. White, and D. Sundaram. Context adaptive visualization for effective business intelligence. In *2013 15th IEEE International Conference on Communication Technology*, pages 786–790, 2013.
- [5] A. Bangor, P. Kortum, and J. Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4:114–123, 03 2009.
- [6] I. Beroš, N. Hlupič, and F. Džapo. Usage of bi tools in analysis of the tomas croatia 2019 survey. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pages 1335–1340, 2021.
- [7] J. Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.
- [8] S. Card. *Information Visualization*, pages 509–543. 01 2008.
- [9] M. Correll, D. Albers Szafir, S. Franconeri, and M. Gleicher. Comparing averages in time series data. pages 1095–1104, 05 2012.
- [10] B. P. Duba. The use of visualization and analytic tools of power bi learning tasks, 2021.
- [11] G. Ellis and A. Dix. A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1216–1223, 2007.

-
- [12] Y. Fang, H. Xu, and J. Jiang. A survey of time series data visualization research. In *IOP Conference Series: Materials Science and Engineering*, volume 782, page 022013. IOP Publishing, 2020.
- [13] A. Gogolou, T. Tsandilas, T. Palpanas, and A. Bezerianos. Comparing similarity perception in time series visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):523–533, 2019.
- [14] S. Haroz, R. Kosara, and S. L. Franconeri. The connected scatterplot for presenting paired time series. *IEEE Transactions on Visualization and Computer Graphics*, 22(9):2174–2186, 2016.
- [15] J. Heer, M. Bostock, and V. Ogievetsky. A tour through the visualization zoo. *Commun. ACM*, 53:59–67, 06 2010.
- [16] J. Heer, N. Kong, and M. Agrawala. Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations. pages 1303–1312, 04 2009.
- [17] A. Jabbari, R. Blanch, and S. Dupuy-Chessa. Composite visual mapping for time series visualization. In *2018 IEEE Pacific Visualization Symposium (PacificVis)*, pages 116–124, 2018.
- [18] W. Javed, B. McDonnel, and N. Elmqvist. Graphical perception of multiple time series. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):927–934, 2010.
- [19] M. Jern. Visual intelligence—turning data into knowledge. In *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)*, pages 3–8, 1999.
- [20] V. S. Khatuwal and D. Puri. Business intelligence tools for dashboard development. In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, pages 128–131, 2022.
- [21] S. M. Kumar and M. Belwal. Performance dashboard: Cutting-edge business intelligence and data visualization. In *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, pages 1201–1207, 2017.
- [22] T. Lammarsch. A compound approach for interactive visualization of time-oriented data. In *2008 IEEE Symposium on Visual Analytics Science and Technology*, pages 177–178, 2008.

- [23] A. Lee, D. Archambault, and M. A. Nacenta. The effectiveness of interactive visualization techniques for time navigation of dynamic graphs on large displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):528–538, 2021.
- [24] Muller and Schumann. Visualization methods for time-dependent data - an overview. In *Proceedings of the 2003 Winter Simulation Conference, 2003.*, volume 1, pages 737–745 Vol.1, 2003.
- [25] T. Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009.
- [26] T. Munzner. *Visualization analysis and design*. CRC press, 2014.
- [27] R. Rad. *Online Book: Power BI From Rookie to Rockstar, Book 2: Visualization in Power BI, Part I: Visuals in Power BI, Ribbon Chart is the Next Generation of Stacked Column Chart*. 09 2017. <https://radacad.com/online-book-power-bi-from-rookie-to-rockstar>, Accessed on 22nd April 2022.
- [28] J. Sauro. Measuring usability with the system usability scale (sus), 03 2011. <https://measuringu.com/sus/>, Accessed on 11th November 2022.
- [29] H. Suematsu, S. Yagi, T. Itoh, Y. Motohashi, K. Aoki, and S. Morinaga. A heatmap-based time-varying multi-variate data visualization unifying numeric and categorical variables. In *2014 18th International Conference on Information Visualisation*, pages 84–87, 2014.
- [30] the Bank of Finland. Mfi balance sheet (loans and deposits) and interest rates, 08 2022. <https://www.suomenpankki.fi/en/Statistics/mfi-balance-sheet/>, Accessed on 18th August 2022.
- [31] A. Toic, P. Poscic, and D. Jaksic. Analysis of selected business intelligence data visualization tools. In *Central European Conference on Information and Intelligent Systems*, pages 25–32. Faculty of Organization and Informatics Varazdin, 2022.
- [32] J. van Wijk. The value of visualization. In *VIS 05. IEEE Visualization, 2005.*, pages 79–86, 2005.
- [33] S. Widjaja and T. Mauritsius. The development of performance dashboard visualization with power bi as platform. *International Journal of Mechanical Engineering and Technology (IJMET)*, 10:235–249, 05 2019.
- [34] W. E. Zhang, R. Chang, M. Zhu, and J. Zuo. Time series visualization and forecasting from australian building and construction statistics. *Applied Sciences*, 12(5):2420, 2022.

- [35] H. Ziegler, M. Jenny, T. Gruse, and D. A. Keim. Visual market sector analysis for financial time series data. In *2010 IEEE Symposium on Visual Analytics Science and Technology*, pages 83–90, 2010.
- [36] H. Ziegler, T. Nietzsche, and D. A. Keim. Visual exploration and discovery of atypical behavior in financial time series data using two-dimensional colormaps. In *2007 11th International Conference Information Visualization (IV '07)*, pages 308–315, 2007.

Appendix A. – Overview of the Final Analysis Tool



Asuntolainakehitys	Asuntolainamarkkinat ja asuntomarkkinakehitys	Asuntolainamarkkinat ja ostoaikomukset	Kotitalouslainat
Asuntomarkkinat	Maturiteetit	Markkinaosuudet	Kotitalouksien velkaantuminen
Omakotitalohintojen muutokset	Omakotitalomarkkinat alueittain	Omakotitalomarkkinat, hajontakuviot	Omakotitalomarkkinat, yhdistelmäkuviot
Omakotitalomarkkinat, taulukkona	Omakotitalohintojen kehitys	Kuluttajabarometrit	Kuluttajabarometrien rinnakkaisvertailu

Suomalaisten lainat ja talletukset

Visualisoinnit avautuvat yllä olevista [painikkeista](#). Painikkeiden alla olevat visualisoinnit vasemmalla alkavat karkealta tasolta siirtyen oikealle kohti pienempiä osakokonaisuuksia.

Data sisältää luottolaitoksilta kerättyjä tietoja lainoista ja talletuksista. **Lainojen ja talletuksien** osalta vastapuolena tarkastellaan **sektoreita** eli suomalaisia/Euroalueen yrityksiä, asuntoyhteisöjä, kotitalouksia ja kotitalouksia palvelevia voittoja tavoittelemattomia yhteisöjä. Markkinaosuuksien osalta vastapuolena on myös yleisö. Kokonaisuus sisältää kaikkien valuuttojen määräisinä myönnettyt lainat ja talletukset euroiksi muunnettuna. Tiedot ilmoitetaan miljoonina euroina.

