LAPPEENRANTA UNIVERSITY OF TECHNOLOGY
SCHOOL OF INDUSTRIAL
ENGINEERING AND MANAGEMENT

IMPROVING REPORTING MANAGEMENT WITH RELATIONAL DATABASE MANAGEMENT SYSTEM

Master's Thesis

Pasi Lehtinen
September 2014
Lappeenranta, Finland

Supervisor: Prof. Timo Kärri
Supervisor 2: Lasse Metso
Instructor: Antti Aromaa

ABSTRACT

Author: Pasi Lehtinen	
Subject: Improving reporting management with relational database management system **Department:** School of Industrial Engineering and Management	
Year: 2014	**Place:** Lappeenranta
Master's thesis. Lappeenranta University of Technology. 71 pages, 3 tables, 14 figures and 12 appendices. Examiner: Prof. Timo Kärri, Examiner 2: Lasse Metso	
Keywords: Data management, performance management, continuous monitoring, performance measurement, database management system **Hakusanat:** Tiedonhallinta, suorituskyvyn johtaminen, jatkuva seuranta, suorituskyvyn mittaaminen, tietokannan hallintajärjestelmä	

Data management consists of collecting, storing, and processing the data into the format which provides value-adding information for decision-making process. The development of data management has enabled of designing increasingly effective database management systems to support business needs. Therefore as well as advanced systems are designed for reporting purposes, also operational systems allow reporting and data analyzing. The used research method in the theory part is qualitative research and the research type in the empirical part is case study. Objective of this paper is to examine database management system requirements from reporting managements and data managements perspectives. In the theory part these requirements are identified and the appropriateness of the relational data model is evaluated. In addition key performance indicators applied to the operational monitoring of production are studied. The study has revealed that the appropriate operational key performance indicators of production takes into account time, quality, flexibility and cost aspects. Especially manufacturing efficiency has been highlighted.

In this paper, reporting management is defined as a continuous monitoring of given performance measures. According to the literature review, the data management tool should cover performance, usability, reliability, scalability, and data privacy aspects in order to fulfill reporting managements demands. A framework is created for the system development phase based on requirements, and is used in the empirical part of the thesis where such a system is designed and created for reporting management purposes for a company which operates in the manufacturing industry. Relational data modeling and database architectures are utilized when the system is built for relational database platform.

TIIVISTELMÄ

Tekijä: Pasi Lehtinen	
Työnnimi: Raportointijohtamisen tehostaminen relaatiotietokannan hallintajärjestelmän avulla **Laitos:** Tuotantotalouden osasto	
Vuosi: 2014	**Paikka:** Lappeenranta
Diplomityö. Lappeenrannan teknillinen yliopisto. 71 sivua, 3 taulukkoa, 14 kuvaa ja 12 liitettä. Tarkastaja: Professori Timo Kärri, Tarkastaja 2: Lasse Metso	
Hakusanat: Tiedonhallinta, suorituskyvyn mittaaminen, jatkuva seuranta, suorituskyvyn mittaaminen, tietokannan hallintajärjestelmä **Keywords:** Data management, performance management, continuous monitoring, performance measurement, database management system	

Tiedonhallinta koostuu tiedon keräämisestä, säilyttämisestä ja prosessoinnista sellaiseen muotoon, mikä tarjoaa johtajille lisäarvoa tuottavaa tietoa päätöksenteon tueksi. Tiedonhallinnan kehittyminen mahdollistaa yhä tehokkaampien tietokannan hallintajärjestelmien kehittämisen liiketoiminnan tueksi. Siinä missä kehittyneitä integroituja järjestelmiä myös operatiivisia järjestelmiä voidaan suunnitella raportointitarkoituksiin ja tiedon analysoimiseen. Teoriaosuudessa on käytetty tutkimusmenetelmänä kvalitatiivista tutkimusta ja empiirisessä osuudessa tutkimustyyppinä on käytetty tapaustutkimusta. Työn tavoitteena on selvittää raportointijohtamisen ja tiedonhallinnan asettamat vaatimukset tietokannan hallintajärjestelmälle. Teoriaosuudessa tutkitaan kyseisiä vaatimuksia sekä relaatiokannan soveltuvuutta järjestelmäalustaksi. Lisäksi tuotannon operatiiviseen seurantaan soveltuvia suorituskyvyn mittareita on kartoitettu teorian pohjalta. Tutkimuksessa käy ilmi, että tuotannon operatiivisen suotituskyvyn seurantaan käytettävien mittareiden tulisi ottaa huomioon aika-, laatu-, joustavuus- ja kustannusnäkökulmat. Tuotannon tehokkuutta korostettiin erityisesti tähän tarkoitukseen sopivana mittarina.

Suorituskyvyn johtaminen ymmärretään jatkuvana suorituskyvyn mittareiden seuraamisena. Tiedonhallinnan työkalun tulee kattaa suorituskyvyn, käytettävyyden, luotettavuuden, skaalautuvuuden ja tietosuojan näkökohdat täyttääkseen raportointijohtamisen asettamat vaatimukset. Vaatimusten pohjalta muodostettua järjestelmäsuunnittelun viitekehystä hyödynnetään työn empiirisessä osuudessa, jossa on suunniteltu ja rakennettu tietokannan hallintajärjestelmä valmistavassa teollisuudessa toimivalle yritykselle tuotannon raportointitarkoituksiin. Järjestelmä on rakennettu relaatiotietokannalle hyödyntäen relaatiotietomallinusta ja -tietokanta-arkkitehtuureja.

TABLE OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

A = Availability

E = Efficiency

l = Length [m]

p = Performance

P = Product

Q = Quality

t = Time [min]

T = Trim

V = Speed [m/min]

% = Percentage

SUBINDEX'S

a = Available

c = Conforming

e = Delivered to clients

m = Mechanical

nc = Non-conforming

p = Planned

po = Production orders

r = Rate

s = Stopped

w = Waste

ABBREVATIONS

BI = Business Intelligence

DBMS = Database Management System

DM = Data Management

DSN = Data Source Name

EE = Equipment Efficiency

ETL = Extract-Transform-Load

IT = Information Technology

KPI = Key Performance Indicator

ODBC = Open Database Connectivity

PLC = Programmable Logic Control

RDBMS = Relational Database Management System

RM = Reporting Management

SME = Small and Medium-size Enterprise

SQL = Structured Query Language

WCM = World-Class-Manufacturing

WIP = Work-in-Process

1 INTRODUCTION

1.1 Background

During the last few years the importance of data management has increased significantly. Simultaneously information technology (IT) hardware systems have developed and memory space has become cheaper. This enables of designing increasingly effective database management systems to support business needs. New operational systems can process data to information and thereby support decision-making. Therefore as well as advanced systems are designed for reporting purposes also operational systems allow reporting and data analyzing.

With effective reporting management, the firm may gain some value adding information to support decision-making. Data management techniques are needed when efficient reporting management is pursued. Data management can be defined in multiple different ways. In this paper data management is understood as a process which is responsible of controlling a database and information flows.

As already mentioned technological progress enables advanced reporting in operational systems, and therefore, when designing the new database, emphasis should be given to reporting management as well. The continuous monitoring of performance measures enables managers to respond to changes even faster. Keeping operational managers informed about the performance of operational processes allows a performance enhancement of internal processes. To keep this kind of information available, operative systems need to be designed not only for collecting and storing the data but also for processing it.

At the moment reporting management and data management is based on excel spreadsheets and therefore forming new reports is time consuming. Secondly, it can be stated that current data management is not enough efficient and reliable since inputting data manually causes data errors. Before the new database project may start, the requirements needs to be defined first.

1.2 Limitations, objectives and research questions

In this paper data management and operational system requirements are studied which are set down by the objectives of reporting management. Also some data mining technologies are considered when determining the ideal state of reporting and data management system. In this thesis, reporting management is considered as a continuous monitoring of performance measures and the main objective is to generate useful information to support decision-making. More importance is paid on data management requirements that have direct impact on reporting management but all parts of data management are introduced in the theory. Only production processes are handled and other areas of business activities are left out of consideration.

Objective of this paper is to design and create a relational database management system (RDBMS) for Tetra Pak Production Oy's reporting purposes. To create a system to fulfill all user demands and to avoid common issues regarding designing process, theories of data management and reporting management are investigated. Then a framework for system design is built based on explored theories. To get a better view about these objectives three research questions were formed:

1. *How should the database management system be designed in order to fulfill the demands of both data management and reporting management?*
2. *What are appropriate key performance indicators in order to support operational decision-making in production?*
3. *What requirements do reporting management and data management have for database management system design, and does relational data model fulfill these demands?*

These questions are first answered and afterwards the resulting information is utilized in the case study. The objective of the case study is to design and create a RDBMS to enhance reporting management.

1.3 Implementation methods and structure

This thesis consists of literature review and empirical case study about designing a data management tool for Tetra Pak Production Oy to improve reporting management. The figure one demonstrates the structure of the thesis. The used research method in the theory part is qualitative research and the research type in the empirical part is case study.

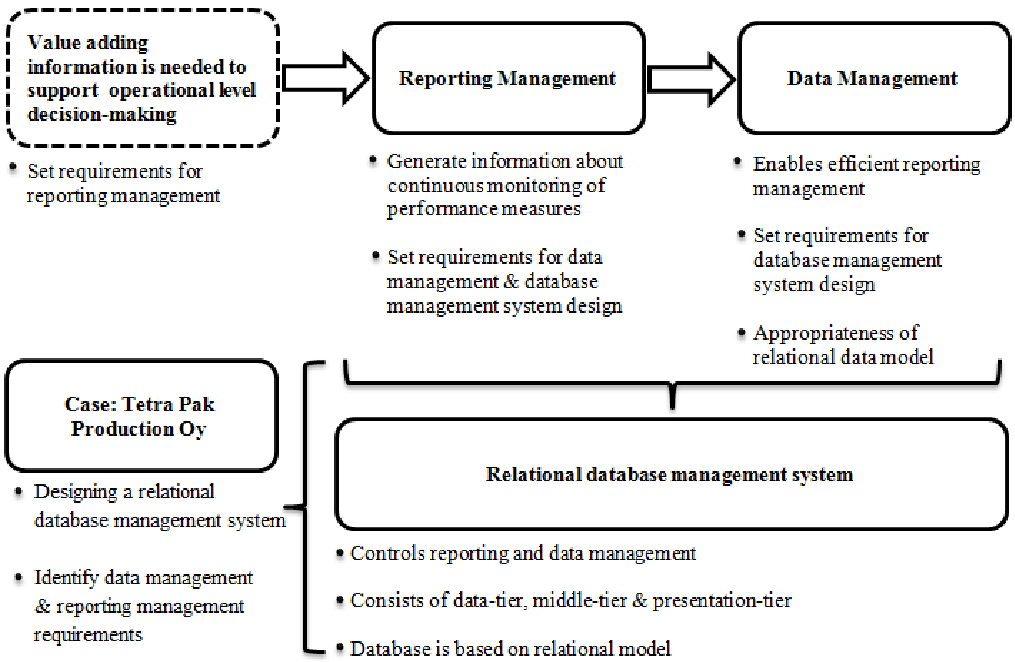


Figure 1. The structure of the thesis

Data management aims to serve several purposes simultaneously. Enhancing efficiency of reporting management is one of the objectives among others. This thesis concentrates on reporting management purposes. As figure one show, the main objective of reporting management is to generate value adding information to support decision-making. Data management has an essential role in decision-making since poor data management may be a responsible of major business damages if managers make decisions based on incorrect information.

The first section of this paper explores requirements regarding reporting management. The second section reveals the requirements of data management in order to fulfill the demands of reporting management. The theory of reporting management and data management is obtained from several different sources, and based on the literature review a framework is formed to support RDBMS designing process. This framework could be determined as an instruction manual for designing a new data management tool. The third section focus on designing data management system, and the framework for system design has been applied for case study company Tetra Pak Production Oy. The database management system consists of three tiers, which together form a wholeness that fulfills the demands of reporting and data management. The three tiers are data-tier, middle-tier, and presentation-tier. System designing is handled in the empirical part according to these three tiers.

Empirical part starts with defining objectives and requirements of reporting. Then, a system can be developed which fulfills these requirements and enhances the efficiency of reporting and data management. The last part of the empirical study introduces how the system needs to be maintained in order to remain gained advantages also in the future.

2 THE REPORTING MANAGEMENT OF THE PRODUCTION

2.1 The operational performance measurement of the production

The main purpose of reporting is considered to support decision-making and moreover to enable the continuous monitoring of the production. Performance measurement can be divided into strategic and operative performance measurement (Ukko et al. 2007, p. 39). This paper handles the operative-level reporting management of the production.

The main idea of performance measuring is that the operative targets would support the strategic targets (Ukko et al. 2007, p. 47). In other words, key performance indicators (KPI's) should be based on the corporation strategy and they should support operative-level decision-making.

Companies that are using performance measurement systems may achieve performance improvements. Successful companies has their focus on continuous improvements and strategic performance measurement. The measurement leads to significant performance improvements if the performance measurement is managed properly. (Bitici et al. 2004, p. 38–39) This can be achieved if performance measurement systems are designed, implemented and used successfully.

The proper usage of the performance measurement system can lead to improved performance. To maximize advantages, information should be understandable and the reporting system should be easy to use. This is why planning of the data management tool is very important. By taking into consideration both front and end users, it is possible to gather and use the data in the way that supports managers. The right information enables managers to allocate resources to right activities. (Ukko et al. 2007, pp. 48–50) Other requirements should be taken into account also in the performance measurement system development process.

Hudson et al. (2001) evaluated the appropriateness of the performance measurement systems for the small- and the medium-sized enterprises (SME). The evaluation is based on the nine criteria of the development process, the seven critical characteristics and the six dimensions of the performance measures (Hudson et al. 2001, p. 1102). System developer should keep in mind these characteristics when developing the new performance measurement system. Although this framework is designed for a development process, these requirements can also be used for the evaluation of the existing performance measurement system. The table one illustrates the characteristics of successfully designed performance measurement system.

Table 1. The characteristics to evaluate the existing performance measurement system or the development process of it (formed from Hudson et al. 2001, p. 1102)

The development process requirements	The performance measure characteristics	The dimensions of the performance
Need evaluation/existing performance measurement audit	Derived from the strategy	Quality
Key user involvement	Clearly defined/explicit purpose	Flexibility
Strategic objective identification	Relevant and easy to maintain	
Performance maintenance structure	Simple to understand and use	Time
Top management support	Provide fast and accurate feedback	Finance
Full employee support	Link operations to strategic goals	Customer satisfaction
Clear and explicit objectives	Stimulate continuous improvement	Human resources
Set timescales		

In addition to the characteristics defined by Hudson et al. (2001) some other performance measure requirements are handled in the paper of Paola Cocca and Marco Alberti (2009). The purpose of their work has been to define an assessment tool of the existing performance measurement system for the SME's.

Besides the characteristics that are defined in the table one, the performance measures should be easy to collect, monitor past performance but also plan future performance, promote integration, and the formula and the source of the data should be defined. Since the environment where firms operate is dynamic, performance measurement systems should reflect to these changes quickly as well. Therefore, the performance measurement systems as a whole should be very flexible, rapidly changeable and maintainable. As already mentioned the performance measures should be easy to collect. This is because effort needed for measuring is supposed to be less than the benefit gained from it. (Cocca & Alberti, 2009, pp. 186, 193–194) There are a lot of indicators that could be measured but if the resources wasted is more than advantages gained from that information, managers should rethink the need of that information or think of ways making the measurement more cost effective. One way of decreasing the costs of data collection could be achieved by improving the data automation in the data supply chain. Planners should keep also these aspects in the mind when designing the new performance measurement system. The system as a whole is supposed to be easy to implement, use and run but moreover easy to maintain.

So far the literate has been focused on the performance measurement process and the indicator requirements. System designers should take into account also the aspects of the performance measurement system as a whole. Some of the criteria were already mentioned. The remaining characteristics have been gathered to the following list:

- The requirements of the performance measurement system as a whole:
  - All stakeholders considered.
  - Flexible, rapidly changeable and maintainable.
  - Balanced.
  - Synthetic.
  - Easy to implement, use and run.
  - Casual relationships shown.
  - Strategically aligned.

- Graphically and visually effective.
- Incrementally improvable.
- Linked to the rewarding system.
- Integrated with the information system. (Cocca & Alberti, 2009, p. 194)

In the situation where the KPI's are given, the characteristics in the list above are crucial for the success of the system design process. Since the managers have already assessed the indicators and decided of their suitability to the corporation strategy, what is left for the planner is to create a system that supports these indicators. Performance measurement projects may lead to failure if the data capturing systems are not able to support measures (Bitici et al. 2004, p. 38).

The performance measurement system should be built in a way that employees would feel like they are "in-control" and not being "controlled". Thereby the system encourages employees to think smarter, rather than just work harder. If the system supports the idea where managers are not trying to control the performance of individuals, extrinsic reward systems to motivate employees are not needed. (Robson 2004, pp. 139–142) On the other hand, if the rewarding systems are used then the performance management system should be linked to them (Cocca & Alberti 2009, p. 194). Despite of whether the rewarding systems are used or not, the KPI's should be chosen carefully keeping the nature of the indicators in the mind.

The criteria defined in the literature focus on the system development process. These characteristics need to be taken into account when designing the system to achieve the long and the short timescale goals. According to research study made by Hudson et al. (2001, p. 1112), based on interviewing the managers of the SMEs, showed that the current performance measurement systems had their weaknesses but none of the managers had made any actions to redesign or update their current performance measurement systems. Although the literature has widely shown for the planners what should be measured, there is not a straight forward framework that identifies how the system should be designed to fulfill the requirements of the performance measurement systems. This suggests that one can't define a step-by-step process that fits for everyone's needs and therefore only volatile evaluation approaches are developed. Later on this paper, the data management requirements are considered from the context of the performance measurement system.

2.2 Continuous operational monitoring

The continuous monitoring of the internal processes allows managers to monitor processes constantly and gain dynamic information that supports operational-level decisions. Continuous monitoring system should include the KPI's that are chosen for the performance measurement.

According to the study of Bourne et al. (2005, p. 385), managers should use the information gained from the performance measurement systems intensively. One of the performance measurement criteria is to provide fast and accurate feedback to its users. Nowadays the importance of real time information is in a major part of an operative-level decision-making process. As mentioned already, the environment is changing rapidly which sets requirements for reporting management. The information is supposed to be available at any time of the day which is why the continuous monitoring of the processes is vital for achieving effective reporting.

In companies, which can be called "average-performance business units", the performance management is based on a simple control approach. These companies collect data through standardized systems, analyze the data, and compare the results to company targets. Companies that are so called "high-performing business units" have a continual interaction with the performance data. Managers have their own data collection systems and KPI's. The information is gained continuously and actions are made throughout the action period rather than waiting for the next meeting. (Bourne et al. 2005, p. 386) This of course insists an understanding of how the performance measures impact on the performance but this is rather simple if operational managers and system users are involved to the system development process and thereby their understanding regarding the performance measures increases.

Company environment is changing and to respond these changes the company has to measure, monitor, and re-evaluate its processes continuously. Business monitoring is not only auditing but also making actions and ensuring process performance. To enable continuous monitoring the system should fulfill below listed three elements:

1. Measuring the actual business process.
2. Comparing and evaluating the actual values of business process to basic standards.
3. Alerting the firm about potential issues. (Mancini et al. 2013, p. 124)

The continuous monitoring and the evaluation of values enables the controlling of business processes and progressive performance. These basic elements are executed with available information technology which can be highly automated or manual systems. The continuous monitoring can be adopted and implemented in a best practice by firms that have a fully automated information system. The information system should be able to collect, store, process and distribute data for management to enable the continuous performance analyzes of the business processes. (Mancini et al. 2013, pp. 125, 131) Usually data that is needed for reporting and monitoring is scattered into multiple different locations. Later in this paper is discussed how the data should be managed in the way that it enables continuous monitoring.

There are four important steps for developing a continuous monitoring approach. First of all, planner needs to define and analyze the processes. After analysis, the trends in the processes need to be identified. Now it is possible to develop the continuous monitoring system that can enhance planning and control. Taken that the KPI's are given, the third part is to build the monitoring system based on the reporting system. The last step in the development process is to design or re-design the IT system. (Mancini et al. 2013, pp. 131–132)

Information flows, systems and networks are relevant parts of the performance management system. Actually operating performance management systems, such as the continuous monitoring system of the production, may be part of the information system and IT infrastructure. The information system can be understood as a platform provider and IT is needed for reporting and performance management purposes. (Ferreira & Otley, 2009, pp. 273–274)

The continuous monitoring enables the creation of dashboards for performance reporting and providing continuous dynamic information rather than static analysis. Therefore managers are able to take immediate actions and achieve better performance in critical processes. (Mancini et al. 2013, p. 133) To gain these advantages, one needs to develop an appropriate system for collecting the data. It is important to create an information system that supports reporting and continuous monitoring without having any side effects on operational processes. For example if entering the data to the system affects on employees concentration on other processes, it makes sense to think of automated data input. There are also other issues to be considered when designing the system for reporting management purposes. Ferreira and Otley (2009, p. 274) highlight issues such as information scope, timeliness, aggregation, integration, level of detail, relevance, selectivity, and orientation. All of these characteristics need an in-depth consideration when such systems are being developed.

2.3 Appropriate operational key performance indicators for measuring production

The performance evaluation can be done from operations perspective, strategic control perspective and management account perspective (Bitici et al. 2012, p. 306). In this paper, the performance measurement is handled from operations perspective, and moreover from manufacturing's point of view. The chosen KPI's should be based on the corporate strategy (Hudson, Smart & Bourne 2001, p. 110; Rausch, Sheta & Ayesh 2013, p. 7). At this point it is taken that common objectives for manufacturing companies is to be as efficient as possible and to reduce waste from its processes. In other words, manufacturers aim to optimize their manufacturing process and to deliver high quality products by pursuing lean production.

No literature has introduced a standard approach for reporting that suits for every situation (Karim & Arif-Uz-Zaman 2013, p. 182). The identification of appropriate KPI's is essential in order to improve the operational performance of manufacturing. In this section, the appropriate KPI's are studied for operational decision-making purposes in production. The literature has introduced several different approaches for measuring the manufacturing performance.

Increasing competition forces manufacturers to optimize their production processes in order to increase production efficiency and quality, and to reduce waste and non-value-adding activities (Karim & Arif-Uz-Zaman 2013, p. 169, 171). Different tools and approaches has been developed for solving manufacturing efficiency problems. Those techniques support the continuous process improvement of manufacturing. These approaches can be implemented more systematically if appropriate performance measures exists in order to support decision-making (Wan & Chen 2009, p. 277; Hicks 2007, p. 234). Researchers have introduced a number of different indicators for evaluating operational manufacturing performance. These indicators are collected to the table two.

