Reports in Transition — The Case of Business Data Visualization in Ideal Keittiöt Oy

Master's Thesis
University of Turku
Department of Future Technologies
Interaction Design
May 2020
Kari Salminen

TURUN YLIOPISTO
Tulevaisuuden teknologioiden laitos

KARI SALMINEN: Reports in Transition — The Case of Business Data Visualization in Ideal Keittiöt Oy

Pro gradu -tutkielma, 72 s., 37 liites.
Vuorovaikutusmuotoilu
Toukokuu 2020

Tämän tutkielman tavoitteena oli kehittää kolme uutta merkityksellistä Qlik Sense -raporttia Ideal Keittiöt Oy:lle, joka on suomalainen keittiö- ja saniteettikalusteita myyvä yritys. Ideal Keittiöt Oy oli käyttänyt Tarla-nimistä erikoisvalmisteista toiminnanohjausjärjestelmää siihen sulautetun räätälöidyn Pentaho Reportingin kanssa vuodesta 2011 alkaen. Vuonna 2017 käynnistettiin liiketoiminnallisten tarpeiden perusteella vaihto Jeeves ERP-iin ja Qlik Senseen. Uusien raporttien loppukäyttäjien oli tarkoitus olla johtohenkilöstö ja harvat valitut työntekijät. Tässä työssä oli tarkoituksena myös osittain dokumentoida Jeeves ERPin tietokanta raportointitarpeita varten.

Pentaho Reportingin raporttien käyttötilastoja analysoitiin ja loppukäyttäjiä haastateltiin heidän raportointitarpeisiinsa liittyen. Maaliskuussa 2018 näiden tietojen perusteella valittiin kaksi uutta raporttia toteutettaviksi. Yksi uusi Qlik Sense -raportti toteutettiin: myyntitoimitusten ABC-analyysi Tarlan dataa käyttäen, koska Jeeves ERPin dataa ei ollut vielä saatavilla. Näköaistitutkimusta käytettiin uudessa raportissa datan tehokkaaseen visualisointiin.

Vuonna 2019 Jeeves ERP otettiin käyttöön Ideal Keittiöt Oy:ssa. Marraskuussa 2019 kirjoitettiin erikoistyö Ideal Keittiöt Oy:n sisäinen ohje Jeeves ERP -raportointiin, mikä osittain täytti tämän tutkielman tavoitteen dokumentointiin liittyen. Alunperin oli tavoiteltu kolmea uutta raporttia, kaksi oli valittu toteutettaviksi ja lopulta yksi oli toteutettu. Tämä uusi raportti näytti melko hyvältä, oli kohtuullisen helppo käyttää, mutta ei ollut enää hyödyllinen vuonna 2020, koska se ei käyttänyt Jeeves ERPin dataa.

Asiasanat: tiedon visualisointi, datan visualisointi, business intelligence, ERP, toiminnanohjausjärjestelmä, PK-yritys

UNIVERSITY OF TURKU
Department of Future Technologies

KARI SALMINEN: Reports in Transition — The Case of Business Data Visualization in Ideal Keittiöt Oy

Master's Thesis, 72 p., 37 app. p.
Interaction Design
May 2020

The aim of this thesis was to develop three new relevant Qlik Sense reports for Ideal Keittiöt Oy, which is a Finnish company selling kitchen and sanitary fixtures. Ideal Keittiöt Oy had been using a custom-built enterprise resource planning system named Tarla with a customized Pentaho Reporting embedded in it since 2011. In 2017, a switch to Jeeves ERP and Qlik Sense was initiated out of business needs. The end-users of the new reports were to be the management staff and a few select employees. The Jeeves ERP's database was also to be partly documented for reporting purposes.

The use statistics of Pentaho Reporting's reports were analyzed and the end-users were interviewed about their reporting needs. In March 2018, based on this information two new reports were selected to be implemented. One new Qlik Sense report was implemented: an ABC analysis of sales deliveries using Tarla's data, because data from Jeeves ERP was not yet available. Research on visual perception was used in the new report for the effective visualization of data.

In 2019, Jeeves ERP was deployed in Ideal Keittiöt Oy. In November 2019, an internal guide for using Jeeves ERP's database in reporting was written for Ideal Keittiöt Oy as a master's project, which partially satisfied this thesis's aim related to documentation. Three new reports had originally been aimed for, two had been selected to be implemented and one had been finally implemented. This new report looked quite good and was moderately easy to use, but was not useful anymore in 2020 because it did not use Jeeves ERP's data.

Keywords: information visualization, data visualization, business intelligence, ERP, enterprise resource planning system, SME

Contents

J Tarla to ABC analysis data model mapping

List of Figures

List of Tables

List of Listings

List of Acronyms

BI	Business Intelligence
CCO	Chief Commercial Officer
CEO	Chief Executive Officer
CFO	Chief Financial Officer
CIO	Chief Information Officer
CSV	Comma-Separated Values
ERP	Enterprise Resource Planning
HTML	Hypertext Markup Language
IT	Information Technology
SME	Small and Medium-Sized Enterprise
SQL	Structured Query Language
SVG	Scalable Vector Graphics
SVN	Subversion
VAT	Value-Added Tax
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations

Chapter 1

Introduction

Ideal Keittiöt Oy () is a medium-sized [1, 2] Finnish limited company retail selling kitchen and sanitary fixtures [3] and the parent company of the Ideal Keittiöt Oy concern [4]. A part of this concern is also TNC-Components Oü [4, 5] whose warehouse in North-Eastern Estonia [6] acts as the concern's central warehouse [7].

The author of this thesis was an employee of Ideal Keittiöt Oy from autumn 2011 until February 2019. In 2017 when this thesis was started the company was transitioning [8] from an old custom Enterprise Resource Planning (ERP) system named Tarla (Section 3.1.1) to the commercially available Jeeves ERP [9], which hereinafter shall be called simply Jeeves. At the same time, also much of the firm's Business Intelligence (BI) solutions were changing. The use of Pentaho Reporting [10,11] was being discontinued in favor of Qlik Sense [12], which had been chosen by the Chief Information Officer (CIO) as the firm's new BI solution.

The old reports [13] based on Pentaho Reporting and the old ERP system's data were not compatible nor directly convertable to Qlik Sense [14–16] and Jeeves's data. Thus, there was a clear need at Ideal Keittiöt Oy for new reports based on the chosen BI tools Qlik Sense and Jeeves.

The original aim of this thesis was to use academic and industry research and best practices as a basis for developing three new relevant Qlik Sense reports for use by Ideal Keittiöt Oy's management and a few select employees. The emphasis was on data visualization and discoverability—discoverability means that one should be able to drill down into the details rather than be constrained to an aggregated view of the data. The number of new reports was restricted to three because of time and resources constraints.

An original ancillary aim of this thesis was the modelling of Jeeves's database to a sufficient degree for making the new reports. Jeeves's database is a relational database using Microsoft SQL Server with no foreign key definitions, hundreds of tables and views, and thousands of columns [17]. The reference guide documenting Jeeves's database tables and their columns [18] was to be the primary source of information for the modelling, while secondary sources were to be database reverse engineering [19] and commercial consultations of business partners Pengon Oy [20] and Staria Oyj [21].

The thesis starts with a review of the selected literature related to human visual perception (Chapter 2) and a presentation of the Ideal Keittiöt Oy's case background (Chapter 3) including their old and new ERP systems and BI solutions (Sections 3.1 and 3.2). Based on interviewing Ideal Keittiöt Oy's management personnel about their reporting needs two new reports are selected to be implemented (Chapter 4). The single new report that was finally implemented with Qlik Sense is described (Chapter 5) and evaluated by Ideal Keittiöt Oy's CIO and the author of this thesis (Chapter 6). The thesis ends with concluding remarks (Chapter 7).

Chapter 2

Human visual perception

This chapter describes scientific research and best practices related to human visual perception as a basis for practical use in information visualization. The aim of this chapter is not to give a full depth description of human visual perception but enough to use it for information visualization purposes in this thesis itself (see e.g. Figures 3.8, 3.12, 4.2 and 5.1b) and in the Qlik Sense reports implemented for this thesis (see Chapter 5).

The main source for information on human visual perception in this thesis is Colin Ware's book *Information Visualization: Perception for Design* (2013) [22]. Sections 2.1, 2.2 and 2.4 to 2.6 are based mainly on this book. If the reader is interested to know more about human visual perception than what is described in this chapter, the author of this thesis recommends the aforementioned Colin Ware's book [22].

The main source for information on how to design effective and intuitively perceivable graphs is William S. Cleveland's book *The Elements of Graphing Data* (1994) [23]. Section 2.7 is based mainly on this book.

2.1 Eye and visual cortex

The lens in the eye focuses the image it receives from outside to the retina which is a layer of photoreceptor cells at the back of the eye. The photoreceptor cells consist of 6 million cones that are very sensitive in dark and 100 million rods that are sensitive in normal light conditions. The sharpest area of vision is in the center of the retina in a small area called the fovea, which consists of more than 100,000 cones and affects a visual angle of 1.5°–2°. [22, p. 49]

The received visual information is preliminarily processed in the retina and then passed through the optic nerve to the lateral geniculate nucleus. From there the information is sent to multiple areas of which one of the most important is the visual area 1 (V1) of the visual cortex at the back of the head. After that most of the information passes from V1 to V2 (Visual area 2). Together V1 and V2 form over 40% of the processing of vision. [22, pp. 71, 143]

Half of the neurons in V1 process signals only from the 10° of central vision. The cells in V1 and V2 process orientation, size, color, stereoscopic depth and motion. Color, form (orientation and size) and motion are processed separately. These separately processed properties do not interfere with each other—e.g. the color of an object does not interfere with the size of an object and vice versa. [22, pp. 53, 145]

2.2 Visual search

Because the sharpest area of vision is only a 1.5°–2° of a visual angle the eyes are moved frequently in order to gather detailed information. Eyes can move in three different ways: saccadic movements, smooth-pursuit movements and convergent movements. Smooth-pursuit movements are smooth eye movements that happen when a person tracks an object moving in the visual field. Convergent movements are either smooth or saccadic eye movements that happen when an object of visual attention changes depth. Saccadic movements are rapid jerky eye movements which can not be stopped once started. A saccadic movement takes 20–180 ms, the eyes dwell at a single place typically for 200–400 ms and the movement length is typically approximately 5° of a visual angle. If eye movement of over 20° is required then also the head is moved which can take over 0.5s. [22, pp. 49, 140–141]

What determines how the eye is moved? First a visual query is formed in a person's mind related to a problem like e.g. finding the nearest postal office. Then a visual search is initiated trying to find patterns to fulfill this query. [22, p. 139]

According to the Jeremy M. Wolfe's Guided Search model color, size and motion are first processed in parallel across the whole visual field (see visual areas 1 and 2 in Section 2.1 and preattentive processing in Section 2.4). A limited number of attributes (1–2 dozen) from this early processing can guide attention. Some attributes—e.g. color—have a stronger ability to guide attention than others—e.g. opacity. [24]

As an example the problem of finding a missing red shirt in a white hospital room can be effectively guided by early visual processing and likely to be rapidly solved if the shirt is anywhere in the person's field of vision.

2.3 Cultural differences in visual search

There are different reading directions in different language scripts: left-to-right (e.g. English in Latin script or Russian in Cyrillic script), right-to-left (e.g. Farsi in Arabic script or Yiddish in Hebrew script) and left-to-right or top-to-bottom (e.g. Mandarin Chinese in Simplified Chinese script or Korean in Hangul script) [25]. So are there cultural differences in visual search?

Westerners, East Asians and Middle Easterners all fixate more on the left than on the right and more on the top than on the bottom. East Asians move their eyes less horizontally and more vertically than Westerners or Middle Easterners. Westerners and Middle Easterners usually move their eyes twice as often horizontally as they do vertically. Westerners move their eyes horizontally more left-to-right than Middle Easterners or East Asians. Middle Easterners move their eyes horizontally more right-to-left than Westerners or East Asians. East Asians move their eyes horizontally almost equally left-to-right and right-to-left. [26] There is also evidence of Chinese people processing visual scenes more holistically than Americans by a more balanced fixation between the foreground object and the background [27].

In practice the fixation more on the left and on the top [26] can be used to inform information visualization by placing the most important elements on the left, on the top or in the top-left corner.

2.4 Preattentive processing of visual features

Preattentive processing means that some aspects of visual objects are processed faster than 10 ms per item [22, p. 153]. Nonpreattentive processing takes typically 40 ms per item or more [22, p. 153]. The primary visual cortex (see Section 2.1) can separately process color, form and motion [22, p. 145]. Mostly preattentive processing works using differences in color, contrast, orientation, size, motion and blinking [22, pp. 155–156]. See Figure 2.1 for examples of different single visual features that are preattentively processed [22, pp. 154–155]. Generally combinations of visual features are not preattentively processed but at least these six are: color and spatial grouping, stereoscopic depth and color, stereoscopic depth and movement, luminance polarity and shape, convexity/concavity and color and motion and shape (see Figure 2.2) [22, pp. 159–162; 28–30].

As an example of a single visual feature search looking at the first subfigure (a) in Figure 2.1 the green circle is found very rapidly i.e. at a glance. The visual feature that is processed preattentively here is the color. For a more complex example of the use of 2–10 distinct colors for information coding see Figure 2.3.

As an example of a combined visual feature search looking at the first subfigure (a) in Figure 2.2 all the combinations of luminance and shape are found rapidly i.e. black circles, white circles, black squares and white squares. The reader is invited to try this out by intentionally looking for a certain combination of luminance and shape and seeing how fast they find them all. The reader is also invited to compare the speed to trying to find all the combinations of color and shape—i.e. red circles, black circles, red squares and black squares—in the last subfigure (d) of Figure 2.2. The former search is preattentively processed and thus is rapid and the latter search is not preattentively processed and thus is slower.

Even though preattentive effects are rapidly perceived they are not all of equal strength and the popout effect of different visual features varies [22, p. 155]. Some effects are hard to ignore and others require attention to be perceived [22, p. 156].

The practical use of preattentive processing is to encode information in preattentively processed visual features or their combinations as they are rapidly perceived. As an example preattentive processing of color is used in ABC analysis report (see Section 5.1) by encoding class A with red, class B with green and class C with blue. This way in the report it can be rapidly perceived whether an information—row in a table or bar in a chart—belongs to class A, B or C.

2.5 Redundant coding of information with multiple visual features

Although combinations of visual features are not generally preattentively processed (see Section 2.4) it still helps in the perception of object groups if redundant coding of information is done using multiple visual features [22, pp. 159–160, 171] (see Figure 2.4). As an example of this see Figure 3.8 on page 32. It redundantly codes the same information using three visual features: color, shape and spatial grouping like in subfigure (c) of Figure 2.4.

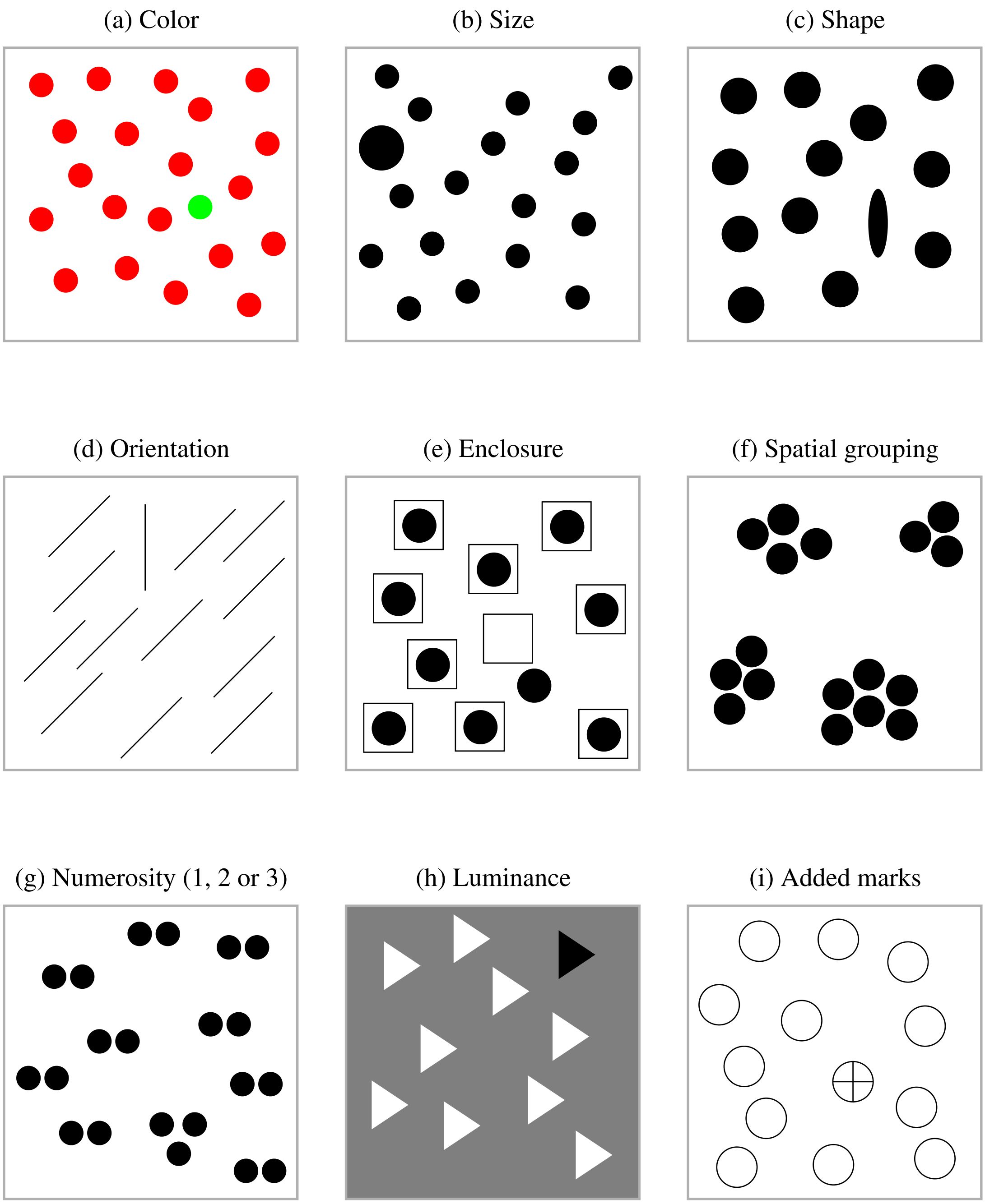


Figure 2.1: Examples of preattentively processed visual features. [22, pp. 154–155]

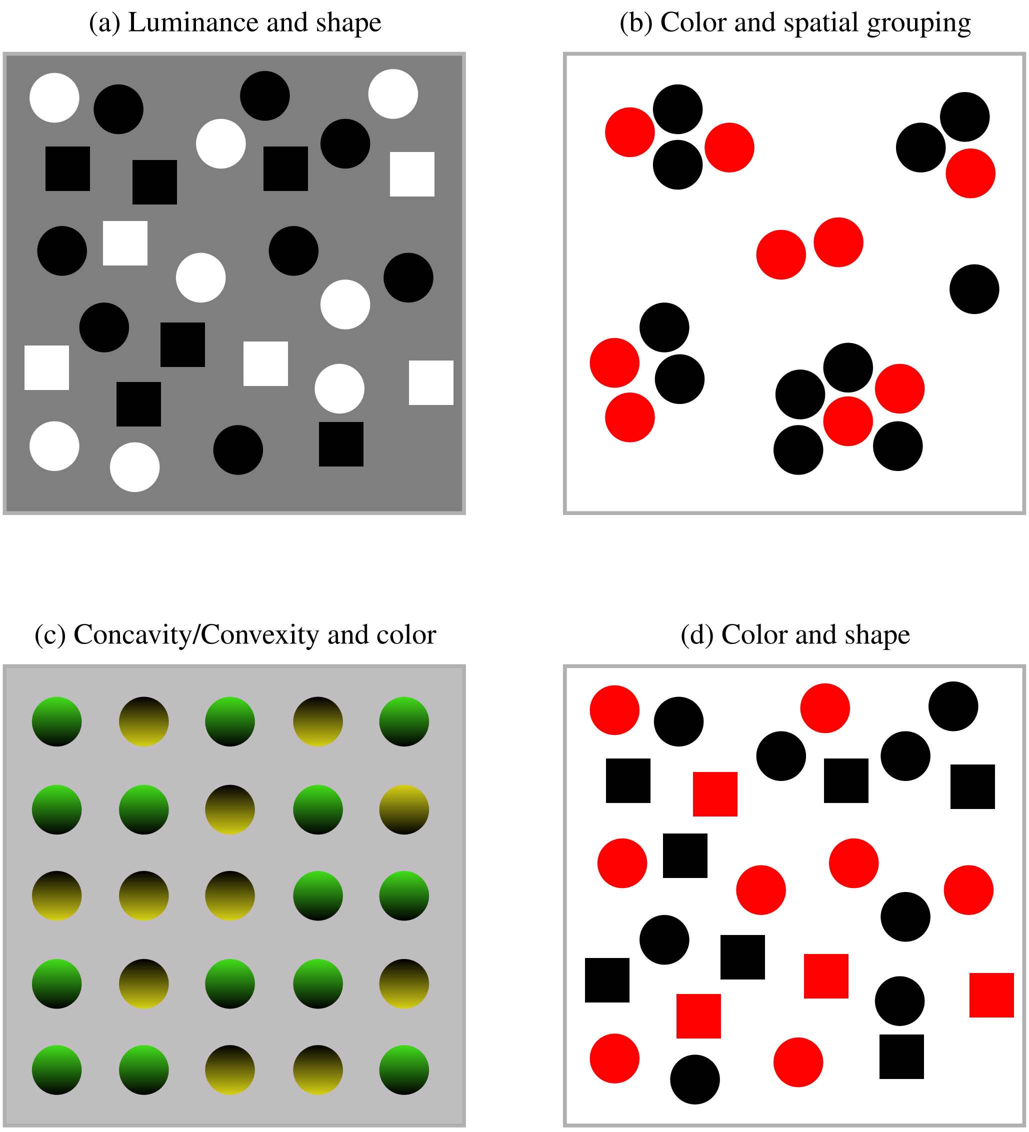


Figure 2.2: Examples of combinations of visual features and their preattentive processing. Combinations of visual features in the subfigures (a), (b) and (c) are preattentively processed [28–30] but not in the subfigure (d) [22, p. 159]. Note that preattentive effects are not all of equal strength [22, p. 155].

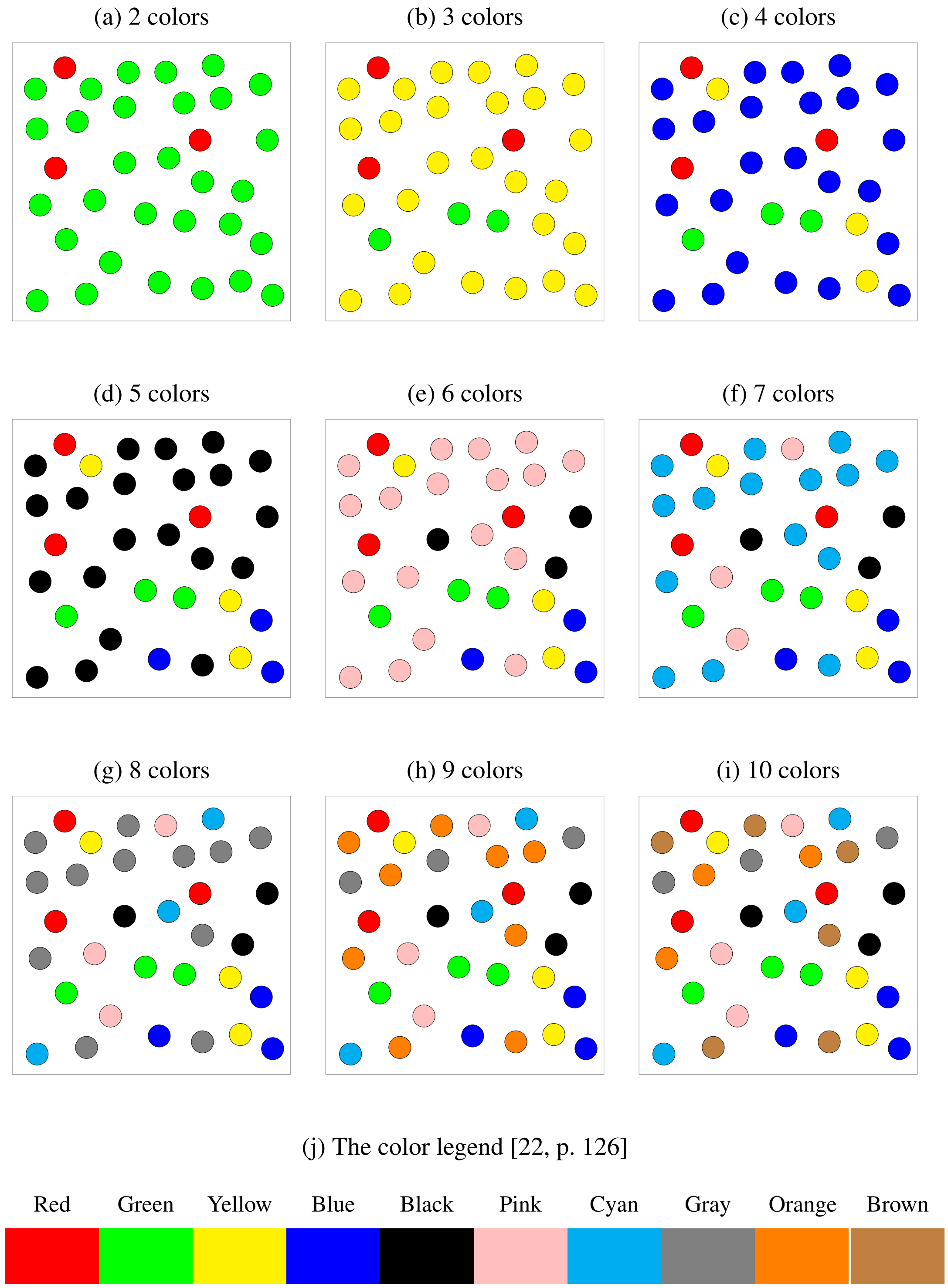


Figure 2.3: Example of using 2–10 distinct colors [22, p. 126] for information coding[a].

