

MASTER DATA MANAGEMENT
AS A TOOL TO IMPROVE BUSINESS PROFITABILITY



Bachelor's thesis

Degree Program in Industrial Management Engineering

Spring 2014

Delia Rodríguez Lucas



Valkeakoski
Final Thesis
Degree Program in Industrial Management Engineering

Author	Delia Rodríguez Lucas	Year 2014
Subject of Bachelor's thesis	Master Data Management	

ABSTRACT

Master Data is a topic whose importance is increasing exponentially due to the fact that it has a severe impact in the overall profitability of companies, even if it is not a direct matter neither in the incomes nor in the benefits. Thus, the aim of this Thesis was to explain from the pillars what is this issue, how it is linked to companies, how it affects the business process and finally which solutions exist to prevent and solve its subsequent problems.

The method to research how Master Data Management impacts profitability was a business case implemented through SAP ERP. In the case, a wholesale company purchased products from Nestlé and sold them to EcoMarket, a supermarket chain. The study focused on how diverse processes developed within the system and how Master Data was involved. Two scenarios were considered: Firstly, a good management was performed where I put into practice different maintenance methods that had been previously exposed. Secondly, the opposite situation was carried out in order to prove the concrete consequences derived from a poor management.

Based on the results of the comparison, it can be said that the implications of a poor Master Data Management are visible even in the most basic procedures. Profitability losses can appear in different shapes, some of them direct (such as expenses increase) and others indirect (such as time waste or customer dissatisfaction). To conclude, I propose an approach that combines IT and business with the aim of framing the Master Data problem. The reason is that most of the current conducted studies focus on the technical scope but, according to current surveys, the main cause of poor Master Data is business data entering. Therefore, I believe the Master Data Management solutions should not be only concentrated in the technical aspect, and spread out covering both angles.

Keywords Master Data Management, Data Quality, Data Maintenance, SAP ERP.

Pages 106 p.



CONTENTS

1	INTRODUCTION.....	1
1.1	Commissioning organization.....	1
1.2	Thesis Objective.....	1
1.3	Reasons for choosing topic.....	2
1.4	Research methodology.....	3
1.5	Master Data concept.....	4
1.6	Related concepts.....	6
2	MASTER DATA MANAGEMENT - THEORY.....	7
2.1	Previous concepts.....	7
2.1.1	Database.....	7
2.1.2	ERP.....	11
2.2	Master Data in a company.....	16
2.2.1	Enterprise Data.....	16
2.2.2	Business Processes.....	18
2.2.2.1.	Aggregate Planning – AP.....	18
2.2.2.2.	Master Production Schedule – MPS.....	20
2.2.2.3.	Bill of Materials – BOM.....	21
2.2.2.4.	Inventory Records.....	22
2.2.2.5.	Materials Requirement Planning – MRP.....	23
2.2.2.6.	Quality.....	26
2.2.2.7.	Routing.....	27
2.3	Implications and consequences.....	28
2.3.1	Data Quality.....	28
2.3.2	Impacts of poor quality data.....	30
2.3.3	Costs of poor quality data.....	33
2.3.4	Sources of poor data quality.....	37
2.4	Master Data Management.....	42
2.4.1	MDM Solution.....	42
2.4.2	Master Data Creation.....	46
2.4.3	Master Data Maintenance.....	47



3	MASTER DATA MANAGEMENT - PRACTICE	50
3.1	Implementation of an ERP	50
3.2	SAP functions.....	54
3.2.1	Purchasing	55
3.2.2	Sales & Distribution	56
3.3	Practical Case	58
3.3.1	Adequate Management.....	59
3.3.1.1.	Materials Management	59
3.3.1.2.	Purchasing	66
3.3.1.3.	Sales and Distribution.....	74
3.3.2	Inadequate management	81
3.3.2.1.	Materials Management	81
3.3.2.2.	Purchasing	87
3.3.2.3.	Sales and Distribution.....	94
4	CONCLUSIONS	101
	SOURCES	103



1 INTRODUCTION

In this section I will put into context this Final Thesis. Thus, I will firstly explain the objectives and the reasons for choosing this concrete topic. After that, I will describe the research methodology used to get the information in which this paper is based. To finish with, the main topic will be described along with related concepts that will help to understand its scope.

1.1 Commissioning organization

HAMK, Degree Program in Industrial Management

- Author: Delia Rodríguez, deliarodriguezlucas@gmail.com
- Supervisor: Tapani Honkanen, Tapani.Honkanen@hamk.fi
- Representative: Kristiina Ranta, Kristiina.Ranta@hamk.fi

1.2 Thesis Objective

The main objective of this Final Thesis is to explain what is Master Data Management and how it affects to the overall profitability of companies.

This general purpose can be divided in two sub-objectives that coincide with the structure of this paper:

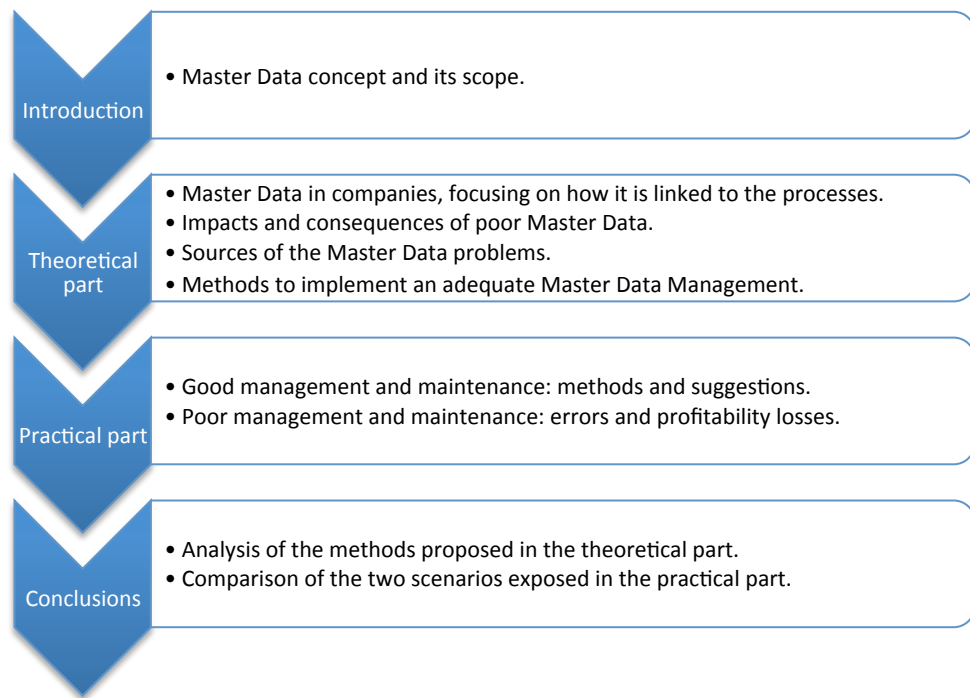
The first one applies to the Theoretical part, and it is to describe the different issues that surround Master Data Management.

Hence, I will begin by properly describing the concept of Master Data. Later, I will explain how it is linked to companies, focusing on how it is used through the different processes of a manufacturing company. Finally, I will analyse its impacts in profitability, exposing which solutions exist to prevent and solve the related problems.

The second one applies to the Practical part, and it is to prove through a business case the real impact of Master Data Management in profitability.

The implementation of the case will be carried out through SAP ERP and it will be based on the different procedures that a wholesale company performs on a daily basis. The analysis will be based on the comparison of a good and poor maintenance of Master Data records.

Hence, the structure used to pursue the Thesis' aim will be as follows:



Graph 1.2. Delia Rodríguez, Objectives.

Note: the Master Data Management topic will be approached with the assumption that the reader has no previous knowledge in the topic.

1.3 Reasons for choosing topic

There are two principal motives of having chosen Master Data Management for the subject of this thesis.

Data Bases, ERPs and other related subjects conform a scope in which the author has not focused during the studies, but in which I am very interested.

Furthermore, Master Data Management is an actual and current topic: Data Base and its management through ERPs is nowadays an emerging issue with increasing importance every year. What is more, Master Data is one of the main problems that big companies face, and therefore it is totally real and practical. Therefore, I believe it is a very useful topic.

1.4 Research methodology

In order to ensure the quality of the information in which this Thesis is based, I will pursue a strict method in terms of researching.

Thus, the objective is to seek accurate and reliable materials. The main sources will be provided by E-books from HAMK library but, if required, I will investigate using further methods -papers from other universities, investigations from research centres, thesis from scholars, information available online, etc.

The process will be:

- Looking for books related to a certain topic.
- Checking if the given topic is approached in the right manner, or in the same scope as the one wanted in the Thesis.
- If possible, search critics or external opinions about the books. Thus, I will be able to check how well/bad considered is this source in the scientific community.
- Select the important parts of the source, so only the most relevant information is used, and there are not useless information.
- Adapt it to the Thesis ambit. Hence, all the text it is coherent and appropriately connected.

If possible, I will use different sources to write about the same sub-topic. Consequently, I will have more than a point of view to support it.

The feedback method will be based on personal research and periodically revisions of my supervisor in order to review the last advances of my project, check if I am using an appropriate approach and set new aims and sub-objectives.

1.5 Master Data concept

Before starting to define what is Master Data, it is necessary to put this matter into context. That is the reason why I will begin this section by describing the types of data used in organizations' systems.

I will use the approach exposed by Wolter & Haselden (2006) from Microsoft. According to it, there are essentially five types of data:

- **Unstructured:** This is data that can be found from different not organized sources. These sources can be e-mails, white papers, magazine articles, intranet portals, product specifications, files, etc.
- **Metadata:** This is data about another data. It may reside in a formal repository or in other forms such as report definitions or column descriptions.
- **Hierarchical:** This is data that stores the relationships between other data. It may be stored as part of an accounting system or separately as descriptions of real-world relationships, such as company organizational structures or product lines.
- **Transactional:** This is data that is temporal because it is linked to a concrete transaction in the company. Thus, it is typically related to sales, deliveries, invoices, claims or other monetary and non-monetary interactions.
- **Master:** Master data are the critical nouns of a business. These records can normally be classified in four groups: people, things, places, and concepts.

The main comparison in order to understand Master Data it is with Transactional Data. Schulz (2012, 62) a redactor from SAP Press, further explains the differences between them:

He states that Master Data is “information that remains unchanged for a long time and is repeatedly required in business processes”. It is the data that, once it has been created, can be available to diverse users, departments or areas in a given organization. Examples within a company are:

Customer: name, address, customer number, terms of delivery, etc.

Material: material number, short text, size, weight, valuation, etc.

On the other hand, Schulz exposes that the key aspect of Transactional Data is that it can be changed. Therefore, it is created, edited and valid only for a limited period of time for a special transaction only. Examples within a company are:

Invoice: document number, debtor, amount, material, quantity, etc.

Purchase order: document number, supplier, material, quantity, conditions, etc.

Graphically, the differences between master and transactional data would be:

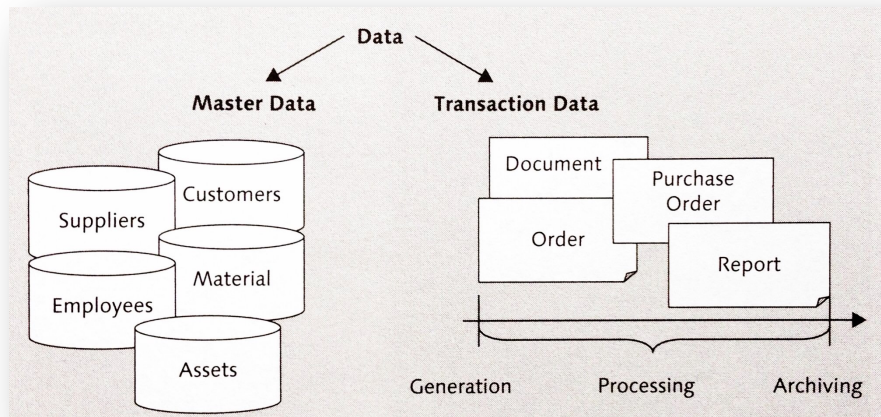


Figure 1.5. O. Schulz, Master Data and Transactional Data.

To sum up, Master Data refers to the data that represents the key information of a system. It is the one that is relevant and shared within different aspects, divisions, departments or areas, linking local issues to the rest of the organization.

Some examples to illustrate this definition are shown below in table 1.5. Note that not all of them are related to the business scope, as far as it is a concept that can also be used in other ambits.

Organization	Master Data	Persons/divisions that need it
Manufacturing company	Clients, Raw Materials, Product records	Production manager, Quality manager, Sales manager and related staff
University	Students basic data: names, degree, courses taken	All the teachers' departments as well as the administration
Retailer company	Clients, Customer requirements, Product records	Sales manager, Purchases manager and related staff
Hospital	Patients' information	Doctors, nurses, administration...

Table 1.5. Delia Rodríguez, Master Data examples.

An adequate Master Data Management is very important since an error in this kind of data could easily spread through the system, causing damage to the information of the entire entity.

Throughout this Thesis, I will expose the different aspects regarding Master Data Management.

1.6 Related concepts

Apart from Master Data, there are many topics that apply to the same area, whose meanings or concepts are important to know.

The most important topics are as follows:

- **Big data:** Big Data is a general term used to describe the voluminous amount of unstructured and semi-structured data a company creates -data that would take too much time and cost too much money to load into a relational database for analysis. Although Big data doesn't refer to any specific quantity, the term is often used when speaking about petabytes and exabytes of data. (Rouse 2011)
- **Data Governance:** Data Governance is the specification of decision rights and an accountability framework to encourage desirable behaviour in the valuation, creation, storage, use, archival and deletion of data and information. It includes the processes, roles, standards and metrics that ensure the effective and efficient use of data and information. In enabling an organization to achieve its goals. (Oracle 2011)
- **Metadata:** Metadata is structured information that describes, explains or locates an information resource (NISO Press 2004, 3).
A more general definition is "Metadata is simply data about data" (Stefanov 2007, 19). She also explains that there are mainly two kinds of Metadata:
Technical: It describes data's structure. For instance: data types, data access restrictions, data traces...
Business: It describes the business context of the core data. For instance: economic rules, good values to ratios...
- **Data Mining:** Data mining is the process of discovering patterns from large-scale data in order to create descriptive, understandable, and predictive models. (Zaki & Meira 2014, 1)
- **Business Intelligence:** Business intelligence is an umbrella term that includes the applications, tools and best practices that enable access to and analysis of information. Its objective is to improve and optimize decisions and performance in a business. (Gartner n.d)

2 MASTER DATA MANAGEMENT - THEORY

In this section, Master Data will be deeply described along with other related concepts that will make it easier to understand its context.

Furthermore, I will expose the main business processes in the case of a Manufacturing company, explaining how Master Data is involved. Later on, I will show the consequences of poor data quality in this issue. Finally, I will introduce the different ways and tools used to adequately manage Master Data and ensure its quality.

This part will be the base for understanding the second one, Master Data Management – Practice (section 3), where I will make a real example in order to prove the consequences of a poor Master Data Management.

2.1 Previous concepts

In order to understand correctly the meaning of Master Data, it is necessary to accurately explain two basic concepts: databases and ERPs.

2.1.1 Database

The concept database can be understood in different ways. For example, Robbins (1995, 2) describes it as “any collection of related data”. He also provides a more restrictive angle, stating that is “a persistent, logically, coherent collection of inherently meaningful data”. But there are scholars that believe that a database is something more general. For instance, Schneider (2005, 1-1) states that “a database can be summarily described as a repository for data”. Nevertheless, the overall idea is that a database is a set of data that has something in common.

Before starting to analyse how databases internally work, it is important to make clear the differences between Data, Information and Knowledge. According to Ackoff (1989, 23) definitions:

Data is a representation of facts that can be usable or not, but it does not have meaning itself. It is normally disorganized and disjointed.

Example: The age of each person in a concrete population.

Information is data that has been given meaning through a relational connection, that can be useful or not. It is properly arranged and classified.

Example: The average age of the country, the average difference between women and men, etc.

Knowledge is the appropriate collection of information whose objective is to be useful.

Example: The tendency of the age and what it implies in terms of demographics, so a government can use it to set a concrete policy.

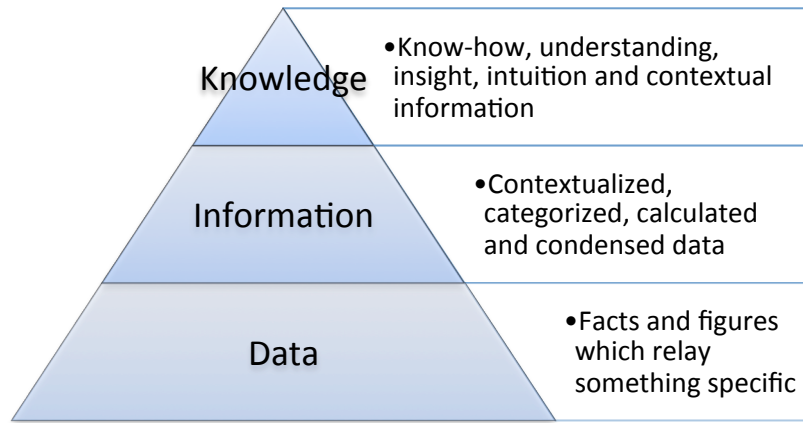


Figure 2.1.1-1. A. Frost, defining data information and knowledge.

And now that the context has been established, the next point to talk about is the structure of the databases. Concretely, I am going to expose two different angles: the structure at the data representation level and the structure at the data management level.

There are many methods and approaches regarding data representation. However, the most common one is the Entity-Relationship Model.

Crane (2011, 2) states that “an Entity-Relationship (ER) model is an abstract and conceptual representation of data.” Therefore, ER modelling is a database modelling method, used to produce a conceptual schema of a system. Silberschatz, Korth & Sudarshan (2010, 7.3) explain that this approach is based on the idea that a database can be simply modelled as a combination of a collection of entities and the relationships among these entities.

An entity is an object that exists and is distinguishable from other objects. These entities have attributes or properties. Finally, entity sets are entities of the same type that share the same properties.