Tietoja on saatavilla kannasta, uusista nostoista ja sopimuksista sekä sovitusta vuosikorosta. **Kotitalouksille** myönnettyt lainat on jaettu käyttötarkoituksittain. **Kulutustuotoista** on lisäksi tietoa vakuuden tyypistä. **Yrityslainojen** uudet nostot on jaettu kokoluokittain. Talletukset on jaettu määräaikais-, sijoitus- ja yon yli -talletuksiin.

Anna palautetta: Mitä muita tietoja olisi tarpeen lisätä? Miten käytettävyyttä voisi kehittää? Muuta? Sähköpostimme on [statistics\(at\)bof.fi](mailto:statistics(at)bof.fi)

Figure A.1: The starting page of the final analysis tool consists of sixteen labeled express transition buttons which each leads to a dashboard concerning its label-corresponding subject.

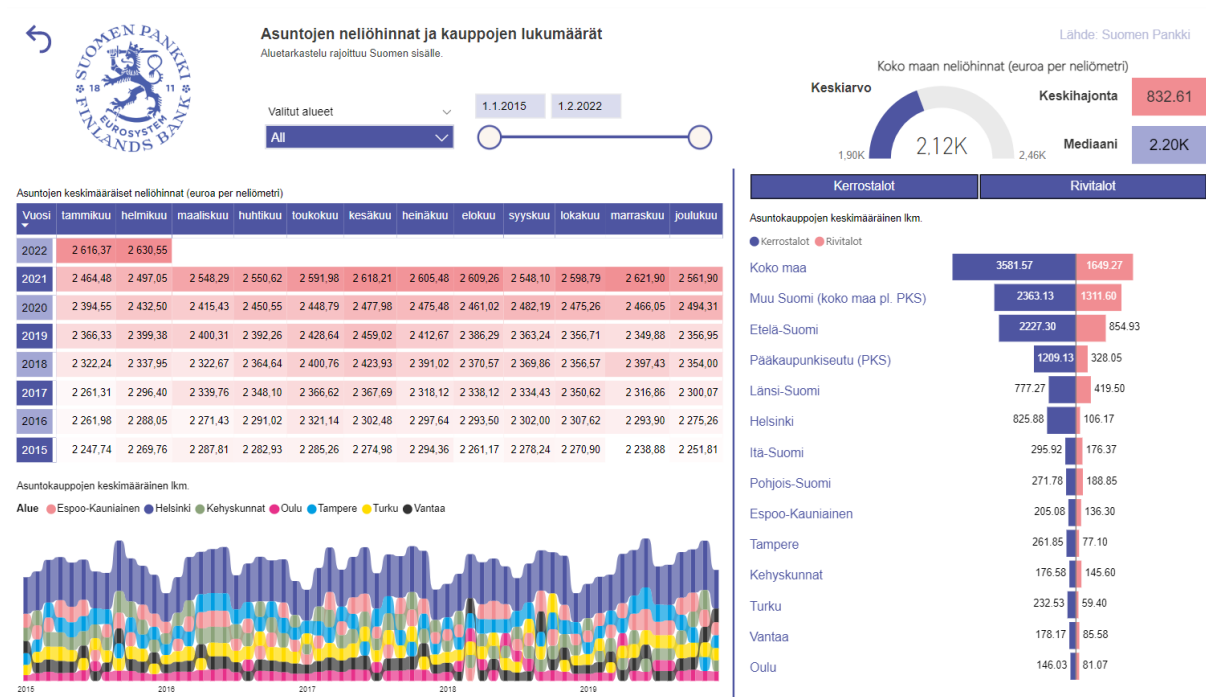


Figure A.2: The dashboard sheet D1 covers the regional dwelling prices per square meter and dwelling transactions.

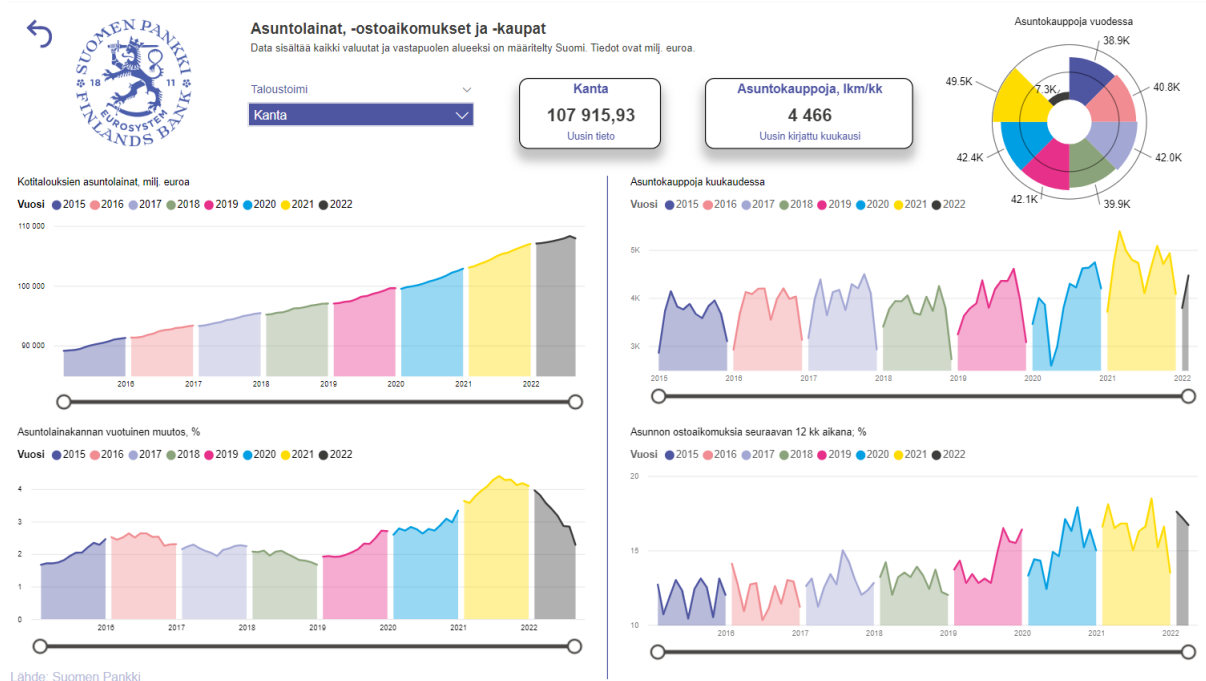


Figure A.3: The dashboard sheet D2 covers the mortgage variables in respect of dwelling purchase decisions and actual dwelling purchases in Finland