[a] At least up to 5 colors can be rapidly perceived if chosen to be sufficiently perceptually distinct. Perception slows down when using 7 colors and further with 9 colors. [31]

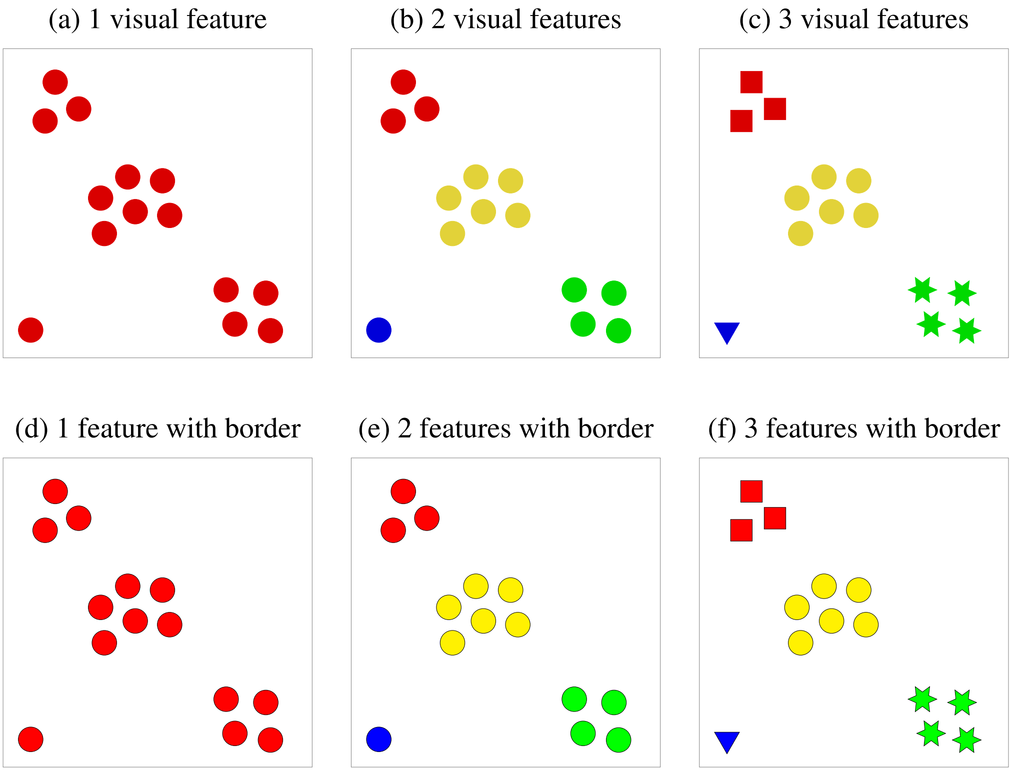


Figure 2.4: Redundant coding of information with multiple visual features (Position, color and shape) enhances perception of object groups [22, pp. 159–160]. Adding a luminance contrast boundary around objects[a] enhances the perception of shapes [22, pp. 113, 123].

[a] Saturation was also added to object colors in the bottom row to make them more distinct from each other.

2.6 Gestalt laws of visual perception

Gestalt laws of visual perception describe how humans recognize patterns visually. Visual pattern recognition was first studied by German psychologists of the Gestalt school in the early part of the 20th century. These laws that they found are general and can easily be used to inform data visualization for effective use of visual objects and patterns. [22, p. 181]

In Table 2.1 are the names of the eight Gestalt laws [22, p. 181], their brief descriptions and links to their example figures.

Table 2.1: The eight Gestalt laws [22, pp. 181–190, 231; 32, pp. 71–88].

Name	**Description**	**Figure**	**Page**
Proximity	Spatial proximity groups objects together.	2.5	13
Similarity	Visual similarity of objects groups them together.	2.6	13
Connectedness	Connection (e.g. a line) between objects groups them powerfully together.	2.7	14
Symmetry	Symmetrical figures are perceived as more holistic than asymmetrical ones.	2.8	14
Relative size	Small areas are more easily perceived as an object in relation to large areas than vice versa.	2.9	14
Continuity	Objects are more easily perceived to be more smoothly curved than sharply cornered.	2.10	15
Closure	Objects are perceived more easily to consist of closed contours rather than broken ones.	2.11	16
Common fate	Motion into the same direction with the same speed groups objects together.	2.12	17

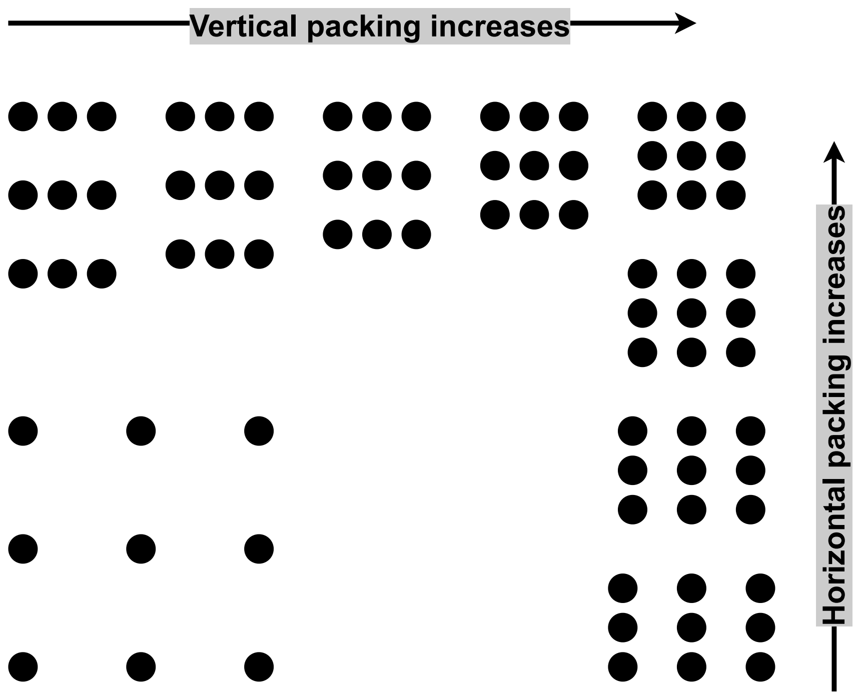


Figure 2.5: Gestalt law of proximity. Spatial proximity groups objects together into rows, columns and more complex formations. [22, pp. 181–182]

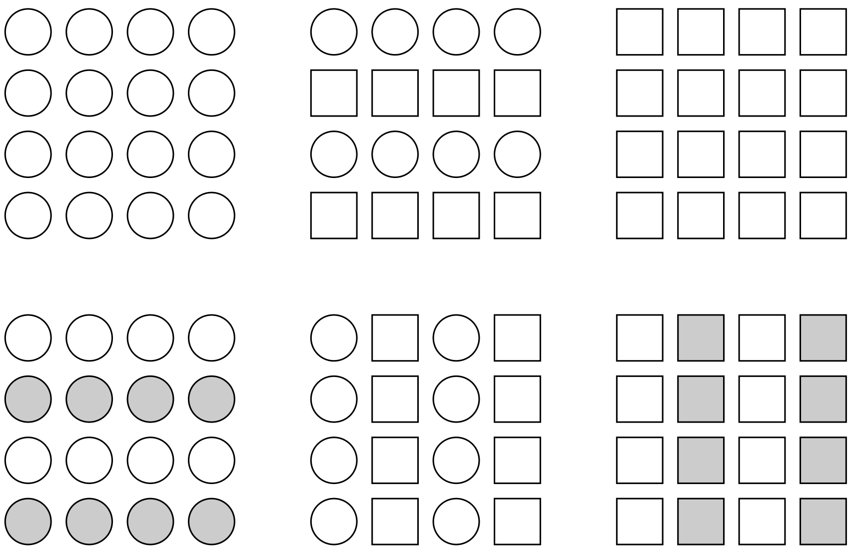


Figure 2.6: Gestalt law of similarity. Similarity of objects makes rows or columns dominate perception. [22, pp. 182–183]

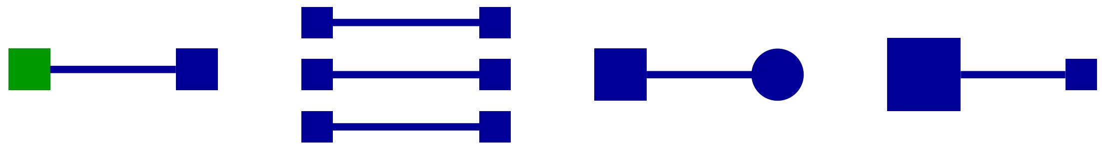


Figure 2.7: Gestalt law of connectedness. Connecting objects groups them together powerfully beyond color, proximity, shape and size. [22, pp. 183–184]

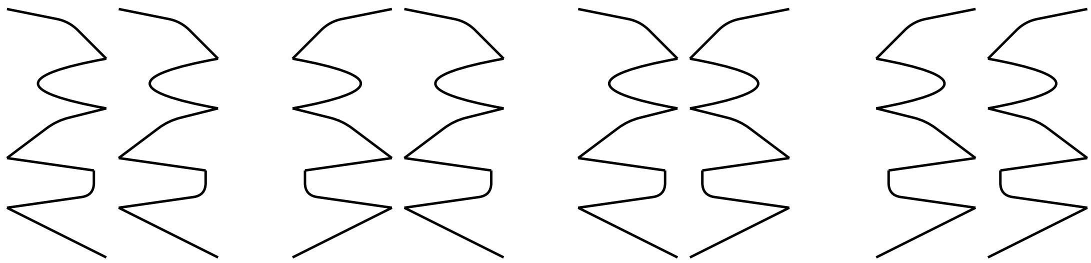


Figure 2.8: Gestalt law of symmetry. Symmetrical figures (near the center) are perceived as more holistic than asymmetrical ones (on the sides). Note that there is only a single silhouette and its mirror silhouette used in all of these four figures. [22, p. 185]

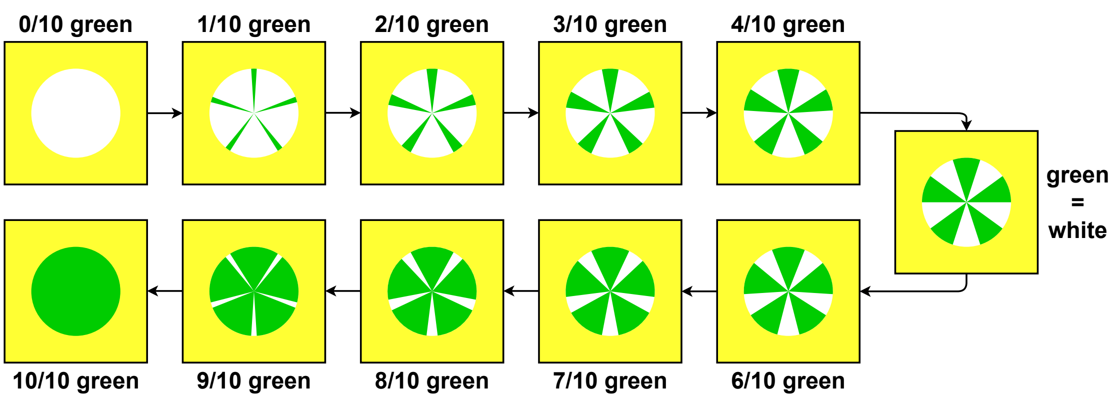
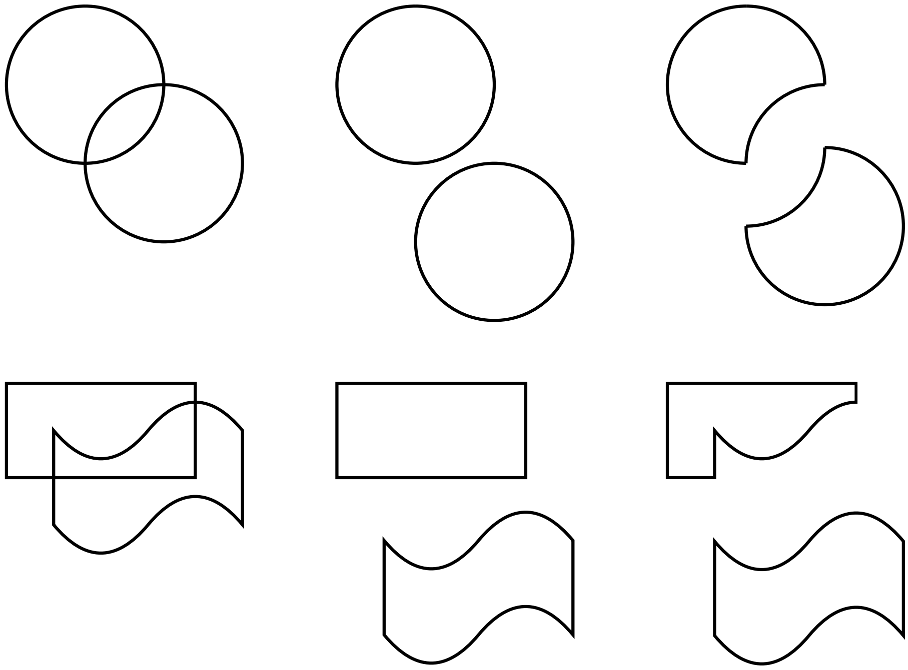


Figure 2.9: Gestalt law of relative size. Small areas are more easily perceived as an object in relation to large areas than vice versa. Therefore on the top row the green area is more easily perceived as an object and on the bottom row the white area. On the right the areas are equal and thus object perception can easily waver. [22, pp. 189–190]

(a) Objects are more easily perceived to be more smoothly curved than sharply cornered.



(b) Smooth connections are more easily perceived than sharply turning ones.

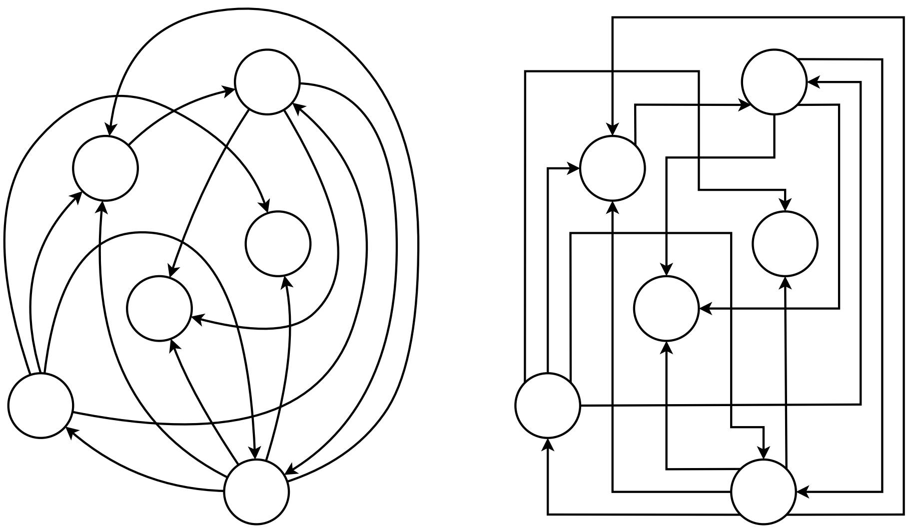
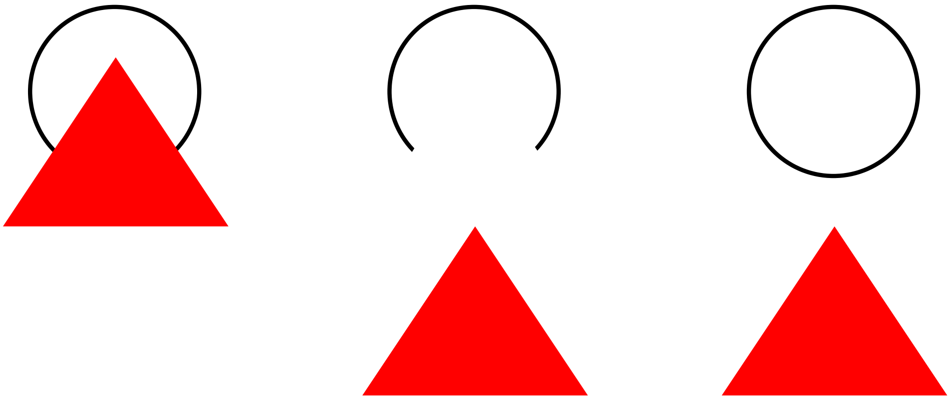


Figure 2.10: Gestalt law of continuity [22, pp. 183–184].

(a) Objects are perceived more easily to consist of closed contours rather than broken ones.



(b) Closed contours organize visual perception into parts. Rectangles are especially powerful examples of this segmentation.

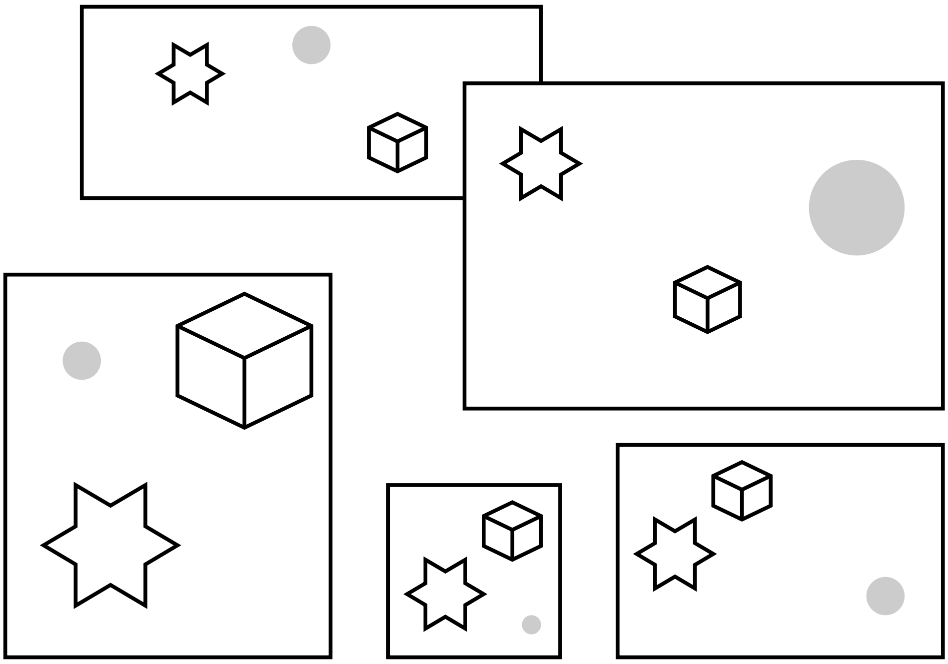


Figure 2.11: Gestalt law of closure [22, pp. 186–190].

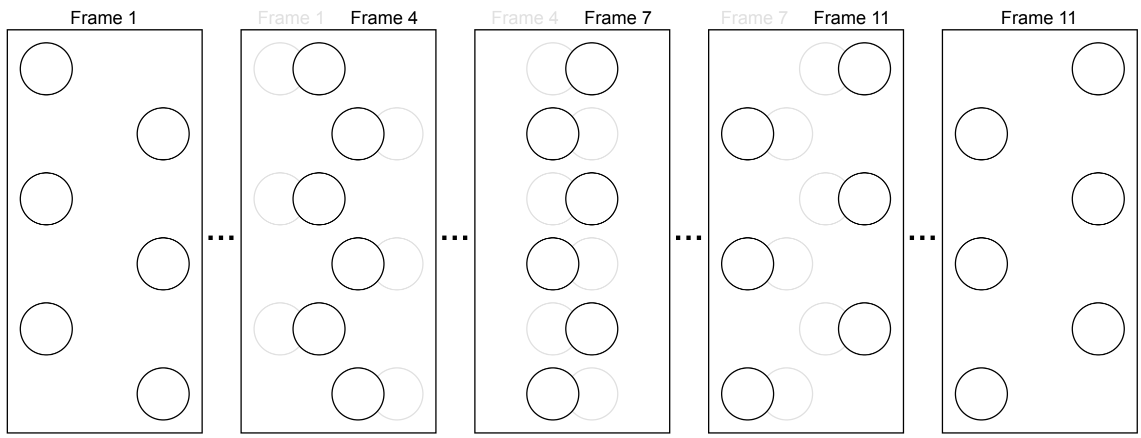


Figure 2.12: Gestalt law of common fate. Objects moving into the same direction with the same speed are perceived as a group. [32, pp. 71–88] Superimposing two frames of animation on top of each other has been used to simulate motion.

2.7 Graphs

To efficiently show the structure of data one should aim for clarity [23, p. 64], avoid showing that which is not absolutely necessary [23, p. 25] and make the data itself prominent [23, pp. 25, 29]. Here are described some of the key elements related to achieving this:

2.7.1 Scale

Each axis on a graph should have a scale line [23, pp. 31–35]. The chosen scale depends on whether the graph is solitary or a part of a compared set. If the graph is solitary then the scale should minimize the scale i.e. maximize the used area of the plotted data in it [23, pp. 81–82]. Some reference values (e.g. zero) may be included in the scale even though their inclusion does not maximize the used area of the plotted data [23, p. 81]. It should however be weighed whether the inclusion of such reference values in the scale merit their inclusion or not—e.g. including zero may arguable if all the data points in the graph are in a radically different value range [23, pp. 92–94]. If the graph is a part of a compared set then a common scale should normally be used for the whole set in order to make it easier to compare its graphs—although sometimes the data sets may be too radically different to use a common scale [23, p. 86].

2.7.2 Tick marks

Tick marks on the scale lines of a graph make it easier to read values from its scale [33, p. 20]. See Figure 2.13 for a comparison of varying the amount of tick marks in a graph. Ticks marks should be made to point outward rather than inward to make sure they do not interfere with the plotted data [23, p. 31]. The range of the plotted data should be fully or almost fully included inside the range of the tick marks [23, p. 80].

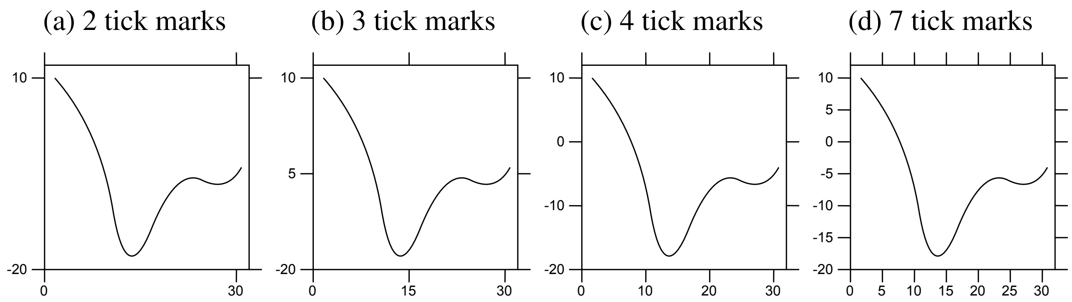


Figure 2.13: A comparison of varying the amount of tick marks in a graph. Too few decrease accuracy of reading values from a scale (see (a) and (b)) and too many clutter the scale (see (d)). A good amount of tick marks (see (c)) increases accuracy of reading values from a scale without cluttering the scale. [23, p. 20; 33, pp. 31–35, 39–41]

2.7.3 Weber's law and visual reference grid

Weber's law (see Figure 2.14) shows that by adding a common frame of reference to objects their lengths can be more easily discerned [23, pp. 240–241]. Visual reference grids (see Figure 2.15) are an application of Weber's law to enhance pattern perception [23, pp. 166–167, 242]. Using the same visual reference grid with multiple graphs increases efficiency of comparing patterns between them [23, pp. 242–243].