On the other hand, a relationship is an association among several entities.

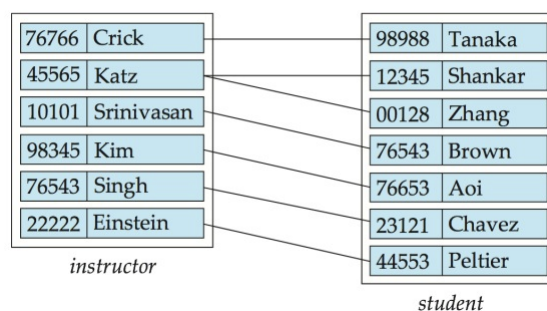


Figure 2.1.1-2. Silberschatz, Korth & Sudarshan, Entity sets and relationships.

In this case, the entity sets are instructor and student, the attributes are the ID and the entities are the persons. The relationship is “advisor”.

This structure is the basis of the real databases. For instance, the following figure recreates a pharmacy system:

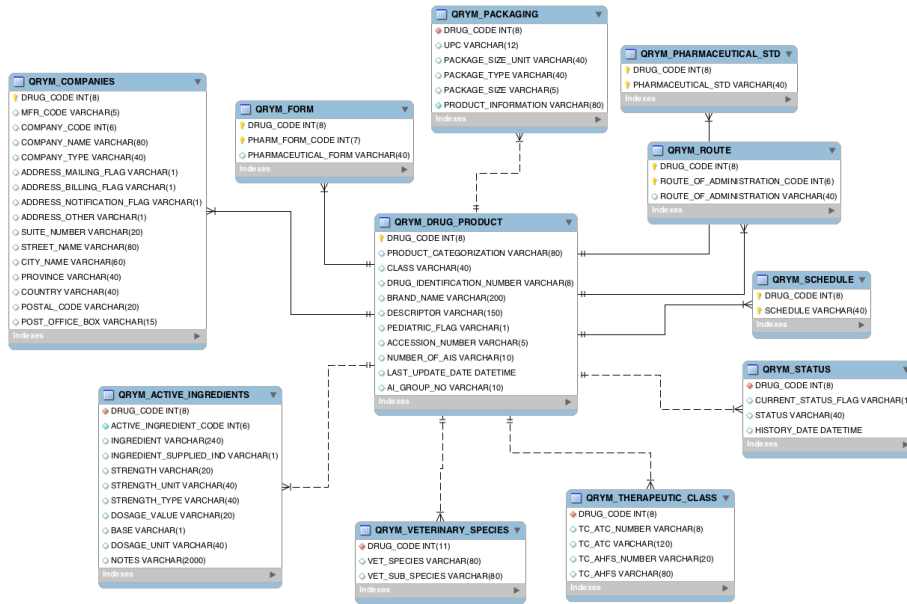


Figure 2.1.1-3. G. McCallum, Health Canada Drug Product Database in MySQL.

Indeed, this way of describing a database is very simple, since it only relates to the very low and basic organization. If we focus on the general architecture of a database, it is necessary to include the management of it - the manner it is possible to get the data and therefore obtain the information. The representation could be something like the following graphic:

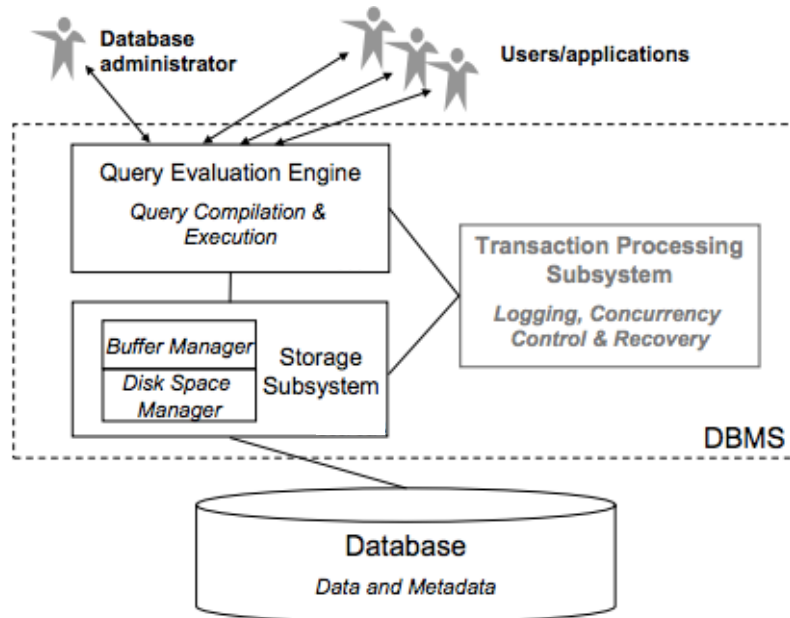


Figure 2.1.1-4. J. Hammer & M. Schneider, Data Base Management System.

Figure 2 shows the map of a typical Database Management System or DBMS.

Hammer and Schneider (2005, 60-2) expose that the components of a DBMS are:

- **Query Evaluation Engine:** Administrators and users interact with the database through different interfaces or applications that are processed by the Query Evaluation Engine, which parses the input and produces an execution plan. The execution refers to the set of instructions that evaluate the input command and create an according response.
- **Storage Subsystem:** Either with queries or with data modifications, there is a need of a system that manages and optimizes the memory space. In this kind of architectures, there is a need for two different memories: the main memory (buffer manager) and the storing memory (disk space manager). The data resides in the second one, but to execute queries or modifications, it has to be firstly transferred to the buffer for processing it and assuring the right location and transfer.
- **Transaction Processing Subsystem:** It is the supporting system that “checks” and controls the activity of the other components, ensuring not only the right execution of queries and modifications, but also the recovering from failure.

To finish with, I will briefly expose the current situation of databases and thereby the implication that they have in our daily life.

Widom (2011), a professor from the University of Stanford, states: “Databases are incredibly prevalent - they underlie technology used by most people every day if not every hour. Databases reside behind a huge fraction of websites; they're a crucial component of telecommunications systems, banking systems, video games, and just about any other software system or electronic device that maintains some amount of persistent information. In addition to persistence, database systems provide a number of other properties that make them exceptionally useful and convenient: reliability, efficiency, scalability, concurrency control, data abstractions, and high-level query languages.”

This importance has had as a result the quickly growth of the knowledge related to this issue, leading to the development of many different kinds of structures for both the managing architecture and the raw data.

Thus, it is a very complex topic and it is not going to be further analysed in this Thesis, as it is not the core point of it. Here there is a brief description of databases in order to merely understand the scope of Master Data and its management.

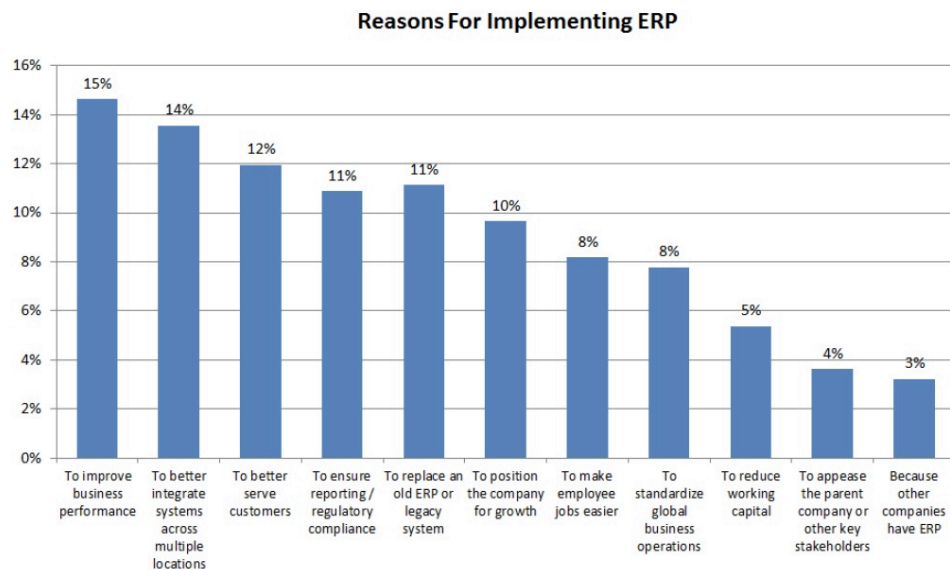
There are many different databases in the market, such as Sql Server, Oracle, Informix, Progress, PostgreSQL, MaxDB, FireBird, etc.

2.1.2 ERP

ERP are the initials of Enterprise Resource Planning. As it usually happens with these relatively new concepts, there are different approaches and definitions of it. In order to present the idea, I will use the introduction of the book Enterprise Resource Planning written Parthasarthy (2007, 1) because it exposes in a very clear manner what ERP really stands for. His book starts as follows:

“Enterprise resource planning software, or ERP, doesn’t live up to its acronym. Forget about planning -it doesn’t do much of that- and forget about resource, a throwaway term. But remember the enterprise part. This is ERP’s true ambition. It attempts to integrate all departments and functions across a company onto a single computer system that can serve all those different departments’ particular needs”.

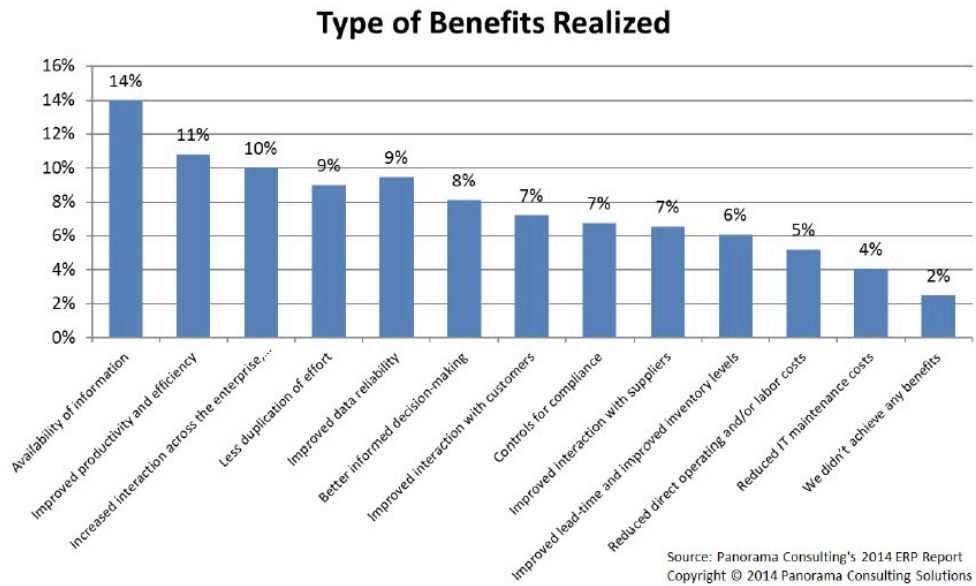
Therefore, ERP is the term used to refer to the software program that manages a single database where the diverse departments, sections and functions of a company are integrated. However, I believe that a good way to expose which are the objectives of ERPs is to analyse the reasons why companies implement them. Panorama Consulting Solutions (2014, 2) conducted a research related to this topic whose results are shown in the following graph:



Graph 2.1.2-1. Panorama Consulting Solutions, Reasons for Implementing ERP.

As it can be seen, the main reasons are both integration and business performance. Note also that the third reason is related to Customers. That shows that the primary application of ERPs is related to improve the service provided to the clients in a company.

Panorama Consulting Solutions (2014, 16) also provide the actual benefits noted by organizations. As it is shown, the main ones are availability of information, improved productivity, increased interaction, less duplication of effort and improved data reliability.



Graph 2.1.2-2. Panorama Consulting Solutions, Type of Benefits Realized.

A more theoretical approach is the one suggested by Shang & Seddon (2000, 1005), who structured the main benefits through the following classification: Operational, Managerial, Strategic, IT infrastructure and Organizational. I believe this is a good way to categorize ERP advantages since it covers every level of the company – from the IT structure to the strategy. The concrete benefits of each dimension are exposed in the following table.

Dimensions	Sub dimensions (21 in total at this stage)
1.Operational	1.1 Cost reduction, 1.2 Cycle time reduction, 1.3 Productivity improvement, 1.4 Quality improvement, 1.5 Customer services improvement
2.Managerial	2.1 Better resource management, 2.2 Improved decision making and planning 2.3 Performance improvement
3.Strategic	3.1 Support business growth 3.2 Support business alliance 3.3 Build business innovations 3.4 Build cost leadership 3.5 Generate product differentiation (including customization) 3.6 Build external linkages (customers and suppliers)
4.IT Infra-structure	4.1 build business flexibility for current and future changes 4.2 IT costs reduction 4.3 Increased IT infrastructure capability
5.Organizational	5.1 Support organizational changes 5.2 Facilitate Business learning 5.3 Empowerment 5.4 Built common visions

Table 2.1.2-1. Shang & P. B. Seddon, Proposed ERP benefits framework.

Hence, it is clear that even if the implantation of an ERP system requires some effort and changes in the company in order to align the processes, it certainly has many advantages. Parthasarthy (2007, 2-4) adds another indirect benefit related to the fact that the ERP vendors are always updating and improving their systems. This happens because the ERP suppliers know that to grow and maintain their businesses they need to embrace the latest technology.

But Parthasarthy (2007, 5) also points out that there are some shortcomings, for instance:

ERPs are rather meant for recording than for planning -they are “a posteriori” systems that are able to extract information from the past data. For this reason, they are a good tool to analyse but they do not provide suggestions about decisions or forecasts.

Furthermore, they are quite rigid in their structure -even if they can be used in very complex processes and workflows, once the application it is configured it is difficult to change it. They have a lack of flexibility in terms of restructuring and reshaping according to the development of the company.

On the economic side, if the ERP used it is not free, the costs derived from its use can actually lead to losses, as it is normally very expensive to implement and maintain.

At this point, I have already exposed what is an ERP and which are its objectives -or rather why companies use it. Now, I will get a little bit deeper in the ERP system itself. Thus, I will answer to internal questions such as “Which are its characteristics?” or “How is its structure?”.

Regarding the attributes of ERPs, Uwizeyemungu & Raymond (2005, 3) expose a complete recapitulation of the main characteristics. The summarizing table is shown below:

Characteristics	Explanatory elements
Integration	Interconnections between functions and hierarchical levels Interaction between the various processes
Completeness (generic function)	Wide range of functions Applicable to various types of firms Connectivity with the outside
Homogenisation	Unique data referential Uniformity of human-machine interfaces Unicity of the system's administration
Real-time	Real-time update and consultation
Adaptability (flexibility)	Capability to follow rule and organisation changes (made possible by parametering)
Openness (evolutionary)	Modularity Portability
Transversality (process-oriented view)	System designed in regard to the business processes necessary to achieve objectives Focus on value rather than authority flows
Best practices	System imbeds best practices in the field
Simulation	Business processes can be simulated

Table 2.1.2-2. S. Uwizeyemungu & L. Raymond, Main characteristics of ERP systems.

Between them, I would like to remark:

- **Integration:** As it has already been exposed, all the departments are linked and therefore the output of one is the input of another.
- **Openness:** They are composed by different “packages” or modules that can be selected according to the requirements of the concrete client.
- **Adaptability:** The modules can be configured or parameterized to adjust them depending on the structure and processes of the concrete business.

These features can be seen in its structure, since it is modular but centralized. There are basically 5 different modules through which it is possible to manage most of the aspects in a company: Supply Chain, Manufacturing, Human Resources, Customer Relationship and Financial.

Nevertheless, these previous modules can be completed with other specialized ones that may be important depending on the concrete focus of the enterprise, such as “Service”, “Sales”, “Project”, etc.

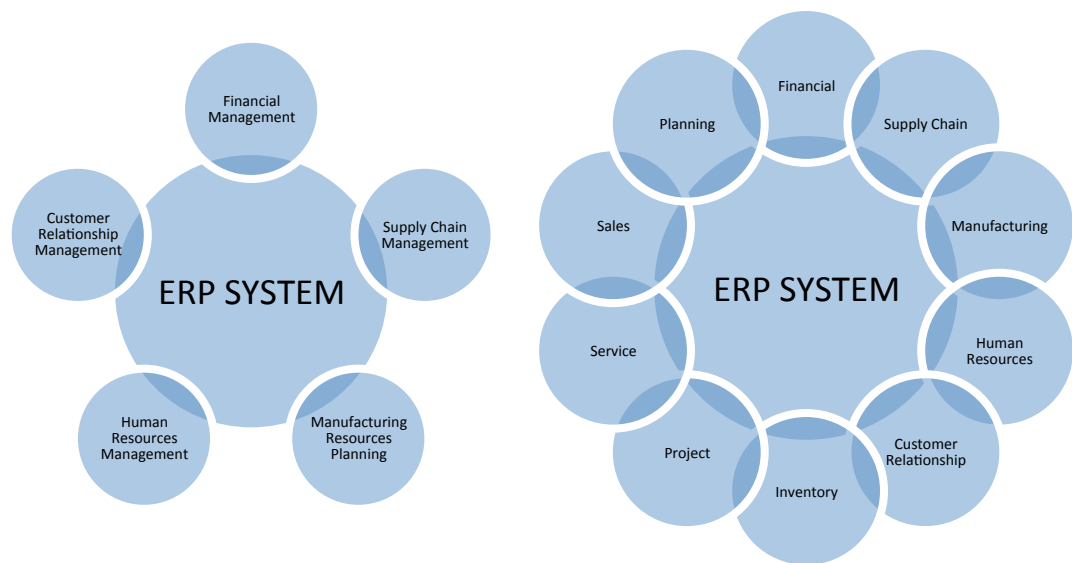


Figure 2.1.2. Delia Rodriguez, ERP modules.

To finish with, I will briefly consider the development of ERPs regarding the history.

Traditionally -and before ERPs began to be used- each department had its own system completely made to satisfy the area’s particular requirements. However, “in the 1980’s a whole trend evolved on company management. Because of increasing interactions between departments, it was no longer as easy to deliver sales orders to the customers cost-effectively and in a timely fashion”. (Van der Hoeven 2009, 3.)