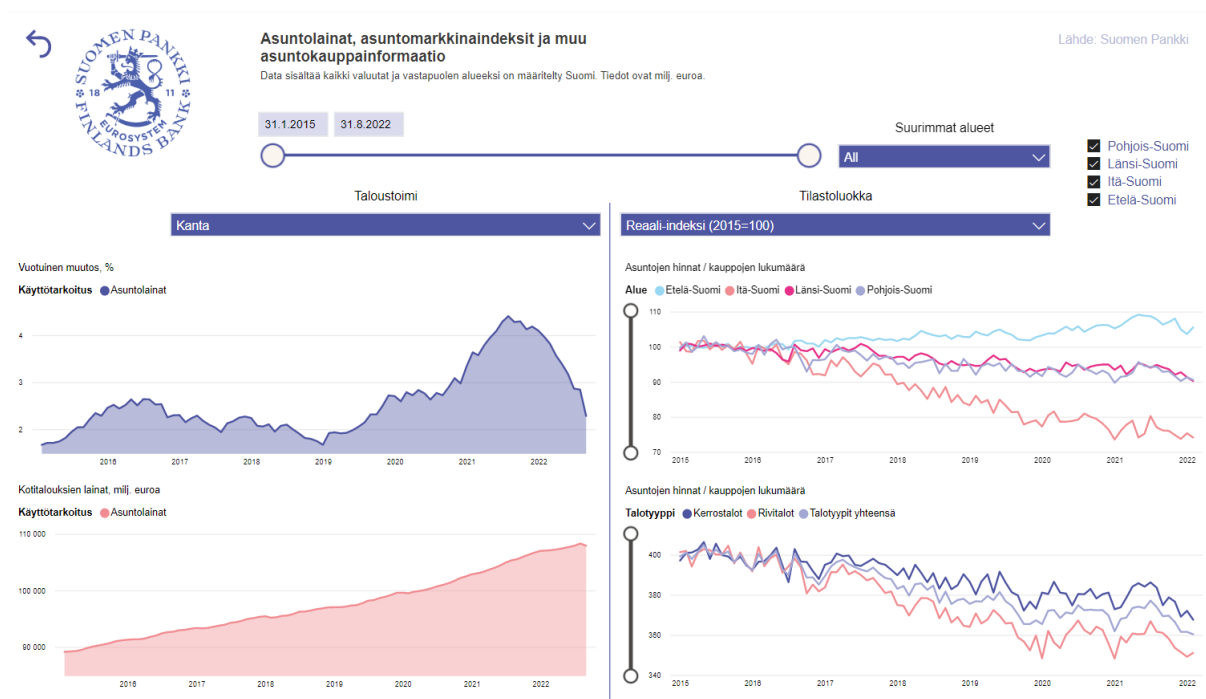


Figure A.4: The dashboard sheet D3 covers the mortgage variables in respect of regional housing market indices and housing market variables

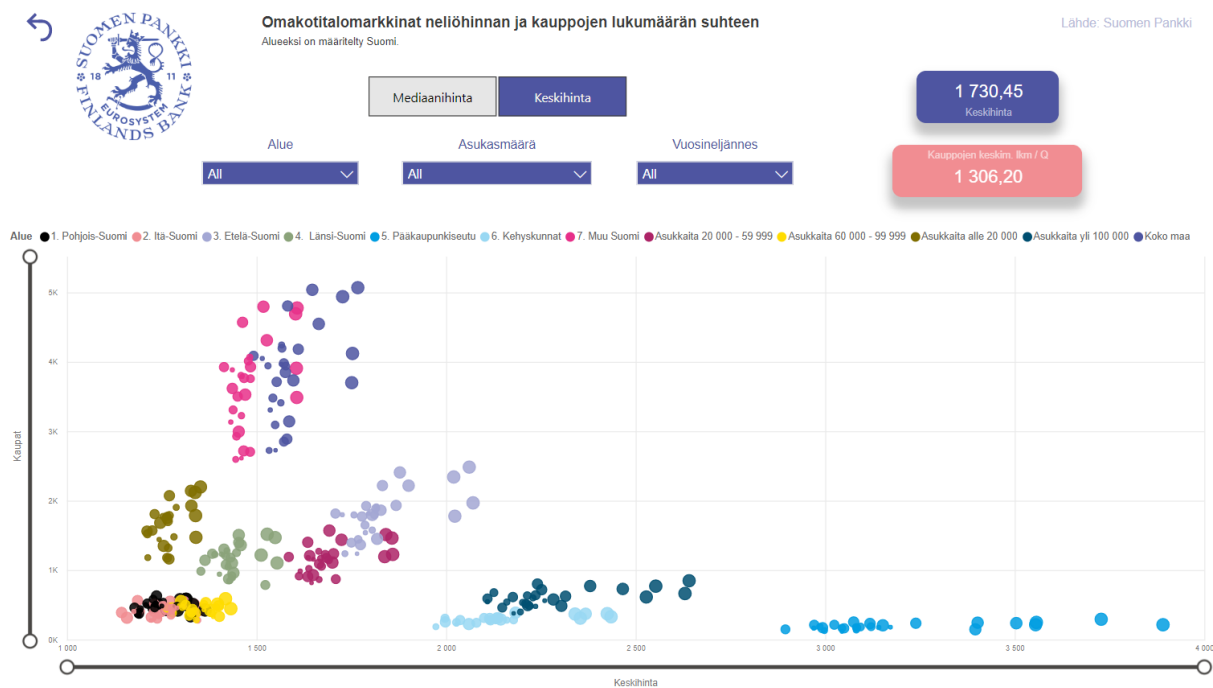


Figure A.5: The dashboard sheet D4 covers the distribution of detached house market