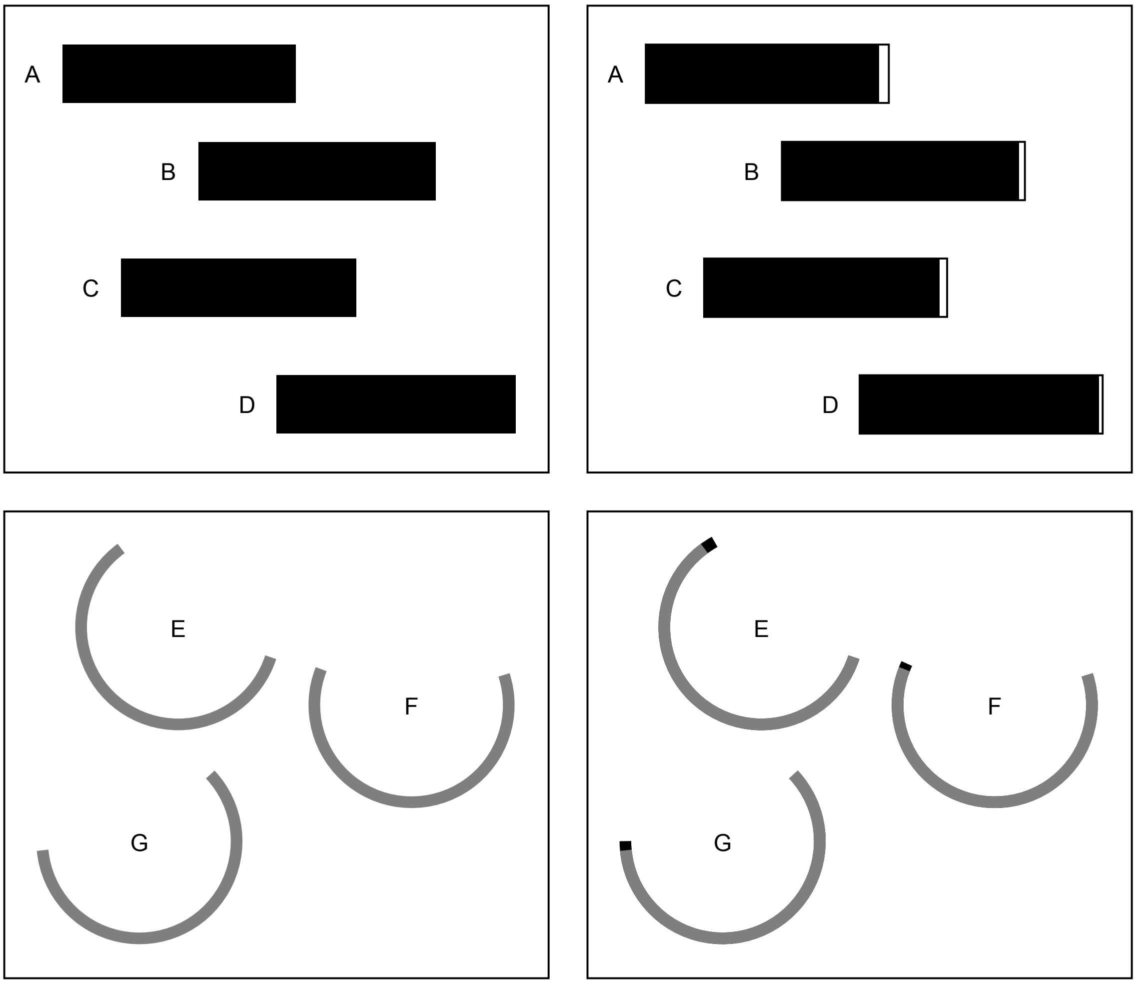


Figure 2.14: Weber's law. The objects A–D and E–G are all of different lengths. On the left it is hard to judge their relative order according to length. On the right adding an identical rectangle around each rectangle or an identical solid arc behind each arc makes this task easier. [23, pp. 240–243] Their order according to length is D > B > C > A and F > G > E.

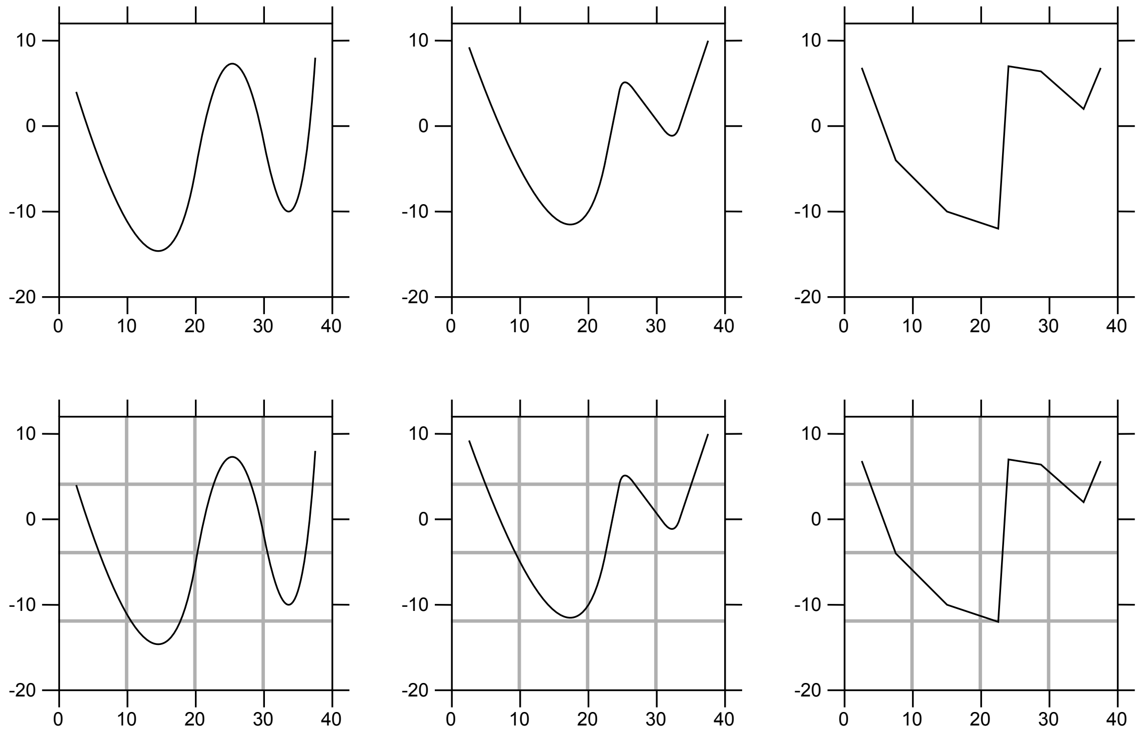


Figure 2.15: Visual reference grid. Visual reference grid makes it easier to compare patterns between two or more graphs. [33, p. 20]

2.7.4 Aspect ratio

The aspect ratio used in graph has a significant effect on how easy it is to judge the rate of change in a graph [23, p. 68] (see Figure 2.16). If it is important that the rate of change in a graph is judged accurately then the aspect ratio should be selected to be such that on average the absolute values of the orientations of the line segments between adjacent data points are 45° [23, p. 70] (see Figure 2.17). The calculation of the optimal aspect ratio is moderately complex [23, pp. 251–256] and is not presented here—there is an open source implementation of the algorithm in the bank_slopes function of the ggthemes [34] package of R [35].

There are however many aspects of a graph to take into account when plotting one. Sometimes other aspects may be more important than judging the rate of change. Examples of this are Figures 3.4 to 3.6 and 3.8 where the author of this thesis judged it more important to be able to read the graphs' X axis values at a glance than judging the rate of change in the graphs. Using the optimal aspect ratios would have made the figures narrower and it impossible to show the full values for the whole X axis horizontally.

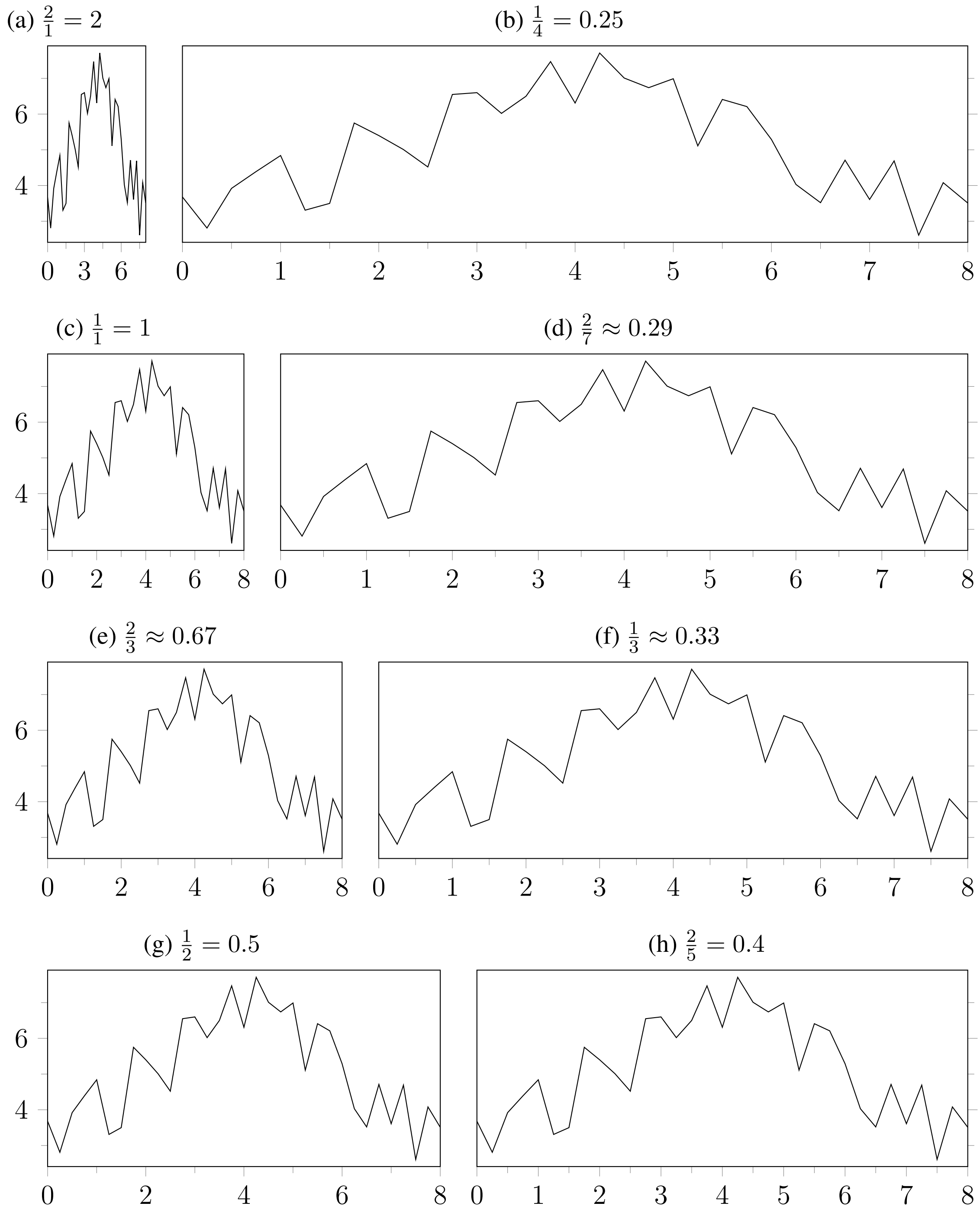


Figure 2.16: Aspect ratio examples.

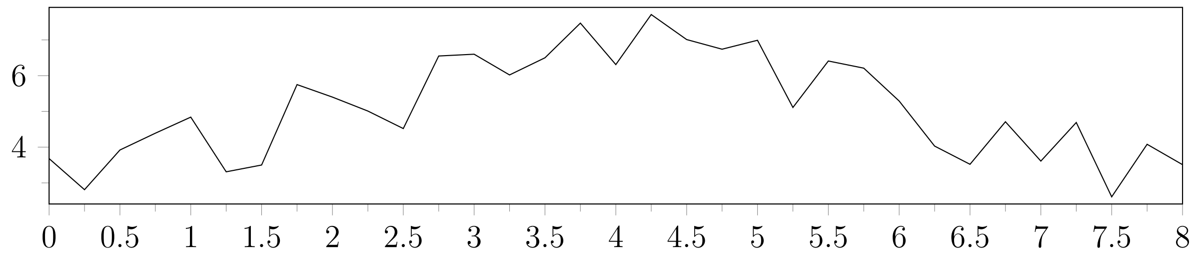


Figure 2.17: Optimal aspect ratio of ≈ 0.173 (see Appendix H).

2.7.5 Graph examples

As an extreme example of how not to design a graph see Figure 2.18. The graph is cluttered, cramped, does not name elements properly and it uses same symbols and colors for different values.

As examples of graphs that use most of the guidelines from this chapter see e.g. Figure 3.8 on page 32 and Figure 4.2 on page 44.

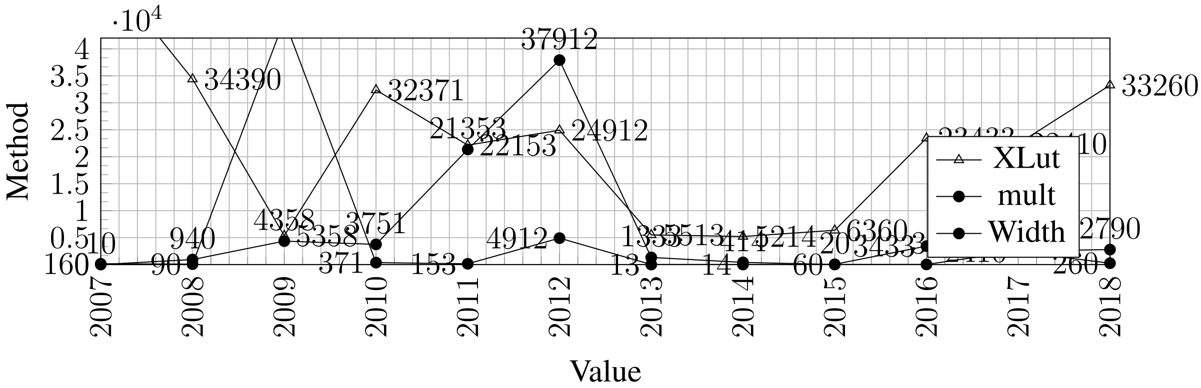


Figure 2.18: Example of a very badly designed graph.

Chapter 3

Case background: Ideal Keittiöt Oy

In this chapter the case background at Ideal Keittiöt Oy is described: the old ERP system named Tarla (see Section 3.1.1), the old BI solution named Pentaho Reporting (see Section 3.1.2), the chosen new ERP system named Jeeves (see Section 3.2.1) and the chosen new BI solution named Qlik Sense® (see Section 3.2.2) are described. The old reports and their use are also described in Section 3.1.2 and this information is later in Chapter 4 used in making an informed judgement about what new reports to implement.

3.1 Old ERP system and BI solution

In this section the development, history of use and future of Ideal Keittiöt Oy's old ERP system named Tarla (see Section 3.1.1) and the old BI solution named Pentaho Reporting (see Section 3.1.2) are described. As this section was mostly written in 2018 the graphs in it do not therefore go beyond the end of year 2017.

3.1.1 Tarla

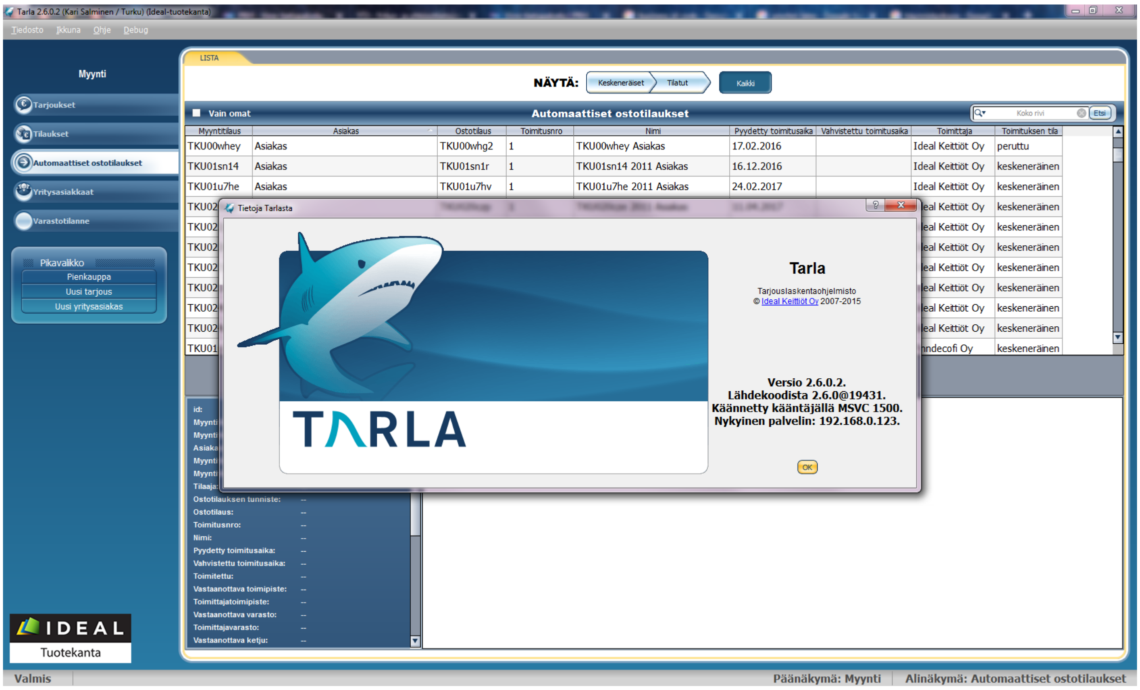


Figure 3.1: Example screenshot of Tarla.

Tarla (see Figure 3.1) is Ideal Keittiöt Oy's [36] custom ERP system. The project for its creation began in 2006 [37] and its software requirements were documented at least already in late 2007[1]. The software development started in 2008[2] with 10 developers and was most vital during 2008–2010, waned from 2011 to 2014 and has been mostly in maintenance mode from 2015 onwards (see Figures 3.2 to 3.4 for details).

History of Tarla's use

The first pilot use of Tarla was at Keittiöjätti Oy's Turku and Espoo offices in the summer of 2009[3]. After the piloting phase Tarla was adopted into nine[4] more offices around the

[1]Earliest found version of Tarla's software requirements document is dated Sep 4, 2007 [38].

[2]First commit to Tarla's source code repository was made on Feb 4, 2008 [13].

[3]These are based on the date of first sale. See Appendix C for a complete list [7].

[4]Salo, Tampere, Kemi, Tornio, Kuopio, Jyväskylä, Oulu, Huittinen and Pietarsaari offices [7].

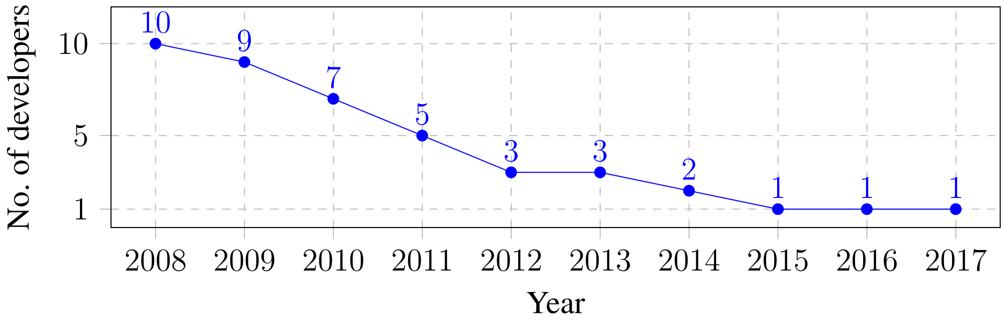


Figure 3.2: Number of active Tarla developers per year [13] (see Appendix B)

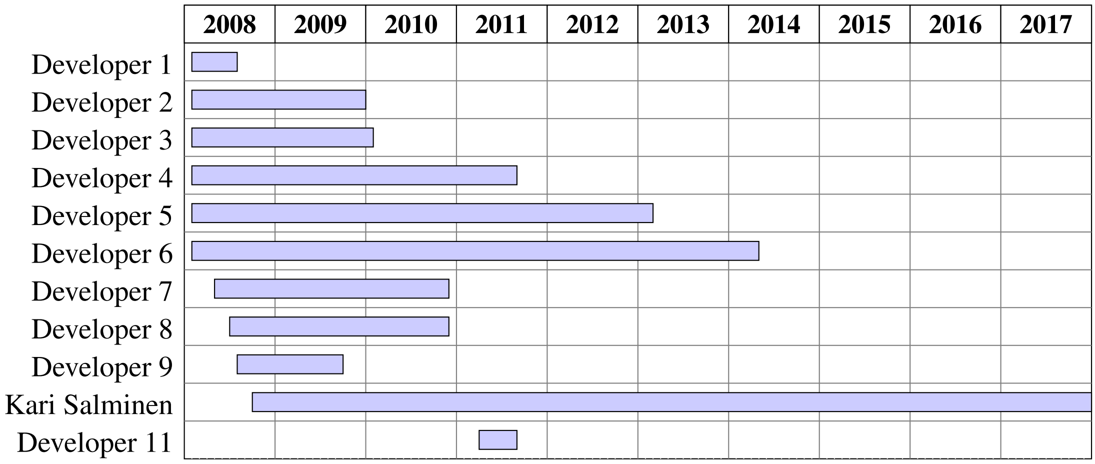


Figure 3.3: Developers' first and last commits to Tarla's code [13] (see Appendix A)

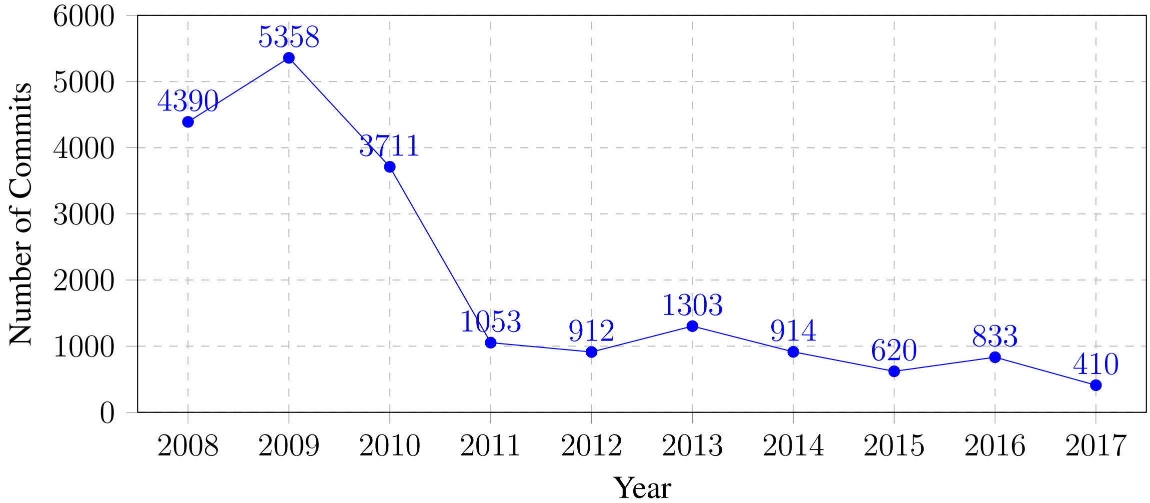


Figure 3.4: Number of commits to Tarla's SVN repository by year [13] (see Appendix B).

turn of the year and all the rest[1] followed suit during 2010 [7].

During 2012 Tarla was taken into use at Rautanet Forssa and Raisio franchise offices and at Keittiö-Kartano Oy's two offices in Pori [7]. The year 2013 was a year of expansion into the capital region as a business alliance between Keittiöjätti Oy and Kitchennet Oy was made and all of the latter's offices[2] started using Tarla. Also, two more franchise offices (Rauma and Värisilmä Forssa) adopted Tarla in 2013.

The year 2014 saw similar expansion but more towards Eastern Finland as a business alliance between Keittiöjätti Oy and Kymen Keittiötukku Oy was made and all of the latter's offices[3] started using Tarla. Kitchennet Oy also opened one more office in Espoo at Suomenoja using Tarla in 2014 [7].

At the very beginning of year 2015 Keittiöjätti Oy, Keittiö-Kartano Oy, Kitchennet Oy and Kymen Keittiötukku Oy were merged into one firm which was named Ideal Keittiöt Oy [39]. An office at Helsinki Kruununhaka was opened in the late summer [7] and in late autumn Ideal Keittiöt Oy's central warehouse moved from Hauho, Finland to Kohtla-Järve, Eastern Estonia[4] which was a major logistics change.

Early 2016 a new unified product catalog[5] was taken into use in Tarla. Previously there had been different product catalogs for different subsidiaries (e.g. one for Kitchennet Oy and one for Keittiöjätti Oy) which had complicated central product management. During 2016 a second office at Oulu was opened, Lahti office changed owners and a franchise office at Seinäjoki was opened [7]. During 2017 Rovaniemi, Tornio and Tammisto offices were opened [7] and three Lakkapää's [7, 41] offices in Northern Finland started selling Ideal Keittiöt Oy's items.

Altogether Tarla's adoption and Ideal Keittiöt Oy's expansion has been relatively

[1] Turku headquarters, Hauho central warehouse, Stella Interior factory outlet and Nummela office [7].

[2] At Helsinki Töölö, Espoo Muurala, Vantaa Varisto and Vantaa Porttipuisto [7].

[3] Kouvola, Lappeenranta, Kotka, Vantaa, Mikkeli and Lahti offices [7].

[4] The central warehouse's postal address was changed in Tarla on Oct 6, 2015 [40, #4881].