This fact pushed companies to analyse the diverse activities within the organization and structure them as inputs and outputs, what is commonly

called “business processes”. This approach aimed to optimize the process in order to be able to face these increasing interactions. The related model is denominated “process orientation” and Van der Hoeven (2009, 4) states that in this model the business process are put in the centre. Thus, the first objective of ERPs was to combine all those business processes in an integrated program that used a single database to manage the whole company.

A business process is a collection of coherent activities that starts with a trigger and ends with the delivery of one or more products that represent a value for the customer. Note that in this ambit the term “customer” can be external (e.g. a client) or internal (e.g. a department). The most common business processes will be exposed in section 2.2.2.

Nowadays, ERPs have become an indispensable part of many companies. They are considered as the software program that manages the single database where the diverse departments, sections and functions of a business are integrated. This allows the whole company to effectively communicate with other parts, share information, align the different processes and ensure that the information it is unique, correct and precise.

Thus, they are used as a tool both to manage the processes and to get information from them. Consequently, they help in the overall optimization of the organization and they support the decision-making process providing accurate and realistic information.

There are many different ERPs in the market, classified as free ERPs (OpenERP, Openbravo, GNUe, Compiere, Adempiere, etc) and not free ERPs (primarily SAP, Oracle and Microsoft Dynamics).

2.2 Master Data in a company

The basic aim for companies is simply to create value, many times in the form of money and profitability. To reach that goal, they have to be efficient and consequently optimize every single process within their system. ERPs were conceived for this purpose –automate procedures and reduce operations, timing and costs.

The challenge increases when the magnitude of the database spreads and it becomes difficult to guarantee that the data is reliable, precise, appropriate, standardized and unique. In this part of the Thesis, I will show how this problem is approached.

Firstly, I will expose the types of data that are used in companies, paying special attention to Master Data and how to prevent its typical errors.

Secondly, the main processes that are carried out in a company will be described. The objective is to see how they interact with the rest of them from a “business process” point of view. Therefore, the inputs and outputs of all of them will be detailed as well as the information flow.

2.2.1 Enterprise Data

In the section Master Data Concept (1.6), the different kinds of data were defined. Now, I will concentrate on the data used concretely in the business scope –the enterprise data.

To begin with, I will briefly describe data types. It is important to see the differences between them in order to be able to prioritize the actions that will be carried out within the system. After that, I will focus on Master Data, exposing the general methods that are used in companies in order to ensure its quality.

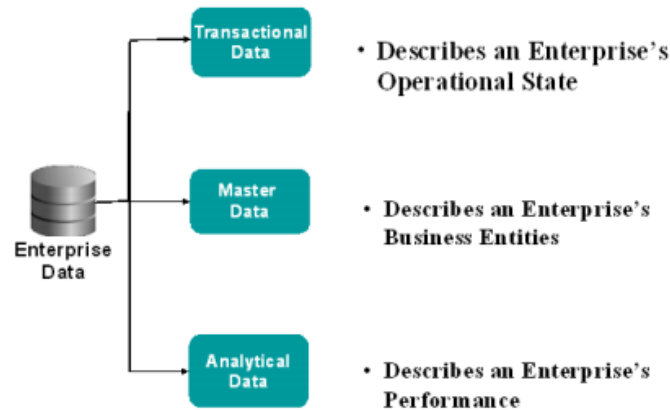
Concerning Enterprise Data, Oracle (2011, 4) explicates that there are three kinds of actual business data: Transactional, Analytical, and Master.

Transactional Data supports the applications that automate a business’ core processes. These applications require significant amounts of data to function correctly, including data about the objects involved and the pure transaction data.

Analytical Data supports decision-making of the business. The most important data of this kind analyses customers, suppliers or products:

- Customer buying patterns are analysed to identify churn, profitability, and marketing segmentation.
- Suppliers are categorized based on performance characteristics over time, for better supply chain decisions.
- Product behaviour is scrutinized over long periods to identify failure patterns.

Master Data represents the business objects that are shared across more than one transactional application. Therefore, it represents the key dimensions around which analytics are done. Technically, Master data creates a unique record of these objects across the whole IT structure as seen in the graph below:



Graph 2.2.1. Oracle, Data in the Enterprise.

To maintain the quality of Master Data, Hillard (2010, 149-151) proposes “the Basic Rules for Master Data” in companies:

- **Unique Description:** The description of Master Data should be unique for the entire entity, which means that every concept should be properly defined and accepted for all the departments.

For instance: The difference to “Client” and “Customer” should be set centrally so every employee in the organization uses the same denomination and therefore they search/create/modify/update data in the correct manner. It also can happen with concrete names, such as calling a certain client “William Black” or “Will Black”. It does not seem to be a big error, but it can lead to inconsistencies in the ERP structure and data.

- **Common Hierarchy:** Every person able to read/write Master Data should know its structure, in order to prevent mistakes and problems in finding the accurate information.
- **Coherent details:** All the Master Data records should have the same level of detail, because it allows using different ways or shortcuts of search and selection over the data.

For instance: If a company wants to contact the customers from a given area, they may use the postal code as a shortcut to look for them. If this rule is not well followed, it can happen that some customers have the address with no postal code and consequently they are not found.

A deeper approach of Master Data will be exposed in section Master Data Management (2.4).

2.2.2 Business Processes

As I have already remarked, Master Data defines the core of the system, because it is shared across different departments in the business.

In this section, the business processes that are carried out in a Manufacturing company will be analysed. The objective is to see how they interact with each other from a Master Data point of view. Therefore, the inputs and outputs of all of them will be detailed, showing therefore which is the crucial data in the overall system.

2.2.2.1. Aggregate Planning – AP

Its objective is to determine the production plan to pursue the strategy of the company. It establishes the quantity of the family products that need to be produced in the following 6-18 months. As Heizer & Render (2001, 540) explain, “it is concerned with determining the quantity and timing of production for the intermediate future”. According to it, it is possible to balance and set the capacity requirements of the factory, the inventory planning and the human resources management.

Basically, it is the way to equal the demand and the production of the company, taking into account the current orders and further forecasts. Normally, the first goal of this balance is to minimize cost over the planning period. However, Heizer et al. (2001, 540) remark that other strategic issues more important than low costs can be involved, such as to smooth employment levels, to drive down inventory levels or to meet a high level of service.

There are two main strategies: Capacity options and Demand options. Capacity options try to absorb the demand fluctuations, whereas Demand options attempt to smooth the demand or modify it.

Thus, Capacity options include Changing inventory levels, Hiring, Firing, Extra hours, Idle hours, Subcontracting, Using Part-Time workers... Depending on which of the previous methods a business uses, the AP can be Level Strategy (maintaining the same production rate and employment level) or Chase Strategy (adapting month by month the working force and consequently the production rate).

On the other hand, Demand options are basically Influencing Demand and Back Ordering during high-demand periods. Hill (2000, 284-285) exposes further ways of managing the demand, such as: Altering price levels, advertising, develop complementary products, arrange reservations and appointments, use fixed service schedules.

The general AP looks like the following table. In this case a Capacity option with a Chase Strategy is been used, since the employment level varies from period to period if required.

Note that in this example there are only four months to make it simpler.

	January	February	March	April
Product Demand Forecast	6350	8800	6700	6400
Required Production	6350	8800	7500	6400
Required Hours	12700	17600	15000	12800
Working days	20	20	22	20
Required workers	79,375	110	85,227	80
Real workers	80	100	86	80
Production hours	12800	16000	15136	12800
REGULAR HOURS				
Production in regular hours	6350	8000	7500	6400
Regular hours cost	152400	192000	180000	153600
FIRING				
Firing	0	0	14	6
Firing cost	0	0	42000	18000
HIRING				
Hiring	0	20	0	0
Hiring cost	0	40000	0	0
IDLE HOURS				
Idle hours	100	0	136	0
Idle hours cost	1600	0	2176	0
INVENTORY				
Backlog Inventory	0	-800	0	0
Possession Inventory	0	0	0	0
Inventory total cost	0	16000	0	0
TOTAL COSTS/MONTH	154000	248000	224176	171600

Table 2.2.2.1 Delia Rodríguez, Aggregate Production.

➔ The Master Data needed to perform this process is the products' forecasts, which is usually provided by the Sales department.

➔ The Master Data from this process that is used in other sections is: Product Demand forecast, Real workers and Production in regular hours. The rest of the parameters are just an approximation, and the real values that will be used to make further calculations should be taken from the Master Production Schedule.

It is important to remark how AG are linked and how they finally lead to get the MRP, as described in the figure below:

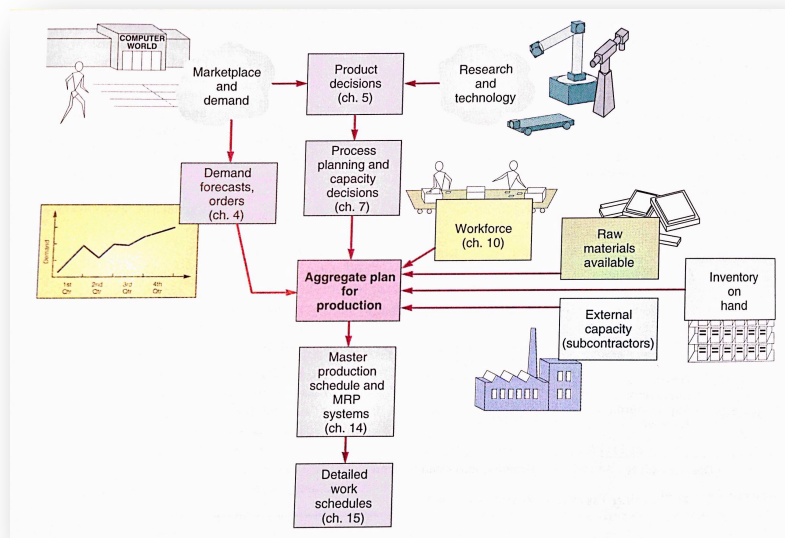


Figure 2.2.2.1. J. Heizer & B. Render, Relationships of the AG.

2.2.2.2. Master Production Schedule – MPS

Hill (2000, 351-352) explains that MPS is “a management commitment to provide or produce certain quantities of services or products in particular time periods”. It is basically the disaggregation of the AP: it is the same information but divided by products (instead of family products), and by weeks (instead of months).

As a result, it provides the planning of the products, their required quantities and when they need to be finished and prepared to be sent for the following 2-4 months. Hill sums up it as “an anticipated schedule of finished services and products”. He points out that it is merely a statement of operations output and not of market demand, because it is not a forecast and should not be confused with it. After doing the MPS, a capacity check will prove if the production that was set in the AP is viable or not.

As I said, MPS provides the schedule that needs to be followed. For example, in this concrete situation, 2000 units of product 1 should be finished in the second week of January.

	% total	Batch size	Current production
P1	70%	2000	1900
P2	20%	680	1000
P3	10%	300	600

	January	February	March
AP - Production in Regular h	6350	8000	7500
AP - P1	4445	5600	5250
AP - P2	1270	1600	1500
AP - P3	635	800	750

	January				February			
	1	2	3	4	5	6	7	8
AP - P1	1111	1111	1111	1111	1400	1400	1400	1400
Inventory Excess	0	789	1678	567	1456	56	656	1256
Current production	1900	0	0	0	0	0	0	0
Net requirements	-789	322	-567	544	-56	1344	744	144
Master Production Schedule	0	2000	0	2000	0	2000	2000	2000

Table 2.2.2.2 Delia Rodríguez, Master Production Schedule.

Note that in this example I have just done it about one product (instead of three) and only for two months (instead of 3 or 4) to make it simpler.

➔ The Master Data needed to perform this process is: the required production calculated in the AG, the proportions of each concrete product out of the whole family product from the Sales department and the current production of each product from the Production department.

➔ The Master Data from this process that will be needed in other sections is basically the last row, the final schedule.

2.2.2.3. Bill of Materials – BOM

According to Kumar & Suresh (2009, 219) definition, “a Bill of Materials is a listing of all the materials, components and subassemblies needed to assemble one unit of an end item”. It is therefore a description of the components of a concrete product, explaining clearly the quantity of each of them and the order they need to be processed.

It is a very important part, especially in order to plan the Materials’ Requirements, because it permits to check which components are needed to produce the quantities determined in the MPS.

The structure of a BOM can be either Indented Bill of Materials or Product tree. The comparison between them is shown below:

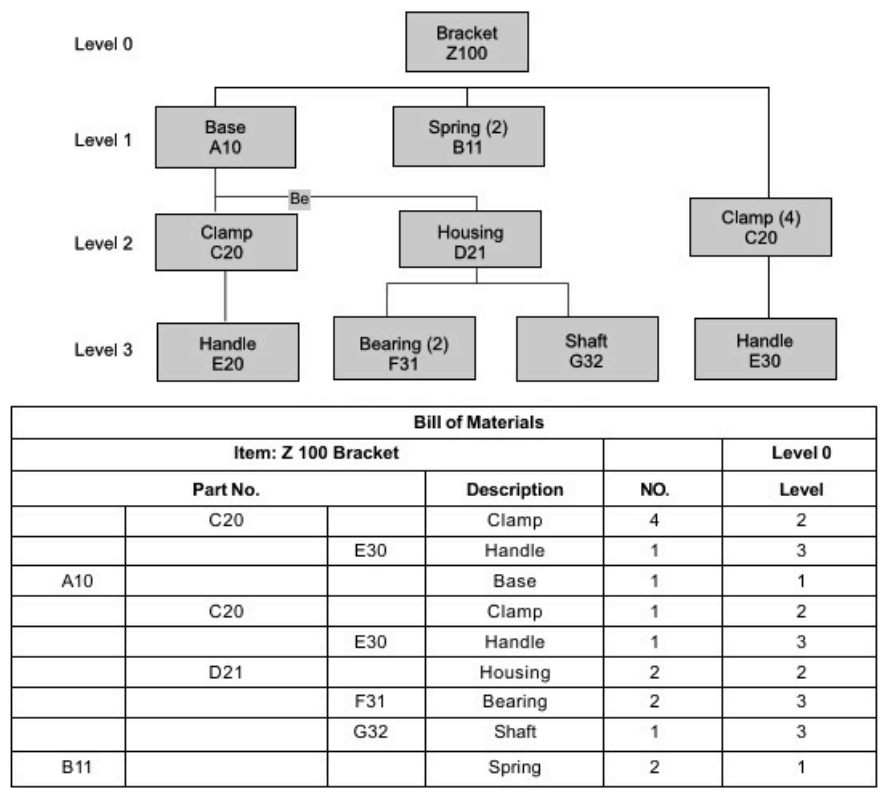


Figure 2.2.2.3-1. S. Kumar & N. Anil Suresh, Bill of Materials.

As an example to explain the BOM further, I will use the tree structure:

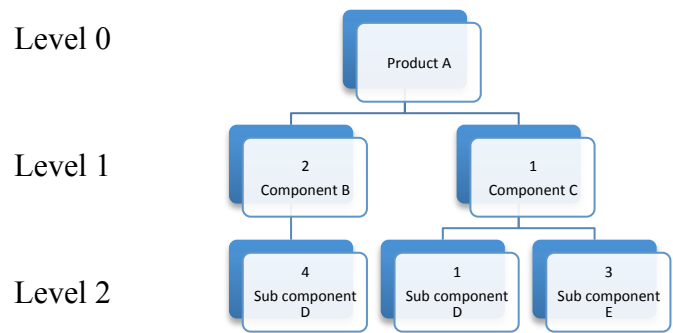


Figure 2.2.2.3-2. Delia Rodríguez, Bill of Materials tree.

In this example, to produce 1 unit of product A, 2 units of B and 1 of C are required. Furthermore, to produce 1 unit of B, 4 units of D are required; and to produce 1 unit of C, 1 unit of D and 3 of E are required.

Note that the same component can be a part of different products (Sub component D is a part of B and C).

Heizer & Render (2001, 579) add that BOM are not only to specify requirements “but also useful for costing, and they can serve as a list of items to be issued to production or assembly personnel”. In this case, they are called Pick Lists.

➔ The Master Data needed to perform this process is the components or subassemblies of each concrete product. This information can be provided from the Design, Manufacturing or Production department.

➔ The Master Data from this process that will be needed in other sections is: the kind of components needed to produce a concrete product and their quantities.

2.2.2.4. Inventory Records

It is the main source of information about the status of the company’s warehouses, it is “knowledge of what is in stock” (Heizer & Render 2001, 579). Hill (2005, 355) defines the records as “all transactions and inventory balances”, and remarks that the main requirements of it are accuracy and timeliness, since they are critical to the running of a MRP system.

Inventory Records basically consist of an updated table that includes the characteristics about the type of storing of every material and component, as well as the quantity in stock and the programmed receptions.

Component	Safety Stock	Order method	Delivery time	Quantity in stock	Programmed receptions
Prod. A					
Comp. B					
...					

Table 2.2.2.4. Delia Rodríguez, Inventory records.

➔ The Master Data needed to perform this process is diverse products’ information: Safety Stock and Quantity in Stock from the Production department and Order method, Delivery Time and Programmed Receptions from the Purchasing department.

➔ The Master Data from this process that will be needed in other sections is actually every field. The point is that many departments are related to the warehouses, for instance Sales, Purchasing, Production, Logistics...”

2.2.2.5. Materials Requirement Planning – MRP

Kumar & Suresh (2009, 217) define a MRP as a “computer-based technique for determining the quantity and timing for the acquisition of dependent demand items needed to satisfy the MPS”. Dependent demand refers to the fact that it is derived from the demand of other items.

Consequently, it is a tool that permits to reduce inventory costs and improve scheduling effectiveness through a continuous and precisely identification of what, how many and when components are needed.

This process is one of the most important ones in the whole Manufacturing system, because it allows the company to set its requirements of components in order to be able to produce the final products. Hence, it is a tool to schedule manufacturing, purchasing and delivering.

The inputs of the MRP are basically the MPS (in order to have the final product demand), the BOM (in order to know the type and quantity of the necessary materials to produce the final products) and the Inventory Records (in order to know the current quantities of the materials).