Table 2. Overview of the KPI's used in the operational level of the manufacturing

Authors, year and article	**The dimensions of measurement introduced in that article**	**Indicators introduced in that article**
Karim, A. & Arif-Uz-Zaman, K. 2013. A methodoly for effective implementation of lean strategies and its performance evaluation in manufacturing organizations.	• Continuous measurement approach • Time, cost, quality & flexibility • The leanness of production	• Continuous performance measurement (CPM) – effectiveness & efficiency • Efficiency, effectiveness, performance, productivity, utilization, value-adding/non-value-adding ratio, throughput & defect rate
Demeter, K. 2013. Operating internationally – The impact on operational performance improvement.	• Operational performance improvement measures • Cost, quality, speed, reliability, flexibility	• Product quality and reliability, capacity utilization, delivery speed • Manufacturing conformance, product customization, volume flexibility, mix flexibility, time to market, product innovativeness, delivery reliability, unit manufacturing cost, manufacturing lead time, labour productivity, inventory turnover, manufacturing overhead cost
Chavez, R., Gimenez, C., Fynes, B., Wiengarten F. & Yo, W. 2011. Internal lean practices and operational performance.	• Operational performance dimensions (quality, delivery, flexibility, cost)	• High product performance, high product reliability • Short delivery time, delivery on due date, on-time delivery • Ability to introduce new products into production quickly, ability to adjust capacity rapidly within a short time period, ability to make design changes in the product after production has started • Labour productivity, production cost, reducing inventory
Niedritis, A., Niedrite, L. & Kozmina, N. 2011. Performance measurement framework with formal indicator definitions.	• Indicator definition (Type of the indicator, reporting period, perspective, success factors, level of details, activities and processes which are being evaluated)	
Phan, C,A. & Matsui, Y. 2010. Comparative study on the relationship between just-in-time production practices and operational performance in manufacturing plants	• The operational indicators of the manufacturing performance	• Manufacturing cost, on-time delivery, volume flexibility, inventory turnover, cycle time
Wan, H. & Chen, F.F. 2009. Decision support for lean practitioners: A web-based adaptive assessment approach	• Lean assessment • Lean indicators • Lean scores	
Bayou, M.E. & de Korvin, A. 2009. Measuring the leannes of manufacturing system – A case study of Ford motor Company and General Motors	• Manufacturing performance	• Efficiency • Effectiveness
Gomes, C.F., Yasin, M.M. & Lisboa, J.V. 2007. The measurement of operational performance effectiveness: an innovative organisational approach	• Manufacturing operational effectiveness	• Manufacturing Operational Effectiveness (MOE) • Efficiency • Availability • Quality

According to Karim & Arif-Uz-Zaman (2013, p. 175) previous theories have focused on measuring cost, quality, lead time, processing time, operations time and value-added time. The table two reveals that the most of the indicators focus on measuring the leanness of the manufacturing processes. The most common performance measurement dimensions are related on time, quality, flexibility and cost. Manufacturing effectiveness and efficiency have been highlighted as well. Choosing the right indicators is crucial in order to succeed in delivering right information for decision-support. As literature review revealed, there is no standard method used for measuring the operational performance of the manufacturing. Therefore, managers need to identify appropriate indicators themselves. Managers may use one of the techniques introduced in the literature or use the indicator definition approach introduced by Niedritis, Niedrite & Kozmina (2011). Managers should also take into account already existing information and used indicators in order to ensure cost effectiveness. This paper does not take a stand which approach should be used in general but in the section four one of the methods is used for the case of Tetra Pak Production Oy.

3 DATA MANAGEMENT

3.1 Introduction to data management

Data integration causes most problems when designing systems for the performance measurement purposes (Rantanen et al. 2007, p. 419). Since reporting management sets certain requirements for data integrity and availability, this paper discusses about the requirements of data management. The figure two demonstrates how top level management, financial management and data management are related. Organizations have an essential need for producing and exploiting data in its processes.

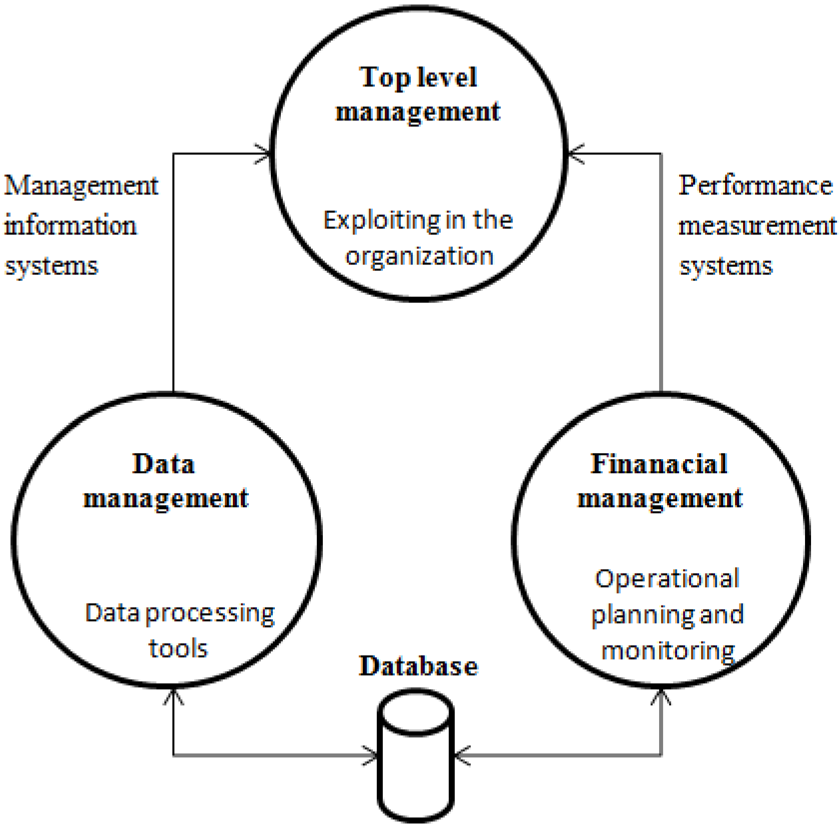


Figure 2. Data management and its relations in organization (formed from Hovi, Ylinen & Koistinen et al. 2001, p. 187).

Data is a resource just like any other. This resource is utilized mostly by financial management departure and top level management. Nevertheless, also operational level managers can use the data resources when the form of the data is understandable. Nowadays obtaining the data is not the problem anymore. More critical issues are related into other parts of the data processing chain, for example to the inability to generate useful information from the data (Lee & Siau 2001, p. 41). In this paper data management is understood as collecting, processing and providing the data for its users in informative form, and therefore the term life cycle of data processing can be used.

In the first part of the data processing chain, the data is collected. This data is saved and stored into a database where it is turned into more informative form. Next phase is the distribution of the information to its users. After distributing the information, it can be handled in the human minds and turned into knowledge and wisdom. (Kaario & Peltola 2008, pp. 8–10) The whole value chain of the data processing can't be automated since the last part requires information processing phase inside human brains. Information itself is quite useless if it is not processed actively in the mind of human being (Mancini et al. 2013, p. 141). Although the tail end of data processing chain is important it can't be automated. Therefore this paper is handling data collection, processing in the information system, and distribution in the form of the performance reporting, leaving out of the consideration data management that occurs inside brains.

3.2 The requirements of data management

Data management requirements depend on the purpose of solution needed. As the business grows, data management becomes more difficult (Granlund & Malmi 2004, p. 24). Business demands differ widely from each other but the structure of the data has also different forms. Information technology provides different solutions for structured and unstructured data. This paper discusses about the requirements of structured homogenous data to meet the demands of reporting management.

Xu and Quaddus (2013 p. 140) have gathered the most common issues related to low quality information:

- Errors in collecting and entering the information;
- Information exists in different systems and different entry standards and formats;
- Information is missing due to wrongly designed systems;
- Information is inconsistent and inaccurate;
- Valuable information can’t be shared since it is trapped in the organization silos;
- Information is lost because of poor system integration;
- Information is not presented in user friendly formats.

This section of the paper reveals the requirements of the system itself and data mining techniques to avoid these above-mentioned issues. Those issues are faced due to the low quality of the information. Besides of the quality aspect, some other information system requirements are handled as well.

3.2.1 Information system requirements

Information systems are needed for supporting decision-making. In other words, they are essential for enhancing reporting management. Business operations generate new data rapidly but to benefit from it, data management and information systems are needed. There is a variety amount of different information systems for decision-making. Information systems are composed of decision support systems, executive information systems, data warehousing and data mining just to mention a few examples (Xu & Quaddus 2013, p. vi).

System solution characteristics depend strongly on the structure of the data and the purpose of storing it. A column-oriented database is the best solution if the data is stored only once and reports demand fast response time from queries that gather information from multiple data stores. (Chandran 2013, pp. 35–36) Common to every system is that they are supposed to improve decision-making.

Three important characteristics of the information systems cover the improved decision-making. First of all, the right information needs to be available at the right time. Secondly, the information needs to be available anywhere and anytime. The third requirement is the same, as Mancini et al. (2013) highlighted in the topic of the continuous monitoring which is that the system needs to alert users if any issues are detected. (Xu & Quaddus 2013, p. 139) To pull together, information system is supposed to generate right information to right people at right time.

The RDBMS could be an appropriate solution to fulfill above-mentioned demands. Actually, the relational database system suits for the situation when the data is clearly defined (Scott et al. 2013, p. 40). Therefore the system itself sets some requirements for the data entry. It needs to be clear, in which form the data can be entered to the system. Users may also set some requirements to the system's data management. If the users want to perform simple searches in the system, interface features need to support these activities, and more over, the data needs to be structured in a way that the performance of a query is acceptable (Scott et al 2013, p. 49). This leads us to a closer review of the system requirements. The following questions need to be considered when talking about the architectural approach of the information system: 'What database should be used?', 'What tools provide the information (reports, dashboards and/or multidimensional cubes)?', and 'How is the data secured?' (Mancini et al. 2013, p. 132). To be more accurate than 'what database should be used', it makes sense to investigate database application's performance, usability and the possibility of data scalability in the chosen system. Another important aspect to consider is the data quality, and therefore reliability is added to the list as well.

Database application performance

The development objectives of data management are usually somehow related to database application performance, for example better performance on searching useful information from a large amount of data or overall a more efficient use of data resources. Data management objectives can be divided into three categories which are efficiency, availability, and quality (Kaario & Peltola 2008, p. 128).

Database application performance is answer to these objectives. Proper indexing may reduce the time of finding certain data contents as well as data mining allows managers to use data resources in a more efficient way since the availability of valuable information increases. The efficiency of data management can also be enhanced by improving the overall system architecture and design. The system itself should be designed in a way that enables automated data processing and the avoidance of duplicated data inputs.

Whereas the overall performance can be enhanced with the data mining, simultaneously the data mining itself has some requirements for other data management features. To make clear, data mining systems are individual systems which are attributable to the database management system and they are able to generate useful information from a large amount of data. Lee & Siau (2001) gathered a list of requirements and challenges according to the data mining. They stated that data mining algorithms should be efficient and scalable, meaning that the time which is used for searching, mining, or analyzing should be predictable and acceptable as the amount of data increases (Lee & Siau 2001, p. 42). In other words, the data mining enables better database performance but in the same time it requires efficient data management from the other parts of data management. In conclusion, the next list of objectives and requirements need to be fulfilled to achieve a desired wholeness:

- The efficient utilization of data resources;
- The accelerated speed of finding useful data contents;
- Efficient and scalable data mining algorithms;
- Avoid duplicated data inputs;
- Efficient and automated data processing;
- Proper designed database content.

Scalability

Changes in business environment, business strategy, processes, organizational structure, law instruments, or in technology affects more likely to firms data management (Kaario & Peltola 2008, p. 145). Nothing lasts forever but at least it can be taken into account when designing the new data management system.

It is necessary to consider the compatibility of data and database systems after changes. Another issue is the data accessibility and scalability in the long run. (Kaario & Peltola 2008, p. 145) It is obvious that incremental changes are needed in the system over time but an important feature of the database management system is the ability to adapt into a new circumstances. Data independence offers flexibility over time to the scalability issue but requires work in the designing phase.

Data independency allows users to create new tables and alter current tables without affecting to other applications. Database application performance can be enhanced by defining new indexes and still old programs can maintain their functions and no changes are required regarding already existing tables. (Hovi, Huotari & Lahdenmäki 2005, p. 12) This is possible due to the data independency and system properties which support scalability.

Scalability has been achieved when multiple users can access same data without affecting the performance of the system. Therefore, designer needs to be aware of the scalability requirements when planning new operational systems. This can be reached rather easy with relational database products because programmer don’t need to decide **how** the data is being retrieved. Access path to certain information is created by an optimization program, and user needs to define only **what** information needs to be retrieved. (Hovi, Huotari, Lahdenmäki 2005, p. 13) This allows the users to get the valuable and needed information, even though the programmer wasn’t aware of all user demands in the first place. Still, it is desirable to be aware of the systems requirements already in the programming phase in order to achieve the desired performance, meaning that the necessary indexes are taken into account.

Data scalability and availability issues become essential when the number of database users increase. Scalability can be ensured for example by using relational database management system that run in the cloud and provide flexibility as data processing requirements vary. Another solution is to scale up the hardware while the users and the volume of requests expand, though this can be quite expensive. More cost-effective solution is to spread the load across additional nodes as the size of the data and users increase. Each node is a database in its own right and this entity is managed with the database management system (DBMS). The last approach is called sharding a database. (McMurtry et al. 2013, p. 62) The database sharding can be implemented in many ways and each one of them has their own advantages and disadvantages. Though the sharding solutions are not getting any closer look, issue that has been often highlighted by McMurtry et al. (2013) related to data availability as a consequence of dividing data into separated databases. The Open Database Connectivity (ODBC) is a technology used for data integration, and it can be utilized when connecting multiple databases via user interface application (McMurtry et al. 2013, p. 70). As a conclusion, data scalability can be improved later as well but to minimize the costs of it, scalability should be considered already in the database designing phase. Features that enables scalability may cause harm to data availability which is why other techniques, such as the ODBC, should be used in parallel. A database design is handled with a greater interest in the fourth section of this paper.

Reliability

The importance of data quality needs to be highlighted in the context of planning and reporting. Data quality plays an important role in reporting since decisions which rely on incorrect information may cause enormous economic damages. The biggest problems in the context of data quality are caused by employees in the phase of data entry. Inputting data manually into the system may cause problems but automated data entry isn’t always an available option. (Rickards & Ritsert 2012, p. 28) Even though data entrance can’t be automated in every situation, it can be arranged in a way that mistakes in the data input phase are reduced. This can be achieved for example by using check constraints which allow the user to entry only certain predefined values. Another issue regarding data quality is an inconsistent definitions of commonly used terms (Rickards & Ritsert 2012, p. 28). Therefore data managers need to pay attention to the employees’ awareness of unambiguous terms.

The quality of the data is acceptable when information meets the user’s expectations. This demand is not as easy to fulfill as it sounds because user’s requirements are not always easy to define and they tend to vary across time. (Rickards & Ritsert 2012, p. 29)

Data quality is essential in the topic of data management. Information doesn’t add any value if the data is not reliable. The person who is responsible of planning and reporting is naturally in response of data quality as well. The data warehouses and the data marts can be used in the data management purposes but that alone isn’t enough. The whole reporting supply chain needs to be examined. (Rickards & Ritsert 2012, p. 27) Actually it would sound a little bit too perfect to create a database system which meets all the reporting management requirements. Still, usually data reliability is in a better shape when the data is managed with software solutions rather than with spreadsheet calculations (Rickards & Ritsert 2012 p. 31).

Usability

In order to gain value from the data stored in to the database, users need to get access to the information. Therefore, usability needs to be considered as well when designing the data management tool. This data management requirement can be achieved with add-on features that provide desired information for the users in understandable format. This kind of software applications are also known as dashboards.

Following four capabilities need to be fulfilled by dashboards. First of all, required information needs to be available on a single screen in understandable formats, such as charts and graphs. This screen should include the real-time values of the earlier defined appropriate KPI's and the system should support both drill-down and slice-and-dice options on each KPI monitored. Drill-down allows the users to gain more accurate information, whereas slice-and-dice option enables users perform what-if and sensitivity analysis. In addition to these characteristics, the dashboard should allow an integrated management of the KPIs. (Bose 2006, p. 57) Taking usability into consideration enables faster and more effective decision-making since the desired information is available in right forms, even though the user demands vary over time.

Such application, as dashboard, provides users the desired information in needed formats. In which case, the information is available but for improving the ease of use also following functionalities need to be taken into account when building the dashboard: user must be able to choose which KPI's are presented and in which forms they are displayed. Therefore, the dashboard should support different kind of easy-to-understand graphics. Another important functionality is to enable proactive alerting in the situations of exceptions and milestones. Exception-based-reporting can be displayed for example in the form of "traffic lights" whereby the colour changes and guides the user in the value analyzing, and simultaneously the system launch a trigger which for its part announce the users about the change, for example via email or alert messages on user interface. (Bose 2006, p. 57)

In addition to the user interface of the dashboard, it is important to ensure also the usability of the system in the data entry phase. Users need to be able to input data to the system without negative impact on the business operations. In other words, automated data entry should be pursued when designing the data management tool. Sometimes the data need to be entered manually, and then it should be as easy as possible which can be achieved for example by using default values.

Data privacy

Data privacy isn't a direct requirement of reporting. Data privacy issues may still cause problems in the critical areas of business features. Another reason for highlighting data privacy is that other characteristics which in other hand enhance reporting may cause a discrimination of the data privacy.

Different features, such as data mining techniques that support data management requirements in one aspect might simultaneously do harm for the second. With data mining techniques, useful information can be generated and presented to the users in forms that are easy to understand. Nevertheless, data mining can set extra demands on the data security and privacy because when data can be viewed from different angles, it threatens data privacy (Lee & Siau 2001, p. 42).

Data privacy needs to be taken into account already in the system programming phase. Different users and their roles in the organization need to be identified. This allows different users examine only certain data and confidential information remains secret. Data privacy can be achieved with user access management. Security parameters should be set based on user, group, or community type (Bose 2006, p. 57). Kaario & Peltola (2008, p. 65) in other hand recommend to bind user privileges to depend on actions rather than organization structure. Nevertheless, when different accounts are created, different permissions can be granted. This also enables monitoring user actions. For example, if certain user has made some changes for the data content, it can be tracked afterwards.

3.2.2 Data mining requirements

Common for the data management requirements is that data analytic processes demand data-intensive processing (Chandran 2013, p. 36). As mentioned earlier, data needs to be processed to retrieve essential information to support managers in business decisions.

Vucetic et al. (2012) studied the methods of finding dependencies between attributes in the relational database system. If managers can’t discover the relationship between data and attributes within the relation, special algorithms need to be developed. Discovering these dependencies may enable the performance improvement of the internal processes. (Vucetic et al. 2012, p. 2738) Reliable data not only creates value but also allows an enterprise to lower costs (Rickards & Ritsert 2012, p. 27). Discovering this hidden and useful knowledge requires data mining techniques. Before data mining, the data need to be prepared for the analysis, for example some preprocessing and data summarization may be necessary to do (Vucetic et al. 2012, p. 2740). Therefore, the form of the tables and the structure of the data need to be designed in a way which allows predictive data mining in the future.

Data mining techniques can discover valuable information from the databases and significantly enhance the ability to analyze the data. To improve reporting management, one needs to integrate data mining technologies within the existing database systems. Data mining techniques vary depending on the used database systems. There is no database that could individually suit the all aspects of data mining. Taken that data is stored into SQL Server database, data mining can be done by using SQL Server Data Mining. The features needed for data mining in SQL Server are Analysis Service and Reporting Services, which can be installed separately. (Aggarwal et al. 2012, p. 164–165, 168–169)

3.3 The appropriateness of the relational data model

A data model defines a logical structure and a format of the data that is stored into the database. In the relational model, data storing and representation is based on tables and logical connections prevailing between them. The relational model is capable of modeling the real world and the logical connections of different data concepts. Database design can be called relational if the data is represented in two-dimensional tables and it supports relational functions. (Mancini et al. 2013, p. 222)

The relational model is based on set theory, mathematics and predicate logics. Almost all new database products are based on the relational model. A structured query language (SQL) is a standardized relational database language. Even though different relational database products, for example MS Access, MySQL, DB2, Oracle, and SQL Server, have their differences they all have almost the same characteristics. (Hovi 2004, pp. 3–5, 11) This next section handles the features of the relational model and relational database products in general level. Even though, all database providers have their own methods of implementation, all the products has a quite similar basic structure to each other.

The main advantages of the relational model are its mathematical punctuality and simplicity. The relational database systems have a numerous of other advantages too. By relational database systems, one can achieve higher flexibility and data independence, relatively easy maintenance and simple data storage conditions. In addition, the relational model can reduce data redundancy and the mistakes of data entry. (Mancini et al. 2013, p. 222) Reduced data redundancy of course demands a normalization of the tables and therefore the designing of the database system requires more hours and resources than non-normalized tables. Still, the designing of the system can’t be praised enough since every hour used during development phase will save time and money on later phases. A well designed relational database allows high flexibility and a user-friendly system (Mancini et al. 2013, p. 223).

As mentioned, the relational database consists of the two-dimensional tables. The other features of the relational databases are indexes, optimizing program, data independence, stored procedures, triggers, user identified functions, and views. These characteristics allow the user to do complicated integrity checks, maintain the timeliness of the information, increase data privacy and independence, and enhance the performance of the queries (Hovi 2004, pp. 11–13). That is to say, the relational database fulfills the requirements of the information system. However, the continuous monitoring and the performance reporting set some additional requirements of their own. SQL and data warehouses together do not form the reporting tool as alone (Hovi 2004, p. 19). Therefore, information needs to be transferred to a separate program. Still, to generate reports that add value to managers decision-making, certain data mining and Extract-Transform-Load

(ETL) -processes are needed which can be achieved with the features of the relational model (Hovi 2004, p. 19).

Data warehouses enables combining data from different operational sources. Then the data is transformed to be compatible with other ones. Data warehouses also allow simultaneous data collection and effective data analyzing and reporting. (Hovi 2004, p. 19) In conclusion, it can be stated that the relational data model enables to design an entity which meets the all requirements of data and reporting management. This entity, in other words the RDBMS, consists of data collection and storing part, data cleansing part, and data presentation part.

Hovi, Ylinen & Koistinen (2001, p. 56–57) investigated how well the relational database fits for data warehouse purposes. The result was that relational databases can be used for data warehousing, though it was originally designed for query purposes. Following list contains pros and cons of the relational database:

(+) Suitable for large amounts of data and users;
(+) Effective and commonly used technique;
(+) Open for SQL-interface;
(+) Works well with combining, processing and summarizing data;
(+) Developed also in the area of data warehousing;
(–) Does not support multi-dimensional processing;
(–) Includes some unnecessary event processing features from data warehousing point of view;
(–) The performance of searching options needs still improvements.

Based on the literature review, it can be stated that the relational model and moreover the RDBMS fit for data management and reporting management purposes. Therefore, the relational data model is used when a data management strategy is formed later on this thesis.

3.4 Framework for developing a data management tool

Any successful database implementation should start from identifying the output. Starting the project from this point of view enables identifying the data types needed for certain business needs. Once the type of the data is clear the developer may continue on confirming the data structures and determining the sources of the data. (Stock 2011, p. 307)

The unreliability of reports generated from the operational systems have been a common reason for starting a new database project. Another reason, closely related to the previous reason is the general goal of improving the quality of the data. Two more short-term goals have been introduced, and they are the reducement of the overlapping duties and enabling the easier management of master data. In addition to these short-term goals, the corporation may achieve long-term goals as well. One long-term goal could be the creation of an enterprise data warehouse. This for sure would be a long-term goal if the project starts from building single data marts which are just the building blocks of the integrated enterprise data warehouse. More accurate information adds value for the decision-making and therefore another long-term goal can be achieved from discovering improvements in the business. (Hovi, Ylinen & Koistinen 2001, pp. 149–152) All of these above mentioned goals fit for Tetra Pak Production Oy as well but more accurate requirements are introduced in the context of forming the data management strategy.

Data management tool can be designed by utilizing the principles of different database architectures. The database architecture depends on the chosen data management strategy. Lopez (2012, p. 18) introduces three steps that should be considered when forming the data management strategy and designing the new data management tool:

1. Identify users’ data requirements;
2. Build a data model which supports the business demands;
3. Choose a right tool for data integration.

The purpose of data management is clear on this thesis. Data requirements are set down by reporting management. The data model has been chosen as well and earlier proved that the relational data model is adequate to the reporting purposes. Data management and reporting management requirements, which have been examined in sections 2.1–2.3 and 3.2.1, are gathered to the figure three. The grey box in the bottom demonstrates the data management requirements whereas other boxes demonstrate the reporting management requirements of the development, the system and the system functionality.