[5] Ideal product catalog was released on Feb 9, 2016 [39].

steady from 2012 onwards although Tarla's userbase has dwindled a little after 2015 (see Figures 3.5 and 3.6 for details).

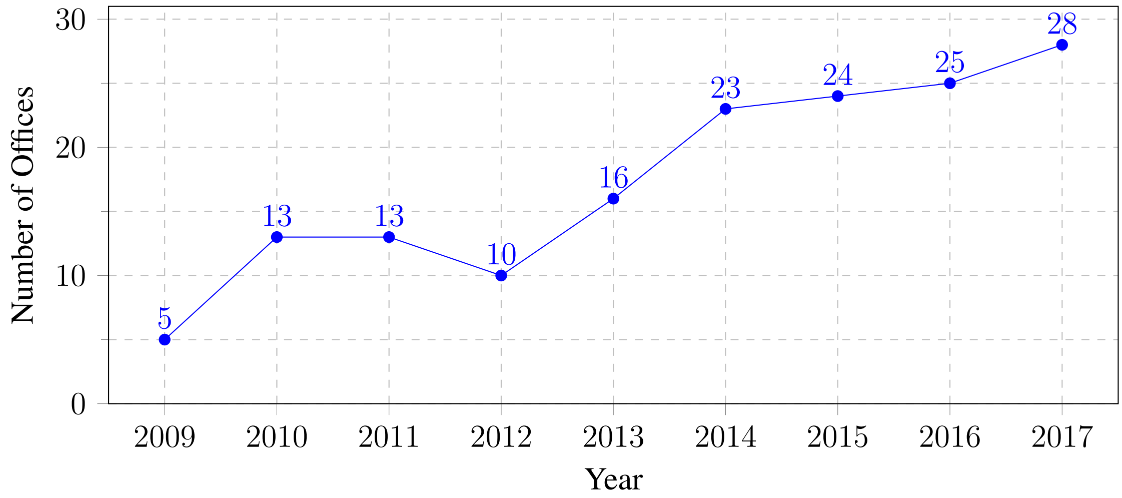


Figure 3.5: Number of offices selling at least €10,000 by year [7] (see Appendix D)

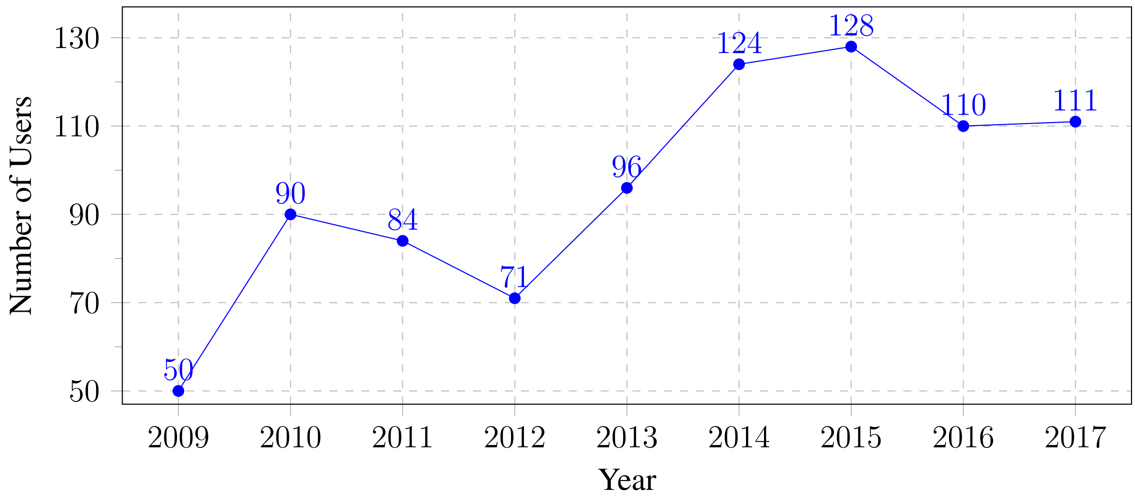


Figure 3.6: Number of different Tarla users per year [7] (see Appendix E)

Future of Tarla

In this section the future of Tarla is considered in 2018, when this thesis was recently started, and in 2020, when this thesis was completed.

Situation in 2018 Tarla is becoming a legacy software. It uses Qt 4.7.4 [13] which is over six years old [42] and it has a dependence on a pre-Qt 4.7.4 component [13]. Tarla's source code is relatively complex and custom made using Qt, MySQL, C++ and XSLT [13] (see Table 3.1) and currently maintained by one developer from a small pool of Tarla experts (see Figures 3.2 and 3.3).

Management has made a decision to change the company's ERP from Tarla to Jeeves [8]. The author of this thesis sees this as an amelioration of the risks involved in maintaining legacy software known to only a few and supporting new ways of doing business where, for example, changing logistical arrangements and growing project sales are taken more readily into account than before.

Some system integration has already been done between Tarla and Jeeves to handle invoice and cash register exports from Tarla to Jeeves and to handle warehouse stock synchronization both ways [13]. The author estimates it likely that Tarla will be used alongside Jeeves for at least half a year after its adoption.

Situation in 2020 Tarla was decommissioned from active use in Ideal Keittiöt Oy during 2019 [43].

Table 3.1: Lines of code in Tarla's source code trunk branch [13] (see Appendix F).

Language	Files	Blank	Comment	Code
SQL [1]	1131	33819	19927	500115
C++ [2]	881	43769	30981	219290
C/C++ Header [2]	1064	28112	55987	72327
XML [3]	300	17	2	42033
Python [4]	54	1180	946	8350
HTML [5]	3	304	9	6695
XSLT [6]	84	650	556	5563
Bourne Shell	86	1219	431	5366
IDL	22	205	0	1556
make	13	167	156	623
Objective C++	1	0	0	287
ASP.Net	6	25	0	149
DOS Batch	2	25	5	63
Javascript	1	18	7	53
Bourne Again Shell	1	5	3	16
	3649	109515	109010	862479

[1] Mostly database patch files.

[2] The main client and server code.

[3] Mostly Pentaho [10] reports (Unpacked .prpt files [14]).

[4] Tools (Build, development, deployment, system integration).

[5] Mostly Tarla's release notes.

[6] Used by the printing system for XML to SVG transformation.

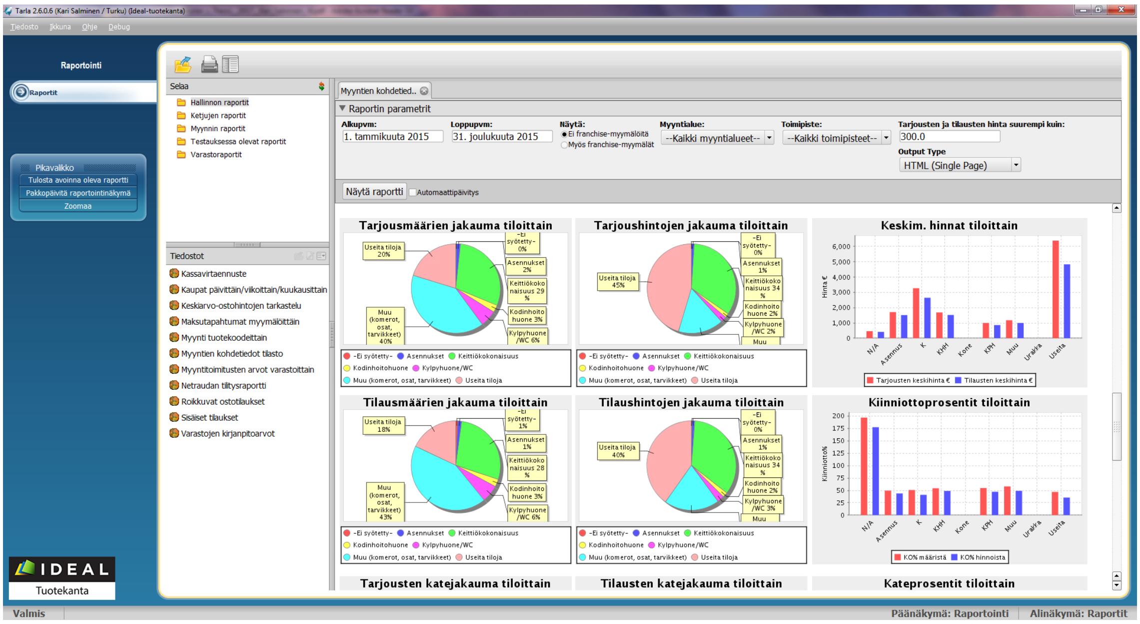


Figure 3.7: Example screenshot of Pentaho Reporting embedded in Tarla.

3.1.2 Pentaho Reporting

Initially reporting was not included in Tarla but less than a week after the release of Tarla 1.0 on Nov 11, 2010 [39] work started on the reporting system's requirement analysis [40, #3189] (see Table 3.2) and candidate surveying [40, #3190]. The community editions of Pentaho [10] and Jaspersoft [44] were chosen as the report system candidates from a pool of seven (Pentaho [10], Jaspersoft [44], OpenReports [45], Eclipse BIRT Project [46], SpagoBI [47], RapidMiner [48] and a self made system) [40, #3190]. Of the chosen two candidates Pentaho was deemed more promising [40, #3191]—albeit it was not readily available in Finnish and its report files were not version controllable (see Table 3.2).

A proof-of-concept report for testing Pentaho reporting was skipped over in favor of implementing the first real report which was payments [40, #3191, #3299]. Pentaho reporting was embedded into Tarla (see Figure 3.7) and its pilot use by the management started with the Tarla 1.3 release on March 3, 2011 [39]. This pilot release contained four reports: Payments, orders, sales by office and sales by seller (Commissions) [39].

Table 3.2: Requirements for Tarla's reporting system [40, #3189] and their fulfillment by Pentaho Reporting.

Requirement	**Fullfilled**	**Out-of-the-box**	**Selfmade**
Available in Finnish language	Yes		☑ [1]
Integrable into Tarla	Yes		☑ [2]
Support for MySQL database	Yes	☑	
Configuration version controllable	No		
Reports version controllable	Yes		☑ [3]
Support realtime reporting	Yes	☑	
Support scheduled reporting	Yes [4]	☑	
Support graphics	Yes	☑	
Creating new reports is easy	Partially	☑	
Has an open source license	Yes	☑	

[1] Translated to Finnish by Tarla's programmers [40, #3329].

[2] By accessing Pentaho BI Server [11] with Qt WebKit [49] [40, #3320].

[3] By unzipping .prpt [14] files to XML files [40, #3332].

[4] An unused feature.

History of Tarla reporting's use

After the management's pilot use of Pentaho reporting the store managers were granted access to it approximately three months later on June 6, 2011 [40, #3654]). For approximately the next two years, the users were the developers, management and store managers [50]. Starting from July 2013 [40, #3659] part of the reports were made available to all Tarla users: 12 total in 2013 and 5 total in 2014 (see Table 3.3's access group All).

Comparing user counts between Tarla and Pentaho reporting (see Figure 3.8) shows that in 2013–2017 most Tarla users have used Pentaho reporting at least once a year [50].

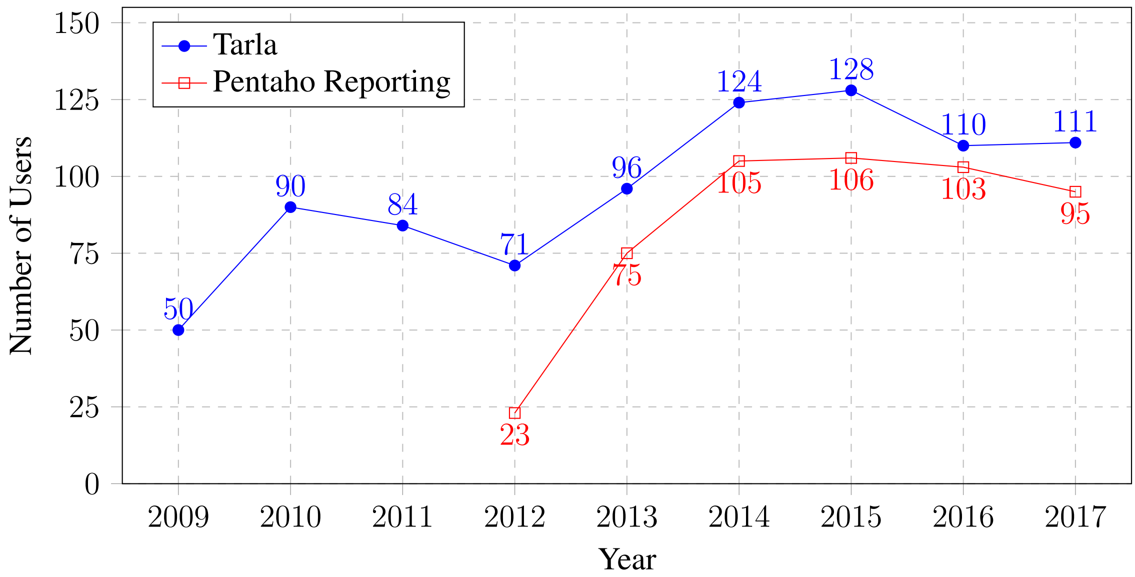


Figure 3.8: Number of different Tarla and Pentaho users per year [7,50] (see Appendices E and G). Pentaho Reporting statistics start from July 12, 2012 [50].

Tarla's reports and their use statistics

Total of 33 Pentaho reports have been made to date (see Table 3.3). During 2017 the top 9 accounted for over 90% of use, the top 16 for over 99% and the top 25 for over 99.9% (see Table 3.4). Reports were published at an approximate pace of 7.5 per year until the end of 2014 after which the development of new reports plateaued (see Figure 3.9).

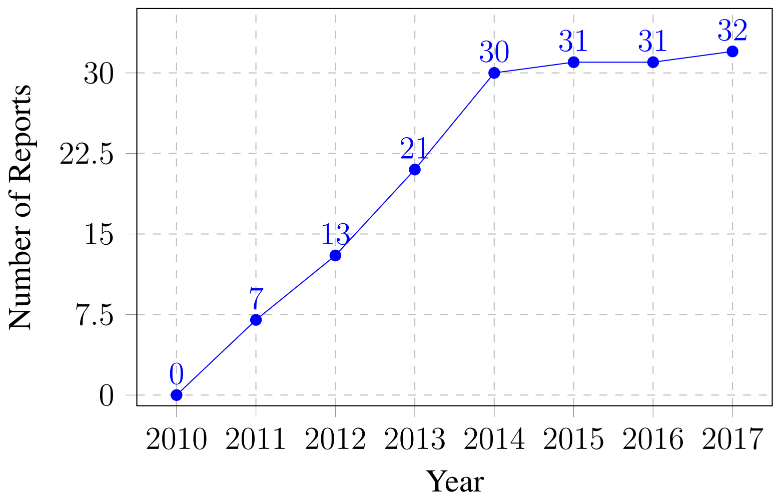


Figure 3.9: Number of reports available for use at the end of each year (see Table 3.3)

Table 3.3: Tarla's Pentaho reports in release order [7, 11, 13, 39, 40, 51]

No.	Finnish Report Name (*Original*)	English Report Name (*Translation*)	Released [1]	Access Group [2]
1.	Maksutapahtumat	Payments	2011-03-03	All
2.	Tilaukset	Orders	2011-03-03	All
3.	Kaupat myymälöittäin	Sales by office	2011-03-03	All
4.	Myynti myyjittäin	Sales by seller (Commissions)	2011-03-03	All
5.	Tarjoukset ja kaupat	Offers and sales	2011-04-18	All
6.	Tilauksen sisältö	Order contents (Order drilldown)	2011-05-09	All
7.	Keskusvarastotilanne	Central warehouse status	2011-06-14	Subsidiary
8.	Kassavirtaennuste	Projected cashflow from deliveries	2012-08-13	Management
9.	Tehdyt kaupat myyjittäin	New sales by seller	2012-08-14	All
10.	Myynti tuotekoodeittain	Sales by product code	2012-08-15	All
11.	Maksutapahtumat myymälöittäin	Payments by office	2012-11-10	Management
12.	Kaupat päivittäin/viikoittain/kuukausittain	Sales by time period	2012-11-10	Management
13.	Varastojen kirjanpitoarvot	Warehouse values	2012-11-13	Management
14.	Tuoteryhmän varastotiedot aikavälillä	Group warehouse data over time	2013-02-25	Warehouse
15.	Varastotuotteiden kiertonopeus	Inventory turnover	2013-02-25	Warehouse
16.	Tuotteen varastotiedot aikavälillä	Product warehouse data over time	2013-02-25	Warehouse
17.	Tuoteryhmien kiertonopeus	Turnover by product group	2013-02-25	Warehouse
18.	Omat maksutapahtumat	My payments	2013-03-08	All
19.	Myynti ja budjetit kuukausittain	Month sales budgets by seller	2013-03-15	All
20.	Varaston kasauskuorma	Warehouse assembly load	2013-06-28	All
21.	Avoimet tarjoukset	Open offers	2013-10-25	All
22.	Inventaariot	Inventories	2014-04-01	Warehouse
23.	Tilausten kohdetiedot	Orders' metadata	2014-04-04	All
24.	Tarjousten kohdetiedot	Offers' metadata	2014-04-07	All
25.	Myyntien kohdetiedot tilasto	Sales events' metadata statistics	2014-04-14	Management
26.	Ideal kampanjaraportti	Ideal campaign purchases [3]	2014-04-14	All
27.	Sisäiset tilaukset	Internal purchase orders	2014-05-28	Management
28.	Kuluvan kuukauden kaupat	Current month sales	2014-08-29	All
29.	Myyntitoimitusten arvot varastoittain	Sales delivery values by warehouse	2014-10-29	Management
30.	Kuukauden kaupat myymälöittäin	Daily sales by office in month	2014-12-12	All
31.	Keskiarvo-ostohintojen tarkastelu	Average purchase price inspection	2015-01-02	Management
32.	Roikkuvat ostotilaukset	Uncompleted purchase orders	2015-03-18	Management
33.	Netraudan tilitysraportti	Netrauta report	2017-06-21	*Special Dual*

[1] Approximate [40] because few were mentioned in release notes [39].

[2] Report access is controlled on access group level. *Special Dual* means Management and Netrauta.

[3] The only report to be removed from production use. Removed on February 11, 2015 [40, #4812].

Table 3.4: Use statistics of Tarla's Pentaho reports for year 2017 [50, 51] (see Appendix G). Top 3 reports in each user category (**Top** = *All users*, **M** = *Management*, **S** = *Sales*, **W** = *Warehouse*) are color coded to support rapid perception [22, p. 152].

Top	M	S	W	Report Name	Views	% of Total	Cumulative Views	% of Total
1	1	5	4	Offers and sales	22621	20.96%	22621	20.96%
2	2	2	2	Payments	20342	18.85%	42963	39.81%
3	3	1	3	Sales by seller (Commissions)	18311	16.97%	61274	56.78%
4	5	6	1	Warehouse assembly load	11792	10.93%	73066	67.71%
5	7	3	6	Current month sales	8262	7.66%	81328	75.36%
6		4	7	New sales by seller	5557	5.15%	86885	80.51%
7	4			Month sales budgets by seller	5334	4.94%	92219	85.45%
8	6	9	9	Sales by office	4158	3.85%	96377	89.31%
9		7		My payments	2781	2.58%	99158	91.88%
10	8			Daily sales by office in month	1671	1.55%	100829	93.43%
11		8		Orders	1516	1.40%	102345	94.84%
12		10		Order contents (Order drilldown)	1387	1.29%	103732	96.12%
13	9		5	Sales by product code	1178	1.09%	104910	97.21%
14	10			Netrauta report	768	0.71%	105678	97.92%
15				Offers' metadata	668	0.62%	106346	98.54%
16				Open offers	502	0.47%	106848	99.01%
17			8	Product warehouse data over time	231	0.21%	107079	99.22%
18				Sales by time period	176	0.16%	107255	99.39%
19				Warehouse values	159	0.15%	107414	99.53%
20				Orders' metadata	147	0.14%	107561	99.67%
21				Internal purchase orders	85	0.08%	107646	99.75%
22				Inventories	74	0.07%	107720	99.82%
23				Sales delivery values by warehouse	56	0.05%	107776	99.87%
24			10	Inventory turnover	25	0.02%	107801	99.89%
25				Sales events' metadata statistics	25	0.02%	107826	99.91%
26				Projected cashflow from deliveries	24	0.02%	107850	99.94%
27				Uncompleted purchase orders	18	0.02%	107868	99.95%
28				Central warehouse status	14	0.01%	107882	99.97%
29				Group warehouse data over time	12	0.01%	107894	99.98%
30				Payments by office	12	0.01%	107906	99.99%
31				Turnover by product group	9	0.01%	107915	100.00%
32				Average purchase price inspection	3	0.00%	107918	100.00%

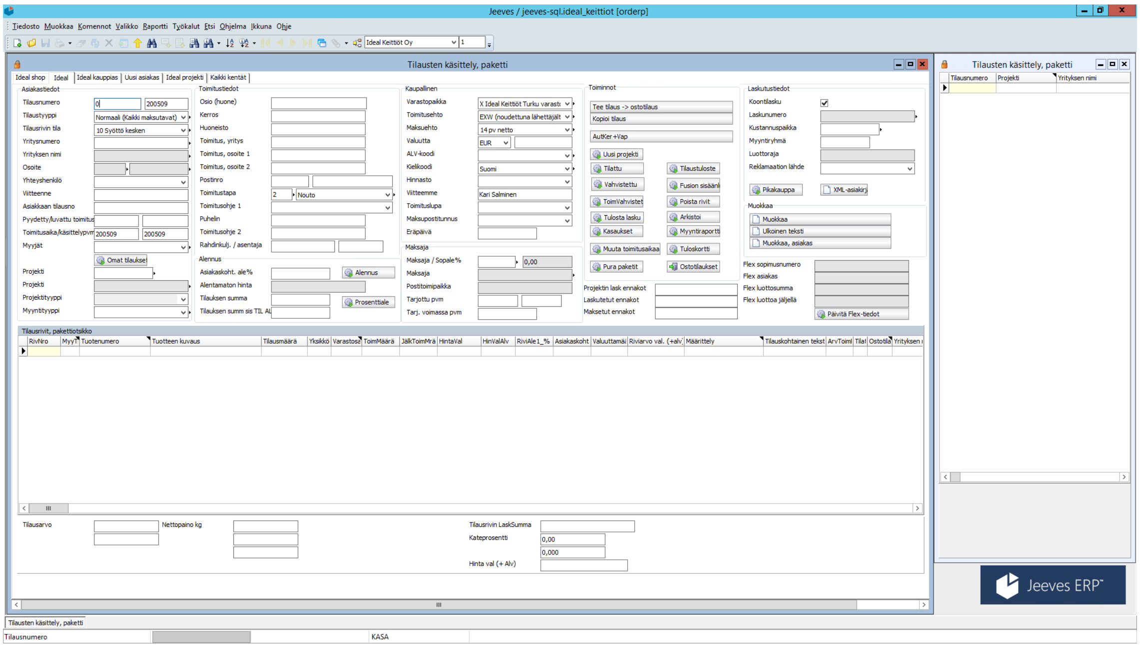


Figure 3.10: Example screenshot of the Jeeves's client program.

3.2 New ERP system and BI solution

In this section Ideal Keittiöt Oy's new ERP system named Jeeves (see Section 3.2.1), its database (see Section 3.2.1) and the new BI solution named Qlik Sense® (see Section 3.2.2) are described. As this section was mostly written in 2020 after these systems were already taken into active use it also contains information about the deployment these systems (Sections 3.2.1 and 3.2.2).

3.2.1 Jeeves

Jeeves (see Figure 3.10) is a commercial enterprise resource planning (ERP) system from Sweden [9]. Ideal Keittiöt Oy's management chose it to be the replacement for the company's former custom ERP system named Tarla [8]. It is the assessment of the author of this thesis that a key advantage of Jeeves over Tarla is its customizability vs. development time factor. Tarla was indeed highly customized but it became more difficult to change it over time as its architecture was quite monolithic and consisted of mostly highly interdependent parts [13]. Jeeves can be highly customized but it tries to keep agility alive by using an architecture that keeps customizations separate from the standard system [52].

Jeeves's database

The information in this section is based mainly on the practical knowledge gained through working on implementing a report from Jeeves's database [17] by the author of this thesis during late 2018 and very early 2019. The author wrote an internal guide to Ideal Keittiöt Oy for Jeeves reporting as his master's project in November 2019 (The original Finnish title *Ideal Keittiöt Oy:n sisäinen ohje Jeeves ERP -raportointiin*).

Using Jeeves as a data source for reporting requires detailed knowledge of its database. Its database is very complex: it has 103,262 columns total, 2659 tables and 514 views (see Appendix I). Table names and column names are normally in English or Swedish and often abbreviated[1]. Tables and views have on average 32 columns within a range of 1–492 columns (see Appendix I). [17]

The author of this thesis is still in May 2020 not aware of any open source or free solution for using a Jeeves database for reporting. The only commercial alternative which the author of this thesis found was the Jeeves Business Intelligence [53, 54] which was unfortunately made for QlikView, not Qlik Sense, and thus not directly usable for Ideal Keittiöt Oy's purposes.