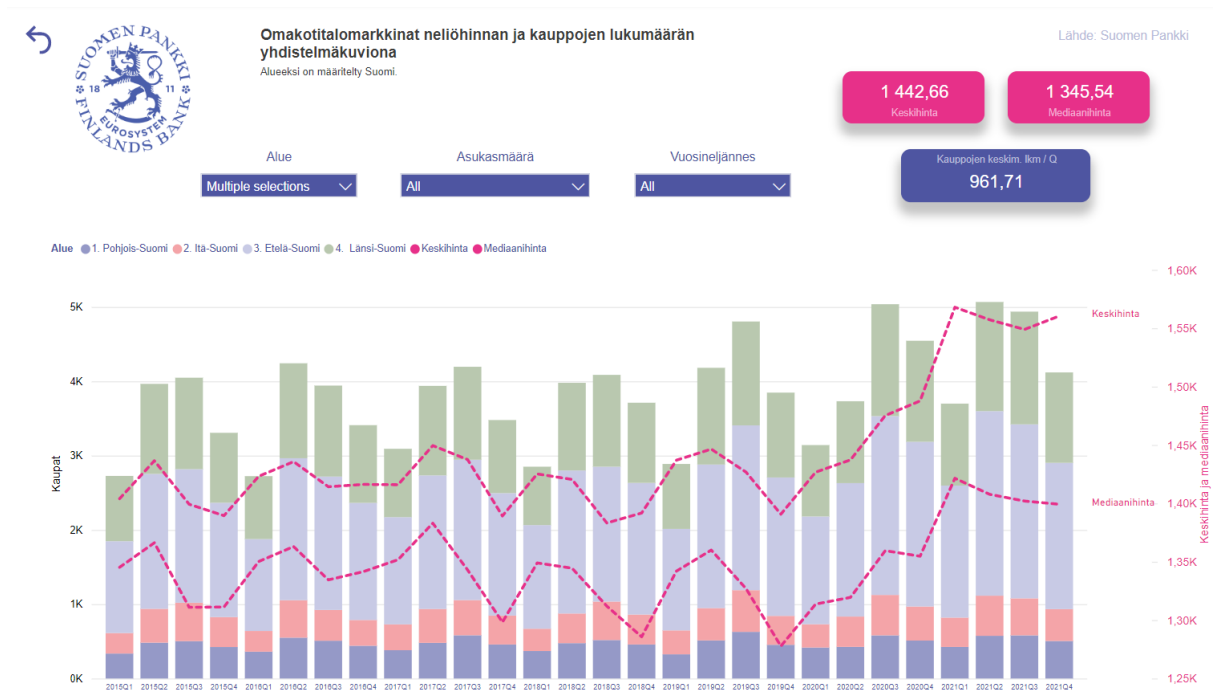


Figure A.6: The dashboard sheet D5 covers the regional outlines of the most essential detached house market variables

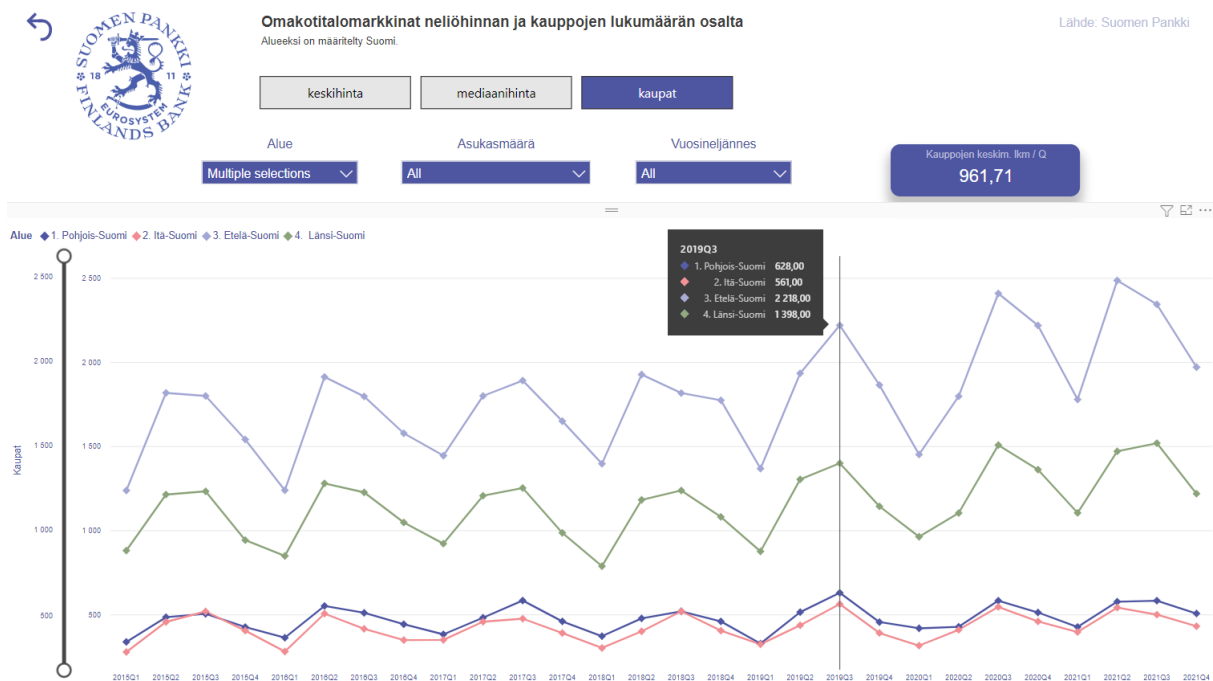


Figure A.7: The dashboard sheet D6 covers the regional comparison of the most essential detached house market variables

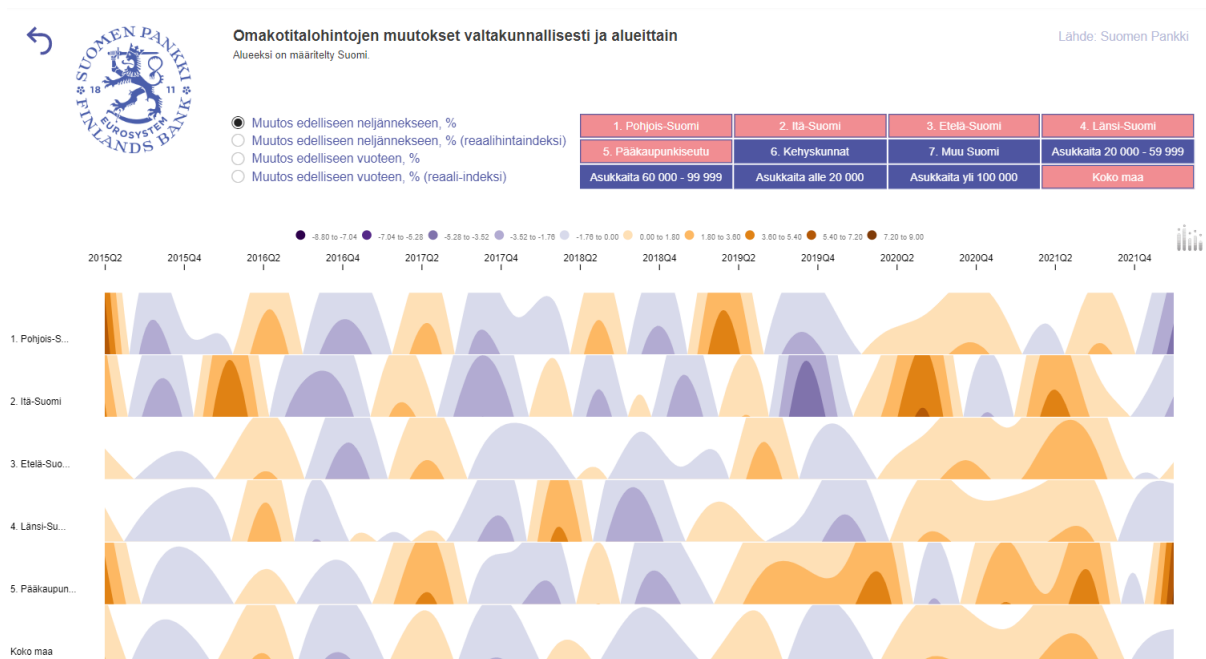


Figure A.8: The dashboard sheet D7 covers the regional changes of detached house prices and market indices