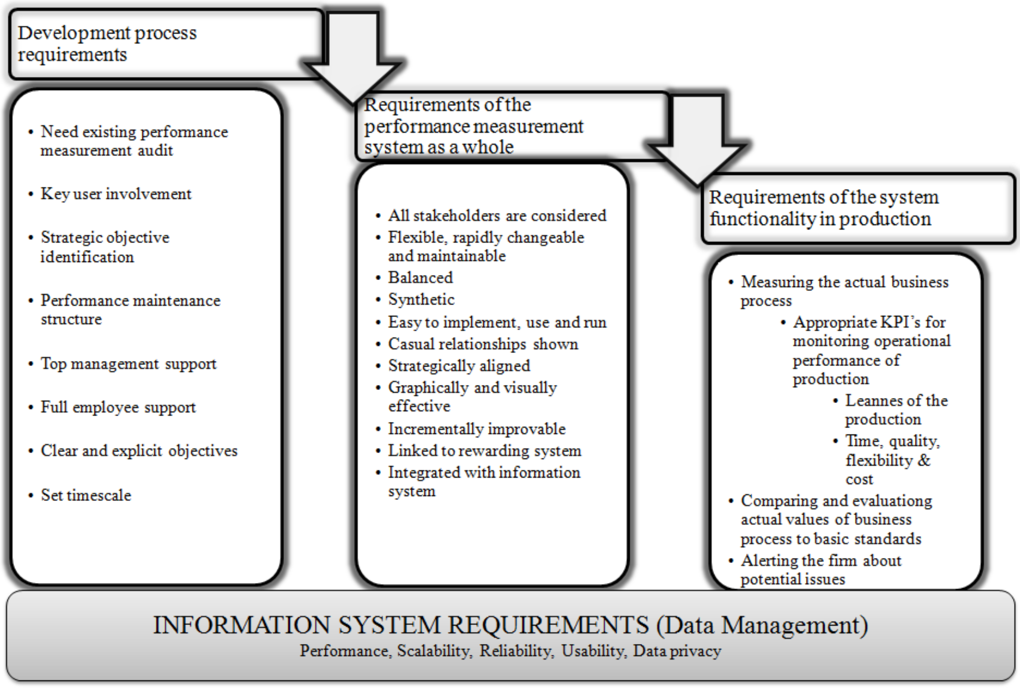


Figure 3. The user requirements regarding data management and reporting management

In this study, the system is designed for reporting purposes and more over for the continuous monitoring and the controlling of production processes. Hence, the purpose of the system is clear, still by far the largest effort remains in analyzing all parts of the system being developed (Stock 2011, p. 307). The system developer needs to arrange numerous meetings with the corporation management team and system users in order to determine the meaning of every internal process to be able to model the structure of the database. Depending on the user requirements, the developer must decide as well whether to normalize fields or not. To ensure that everything will be taken into account, a data management strategy should be formed and for example data dictionary's and data models should be constructed.

In order to follow instructions of Lopez (2012), different data models are introduced next. The third phase is to choose a tool for data integration. Therefore, this section of the paper examines also different data architecture approaches.

Data models

Data models are commonly used method for designing the structure of the database especially when designing operational databases. Operational data model should try to imitate only the most critical aspects of the business processes. (Hovi, Ylinen & Koistinen 2001, p. 92) Still, the model should contain at least the data requirements which have been identified in the previous step of forming data management strategy. In this paper, the requirements consists of the reporting and the monitoring of the operational performance measurements. Therefore, the model should be built from the viewpoint of reporting.

The star schema is commonly used data modelling technique. This technique is used especially when designing local data marts. The goal of building the star schema is to make the use of queries and reporting as easy as possible. (Hovi, Ylinen & Koistinen 2001, p. 94) In other words, modelling tool enables the consideration of data types, data requirements, user requirements and relations between different data contents.

The star schema reminds a shape of star which is why it is called star schema. This means that tables are connected into each other by relations. In the middle of the model is a fact table which is connected into multiple dimension tables. Usually fact table's rows consist from transactions, such as orders and sales, and the primary key of the fact table is a combination of dimension tables' primary keys. The figure four demonstrates the form of the star schema.

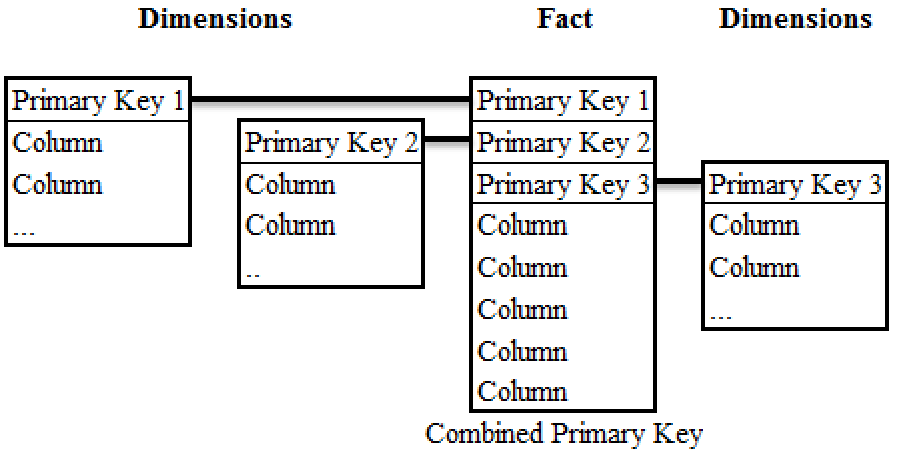


Figure 4. The basic idea of the star schema data model (formed from Hovi, Ylinen & Koistinen 2001, p. 95)

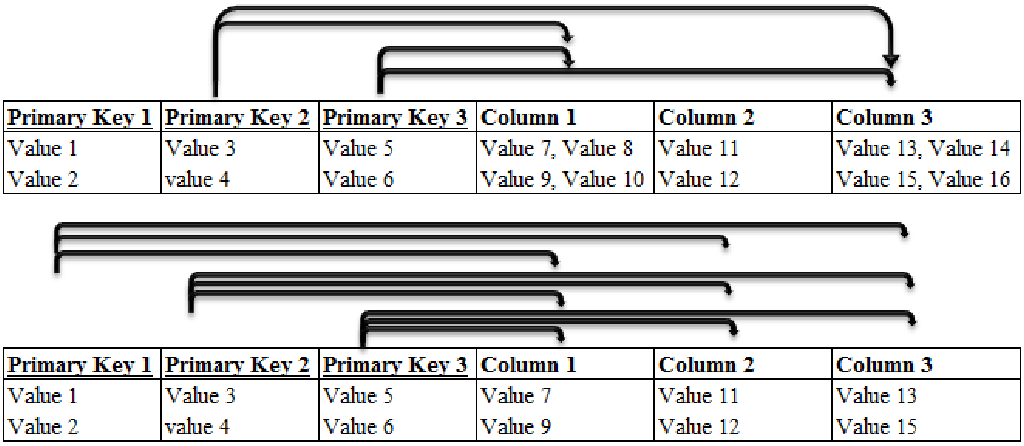
The dimension tables in the tar schema could contain for example master data which is used in the fact table. Another reason for using multiple tables is due to the performance of the queries. The dimension tables could be used as a search condition in queries. The dimension tables are usually smaller than fact table, and therefore, it makes sense to use dimension tables as a search conditions rather than going through the entire fact table.

In the star schema model, the dimension tables contain data repetition. Data redundancy can be reduced by modifying dimension tables into a third normalization form. When dimension tables are normalized, the model is so called snowflake schema. (Hovi, Ylinen & Koistinen 2001, p. 100) By normalizing the tables, data redundancy can be minimized but more over, updating will become more effective since data needs to be updated only once.

A simple approach for normalizing tables into the third normalized form is introduced next. The figure five demonstrates the normalization as well in action. Steps for forming a normalized table is listed below:

- Eliminate repeating groups from columns.
- Eliminate columns that contain multivalues.
- All columns of the table needs to be functionally dependent on the primary key, even if the primary key is a combined primary key.

- All columns of the table needs to be functionally dependent only on the primary key. (Hovi, Huotari & Lahdenmäki 2005, p. 87–94)



Primary Key 1	Primary Key 2	Primary Key 3	Column 1	Column 2	Column 3
Value 1	Value 3	Value 5	Value 7, Value 8	Value 11	Value 13, Value 14
Value 2	value 4	Value 6	Value 9, Value 10	Value 12	Value 15, Value 16

Primary Key 1	Primary Key 2	Primary Key 3	Column 1	Column 2	Column 3
Value 1	Value 3	Value 5	Value 7	Value 11	Value 13
Value 2	value 4	Value 6	Value 9	Value 12	Value 15

Figure 5. Modifying table into the third normalized form

In the figure five the first table contains column 1 and column 3, which includes multivalues. Arrows pictures dependencies between different rows. The first table format has combined primary key but all of the columns are not functionally dependent on the primary key. The table below has no multivalues and all columns are functionally dependent to the primary keys. Therefore, data redundancy is reduced as well as updating the values in the future is easier.

Data architectures

A successful data management strategy takes into account scalability and cost effectiveness. New database architectural approaches are needed when scaling data volumes (Lopez 2012, p. 18).

Different database architectures need to be investigated in order to evaluate which of them fulfils the database system demands which are scalability, performance, reliability, usability and data privacy. The architectures which have been investigated are centralized, partitioned and replicated models. In the centralized database architecture all data is stored and controlled by the centralized database servers. In the partitioned database architecture, data processing is distributed and computers have peer relationships. The database can either be partitioned horizontally or vertically. The replicated database architecture has basically copied parts of the database into other locations. All database architectures have their advantages and disadvantages. Portioning architecture provides the best performance, whereas peer-to-peer replication offers enhanced data reliability. (Chen et al. 2012, p. 1525–1527, 1545) Therefore, in order to ensure scalability, one should either use partitioning or replicated database architecture. Data privacy on the other hand can be ensured by using different database features.

To understand the database management system design, different kind of approaches have been built to perceive a better understanding of the architecture of the database systems. A bottom-up approach is used when the database management system is evolved bottom-up (Bose 2006, p. 47). The figure six demonstrates the nature of the bottom-up built database. This approach is commonly used when separate data mart solutions are designed. The data mart is a database solution which is built only for a part of the organization rather than for covering the whole corporation needs (Granlund & Malmi 2004, p. 40).

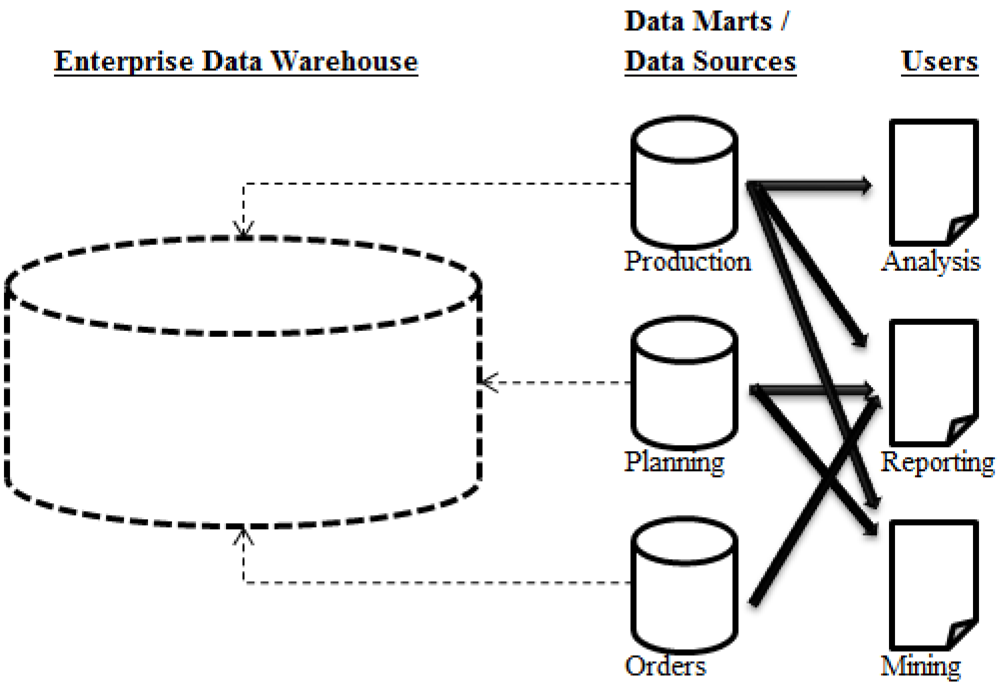


Figure 6. The architecture of the bottom-up approach database system (formed from Bose 2006, p. 46–47)

In the figure six, the data is loaded from the data marts into the enterprise warehouse. Users get the needed information also directly from the source systems. This approach is common for situations when data marts are built before the enterprise warehouse. Therefore, the development phase doesn't usually take long but on the other hand it might bring some problems up in the future if data marts are not developed in a standardized model. Best solution would be to model both enterprise data warehouse and data marts in parallel. (Bose 2006, p. 47) Enterprise data warehouse is dashed in the figure six because it doesn't exist yet and only the data marts are being developed in the first place. A multi-tier warehouse solution is the best architecture, in which both the data marts and the enterprise data warehouse are built in parallel but because it is more difficult to manage and build many organizations start with data marts only (Bose 2006, p. 47). As a conclusion could be said that in order to create the database system, the organization should start from a single data mart solution but develop it in a way that enables further data integration in the future.

Li et al. (2006, p. 244) introduced another architectural approach for designing the relational database management system. This approach can be used in parallel with the bottom-up approach. The main idea in this approach is that the database is just one part of the DBMS. The figure seven demonstrates the system layout in three tiers.

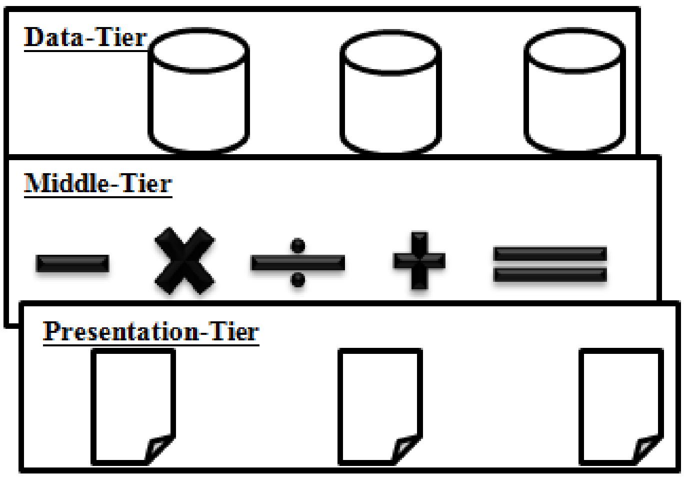


Figure 7. The three-tier system design approach (formed from Li et al. 2006, p. 244)

A data-tier includes databases, both operational and historical ones. Old data, which isn't anymore in the focus of surveys will be transformed to a history database. This procedure minimizes the load on operational databases and thereby enhances query performance (Hovi, Ylinen & Koistinen 2001, p. 87). A middle-tier includes an option for input, edit and delete data which is stored in the data-tier. The middle-tier contains also all the functionality like evaluation, reasoning and forecasting (Li et al. 2006, p. 244). In other words, the middle-tier works as an interaction between the users and the databases. To make the system easy-to-use, a presentation-tier is built, which contains the user interfaces for both the front and the end users (Li et al. 2006, p. 244, 247). In the fourth section of this thesis, a RDBMS is designed, and elements from both the bottom-up approach and the three-tier approach are utilized in order to create a system that is adequate for reporting purposes.

4 CASE STUDY: DESIGNING A RELATIONAL DATABASE MANAGEMENT SYSTEM

4.1 Company introduction: Tetra Pak Production Oy

Factory of the Tetra Pak Production Oy locates in Imatra and is a part of Tetra Laval Group. Tetra Pak produces packaging for the food industry. Tetra Pak Production Oy press the carbon with customer-specific and product-specific designs. These designs contain different colors and pictures. Strict hygiene requirements set own demands for the production as well. Quality management can ensure the fulfillment of these hygiene requirements.

Production starts with pressing the carbon with certain designs. Then the carbon is cut to the right shape and after this a robot piles blankets on skids. These skids are stored in the work-in-process (WIP) -area for a while. From the WIP -area, the skids continues to the side sealer where blankets are transferred from the skids to the machine. A certain order may contain several skids. Therefore production workers at side sealer need to be aware of the amount of the skids produced by printing press or otherwise they could accidentally start another order even though they haven’t finished the previous one yet. This can lead to increased installation costs because different designs may need other machine adjustments. In the side sealer the blankets are sealed and in the end sealed blankets are piled on pallets.

Both the side sealer and the printing press contain a quality check camera which rejects blankets that doesn’t fulfill quality demands. In addition to the camera’s quality monitoring some quality checks are performed by the production workers as well. The high quality of products is ensured by multiple quality checks which are conducted at certain time frame. The quality checks should be performed also when certain events occur, for example when a new production design is adjusted.

4.2 Methods used for collecting the empirical data

So far Tetra Pak Production Oy has used excel spreadsheets in order to monitor their performance. This excel file contains a lot of information about the current situation of reporting and data management. Exploring the current data management tool (excel spreadsheets) enables the better understanding of the current situation of data collection and storing but as well identifying the weaknesses of it. Besides of that, the user requirements have been acquired by interviewing both end and front users. Basically, the front users consists of manufacturer workers and the end users consists of both operational and executive managers. Several meetings with the management team enabled identifying both the short and the long-term objectives.

When the user demands have been identified, it is important to see how the current system works in practice. Therefore, production processes have been explored by spending time on the production line for a while. This has given a deeper understanding of the internal processes but the value chain of data processing has been investigated as well.

Tetra Pak's other factory in Sweden, in Sunne, is using a single software solution called FlexNet for monitoring and controlling internal processes. Therefore, a trip has been managed to see how their information system works and if it is possible to make use of their knowledge and apply some of their features in to the system which is being developed in Imatra. After the trip, still a close connection is maintained with Sunne's IT-manager in order to get the needed support for certain IT-solutions.

Some of the needed and valuable data is trapped in to the machine's programmable logic control (PLC) -systems. Fortunately, Tetra Pak Production Oy has another ongoing database project in order to reach this data from PLC -systems. By acquiring this data into the new system enables programming the needed functionality in order to secure certain user demands. Several e-mail conversations have been arranged with PLC's project manager in order to let their team know of Tetra Pak Production Oy's data requirements and vice versa.

4.3 The current state of reporting and data management

Tetra Pak Production Oy arranges meetings every morning where management team go through production reports. In these meetings managers need information to support operational decision-making. At the moment, data is collected, processed and demonstrated in excel spreadsheets. There is a slight chance for mistakes already in the data entry phase. Therefore, information is not accurate and reliable enough. Another problem is the performance of the system. Calculations and data analyzing is not cost effective enough. Because of this, real-time information is not available.

Tetra Pak Production Oy don't have information system which would enable effective data management. There is no responsible person in the organization who would manage data as a fulltime job. This makes it hard to do any kind of data mining to identify cost saving and operation improvement opportunities. At the moment, reporting management neither enables proactive actions nor continuous monitoring. Therefore, managers are only able to reflect on what has already happened.

Tetra Pak Production Oy is measuring production's performance based on the KPI's which are standardized across Tetra Pak factories all over the world. These measures need to be included into the new system as well. The most important indicators are equipment efficiency and waste.

Sweden's factory in Sunne is using the system called FlexNet as the data management tool. FlexNet provides an accurate and a visible information about operations. FlexNet has some other features as well but it can be used as a benchmark for Imatra's development process. In order to identify the advantages of FlexNet, a trip was made to Sunne. Their IT Manager introduced the functionality of FlexNet and gave some essential guidance in order to succeed in the system development process. FlexNet fulfils the demands of both data management and reporting management.

4.4 The specifications of reporting and data management requirements

Before starting the system development, some key points could be highlighted as a reminder regarding database projects. First of all, one needs to ensure to be aware of the total data picture and make sure that the system fulfills the needs of data management requirements. To be absolutely sure of the suitability, one should start the analysis from the output side of the picture. The database development project will take its time and effort, and to reduce pitfalls at later phases, it is important to pay maximum attention at the analysis phase. Once the analysis of the operational system is done and development may start, system should be built in enough flexible manner in order to be able to handle even the complex calculations and data manipulations. (Stock 2011, p. 315) For sure, these points needs to be taken into account in the data management strategy as well.

The information provided by the new system needs to be accurate and available at any time. Data management tool needs to be developed to support valid and consistent reporting as well as operational decision-making. The formed reports should consist important KPI’s which are appropriate for the operational management purposes of production. Nevertheless, users need to be able to create dynamic reports as well, whereas user demands vary across the time.

Features identified in the literature review are applied for the case of Tetra Pak Production Oy. As literature has shown, the system should fulfill long-term application performance, scalability, availability, usability, reliability and data privacy. To meet the demands of the users as well as system design requirements, a data management strategy needs to be created. According to Lopez (2012, p. 17) the data management strategy consists of three steps. These steps have been handled with a great interest in the section 3.4 of this paper. The table three demonstrates the data management strategy that is created for Tetra Pak Production Oy.

Table 3. Tetra Pak Production Oy’s data management strategy

Step of the data management strategy	How is this step considered in the case
1. Identify users’ data requirements	End users are interviewed in order to identify the used key performance indicators in the organization. Literature review supports also in choosing the appropriate key performance indicators for operational management purposes.
2. Build a data model	A relational data model is built by using star schema model.
3. Choose a data integration tool	Different database architectures and platforms are evaluated. Three-tier and bottom-up approaches are utilized in the integration tool decision.

Relational data model which is built for Tetra Pak Production Oy and the decision of the integration tool are introduced in the next section of this paper because the system design is being considered with greater interest later. This section reveals the results of the survey regarding data requirements. The data management and the reporting management requirements of Tetra Pak Production Oy are presented in the figure eight. Requirements which were especially highlighted by the users are related to the usability and the availability aspects.

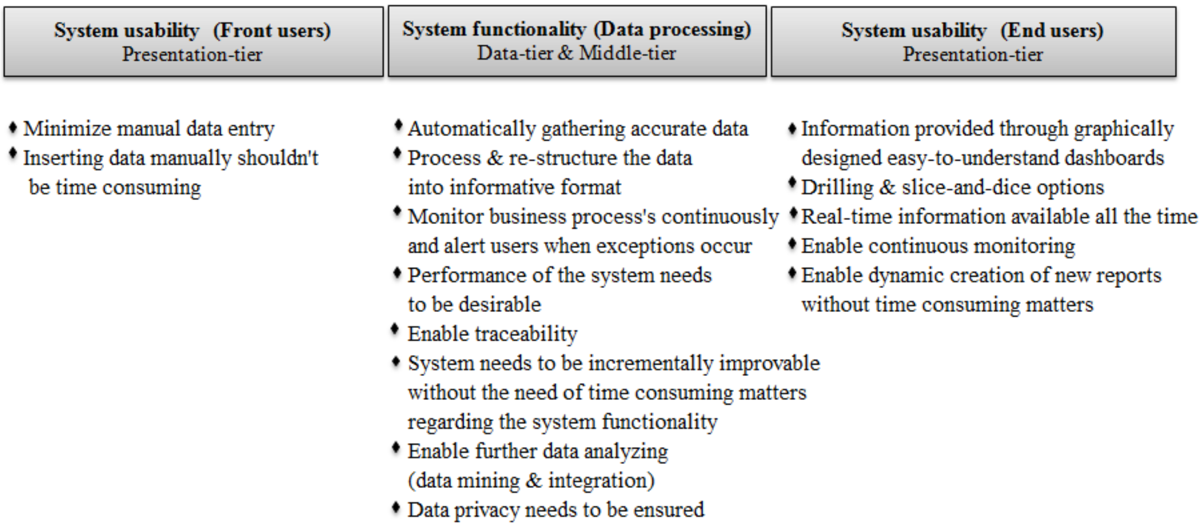


Figure 8. The data management and reporting management requirements of Tetra Pak Production Oy

Whether the company decides to purchase a complete business intelligence tool or build a system of their own, the process should always start from identifying the needs. System requirements can be identified for example by interviewing the managers and other employees whose demands the new system will be serving. (Granlund & Malmi 2004, p. 133) In this case, the user demands have been identified by interviewing the managers and employees.