There is however a basic table level and table column level documentation in English that comes with a Jeeves installation [18]. This documentation was very useful in trying to understand the Jeeves's database although it in itself was not enough—a lot of additional work had to be done to understand the database because of its complexity. For example, knowing what a column does may not be enough because there may be several columns that serve a similar looking function in a single table and it is not necessarily at all readily apparent which one of the columns should be used in practice. Also the

[1] E.g. FsgArtOmvFaktor, MomsGrundBelopp, orp_delivtopolnno, ProjCostTime, rabatt, xwvw. [17]

relationships between tables are not readily apparent because Jeeves's database has no foreign keys [17].

The author of this thesis furthered his understanding of the Jeeves's database by using the Jeeves client program's developer support features, e.g. SQL logging, macro editor and ALT-CTRL-F12 for switching between showing the user interface elements and their data source. Also the reading of stored procedures, triggers, keys and indices from the Jeeves's database [17] and commercial consultations of Ideal Keittiöt Oy's partners [20, 21] were used. Without a doubt the consultation of the Jeeves solution partner Staria Oyj [21] was instrumental in this process—without their help the process would have definitely taken longer.

Through combining the knowledge gained from all of the above methods the author of this thesis was able to make his first report *Tuloskortti* (i.e. *Scorecard*) near the end of 2018 using the databases of both Jeeves and Tarla as its sources. This was an office level sales report and it was made using SAP Crystal Reports 2016 [55] because Jeeves supported it directly and it did not require token licensing.

Unfortunately soon after this the employee contract of the author of this thesis was terminated on February 4, 2019 on financial and production-related grounds. As mentioned earlier the author of this thesis did however write an internal guide to Ideal Keittiöt Oy for Jeeves reporting as his master's project in November 2019 via which as least part of the gained knowledge of the Jeeves's database was documented.

Deployment of Jeeves

Jeeves was taken into active use in Ideal Keittiöt Oy on January 1, 2019 [43]. Non-franchise business users in Ideal Keittiöt Oy switched to it immediately and most franchise business users switched to it during spring and summer of 2019 [43]. In May 2020 the CIO of Ideal Keittiöt Oy assessed that the use of Jeeves will probably continue for many years to come and that there are currently no plans for replacing it with anything else.

Even if 2020 Fusion [56] or Fusion Sales [57] were to be taken into use in Ideal Keittiöt Oy—and this was not a given—Jeeves would still remain underneath. [43]

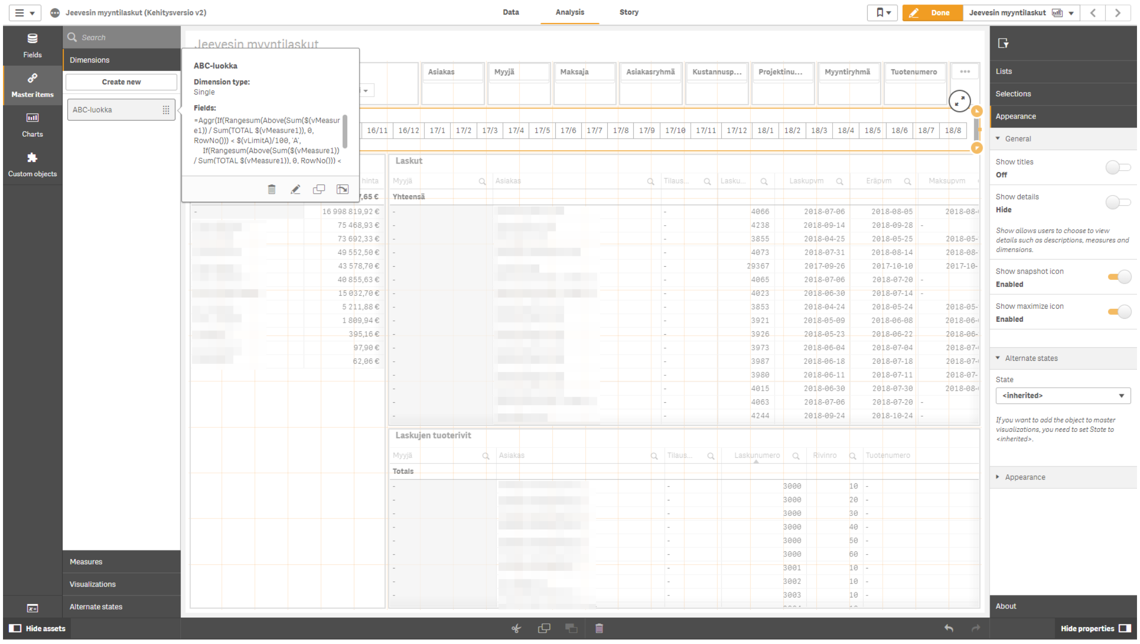


Figure 3.11: Example screenshot of Qlik Sense. Sensitive information has been pixelized.

3.2.2 Qlik Sense

Ideal Keittiöt Oy's preferred reporting system for Jeeves is Qlik Sense® (see Figure 3.11) which is a business intelligence tool for data discovery. It uses an associative engine which works in-memory if not limited by the data size—for that use case Qlik® offers Qlik Big Data Index™. The in-memory model supports changing filters on the fly and calculation on demand. This supports an experimental way of working with data. Users can try out different filters and see how the data looks with them—this is what is meant by data discovery. [12, 15, 16]

A clear difference between Pentaho Reporting (see Section 3.1.2) and Qlik Sense® is that previously the reports (see Section 3.1.2) were more rigid, the adjustable filters were chosen by the report developers and the calculation on demand was on average slower[1]

[1] Anecdotal evidence only. Speeds might be compared using data from both systems' logs [16, 50].

than with Qlik Sense. With Qlik Sense® the choice of filters can be changed on the fly and calculation on demand is on average faster[1] than with Pentaho Reporting. [10, 12, 15, 16, 51]

On the whole Qlik Sense® is more an exploratory data tool than Pentaho Reporting—at least in the way in which Pentaho Reporting was used in Ideal Keittiöt Oy. Pentaho Reporting was used for traditional reports by querying a database with SQL queries and having a fixed set of filters per report while Qlik Sense® can be used more as a tool to explore data because of its capability to change the set of filters on the fly. [10, 12, 15, 16, 51]

Deployment of Qlik Sense

Qlik Sense® based sales reports using Jeeves as their data source were developed by Pengon Oy [20] during 2019 and deployed for use by Ideal Keittiöt Oy's management in October 2019 [43]. In May 2020 the CIO of Ideal Keittiöt Oy assessed it probable that Qlik Sense® will continue to be used by the management, its use will not expand extensively nor become a tool for sales persons [43].

3.3 Ideal Keittiöt Oy's timeline 2009–2017

The timeline in Figure 3.12 is a concise representation of the major events described in Chapter 3 until the end of 2017 (events before 2009 have been left out to fit events onto the timeline):

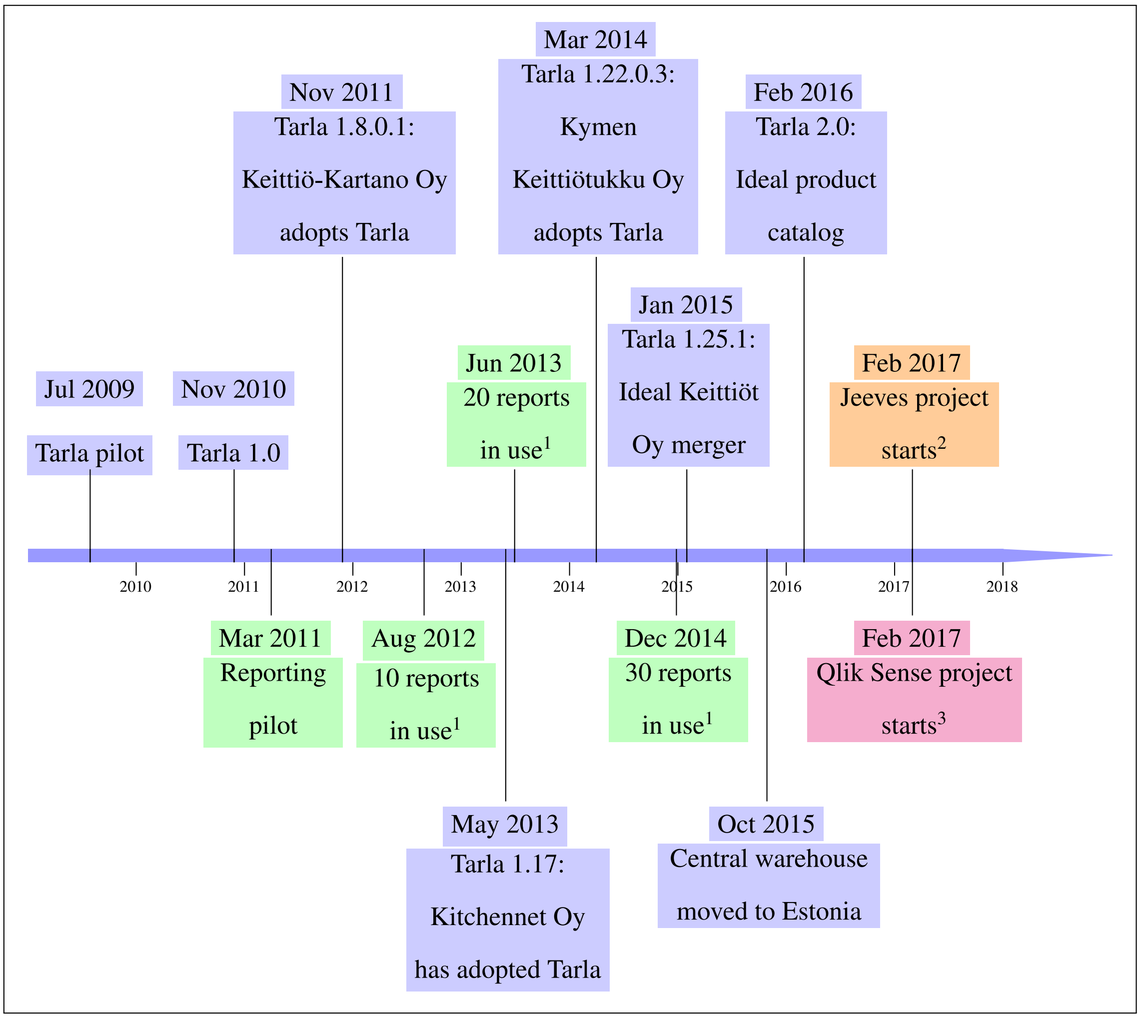
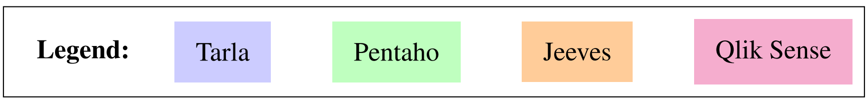


Figure 3.12: Ideal Keittiöt Oy's timeline 2009–2017 [7, 39, 40], Table 3.3[1], [8][2], [58][3]



Chapter 4

Reporting needs at Ideal Keittiöt Oy

In this chapter the intended end-users (see Section 4.1) of the new reports in Ideal Keittiöt Oy are interviewed about their needs for new reports (see Section 4.2) and based on their needs (see Figure 4.1) and previous report use (see Section 4.3.1) the new reports to be implemented are chosen (see Section 4.3.2).

What are the *key parameters* to consider in creating new reports for Ideal Keittiöt Oy?

1. Intended end-users
2. User requirements
3. How to fulfill user requirements most efficiently

In this chapter these key parameters are considered each one in its own section.

4.1 Intended end-users

Tarla's reporting was based on an open source software solution [11] and had no license costs at all. Thus the size of its target audience was not limited by the reporting system's license costs and it was 95 users in 2017 (see Figure 3.8). In constrast the new BI tool Qlik Sense is proprietary software and has annual license costs [59] which limit its adoption. The selected target audience is the management [60] and the amount of user access passes [59] is limited to approximately ten [60].

Table 4.1: The new reports' intended end-users [61].

No.	Management Section	Intended End-User's Job Title
1	Sales and Marketing	Chief Commercial Officer (CCO)
2	Sales and Marketing	Marketing Coordinator
3	Product	Product Coordinator
4	Information Technology	Chief Information Officer (CIO)
5	Project Sales	Project Manager
6	Financial	Chief Financial Officer (CFO)
7	Product	Technical Product Manager
8	Financial	Management Assistant / Liaison
9	Audit / Reclamations	Bill Auditor / Reclamation Manager
10	Executive	Chief Executive Officer (CEO)
11	Area Management	Western and Northern Finland Manager
12	Area Management	Helsinki Metropolitan Area Manager

4.2 User requirements

The intended end-users were interviewed during week 10 of 2018 about their reporting needs for Jeeves. Most interviews were done by phone as e-mail proved an ineffective interview method, a few were done face to face. All but the product coordinator (Intended end-user no. 3) were available at the time and were interviewed. Most interviews took approximately 10–15 minutes, a few less and a few more with the longest two being 30 and 45 minutes.

Notes were written down on a computer while conducting the phone interviews resulting in approximately 1800 words. Subsequently a mind map (see Figure 4.1) was constructed from the interview notes in order to distill their contents. To include it in this thesis it was translated to English and edited for space and confidentiality.

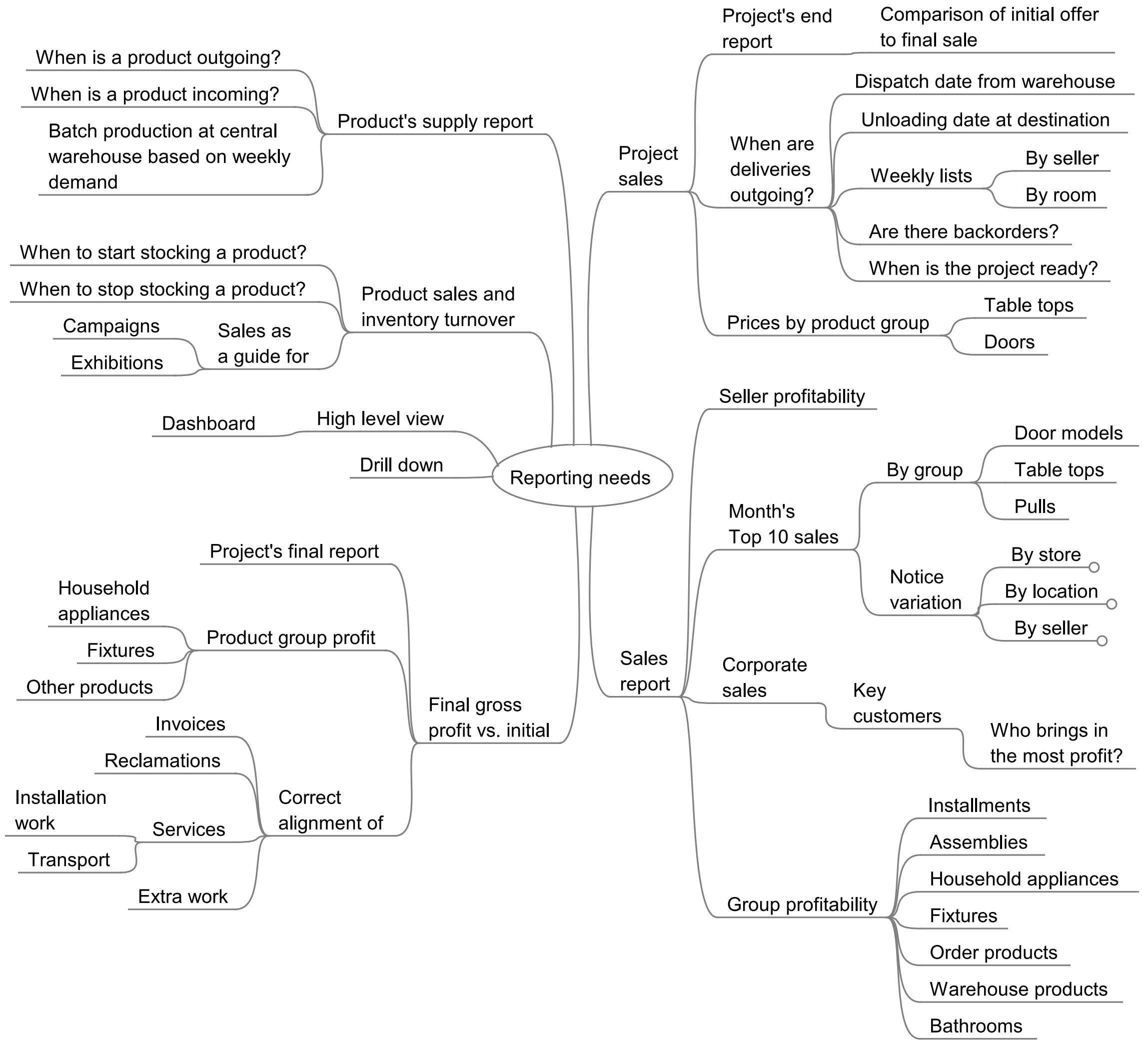


Figure 4.1: Intended end-users' reporting needs mind map. Translated from Finnish to English, edited for space and confidentiality. Three nodes (By store, By location, By seller) on the right side have been collapsed to conserve horizontal space.

4.3 How to fulfill user requirements most efficiently

In this section based on the previous report use (see Section 4.3.1) and reporting needs of the intended end-users the new reports to be implemented are chosen (see Section 4.3.2).

4.3.1 Reflection on previous report use

From the use statistics of Tarla's Pentaho reports in 2017 (see Table 3.4) it is clear that the bulk of the most used reports has been the monitoring of sales and of incoming cash. The only very significant addition to this is the monitoring of warehouse assembly load which accounts for approximately 11% of all report use. It is notable that the Pareto principle [62] (colloquially the 80/20 rule) applies to the use statistics because the top 20% ($= 32 * 0.2 = 6.4 \approx 6$) of reports account for approximately 80% (80.51%) of use.

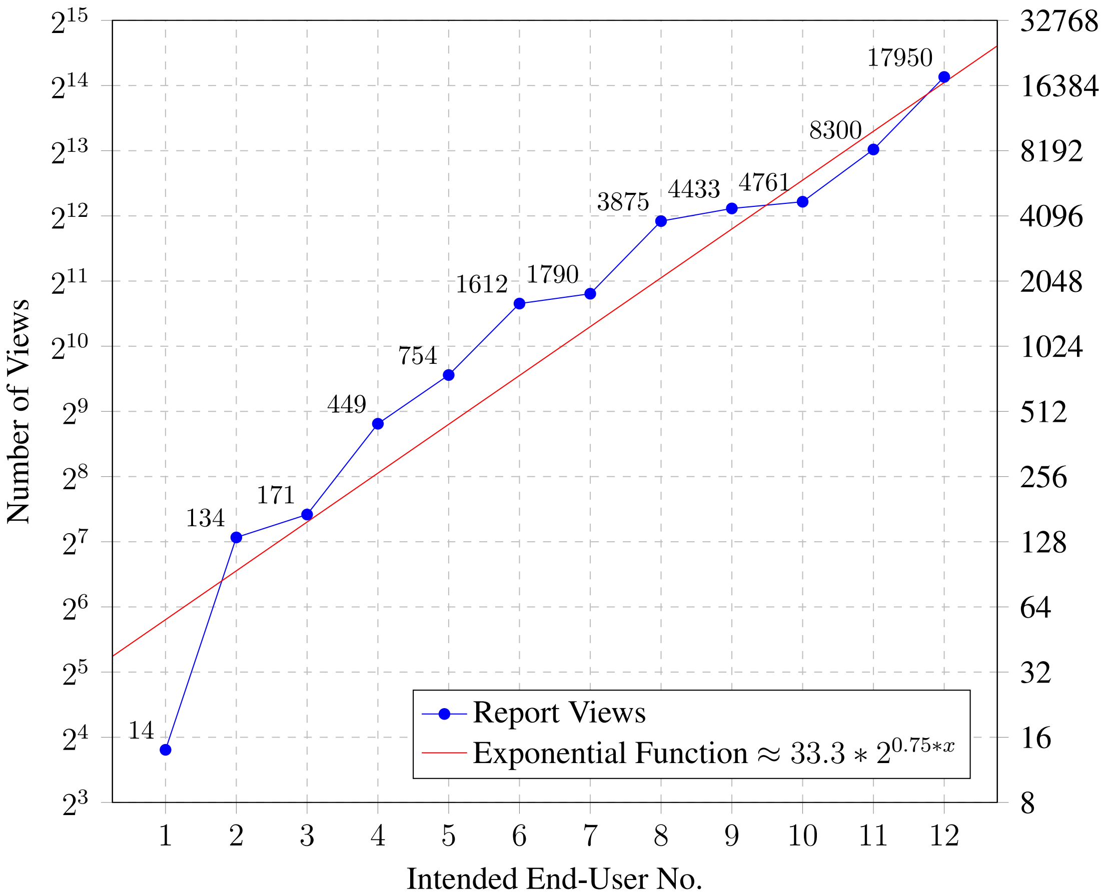


Figure 4.2: The intended end-users' report views in year 2017 in ascending order. The fitted exponential function is $y = 33.2529 * 2^{0.749431*x}$, mean absolute error $= 717.87$ [35, 63]. This shows that the growth of report use in this group is exponential.

Judging by the interview notes and the management's report use in 2017 (see Tables 3.4 and 4.1, and Figure 4.2) the area managers (i.e. intended end-users 11–12) are extremely eager for timely sales data as their report use is on average at least 22 reports each day of the year. Also the financial management assistant, the bill auditor / reclamation manager and the CEO (i.e. intended end-users no. 8–10) are avid report users as their report use is on average approximately 12 reports each day of the year. The CFO and the technical product manager (i.e. intended end-users no. 6–7) are moderate users of reporting as their report use is on average approximately 5 reports each day of the year. All the other management users (i.e. intended end-users 1–5) seem to be not very heavy users of reporting—although it is notable that a couple of them have only been employed recently by Ideal Keittiöt Oy which also reflects on their report use.

4.3.2 Reflection on intended end-users' reporting needs

The mind map of the intended end-users' reporting needs (see Figure 4.1) was presented at a meeting held between the author of this thesis and Ideal Keittiöt Oy's Chief Commercial Officer (CCO) during week 11 of 2018. The presented ideas were weighted and discussed, and a consensus was reached that given the time and resources constraints of the thesis two types of reports would serve the intended end-users' reporting needs most efficiently:

1. ABC analysis report ("Product sales and inventory turnover" in the mind map):
   Uses ABC analysis [64] to determine what products to stock and what not to.
   ⇨ *See the second highest left branch in Figure 4.1 for the mind map source*
   ➡ *See Section 5.1 for the implementation*
2. Sales report:
   Includes the possibility to drill down into the details at sold product level.
   ⇨ *See the lower right branch in Figure 4.1 for the mind map source*
   ➡ *See Section 5.2 for the implementation*

Chapter 5

Implementation of the new reports

In this chapter, the new reports that were implemented are described and examined in detail. It was deemed acceptable to use Tarla's data as the new reports' source given that Jeeves was not yet in widespread use in Ideal Keittiöt Oy at the time of their implementation [60]. The reasoning was that Jeeves did not have real data yet to analyze [60].

5.1 ABC analysis report

The ABC analysis report (see Section 4.3.2 and Figure 5.1) was implemented using Qlik Sense® (see Section 3.2.2), Tarla's data (see Section 3.1.1) and in Finnish language. It was named *Toimitusten ABC-analyysi* in Finnish which means ABC analysis of deliveries. The report was published into the Development stream of Ideal Keittiöt Oy's Qlik Sense Hub on April 9, 2019. [65]

Because the report was implemented in Finnish, many terms are given in both Finnish and English in this section. All the Finnish terms are written in cursive, e.g. *Esimerkki*.

5.1.1 ABC analysis

ABC analysis [64] is a generalization of the Pareto principle [62] (colloquially the 80/20 rule). The Pareto principle is that usually a small portion of a set's elements have a large effect and a large portion of its elements have a small effect [62], e.g. 80% of report use comes from the 20% of most used reports and 20% of report use comes from the 80% of least used reports. The Pareto principle divides data into two parts but ABC analysis divides data into three parts: A, B and C where the effect per element ratio is highest in A, moderate in B and lowest in C [62].

In this report, single-criterion ABC analysis [66] is used for analyzing the importance of a dimension of sold, delivered products according to their measure, e.g. models (i.e. the chosen dimension) of sold, delivered products according to their profit value (i.e. the chosen measure). The exact details of how elements are evaluated to be in A, B or C class in this report (see the next paragraph and Section 5.1.6) differ from the given sources [62, 64, 66]. This is a consequence of the author of this thesis having created the implementation mostly in 2018 without any reference book for ABC analysis.