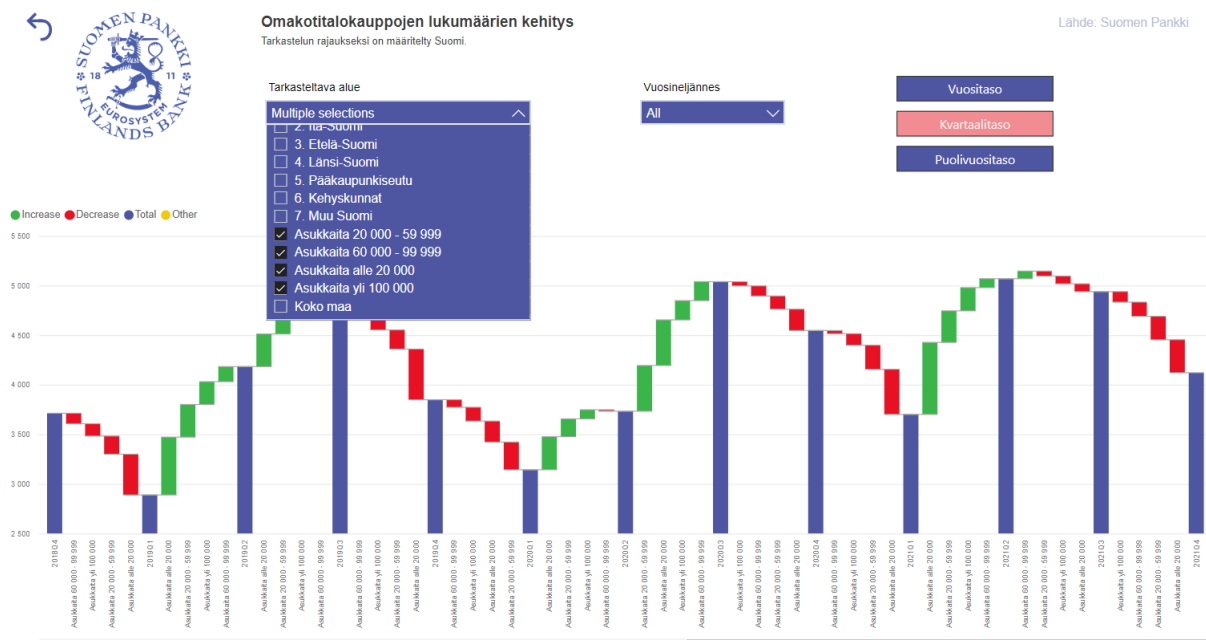


Figure A.9: The dashboard sheet D8 covers the national and regional development of detached house transactions

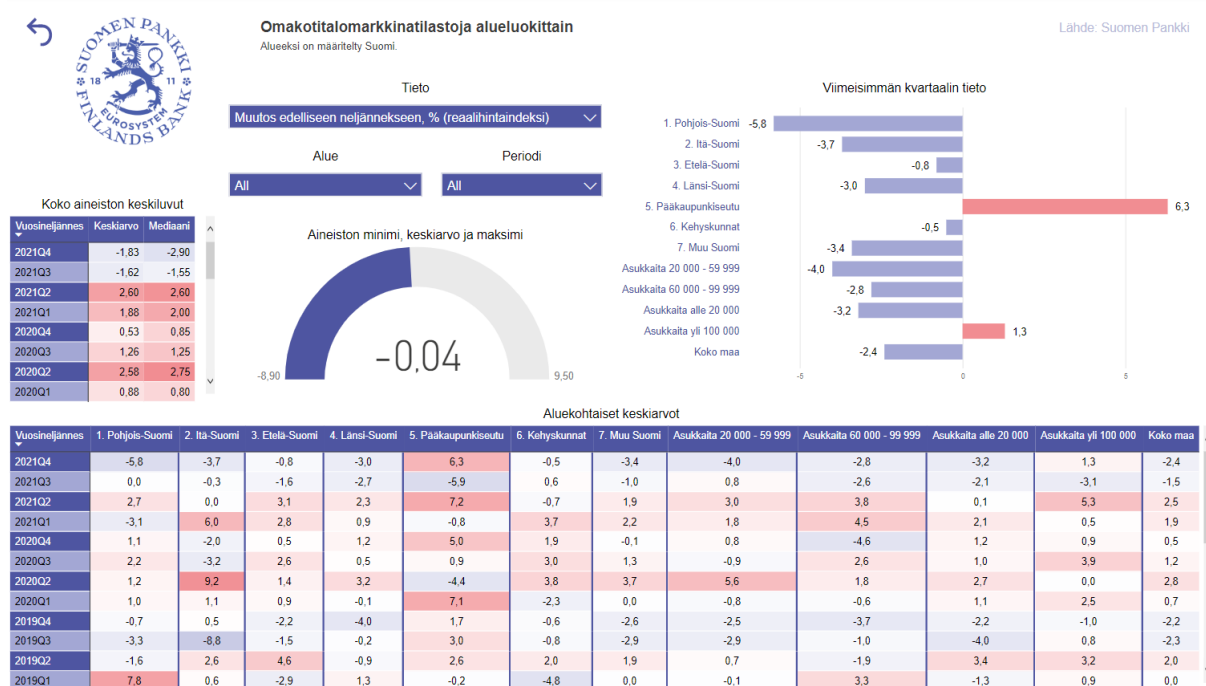


Figure A.10: The dashboard sheet D9 covers the rest detached house market statistics