According to the managers and the employees of Tetra Pak Production Oy, the data management strategy should help in achieving the objectives of the business strategy. Managers are using certain KPI’s in the operational performance measurement monitoring of production. These KPI’s are equipment efficiency and waste. Equipment efficiency provides information about the efficiency of the production by measuring the time of the machine. Waste instead measures the quality of the process. Waste is formed when the product doesn’t fulfil the quality demands.

Based on already available information and literature review, the KPI’s used for reporting are chosen. The literature review highlights the importance of manufacturing effectiveness and efficiency. Most commonly measured dimensions are time, quality, flexibility and cost. Gomes, Yasin & Lisboa (2007, p. 341) has introduced operational performance measure which suits best for Tetra

Pak Production Oy's continuous monitoring purposes. The indicator is called as Manufacturing Operational Effectiveness (MOE) which consists of efficiency, availability, and quality aspects.

$$MOE = A \times Q \times E = \left(\frac{t_a - t}{T_a}\right) \times \left(\frac{QP_c}{QP_c + QP_{nc}}\right) \times \left(\frac{P_e}{P_p}\right) \quad (1)$$

A = Availability,

Q = Quality,

E = Efficiency,

t_a= Available manufacturing time,

t_s= Time when all manufacturing processes are stopped,

QP_c= Quantity of conforming manufactured products,

QP_{nc}= Quantity of non-conforming manufactured products,

P_e= Quantity of manufactured products delivered to clients,

P_p= Quantity of planned products to be manufactured (Gomes, Yasin & Lisboa 2007, p. 341).

The formula one presents how the MOE indicator can be calculated. As already mentioned, the managers of Tetra Pak Production Oy are interested of measuring especially waste and equipment efficiency. In the formula of MOE, quality represents waste and efficiency represents equipment efficiency. Managers have not been measuring availability aspect before but taking into account the nature of the business, availability is an important aspect which should be measured. Other indicators have been introduced also in the literature review part but not all of them fit for continuous monitoring but rather for one-time calculation purposes.

Because Tetra Pak Production Oy is following strict standards regarding which KPI's are necessary, MOE will be modified to corresponding Tetra Pak production Oy's needs. In other words, availability, quality, and efficiency aspects are being monitored but with slight modifications.

Availability aspect is considered by comparing available capacity to production orders and forecasts. Also estimated production will be calculated based on the actual quantity of production. These three indicators will be presented as a cumulative numbers in the same graphic, and thereby managers are able to make decisions based on availability. Though, if managers want to examine production volumes with a greater interest, they can drill down into each designs produced on a daily basis level.

$$Waste\text{-}\% = \left(\frac{QP_{nc}}{QP_c + QP_{nc}}\right) + T_w\text{-}\% \qquad (2)$$

QP_c= Quantity of conforming manufactured products,

QP_{nc}= Quantity of non-conforming manufactured products,

T_w-% = Trim waste percentage (Tetra Pak Production Oy 2014).

The formula two presents the components which are taken into account when calculating waste. Whenever waste occurs, the operators enter the information into the system. The information includes amount of the waste, the part of the process where waste has been produced and the reason for the event. This enables drilling down into the processes which produces the most waste.

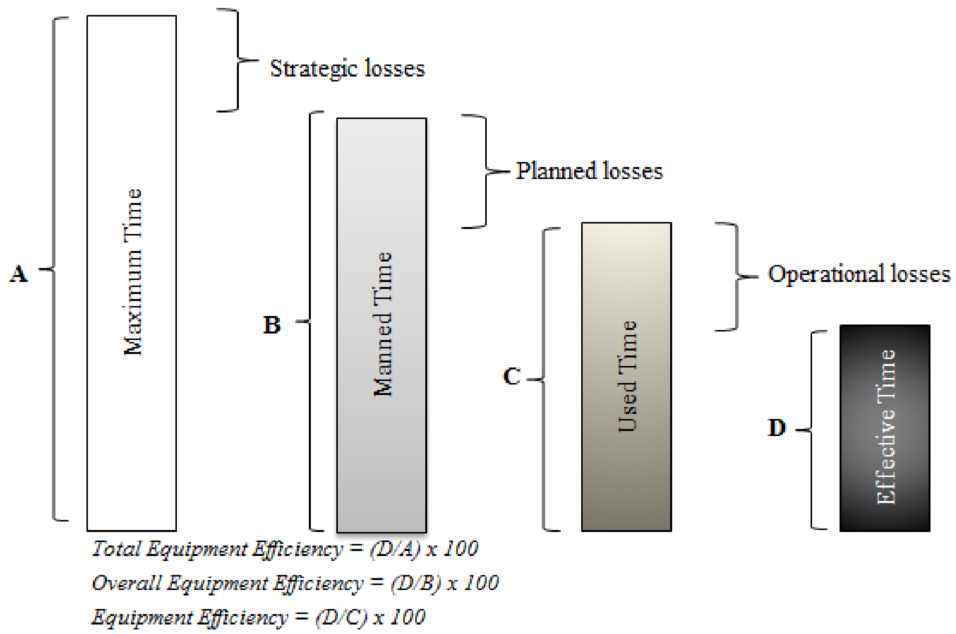


Figure 9. Total equipment efficiency, overall equipment efficiency and equipment efficiency based on time (formed from Tetra Pak Production Oy 2014)

The figure nine demonstrates how total equipment efficiency, overall equipment efficiency and equipment efficiency consists based on time. First of all, strategic losses are reduced from the maximum available time remaining the manned time. The strategic losses consists of legal restrictions, religious days, bottlenecks and lack of market demand. Planned loss in other hand consists of planned activities utilizing time to other matters than production, such as education and maintenance. Because the aim of this case is to create a system for operational monitoring, equipment efficiency is suitable for this purpose since it measures operational efficiency. Operational losses consists of availability, performance and quality aspects.

$$EE = A_r \times P_r \times Q_r = \left(\frac{t_{run}}{t_{used}} \times \frac{v_{run}}{v_{mec}} \times \frac{l_{po} - l_w}{l_{po}}\right) \times 100 \tag{3}$$

EE = Equipment Efficiency,

A_r = Availability rate,

P_r = Performance rate,

Q_r = Quality rate,

t_{run} = Run time,

t_{used} = Used time,

v_{run} = Run speed,

v_{mec} = Mechanical speed,

l_{po} = Meters into production orders,

l_w = Production waste in meters (Tetra Pak Production Oy 2014).

Equipment efficiency is calculated by using the formula three. This formula has been customized by Tetra Pak Production Oy for their needs. Equipment efficiency enables drilling-down option as well, when the data is captured with the most accurate level of detail. For example overall equipment efficiency, total equipment efficiency can be calculated but managers can as well drill-down and analyze which are the reasons causing the greatest lack in efficiency. These features set their own requirements for the front user's interface and system functionality.

Used indicators need to be considered carefully since they might have an impact on the psychological behavior of production workers. The end of section 2.1 examined how chosen performance measures affects on employees. This thought can be applied into Tetra Pak Production Oy's circumstances as well. For example, when printing press is starting a new design, the machine and new colors need to be adjusted. The setup can be run with different machine speeds. And it is clear that a faster machine speed correlates with a better performance of KPI's if the corporation is measuring machine's efficiency by time. Nevertheless, faster run speeds cause more waste since adjustments don't catch the quality standards if the machine setup is being run with too fast machine speed. If the employees are being controlled with equipment efficiency, they try to run the setup as fast as they can in order to look good on these measures. On the other hand, if managers wouldn't measure the time of the machine and instead highlight the importance of waste and quality, perhaps the employees would think smarter and try to find an optimal setup run speed.

Nevertheless, equipment efficiency is chosen as a performance measure but in parallel with waste and availability. Anyway, formula three takes into account quality rate Q_r which is why the value of EE is effected by quality in addition to time and speed. Also used machine speeds are monitored since the new system enables of capturing more accurate data regarding machine events.

4.5 System design

The database management system consists of the data-tier, the middle-tier, and the presentation-tier which have been introduced in the section 3.4 of this thesis. The database which locates in the data-tier provides a platform for the data. The data needs to be processed into easy-to-understand formats of dashboards and management reports. These reports and data input interfaces locate in the presentation-tier. SQL is used for interaction between data-tier and presentation-tier. The middle-tier consists of user-defined functions and other features which enables efficient data analyzing and monitoring.

SQL Server is used as a database platform in the case of Tetra Pak Production Oy because it supports indexes and partitioning of tables. Then scalability can be achieved with a good performance as well. SQL Server supports also triggers and user-defined functions, whereby the functionality of the middle-tier can be confirmed. Microsoft Access is also utilized because it supports usability via data entry forms and reports. From system designs point of view, the data-tier and the middle-tier are located and controlled in SQL server, whereas the presentation-tier of front users is managed with Microsoft Access. End users reports are still generated with Excel. The middle-tier uses structured query language to provide information from the data-tier to the presentation-tier. Then visual basic is used for generating interfaces, reports and dashboards in the presentation-tier. The ODBC connections are used to maintain reliability and availability between data systems components. The figure ten represents the system design.

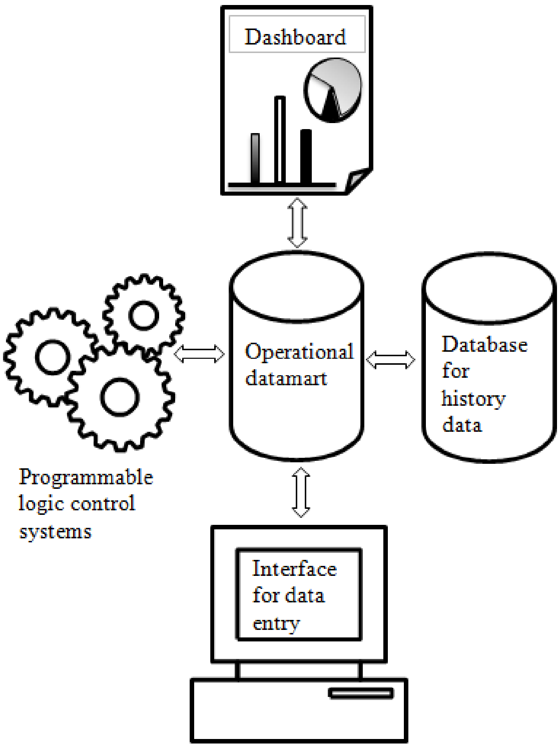


Figure 10. The structure of the database management system design

The system as a whole serves both the end and the front user's requirements. The arrows in the figure ten demonstrates the ODBC connections between different system components. The operational data mart is located in the middle of the picture and is surrounded by other features enabling a desired wholeness for the monitoring and the controlling purposes of production processes.

As already mentioned, relational database locates in the server and other applications are connected to this server in order to enable users execute queries and work with the server in user friendly format. The type of the connection needs to be considered carefully since it might have an impact on the performance. The ODBC is a standardized data integration technique which enables connecting to any RDBMS that has an ODBC driver (McMurtry et al. 2013, p. 70). Despite the age of the ODBC technique, it is used for connecting different applications and systems surrounding the database, and tests proved an excellent performance of the queries. The ODBC has been built by using a file data source name (DSN), and thereby any user who has access to MS Access file can view the contents. A user specific DSN would have enabled better data privacy but on the other hand the user DSN is not as flexible as the file DSN. The file DSN connection can be protected with a password common for all users, who shall access the file contents.

4.5.1 The data-tier

The bottom-up approach is utilized for designing the data-tier. Only one data mart is built in the first place but it is built in a way which allows further data integration in the future. Therefore, the system is built by using standards provided by world-class-manufacturing (WCM). This means that for example standardized defect codes are used and master data is formed in a structure that is incrementally improvable. The literature review proved that relational database fits for reporting purposes and meets the data management requirements and therefore the system is built on the relational data platform.

The database is built on already existing SQL Server and thereby saves costs. In the first place, the database needs to be planned in a way that enables the needed functionality for achieving the data requirements which have been identified earlier. This means that the data dictionary is built based on the arranged meetings with management team. After this, the data models can be built by using already introduced star schema method. By planning the structure of the database properly, unnecessary adjustments can be reduced in the later phases.

Data dictionary is presented in the appendix one. The data dictionary demonstrates the data value chain. By creating the data dictionary, the system developer can investigate which data items are utilized by different processes. For example state of order can be utilized in the side sealer when operators are deciding which order they should run next. Whenever a skid finishes printing press and is being transferred into WIP -area, the system refreshes the state of order. Then operators at side sealer are aware whether they can finish certain orders and move on to next orders or if there are still skids left in the WIP -area. Even more important, the system can be designed in a way that required data for the performance measurements will be stored and processed.

The next phase is to build the data models based on the data requirements and the identified data value chain. Since the data management tool is based on the relational two-dimensional tables, the relational data modelling can be utilized. The star schema and the snowflake schema have been introduced in the literature part. Both models could be used for Tetra Pak Production’s needs but since the managers want to store printing press data and side sealer data into separate tables, the data model is built according to the star schema approach. By storing the data into separate tables will enable easier master data management in the future but on the other hand it increases certain data redundancy. Nevertheless, the remainder of the tables are built in to the third normalization form, and therefore the system performance won’t suffer.

In the star schema approach, tables are connected into each other’s and the fact table is surrounded by multiple dimension tables. The fact table contains usually data from certain events. Because users insisted drilling-down options for used performance measures, the data needs to be collected in a more accurate level of details. Therefore, front users need to insert data into the system more frequently and to minimize the effort, machine events are being stored into a single fact table which is being supported by dimensional tables, such as quality target table, raw-material table, and production plan table. This enables the users to insert only reason codes whenever the machine events occur, such as produced waste and production line stops, and all other data is gathered automatically from dimension tables. The appendix two presents the built data models for side sealer. Data models are similar in printing press with only slight differences.

Advances in information technology has led to data warehouse solutions, which are database solutions built for pure analyzing purposes. This data warehouse is separate from the operational data mart and is used only for already cleaned and structured data. From time to time, structured and pre-processed data is loaded into the data warehouse where it is then analyzed. (Granlund & Malmi 2004, p. 40) Nevertheless, data transmissions can be reduced and no separate data warehouse solutions are needed if the data is analyzed straight in the operational data marts (Hovi, Huotari & Lahdenmäki 2005, p. 316). Therefore, features supporting reporting and analysis need to be inserted or built straight into the operational data mart solution. Taking into account the volume of the data and the number of the users in Tetra Pak Production Oy, a separate data warehouse solution is not needed. The performance of the queries is acceptable even though the capacity is used for operational purposes but as well for reporting purposes in parallel. Anyway, in order to enable continuous real-time monitoring, the queries need to be focused straight in to the operational databases rather than waiting for the data to be first transferred into separate data warehouse which might be done only once in a day. Nevertheless, the data warehouses can be useful when the analysis and the reports utilize data which is located in different data marts (Hovi, Huotari & Lahdenmäki 2005, p. 16). This way queries could return information regarding production, sales and finance for example.

To maintain the performance of the queries, historical data is transferred into a separate database after a while. This approach maintains the size of the operational data mart rather small and therefore the performance of the queries remain desirable (Hovi, Ylinen & Koistinen 2001, p. 87). The performance can be enhanced as well by aggregating the data regularly into summarized tables. The aggregated data is utilized continuously in reports and thereby the speed of queries can be accelerated by pre-processing the data into the correct form and then storing the pre-processed data into individual tables.

4.5.2 The middle-tier

The middle-tier of the RDBMS consists of the functionality part. This tier enables the required operational functionality in order to gather the data automatically without manual data entries and as well the data aggregation for data analyzing purposes. The database objects utilized for this part are PLC -systems, triggers, user defined functions and aggregation tables.

PLC data is utilized when machine events occur. The PLC -system stores the machine event data and sends the signal to the operational data mart based on either “on time” or on the combination of “on time and data change”. The “on time” saves data every 10 seconds as a default but the frequency can be changed, no matter if the data content is changed or not. The “on data change” saves data when the value has changed more than a configurable dead band, though the maximum frequency is 500 ms. This method enables for example storing the actual time when a production line stops or continues running. Therefore, the production workers don’t need to evaluate the time of the machine events anymore. Besides of storing data for machine events such as stops and runs, PLC stores data as well about the line speed. The speed is saved into the system based on the “on time” method. This enables calculating the average line speed but monitoring the used line speeds of the setup adjustments as well. Waste is a problematic machine event because the amount of waste can’t be stored into the operational system automatically by using the data provided by the PLC -systems.

This is because of the fact that whenever a production line starts detecting waste, the “on data change” signal will be sent to the system and the data will be saved but the waste will be produced for next seconds as well and then the system might get over loaded when signals are sent in every 500 ms. A solution for this problem could be achieved by building a logic into the PLC -system but Tetra Pak Production Oy don’t have this kind of logical manner at the moment, and therefore the amount of waste will be entered into the system manually by the production workers.

The PLC -systems enable achieving the better quality of the data since manual data entries are reduced. Other data management requirements can be fulfilled by utilizing relational database features, such as triggers, user defined functions, and data pre-processing into aggregated tables. Triggers for example are utilized when production plans are generated. When production planner imports plan file and order file into the system, a trigger will be launched and the production plan will be generated. Triggers are used as well for enabling the monitoring and the controlling of the WIP -area. When a skid is completed at the printing press, the system generates personalized skid id and a sequence of numbers used for a barcode. At this point, the state of the WIP view will be refreshed based on the trigger which was launched by completion of the skid. Now production workers at the side sealer are aware of the skids in the WIP -area belonging to certain orders. Triggers are used as well, when the skid is transferred from the WIP -area to the side sealer and therefore the state of skids in the WIP -area view remains up to date. The state of WIP -area is an important aspect to the system because it can reduce the time used for setup adjustments at the side sealer. This can be explained by using the following example which demonstrates the problem of the current situation. This problem is fixed in the new system by acquiring the real-time information regarding the state of the WIP -area.

> *Side sealer is running a certain order A, which is using design A1. Production workers think that no more skids belonging to order A exists anymore and they pick up the next order B, which is using design*

B1, from the production plan. At this point production workers need to make some adjustments for the machine. When production workers have already started running order B, someone discovers more skids belonging to order A. At this point, either more machine adjustments are done to correspond design A1, or remaining skids are left and order A will be delivered with a short amount of skids.

Another data management requirement which is traceability can be achieved with a proper designing. For sure, the SQL queries are needed as well but traceability can’t be achieved when the needed data is not taken into account when the system is designed and therefore the data has never been stored into the system. Importance of traceability can’t be emphasized enough. This is due to the fact that the food industry has high hygiene requirements. If already used raw-material contains stains, it would be essential to be able to trace which orders used this certain raw-material. For this kind of case, users can execute a query and find out which orders need to be pulled back. This kind of cases demand users to know how to use SQL queries but when they have the needed knowledge, generation of new reports or execution of such queries is easy and not time consuming at all. For more common issues, such as master data management purposes, a separate user interface can be built. Master data management is important due to the functionality of the whole system, and therefore incremental improvements need to be enabled without time consuming matters. User interfaces are handled with a greater detail in the next section.

The last data management requirement regarding the system functionality is the performance of the queries. This can be achieved by pre-processing the commonly used data into the aggregation tables. For example, the KPI’s can be automatically calculated into the specific table, and thereby continuous monitoring can be achieved with the reduced amount of effort and time. The aggregation tables enable an opportunity of further data analyzing purposes since the data mining techniques can be attached without major data cleansing actions. Data analyzing can be done without such data mining techniques as well. Business intelligence (BI) is a supportive process for decision-making, pursuing to analyze, refine, and present data which is aggregated from multiple sources (Peltola & Kaario 2008, p. 61). In this case BI is achieved with three-step process. First of all, data cleansing and aggregation is achieved with the functionality integrated into the operational data mart. Once the data is cleansed, it can be structured into easy-to-understand formats, which enable the monitoring and the controlling of the internal processes. The third part is only needed if managers want to drill-down into the deeper level of details, and at that point data is imported into a pivot table for further analyzing. It is obvious that data mining and business intelligence analysis can add value to decision-making. Nevertheless, such techniques are not substitutes for traditional cost management systems but rather a method for generating information from such dimensions that are not included into those cost management systems (Granlund & Malmi 2004, p. 115).

4.5.3 The presentation-tier

The presentation-tier of the RDBMS consists of the part which is visible for its users, both the front and the end users. With the user interface, the users are able to communicate with the server and execute queries whereas insert new data into the system as well. In this part of the thesis, the user interfaces and access management are handled.

The importance of usability has been highlighted multiple times in the literature review. Still, mainly end users' interfaces has gained greater interest than front users' interfaces. Despite of this, both the front users' and the end users' usability need to be ensured. From the end users point of view, it is important to create easy-to-understand dashboards which are visually effective. Therefore, it is important to create a tool which enables monitoring performance measures that are essential for the business. Nevertheless, the user requirements tend to vary across the time, and therefore, users should be able to create new reports without time consuming matters. For achieving the flexibility of the reporting tool, one could for example create reports with parametric values (Hovi, Ylinen & Koistinen 2001, p. 118). This way the users don't need to go through massive reports, and for example just give a certain time period which they are interested in. Another aspect that enables reports to be created from the fly, is the structured query language, which enables the creation of new queries and one-time reports, without the need to make any changes in the database structure. This of course demands knowledge from the SQL programming.

If reporting requirements are clear, no SQL programming studies are needed since these pre-constructed queries can be embedded into the user interface if the interface is connected into the SQL Server. This has been already discussed in the previous sections but as a reminder this can be achieved with the ODBC tools. In the case of Tetra Pak Production Oy’s, certain KPI’s were given, and therefore the dashboards could be created with embedded query opportunities. For example, in the morning meeting, if one wants to know what was the performance of previous days, no time consuming data cleansing and importing actions are needed anymore. At this kind of situation, the pre-defined KPI’s are calculated already in the system, and by executing the embedded query from the dashboard, fresh data is imported into the graphics.

As long as the database contains the required data, it can be imported from the SQL Server into the excel spreadsheets for further data analyzing purposes. In order to enable efficient analyzing, the data should be already re-structured and cleansed. This is achieved in the middle-tier part. For example, one can import a batch of data into pivot table, and go through it in there. This enables for example finding causalities between the operative actions and the performance.

To enable monitoring, the data needs to be in the database, and therefore, the interface needs to be built for the front users as well. By knowing, what information is needed for monitoring purposes, easy-to-use interfaces for data entering purposes can be designed. The KPI’s being monitored are waste, equipment efficiency and the availability of the capacity. In order to enable drill-down and slice-and-dice options, the data needs to captured with as accurate level as possible. Therefore, when ever machine events occur, operators need to enter the reasons for the events. The next example demonstrates how operators are interacting with the system through the interface.

> *When a machine stops, a signal is received from the PLC -systems and a trigger will be launched. This trigger creates a new machine event into the front user's interface and captures the timestamp as well. Thereby, when the machine stops, the production workers can concentrate on repairing and no more manually inserted timestamps are needed. When the production workers are done with the repairing actions, all they need to do is to give the reason for the stop event and the name of the machine's part where the event occurred. Thereby, the data is reliable since it does not rely on estimates any more but still managers can drill-down into the reasons which are causing the main disadvantages of the performance.*

The same principles applies to machine events such as waste and cases when the machine continues to operate. Quality check forms are filled in the user interface as well. In order to ensure usability of the system and to reduce the need of the manual data inputs, certain triggers and default values are used. For example, in order to monitor order number specific performance, it is required to capture the information of the orders which are being under manufacturing at certain times. To make it as easy as possible for the front users, the system assumes that whenever a machine event occurs, it belongs to the same order than the previous one. When order is completed, users enter new order number, and thereby they don't need to insert order numbers for every single machine event separately. The same principle applies to shift id's, and operator id's as well.