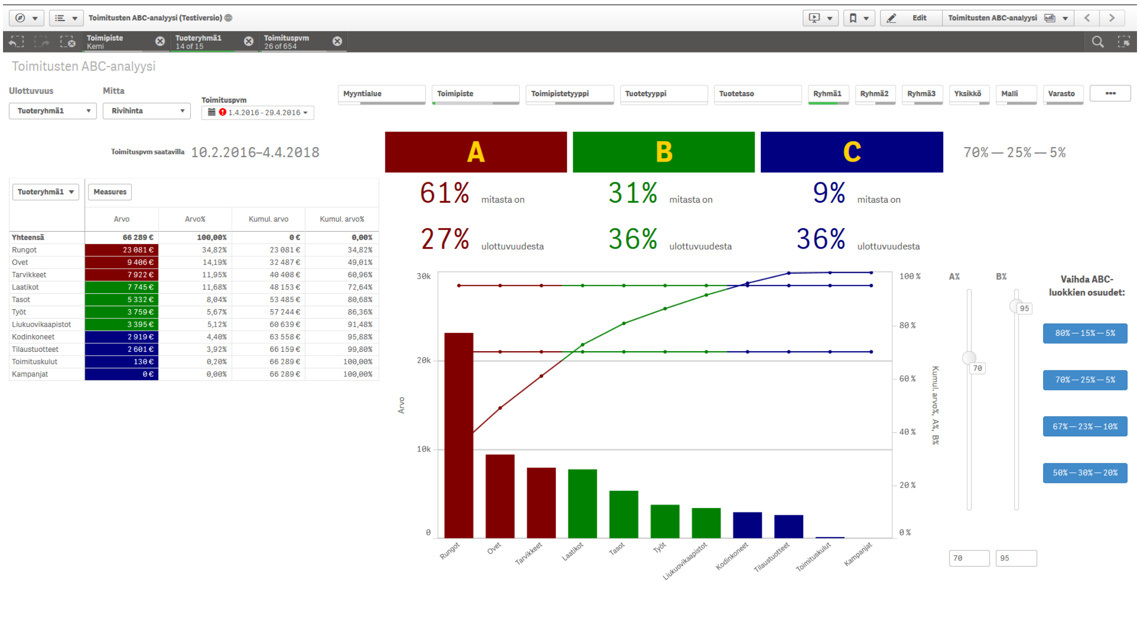
In this paragraph, dimension means the currently chosen dimension and measure means the currently chosen measure. The ABC classes are defined by controlling two values: A% and B%. In class A are the products grouped by dimension whose effect according to measure in descending order of magnitude is less than A% of the total effect of all products. In class B are the products grouped by dimension whose effect according to measure in descending order of magnitude is less than B% of the total effect of all products and that do not belong in class A. In class C are the products grouped by dimension that do not belong in A or B. [65]

5.1.2 User interface

The user interface of the ABC analysis report is described in this section. The reader is invited to refer to Figure 5.1 when reading this section.

The dimensions available for use in the ABC analysis report are *Tuoteryhmä1–Tuoteryhmä3* (Product group hierarchy levels 1–3), *Tuotekoodi* (Product code), *Malli* (Model), *Oletustoimittaja* (Default supplier), *Asiakas* (Customer), *Kohdetila* (Target space), *Ko-*

(a) Example screenshot of the ABC analysis Qlik Sense report.



(b) ABC analysis report with added explanations.

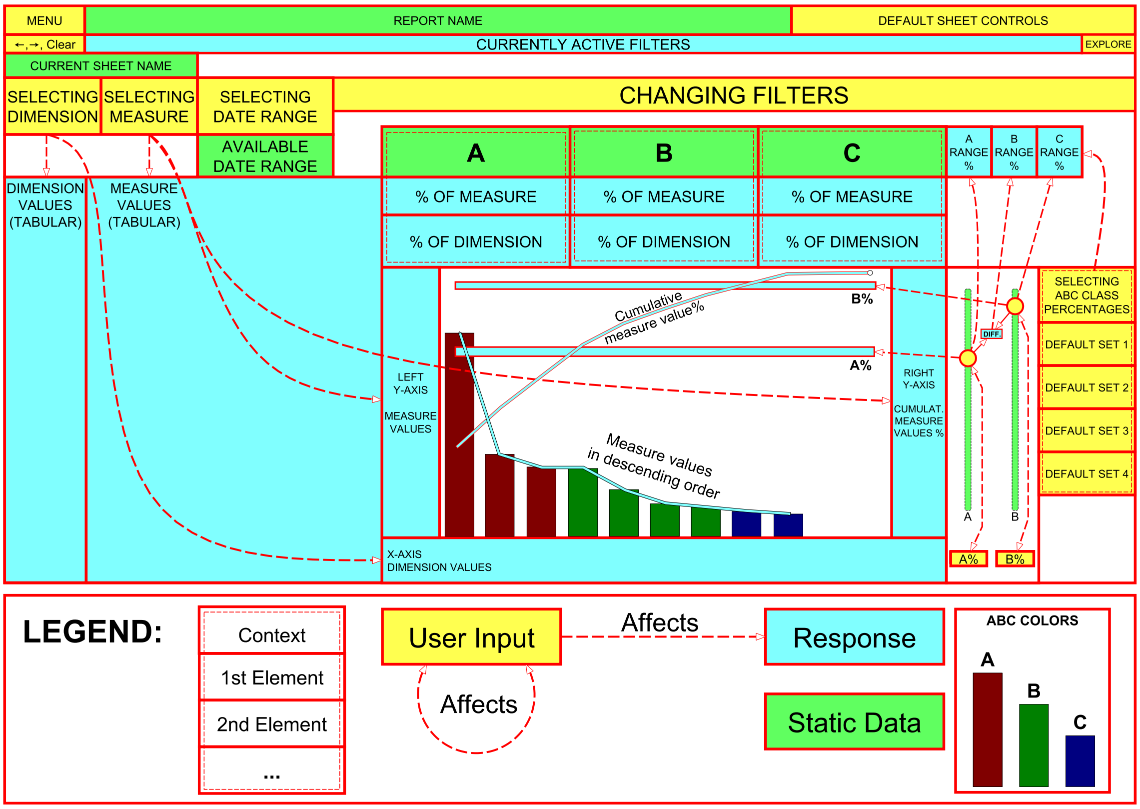


Figure 5.1: ABC analysis report visually explained. Because the vector graphics in (b) were originally drawn on top of the screenshot in (a) they are directly comparable.

hdetyyppi (Target type), *Asiakastyyppi* (Customer type), *Myyjä* (Seller), *Tarjousnro* (Offer number), *Toimipiste* (Office) and *Varasto* (Warehouse). A dimension can be chosen by the drop down box *Ulottuvuus* (Dimension) in the top left corner of the report. [65]

The measures available for use in the ABC analysis report are *Rivihinta* (Product row sales price), *Rivikate* (Product row profit), *Hankintarivihinta* (Product row acquirement price) and *Määrä* (Product quantity). The price and profit measures are without VAT. A measure can be chosen by the drop down box *Mitta* (Measure) near the top left corner of the report. [65]

The ABC classes can be changed by moving the A% or B% vertical slider (on the right side of the graph), by clicking any of the default ABC class selection buttons (on the right) or by writing a new value to the text box under the A% or B% vertical slider. [65]

The date range can be changed by selecting a date range in the *Toimituspvm* (Date of delivery) date range selector near the top left corner of the report. [65]

Filters for *Myyntialue* (Sales area), *Toimipiste* (Office), *Toimipistetyyppi* (Office type), *Tuotetyyppi* (Product type), *Tuotetaso* (Product hierarchy level), *Ryhmä1–Ryhmä3* (Product group hierarchy levels 1–3), *Yksikkö* (Unit), *Malli* (Model), *Varasto* (Warehouse) can be changed at the top of the report and more filters can be accessed by pressing the ellipsis button in the top right corner of the report. [65]

The currently chosen dimension can be filtered by clicking on a box with its name in the top left corner of the table on the left, by selecting values from the leftmost column of the table on the left or by clicking on the bars in the graph. [65]

According to the A% and B% values the real percentage portions of the A, B and C classes as calculated from the data are shown in the report above the graph. The dimension values (under *Yhteensä*), the measure values (under *Arvo*), the percentages of the measures from total measure (under *Arvo%*), the cumulative measure sums (under *Kumul. arvo*) and the percentages of the cumulative measure sums from the total measure (under *Kumul. arvo%*) are shown in the table on the left side of the report. In this table on the first data row to the right of *Yhteensä* (Total) are the total values for each column. [65]

The graph in the middle of the report shows a subset of the information shown in the table on the left of the report. The dimension values are shown on the graph's X axis. The measure values are shown in the graph as the bars' vertical heights and their scale is on the left side of the graph. The cumulative measure sum is shown as curve from left to right in the graph and its scale is on the right side of the graph. [65]

The ABC class percentage values based only on the values of A% and B% are shown in the top right corner of the report. The shown values are A%, B%−A% and 100%−B%. For an example, if A% is 75% and B% is 95% then values 75%, 20% and 5% are shown in the top right corner. These values usually differ from the ones above the graph because they are based also on the product data in addition to the values of A% and B%. [65]

5.1.3 Visual design

The ABC classes were color coded using clearly distinct colors red, green and blue [22, p. 111] to support their rapid perception [22, p. 152]. These colors were used to preattentively [22, pp. 154–155] separate rows of data, vertical bars in the graph and single values in the report to either the A, B or C class (see Section 2.4).

The real percentage values of the A, B and C classes and their portions of the chosen dimension were chosen to be shown as big numbers near the center of the report above the graph. This was done to draw attention to these important values because humans tend to look at big numbers early on when looking at a dashboard [67].

A graph with two Y axes, the left one for the measure values and the right one for the cumulative measure values, was chosen to visually represent the data that is being inspected. The graph was positioned in the middle of the report in order to make it and its data prominent (see Section 2.7).

The most important user interface elements requiring input were put in the top left corner of the report because humans—regardless of whether they are East Asians, Middle

Easterners or Westerners—tend to fixate more on the left and on the top than on the right or the bottom [26] (see Section 2.3). These user interface elements were laid out in a left-to-right order to accommodate Westerners [26], which was justified by the intended end-users (see Section 4.1) all being Westerners.

The dimension level data was put in a table on the left side of the report where the measure values and their cumulative sum are shown in descending order. The left side of the report was chosen to make the data more prominent as people tend to fixate more on the left than on the right [26].

The Gestalt law of proximity (see Section 2.6 and Figure 2.5) was used for making the connection between the position of the vertical sliders (A% and B%) on the right side of the graph and the vertical position of the horizontal reference lines in the graph more apparent. As an example, if the user moves the vertical slider A% up or down the horizontal reference line A% moves similarly in the graph. The connective effect is not very strong though because the vertical positions of the reference lines and the vertical sliders differ slightly and a reference line only moves after a vertical slider has been released. If the reference line and vertical slider would have been synchronized even while moving the vertical slider then the Gestalt law of common fate (see Section 2.6 and Figure 2.12) would have made their connection more apparent.

The Gestalt law of closure (see Section 2.6 and Figure 2.11) is used in many parts of the report to segment the visual space into parts by having a closed contours around objects. For example, the graph in the middle has a rectangle enclosing it, the data grid on the left has a rectangle enclosing it and the drop down boxes in the top left corner have rectangles enclosing them. In fact the whole report is enclosed in a rectangle which defines a common area.

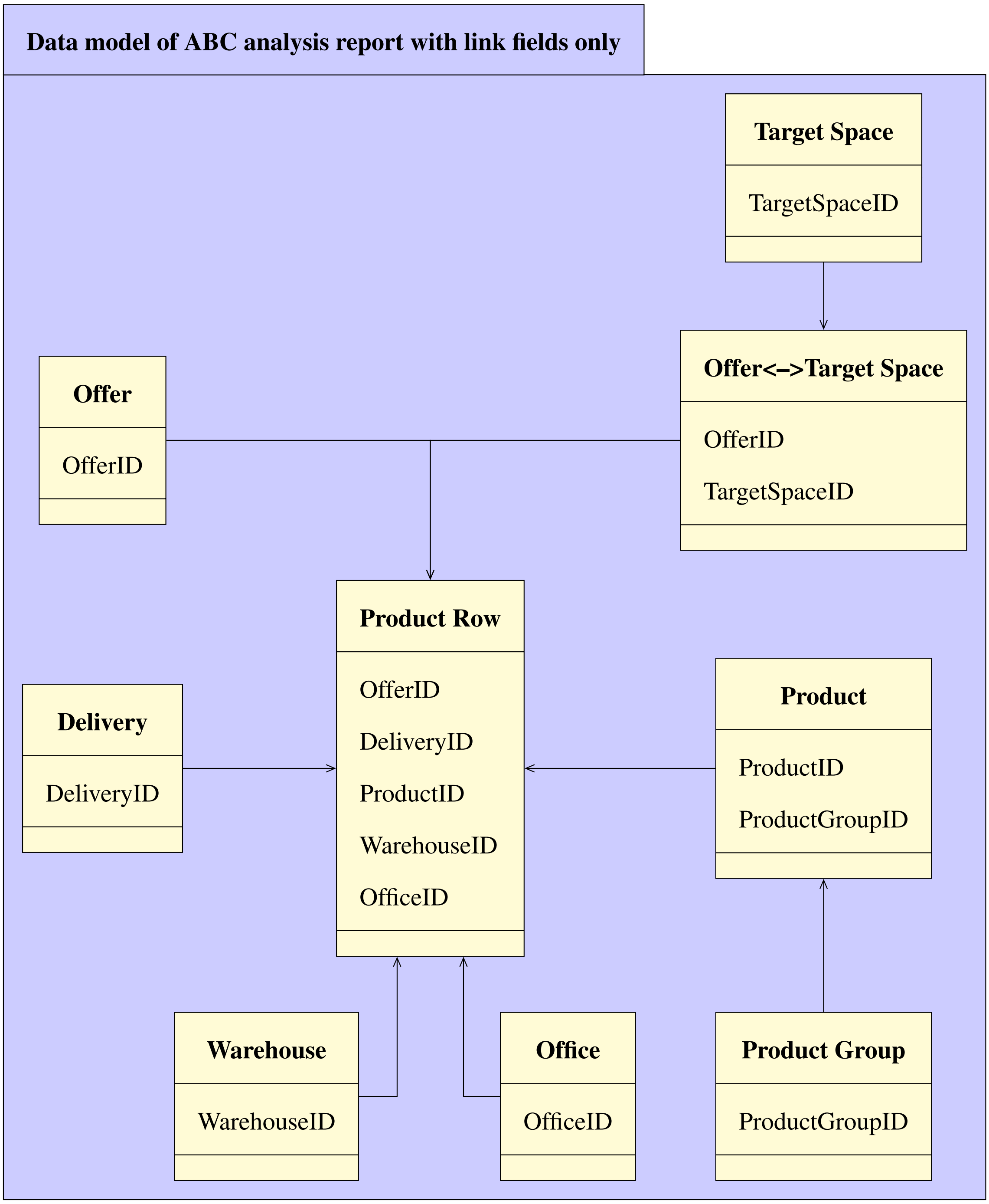


Figure 5.2: Data model of ABC analysis report with link fields only. Translated from Finnish original to English (see Table J.1 for a partial translation). The used data model is a snowflake schema [68].

5.1.4 Used data model

The ABC analysis report's data model (see Figure 5.2) in Qlik Sense® was chosen to be a snowflake schema [68], with the central fact table being the (sold and delivered) product row, because Qlik® recommends the use of a star schema or a snowflake schema for efficient handling data in Qlik Sense® [69].

The central fact table has as many rows as there are delivered product rows in sales deliveries (see sales_delivery_product table in Figure 5.4). Only sales made using the Ideal product catalog, which was released in February 9, 2016 (see Section 3.1.1), are taken into account. [65]

The main measures in the central fact table are sales price, list price, profit, acquirement price and quantity. All of these price and profit measures are without VAT except that there is also sales price including VAT. All of them have two versions: a unit price and a row price. A row price is a unit price multiplied by quantity. Although the row prices are redundant information they have been included to help ABC class calculation in Qlik Sense®. [65]

The dimension tables [68] around the central fact table provide additional information about the related deliveries, offers, offers' target spaces, offices, product groups, products and warehouses.

Figures 5.3 and 5.4 describe the most important tables and their interrelations in Tarla's database related to sales delivery and product handling. Many of them were used in the mapping of data from Tarla's database to the ABC analysis report's data model (see Appendix J for a detailed list of the used table mappings).

5.1.5 Qlik Sense extensions

Because of the unavailability, limitedness or non-traditional nature of certain user interface elements in the base Qlik Sense® distribution in 2018 the author of this thesis chose to use a few Qlik Sense® extensions: Climber Selection Bar [70], Variable extension [71] and Simple KPI extension [72]. Climber Selection Bar was used to give a traditional looking user-friendly date range selection functionality, which was otherwise missing. Variable extension was used e.g. for the traditional looking drop down boxes for choosing variable values and the vertical sliders on the right side of the graph. Simple KPI extension was used e.g. for the large ABC class values above the graph.

5.1.6 ABC class calculation

Calculating the ABC classes dynamically in Qlik Sense® was a challenge for the author of this thesis. Initially he was not even sure if this could be done at all because of the limitations of Qlik Sense® in 2018. After experimentation and a lot of documentation reading however a working solution using the StructuredParameter parameters of the Aggr chart function [16] was found (see Listing 5.1).

Listing 5.1: Choosing a color based on dynamic calculation of an ABC class.

```
Aggr(If(Rangesum(Above(Sum($(vMeasure1)) / Sum(TOTAL $(vMeasure1)), 0, RowNo())) <
        $(vLimitA)/100, $(vColorA),
     If(Rangesum(Above(Sum($(vMeasure1)) / Sum(TOTAL $(vMeasure1)), 0, RowNo())) <
        $(vLimitB)/100, $(vColorB), $(vColorC))),
    ($(vDim1), (=Sum($(vMeasure1)), Desc)))
```

In Listing 5.1 the cumulative sum of the measure defined by the variable vMeasure1 is dynamically calculated in the dimension defined by the variable vDim1 in a descending order of magnitude of measure. The percentage portion p of this cumulative sum from the total sum of the same measure in the same dimension is then calculated. Then it is checked if $0 \leq p < \frac{vLimitA}{100}$, if it is then the color defined by the variable vColorA is chosen. If it is not then it is checked if $\frac{vLimitA}{100} \leq p < \frac{vLimitB}{100}$, if it is then the color defined by the variable vColorB is chosen. If it is not then the color defined by the variable vColorC is chosen. The ABC classes' percentage limits are defined by the vLimitA and vLimitB variables and their order is $0 \leq vLimitA \leq vLimitB \leq 100$.

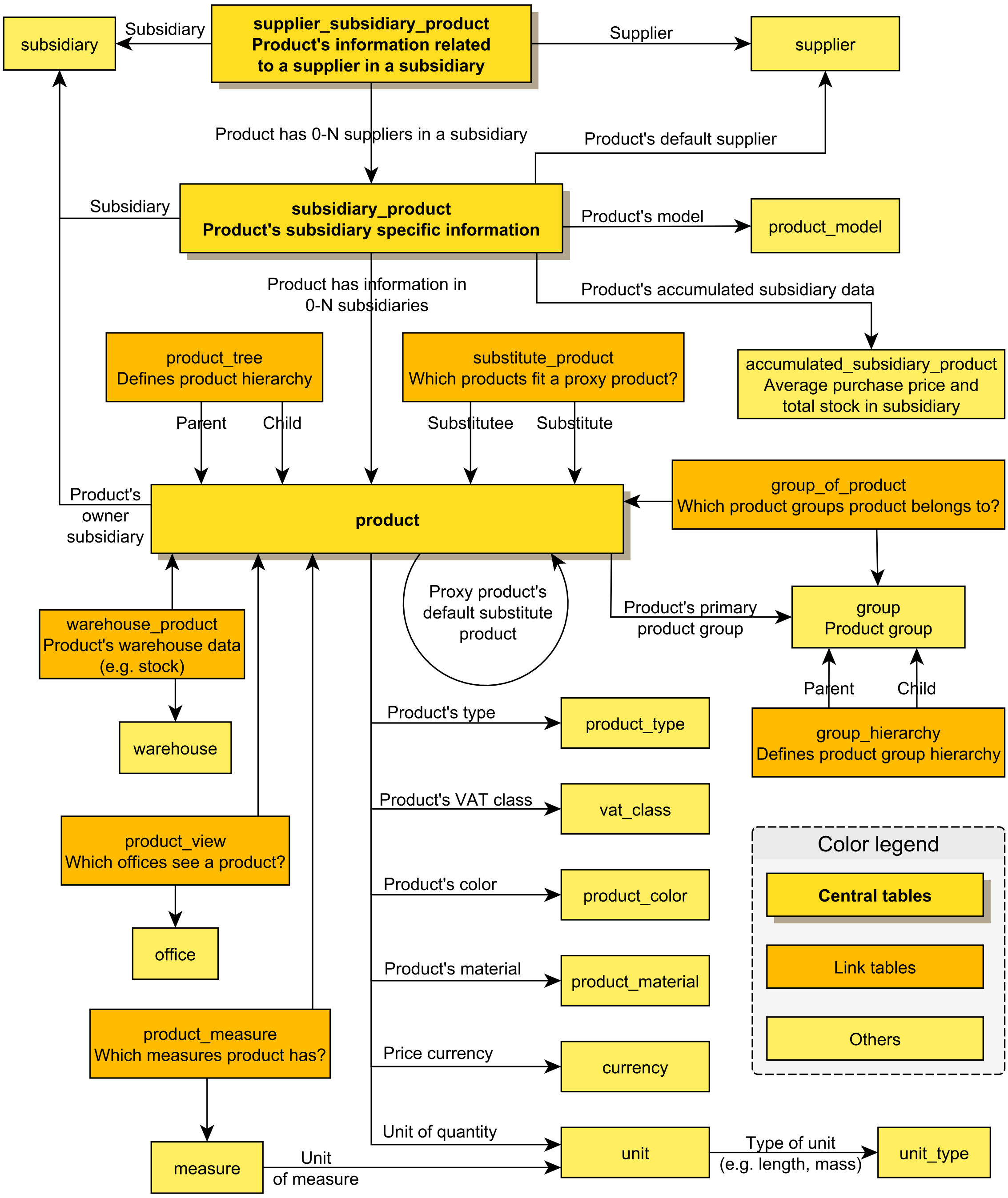


Figure 5.3: Most important product related database tables in Tarla's database [7]

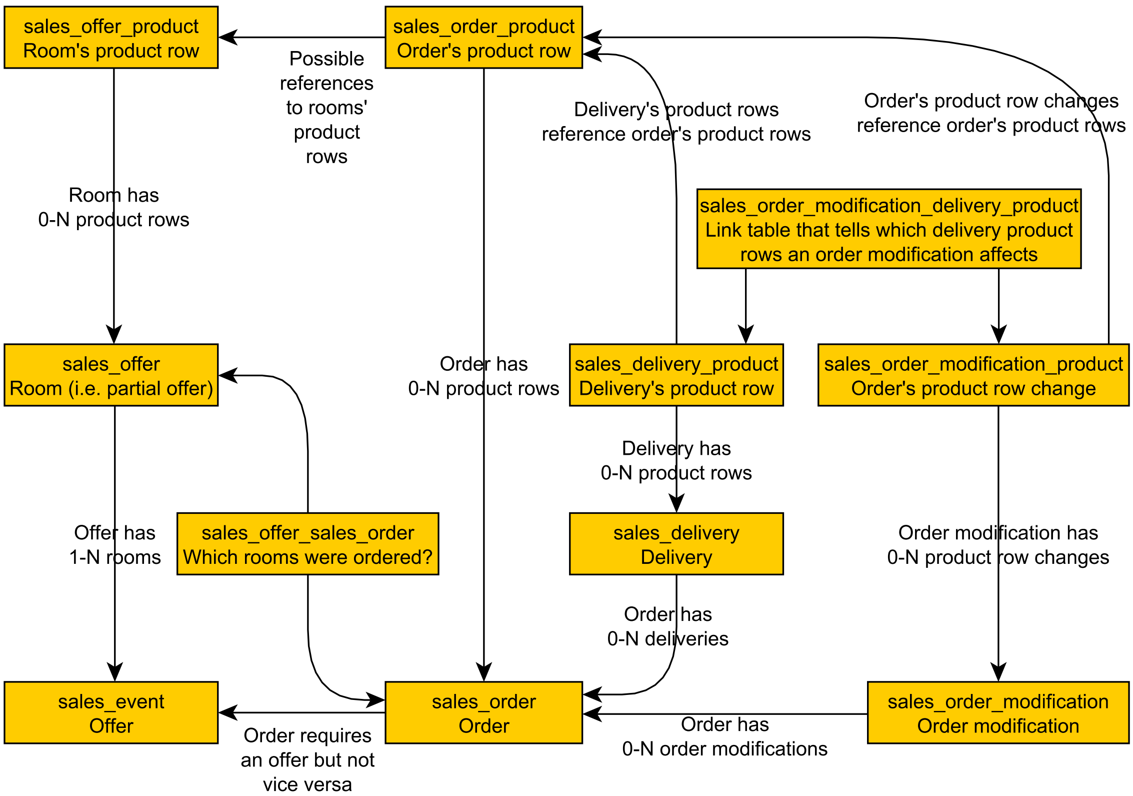


Figure 5.4: Most important sales order related database tables in Tarla's database [7]

5.2 Sales report

Before implementing the sales report it was confirmed [73] that including only top level products and thus excluding the component parts of hierarchically structured products was a desired and sufficient level of information for the sales report.

Unfortunately because the author of this thesis was laid off as an employee of Ideal Keittiöt Oy in February 2019 on financial and production-related grounds, the sales report was never implemented. During February–March 2019 there were negotations with the CIO and the Executive Vice President about a possibility of implementing a third alternative report not described here. The report was to be made using Jeeves as its data source. Unfortunately because of maintenance related concerns the negotations ended near mid-March without reaching agreement. Therefore, the ABC analysis report (see Section 5.1) is the only implemented new report related to this thesis.