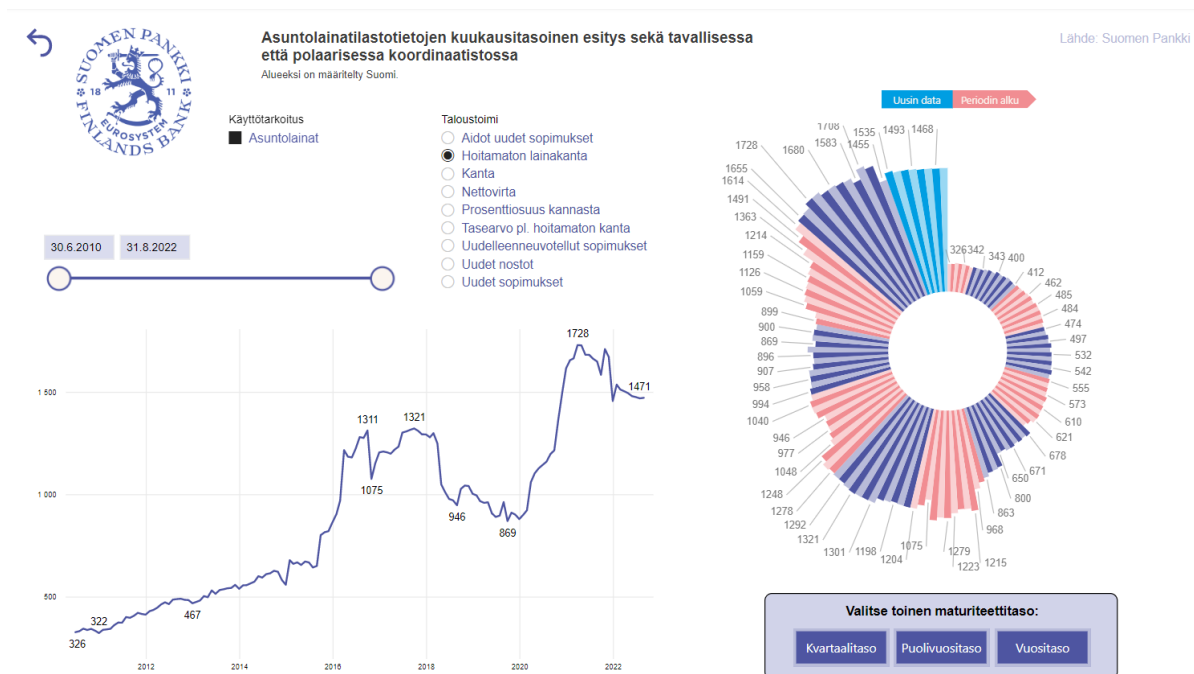


Figure A.11: The dashboard sheet D10 covers the outlines of housing-related loan market in Finland. The figures illustrate the development of untreated loan base.

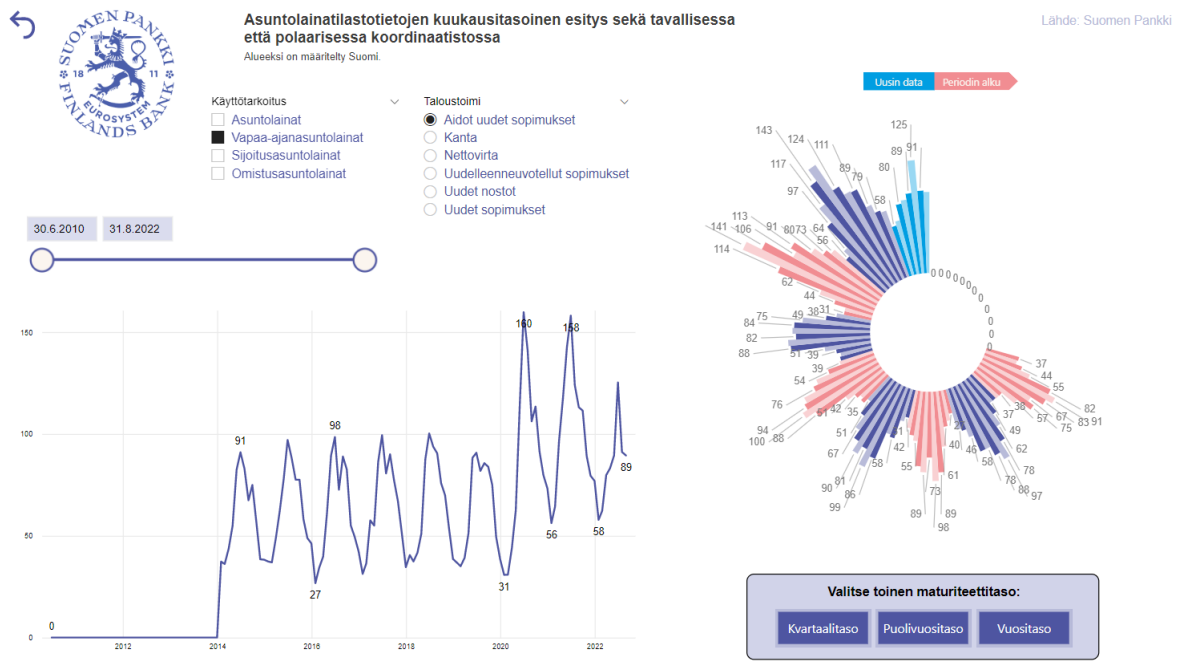


Figure A.12: The seasonality of summerhouse loans illustrated in the dashboard sheet D10.

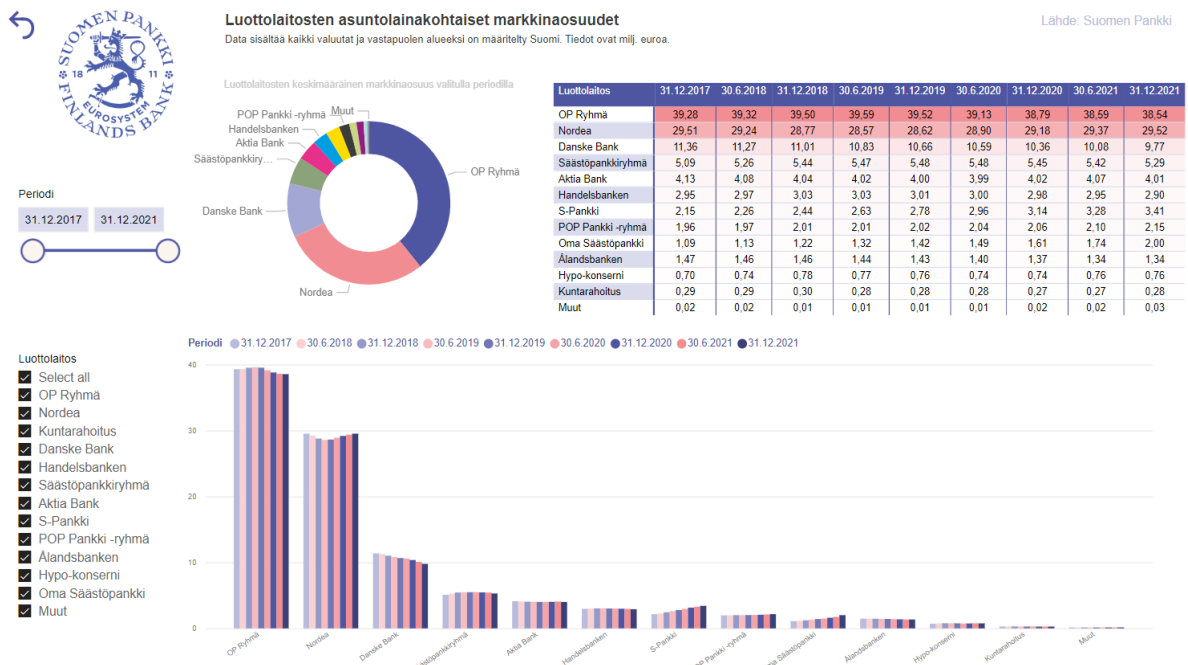


Figure A.13: The dashboard sheet D11 covers the mortgage market shares of Finnish creditors