Access management enables ensuring that classified information maintains secret. This has been achieved in the system by creating user logins for different roles in the organization. These logins are able to view different information and these certain roles has limited privileges regarding inserting, deleting or updating data in the database. All parts of the RDBM are protected with passwords, and thereby data privacy is assured.

The print screens of the user interfaces and the digital dashboards are attached to the appendices 3–12. The user interfaces are in Finnish since all production workers don't speak English. The dashboards on the other hand are also in English because it is the used reporting language in Tetra Pak Production Oy. The front users' interfaces are generated in MS Access whereas end users' dashboards are created in excel.

4.6 The maintenance of the system

When business environment changes the system needs to be able to fit into new circumstances. This kind of situations should neither harm the performance of the system nor do harm for any other feature of the system. Changes are not only exceptions and therefore one can prepare the system already in the designing phase by enabling incremental improvements. From Tetra Pak Production Oy's point of view, it is important to enable adding new product designs and customers without time consuming matters. There might also occur changes in the production shift schedules. Another aspect that is highlighted in the literature review is the scalability aspect. The amount of data and users might increase across the time, and at that point the system needs to be scalable. Neither the performance is allowed to decrease nor the reliability of the results. Fortunately the relational data model allows such modifications without affecting results or the performance. Thereby, the system is incrementally improvable.

In the case of Tetra Pak production Oy a separate interface has been built for master data management purposes. In that interface users are able to make changes to master data content when changes occur. Because of the flexibility of the system, no actions are needed regarding the systems structure. In order to ensure that these changes will be made when these changes occur, a written rule is made regarding responsibilities. These rules contain the person who is responsible of certain database area, and as well guidelines for making such changes. The user interface contains easy-to-navigate feature since database areas are named according to these responsibilities, and separate tabs are made for those responsibilities. The appendix three demonstrates how the user interface of master data management is constructed.

Though, the system is scalable, it has not been tested with a huge amount of data. This can be done only at the point, when the system has been used for certain amount of time, and when the needed data is in the database. At that point, if the systems performance is not anymore desirable, more proper indexing can be added. Though, the system is already indexed, it can be enhanced at later phases. Actually Hovi, Huotila & Lahdenmäki (2005, p. 202) stated that indexing should not be done if it is not necessary. Therefore, indexes has been added to only those rows which are the most commonly used in the queries.

As already mentioned, indexing can be added after the implementation. Users will recognize when those indexes are needed because the performance of the query won't be as good as before. In this kind of situation, the structure of the tables need to be modified, but the framework defined by Hovi, Huotila & Lahdenmäki (2005, pp. 186–187) can be followed:

- The rows which are retrieved with the certain query, should be located next to each other in the table structure.
- The rows which are retrieved with the certain query, should be in the same order as it is in the query.
- The indexing is the best available when it contains all of the rows which are being retrieved. In that kind of situation, the system doesn't need to go through the table at all.

The employees of Tetra Pak Production Oy are educated for all situations mentioned. Also separate instructions are handed for them and these instructions contain the guides for modifying each part of the database. The whole system is documented as well for employees for possible future improvement purposes.

5 CONCLUSIONS

5.1 Designing the database management system

How should the database management system be designed in order to fulfill the demands of both data management and reporting management?

In order to build a desired wholeness to fulfill the requirements of data management and reporting management, the framework introduced in the figure eleven can be utilized in the system designing phase. The arrows demonstrate the process and the requirements. System designer should start from identifying appropriate KPI's to be monitored since they set requirements for reporting management and data management. The requirements put in place by chosen performance measures are pretty much the same than the ones required by reporting management in general. For example, in order to enable monitoring aspects such as time, quality, and flexibility demand highly sophisticated techniques to be utilized. In the case of Tetra Pak Production Oy monitoring of these dimensions required automation from the system in order to ensure simultaneously quality and availability of the information. Automation is achieved in the case study with the combination of the features provided by PLC -system and the functionality built into the SQL Server. Automation is also required since generally reporting management is requiring highly accurate real-time information for continuous monitoring purposes. Therefore, it can be stated that the system requirements depend on the general requirements of reporting management and on the nature of the chosen KPI's as well.

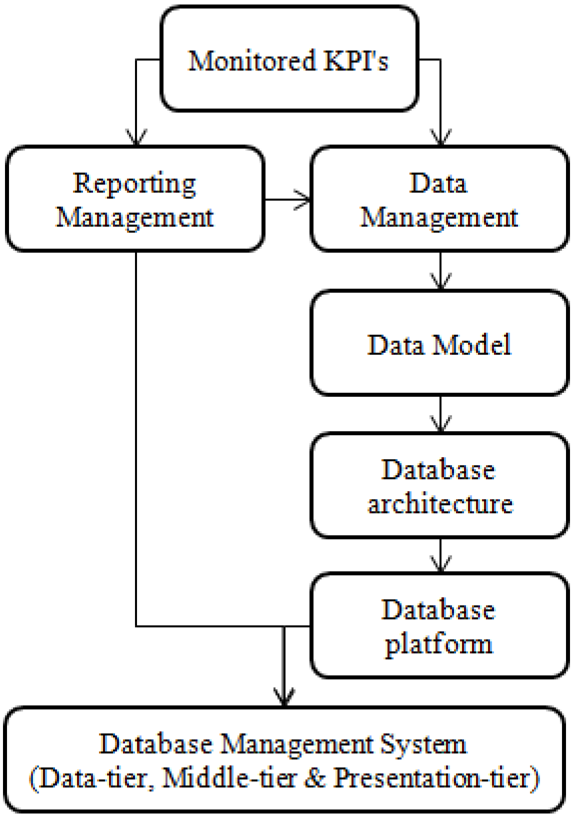


Figure 11. The framework constructed for designing the desired database management system taking into account the requirements of reporting management and data management

Reporting management and data management should be handled in parallel since reporting management is setting requirements for data management. To keep these demands on mind, one should identify both the requirements of reporting management and data management in general before building data models. These requirements were found from the literature but some of them were highlighted when employees and managers were interviewed as well. Even though, the requirements can be identified from the literature, it is highly recommended to involve key users in the designing phase. This decreases the change resistance but simultaneously allows the system developer to actually hear from the users how the system should work in action. On the other hand, by interacting closely with the key users, the current situation of the used system can be identified, and thereby the major problems of it can be identified as well. These pitfalls can be the result of logic rather than programming skills, and thereby the key users are the experts who may help avoiding these problems already in the designing phase.

When the current situation and the desired goal is clear, one can build a data dictionary to ensure that all parts of the system are taken into account.

Data models should be constructed after the system requirements are clear. Before modeling work, the appropriateness of the chosen data model should be evaluated in general. In the case of Tetra Pak Production Oy's, the relational data modeling was chosen since the relational data model fulfills the requirements of both reporting management and data management. The data model enables system developer to take all parts in to the consideration. Though, business processes and needs are unique, still the standardized data models can be utilized with slight modifications. In the case of Tetra Pak Production Oy, the data models were built from the scratch but these models can be later re-used in similar cases. By modeling the business process, one can ensure that everything is taken into account before the actual programming starts. This reduces the need of re-work in the programming phase.

Once the data models are created and the structure of the database is clear, including normalized and indexed tables, the database architectures can be evaluated. Three-tier and bottom-up approaches were utilized in parallel in the case of Tetra Pak Production Oy. The chosen architecture should fulfill the requirements of reporting management and data management. Different architectures have for example different capabilities regarding scalability and data privacy. Different database platforms can be examined once the database architecture is chosen. Database platforms have different features but in general most of them fit for every situation with a few expectations. The purpose of the study made in this paper was not to evaluate different database platforms but they differ in available features and price. For example triggers and stored procedures are not supported in MS Access, whereas in SQL Server both of them can be utilized.

Once the platform is chosen and licenses have been purchased, the real programming may start. A proper background work increases the chances to actually succeed in the database management system development process. Though, the requirements are clear and everything should have been taken into account, some problems might occur during the development phase. In order to stay in the schedule, these potential problems should be considered in the project plan by adding flexibility to the schedule. In the case of Tetra Pak Production Oy, a highly potential issue could have been the failure of integration of PLC -systems to the DBMS. Fortunately, everything went according to the plans and no major problems were faced. Nevertheless, some minor problems were faced in the logic part of the system. Though, the ETL-process was working, in the testing phase came out that the system was not working functionally in a desired way. This kind of problems are easy to correct but they are not so easy to detect since they don't appear as errors or warnings. To detect such problems in the logic, new system should be tested while the old system is still in use and then the results between them should be compared in order to make sure that the logic is right.

5.2 The operational reporting management of the production

What are appropriate key performance indicators in order to support operational decision-making in production?

One goal of the paper is to examine the appropriate KPI's for monitoring the operational performance of production. Though, the KPI's are chosen based on their ability to add information for decision-making, system developer should keep in mind the influence of the KPI's on the system and the requirements of the KPI's. Continuous monitoring and the nature of the KPI's set some requirements for the information system. The requirements of the system's functionality is especially highly affected by the chosen KPI's.

For monitoring the operational performance of production, such dimensions as time, quality, flexibility and cost should be taken into account. Manufacturing effectiveness and efficiency are highlighted as well. Several indicators were found from the literature which can be used for supporting operational decision-making in production. Nevertheless, managers need to evaluate appropriateness of each KPI individually by taking into account the nature of the production. Operational measures differ from the strategic ones by the fact that the aggregation level of the data is lower. Just like strategic KPI's as well operational ones should be based on the corporation strategy and support decision-making. Manufacturing Operational Effectiveness (MOE) indicator was modified in the case study to correspond Tetra Pak Production Oy's needs.

The literature review also commented the nature of appropriate KPI's. In order to encourage employees to think smarter rather than work harder, better performance may be achieved due to the fact that employees don't feel like being controlled and rather be in control. This can be achieved by keeping the nature of KPI's in mind and in the case of Tetra Pak Production Oy's chosen KPI's should encourage employees for example to find optimal circumstances to minimize produced waste during the machine adjustments. This can be achieved by highlighting the importance of quality and by measuring the used machine speeds during the machine adjustments. Thereby, external rewarding systems are not necessary to achieve desired performance because employees are pursuing enhanced performance by themselves. Carefully chosen measures will provide managers accurate value adding information regarding internal activities, and thereby enable the enhancement of actions affecting the performance of the production.

5.3 Fulfilling the requirements of data management and reporting management

What requirements do reporting management and data management have for database management system design, and does relational data model fulfill these demands?

The study revealed that the reporting management system has quite similar requirements than the information systems. This makes sense since it is recommended that the reporting management system should be integrated with the information systems. The figure 12 demonstrates data management and reporting management requirements in the form of data processing value chain and three-tier architecture approach. These two frameworks seem to have same phases and requirements. Some of these requirements are due to the reporting management and some of them due to the data management. In the figure, DM stands for data management and RM for reporting management. One can identify as well that some of the requirements fit for both reporting management and data management. The relational data model fulfills all of those requirements and therefore it can be used as a database platform when such systems are built. The main advantages of the relational data model are the high flexibility and the data independence. These features among others enables the system to be incrementally improvable and to fit into the new circumstances without time consuming matters regarding the system functionality.

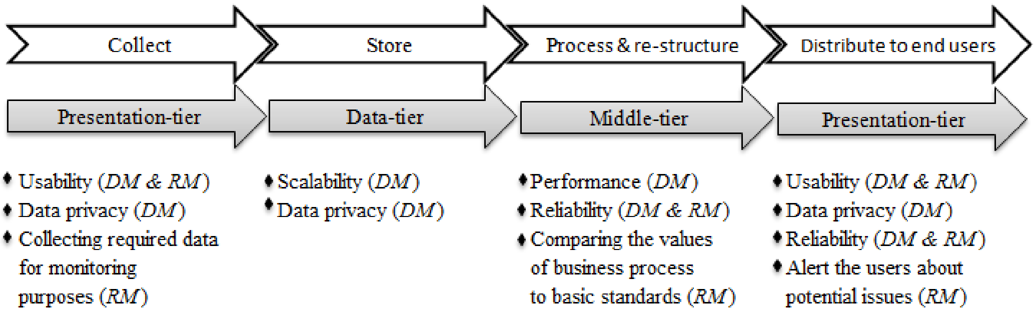


Figure 12. The requirements of data management and reporting management demonstrated in the form of data processing value chain and three-tier architecture approach

Scalability can be fulfilled rather easy if amount of users and data-flows are small. Nevertheless, the business environment might change or the system is on a big load already in the first place. Therefore, data models such as star- and snowflake-schemas should be utilized whereas different database architecture approaches as well. The combination of these techniques enable designing a desired system for such situations without the fear of weakened performance.

Though, the user requirements may be clear after the KPI's are chosen and the features of the system are desirable, system developer should simultaneously be aware of the future needs. Therefore, besides of the data required for continuous monitoring purposes, other needs should be considered as well. When the database project is implemented with data mart solution approach, the goal is to create predefined solutions for managers, such as dashboards for monitoring the performance of certain business area. Nevertheless, managers might think if these systems are capable of executing sophisticated data mining in the later phases. This may cause problems because information required regarding predictive forecasting for example is not available. Such data mining techniques can be integrated to the system in the later phases but the problem arises when the required data has not been stored in to the system because it was not essential due to the monitoring of predefined KPI's. Taking into consideration such data items is hard since the future interests may vary across the time, and once a certain business objective is achieved, the next one is waiting around the corner. Nevertheless, it is not recommended to include every single bit of data to the system but rather consider in which situations such data should be stored to the system. For example in the case of Tetra Pak Production Oy, PLC -systems enable collecting a huge amount of data in to the system but it was decided to collect and store data only once the value in sensors changes. Therefore, data is not stored absurdly every five seconds and rather when for example the machine stops.

5.4 The results of the empirical study

The purpose of creating the new system was to enhance reporting management. Reporting management has been assumed to be efficient when users are able to gain advantages from it. The created system brought several short-term advantages but it also enables pursuing some long-term advantages. The old system which was built on excel caused most problems on quality and availability. First of all, when front users were entering data to the old system, certain quality issues might have occurred since data was entered manually and some information was based on estimates. Secondly, end users didn't have real-time information available without time consuming data aggregation actions. These problems are solved by gathering accurate data automatically in the new system. After that the system processes and re-structures the data into informative format automatically. Thereby, real-time information is available all the time and both end and front users can save time.

With the old system, reports demanded manual data aggregation which took approximately 10–15 minutes. The new system provides real time information continuously and refreshing takes less than 10 seconds. The new user interface is clear and informative and thus easy-to-use according to the feedback given by front users. Production planner can as well save time when new production plans are imported to the system. With the old system, the activity took approximately 10–15 minutes where the new system is able to do it in less than 10 seconds. Implementing these plans to the production is also improved since with the old system, production workers had to wait 10 minutes when production plan changes. The performance of the new system is enhanced and no more waiting is needed.

Long-term advantages can be achieved as well since the new system is generating certain information that was not even available in the previous one. The new system enables pursuing information regarding causalities between the internal processes and the performance. Thereby, generated information is supporting operational decision-making but as well enables the enhancement of the internal processes. Such causalities could be found for example between the quality of the production and the used machine speeds while set up adjustments.

To improve the system created in the case of Tetra Pak Production Oy, one could for example create an integrated enterprise data warehouse for covering the whole corporation needs and then implement some data mining techniques such as predictive forecasting. Data mining could enable pursuing the information of the machine’s lifetime, and thereby prevent machine’s to crash by maintaining them predicatively. This requires that the data marts are built in a way which enables efficient data integration. When the bottom-up approach is utilized in the RDBMS creation, single data marts are built individually in the first place for covering single business unit needs. Therefore, when a data warehouse is created for fulfilling the whole corporation needs, those data marts need to be able to interact with each others.

Future propositions depend on the amount of available data resources as well. If there is enormous amount of available data, one could as well think of utilizing some commonly known big data handling techniques such as Hadoop.

6 SUMMARY

This paper has examined the data management requirements from the viewpoint of reporting management. Reporting management is understood as a continuous monitoring of the performance indicators. The study has been limited to handle only operational level data management and reporting management. Reporting management has been handled by taking into account only production whereas other parts of business have been left out of consideration. In this paper the appropriate KPI's have been examined, and the performance management system has been built to support these predefined measures.

In order to enable efficient reporting management, not only the end users should be considered but the front users as well. The end users are the managers who utilize the information provided by the system, whereas the front users enter the data into the system. Literature states that the performance measurement system should be integrated with the information system. Therefore, the whole DBMS should be created rather than building only a separate reporting tool. The figure 13 demonstrates the data management process. The DBMS covers the ETL-process of data management. When a desired system is created for collecting, storing and processing the data which is needed for monitoring purposes, then some separate data mining techniques can be integrated to find more value-adding information for achieving long-term goals, such as an enhanced performance of the internal processes.

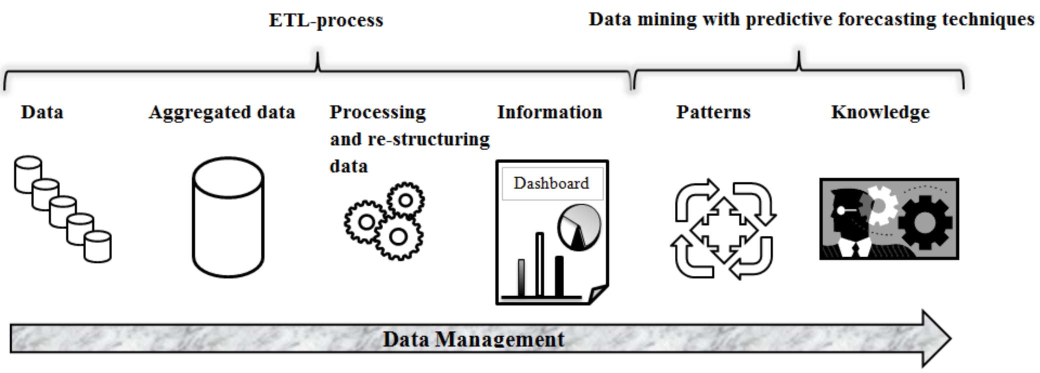


Figure 13. The data management process

The database architecture depends on the chosen data management strategy. The data management strategy should be formed when new database management systems are being developed. Building the data management strategy starts by identifying users' data requirements. In this paper, the users' demands have been seen from the reporting managements point of view. The next phase is to build the data model which supports the business needs. Star- and snowflake schemas have been introduced which can be utilized when the data models are created. The last step is to choose a right tool for the data integration. At this point, different database platforms and architectures should be evaluated. The database architectures of three-tier and bottom-up approaches have been introduced in this study and they could be utilized when designer wants to ensure the desired wholeness to fulfill the information system requirements. In order to make the right decision, one should examine the data management requirements and then make the decision whether the tool is appropriate or not. In general, the relational data model fulfills the information system requirements of performance, scalability, reliability, usability and data privacy.

The data model should imitate only the most critical aspect of the business processes. Therefore this paper has first examined the requirements of efficient reporting management. The study reveals that the reporting management system requirements have a lot of common aspects with the data management requirements. First of all, the system needs to be easy to implement, use and run, but moreover maintainable and incrementally improvable.

The elements of reporting management that needs to be fulfilled by the information system are:

1. Measuring the actual business process. Aspects of time, quality, flexibility and cost should be covered with the chosen KPI's monitoring the operational performance of production.
2. Comparing and evaluating the actual values of business process to basic standards.
3. Alerting the firm about potential problems.

The successful data management strategy takes into account scalability and cost effectiveness. Architectural database structures need to be examined when scaling data volumes. This paper has revealed that scalability can be achieved at any time of the systems lifetime but in order to do it in cost effective way it should be considered already in the database designing phase.

Commonly faced issues regarding data management are availability and reliability of the data. Data availability decreases when data is scattered into several systems and data integration is not managed in appropriate way. When creating the new database system, it is easier for the organization to start from a single data mart solution rather than building a completely integrated data solution. Nevertheless, developer should build the local data marts in a way that enables further data integration in the future.

Data reliability, in other words data quality and availability, can be assured with a proper system design, and moreover user interface design. Manual data input should be minimized since entering the data manually may cause mistakes. Because all data can’t be entered automatically, the system should check if the manually entered value is correct. This can be achieved for example by using check constraints which allow the user to enter only certain predefined values. The information fulfills the quality requirements when the information meets the user’s expectations. Nevertheless, user demands tend to vary across the time and it is time consuming to create new reports in excel spreadsheets. Therefore, the relational data structure and the data mining techniques should be used.

Other issues that database developer should consider are performance aspect, data privacy, and usability. The data management system should be suitable for a large amount of data and users. The purpose of the system is to add value to decision-making by offering valuable information to its users. In other words, the purpose of such a system is to collect, store, and process the data in a way which enable the creation of easy-to-understand dashboards and the continuous monitoring of performance.

In the empirical part of this paper, such a system has been designed and created for Tetra Pak Production Oy. The old system didn't fulfill demands which had been identified from the literature regarding information system requirements. Most problems were faced with availability and quality of the information. The new system has been built on SQL Server platform, but PLC -systems have been utilized as well. Front users' interfaces have been built to MS Access with the visual basic. End users' dashboards have been constructed to excel. Tetra Pak Production Oy is following the principles of lean manufacturing, and therefore the dimensions chosen for the monitoring of production are availability, quality and effectiveness.

According to the feedback given by front and end users, the new system fulfills the requirements of data management and reporting management. Most advantages have been gained from performance and usability aspects. Both front and end users can save time with the new system because accurate data is gathered and processed automatically. Real-time information is available as well and no more time-consuming data aggregation actions are needed. The system is able to fit into the new business circumstances as well because users can make changes to master data easily through the separate user interface. The new system enables also traceability, drilling-down, and slice-and-dice options. The biggest changes between the old and the new system are presented in the figure 14.

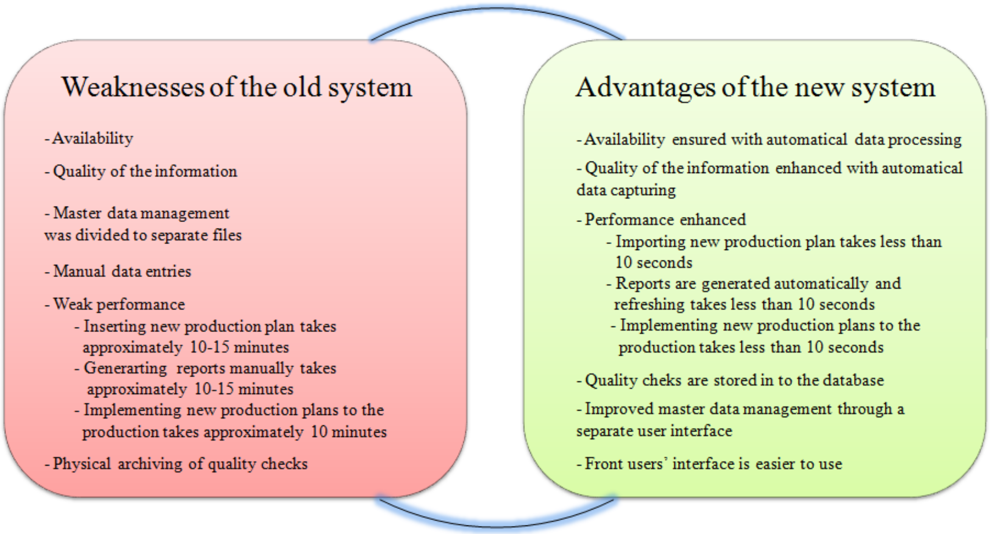


Figure 14. Changes between the old and the new system

In overall the relational database management system has contributed the operational performance measurement of production. To further improve the system created in this study, an integrated data warehouse could be created and data mining techniques could be implemented as well. Thereby long-term advantages could also be pursued by enhancing the performance of the internal processes since causalities could be found easily between the internal actions and the performance.