Chapter 6

Evaluation of the new reports

Initially the author of this thesis planned to interview the end-users (see Section 4.1) of the new reports. When the time to execute this evaluation came in late spring 2020 many of the variables related to this plan had changed. First the author of this thesis was no longer an employee of Ideal Keittiöt Oy and thus not privy to all of the internal information—e.g. employee job positions, changes or contact information—in the firm. Second the deadline for the completion of this thesis was less than a month ahead from the start of this evaluation. Third the contact person of the author of this thesis in Ideal Keittiöt Oy was generally unavailable during spring 2020 because of heavy preoccupation with work.

Taking into account all of the aforementioned changes in the circumstances the author of this thesis chose to drastically simplify the evaluation plan and only evaluate the new reports by himself and by Ideal Keittiöt Oy's CIO. The justification for this action was that the initial plan was not anymore practical given the new circumstances.

Only one new report was finally implemented (see Section 5.1) and it is the report that is hereinafter referenced in this chapter.

6.1 Evaluation by Ideal Keittiöt Oy's CIO

Ideal Keittiöt Oy's CIO was interviewed [43] by phone for approximately 1h 15min on May 7, 2020 in relation to the new reports. Notes were taken during the interview and the relevant parts are summarized here:

The report looks quite good—e.g. the color use is distinct—and it is moderately easy to use but is not really useful anymore because it uses outdated data from Tarla (see Section 3.1.1) instead of current data from Jeeves (see Section 3.2.1). The business model of Ideal Keittiöt Oy has changed since 2018 especially related to warehousing which makes the use of the report also less useful in 2020 than in 2018. To make the report useful it would have to be changed to use data from Jeeves and add a drill down functionality to product level. The report could probably be used directly by some management personnel (at least by the CIO and the CFO) but would require training to be used in local warehouses. [43]

With the required changes and training the report could be used for inspecting what has been sold a lot and what has not been sold almost at all although the CIO does not see it probable that the report would be taken into continuous use or used for planning of action. [43]

6.2 Self-evaluation

In this section the author of this thesis evaluates the ABC analysis report (see Section 5.1) and acknowledges that he may be at least partially biased in his evaluation:

The aesthetics and the usability of the report are good. The user interface has been laid out in a logical and consistent manner: the most important user interface elements that need interaction are near the upper left corner, the most important user interface elements that show information are near the center and elements that are interrelated are near each other. The report also follows conventions related to Qlik Sense reports already in place in Ideal Keittiöt Oy by placing the filters either horizontally at the top of the report or vertically at the left side of the report which makes it easier to use than without following these conventions.

The use of color coding for identifying the ABC classes is effective and makes their perception very easy. The graph in the center of the report is a moderately complex with two different Y axes which makes the information in it somewhat hard to grasp initially. The conservation of space for imparting information justifies the complexity of the graph.

Use of Tarla's data is unfortunate but understandable taking into account the time when the report was initially conceived and its implementation started. By using Tarla's data the report is practically useless in 2020 as Tarla is not used anymore [43] in Ideal Keittiöt Oy. Drill down functionality would make the report more useful by being able to see of what products each ABC class consists of—this was initially planned but was not finally implemented. Even without the drill down functionality the report has utility for inspecting e.g. ABC classes of door models.

To summarize the report has a good and intuitive user interface, is relatively easy to use but is not practically useful for Ideal Keittiöt Oy in 2020 anymore without changes.

Chapter 7

Conclusions

The original aim of this thesis (see Chapter 1) of developing three new reports for Ideal Keittiöt Oy first changed to two reports (see Chapter 4) when interviewing the intended end-users about their reporting needs and finally to one (see Chapter 5) at the implementation stage. The implemented report (see Section 5.1) was partially successful as it was evaluated (see Chapter 6) that it looks quite good, is moderately easy to use and could be somewhat useful but in 2020 is not because it was made using the old ERP system's data (see Section 3.1.1) and it did not include a drill down functionality to the product level.

The original ancillary aim of this thesis (see Chapter 1) of modelling of Jeeves's database was partially successful (see Section 3.2.1) and resulted in an internal guide to Ideal Keittiöt Oy for Jeeves reporting.

As the author of this thesis initially had hoped it would, the theoretical background (see Chapter 2) helped him to create easily perceivable, graspable visualizations of data both in the new reports (see Chapter 5) and in this thesis itself (see e.g. Figures 3.8, 3.12 and 5.1). He found the research about the preattentive effects (see Section 2.4) and the Gestalt laws (see Section 2.6) both informative and helpful when designing the user interface for the ABC analysis report and the figures in this thesis.

For new developers creating Qlik Sense® reports the author of this thesis has a few recommendations: In order to alleviate the maintenance concerns related to using nonofficial extensions use the officially supported [74] extensions in the Dashboard controls bundle [75] and the Visualization charts bundle [76]. For the same reason check out the extensions using Qlik's Trusted Extension Developer (TED) Program [77]. For extending Qlik Sense's functionality even further check out the open source libraries [78], e.g. picasso.js [79], of Qlik. Also make sure to verify expectations whether a Qlik Sense component can do something or not as they are—or at least were in 2018—inconsistent, e.g. one component supports changing a name using a variable but another component does not. Finally make sure to use all the available peer support whether it be commercial consulting services, Qlik's public Slack channel or your coworkers—two heads are better than one.

References

[1] Suomen Asiakastieto Oy. Financial Information of Ideal Keittiöt Oy. URL: https://www.asiakastieto.fi/yritykset/fi/ideal-keittiot-oy/08935302/taloustiedot, March 2018. Accessed on March 18, 2018. Latest staff headcount was 102 in December 2016. Latest turnover was € 15,481,002 in December 2016.

[2] European Commission. Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. *Official Journal of the European Union*, L 124:36–41, May 2003. URL: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32003H0361&from=EN. Accessed on March 18, 2018.

[3] Business Information System (BIS). Ideal Keittiöt Oy (Business ID: 0893530-2). URL: https://tietopalvelu.ytj.fi/yritystiedot.aspx?yavain=731388&tarkiste=63E9CBA26109E9804C45CD803018E81FF7A2EA55, March 2018. Accessed on March 18, 2018.

[4] Ideal Keittiöt Oy. Ideal Keittiöt Oy's LinkedIn profile. URL: https://www.linkedin.com/company/keitti-j-tti-oy, March 2018. Accessed on March 18, 2018.

[5] Centre of Registers and Information Systems: e-Business Register. Osaühing TNC-Components (Registry code: 10713936). URL: `https://ariregister.rik.ee/ettevotja?id=9000081087&lang=eng`, March 2018. Accessed on March 18, 2018.

[6] TNC-Components Oü. Company page of TNC-Components Oü's website. URL: `http://www.tnc.ee/en/company`, March 2018. Accessed on March 18, 2018.

[7] Ideal Keittiöt Oy. Tarla's production MySQL database. URL: `mysql://reporting.jatti.local/tarla_replicated`, December 2017. Private information of Ideal Keittiöt Oy. Accessed on December 3, 2017.

[8] Pekka Viikari (Chief Information Officer of Ideal Keittiöt Oy). Announcement in Ideal Keittiöt Oy's Intranet About Jeeves ERP Project Starting. URL: `https://keittiojatti.sharepoint.com/sites/Ideal-intra/ajankohtaista/ajankohtaista/Lists/Viestit/Post.aspx?ID=244`, February 2017. Private information of Ideal Keittiöt Oy. Accessed on February 15, 2018. Announcement was made on February 15, 2017.

[9] Jeeves Information Systems AB. Jeeves ERP. URL: `http://www.jeeveserp.com/`, December 2017. Accessed on December 2, 2017.

[10] Hitachi Vantara. Pentaho website. URL: `http://www.pentaho.com/`, December 2017. Accessed on December 3, 2017.

[11] Ideal Keittiöt Oy. The Git repository of Ideal Keittiöt Oy's custom version of Pentaho BI Server Community Edition 3.7.0. URL: `git+ssh://trac.jatti.local/var/repos/tarlasystem/unumpentaho`, December 2017. Private information of Ideal Keittiöt Oy. Accessed on December 3, 2017.

[12] Qlik. Qlik Sense website. URL: https://www.qlik.com/us/products/qlik-sense, May 2020. Accessed on May 9, 2020.

[13] Ideal Keittiöt Oy. Tarla's SVN source code repository. URL: svn+ssh://trac.jatti.local/var/repos/tarlasystem/source, January 2018. Private information of Ideal Keittiöt Oy. Accessed on January 1, 2018. Last commit revision 19507 on December 28, 2017.

[14] Will Gorman. *Pentaho Reporting 3.5 for Java Developers*. Packt Publishing Ltd, 2009.

[15] Christopher Ilacqua, Henric Cronström, and James Richardson. *Learning Qlik Sense®: The Official Guide*. Packt Publishing Ltd, 2015.

[16] Qlik. Qlik Sense's Online Manual. URL: https://help.qlik.com/en-US/sense, May 2020. Accessed on May 9, 2020.

[17] Jeeves Information Systems AB and Ideal Keittiöt Oy. Jeeves ERP's Microsoft SQL Server database. URL: jdbc:sqlserver://jeeves-sql.jatti.local;databaseName=Ideal_keittiot, May 2020. Private information of Jeeves Information Systems AB and Ideal Keittiöt Oy. Accessed using Microsoft SQL Server 2014 Management Studio. Accessed on May 9, 2020.

[18] Jeeves Information Systems AB. Jeeves Application services (JAS) 14.0's English Reference Guide for Jeeves Programs and Database. Compiled HTML Help file DocAreas_sw.chm from the Jeeves ERP's application directory /Jeeves/lang/sw/, October 2016.

[19] Jean-Luc Hainaut. Introduction to database reverse engineering. *LIBD Lecture Notes*, 2002.

[20] Pengon Oy. Pengon Oy's website. URL: https://www.pengon.fi/, January 2018. Accessed on January 16, 2018. Pengon Oy's Finnish Business ID: 2170265-0.

[21] Staria Oyj. Staria Oyj's website. URL: `http://www.staria.fi/`, January 2018. Accessed on January 16, 2018. Staria Oyj's Finnish Business ID: 0696081-1.

[22] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 3rd edition, 2013.

[23] William S. Cleveland. *The Elements of Graphing Data*. Hobart Press, Summit, New Jersey, United States of America, revised edition, 1994.

[24] Jeremy M Wolfe and W Gray. Guided search 4.0. *Integrated models of cognitive systems*, pages 99–119, 2007.

[25] Tex Texin and W3C®. Script direction and languages. URL: `https://www.w3.org/International/questions/qa-scripts.en`, September 2003. Accessed on May 11, 2020.

[26] Farough Abed. Cultural influences on visual scanning patterns. *Journal of Cross-Cultural Psychology*, 22(4):525–534, 1991.

[27] Hannah Faye Chua, Julie E Boland, and Richard E Nisbett. Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences*, 102(35):12629–12633, 2005.

[28] Jan Theeuwes and Frank L Kooi. Parallel search for a conjunction of contrast polarity and shape. *Vision Research*, 34(22):3013–3016, 1994.

[29] Anne Treisman and Stephen Gormican. Feature analysis in early vision: evidence from search asymmetries. *Psychological review*, 95(1):15–48, 1988.

[30] Michael D'Zmura, Peter Lennie, and Carlo Tiana. Color search and visual field segregation. *Perception & psychophysics*, 59(3):381–388, 1997.

[31] Christopher G Healey. Choosing effective colours for data visualization. In *Proceedings of Seventh Annual IEEE Visualization'96*, pages 263–270. IEEE, 1996.

[32] Willis D Ellis. *A source book of Gestalt psychology*. Kegan Paul, Trench, Trubner & Company, 1938.

[33] William S. Cleveland. *Visualizing Data*. Hobart Press, Summit, New Jersey, United States of America, 1993.

[34] Jeffrey B. Arnold. *ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'*, 2019. R package version 4.2.0.

[35] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. R version 4.0.0.

[36] Ideal Keittiöt Oy. Ideal Keittiöt Oy's website. URL: `https://ideal-keittiot.fi/`, December 2017. Accessed on December 2, 2017. Ideal Keittiöt Oy's Finnish Business ID: 0893530-2.

[37] Eppu Tuominen. Private interview on December 13, 2017. Private spoken interview face-to-face, December 2017. Eppu Tuominen was involved in the start of the Tarla project.

[38] Ideal Keittiöt Oy. Tarla's SVN document repository. URL: `svn+ssh://trac.jatti.local/var/repos/tarlasystem/doc`, December 2017. Private information of Ideal Keittiöt Oy. Accessed on December 3, 2017.

[39] Ideal Keittiöt Oy. Tarla's release notes. URL: `svn+ssh://trac.jatti.local/var/repos/tarlasystem/source/trunk/ReleaseNotes_fi.html`, February 2018. Private information of Ideal Keittiöt Oy. Accessed on February 14, 2018.

[40] Ideal Keittiöt Oy. KjTrac (Tarla project's project management and bug tracking system). URL: `https://trac.jatti.local/kj`, February 2018. Private information of Ideal Keittiöt Oy. Accessed on February 14, 2018.

[41] Lakkapää Oy. Lakkapää Oy's website. URL: https://www.lakkapaa.com/, February 2018. Accessed on February 11, 2018.

[42] Jarmo Bäcklund. Qt 4.7.4 release announcement. URL: http://blog.qt.io/blog/2011/09/01/qt-4-7-4-released/, September 2011. Accessed on February 11, 2018. Release Announcement was made on September 1, 2011.

[43] Pekka Viikari (Chief Information Officer of Ideal Keittiöt Oy). Private interview on May 7, 2020. Private spoken interview by phone, May 2020.

[44] TIBCO Software Inc. Jaspersoft community edition. URL: https://community.jaspersoft.com/, February 2018. Accessed on February 14, 2018.

[45] OpenReports Development Team. OpenReports. URL: http://oreports.com/, February 2018. Accessed on February 14, 2018.

[46] Eclipse Foundation. Eclipse BIRT Project. URL: http://www.eclipse.org/birt/, February 2018. Accessed on February 14, 2018.

[47] OW2 Consortium. SpagoBI. URL: https://www.spagobi.org/, February 2018. Accessed on February 14, 2018.

[48] RapidMiner. RapidMiner. URL: https://rapidminer.com/, February 2018. Accessed on February 14, 2018.

[49] Nokia Corporation and/or its subsidiaries. Qt 4.7 Reference Documentation of QtWebKit Module. URL: https://doc.qt.io/archives/qt-4.7/qtwebkit.html, February 2018. Accessed on February 15, 2018.

[50] Ideal Keittiöt Oy's custom version of Pentaho BI Server Community Edition 3.7.0. Pentaho audit log. URL: ssh://reporting.jatti.local/opt/pentaho/biserver-ce/pentaho-solutions/system/logs/

audit/PentahoAuditLog.log, February 2018. Private information of Ideal Keittiöt Oy. Accessed on February 18, 2018. Log was found split into two files. Combining them resulted in data from July 12, 2012 to February 18, 2018.

[51] Ideal Keittiöt Oy. Tarla's Pentaho reports. URL: svn+ssh://trac.jatti.local/var/repos/tarlasystem/source/trunk/tools/reporting/report_templates, February 2018. Private information of Ideal Keittiöt Oy. Accessed on February 17, 2018.

[52] Jeeves Information Systems AB. Genius Architecture. URL: https://www.jeeveserp.com/en/genius-architecture, May 2020. Accessed on May 9, 2020.

[53] Jeeves Information Systems AB. Additive Solutions and Integrations. URL: https://www.jeeveserp.com/en/additive-solutions-and-integrations, May 2020. Accessed on May 9, 2020.

[54] Jeeves Information Systems AB. Jeeves Business Intelligence (BI) Powered by QlikView. URL: https://www.jeeveserp.com/en/jeeves-business-intelligence-bi-powered-qlikview, May 2020. Accessed on May 9, 2020.

[55] SAP SE. SAP Crystal Reports. URL: https://www.crystalreports.com/reports/, May 2020. Accessed on May 9, 2020.

[56] DB-Manager Oy. 2020 Fusion product's website. URL: https://www.dbmanager.fi/en/ratkaisut/fusion/, May 2020. Accessed on May 9, 2020.

[57] DB-Manager Oy. Fusion Sales product's website. URL: https://www.dbmanager.fi/en/ratkaisut/fusion-sales/, May 2020. Accessed on May 9, 2020.

[58] Veli-Pekka Vihonen (BI Consultant at Pengon Oy). E-mail invitation to Qlik Sense application workshop. E-mail with subject "Qlik sovellus workshop 15.2/9.30", February 2017. Private e-mail conversation. E-mail received on February 10, 2017. Invitation for date February 15, 2017.

[59] Qlik. Qlik® Sense Token Licensing. URL: `https://www.qlik.com/us/resource-library/qlik-sense-token-licensing`, November 2014. Accessed on March 1, 2018.

[60] Pekka Viikari (Chief Information Officer of Ideal Keittiöt Oy). Conversation at Ideal Keittiöt Oy, Lemminkäisenkatu 36, 20520 Turku, Finland, January–March 2018.

[61] Ideal Keittiöt Oy. Ideal Keittiöt Oy's internal contact list (February 16, 2018). URL: `https://keittiojatti.sharepoint.com/sites/Ideal-intra/organisaatio/Documents/Ideal%20Sis%C3%A4inen%20yhteystietoluettelo%2016.2.18.pdf`, February 2018. Private information of Ideal Keittiöt Oy. Accessed on March 1, 2018.

[62] Joseph M Juran. Universals in management planning and controlling. *Management Review*, 43(11):748–761, 1954.

[63] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer, 7 edition, 2013.

[64] H Ford Dickie. ABC inventory analysis shoots for dollars not pennies. *Factory Management and Maintenance*, 109(7):92–94, 1951.

[65] Kari Salminen. Toimitusten ABC-analyysi (Testiversio) Qlik Sense® report. URL: `https://qlikview.jatti.local/sense/app/231cc5f7-cbd4-4ea9-9d61-cae7913555e0`, April 2019. Private information of Ideal Keittiöt Oy. Published to Development stream on April 9, 2019. Accessed on May 12, 2020.

[66] Thomas E.. Vollmann, William L.. Berry, and D Clay Whybark. *Manufacturing planning and control systems*. Irwin/McGraw-Hill, 1997.

[67] Amy Alberts. Eye-tracking study: 5 key learnings for data designers everywhere. URL: `https://www.tableau.com/about/blog/2017/6/eye-tracking-study-5-key-learnings-data-designers-everywhere-72395`, June 2017. Accessed on April 29, 2020.

[68] Surajit Chaudhuri and Umeshwar Dayal. An overview of data warehousing and olap technology. *ACM Sigmod record*, 26(1):65–74, 1997.

[69] Qlik. Understanding script syntax and data structures. URL: `https://help.qlik.com/en-US/sense/Subsystems/Hub/Content/Sense_Hub/LoadData/understand-data-structures.htm`, April 2020. Accessed on May 11, 2020.

[70] Karl Fredberg Sjöstrand and David Alcobero. Climber Selection Bar extension for Qlik Sense®. URL: `https://github.com/ClimberAB/ClimberSelectionBar`, May 2020. Accessed on May 11, 2020.

[71] Erik Wetterberg. Variable extension for Qlik Sense®. URL: `https://github.com/erikwett/qsVariable`, May 2020. Accessed on May 11, 2020.

[72] Alexander Nerush and Håkan Axelsson. Simple KPI component for Qlik Sense®. URL: `https://github.com/alner/qsSimpleKPI`, May 2020. Accessed on May 11, 2020.

[73] Two employees of Ideal Keittiöt Oy (A sales area manager and a sales person). Private interview on January 10, 2019. Private spoken interview face-to-face. Names have been anonymized for privacy, January 2019.

[74] Qlik. How are qlik visualization and dashboard bundle extensions supported? URL: https://support.qlik.com/articles/000076030, June 2019. Accessed on May 13, 2020.

[75] Qlik. Dashboard bundle. URL: https://help.qlik.com/en-US/sense/February2020/Subsystems/Hub/Content/Sense_Hub/Visualizations/DashboardBundle/dashboard-bundle.htm, February 2020. Accessed on May 13, 2020.

[76] Qlik. Visualization bundle. URL: https://help.qlik.com/en-US/sense/February2020/Subsystems/Hub/Content/Sense_Hub/Visualizations/VisualizationBundle/visualization-bundle.htm, February 2020. Accessed on May 13, 2020.

[77] Qlik. Trusted extension developer (ted) program. URL: https://developer.qlik.com/ted, May 2020. Accessed on May 13, 2020.

[78] Qlik. Qlik - open source software. URL: https://github.com/qlik-oss, May 2020. Accessed on May 14, 2020.

[79] Qlik. picasso.js. URL: https://picassojs.com/, May 2020. Accessed on May 13, 2020.

[80] Ben Collins-Sussman, Brian Fitzpatrick, and Michael Pilato. Version control with subversion. URL: http://svnbook.red-bean.com/, February 2018. Accessed on February 10, 2018.

[81] Al Danial. GitHub source code repository of cloc program (Count Lines of Code). URL: https://github.com/AlDanial/cloc, February 2018. Accessed on February 10, 2018.

Appendix A

Developers' first and last commits calculation

SVN log was read with **svn log** command from Tarla's source code repository [13] and the output was redirected to svn.log file. Then the file was parsed with this Python script:

Listing A.1: Parsing developers' first and last commits from SVN log file for pgfgantt.

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Tested with Python 2.7, may not be 3.x compatible
from __future__ import print_function
import re, datetime as dt

def main(argv = None):
    data = dict()
    log = open('svn.log', 'rt')

    # Only read data in year range 2008-2017
    min_year, max_year = 2008, 2017

    # Read the SVN log file and parse it
    for line in log.readlines():
```

```
        # Parse revision number, user and date
        pattern = r'^r(\d+)␣\|␣([^|]+)␣\|␣(.{7})'
        match = re.match(pattern, line)
        if match:
            revision, user, date_text = match.groups()
            # Parse date at year and month granularity
            date = dt.datetime.strptime(date_text, '%Y-%m')
            if min_year <= date.year <= max_year:
                if user not in data:
                    data[user] = []
                data[user] += [(int(revision), date)]

    # Remove non-personal users (Total three commits)
    for user in ['builder', 'nerp']:
        del data[user]

    # Combine same users with multiple usernames
    data['kari'] += data['kari.salminen']
    del data['kari.salminen']

    # Sort users by their first revision dates
    starts = sorted([(
        min(data[user])[1], # First revision date
        max(data[user])[1], # Last revision date
        user)               # User name
        for user in data.keys()])
    # Anonymize users
    anon_users = []
    for pos, (mindate, maxdate, user) in enumerate(starts):
        num = pos+1
        if user == 'kari':
            anon_user = 'Kari␣Salminen'
        else:
```

```
            anon_user = ('Developer␣%d' % (num,))
        anon_users += [(num, anon_user, mindate, maxdate)]
    anon_users = sorted(anon_users)

    # Format the users' first and last revision dates
    # (Month granularity) as Gantt chart bars for use
    # by the pgfgantt package in LaTeX
    def date_to_month_num(date):
        return (date.year - min_year) * 12 + date.month
    max_pos = max(anon_users)[0]
    for (pos, anon_user, mindate, maxdate) in anon_users:
        bar_start = date_to_month_num(mindate)
        bar_end = date_to_month_num(maxdate)
        print((r'\ganttbar{%s}{%d}{%d}␣' +
            r'%s%%␣%4d-%02d␣-␣%4d-%02d') % (
            anon_user, bar_start, bar_end,
            r'\\␣' if pos != max_pos else '',
            mindate.year, mindate.month,
            maxdate.year, maxdate.month))
    # Exit the program
    sys.exit(0)

# The main function
if __name__ == '__main__':
    import sys
    main(sys.argv)
```

Listing A.2: Result of running listing A.1.