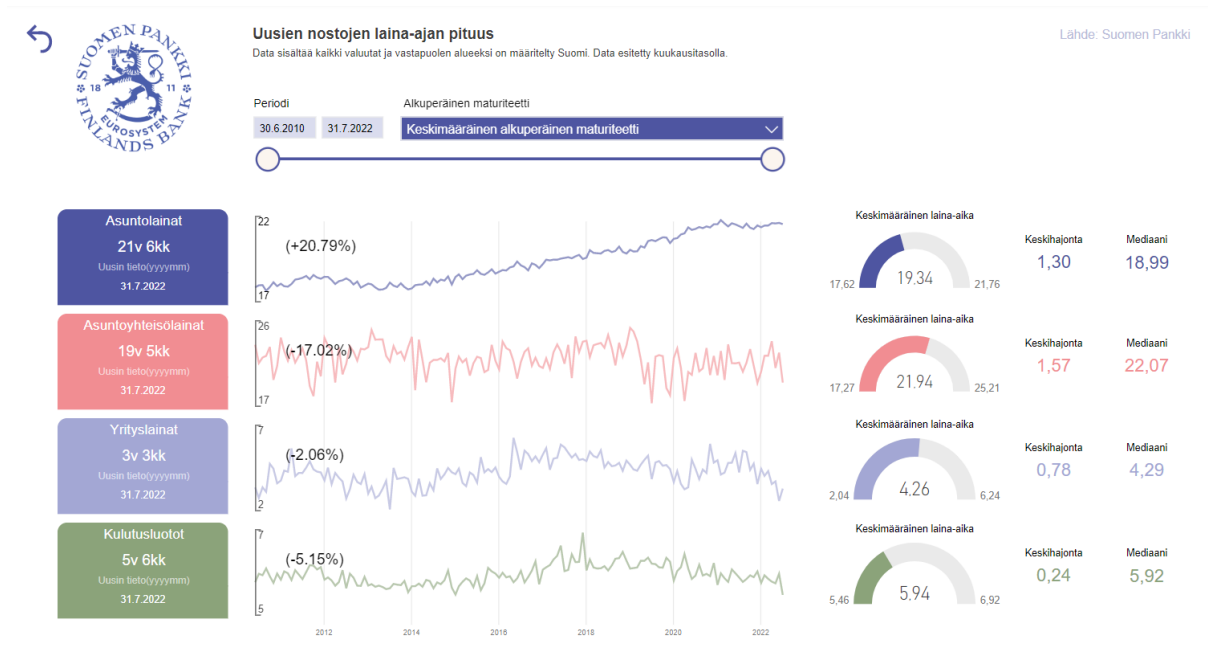


Figure A.14: The dashboard sheet D12 covers the maturities of housing-related loans

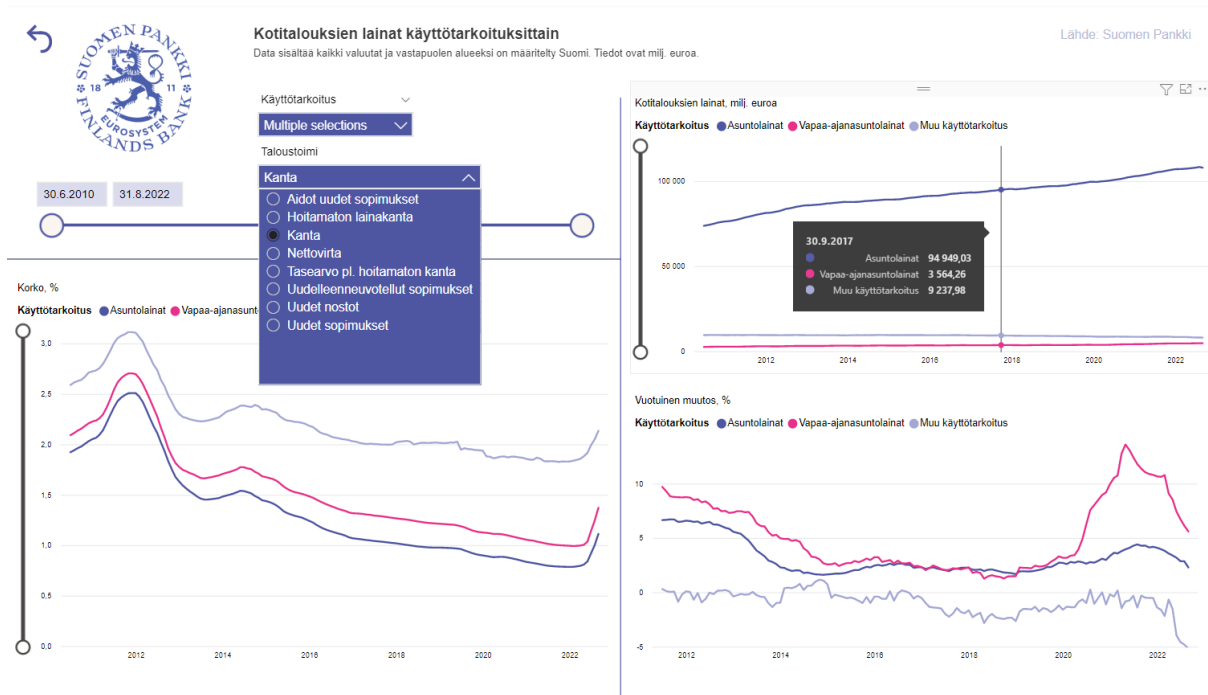


Figure A.15: The dashboard sheet D13 covers the Finnish domestic loans in respect of the most essential loan variables