REFERENCES

Aggarwal, N., Kumar, A., Khatter, H. & Aggarwal, V. 2012. Analysis the effect of data mining techniques on database. *Advances in Engineering Software 47.* pp. 164–169.

Bayou, M.E. & de Korvin, A. 2008. Measuring the leanness of manufacturing systems – A case study of Ford Motor Company and General Motors. *Journal of Engineering Technology Management, vol. 25.* pp. 287–304.

Bitici, U.S., Mendibil, K., Nudurupati, S., Turner, T. & Garenzo, P. 2004. The interplay between performance measurement, organizational culture and management styles. *Measuring Business Excellence, vol. 8 No 3.* pp. 28–41.

Bitici, U., Garengo, P., Dörfler, V. & Nudurupati, S. 2012. Performance Measurement: Challenged for Tomorrow. *International Journal of Management Reviews, vol. 14.* pp. 305–327.

Bose, R. 2006. Understanding management data systems for enterprise performance management. *Industrial Management and Data Systems, vol. 106 No. 1.* pp. 43–59.

Bourne, M., Kennerley, M. & Franco-Santos, M. 2005. Managing through measures: a study of impact on performance. *Journal of Manufacturing Technology Management, vol. 16 No 4.* pp. 373–395.

Chandran, R. 2013. Big data management platforms: Architecting heterogeneous solutions. *Business Intelligence Journal, vol. 18 No 4.* pp. 32–38.

Chavez, R., Gimenez, C., Fynes, B., Wiengarten, F. & Yu, W. 2011. Internal lean practices and operational performance – The contingency perspective of industry clockspeed. *International Journal of Operations and Production Management, vol. 33 No. 5.* pp. 562–588.

Chen, S., Ng, A. & Greenfield, P. 2012. A performance evaluation of distributed database architectures. *Concurrency and Computation: Practices and Experience, vol. 25.* pp. 1524–1546.

Cocca, P. & Alberti, M. 2009. A framework to assess performance measurement systems in SMEs. *International Journal of Productivity, vol. 59 No 2.* pp. 186–200.

Demeter, K. 2013. Operating internationally – The impact on operational performance improvement. *International Journal of Production Economics, vol. 149.* pp. 172–182.

Ferreira, A. & Otley, D. 2009. The design and use of performance management system: An extended framework for analysis. *Management Accounting Research 20.* pp. 263–282.

Granlund, M & Malmi, T. 2004. Tietotekniikan mahdollisuudet taloushallinnon kehittämisessä. Jyväskylä: WSOY. 167 p. ISBN: 951-0-27703-7.

Gomes, C.F., Yasin, M.M. & Lisboa, J.V. 2007. The measurement of operational performance effectiveness: an innovative organisational approach. *International Journal of Business Innovation and Research, vol. 1 No. 4.* pp. 337–364.

Hicks, B.J. 2007. Lean information management: Understanding and eliminating waste. *International Journal of Information Management, vol. 27.* pp. 233–249.

Hovi, A. 2004. SQL-opas. Jyväskylä: WSOY. 280 p. ISBN: 941-846-228-3.

Hovi, A., Huotari, J. & Lahdenmäki, T. 2005. Tietokantojen suunnittelu & indeksointi. Porvoo: Docendo Finland Oy, WSOY. 353 p. ISBN: 951-846-262-3.

Hovi, A., Ylinen, J. & Koistinen, H. 2001. Tietovarastot liiketoiminnan tukena. Jyväskylä: Talentum Media Oy. 276 p. ISBN: 951-762-777-7.

Hudson, M., Smart, A. & Bourne, M. 2001. Theory and practice in SME performance measurement systems. *International Journal of Operations & Production Management, vol. 21 No. 8.* pp. 1096–1115.

Kaario, K. & Peltola, T. 2008. Tiedonhallinta. Avain tietotyön tuottavuuteen. Porvoo: WSOY. 164 p. ISBN: 978-951-0-34793-5.

Karim, A. & Arif-Uz-Zaman, K. 2013. A methodology for effective implementation of lean strategies and its performance evaluation in manufacturing organizations. *Business Process Management Journal, vol. 19 No. 1.* pp. 169–196.

Lee, S.J. & Siau, K. 2001. A review of data mining techniques. *Industrial Management & Data Systems, vol. 100 No. 1.* pp. 41–46.

Li, J., Moselhi, O. & Alkass, S. 2006. Internet-based database management system for project control. *Engineering, Construction and Architectural Management, vol. 13 No. 3.* pp. 242–253.

Lopez, J.A. 2012. Best Practices for Turning Big Data into Big Insights. *Business Intelligence Journal, vol. 17 No. 4.* pp. 17–21.

McMurtry, D., Oakley, A., Sharp, J., Subramanian, M. & Zhang, H. 2013. Data access for highly-scalable solutions: using SQL, NoSQL, and polyglot persistence. Microsoft. 244 p. ISBN: 978-1-62114-030-6.

Mancini, D., Vaassen, E.H.J. & Dametri, R.P. 2013. Accounting information systems for decision making. Springer. 349 p. ISBN 978-3-642-35760-2.

Niedritis, A., Niedrite, L. & Kozmina, N. 2011. Performance Measurement Framework with Formal Indicator Definitions. *Perspectives in Business Informatics Research, Lecture Notes in Business Information Processing, vol. 90.* pp. 44–58.

Phan, C.A. & Matsui, Y. 2010. Comparative study on the relationship between just-in-time production practices and operational performance in manufacturing plants. *Operational Management Research*. pp. 184–198.

Rantanen, H., Kulmala, H.I., Lönnqvist, A. & Kujansivu, P. 2007. Performance measurement systems in the Finnish public sector. *International Journal of Public Sector Management, Vol. 20 No. 5.* pp. 415–433.

Rausch, P., Sheta, A.F. & Ayesh, A. 2013. Advanced Information and Kowledge Processing. Theory, Systems and Industrial Applications. Business Intelligence and Performance Management. 269 p. ISBN 978-1-4471-3865-4

Rickards, R.C. & Ritsert, R. 2012. Data governance challenges facing controllers. *International Journal of Business, Accounting and Finance, vol. 6 No. 1.* pp. 25–42.

Robson, I. 2004. Implementing a performance measurement system capable of creating a culture of high performance. *International Journal of Productivity, vol. 52 No. 2.* pp. 137–145.

Scott, M., Boardman, R.P., Reed, P.A., Austin, T., Johnston, S.J., Takeda, K. & Cox, S.J. 2013. A framework for user driven data management. *Information Systems 42.* pp. 36–58.

Stock, T. 2011. Using a data warehouse to solve risk, performance, reporting and compliance-related issues. *Journal of Securities Operations and Custody, vol. 3 No. 4*. pp. 1753–1810.

Tetra Pak Production Oy. 2014. An internal WCM education material.

Ukko, J., Tenhunen, J. & Rantanen, H. 2007. Performance measurement impacts on management and leadership: Perspectives of management and employees. *International Journal of Production Economies.* pp. 39–51.

Vucetic, M., Hudec, M. & Vujosevic, M. 2012. A new method for computing fuzzy functional dependencies in relational database systems. *Expert Systems with Applications, vol 40.* pp. 2738-2745.

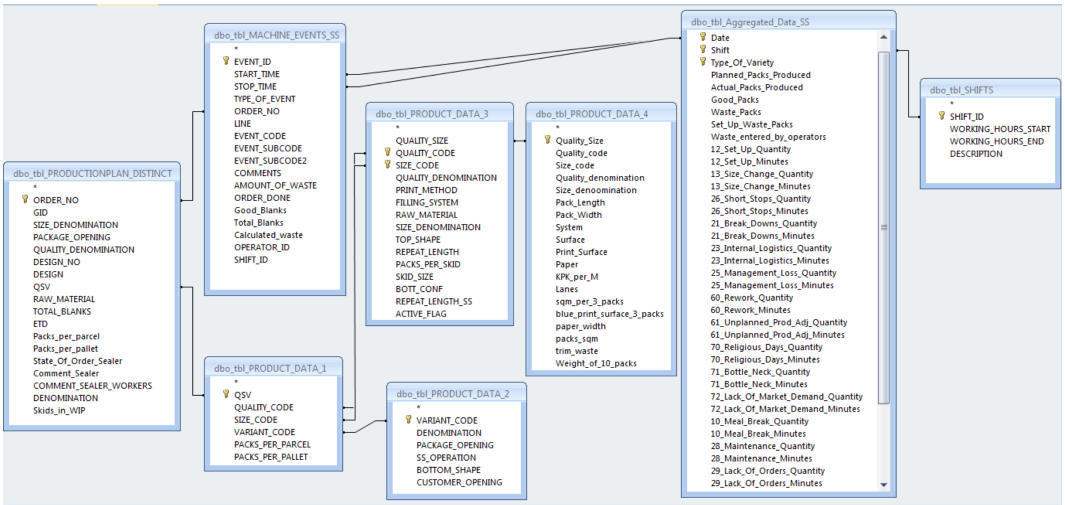
Wan, H.-d. & Chen, F.F. 2009. Decision support for lean practitioners: A web-based adaptive assessment approach. *Computers in Industry, vol. 60.* pp. 277–283.

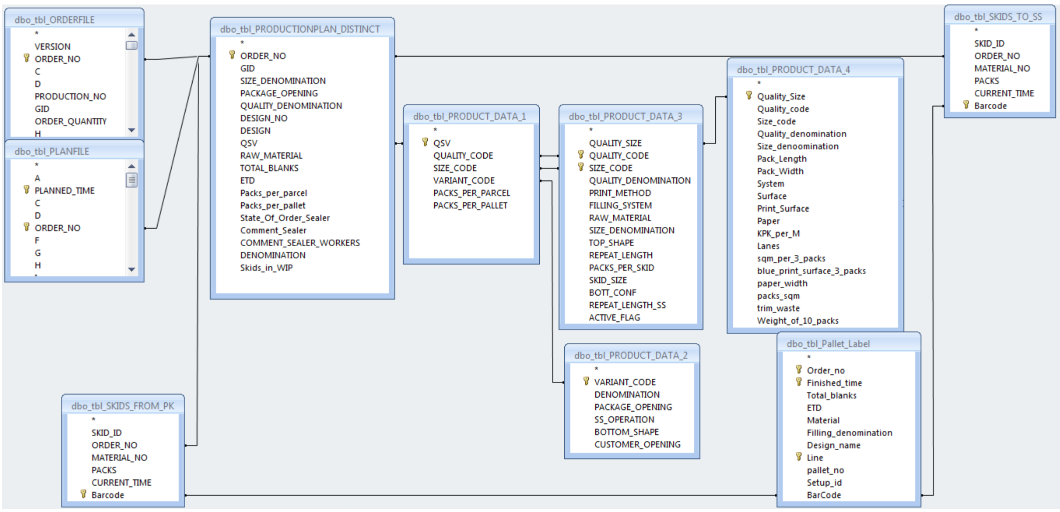
Xu, J. & Quaddus, M. 2013. Managing information systems: Ten essential topics. Paris: Atlantis Press. 166 p. ISBN: 978-94-91216-88-6.

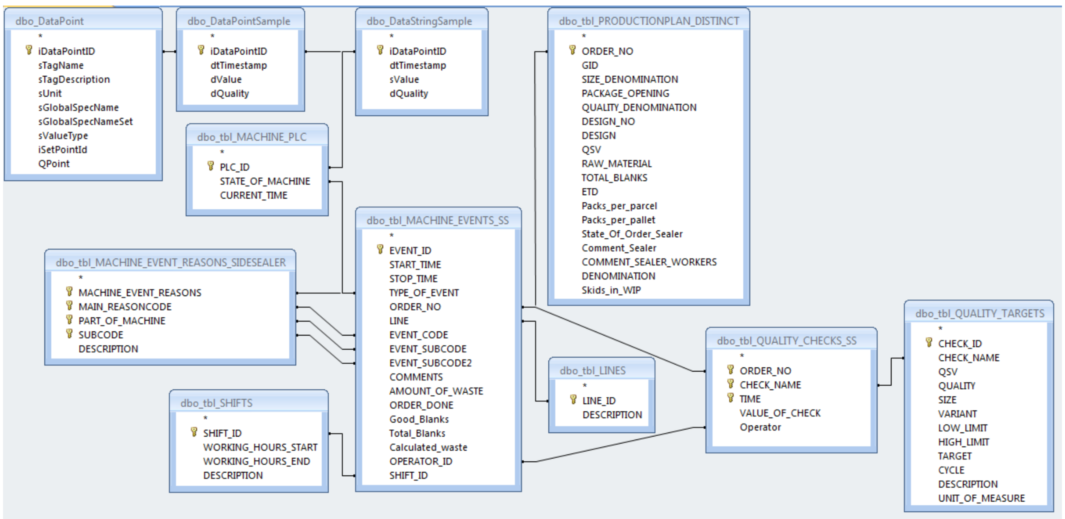
APPENDIX 1: Data Dictionary

Data Item \ Process	Production planner	Prepress	Printing press 1	Printing press 2	Printing press 3	Side sealer 1	Side sealer 2	Side sealer 3	Reporting manager
Order number	Insert		Use	Use	Use	Use	Use	Use	Use
Global order number	Insert							Use	Use
Size denomination	Insert		Use	Use	Use	Use	Use	Use	Use
Package opening denomination	Insert		USe	USe	Use	Use	Use	Use	Use
Quality denomination	Insert		Use	Use	Use	Use	Use	Use	Use
Design	Insert	Use	Use	Use	Use	Use	Use	Use	Use
Design name	Insert		Use	Use	Use	Use	Use	Use	Use
Quality - Size - Variation denomination	Insert		Use	Use	Use	Use	Use	Use	Use
Raw-material	Insert		Use	Use	Use	Use	Use	Use	Use
Order quantity	Insert		Use	Use	Use	Use	Use	Use	Use
Exchange-traded derivative	Insert		Use	Use	Use	Use	Use	Use	Use
Production line	Insert		Use	Use	Use	Use	Use	Use	Use
Production planners comments	Insert		Use	Use	Use	Use	Use	Use	Use
Operators comments	Insert		Use	Use	Use	Use	Use	Use	Use
Barcode of used raw-material roll	Insert		Use						Use
Used time per raw-material roll			Insert						Use
Order starting time			Use	Use	Insert				Use
Order ending time			Use	Use	Insert				Use
Produced waste				Insert			Insert		Use
Reason for the produced waste				Insert			Insert		Use
Time when waste occured				Insert			Insert		Use
Machine events (stops and runs)				Insert			Insert		Use
Time of machine events				Insert			Insert		Use
Reasons for machine events				Insert			Insert		Use
Quality tests				Insert			Insert		Use
Machine run speed				Insert			Inset		Use
Shift changes				Insert			Insert		Use
Skids transformed in work-in-process area					Insert	Use			
Skid transformed from work-in-process area						Insert / Use			
State of the order	Insert		Insert		Insert	Insert / Use			Use

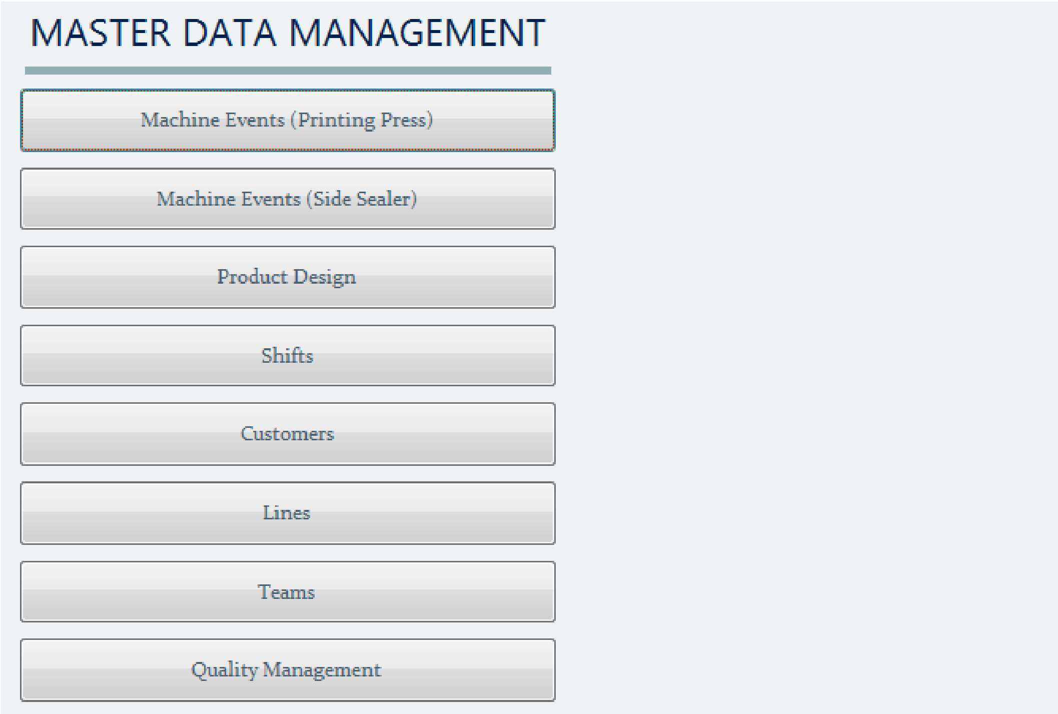
APPENDIX 2: Data models – Performance Data, Production Plan and Machine Events







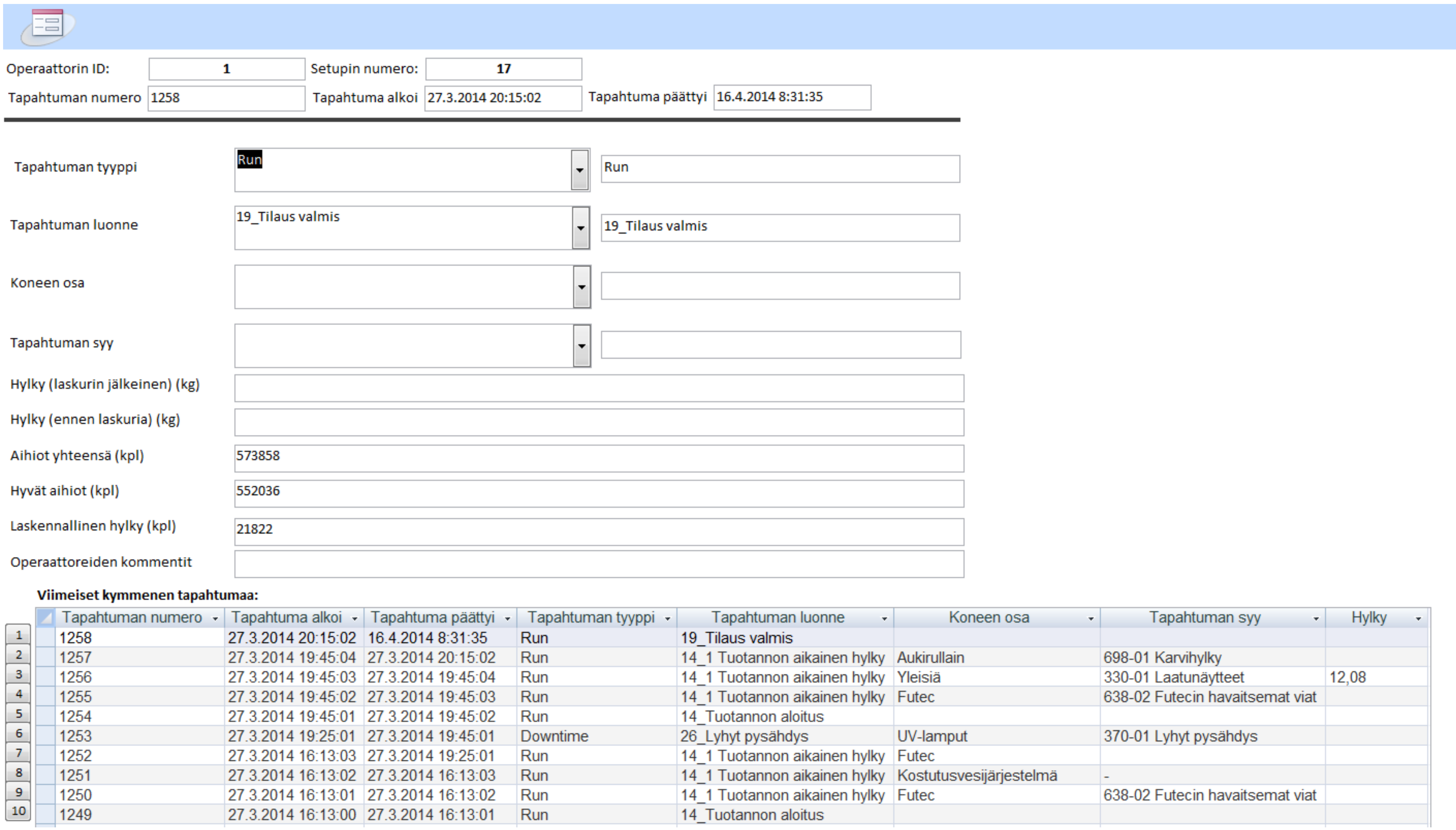
APPENDIX 3: User Interface – Master Data Management



APPENDIX 4: User Interface – Printing Press 1

ID	DAT E_AND_TIME:	SETUP	RULLAN TILAUSNUMERO	RULLA	SETUPILLE AJETUT RULLAT	AJETUT METRIT (m)	Koko (l)	UIB	MATERIAL_CODE
48	13.5.2014 12:48:00		K4FI-400049	50155	1	4075	1		24604000490155126603301M 0040
37	13.5.2014 11:44:00	4	K4FI-400049	50155	2	4092	1	30053143140201	24604000490155126603301M 0040
44	13.5.2014 11:44:00	4	K4FI-400049	50155	3	4092	1	30053143140201	24604000490155126603301M 0040
45	13.5.2014 11:44:00	4	K4FI-400049	50155	4	4092	1	30053143140201	24604000490155126603301M 0040
40	13.5.2014 11:44:00	4	K4FI-400049	50155	5	4088	1	30053143140201	24604000490155126603301M 0040
41	13.5.2014 11:44:00	4	K4FI-400049	50155	6	4087	1	30053143140201	24604000490155126603301M 0040
42	13.5.2014 11:44:00	4	K4FI-400049	50155	7	4092	1	30053143140201	24604000490155126603301M 0040
43	13.5.2014 11:44:00	4	K4FI-400049	50155	8	4077	1	30053143140201	24604000490155126603301M 0040
									Skannaa rulla tähän. Muut tiedot tulevat itsestään

APPENDIX 5: User Interface – Printing Press 2



Operaattorin ID: 1 | Setupin numero: 17

Tapahtuman numero 1258 | Tapahtuma alkoi 27.3.2014 20:15:02 | Tapahtuma päättyi 16.4.2014 8:31:35

Tapahtuman tyyppi	Run	Run
Tapahtuman luonne	19_Tilaus valmis	19_Tilaus valmis
Koneen osa		
Tapahtuman syy		
Hylky (laskurin jälkeinen) (kg)		
Hylky (ennen laskuria) (kg)		
Aihiot yhteensä (kpl)	573858	
Hyvät aihiot (kpl)	552036	
Laskennallinen hylky (kpl)	21822	
Operaattoreiden kommentit		

Viimeiset kymmenen tapahtumaa:

	Tapahtuman numero	Tapahtuma alkoi	Tapahtuma päättyi	Tapahtuman tyyppi	Tapahtuman luonne	Koneen osa	Tapahtuman syy	Hylky
1	1258	27.3.2014 20:15:02	16.4.2014 8:31:35	Run	19_Tilaus valmis			
2	1257	27.3.2014 19:45:04	27.3.2014 20:15:02	Run	14_1 Tuotannon aikainen hylky	Aukirullain	698-01 Karvihylky	
3	1256	27.3.2014 19:45:03	27.3.2014 19:45:04	Run	14_1 Tuotannon aikainen hylky	Yleisiä	330-01 Laatunäytteet	12,08
4	1255	27.3.2014 19:45:02	27.3.2014 19:45:03	Run	14_1 Tuotannon aikainen hylky	Futec	638-02 Futecin havaitsemat viat	
5	1254	27.3.2014 19:45:01	27.3.2014 19:45:02	Run	14_Tuotannon aloitus			
6	1253	27.3.2014 19:25:01	27.3.2014 19:45:01	Downtime	26_Lyhyt pysähdys	UV-lamput	370-01 Lyhyt pysähdys	
7	1252	27.3.2014 16:13:03	27.3.2014 19:25:01	Run	14_1 Tuotannon aikainen hylky	Futec		
8	1251	27.3.2014 16:13:02	27.3.2014 16:13:03	Run	14_1 Tuotannon aikainen hylky	Kostutusvesijärjestelmä	-	
9	1250	27.3.2014 16:13:01	27.3.2014 16:13:02	Run	14_1 Tuotannon aikainen hylky	Futec	638-02 Futecin havaitsemat viat	
10	1249	27.3.2014 16:13:00	27.3.2014 16:13:01	Run	14_Tuotannon aloitus			

APPENDIX 6: User Interface – Printing Press 3

Syötä allaolevaan laatikkoon Setupin numero, jolle haluat tulostaa uudet lavalaput.