The output of the MRP is a plan with the information about when it is necessary to order a component from a Supplier, when it should arrive, when it is necessary to start producing a final product, when it should be finished, etc.

Kumar & Suresh (2009, 219) provide a schema with the essential information surrounding a MRP:

Inputs	Outputs
<ul style="list-style-type: none"> • MPS of end items required. • Inventory status file of on-hand and on-order items, lot sizes, lead times, etc. • Product structure (BOM) file of what components and subassemblies go into each end product. 	<ul style="list-style-type: none"> • Order release data to CRP for load profiles • Orders to purchasing and in-house production shops. • Rescheduling data to MPS. • Management reports and inventory updates.

Table 2.2.2.5-1. S. Kumar & N. A. Suresh, Essential inputs and outputs.

Firstly, the MRPs about the final products are fulfilled and later on, from their schedules, the MRPs about the related components. I will give a simple example to understand the whole sequence, where the final product is “Product A” and its requirements are the same as those shown in the BOM tree (section 2.2.2.3). The initial MRP only includes the information from the Inventory Records and from the MPS, and it is coloured in green. The cells that need to be filled in order to get the final schedule are coloured in blue. The required information to continue to the next level is “Planned Orders Reception” and “Planned Orders Launching”.

Further information required:

Order method: Batch to batch (any quantity can be ordered)

Delivery time: 1 week

Safety Stock: 350

Week	January				February			
	1	2	3	4	5	6	7	8
Product A - Gross Requirements	0	1500	0	1500	0	1500	1500	1500
Stock	300	1800	350	350	350	350	350	350
Programmed Receptions	1500	0	0	0	0	0	0	0
Net Requirements	0	50	0	1500	0	1500	1500	1500
Planned Orders Reception	0	50	0	1500	0	1500	1500	1500
Planned Orders Launching	50	0	1500	0	1500	1500	1500	0

Table 2.2.2.5-2. Delia Rodríguez, MRP Product A

Secondly, a Requirements Explosion is carried out. A Requirement Explosion is “the breaking down of parent items into component parts that can be individually planned and scheduled” (Kumar & Suresh 2009, 218). Therefore, it is done to get to know the quantities of the subsequent materials, in this case the Component B and Component C. It is made simply multiplying the Planned Orders Launching according to the proportions given by the BOM.

Week	January				February			
	1	2	3	4	5	6	7	8
Gross Requirements from Product A (x2)	100	0	3000	0	3000	3000	3000	0
External Demand	0	0	0	0	0	0	0	0
Comp B - Total Gross Requirements	100	0	3000	0	3000	3000	3000	0

Table 2.2.2.5-3. Delia Rodríguez, Component B Requirements Explosion.

Note that in this example the Component B is only required by the Product A because there is only one final product. Usually, a concrete component is needed for more than one product and therefore the Total Gross Requirements are the addition from all the requirement’s sources.

The MRP of Component B has exactly the same structure than the one of the Product A as it can be seen below.

Further information required:

Order method: Minimum lot size 3000 (only orders above 3000)

Delivery time: 2 weeks

Safety Stock: 400

Week	January				February			
	1	2	3	4	5	6	7	8
Component B - Gross Requirements	100	0	3000	0	3000	3000	3000	0
Stock	450	650	650	650	650	650	650	650
Programmed Receptions	300	0	0	0	0	0	0	0
Net Requirements	0	0	2750	0	2750	2750	2750	0
Planned Orders Reception	0	0	3000	0	3000	3000	3000	0
Planned Orders Launching	3000	0	3000	3000	3000	0	3000	0

Table 2.2.2.5-4. Delia Rodríguez, MRP Component B.

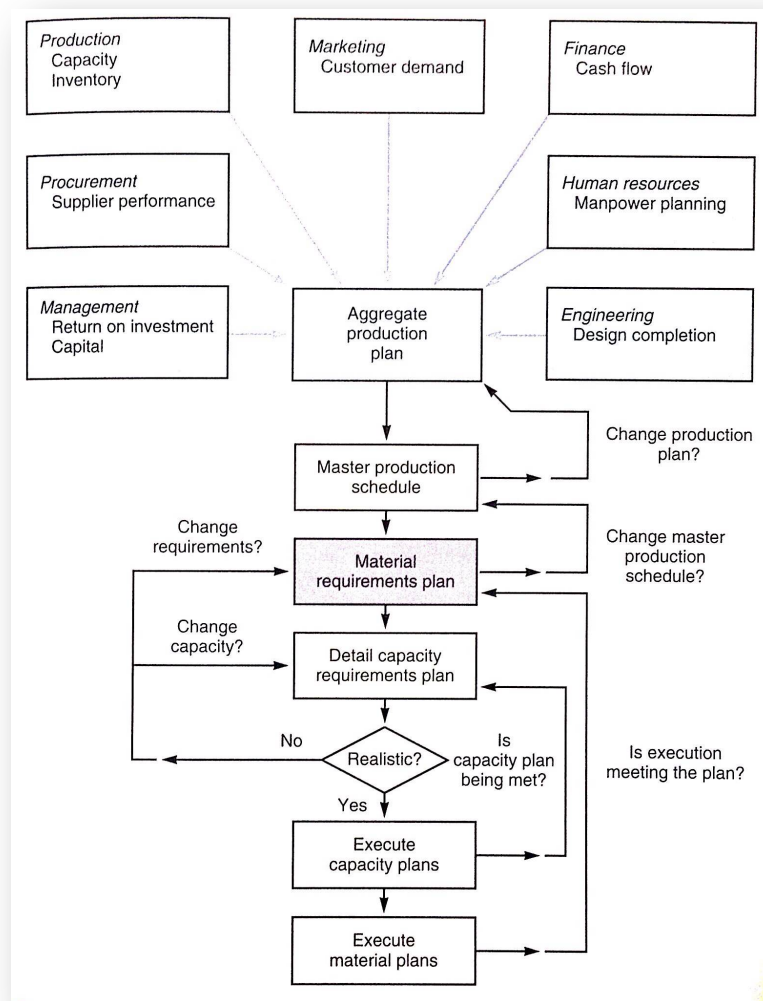
At this point, the Requirements Explosion and its related MRP of Component C should be carried out. Thus, the level 1 of the BOM will be completed and the next step would be to continue with level 2. The process has to be fulfilled for each material in the system.

It is important to remark that MRPs are not only for materials that are purchased from outside suppliers, but also for the final products and other intermediate items that are produced internally.

➔ The Master Data needed to perform this process is: the required and scheduled production from the MPS, the products' components and quantities from the BOM and the materials' stock from the Inventory Records.

➔ The Master Data from this process that will be needed in other sections is the final schedules, given in the Planned Orders Reception and the Planned Orders Launching of every product and component.

Regarding MRP, it is important to summarize the processes that have already been studied that surround this process. A general diagram of the whole planning process is shown below:



Graph 2.2.2.5. J. Heizer & B. Render, The Planning Process.

2.2.2.6. Quality

Six Sigma is a method that tries to ensure the appropriate quality of the products. As Burton & Sams (2004, 33) state, “it is a data-driven methodology that strives for perfection and value added in every aspect of the total value chain”.

They further expose that Six Sigma is based on statistics, due to the fact that most of the productive processes follow a Gauss Distribution. In this distribution the bulk of the pieces are around the average and between the upper and lower limits, where the variance of the items is acceptable. The statistical objective is “to drive down process variation so that plus or minus six standard deviations lie between both the upper and lower specification limits”.

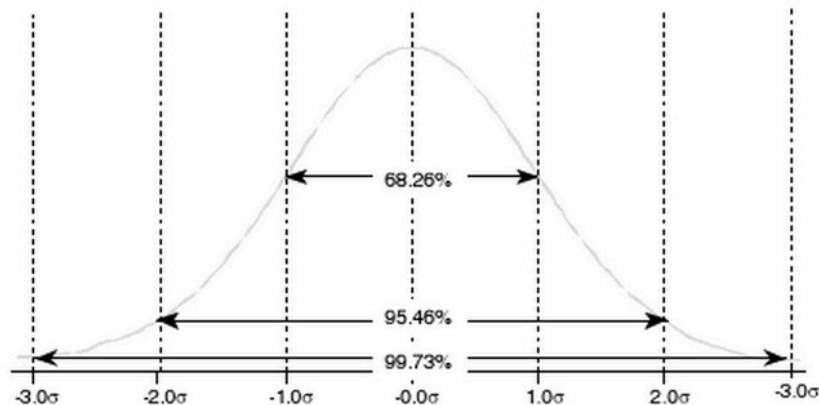


Figure 2.2.2.6. T.T. Burton & J. L. Sams, Gauss Distribution.

The method basically consists of recording the measure of every product and place those marks in a temporal graph. After a short period, the average is calculated and, through some statistic calculus, the areas “ 2σ ” and “ 3σ ” are defined. In addition, the range graph is also carried out in order to study the fluctuations in terms of variability.

At this point, both graphs are analysed, beginning with the range’s one, because it clearly shows whether the process is stable or not:

- If there are marks over the 3σ range limit, then the process is very unstable and probably there are defective pieces. The production should be stopped and all the items in the given batch should be inspected.
- If there are marks below the 3σ range limit, then the process’s variability is decreasing and the causes should be analysed to implement them.
- If there are many points in one side of the 2σ range limit, then the process can continue, but it should be concurrently checked.

After this analysis, the average’s graph is studied in order to check if there is any anomaly in the process, but if there are marks over/below the limits the process does not need to be stopped.

- ➔ The Master Data needed to perform this process is the products' information from the Production department.
- ➔ The Master Data from this process that will be needed in other sections is the results of the analysis –if there is any quality matter it should be immediately communicates to other departments, such as Production or Sales.

2.2.2.7. Routing

According to Dickersbach & Keller (2011, 129), routing “defines how a component or input material is to be converted from its original status into the final status”. Therefore, it is the process that selects the best or optimized paths in a production system.

It is a part of work scheduling, and during this process there are different tasks that need to be done:

The initial part consists of including the product information provided by the product development and design departments into instructions for production and assembly.

After that, the actual routing begins:

- Definition of the concrete sequence of operation series to transform an input to its final status.
- Machine selection to each operation.
- Assignment of resources/tools to each operation.
- Determination of the standard time and standard values. This basically means identifying target times for carrying out the individual operations.

Dickersbach et al. (2011, 131) add that routing management is very relevant because “it efficiently supports archiving”. Those archives permit to have organized classifications that are used to store descriptive information on work-piece groups or similar production processes. This can be very useful when setting up a new routine or when adapting the existing ones.

- ➔ The Master Data needed to perform this process is the products' information from the Design and Production department.
- ➔ The Master Data from this process that will be needed in other sections is basically the whole schedule for the Production department.

2.3 Implications and consequences

After having explained the diverse processes that take place in a Manufacturing company to satisfy the customer demand, it is the moment to analyse how the corresponding Master Data affects the whole system in a tremendous way.

Thus, the diverse consequences linked to poor data quality and their results will be described in this section. Moreover, I will explain how these problems and errors emerge –what are the sources of poor data quality.

2.3.1 Data Quality

Before proceeding to evaluate the impacts of poor Master Data within a company, it is necessary to clarify what exactly defines the quality of the data –to improve Data Quality, it is essential to understand what it means to its consumers.

There are many studies where the topic Data Quality has been discussed. There is not an unique definition, but an accurate one would be “Data is of high quality if it fits for its intended use in operations, decision-making, and planning. Data is fit for use if it is free of defects and possess desired features.” (Redman 2001, 241.)

Nevertheless, it is very difficult to have data totally accurate. Hence, “the real concern with data quality is to ensure not that the data are perfect, but that they are good enough for the organisation to make appropriate and reliable decisions” (Kerr, Norris & Stockdale 2007, 1017).

Consequently, the term “quality” applied to data is subjective, because it is bonded to the concrete organization/person’s requirements.

For this reason, some scholars have tried to get deeper in the field in order to provide more concrete definitions. Haug, Zachariassen & Van Liempd (2010, 168-193) did a research where they compiled some of the main studies of the topic, highlighting:

- Ballou and Pazer divide data quality into four dimensions: accuracy, timeliness, completeness, and consistency. The accuracy dimension is the easiest to evaluate, as it is merely a matter of analysing the difference between the correct value and the actual value used. Timeliness and completeness are also tested in an unproblematic manner. On the other hand, the evaluation of consistency is more complex, since this requires two or more representation schemes in order to be able to make a comparison.
- Wand and Wang limit their focus to intrinsic data qualities, of which they define four intrinsic dimensions: completeness, unambiguousness, meaningfulness and correctness.

So far I would like to remark that, even if it is obvious that Data Quality depends on the requirements of the concrete user, most of the approaches focus on the internal characteristics of the data. Therefore, I think that these academic researches are very precise and complete, but they may have a lack of practical application.

A different proposition is the one that Wang did along with Strong (1996, 5-33). In this case, the objective was to compile data from the current studies, company problems, etc., in order to get a realistic list of data characteristics. Their study is divided in two surveys:

The first one produced a list of possible data quality attributes, attributes that came to mind when the data consumer thought about data quality.

The second survey assessed the importance of these possible data quality attributes to data consumers. The importance ratings from the second survey were used in an exploratory factor analysis to yield an intermediate set of data quality dimensions that were important to data consumers.

The outcomes from the survey are concentrated in a list of 20 dimensions of the Data Quality. Those characteristics were gathered into 4 target categories:

Target category	Dimension
Accuracy of data	Believability Accuracy Objectivity Completeness Traceability Reputation Variety of Data Sources
Relevancy of data	Value-added Relevancy Timeliness Ease of operation Appropriate amount of data Flexibility
Representation of data	Interpretability Ease of understanding Representational consistency Concise representation
Accessibility of data	Accessibility Cost-effectiveness Access security

Table 2.3.1. R. Wang & D. Strong, Four Target Categories for the 20 Dimensions.

I personally think that this categorization is very complete since it does not only covers the basic concepts that surround the term “quality” (such as accuracy, completeness, etc.) but also consider other issues that affect companies. For instance, value-added, cost-effectiveness or appropriate amount of data are totally bounded to the business scope, and therefore to the aim of optimizing the system.

Furthermore, based on those dimensions, they developed a model with the following structure, which it is a comprehensive definition of Data Quality:

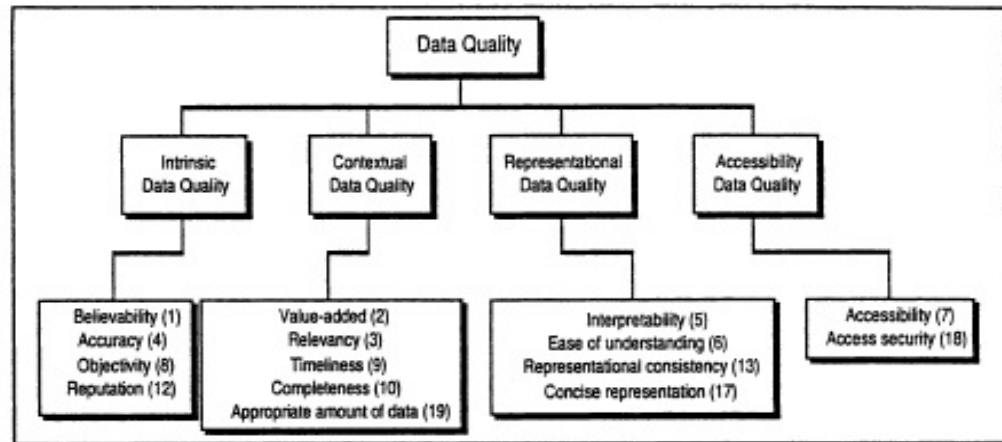


Figure 2.3.1. R. Wang & D. Strong, Conceptual Framework of Data Quality

I would like to remark that in this approach only one of the four main divisions of the classification refers to the internal characteristics of the data. The other concepts are primarily practical, and therefore many of them apply to the fact that data is managed by users (e.g. interpretability, ease of understanding or accessibility). For all these reasons, I believe that this categorization is a very good tool that could be used to analyse and check the data quality of a real company's system.

There are many discussions about the subject for the reason that is a relatively new topic that can be approached from different angles and perspectives. Nevertheless, I think that the general idea should be that data needs to fulfil the concrete requirements of the user. To reach that aim, it should be considered in a double way: Firstly, ensure that is accurate and reliable. Secondly, analyse the social aspect and align it with the data quality plan, so the result is not only correct, but also optimal.

2.3.2 Impacts of poor quality data

The origin of the Master Data Management problems is due to the fact that data quantity has been exponentially growing during the last decades and, as a direct consequence, its complexity. Watts & Shankaranarayanan (2009, 202-211) explain that, in organizations today, the risk of poor information quality is becoming increasingly high as larger and more complex information resources are being collected and managed.

Furthermore, it is important to understand that businesses have switched from an industrial to an information economy base in last decades. Companies now compete on the ability to absorb and respond to information, not just manufacture and distribute products. Hence, it is essential to manage well data, as it is the core of information and this management can be the critical factor to achieve success.

Consequently, poor quality data can lead to severe and harmful results on the health of a company. Specially when it is not identified and corrected early on, it can contaminate all downstream systems and information assets, jacking up costs, jeopardizing customer relationships, and causing imprecise forecasts and poor decisions.

(Eckerson 2002, 5.)

There are many negative implications that defect data can carry in an organization. Haug, Zachariassen & Van Liempd (2010, 168-193) summarized the research of Kahn, Pipino and Redman, obtaining the following major impacts:

- Less customer satisfaction.
- Increased running costs.
- Inefficient decision-making processes.
- Lower performance.
- Lowered employee job satisfaction.
- Increased operational costs.

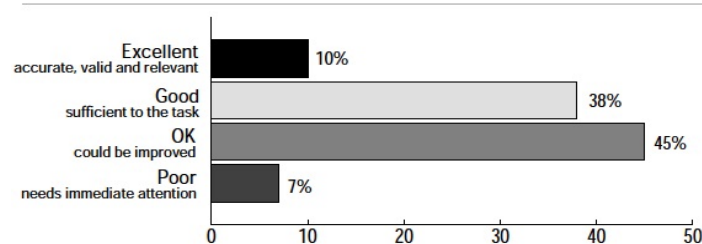
Therefore, based on theoretical studies about the topic, these are the most significant implications of poor data quality. However, now I would like to compare these results with what actual surveys reveal. Thus, I will use a research of The Data Warehousing Institute.