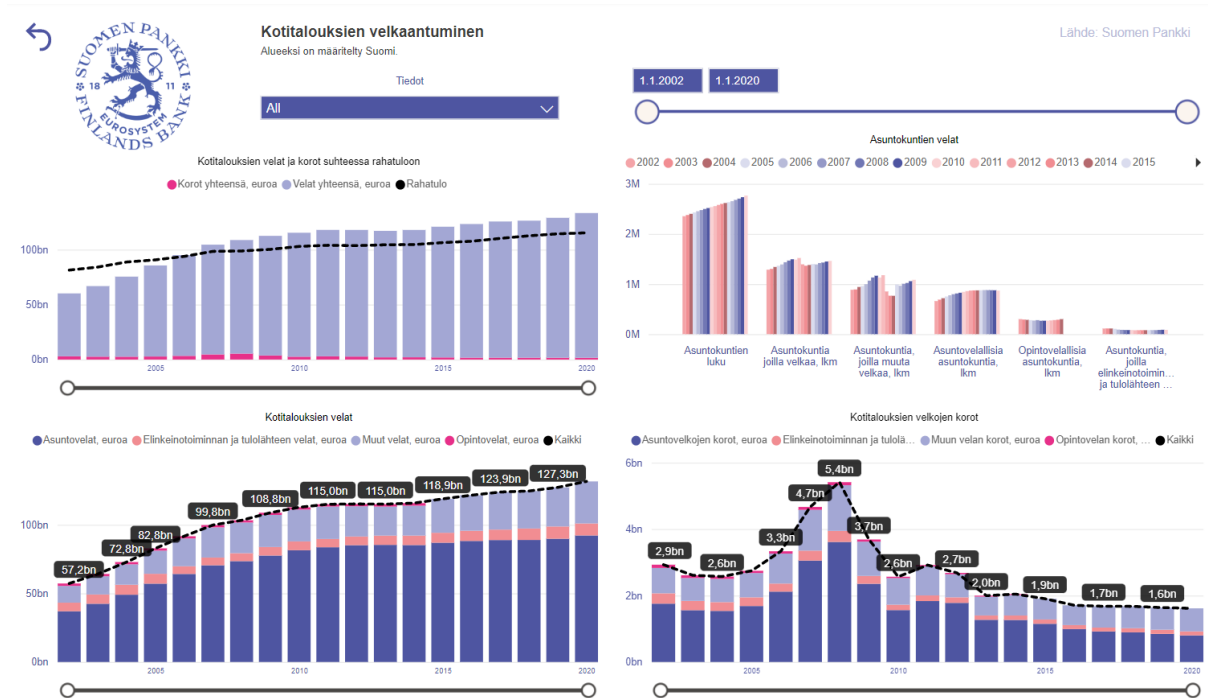


Figure A.16: The dashboard sheet D14 covers the domestic indebtedness and income

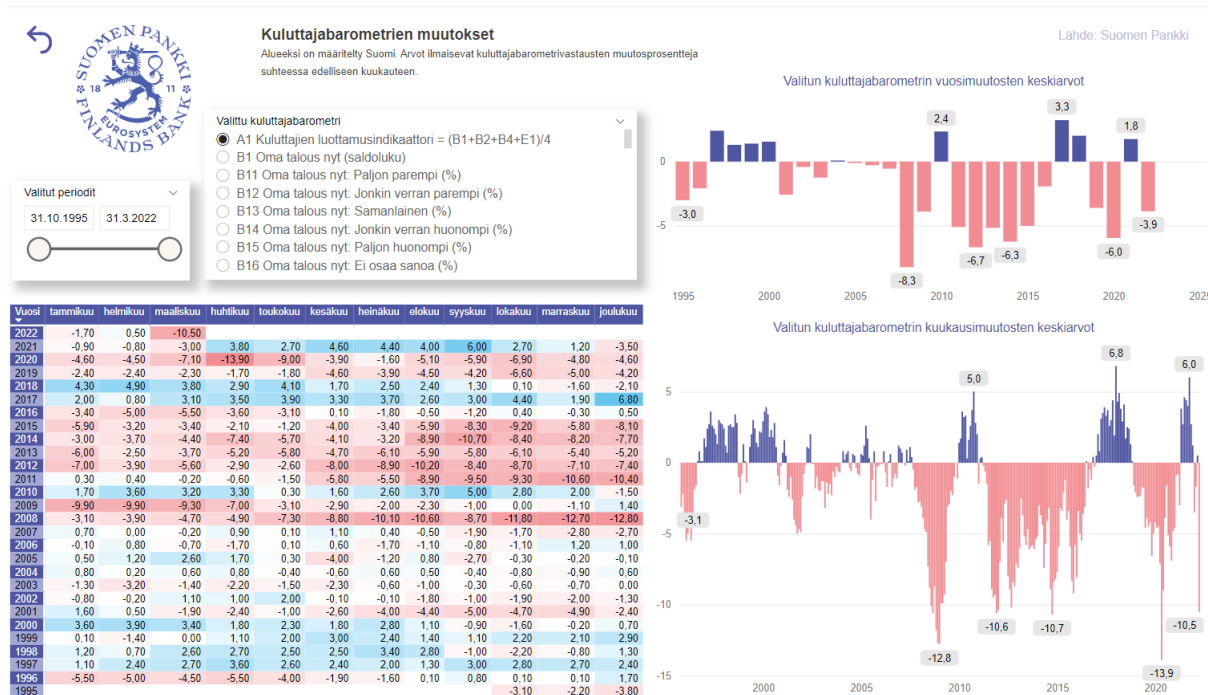


Figure A.17: The dashboard sheet D15 covers the atmosphere changes of the economic consumer barometers

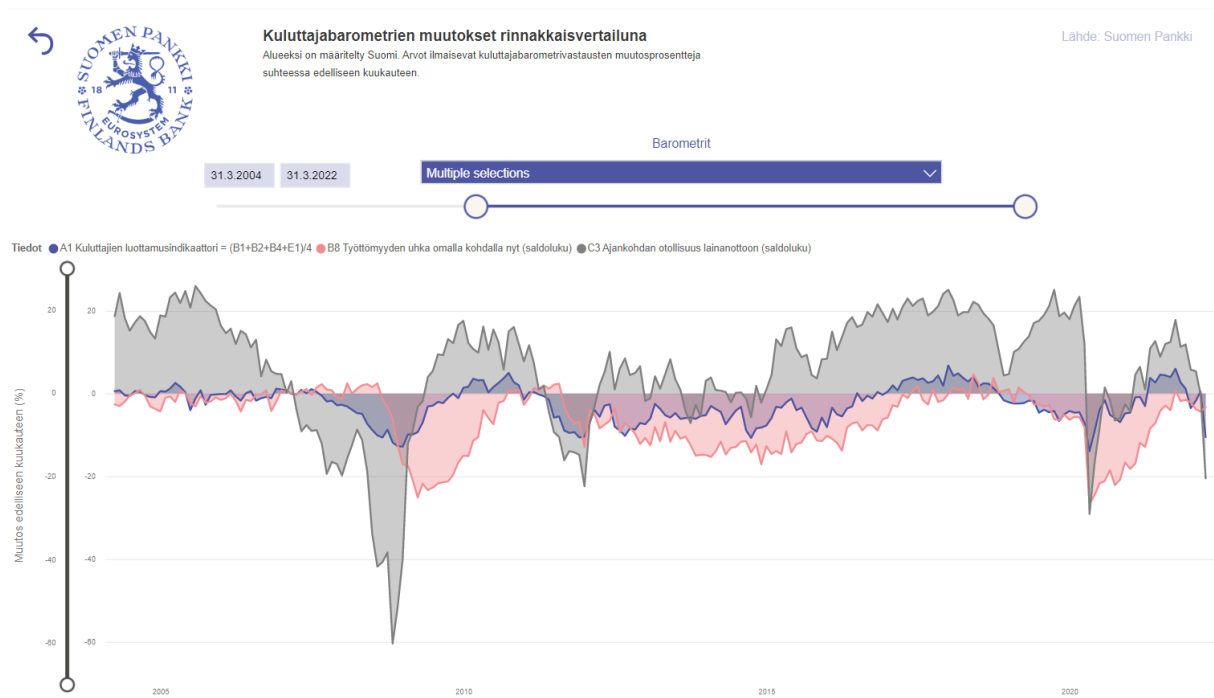


Figure A.18: The dashboard sheet D16 covers the comparison between the atmosphere development of the economic consumer barometers and the actual economic development