```
\ganttbar{Developer 1}{2}{7} \\ % 2008-02 - 2008-07
\ganttbar{Developer 2}{2}{24} \\ % 2008-02 - 2009-12
\ganttbar{Developer 3}{2}{25} \\ % 2008-02 - 2010-01
\ganttbar{Developer 4}{2}{44} \\ % 2008-02 - 2011-08
\ganttbar{Developer 5}{2}{62} \\ % 2008-02 - 2013-02
```

```
\ganttbar{Developer 6}{2}{76} \\ % 2008-02 - 2014-04
\ganttbar{Developer 7}{5}{35} \\ % 2008-05 - 2010-11
\ganttbar{Developer 8}{7}{35} \\ % 2008-07 - 2010-11
\ganttbar{Developer 9}{8}{21} \\ % 2008-08 - 2009-09
\ganttbar{Kari Salminen}{10}{120} \\ % 2008-10 - 2017-12
\ganttbar{Developer 11}{40}{44} % 2011-04 - 2011-08
```

Appendix B

Commits and developers per year calculation

SVN log was read with **svn log** command from Tarla's source code repository [13] and the output was redirected to svn.log file. Then the file was parsed with this Python script:

Listing B.1: Calculating commit count and developer count per year from SVN log file.

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Tested with Python 2.7, may not be 3.x compatible
from __future__ import print_function
import re, datetime as dt

def main(argv = None):
    data = dict()
    log = open('svn.log', 'rt')

    # Only read data in year range 2008-2017
    min_year, max_year = 2008, 2017

    # Read the SVN log file and parse it
    for line in log.readlines():
```

```
        # Parse revision number, user and year
        pattern = r'^r(\d+)␣\|␣([^|]+)␣\|␣(\d{4})'
        match = re.match(pattern, line)
        if match:
            revision, user, year = match.groups()
            year = int(year)
            if min_year <= year <= max_year:
                if user not in data:
                    data[user] = []
                data[user] += [(int(revision), year)]

# Remove non-personal users (Total three commits)
for user in ['builder', 'nerp']:
    del data[user]

# Combine same users with multiple usernames
data['kari'] += data['kari.salminen']
del data['kari.salminen']

# Calculate yearly user and commit counts
yearly_users = dict()
yearly_commits = dict()
for user in data.keys():
    for (revision, year) in data[user]:
        if year not in yearly_users:
            yearly_users[year] = set()
        yearly_users[year] |= set([user])
        if year not in yearly_commits:
            yearly_commits[year] = set()
        yearly_commits[year] |= set([revision])

# Print years and their commit and user counts
print('Year,␣Commits,␣Developers')
```

```
    for year in xrange(min_year, max_year + 1):
        print('%d,␣%d,␣%d' % (year,
            len(yearly_commits[year]),
            len(yearly_users[year])))

    # Exit the program
    sys.exit(0)

# The main function
if __name__ == '__main__':
    import sys
    main(sys.argv)
```

Table B.1: Result of running listing B.1 shown as a table.

Year	Commits	Developers
2008	4390	10
2009	5358	9
2010	3711	7
2011	1053	5
2012	912	3
2013	1303	3
2014	914	2
2015	620	1
2016	833	1
2017	410	1

Appendix C

First sale per office calculation

Listing C.1: First sale per office SQL query.

```
SELECT
MAX(s.abbreviation) AS "Subsidiary",
TRIM(TRAILING '␣(Lopetettu)'
     FROM o.short_name) AS "Office",
DATE(FROM_UNIXTIME(MIN(se.created_timestamp) / 1000))
    AS "First␣Sale"
FROM sales_event AS se
INNER JOIN office AS o
ON se.office_id = o.id
INNER JOIN subsidiary AS s
ON o.subsidiary_id = s.id
GROUP BY o.group_num
ORDER BY "First␣Sale";
```

Table C.1: Result of listing C.1's query (Years 2009–2012).

Subsidiary	**Office**	**First Sale**
KJ	Turku	2009-07-10
KJ	Espoo	2009-07-15
KJ	Salo	2009-11-24
KJ	Tampere	2009-12-01
KJ	Kemi	2009-12-07
KJ	Tornio	2009-12-10
KJ	Kuopio	2010-01-04
KJ	Jyväskylä	2010-01-07
KJ	Oulu Alppila	2010-01-11
KJ	Pietarsaari	2010-01-14
KJ	Huittinen	2010-01-14
KJ	Nummela	2010-03-16
KJ	Pääkonttori	2010-04-07
KJ	Stella Keuruu	2010-06-30
KJ	Keskusvarasto	2010-11-05
KJ	Rautanet Forssa	2012-01-31
KK	Pori Harmaalinna	2012-04-16
KK	Pori Antinkatu	2012-10-09
KJ	Raisio	2012-11-26

Table C.2: Result of listing C.1's query (Years 2013–2017).

Subsidiary	Office	First Sale
KJ	Rauma	2013-03-26
KN	Töölö	2013-04-24
KN	Muurala	2013-05-02
KN	Varisto	2013-05-29
KN	Porttipuisto	2013-06-06
KJ	Värisilmä Forssa	2013-10-28
KP	Kouvola	2014-03-19
KP	Lappeenranta	2014-03-25
KP	Kotka	2014-03-27
KP	Vantaa	2014-04-01
KN	Suomenoja	2014-06-03
KP	Mikkeli	2014-09-25
KP	Lahti Avant	2014-11-06
KN	Kruununhaka	2015-08-11
IK	Oulu Limingantulli	2016-02-23
IK	Lahti	2016-03-02
IK	Seinäjoki	2016-12-12
IK	Rovaniemi	2017-02-07
IK	Tornio	2017-02-21
IK	Tammisto	2017-04-24
IK	Lakkapää Rovaniemi	2017-09-20
IK	Lakkapää Ylitornio	2017-09-20
IK	Lakkapää Tornio	2017-09-20

Table C.3: Mapping from subsidiary to company information [7].

Subsidiary	**Company Name**	**Auxiliary Company Name**	**EU Vat No.**
IK	Ideal Keittiöt Oy		FI08935302
KJ	Keittiöjätti Oy		FI08935302
KK	Keittiö-Kartano Oy	Kartanon Keittiötukku	FI23458161
KN	Kitchennet Oy	Keittiönet	FI22585548
KP	Kymen Keittiötukku Oy	Vantaan Keittiöpiste	FI10937577

Appendix D

Offices per year calculation

Listing D.1: Offices selling at least €10,000 per year SQL query.

```sql
SELECT
    y AS "Year",
    COUNT(DISTINCT office_name) AS "Offices"
FROM
(
    SELECT
        YEAR(FROM_UNIXTIME(
            p.created_timestamp / 1000)) AS y,
        o.short_name AS office_name
    FROM payment AS p
    INNER JOIN sales_event AS se
    ON p.sales_order_id = se.id
    INNER JOIN office AS o
    ON se.office_id = o.id
    INNER JOIN office_type AS ot
    ON o.office_type_id = ot.id
    WHERE
        -- Include only customer payments
        NOT p.is_interoffice AND
        -- Include non-franchise and franchise offices
```

```
        -- but exclude central warehouse and head quarters
        ot.locked_sourcecode_reference IN
        (
            'locked_ref_shop',
            'locked_ref_franchise_shop'
        )
    GROUP BY y, office_name
    HAVING SUM(p.paid_amount) >= (10 * 1000)
) AS subquery
WHERE y BETWEEN 2009 AND 2017
GROUP BY y
ORDER BY y;
```

Table D.1: Result of listing D.1's query.

Year	**Offices**
2009	5
2010	13
2011	13
2012	10
2013	16
2014	23
2015	24
2016	25
2017	28

Appendix E

Users per year calculation

This SQL query was used for listing all foreign keys pointing to tarla_replicated.user [7] table:

Listing E.1: User foreign keys SQL query.

```
-- Foreign keys in tarla_replicated database to user.id
SELECT TABLE_NAME, COLUMN_NAME
FROM information_schema.KEY_COLUMN_USAGE
WHERE
    CONSTRAINT_SCHEMA = 'tarla_replicated' AND
    TABLE_SCHEMA = 'tarla_replicated' AND
    REFERENCED_TABLE_NAME = 'user' AND
    REFERENCED_COLUMN_NAME = 'id' AND
    -- Omit history, change, serial number and patch tables
    TABLE_NAME NOT LIKE '%_history' AND
    TABLE_NAME NOT LIKE '%_changes' AND
    TABLE_NAME NOT LIKE '%_serial_number' AND
    TABLE_NAME NOT REGEXP '^p[0-9]+_.*$'
ORDER BY TABLE_NAME, COLUMN_NAME;
```

The information from the above query was filtered and used to make this SQL query used to produce the data for Figure 3.6 (Note that the timestamps in tarla_replicated database [7] are Unix timestamps in milliseconds rather than seconds):

Listing E.2: Users per year SQL query.

```
SELECT
    users.y AS "Year",
    COUNT(DISTINCT
        -- Combine two names of single user into one
        REPLACE(ufn.full_name, '␣(Avant)␣', '␣')
    ) AS "Users"
FROM
(
    -- --------------------------------------------------
    -- Creators and creation times from various tables
    -- --------------------------------------------------
    SELECT DISTINCT creator_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "account"
    UNION DISTINCT
    SELECT DISTINCT creator_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "bank_transfer_file_info"
    UNION DISTINCT
    SELECT DISTINCT creator_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "customer"
    UNION DISTINCT
    SELECT DISTINCT creator_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "customer_info"
    UNION DISTINCT
    SELECT DISTINCT creator_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "freight"
```

```
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "group"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "inventory_instance"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "inventory_template"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "payment"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "product"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "public_customer"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "purchase_delivery"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "purchase_order"
UNION DISTINCT
```

```
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "purchase_order_customer_info"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "quick_inventory"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "sales_delivery"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "sales_event"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "sales_offer"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "sales_order"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "sales_order_modification"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "shrinkage"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
```

```
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "subsidiary"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "subsidiary_product"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "supplier"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "supplier_subsidiary_product"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "user"
UNION DISTINCT
SELECT DISTINCT creator_user_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "warehouse_product"
UNION DISTINCT
-- ----------------------------------------------------
-- Person in charge of a warehouse inventory
-- ----------------------------------------------------
SELECT DISTINCT user_in_charge_id AS user_id,
    YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
FROM "inventory_instance"
UNION DISTINCT
-- ----------------------------------------------------
-- Person starting the picking of a purchase delivery
-- ----------------------------------------------------
```

```
(
    SELECT DISTINCT pd.picker_user_id AS user_id,
        YEAR(FROM_UNIXTIME(
            pdsh.created_timestamp / 1000)) AS y
    FROM purchase_delivery AS pd
    INNER JOIN purchase_delivery_state_history AS pdsh
    ON pd.id = pdsh.purchase_delivery_id
    INNER JOIN purchase_delivery_state_type AS pdst
    ON pdsh.purchase_delivery_state_type_id = pdst.id
    WHERE pdst.locked_sourcecode_reference =
        'locked_ref_being_picked'
)
UNION DISTINCT
-- ----------------------------------------------------
-- Person ordering a purchase order
-- ----------------------------------------------------
SELECT DISTINCT orderer_user_id AS user_id,
    YEAR(FROM_UNIXTIME(
        ordering_day_timestamp / 1000)) AS y
FROM "purchase_order"
-- ----------------------------------------------------
-- Person starting the picking of a sales delivery
-- ----------------------------------------------------
UNION DISTINCT
(
    SELECT DISTINCT sd.picker_user_id AS user_id,
        YEAR(FROM_UNIXTIME(
            sdsh.created_timestamp / 1000)) AS y
    FROM sales_delivery AS sd
    INNER JOIN sales_delivery_state_history AS sdsh
    ON sd.id = sdsh.sales_delivery_id
    INNER JOIN sales_delivery_state_type AS sdst
    ON sdsh.sales_delivery_state_type_id = sdst.id
```

```
        WHERE sdst.locked_sourcecode_reference =
            'locked_ref_being_picked'
    )
    UNION DISTINCT
    -- ---------------------------------------------------
    -- Person making a sale
    -- ---------------------------------------------------
    SELECT DISTINCT seller_user_id AS user_id,
        YEAR(FROM_UNIXTIME(created_timestamp / 1000)) AS y
    FROM "sales_event"
) AS users
INNER JOIN user_full_name AS ufn
ON users.user_id = ufn.user_id
WHERE
    -- Limit years to the wanted range 2009-2017
    (users.y BETWEEN 2009 AND 2017)          AND
    -- Omit user names that are not personal
    -- e.g. Asennus means installation and
    --      varasto means warehouse.
    ufn.full_name NOT LIKE '%administrator%' AND
    ufn.full_name NOT LIKE '%asennus%'       AND
    ufn.full_name NOT LIKE '%varasto%'       AND
    ufn.full_name NOT LIKE '%projekti%'      AND
    ufn.full_name NOT LIKE '%malli%ryhmä%'   AND
    ufn.full_name NOT LIKE '%messu%tunnus%'  AND
    ufn.full_name NOT LIKE '%myymälä%'       AND
    ufn.full_name NOT LIKE '%johtaja%'       AND
    ufn.full_name NOT LIKE '%raportointi%'   AND
    ufn.full_name NOT LIKE '%hallitus%'      AND
    ufn.full_name NOT LIKE '%testaaja%'      AND
    ufn.full_name NOT LIKE '%tarkastus%'     AND
    ufn.full_name NOT LIKE '%kauppa%'        AND
    ufn.full_name NOT LIKE '%myynti%'
```

```
GROUP BY users.y
ORDER BY "Year" ASC;
```

Table E.1: Result of listing E.2's query.

Year	**Users**
2009	50
2010	90
2011	84
2012	71
2013	96
2014	124
2015	128
2016	110
2017	111

Appendix F

Lines of code calculation

The last revision of year 2017 (Revision 19507 from December 28th, 2017) of the main development branch (i.e. trunk in SVN [80]) was checked out from Tarla's SVN repository [13]. Then the Linux program cloc [81] was used to calculate the lines of code in the source code excluding third party libraries:

```
cloc . --exclude-dir=thirdparty
```

Appendix G

Pentaho report usage statistics

The Pentaho audit log file [50] of Ideal Keittiöt Oy's Pentaho BI Server 3.7.0 is the source of Pentaho report usage statistics. Its format hasn't been customized by Ideal Keittiöt Oy [11] but is the same as in an uncustomized version of Pentaho BI Server 3.7.0. It was converted to Comma-Separated Values (CSV) file using the following Python script:

Listing G.1: Converting Pentaho audit log to CSV file.

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Tested with Python 2.7, may not be 3.x compatible
from __future__ import print_function
from decimal import Decimal
import codecs


base = 'org.pentaho.reporting.platform.plugin.'
marks = [base + 'ExecuteReportContentHandler',
         base + 'ReportContentGenerator',
         'instance_end']
delim = ','
cols = [u'Ajohetki', # When report was run
        u'Hakemisto',  # Report directory
        u'Tiedosto', # Report filename
```

```
        u'Käyttäjätunnus', # Username
        u'Sekuntikesto'] # Duration in seconds

def format_duration(x):
    return str(Decimal(x).quantize(Decimal('0.0001')))

def main(argv = None):
    sys.stdout = codecs.getwriter('utf8')(sys.stdout)
    if len(argv) < 2:
        return
    filename = argv[1]
    print(delim.join(cols))
    with open(filename, 'rt') as file:
        lines = file.readlines()
        for line in lines:
            if all([mark in line for mark in marks]):
                data = line.split('\t')
                date, report, username, duration = \
                    data[0], data[3], data[5], data[10]
                report_dir, report_file = report.split('/')
                print(delim.join([
                    date, report_dir, report_file,
                    username, format_duration(duration)]))
    exit

if __name__ == '__main__':
    import sys
    main(sys.argv)
```

The resulting CSV file was manually reduced to years 2015–2017 to fit it into 25MB so it could be easily uploaded to Qlik Sense. The data was read with Qlik Sense into a custom report and combined with user data from Tarla's database [7] and Pentaho reports' directories, filenames and translated names [13]. Management, sales and warehouse personnels' report uses were compared using the Qlik Sense report (See Tables 3.4 and G.1 to G.4).

Table G.1: All management personnel's top 10 reports in year 2017.

Report Name	**Views**
Offers and sales	17530
Payments	9682
Sales by seller (Commissions)	5320
Month sales budgets by seller	5010
Warehouse assembly load	3556
Sales by office	3085
Current month sales	1154
Daily sales by office in month	1139
Sales by product code	803
Netrauta report	768

Table G.2: Target users' (see Table 4.1) top 20 reports in year 2017.

Report Name	**Views**
Offers and sales	15665
Payments	9447
Month sales budgets by seller	4522
Sales by seller (Commissions)	3604
Warehouse assembly load	3501
Sales by office	3071
Current month sales	1150
Daily sales by office in month	1087
Sales by product code	886
Order contents (Order drilldown)	261
Netrauta report	259
Orders	247
Product warehouse data over time	197
Sales by time period	112
Warehouse values	87
New sales by seller	81
Open offers	70
Inventories	58
Orders' metadata	52
Projected cashflow from deliveries	20

Table G.3: All sales personnel's top 10 reports in year 2017.

Report Name	**Views**
Sales by seller (Commissions)	11778
Payments	9104
Current month sales	7040
New sales by seller	5398
Offers and sales	4676
Warehouse assembly load	2908
My payments	2767
Orders	1143
Sales by office	1045
Order contents (Order drilldown)	846

Table G.4: All warehouse personnel's top 10 reports in year 2017.

Report Name	**Views**
Warehouse assembly load	5324
Payments	1546
Sales by seller (Commissions)	1201
Offers and sales	410
Sales by product code	116
Current month sales	63
New sales by seller	58
Product warehouse data over time	53
Sales by office	21
Inventory turnover	18

In addition to the Qlik Sense report the yearly statistics for personal users during the whole available time range from Jul 12, 2012 to Feb 18, 2018 were calculated directly from the Pentaho audit log file [50] with the following Python script:

Listing G.2: Calculating yearly statistics from Pentaho audit log to CSV file.

```
#!/usr/bin/python
# -*- coding: utf-8 -*-
# Tested with Python 2.7, may not be 3.x compatible
from __future__ import print_function
from decimal import Decimal
import codecs

base = 'org.pentaho.reporting.platform.plugin.'
marks = [base + 'ExecuteReportContentHandler',
         base + 'ReportContentGenerator',
         'instance_end']
delim = ','
setcols = ['Reports', 'Users']
numcols = ['Views']
cols = ['Year'] + setcols + numcols

# Remove substrings from report names
# to combine multiple entries into one
# (e.g. "x" and "x_orderby" are counted as one)
report_removable_substrings = [
    '_orderby',
    '_testing',
    '_with_average_support_enabled']

# Remove substrings from usernames
# to combine multiple entries into one
# (e.g. "kari2" and "kari" are counted as one)
username_removable_substrings = [
```

```
    '2',
    '_kaikki',
    '_kj']

# Non-personal usernames
# These are left out from the calculated statistics
# in order to more accurately count real users
nonpersonal_usernames = set([
    'admin',           # Administrator
    'autoupdater',     # Pentaho update script username
    'espoovar',        # Espoo warehouse
    'ketjujohto',      # Subsidiary management
    'kjhallitus',      # Keittiöjätti management
    'kpraportointi',   # Keittiöpiste reporting
    'malliryh',        # Model group
    'mestun',          # Fair username
    'myymuu',          # Muurala outlet
    'myysal',          # Salo outlet
    'myytam',          # Tampere outlet
    'projektimyynti',  # Project sales
    'tilintarkastus',  # Auditing
    'tncvar',          # TNC-Components Oü warehouse
    'varmyy',          # Varisto outlet
    'verkkokau',       # Online store
    'yritysjyv'        # Jyväskylä corporate sales
])

def main(argv = None):
    sys.stdout = codecs.getwriter('utf8')(sys.stdout)
    if len(argv) < 2:
        return
    filename = argv[1]
    stats = {}
```

```
print(delim.join(cols))
with open(filename, 'rt') as file:
    lines = file.readlines()
    for line in lines:
        if all([mark in line for mark in marks]):
            # Parse data from tab limited source row
            data = line.split('\t')
            date, report, username, duration = \
                data[0], data[3], data[5], data[10]
            report_dir, report_file = report.split('/')
            # Combine multiple report names into one
            for s in report_removable_substrings:
                report_file = \
                    report_file.replace(s, '')
            # Combine multiple usernames into one
            for s in username_removable_substrings:
                username = username.replace(s, '')
            # Remove trailing "a" from special case
            # in order to combine duplicate usernames
            # while trying to anonymize the user
            if username.endswith('atuoa'):
                username = username[:-1]
            # Skip over non-personal usernames' data
            if username in nonpersonal_usernames:
                continue
            # Parse date (e.g. "2014/10/23")
            y, m, d = date[0:4], date[5:7], date[8:10]
            # Initialize year's statistics if needed
            if y not in stats:
                stats[y] = dict()
                for col in setcols:
                    stats[y][col] = set()
                for col in numcols:
```

```
                        stats[y][col] = 0
                # Add entry to year's statistics
                stats[y]['Reports'] |= set([report_file])
                stats[y]['Users'] |= set([username])
                stats[y]['Views'] += 1
    # Print out the yearly statistics in CSV format
    for y in sorted(stats.keys()):
        print(delim.join([
            str(y),
            str(len(stats[y]['Reports'])),
            str(len(stats[y]['Users'])),
            str(stats[y]['Views'])]))
    exit

if __name__ == '__main__':
    import sys
    main(sys.argv)
```

Table G.5: Result of running listing G.2's script on Pentaho audit log [50].

Year	Reports	Users	Views
2012	13	23	22661
2013	21	75	80875
2014	30	105	186612
2015	32	106	147057
2016	31	103	125580
2017	32	95	106430
2018	30	76	11883

The Pentaho user counts from Table G.5 were compared with Tarla's from Table E.1 in Figure 3.8.

Appendix H

Optimal aspect ratio calculation

The following listing calculates the optimal aspect ratio using R's [35] bank_slopes function from the ggthemes package [34]:

Listing H.1: Calculating optimal aspect ratio using R's bank_slopes function.

```
library(ggthemes, ggplot2)
x <- seq(0, 8, by=0.25)
y0 <- c(3.68, 2.81, 3.92, 4.39, 4.84, 3.31, 3.50, 5.75, 5.40, 5.01)
y1 <- c(4.52, 6.55, 6.60, 6.02, 6.50, 7.47, 6.31, 7.71, 7.01, 6.74)
y2 <- c(6.99, 5.11, 6.41, 6.21, 5.29, 4.03, 3.52, 4.71, 3.61, 4.69)
y3 <- c(2.61, 4.08, 3.51)
y <- c(y0, y1, y2, y3)
# Calculate aspect ratio (i.e. y/x)
# with Average Absolute Slope and culling
aspect_ratio <- bank_slopes(x, y, method = "as")
aspect_ratio
# aspect_ratio is 0.1725888
```

Appendix I

Jeeves database's information schema statistics

Statistics about tables, views and columns were calculated on May 9, 2020 from Jeeves's Ideal_keittiot database [17]:

Listing I.1: Calculating column statistics per table/view.

```
SELECT
    MIN(column_count) AS minimum_column_count,
    AVG(column_count) AS average_column_count,
    MAX(column_count) AS maximum_column_count
FROM
(
    SELECT TABLE_NAME, COUNT(DISTINCT COLUMN_NAME) AS column_count
    FROM Ideal_keittiot.INFORMATION_SCHEMA.COLUMNS
    WHERE TABLE_CATALOG='Ideal_keittiot' AND
    TABLE_SCHEMA='dbo'
    GROUP BY TABLE_NAME
) AS subquery;
```

The results of Listing I.1 were a minimum column count of 1, an average column count of 32 and a maximum column count of 492.

Listing I.2: Calculating total columns in a database.

```
SELECT COUNT(*)
FROM
(
    SELECT TABLE_NAME, COLUMN_NAME
    FROM Ideal_keittiot.INFORMATION_SCHEMA.COLUMNS
    WHERE TABLE_CATALOG='Ideal_keittiot' AND
    TABLE_SCHEMA='dbo'
    GROUP BY TABLE_NAME, COLUMN_NAME
) AS subquery;
```

The result of Listing I.2 was a total of 103262 columns.

Listing I.3: Calculating total tables and total views in a database.

```
SELECT TABLE_TYPE, COUNT(DISTINCT TABLE_NAME) AS amount
FROM Ideal_keittiot.INFORMATION_SCHEMA.TABLES
WHERE TABLE_CATALOG='Ideal_keittiot' AND
TABLE_SCHEMA='dbo'
GROUP BY TABLE_TYPE;
```

The results of Listing I.3 were 2659 tables (BASE TABLE) and 514 views (VIEW).

Appendix J

Tarla to ABC analysis data model mapping

Tables J.1 and J.2 are the table mappings from Tarla's database (see Figures 5.3 and 5.4) to ABC analysis report's data model (see Figure 5.2). E.g. Warehouse (i.e. target table no. 9) is created from firm, office, sales_area, subsidiary, warehouse and warehouse_type.

Table J.1: Enumeration of tables in ABC analysis data model (see Figure 5.2).

Target table no.	**English name**	**Finnish name**
1	Target Space	Kohdetila
2	Offer	Tarjous
3	Offer<–>Target Space	Tarjous <-> Kohdetila
4	Office	Toimipiste
5	Delivery	Toimitus
6	Product	Tuote
7	Product Row	Tuoterivi
8	Product Group	Tuoteryhmä
9	Warehouse	Varasto

Table J.2: Mappings from Tarla's tables to ABC analysis data model [65]. See Table J.1 for mapping the numbers 1–9 to target tables in ABC analysis data model (see Figure 5.2).

Tarla's table	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
assembly_method					x				
customer_source		x							
delivery_method					x				
firm				x					x
office		x	x	x	x		x		x
office_type				x					
product						x			
product_catalog								x	
product_model						x			
product_type						x	x		
root_expanded_group_hierarchy								x	
sales_area				x					x
sales_delivery					x		x		
sales_delivery_product							x		
sales_delivery_state_type					x		x		
sales_event		x	x		x		x		
sales_event_customer_info		x							
sales_event_state_type		x							
sales_event_target_space		x	x						
sales_order		x	x		x		x		
sales_order_modification		x							
sales_order_product							x		
subsidiary		x	x	x	x	x	x		x
subsidiary_product						x			
supplier						x			
target_space	x	x	x						
target_type		x							
unit						x			
unit_type						x			
user_full_name		x			x				
vat_class							x		
warehouse							x		x
warehouse_type							x		x