TDWI conducted a survey where 286 respondents from different companies in U.S.A. answered to diverse questions, including issues as “Status of Data Quality Plans?”, “How does your firm think its data quality is?”, “How is the data quality in reality?” or “Does your organization need more education in the area?”.

Some of the graphs and conclusions referring to the results of those questions are shown below.

The first part of the survey focuses on the perception of the quality of the data. As it can be seen, almost half of the organizations believe that their data is just “OK” and that it could be improved. Note that this option represents that the quality is less than “sufficient to the task” (which is the option “Good”).

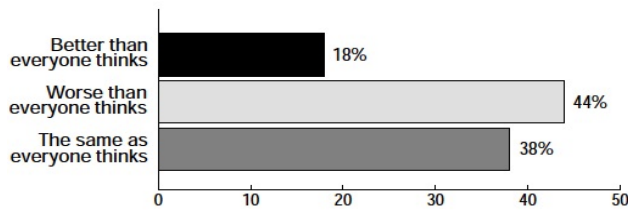
Our Firm Thinks Its Data Quality Is:



Graph 2.3.2-1. The Data Warehousing Institute, Data Quality Survey, Company beliefs about DQ.

However, the most alarming fact is that actually, the quality is usually even worse than what people think it is. This is especially worrying because of the already mentioned idea that exposed that the effects of defective data when it is not identified are much more severe, since they quickly spread over the whole system.

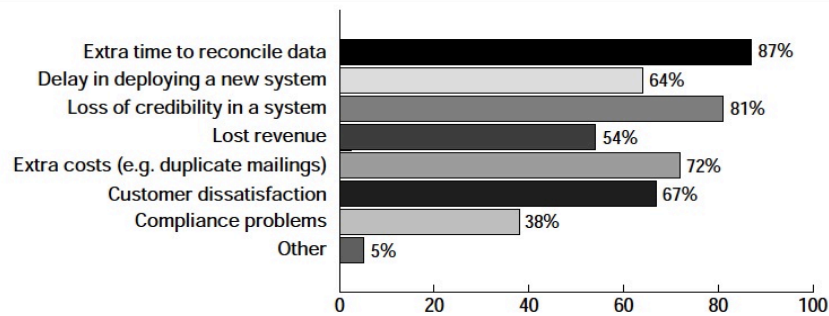
In Reality, the Quality of Our Data Is:



Graph 2.3.2-2. The Data Warehousing Institute, Data Quality Survey, Reality DQ.

About the problems caused by the defective data, most of them are related either to costs or to time delays. In fact, delays are also a concrete type of costs, so it could be said the main effect of poor data is primarily higher costs. This will be further discussed in section Costs of poor data quality (2.3.3.)

Problems Due to Poor Data Quality



Graph 2.3.2-3. The Data Warehousing Institute, Data Quality Survey, Problems due to DQ.

In the graph, I believe all the problems are understandable and thereby there is no need for further explanation. Nevertheless, I would like to concretize the “extra costs”. These costs can appear from very different reasons: duplicate mailings, excess inventory, inaccurate billing, lost discounts, delays in deploying new systems, etc.

According to these findings, Eckerson (2002, 9-10), a member of the TDWI, states that the two main implications of poor data are extra time to reconcile data and loss of credibility in the system, and explains that those two problems are deeply linked since the inability to reconcile data between the data warehouse and the source systems causes end users to lose confidence in the data warehouse.

His final conclusion points out the ultimate impact of poor Data Quality, that is basically the inability to make decisions: “Without good data, organizations are running blind. They can’t make good decisions because

they have no accurate understanding of what is happening within their company or the marketplace. They rely on intuition, which is dangerous in a fast-moving market with nimble competitors and finicky customers.”

2.3.3 Costs of poor quality data

As it has been already exposed, the principal implication of defective data is costs. It is a really alarming issue because, according to Redman (1998, 80), the total cost of poor data quality is around the 8-12% of revenue range. What is more, he estimates that 40-60% of a service organization's expense is a consequence of poor data.

As this is an actual and real problem, it is constantly developing, and much of that progress is carried out by technical consultancies. This kind of consultancies advice companies in subjects related to their data structure and quality with the objective of decreasing the corresponding costs.

Hence, in this topic the researches are not only from academic or theoretical sources. Rather, there are parallel studies in the private and the educational field. Therefore, I will try to provide both points of view throughout this section in order to have the whole picture of the costs of poor data quality.

Costs can come from different sources, have different and therefore they can be structured in different ways.

A good structure is the one provided by Eppler & Helfert (2004, 316), that considers two main costs: the ones caused by low Data Quality and the ones of assuring the Data Quality.

Low Data Quality costs are further subdivided into Direct and Indirect.

- Direct costs are those negative monetary effects that arise immediately out of low data quality, such as: the costs of verifying data because it is of questionable credibility, the costs of re-entering data because it is wrong or incomplete, and the costs of compensation for damages to others based on bad data.
- On the other hand, Indirect costs are those negative monetary effects that arise, through intermediate effects, from low quality data. For instance, the costs derived from a bad reputation, the costs incurred by sub-optimal decisions based on poor data, etc.

Assuring Data Quality costs are further subdivided into Prevention, Detection and Repair. Their objective is basically to lower the costs provoked from low data quality. Häusermann & Pilz (n.d., 1-3), from Business Intelligence Consultants, explain these concepts further:

- The necessary measures in order to reduce errors from the beginning are the Prevention costs. For instance, improvement of data capture through plausibility tests, documented standards, better training of personnel or better coordination between sub-processes.

- Detection costs are the methods that lead to the discovery of already existing errors in the data. The most common tool is database analysis.
- Repair costs include all activities that are necessary to correct the detected errors in the databases. Examples are elimination of duplicates, insertion of missing data, data cleaning, etc.

A blueprint of the classification:

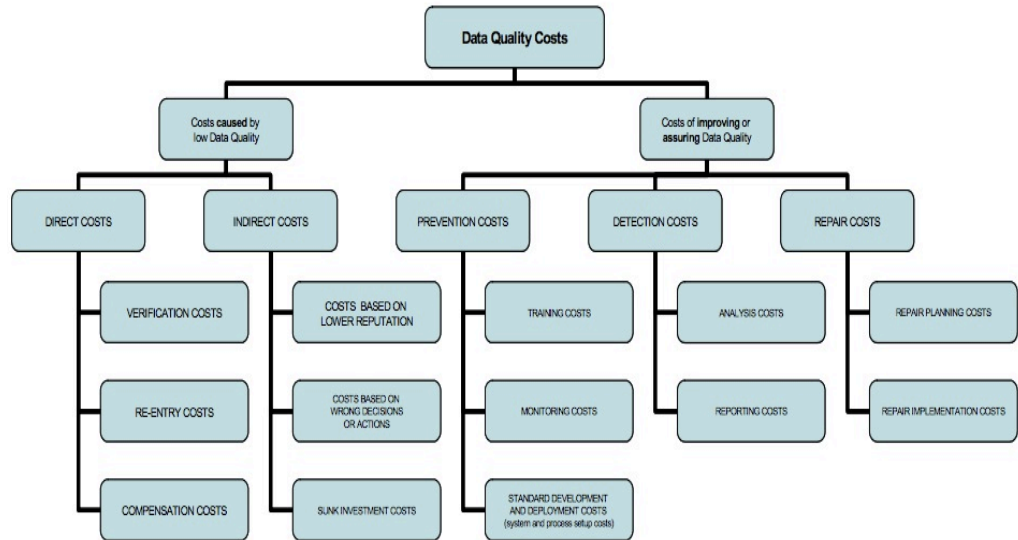


Figure 2.3.3-1. M. J. Eppler & M. Helfert, A data quality cost taxonomy.

However, Eppler et al. (2004, 317) explain that the previous classification is merely conceptual and it does not provide a method to analyse the real costs in a company. The reason is that in a business the importance resides not in the data, but in the information. Thus, the result will be a failure even when the data is correct, if the derived information is not.

Consequently, they expose a framework to manage the cost of Information Quality.

In the diagram there are three sections (data entry, data processing and data use), which are further subdivided into Low Information Quality Costs and High Information Quality Costs.

The main idea is that the lower costs should reduce the upper costs - Investments in preventive costs should lead to reductions in corrective costs and costs in process improvements should lead to cost reductions in re-processing.

Thus, the upper part of the framework must be continually compared with the lower one to determine the overall cost impact of maintaining the data quality.

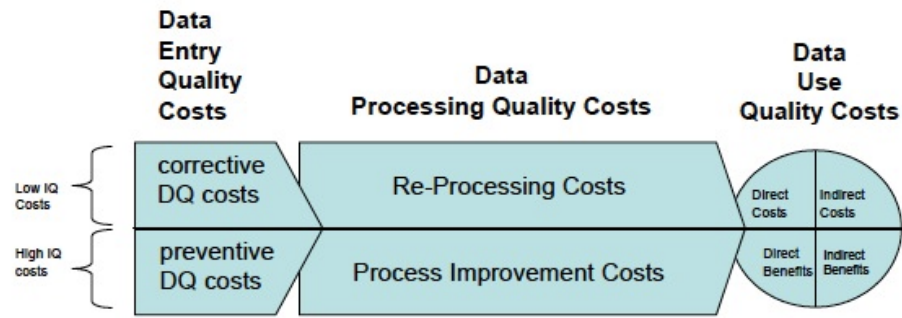


Figure 2.3.3-2. M. J. Eppler & M. Helfert, A framework for a classification of data quality costs.

As I exposed at the beginning of this section, I want to have both theoretical and practical approaches to this topic in order to have a rich idea. Thus, I will now explain a totally different point of view of the costs of poor data quality, based on the experience of current consultants.

The structure proposed by Dietze & Fischer (2013), members of the Roland Berger Consultancy focus on the most common impacts that real companies face. They highlight the following costs:

- Location information: Mistakes in geographical data that lead to errors in logistics planning and thereby to higher freight costs.
- Supplier data: Duplicates in the supplier master mean that incorrect sales figures are calculated for a specific supplier. This can be a drawback when negotiating new terms.
- Customer data: Missing or incorrect information in the customer hierarchies, for instance regarding the relationships between parent companies and subsidiaries, could result in a lack of information about the client's credit. This could drive to unpaid that would provoke losses.

Furthermore, they point out that, apart from these direct costs, there are indirect costs derived from them. The reason is that to correct such mistakes, departments and IT normally require considerable human resources that have to leave other areas of the organization. This can incur serious costs, since it can lead to a lack of personnel in other important strategic projects in which the company is involved.

To finish with, I will expose a much simpler classification proposed by English (1999, 24). He believes that the costs of poor data quality are either Process Costs or Opportunity Costs. Process costs are those linked to the re-execution of a process due to data problems, while Opportunity Costs are those due to lost and missed revenues.

Nevertheless, he also remarks that costs of poor data quality are strongly dependent on the context and the customer of the data. That is the reason why evaluating the costs can be really difficult, as the same data value and corresponding level of quality has a different impact depending on the recipient.

In my opinion, the approach presented by Eppler & Helfert is very interesting because of the framework they propose. It is a different angle, since they assume that costs related to data quality both in preventing and correcting are going to appear. The focus should be not in the absolute effort, but rather in its comparison -they focus on the real, overall costs, not in the local. I believe that this is a middle point between the theory and the practice that can be very useful to manage Master Data.

On the other hand, the point of view of Dietze & Fischer is very practical, as they use as classification the most common problems in real companies. I think this can be a good approach in order to prioritize the areas to focus on when starting a new project.

Regarding English classification proposal, I believe it helps to point out the fact that profitability losses are not only derived from regular costs, but also from opportunity costs. It is a very important issue that can be easily forgotten, since there are no visible effects.

To illustrate the problems and costs explained in the previous sections, it is interesting to expose some of the researches that reflect how big this problem can be.

Capgemini (2008), one of the world's foremost providers of consulting, discovered that a broken information culture (how they collect, use, manage and share data) is endemic in the UK and is believed to suppress performance by an average of 29 per cent. This equates to an annual £46 billion missed opportunity for private sector profits, and £21 billion in administrative costs across the public sector.

The situation is similar across the big pond, as a study published by Friedman & Smith (2011, 1), consultants from Gartner, who expose that poor data quality is a primary reason for 40% of all business initiatives failing to achieve their targeted benefits. Besides, Data Quality effects overall labour productivity by as much as a 20%.

Another survey conducted by The Data Warehousing Institute supports the previous findings. Eckerson (2002, 9) shows the severe losses that have taken place in U.S.A. during the last years. Some of the examples are:

- A telecommunications firm lost \$8 million a month because data entry errors incorrectly coded accounts, preventing bills from being sent out.
- An insurance company lost hundreds of thousands of dollars annually in mailing costs (postage, returns, collateral, and staff to process returns) due to duplicate customer records.
- An information services firm lost \$500,000 annually and alienated customers because it repeatedly recalled reports sent to subscribers due to inaccurate data.

- A global chemical company discovered it was losing millions of dollars in volume discounts in procuring supplies because it could not correctly identify and reconcile suppliers on a global basis.

As it can be seen, the implications of poor Quality Data can be devastating, provoking different problems in companies that always end up having their consequence in profitability.

2.3.4 Sources of poor data quality

After having analysed the impacts of poor data quality and concretely the costs derived from them, it seems logical to make one related question – How do those problems appear? It is necessary to study the roots of poor data quality in order to be able to reduce or eliminate the subsequent impact, and that is the purpose of this section.

It is interesting to begin by structuring the possible sources of a data quality problem. In this field, Rahm & Hai Do (n.d., 3) propose a blueprint that classifies the problems in two ways, through the source (single or multiple) and through the level (schema or instance).

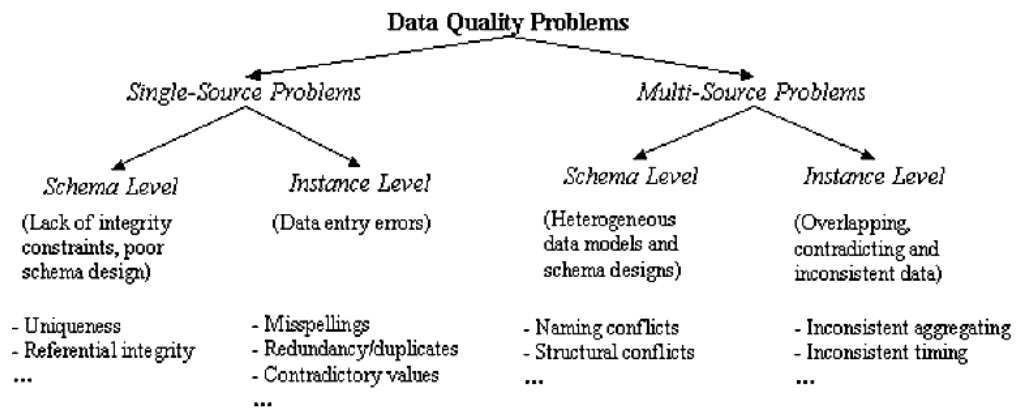


Figure 2.3.4. E. Rahm & H. Hai Do, Classification of DQ problems in data sources.

Single-Source problems occur only within one system.

- The Schema level defines the constraints that control the permissible data values in the different fields. Therefore, the related errors will emerge if the design of the data structure is poor.

Scope/Problem		Dirty Data	Reasons/Remarks
Attribute	Illegal values	bdate=30.13.70	values outside of domain range
Record	Violated attribute dependencies	age=22, bdate=12.02.70	age = current year - birth year should hold
Record type	Uniqueness violation	emp ₁ =(name="John Smith", SSN="123456"); emp ₂ =(name="Peter Miller", SSN="123456")	uniqueness for SSN (social security number) violated
Source	Referential integrity violation	emp=(name="John Smith", deptno=127)	referenced department (127) not defined

Table 2.3.4-1. E. Rahm & H. Hai Do, Examples for single-source schema problems.

- The Instance level applies to the problems that appear in the data itself, such as errors or inconsistencies that cannot be prevented through the schema.

Scope/Problem		Dirty Data	Reasons/Remarks
Attribute	Missing values	phone=9999-999999	unavailable values during data entry (dummy values or null)
	Misspellings	city="Liipzig"	usually typos, phonetic errors
	Cryptic values, Abbreviations	experience="B"; occupation="DB Prog."	
	Embedded values	name="J. Smith 12.02.70 New York"	multiple values entered in one attribute (e.g. in a free-form field)
	Misfielded values	city="Germany"	
Record	Violated attribute dependencies	city="Redmond", zip=77777	city and zip code should correspond
Record type	Word transpositions	name ₁ ="J. Smith", name ₂ ="Miller P."	usually in a free-form field
	Duplicated records	emp ₁ =(name="John Smith",...); emp ₂ =(name="J. Smith",...)	same employee represented twice due to some data entry errors
	Contradicting records	emp ₁ =(name="John Smith", bdate=12.02.70); emp ₂ =(name="John Smith", bdate=12.12.70)	the same real world entity is described by different values
Source	Wrong references	emp=(name="John Smith", deptno=17)	referenced department (17) is defined but wrong

Table 2.3.4-2. E. Rahm & H. Hai Do, Examples for single-source instance problems.

On the other hand, Multiple-Source problems are those that happen when there is more than one system. Rahm et al (n.d., 4) explain that the same errors that appear in single-source systems can appear in these cases as well but leading to more severe results, due to the fact that the integration between different sources is a complex process. In this cases, their data can overlap, contradict or be represented differently.

- The Schema level problems are derived from two main causes: naming and structuring. The naming errors can be either Homonyms (same name used for different objects) or Synonyms (different names used for the same object). The structure conflicts are those addressed to different representations of the same object in the sources in terms of constraints, tables, data types, etc.
- The Instance level errors are basically data conflicts. These conflicts can be due to several reasons, mainly different value representations and different value interpretations.