Setupin numero: 9

Tulostamaan lavalappuja | Skannaamaan lavoja | Tulostamaan tilauslappuja

Poikkeustapauksessa lavalapun tulostaminen

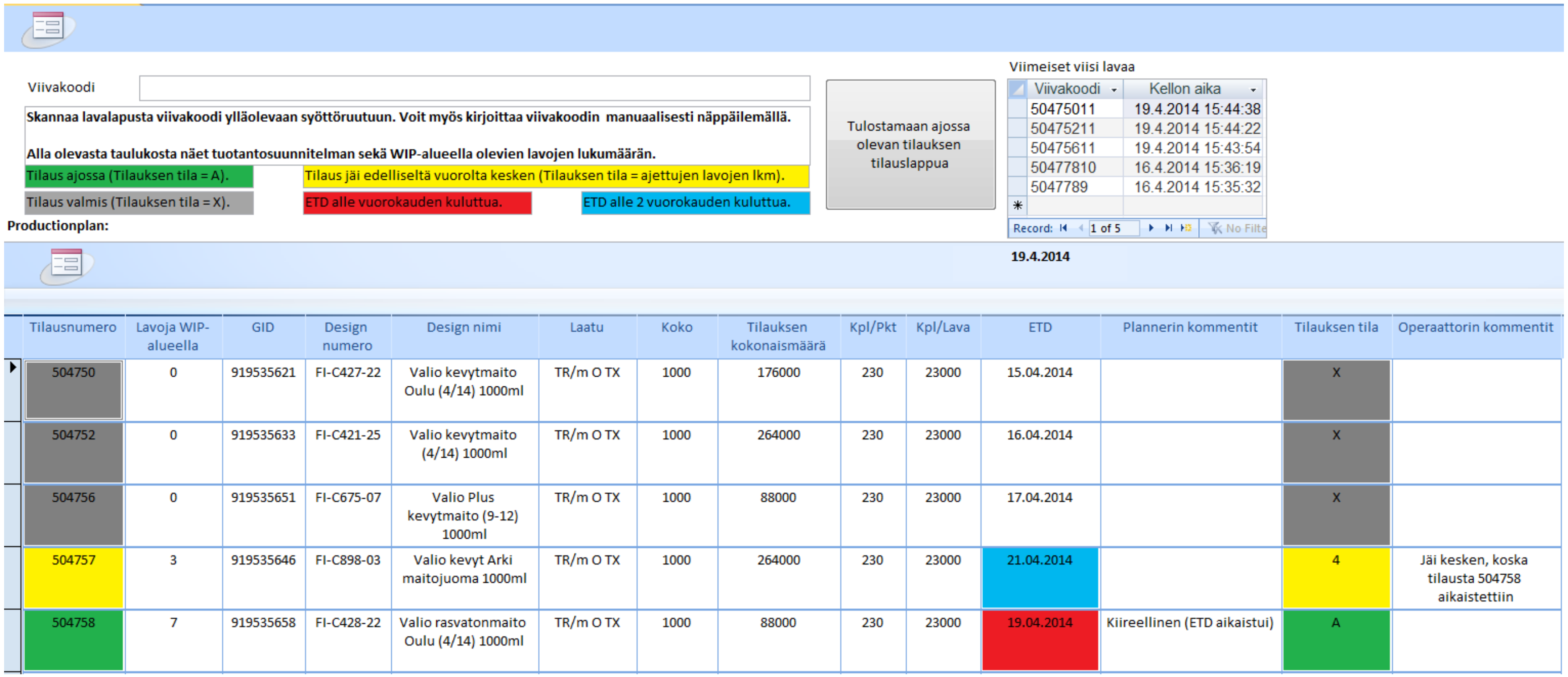
- Jos lavat eivät valmistu samanaikaisesti jokaiselta linjalta. Tällöin sinun tulee poistaa Setupin numero siltä linjalta, jolta lavaa ei valmistunut. Tämän jälkeen täytyy muistaa siirtää hiiren kursori pois siltä Setupilta, jolle muutokset tehtiin.

- Kun lavoja taas valmistuu jokaiselta linjalta samanaikaisesti, voit kirjoittaa Setupin numeron takaisin sille varatulle alueelle. Myös tällöin hiiren kursori kannattaa siirtää jollekkin muulle Setupille.

- Jos lavalappuja ei muodostunut kaikille halutuille linjoille, voit käyttää poikkeustapauksiin rakennettua tulostustoimintoa. Tänne on listattu kaikki lavalaput, jotka ovat muodostuneet.

Tilausnumero	Suunnittelijan kommentit	Setupin numero	Linja	GID	Koko	Avaus	Laatu	Designin numero	Designin nimi	QSV	Materiaali	Tilauksen koko	ETD	Operaattorin kommentit
504778		1	1	919535773	1000	EO	GT/m O TX-H	FI-C936-01	GT Mansikkajogurtti 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504777		1	2	919535767	1000	EO	GT/m O TX-H	FI-C714-03	GT lakt.vanilja jogurtti (2-4/14) 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504774		1	3	919535753	1000	EO	GT/m O TX-H	FI-C935-01	GT Banaanijogurtti 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504772		2	1	919535747	1000	EO	GT/m O TX-H	FI-C695-01	GT A+ luonnon jogurtti 1000ml	B051-810-30	1242-033-01	84000	2014-03-26	
504772		2	2	919535747	1000	EO	GT/m O TX-H	FI-C695-01	GT A+ luonnon jogurtti 1000ml	B051-810-30	1242-033-01	84000	2014-03-26	
504782		2	3	919535791	1000	EO	GT/m O TX-H	FI-C941-01	GT rasv.hedelmÃ¤mix jog. Hyla 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504780		3	1	919535781	1000	EO	GT/m O TX-H	FI-C938-01	GT Mustikka-vanilja jogurtti 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504775		3	2	919535758	1000	EO	GT/m O TX-H	FI-C713-03	GT lakt.mansikka-herukka jog.(2-4)1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504771		3	3	919535744	1000	EO	GT/m O TX-H	FI-C841-01	GT YO-ghurt sommarbÃ¤r 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504783		4	1	919535794	1000	EO	GT/m O TX-H	FI-C937-01	GT Uuniomenajogurtti 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504781		4	2	919535788	1000	EO	GT/m O TX-H	FI-C934-01	GT rasv. metsÃ¤m.jog. Hyla 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504779		4	3	919535778	1000	EO	GT/m O TX-H	FI-C942-01	GT marjapommi jogurtti 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504785		5	1	919535799	1000	EO	GT/m O TX-H	FI-C939-01	GT Vaniljajogurtti Hyla 1000ml	B051-810-30	1242-033-01	42000	2014-03-26	
504784		5	2	919535801	1000	EO	GT/m O TX-H	FI-C707-03	GT YO-ghurt skogsbÃ¤r 1000ml	B051-810-30	1242-033-01	84000	2014-03-26	
504784		5	3	919535801	1000	EO	GT/m O TX-H	FI-C707-03	GT YO-ghurt skogsbÃ¤r 1000ml	B051-810-30	1242-033-01	84000	2014-03-26	
504842		6	1	919552856	1000	EO	TR/m O TX-H	FI-C611-01	A+ luonnon jogurtti 1000ml	8450-810-15	1242-033-01	88000	2014-04-03	
504842		6	2	919552856	1000	EO	TR/m O TX-H	FI-C611-01	A+ luonnon jogurtti 1000ml	8450-810-15	1242-033-01	88000	2014-04-03	
504842		6	3	919552856	1000	EO	TR/m O TX-H	FI-C611-01	A+ luonnon jogurtti 1000ml	8450-810-15	1242-033-01	88000	2014-04-03	
504757		7	1	919535646	1000	EO	TR/m O TX	FI-C898-03	Valio kevyt Arki maitojuoma 1000ml	8226-810-15	1266-033-01	264000	2014-03-27	
504761		7	2	919535675	1000	Twist cap 33,5 mm	TR/m O TX	FI-C277-17	Valio lakt. kev.maitoj.(4/14)1000ml	8226-810-75	1266-033-01	176000	2014-03-28	
504750		7	3	919535621	1000	Twist cap 33,5 mm	TR/m O TX	FI-C427-22	Valio kevytmaito Oulu (4/14) 1000ml	8226-810-75	1266-033-01	176000	2014-03-27	
504757		8	1	919535646	1000	EO	TR/m O TX	FI-C898-03	Valio kevyt Arki maitojuoma 1000ml	8226-810-15	1266-033-01	264000	2014-03-27	
504756		8	2	919535651	1000	Twist cap 33,5 mm	TR/m O TX	FI-C675-07	Valio Plus kevytmaito (9-12) 1000ml	8226-810-75	1266-033-01	88000	2014-03-27	
504758		8	3	919535658	1000	Twist cap 33,5 mm	TR/m O TX	FI-C428-22	Valio rasvatonmaito Oulu (4/14) 1000ml	8226-810-75	1266-033-01	88000	2014-03-27	
504760		9	1	919535655	1000	EO	TR/m O TX	FI-C899-03	Valio rasvaton Arki maitojuoma 1000ml	8226-810-15	1266-033-01	264000	2014-03-27	
504786		9	2	919536233	1000	Twist cap 33,5 mm	TR/m O TX	FI-C422-24	Valio rasvatonmaito (4/14) 1000ml	8226-810-75	1266-033-01	352000	2014-03-28	
504752		9	3	919535633	1000	Twist cap 33,5 mm	TR/m O TX	FI-C421-25	Valio kevytmaito (4/14) 1000ml	8226-810-75	1266-033-01	264000	2014-03-27	
504763		10	1	919535677	1000	EO	TR/m O TX	FI-C304-07	Valio luomu rasvatonmaito(2-4/14)1000ml	8226-810-15	1266-033-01	88000	2014-03-28	
504786		10	2	919536233	1000	Twist cap 33,5 mm	TR/m O TX	FI-C422-24	Valio rasvatonmaito (4/14) 1000ml	8226-810-75	1266-033-01	352000	2014-03-28	
504759		10	3	919535663	1000	Twist cap 33,5 mm	TR/m O TX	FI-C652-06	Valio Plus rasvaton maito (Teemu) 1000ml	8226-810-75	1266-033-01	88000	2014-03-28	

APPENDIX 7: User Interface – Side Sealer 1



Viivakoodi

Skannaa lavalapusta viivakoodi ylläolevaan syöttöruutuun. Voit myös kirjoittaa viivakoodin manuaalisesti näppäilemällä.

Alla olevasta taulukosta näet tuotantosuunnitelman sekä WIP-alueella olevien lavojen lukumäärän.

Tilaus ajossa (Tilauksen tila = A).

Tilaus jäi edelliseltä vuorolta kesken (Tilauksen tila = ajettujen lavojen lkm).

Tilaus valmis (Tilauksen tila = X).

ETD alle vuorokauden kuluttua.

ETD alle 2 vuorokauden kuluttua.

Tulostamaan ajossa olevan tilauksen tilauslappua

Viimeiset viisi lavaa

Viivakoodi	Kellon aika
50475011	19.4.2014 15:44:38
50475211	19.4.2014 15:44:22
50475611	19.4.2014 15:43:54
50477810	16.4.2014 15:36:19
5047789	16.4.2014 15:35:32

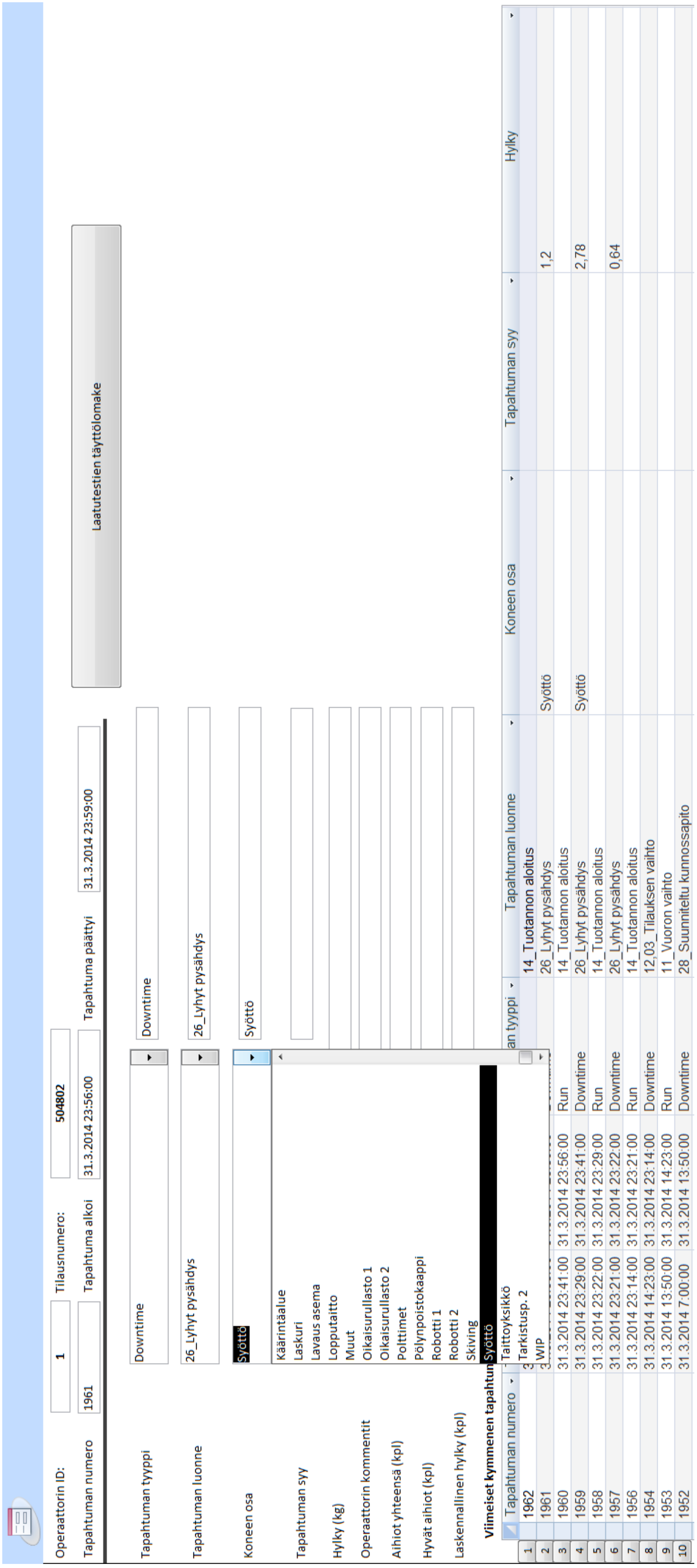
Record: 1 of 5 No Filte

Productionplan:

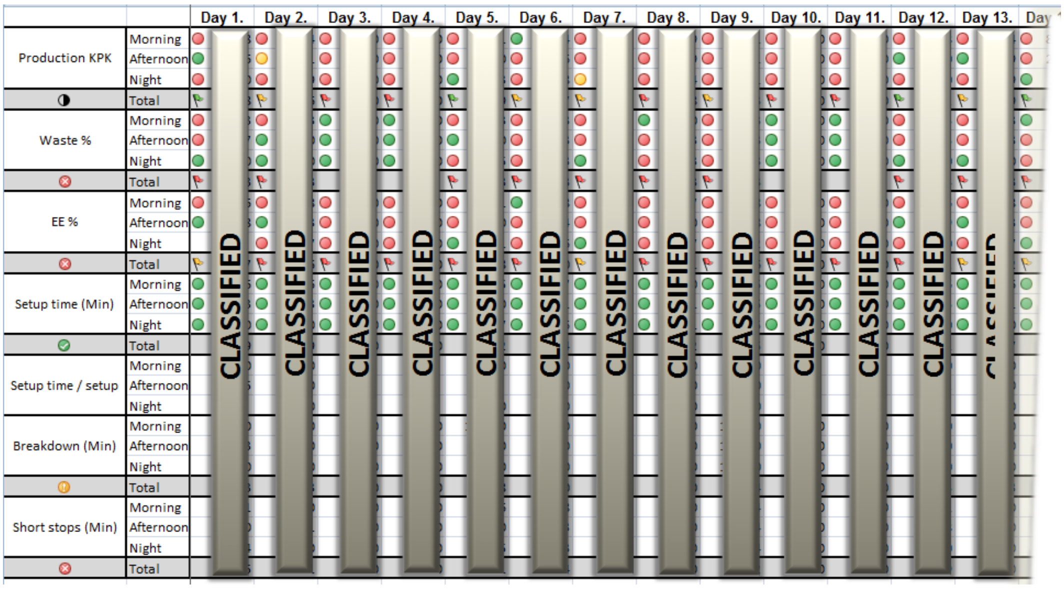
19.4.2014

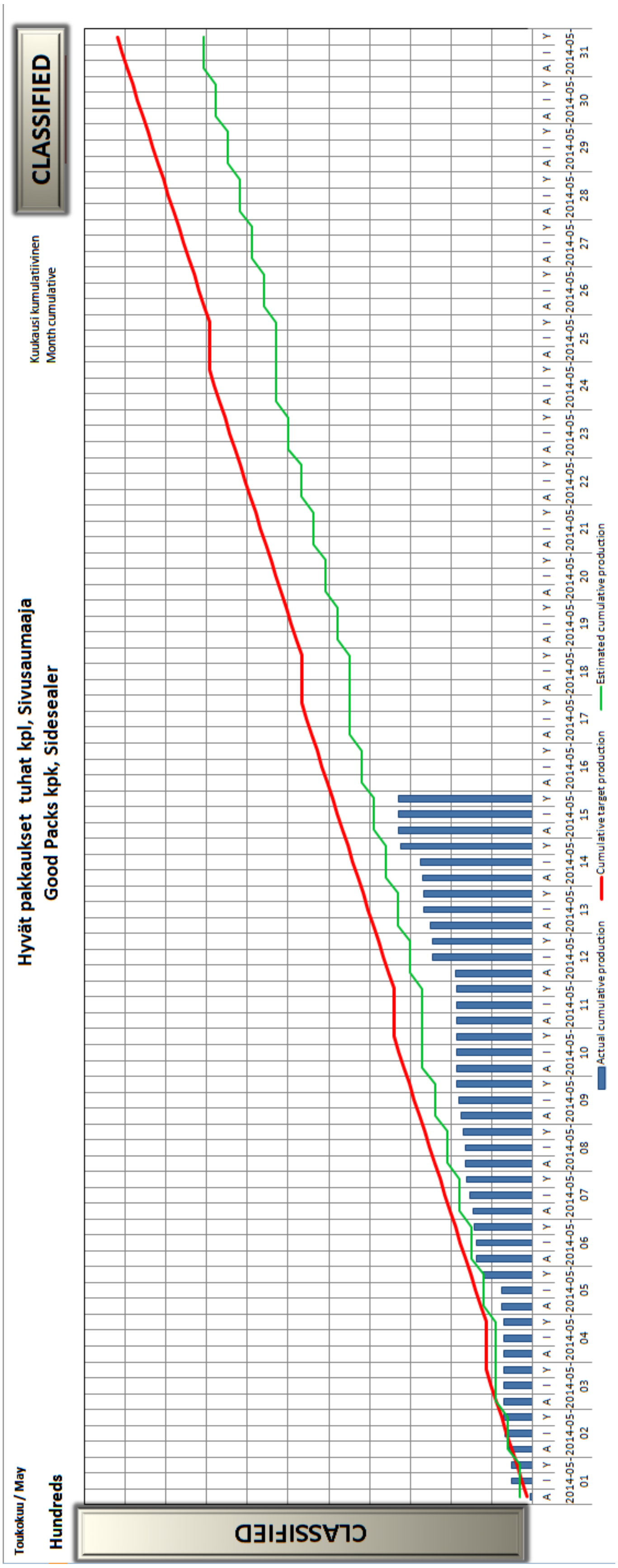
Tilausnumero	Lavoja WIP-alueella	GID	Design numero	Design nimi	Laatu	Koko	Tilauksen kokonaismäärä	Kpl/Pkt	Kpl/Lava	ETD	Plannerin kommentit	Tilauksen tila	Operaattorin kommentit
504750	0	919535621	FI-C427-22	Valio kevytmaito Oulu (4/14) 1000ml	TR/m O TX	1000	176000	230	23000	15.04.2014		X	
504752	0	919535633	FI-C421-25	Valio kevytmaito (4/14) 1000ml	TR/m O TX	1000	264000	230	23000	16.04.2014		X	
504756	0	919535651	FI-C675-07	Valio Plus kevytmaito (9-12) 1000ml	TR/m O TX	1000	88000	230	23000	17.04.2014		X	
504757	3	919535646	FI-C898-03	Valio kevyt Arki maitojuoma 1000ml	TR/m O TX	1000	264000	230	23000	21.04.2014		4	Jäi kesken, koska tilausta 504758 aikaistettiin
504758	7	919535658	FI-C428-22	Valio rasvatonmaito Oulu (4/14) 1000ml	TR/m O TX	1000	88000	230	23000	19.04.2014	Kiireellinen (ETD aikaistui)	A	

APPENDIX 8: User Interface – Side Sealer 2



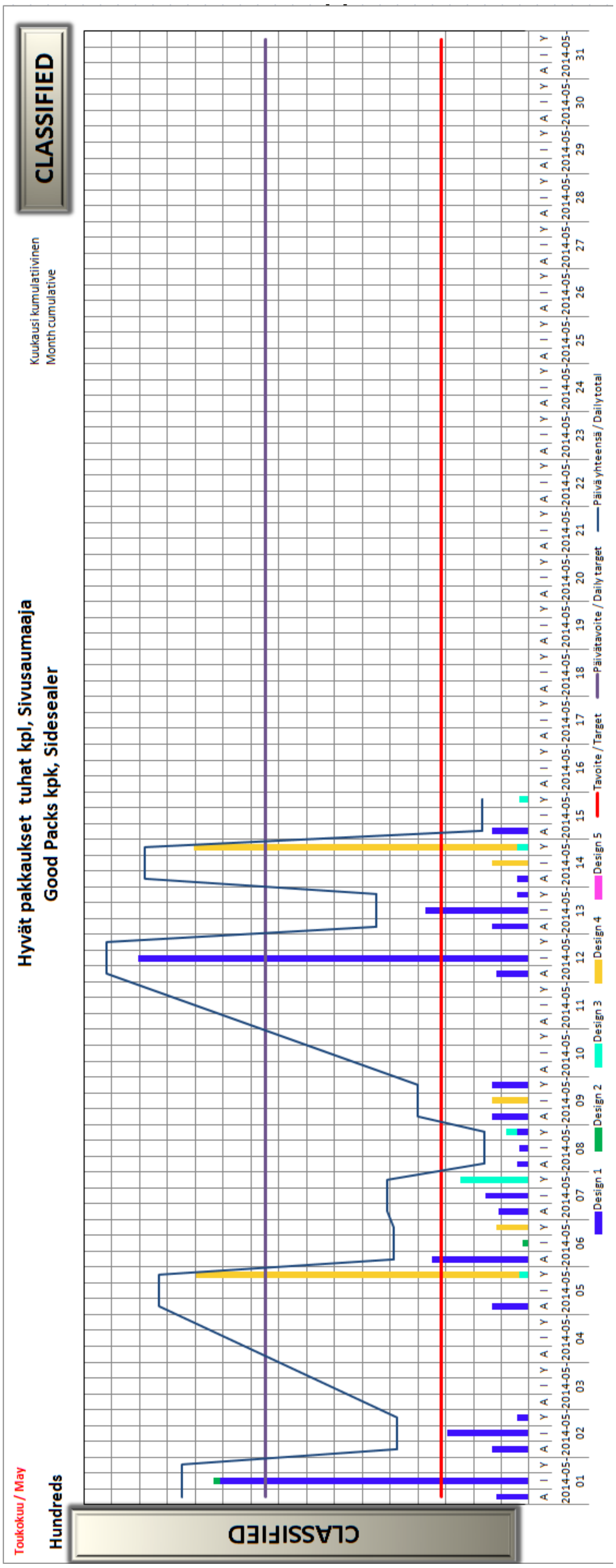
APPENDIX 9: Dashboard – General Performance

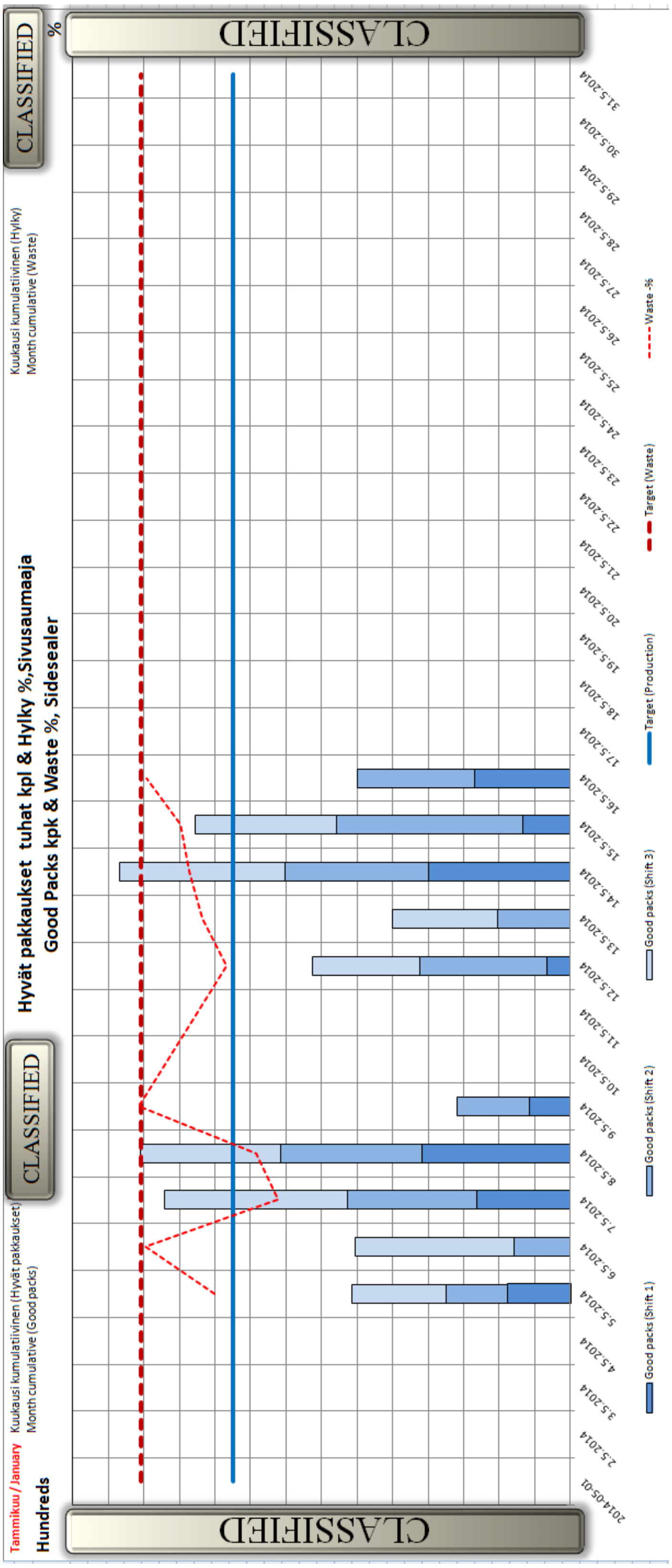



CLASSIFIED
Toukokuu / May
Hundreds
Hyvät pakkaukset tuhat kpl, Sivusaumaaja
Good Packs kpk, Sidesealer
Kuukausi kumulatiivinen
Month cumulative
CLASSIFIED
Actual cumulative production
Cumulative target production
Estimated cumulative production

(Continued)

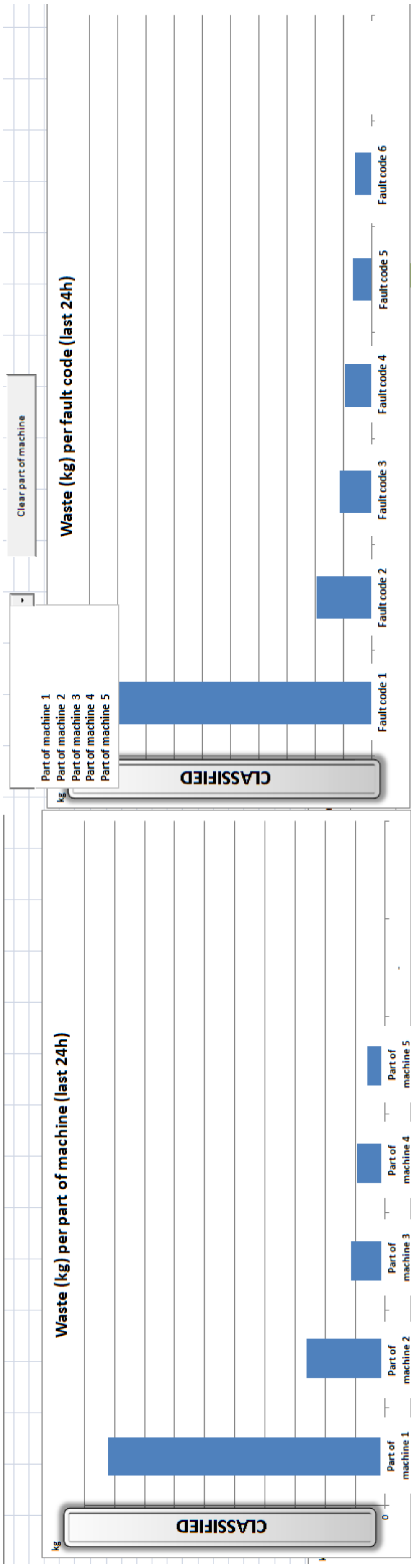
APPENDIX 10: Dashboard – Availability (Continues)




Tammikuu / January
Hundreds
Kuukausi kumulatiivinen (Hyvät pakkaukset)
Month cumulative (Good packs)
CLASSIFIED
Hyvät pakkaukset tuhat kpl & Hylky %,Sivusaumaaja
Good Packs kpk & Waste %, Sidesealer
Kuukausi kumulatiivinen (Hylky)
Month cumulative (Waste)
CLASSIFIED
%
CLASSIFIED
CLASSIFIED
2014-05-01
2.5.2014
3.5.2014
4.5.2014
5.5.2014
6.5.2014
7.5.2014
8.5.2014
9.5.2014
10.5.2014
11.5.2014
12.5.2014
13.5.2014
14.5.2014
15.5.2014
16.5.2014
17.5.2014
18.5.2014
19.5.2014
20.5.2014
21.5.2014
22.5.2014
23.5.2014
24.5.2014
25.5.2014
26.5.2014
27.5.2014
28.5.2014
29.5.2014
30.5.2014
31.5.2014
Good packs (Shift 1)
Good packs (Shift 2)
Good packs (Shift 3)
Target (Production)
Target (Waste)
Waste %

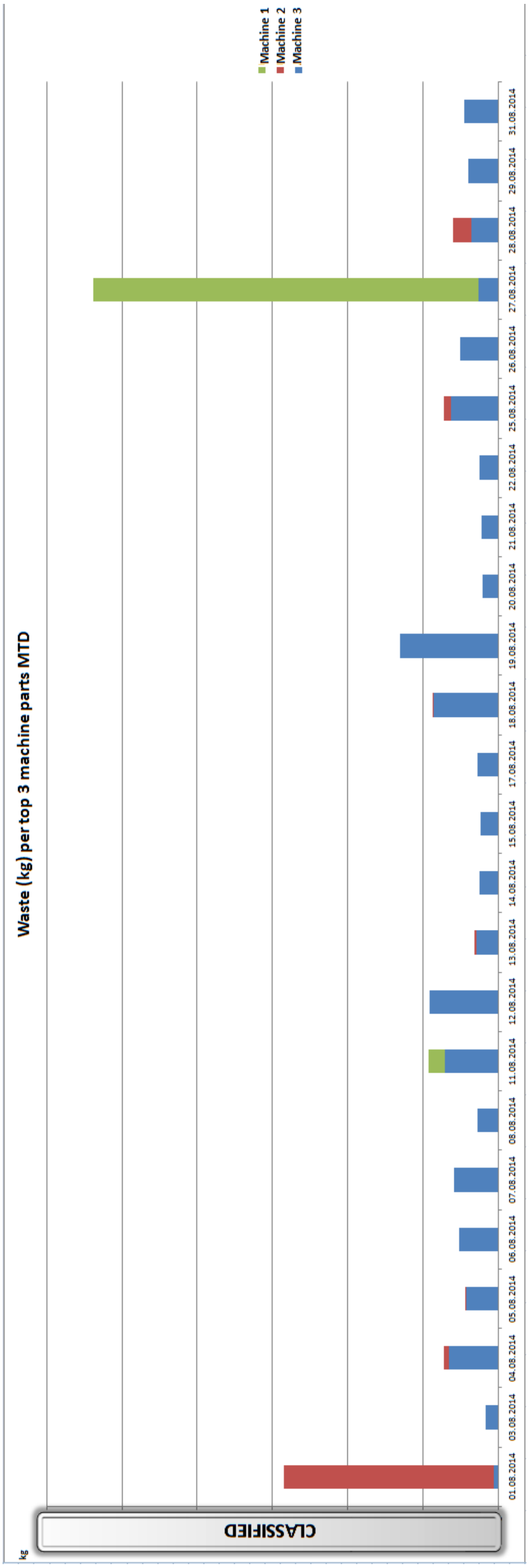
(Continued)

APPENDIX 11: Dashboard – Quality (Continues)

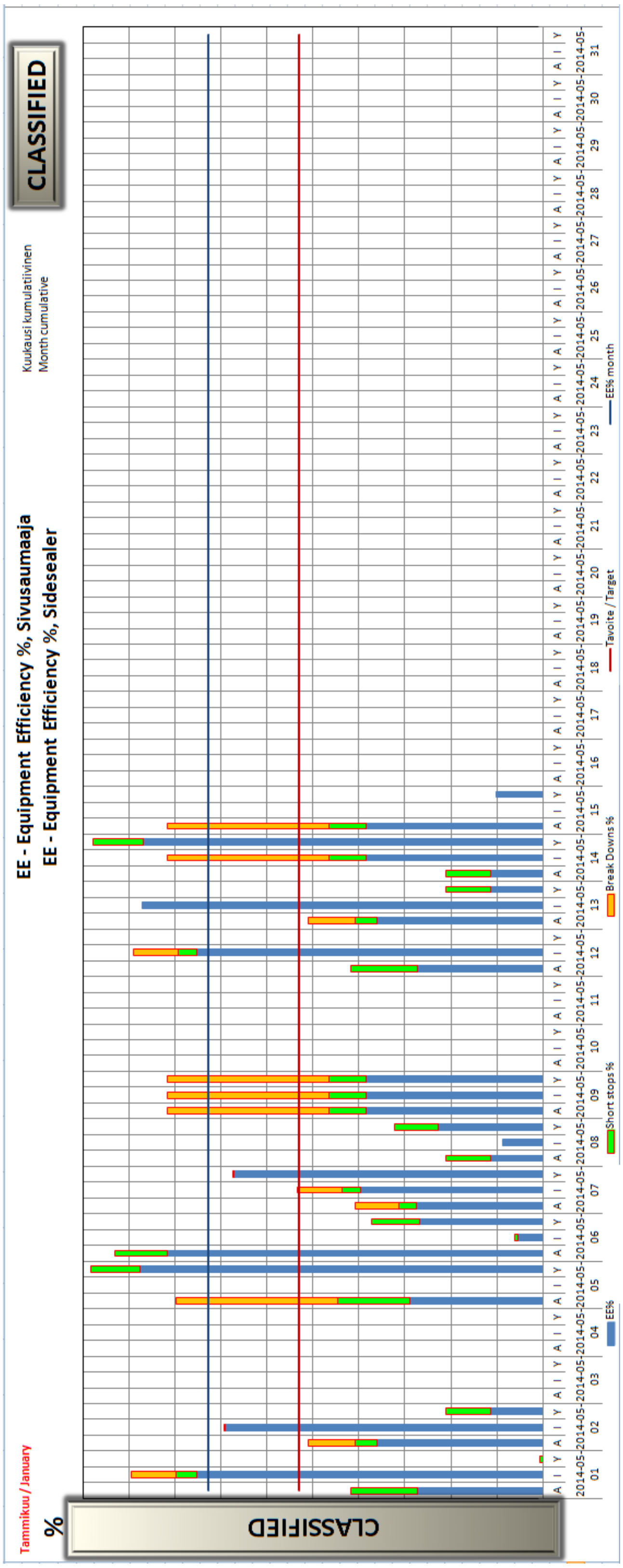


(Continued)

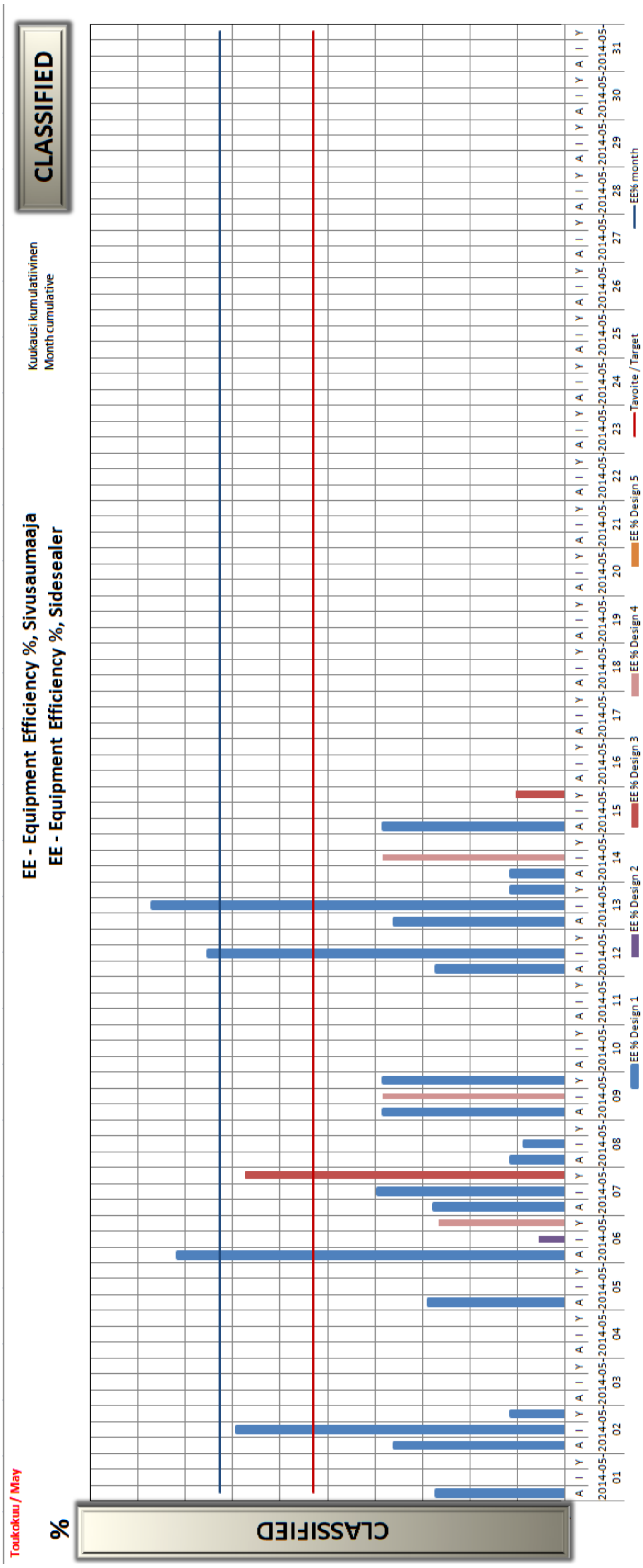
APPENDIX 11: Dashboard – Quality (Continues)



APPENDIX 12: Dashboard – Efficiency

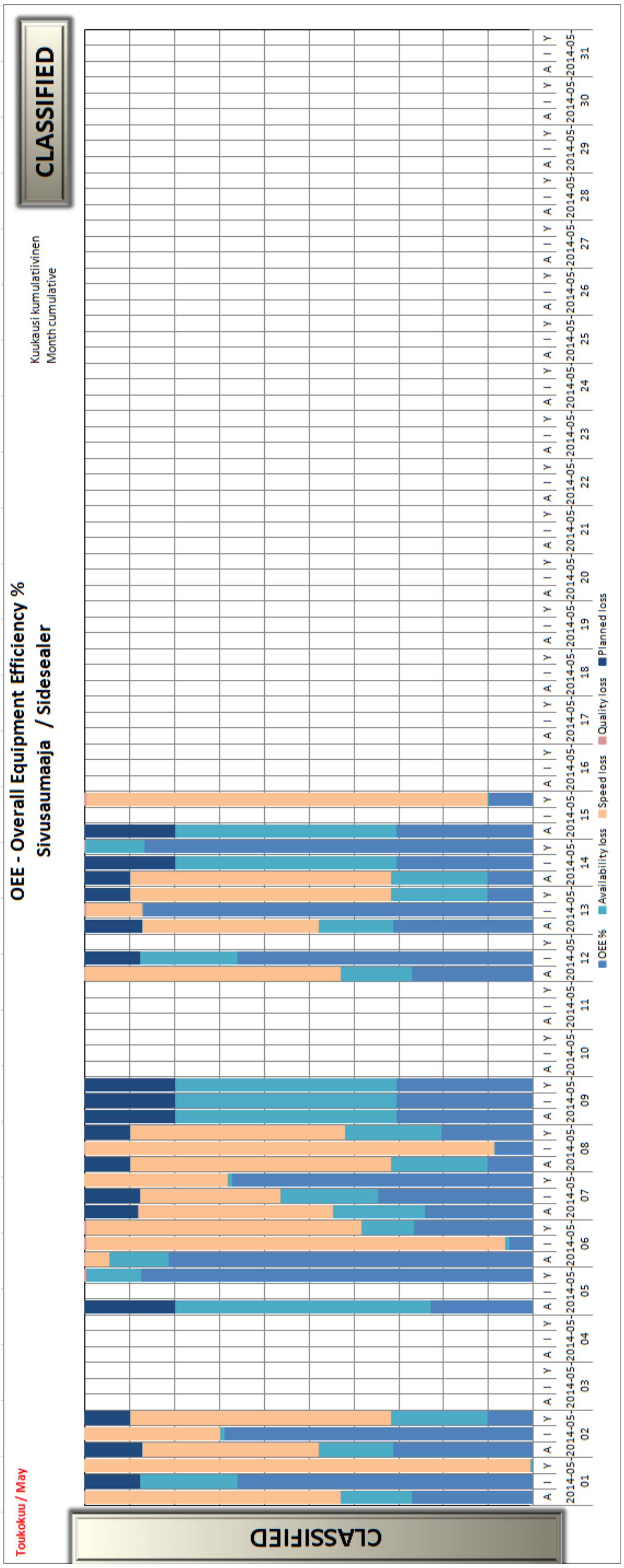


(Continued)

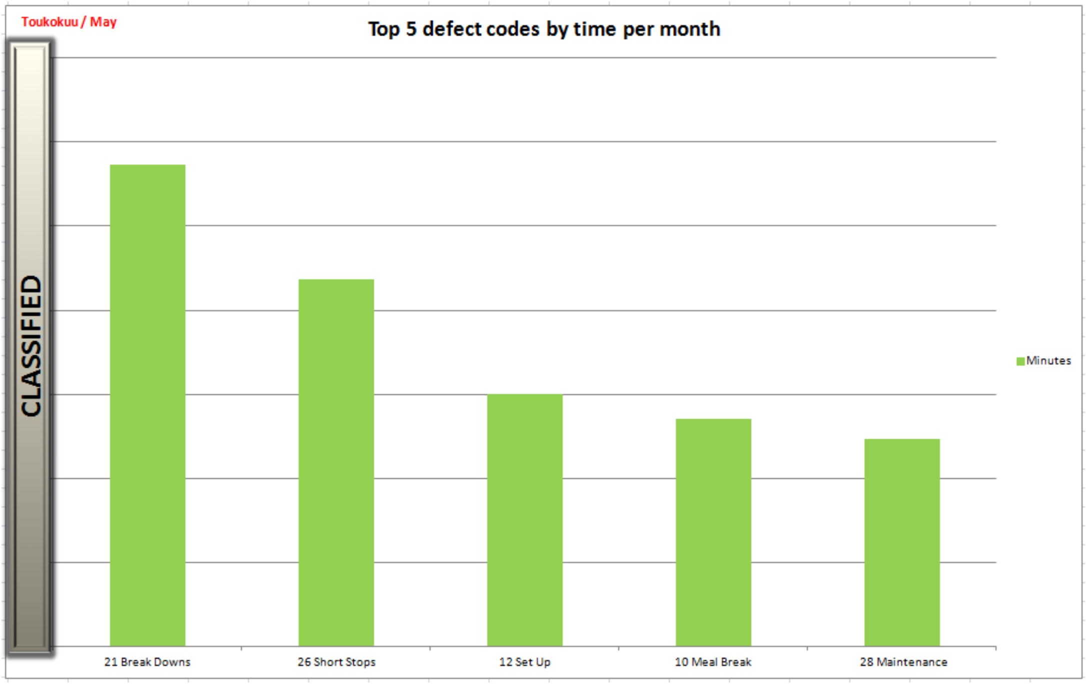


EE - Equipment Efficiency %, Sivusaumaaja
EE - Equipment Efficiency %, Sidesealer
Kuukausi kumulatiivinen
Month cumulative
CLASSIFIED
Toukokuu / May
%
CLASSIFIED
EE % Design 1
EE % Design 2
EE % Design 3
EE % Design 4
EE % Design 5
Tavoite / Target
EE% month

APPENDIX 12: Dashboard – Efficiency (Continues)



(Continued)



Time divided into Strategic losses, Planned losses & Operational losses