I believe that it is very important to be totally aware of this kind of problems, because in the business ERPs the data can be entered from diverse sources. I would like to remark that these sources can be either internal or external. On the one hand, different departments or modules enter data within the company and therefore it is very likely to have some inconsistencies or related issues. On the other hand, there is a part of the data that is shared with suppliers, clients or other stakeholders, and various problems can appear if they are not properly connected and integrated.

In order to understand better this matter, a multiple-source example is shown below. The tables represent two different sources that need to be integrated in the same system.

Regarding the schema level, there are name conflicts (synonyms Customer/Client, Cid/Cno, Sex/Gender) and structural conflicts (different representations for names and addresses).

On the other hand, at the instance level there are different gender representations ("0"/"1" vs. "F"/"M") and presumably a duplicate record (Kristen Smith). The latter observation also reveals that while Cid/Cno are both source-specific identifiers, their contents are not comparable between the sources; different numbers (11/493) may refer to the same person while different persons can have the same number (24).

Hence, solving these inconsistencies requires both schema integration and data cleaning.

Customer (source 1)

<i>CID</i>	<i>Name</i>	<i>Street</i>	<i>City</i>	<i>Sex</i>
11	Kristen Smith	2 Hurley Pl	South Fork, MN 48503	0
24	Christian Smith	Hurley St 2	S Fork MN	1

Client (source 2)

<i>Cno</i>	<i>LastName</i>	<i>FirstName</i>	<i>Gender</i>	<i>Address</i>	<i>Phone/Fax</i>
24	Smith	Christoph	M	23 Harley St, Chicago IL, 60633-2394	333-222-6542 / 333-222-6599
493	Smith	Kris L.	F	2 Hurley Place, South Fork MN, 48503-5998	444-555-6666

Table 2.3.4-3. E. Rahm & H. Hai Do, Examples for multiple-source problems.

A possible integration could be as follows:

Customers (integrated target with cleaned data)

<i>No</i>	<i>LName</i>	<i>FName</i>	<i>Gender</i>	<i>Street</i>	<i>City</i>	<i>State</i>	<i>ZIP</i>	<i>Phone</i>	<i>Fax</i>	<i>CID</i>	<i>Cno</i>
1	Smith	Kristen L.	F	2 Hurley Place	South Fork	MN	48503-5998	444-555-6666		11	493
2	Smith	Christian	M	2 Hurley Place	South Fork	MN	48503-5998			24	
3	Smith	Christoph	M	23 Harley Street	Chicago	IL	60633-2394	333-222-6542	333-222-6599		24

Table 2.3.4-4. E. Rahm & H. Hai Do, Multiple sources Integration.

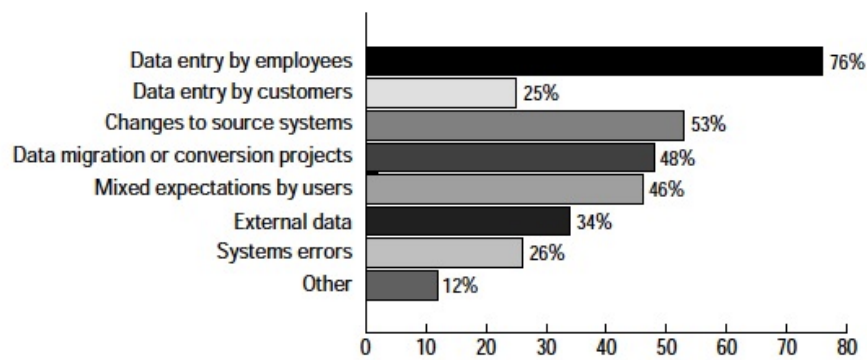
Where it can be seen that the duplicated record has been fixed, the address has been standardized, the gender representation has been unified, and both CID and Cno data has been included.

Note that, even if most of the problems have been solved, there is a lack of information in some of the fields (such as Phone or Fax) that could lead to other indirect problems.

So far I have explained the poor quality sources in a purely theoretical manner. However, another way to get to understand the possible causes is through an analysis of what actually happens nowadays in real companies.

Hence, I will refer to the already mentioned TDWI survey, where 286 respondents from different companies in U.S.A. exposed why they had problems with Data Quality. The results about sources of poor data quality are below:

Sources of Data Quality Problems



Graph 2.3.4. The Data Warehousing Institute, Sources of Data Quality Problems.

Eckerson (2002, 14) analyses these findings and explain the reasons why these problems occur:

The first and easiest-to-fix cause is simply a lack of validation routines. These routines check data as it is entered in the system, so it prevents the error from the very beginning.

However, there are sometimes when the field can be valid but incorrect, and in this case the way to prevent the derived problems is through Database Integrity rules and further complex methods.

Another option is that the field is both valid and correct but it appears more than once because it can have different syntaxes, formats or structures. In these cases the solution would be a data cleansing or further tools to standardize the data.

In terms of structure, if the system design is poor or not robust, it can drive to several data problems. This happens very often nowadays for the reason that companies rush to develop new systems. On the other hand, when changes in the structure of source systems take place (such as adding new fields, varying the code value, etc), it can lead to multiple errors downstream if front-line people are not aware about the modifications. On the other hand, the opposite error can happen when there are changes in source systems that do not modify the subsequent processes because they are not integrated.

To finish with, there are other actions that can provoke problems, such as data conversion or data migration from one system to another. The errors appear because, even if the data itself have been correctly converted, the business processes are not implemented properly in the new system.

In my opinion, both approaches are very interesting, especially if they are combined –using the theoretical frame of Rahm et al., and attacking the problems through the recommendations of Eckerson.

The classification proposed by Rahm et al. covers all the possible sources of errors from two points of view: internally (the level within the database) and externally (the number of sources that need to be merged). I personally like this manner because it is very structured and thereby per-

mits to frame the problems and apply the adequate tools to solve them. I would suggest using this categorization to separate the errors and to prioritize them.

On the other hand, the study of Eckerson is very useful since it focuses on the most common causes of poor data quality. I believe that it could be the base when implanting Data Maintenance procedures. Thus, the first approach should regard data entry by employees, the second the change to source systems, etc. Moreover, the explanations of the errors and the possible practices to avoid them are useful, and therefore should be also considered.

Nevertheless, I would like to point out that, even if the solutions proposed by Eckerson are helpful, maybe they do not focus enough in the main source of problems - data entries. He explains that there are different methods to prevent these errors, such as validation routines, integrity rules and other technical methods. But he does not concentrate on why the employees commit so many mistakes.

Is the process too complex? Could it be simplified? Are the employees able to understand it? Do they need more education?

I would suggest further research on the social aspect of this issue, since it is a big source of errors that quickly spread within the system. Furthermore, it may happen that these “external” or rather “non-technical” solutions are able to provide an easier or cheaper way to improve the quality of the Master Data. Thus, even if they do not solve entirely the problem, in terms of cost-efficiency or profitability, they can be a good approach.

2.4 Master Data Management

In the section Implications and consequences (2.3), it is been proved through different ways how defective data can affect a company provoking several damages. Thus, in this part I will explore the options to solve that problem, or rather to keep it under control.

First, in section MDM Solution (2.4.1) I will expose diverse approaches about the management itself – how Master Data needs to be handled, which solutions exist, how a plan should be implanted, etc.

Secondly, in section Master Data Creation (2.4.2) I will explain how the Master Data should be created in order to avoid future errors that could spread downstream the system.

To finish with, in section Master Data Maintenance (2.4.3) I will analyse different methods to maintain the Master Data in a proper way.

2.4.1 MDM Solution

What is Master Data Management? What does it imply? How should it be implemented? These questions will be answered throughout this part.

Hence, the first step is to define Master Data Management. Wefer (2008, 41), a manager of product marketing, states “think of it as a highly accurate tool for making decisions and gaining insight about customers, products, employees, vendors, and other aspects of a business with laser-like precision”.

A more technical definition could be “a combination of applications and technologies that consolidates, cleans, and augments this corporate master data, and synchronizes it with all applications, business processes, and analytical tools. This results in significant improvements in operational efficiency, reporting, and fact based decision-making.” (Oracle 2011, 5).

The general idea should be that MDM is a tool that links IT and business in order to generate and maintain accurate, reliable and consistent information across the whole company. This is very important because it is necessary to have proper Master Data in order to carry out correctly the diverse processes, optimize the operations and take well-based decisions.

An overall structure of the MDM process is shown in the following schema. It shows the initial requirements or procedures that need to be done before starting the actual management and the outputs in terms of business advantages.

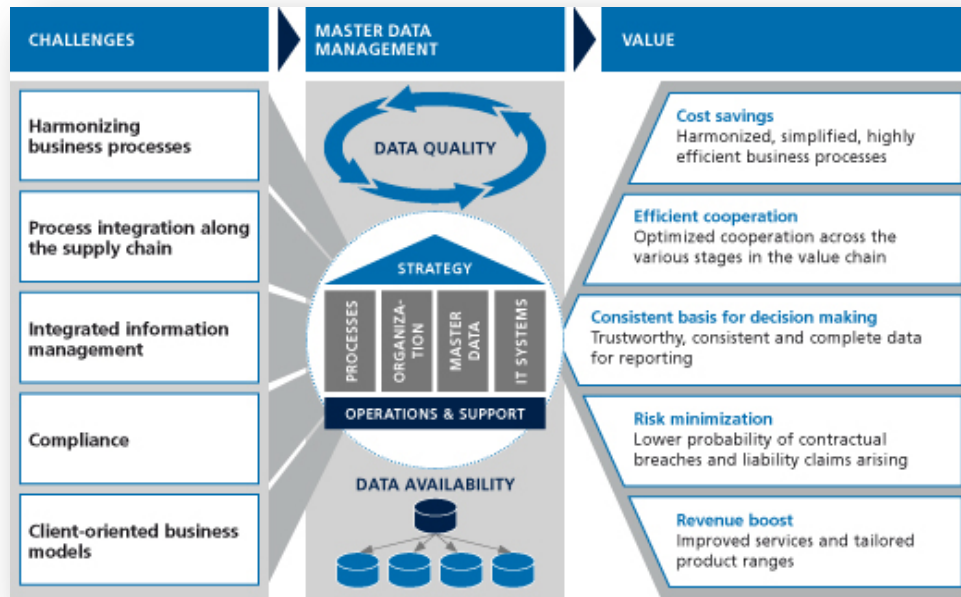


Figure 2.4.1. Camelot Management Consultants, Study on strategic management of master data.

Now, I would like to expose more concretely what this tool implies – what is the real meaning of implementing a MDM solution in a company. Halpern (2007, 2), the senior principal of value engineering at SAP, explains that to achieve effective master data management and improve operating performance, the solution has to fulfil the following requirements:

Master Data consolidation means that the Master Data needs to be identified along the whole company system, then it has to be consolidated, and finally located into a “map” known for all the company in order to unify analytics and reporting. Note that the term “consolidate” implies several activities such as matching, normalizing, cleansing, and storing master data imported from client systems.

Master Data harmonization ensures that master data is synchronized across the different areas or departments of the company. This step basically extends the consolidated data, distributing the relevant parts along the system. Therefore, it enriches the client application systems with locally relevant information.

Central Master Data Management refers to the management itself – it means that the maintenance, storing and distribution mechanisms of Master Data should be handle in a central manner. Hence, this activity differs from Master Data Harmonization in that Master Data is created and delivered centrally using a rich client.

As it can be seen, this approach is mainly a theoretical idea, since it does not provide the concrete steps to fulfil the exposed requirements.

This leads to the last question that I exposed at the beginning of this section, “How should it be implemented?”.

Wolter & Haselden (2006), from Microsoft Developer Network, explain that every MDM project is customized to each company according to the concrete requirements and resources, but the common steps are:

- **Identify sources of Master Data:** It includes analysing the data system and determining which parts of it match the characteristics of Master Data.
- **Identify producers and consumers of Master Data:** Analyse further the Master Data and get to know where it is produced (or rather which applications produce it) and where it is used (or rather which applications need it).
- **Collect and analyse Metadata about Master Data:** Analyse further the sources of step one in order to determine what Master Data really means. Different aspects should be defined, such as attribute name, data type, allowed values, constraints, default values, dependencies and owner of the definition and maintenance.
- **Design data responsible:** The owner of each data source designs a person to be responsible of the data. Hence, the steward should have the knowledge of the current source data and the ability to determine how to transform the source into the master-data format.
- **Establish a data-governance program and council:** A group of people with the knowledge and authority to make decisions about Master Data should be appointed. Their tasks include to determine how the master data is maintained, what it contains, how long it is kept and how changes should be made.
- **Develop the Master Data model:** This is a critical step since it establishes the characteristics of Master Data -what attributes are included, which data type they are, what values are allowed and so forth. This is normally both the most important and most difficult step in the process, since every party wants its attributes to be included. However, this could lead to have Master Data that is too complex to be useful. Therefore, it is important to select from the beginning the final decision maker and set the priorities.
- **Select a toolset:** In order to have proper master records, it is necessary to create them adequately. Therefore, it is important to select a tool that is able to create the master lists by cleaning, transforming, and merging the source data.
- **Design the infrastructure:** Through the infrastructure, Master Data will be exposed to the different areas and applications of the company. Depending on how the company requires the data to be managed and maintained, different types of infrastructures can be selected.
- **Generate and test the Master Data:** In this step the creation of Master Data takes place through the tool designed before. This is often an iterative process that requires a lot of manual inspection to ensure that the results are correct and meet the requirements established for the project. A more detailed description about Master Data creation is available in section 2.4.2.

- **Modify the producing and consuming systems:** Depending on the infrastructure selected, different changes in the systems that produce, maintain, or consume Master Data may be required. The reason is that it is necessary to ensure that they work with the new source of Master Data.
- **Implement the maintenance processes:** A maintenance plan should be established, and it must incorporate tools, processes, and people to maintain the quality of the data. This maintenance plan depends on the infrastructure designed before and it includes different processes that will be exposed more in detail in section 2.4.3.

Thus, theory affirms that MDM solution should be applied to the whole system in order to have a unique and integrated source of Master Data, so many companies assume that they need to merge all data before taking advantage of it. The problem of this is that, as it can be seen, the implementation of a MDM is a complex process -it is a long and expensive procedure that impedes to have an attractive ROI (Return On Investment). Therefore, sometimes a better approach is to implement the MDM project in a phased way.

In that line of MDM cost-efficiency, Wefer (2008, 41) exposes that an incremental implementation enables enterprises to develop an effective working model for future development, while allowing a fast ROI on the first phase of the project. She suggests that a possible start could be a hybrid implementation where the data stays where it currently resides rather than consolidating it into one master source.

Another option could be to begin implementing a MDM solution in one area of the company, typically the customer section. The reason is that the customer Master Data is normally the most used one. Afterwards, the company can expand the MDM project to other sectors, taking advantage of the previous implantation and therefore giving them the required intelligence to improve business processes and increase competitiveness.

She ends proposing some procedures in order to select where to start:

- Sketch the big picture, but concentrate on short-term gain.
- Understand the effects on major business processes.
- Evaluate projects from both an IT and business case point of view.
- Demonstrate short-term gains to sell long-term opportunities.

To finish with, I would like to remark that these projects imply a lot of effort not only of the team that carries out the implantation, but also of all employees in the company that have contact with the system:

"A Master Data Management program is not a project but a commitment by the business to leverage information for reuse in order to improve business process outcomes" (Gartner, n.d.).

2.4.2 Master Data Creation

The creation of Master Data is one of the most important steps when implementing a Master Data Management solution. The reason is that the whole process will be non-sense if the data that is going to be shared along the different areas is not the correct one.

Before explaining further how it should be done, I would like to point out that normally, the creation of Master Data does not imply to enter the data. In reality, companies already have the core data but this data is not well managed –it is neither shared between the systems nor well maintained. Therefore, the creation of Master Data is not a creation of data (since it is actually already in the system) but rather a conversion from data to Master Data.

According to Wolter et al. (2006), this process is divided in two parts: normalizing and matching.

Normalizing basically refers to the actions of cleaning and standardizing, which means transforming it into the arranged model of Master Data. Consequently, before starting this step, it is crucial to understand the model in terms of attributes, contents, and mapping of each source to the final Master Data system.

They explain that this step is very similar to the Extract, Transform, and Load (ETL) process that is used in data warehousing. It includes:

- Normalize data formats to ensure that all the data has the same structure.
- Replace missing values.
- Standardize values using common or standard measurements.
- Map attributes.

On the other hand, the matching part aims to eliminate duplicates that appear when two or more sources are merged. They state that it is very hard, since a balance between “false matches” and “missed matches” has to be reached -false matches lead to lose data and missed matches reduce the value of maintaining a common list.

The way this process works is through a tool that calculates a confidence factor for each match, taking into account the number of attributes matching and the proximity of the match. This degree of confidence is compared to a given limit that is adjusted according to the importance or consequences of a false match. The higher degree of confidence, the higher possibility to surpass the limit and create a new match.

2.4.3 Master Data Maintenance

Implanting a Master Data Management system in a company is not a punctual action, but rather a continuous process that needs to be well maintained over the time in order to ensure its reliability and quality.

There are many different ways to maintain the Master Data, some of them referring to the final-user actions and other to the technical scope of the ERP system. In this section I will expose through a general approach what Master Data maintenance implies.

Note that this part is totally linked to the MDM Solution section (2.4.2), since some of the steps that were detailed there have much importance in this procedure.

To begin with, the data responsible or steward should be the main person involved in maintaining the Master Data quality. Regarding this person, Wolter et al. (2006) expose that he/she “is normally a business person who has knowledge of the data, can recognize incorrect data, and has the knowledge and authority to correct the issues”.

Besides, they suggest that the MDM infrastructure should include tools that help the data steward to recognize issues and simplify corrections. For instance, a program whose objective is to point out questionable matches.

Other tasks that the data responsible may have are:

- Reviewing the second part of the Master Data creation (section 2.4.2), the matching between different Master Data records. It is important to check the cases where the data match criteria were close but below the threshold.
- Pulling changes and additions into the MDM system, and to distribute the cleansed data to the required places.

In my opinion, this approach suggests that the maintenance should be done backwards in the sense that the basis is a feedback and control carried out by the Data Responsible.

Another approach is exposed by Hillard (2010, 151-153), who establishes a forward method to adequately implant the Master Data in the system. He explains that Master Data usually uses a central database that can be accessed and updated by all the required parties. Hence, the key is to manage it through common and standardized steps to ensure its accuracy.

Thus, the first step is to agree the characteristics of the data centrally, having into account all the areas of the company and their related requirements. After this, the particular departments fill in their parts of the data in a concrete and predefined order.

Robert Hillard further provides some technical norms that should be always followed in order to ensure that the Master Data is sustainable and robust over the time. These norms should be applied continuously as a cycle, as it can be seen below:

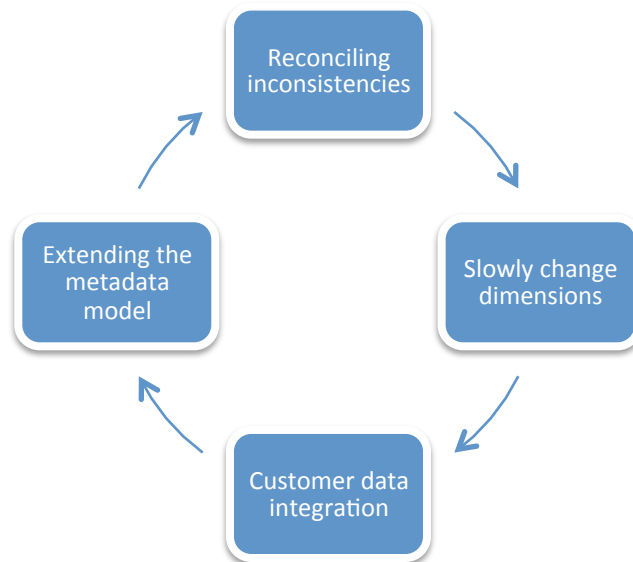


Figure 2.4.3. Robert Hillard, Master Data Management.

Reconciling inconsistencies: When the amount of data is quite wide and spreads across diverse departments and systems, it is sure that some inconsistencies will appear. The term “inconsistencies” refers to the fact when an errant process causes two different updates to be applied to the same record, which cannot be correct.

In those cases, it is very important to have a robust system that never overwrites the data, but rather tracks the changes. In addition, it must have a mechanism to notify all the parts involved or stakeholders when data is updated.

Slowly changing dimensions: It is completely normal that in a database the relations, names, hierarchies and other details change periodically. In these cases, it can be really difficult to decide whether to apply the change only to the future records or to the past ones as well.

The problem of applying only to future updates is that, when analysing past data in order to get information, several problems can occur. There are three ways to deal with these changes:

- Show data as it is now: Cast all data (in terms of Master Data and hierarchies) as it is now.

This is the most used method as far it ensures that there are not problems when analysing the data and its related reports.

- Show data according to its date: Describe any time series information (such as this year versus last year) and therefore show the data in the manner it was described at the time.

This is used when the data is very associated with its concrete management (in terms of relations or hierarchies).

- Offer both options: Depict all results in terms of how they could have looked at a particular, specified, point in time. So it could be possible to show future data with a past used method.

This is used especially in issues related to audits, analysis and comparison of the data.

Customer data integration: It integrates the master data that refers to the relation between the organization and its clients. It is a very relevant part of the ERP due to the fact that clients' management is a crucial issue in any business –they are the source of profit.

However, it has always been an application not very integrated with the rest of the system. This fact does not permit to optimize the rest of the business processes according to the customer needs and requirements. That is the reason why it is one of the main steps in Data Maintenance.

Extending the metadata model: Master Data has become a very used way to manage organizations, since it is a good manner of navigating through the different business' dimensions.

Therefore, it is crucial to have a tool that creates a model where the Master Data is implanted. The best way it is the "Metadata model". The definition of this term is exposed in section Related Concepts (1.6).

The model consists of a concept diagram that shows the relation between Master Data and other parts of the system. It is very important to update this diagram as the company develops its structure and hierarchy, and to accurately describe the new concepts (their dimensions, their use and so on). In this manner, it will be a frame to look at when managing Master Data through the different dimensions of the company.

3 MASTER DATA MANAGEMENT - PRACTICE

In order to understand and prove the information presented in the first part of this Thesis, a Business Case will be put into practice. In this way it will be easier to understand all the concepts that have appeared throughout this paper, and check what they imply –how they affect business profitability.

Hence, in this section I will firstly introduce how the implementation “business + SAP ERP” works, then I will briefly explain the areas that will be used in the Case, and finally I will put into practice all the knowledge acquired throughout this Thesis.

The tool that will be used to carry on this practical example is SAP ERP. SAP it is a pioneer in business related software, having over 42 years experience in this field. Its headquarter is established in Walldorf, Germany, but due to its huge expansion, it counts with locations in more than 130 countries (SAP n.d.). Nowadays, many of the biggest companies use this tool to manage their processes, being the world leader in enterprise software and software-related services.

Note that this second section attempts to be completely related to the theoretical part. However, it is important to remark that the processes explained previously were associated with a manufacturing Company, whereas from now I will be referring to a wholesale company. Many of the systems are similar, but there are some parts linked to the production itself that are not going to be mentioned.

3.1 Implementation of an ERP

Implementing SAP or any other ERP in a business is not a simple action. It is not just installing the program and starting to use it. Rather, it implies a complex project that can lead a big company to break.

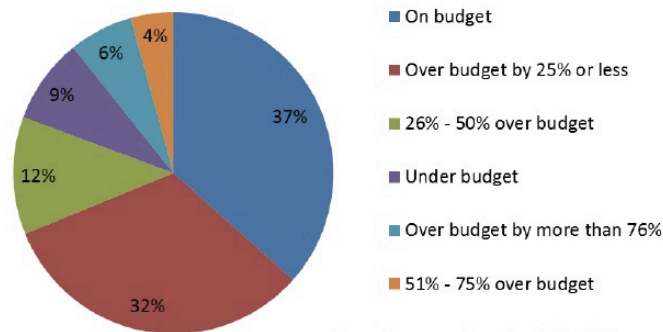
Panorama Consulting Solutions (2014, 2) carried out a research to analyse deeper these implementations. According to it, it is possible to see that adjustments in a company in order to align the whole system with the structure provided by an ERP are costly and require a long duration. The average results of the last years are:

YEAR	COST	% OF COST OVERRUNS	DURATION	% OF DURATION OVERRUNS	% RECEIVING 50% OR LESS BENEFITS
2013	\$2.8MM	54%	16.3 months	72%	66%
2012	\$7.1MM	53%	17.8 months	61%	60%
2011	\$10.5MM	56%	16 months	54%	48%
2010	\$5.5MM	74%	14.3 months	61%	48%

Table 3.1. Panorama Consulting Solutions, Data Summary by year.

Regarding the costs, they are normally out of the budget. This happens because it is very difficult to estimate the unexpected events that may appear during the 16 months of implementation. Panorama Consulting Solutions also provide details about what companies finally spend in an ERP implementation compared to the previous expectations.

Implementation Costs

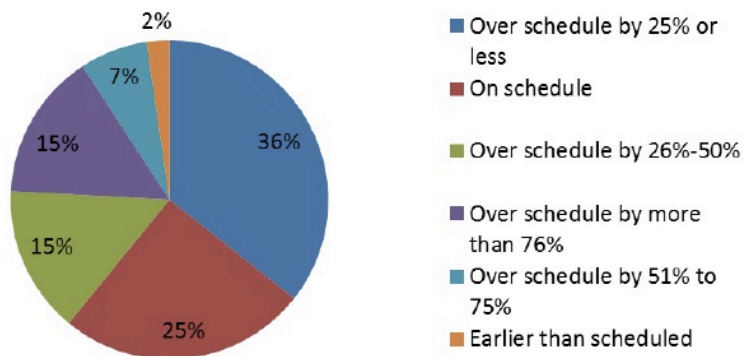


Graph 3.1-1. Panorama Consulting Solutions, Data Summary by year.

As it is shown, the majority of the implementations are on budget but almost the same proportion is over it (even if it is by less than 25%). That is a big amount of money taking into account that normally the budgets are over \$3 million.

In terms of duration, they also provide the comparison between the expectation and the actual one. In this case it is obvious how tough is to forecast the duration of each step in the implementation process.

ERP Project Duration



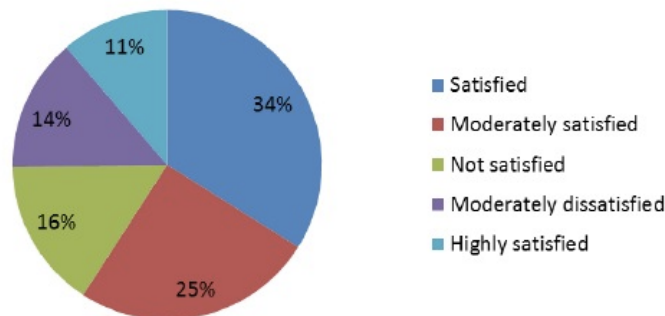
Graph 3.1-2. Panorama Consulting Solutions, Data Summary by year.

So, in general it could be said that the implementation usually costs and lasts more than what it was expected. That is one of the reasons why, even when they have an adequate budget and the required time, many enterprises think twice before beginning an ERP project.

To finish with, I would like to remark that 60% of the companies are satisfied with the overall result of the ERP. This percentage represents more than the half of the cases, but in my opinion it is small taking into account the tremendous efforts that they have to do.

I think it would be very interesting to know which proportion of the dissatisfied customers are not convinced of the results because of the ERP itself or rather because of the implementation project (that is, because of the incremented costs and duration of it).

Overall Experience with ERP Vendors



Graph 3.1-3. Panorama Consulting Solutions, Data Summary by year.

So far I have exposed the implications of the implementation regarding the resources that a company has to spend. Now I will analyse how the implementation is done inside the ERP, concretely the SAP one. Therefore, I will focus on how the business processes run within the system and how the general structure of the companies is adapted.

Regarding the alignment of the enterprise, Schulz (2012, 55) explains that every company has its concrete business processes that are represented in the structure of the organization. A good way to show it is through an organization chart:

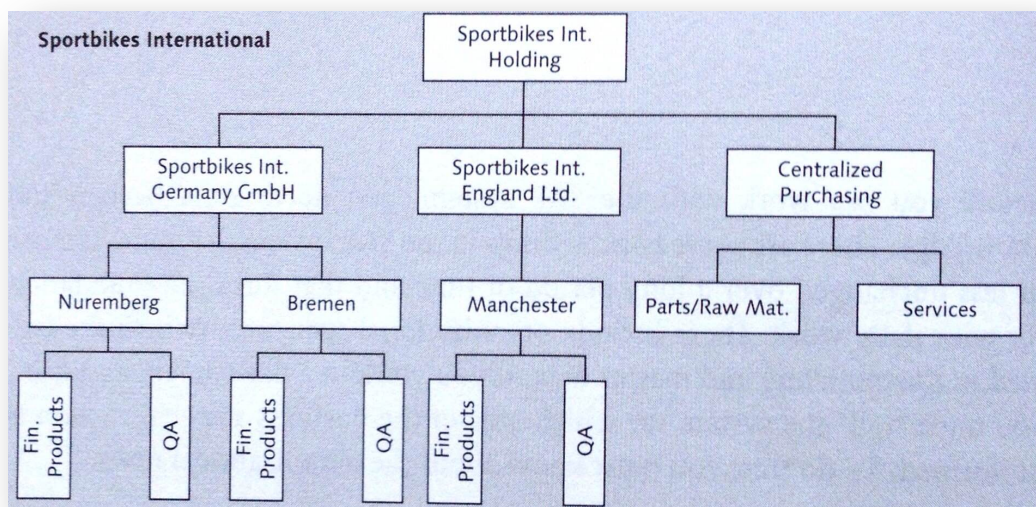


Figure 3.1-1. O. Schulz, Example of an Enterprise Structure from the Logistics Perspective.

This example shows a fictional bicycle manufacturer, Sportbikes International. The three different international companies are pooled in a holding and the production happens at two plants in Germany and one plant in England. Besides, the subsidiaries' warehouses are subdivided into two areas: Finished Products and Quality Assurance. Finally, there is a Purchasing department that controls all the purchase within the whole holding: a group of employees is responsible for procuring raw materials and other subassemblies, and another one for purchasing services.

This company structure would be mapped in SAP through different organization units. Translating the previous structure to the corresponding in SAP, the chart would be as follows:

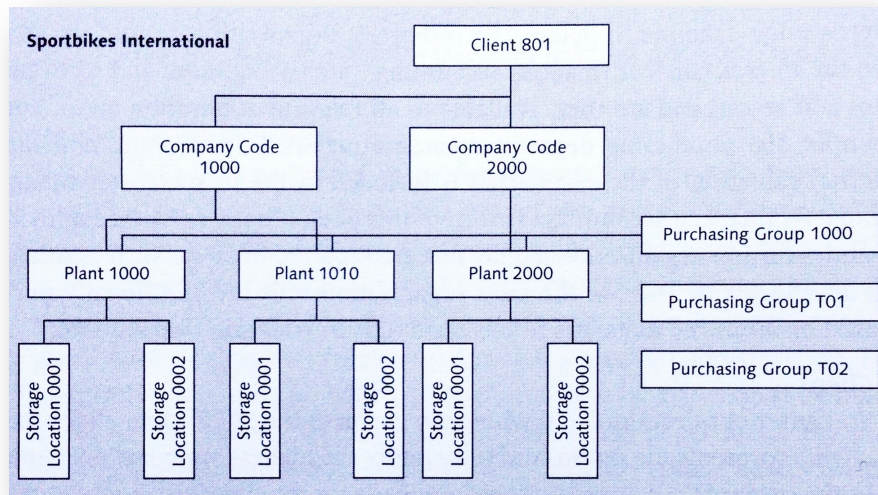


Figure 3.1-1. O. Schulz, Organizational Units of Sportbikes International in the SAP System.

Thus, there are different kinds of organizational units:

- The client (801) represents the highest level in the enterprise. It is the whole companies' holding, such as the group or the corporate headquarters.
- Company codes (1000, 2000) represent the companies or subsidiaries in Germany and England.
- Plants (1000, 1010, 2000) represent the production plants of the companies.
- Storage Locations (0001, 0002) are a further division of the plants in different warehouses.
- The Purchasing Group (1000) is the organization that manages and controls all the purchasing activities of the Client, and that is the reason why it is not linked to any Company Code.
- Purchasing Groups (T01, T02) represent different purchasing teams that have different objectives, purposes or tasks. Purchasing Groups are not assigned directly to other organization unit.

Schulz (2012, 31) also exposes that the way these different organizational units are involved in a business activity is through processes. In SAP terminology, a business process is a defined sequence of steps in the enterprise. It is structured this way because every activity needs to be mapped within the system.

Hence, it is very important to have in mind the whole process when carrying out a given transaction. If not followed, different problems and misunderstandings can appear easily as, in general, SAP processes are quite complex. In fact, it is very common that multiple departments participate in a single business process.

Inside the SAP ERP system, each step triggers another one that triggers another one, etc. Furthermore, it is important to point out that every step, process or activity is always recorded in the system. This is important because it makes the process transparent, continuous and reproducible at any time. That is the reason why the basic principle inside SAP is “don’t post without documentation”.

A simple example of a process is shown below. In this case is the Sales Order process, and it can be seen how it is divided in four different steps. Each step triggers the next one, and in every step a document or other relevant information is recorded within the system, so it is easy to trace all processing steps and documents.

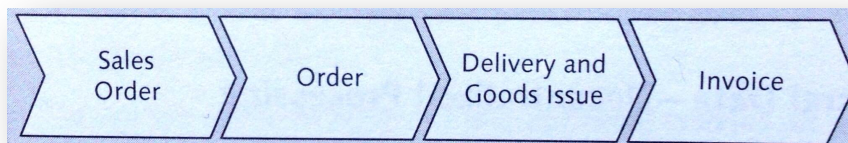


Figure 3.1-3. O. Schulz, Sales Process as Mapped in the SAP System.

3.2 SAP functions

In this section I will briefly introduce the areas in SAP ERP that will be managed in the Business Case. The objective is to understand the flow that each of these processes has in terms of Master Data, Documents, Inputs and Outputs.

In each of the areas' explanations one or more diagrams will be exposed. When handling the program (or going through the explanations), it is important to continuously have them in mind, as far the process' structure can be sometimes complicated and not obvious.

It is also relevant not to forget to think about them in a real manner - they all apply to existent processes that happen everyday in a general business. Therefore, there is much information required in the system whose source it is external -a real document, status, order, delivery, etc.

3.2.1 Purchasing

Purchasing is basically the company procurement process. As Magal & Word (2012, 7) explain, “it includes all of the tasks involved in acquiring needed materials externally from a vendor”. This process provides to the company the required materials from the corresponding suppliers.

It is deeply interrelated with the Materials Requirement Planning that was previously exposed (section 2.2.2.5. Materials Requirement Planning – MRP), since it sets the schedule for ordering or “purchase requisitions”. Nevertheless, there are other sources apart from MRP that can trigger the procurement process, such as Production Orders, Maintenance Orders or Sales Orders (Wagner, Weidner & Tracy 2009, 116).

Simha et al. provide a simple scheme where it can be seen how the requisition evolves along the different departments of the entity.



Figure 3.2.1-1. S. R. Magal & J. Word. Procurement process.

In the basic diagram the second step is already the purchase order, but normally there are previous activities, especially when there are more than one possible supplier. The following diagram shows that process more in detail:

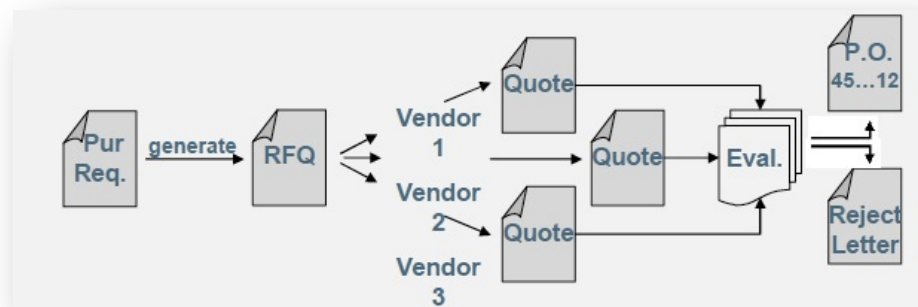


Figure 3.2.1-2. Wagner, S. Weidner & S. Tracy. Purchase Order.

An RFQ is an invitation to a vendor by a Purchasing Organization to submit a bid for the supply of materials or services. They are sent to different vendors that will send back their concrete quotations. After having evaluated them, the company selects the final supplier and replies to the rest of them in order to inform that their quotations were rejected.

The complete flow in SAP ERP:

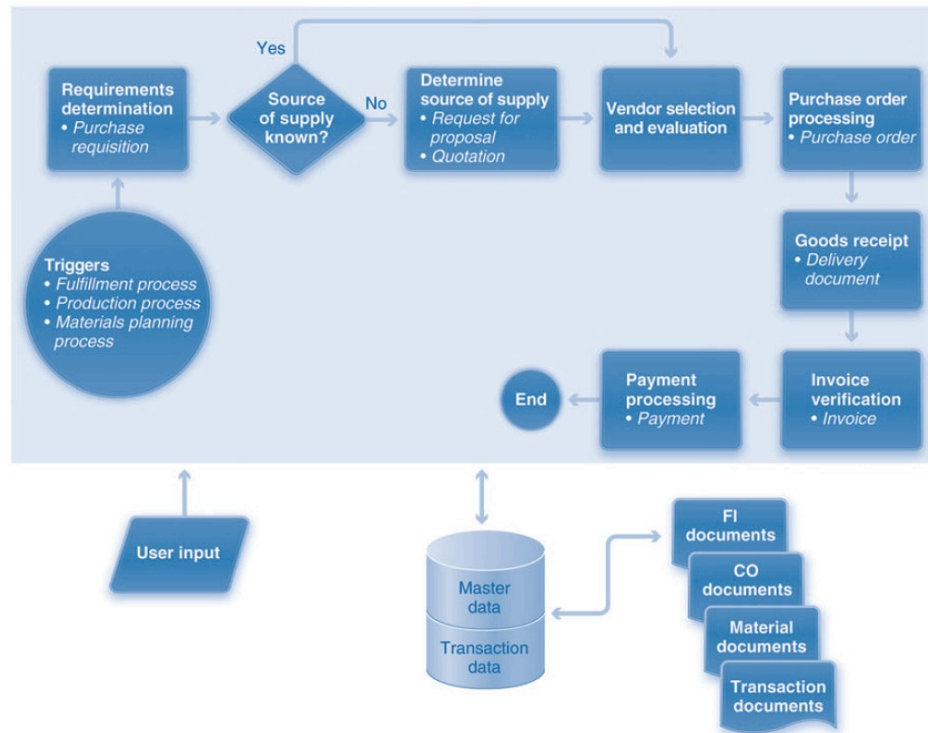


Figure 3.2.1-3. S. R. Magal & J. Word A detailed procurement process.

In the diagram is shown the complete process that a typical purchase follows. Nevertheless, it is important to remark that the process can begin with other transactions apart from a Purchase Requisition. For instance, it can start directly with a Purchase Order.

3.2.2 Sales & Distribution

Sales and distribution is the area whose objective is fulfilling customer orders. According to Magal et al. (2012, 9), it is “efficiently processing customer orders”. This process is the way a company has to get revenues and therefore profit -it is the basis of the business.

It is deeply interrelated with the Aggregate Planning that was previously exposed (section 2.2.2.1. Aggregate Planning - AG), since the Sales department is the one that -based on previous orders, current contracts and other methods- provide the Product Demand Forecast that is the main input in the AG. Consequently, this department also affects to the Master Production Schedule (section 2.2.2.2.) and the Materials Requirement Planning (section 2.2.2.5.). The current and future sales are one of the most important pillars in the global management of a company.

Magal et al. provide a simple scheme where it can be seen how the customer purchase order evolves along the different departments of the entity.



Figure 3.2.2-1. Simha R. Magal & Jeffrey Word, A basic fulfillment process.

In the basic diagram the first step is already the sales order, but normally there are previous activities such as Inquiries or Quotations.

The complete flow in SAP ERP:

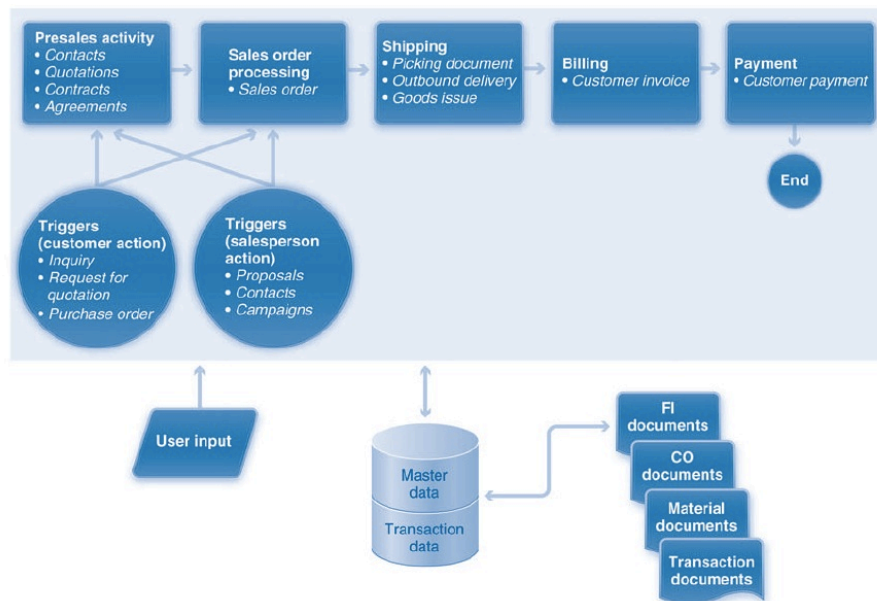


Figure 3.2.2-2. Simha R. Magal & Jeffrey Word, Fulfillment process steps.

At this point it is important to make sure that the process flow is clear. The main idea is to understand that transactions are related to other transactions. In the case of Purchasing and Sales and Distribution, the integration is as it is shown in the following diagram:

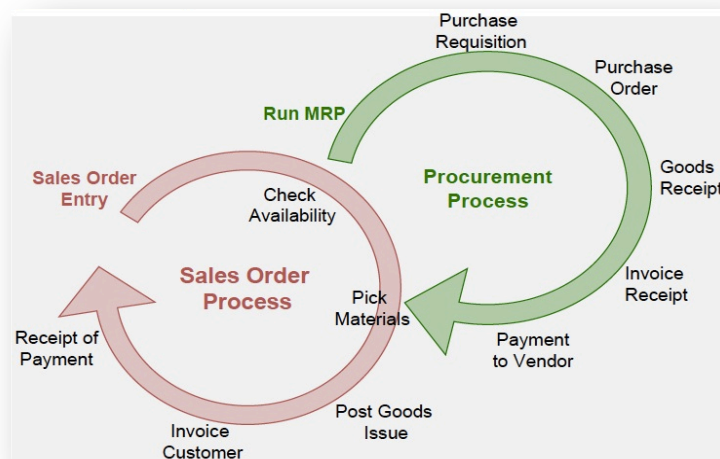


Figure 3.2.2-3. SAP University Alliances, Process Integration.

3.3 Practical Case

In this section a business case will be exposed in order to show how much master data can affect a company in terms of profitability. The way this is going to be proved is through data maintenance –handling different processes that are normal in a business. Concretely, the organization will be a Wholesale company thereby the data maintenance will be mainly focusing on the processes “Purchasing” and “Sales and Distribution”.

In the first part, Adequate Management (3.3.1), I will do a correct maintenance. In the text, there will be arrows to remark good practices in order to improve the data maintenance process.

In the second part, Inadequate Management (3.3.2), I will do a poor maintenance, having the common errors that happen in a company. In the text, there will be arrows to remark the errors and the profitability losses due to poor Master Data Management.

Hence, I will be able to compare in a practical way how these problems appear and how they affect to the overall process.

Note: The case scenario will appear in italics and the processes that are made by the company in normal typography.

The frame of the case begins as follows:

Our company is a Wholesale business that was established in Germany in 2001. We focus on food that typically is sold in supermarkets or groceries stores. Therefore, our activity is basically buying diverse kinds of food from different suppliers and selling them to Retail companies, normally supermarket chains. So far, we just perform our functions in Germany, although we are thinking about expanding abroad.

Four months ago the Sales Manager stated that the demand for chocolate bars was growing and the potential profits from buying and selling these products were very promising. Therefore, the Purchasing Manager has been inquiring from different possible suppliers and after some negotiations, he has exposed that the best option (according to the different vendors, plant distances and prices) is to buy Kik-Kat chocolate bars from the German office of Nestlé.

3.3.1 Adequate Management

In this first section I will show how to correctly perform the data maintenance of the most basic processes in a Wholesale Company – Materials Management, Purchasing and Sales and Distribution.

Throughout the case, I will therefore carry out these business transactions in a proper way, taking into account the management tools and methods that have been shown in the section Master Data Management (2.4).

As it has already been exposed, there will be arrows to remark good practices in order to improve the data maintenance process.

3.3.1.1. Materials Management

The first necessary step is to create the new material in our SAP ERP system. Hence, our employee creates a new Material Master record that is going to be a trading good since we do not produce it nor we sell it to the final customer –we just trade with it.

The employee uses the transaction MMH1 and begins to fill the required fields.

Basic data:

➔ Suggestion: In the case of Trading Goods, the basic data is normally provided by the supplier due to the fact that the items are not produced in the company. As exposed in section Sources of Poor Data Quality (2.3.4), the most reliable data in a company system is normally the provided by the customer. Hence, any company should always try to get data directly from the client through a questionnaire or a shared database.

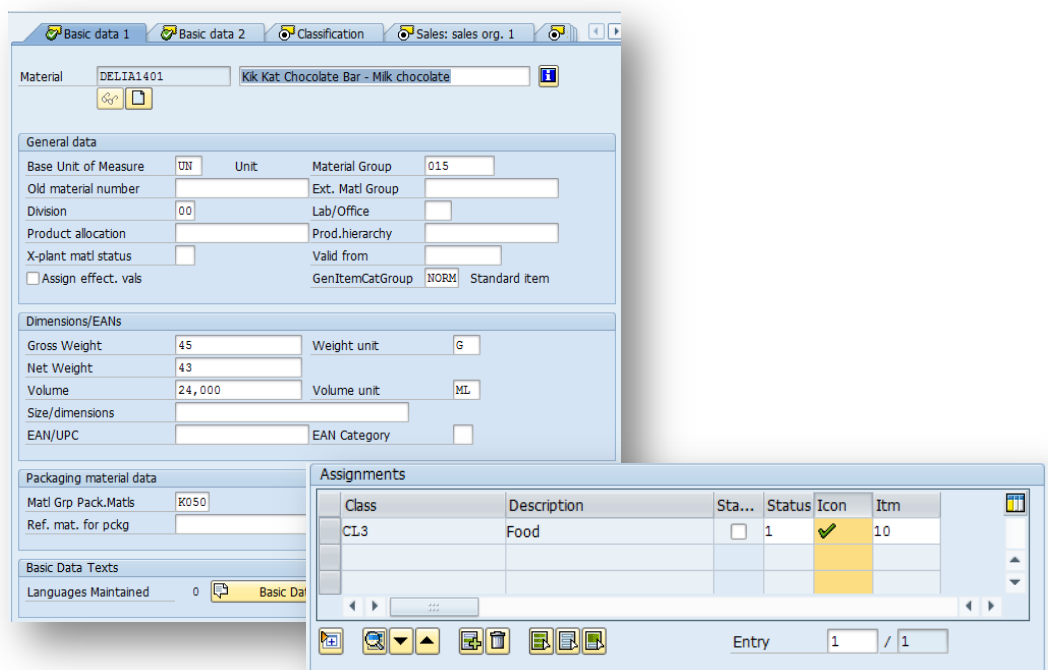


Figure 3.3.1.1-1. SAP, Basic Data.

➔ Suggestion: The Master Data structure should be agreed and unified for all the products in the company. Thus, there should be a concrete set of fields that should be always filled to ensure a correct future management. The employee who enters the data should be concerned about this issue.

In this case, the agreed fields to fill are: Unit of Measure, Material Group (food), gross weight, net weight and volume.

He also adds the Material Group Pack because he knows that the supplier provides the chocolates bar in palettes.

Sales:

This section is very important because it sets basic Master Data information related to the sales orders and their consequent deliveries.

➔ Suggestion: Accuracy is one of the first characteristics that good Master Data should have. In this section, this is particularly relevant because it sets the basic characteristics of the sales orders, such as the order quantities. The employee who enters the data should never (except products with no prior fix costs) leave the fields blank, he/she should rather ask the Sales and Distribution department.

Sales and Distribution department has analysed the costs to pick and load the Kit-Kat chocolate bars, and they have found that the breaking point is 100 units. That is why this is the minimum delivery quantity. Furthermore, they have estimated the sales orders processing costs and determined that the minimum order quantity is 500 units.

The screenshot displays the SAP Sales organization configuration for material DELIA1401 (Kik Kat Chocolate Bar - Milk chocolate). The interface is divided into several sections:

- General data:** Base Unit of Measure (UN), Unit (00), Division (00), Cross-division, Sales unit, Unit of Measure Grp, X-distr.chain status, Dchain-spec. status, Delivering Plant, Material Group (015 Food), and Cash discount (checked).
- Tax data:** A table showing tax information for Germany (DE) with tax category 'MWS Output Tax' and tax classification '1 Full Tax'.
- Quantity stipulations:** Min. order qty (500 UN), Min. dely qty (100 UN), and Delivery unit (50).

Figure 3.3.1.1-2. SAP, Sales organization.

➔ Suggestion: The same applies for the shipping data –it is Master Data that will be used in the Delivery process so it is very important that is to-tally accurate

The employee should know perfectly the differences between:

- Shipping setup time: time to set up the work centres where the material is processed. It is independent of quantity.
- Shipping processing time: time to process a specific quantity of a material.

The employee asks the Delivery section how long would it take to prepare the work centre to begin the shipping and he is told that, due to the complexity of the plant, they need one day to set the transportation, machines and workers that will carry out the shipping. The Delivery department further explains that, with Kit-Kats, they will need half a day for loading 100 units.

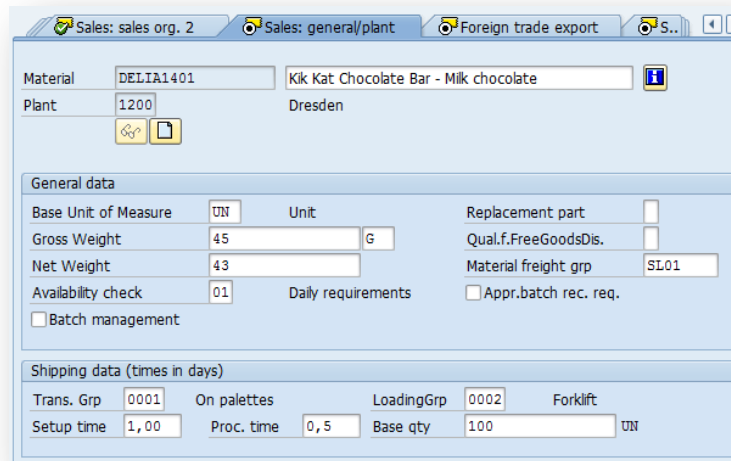


Figure 3.3.1.1-3. SAP, Sales plant.

To finish this section, the employee adds text about the product and about the brand that will be used by the Sales department.

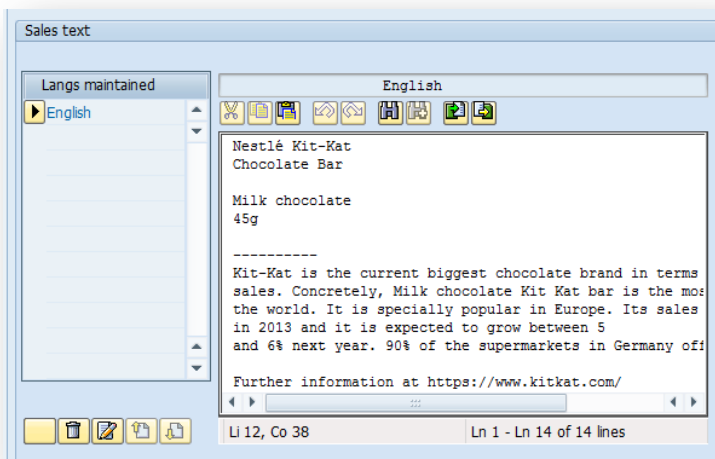


Figure 3.3.1.1-4. SAP, Sales text.

Purchasing:

This section is very important as well, since it sets basic Master Data information related to the purchase orders and their related goods receipts. There are basically two fields that should be filled:

- Purchasing Value key: Determines the number of days representing the time interval at which reminders or urging letters/messages are to be issued to the vendor. It also stipulates the minimum proportion of the whole quantity that could be delivered.
- GR Processing time: Number of workdays required after receiving the material for inspection and placement into storage.

The employee asks the Purchasing department about the frequency of reminders, which is 10 days. Besides, he is told that, about the goods receipts, our company requires 1 day in order to ensure the quality and place the items into their right place.

➔ Suggestion: The data required in “Other data/manufacturer data” that estimates the time required to process a Good Receipt should be actively maintained. That is, that it should be updated with a relatively high frequency. The reason is that this time value is used to determine the availability of the materials and consequently, an error in this number could drive to inconsistencies in the schedule. Besides, the processing time can increase/decrease over the time according to the experience or the workers, the number of products in the warehouses, the batch sizes, etc.

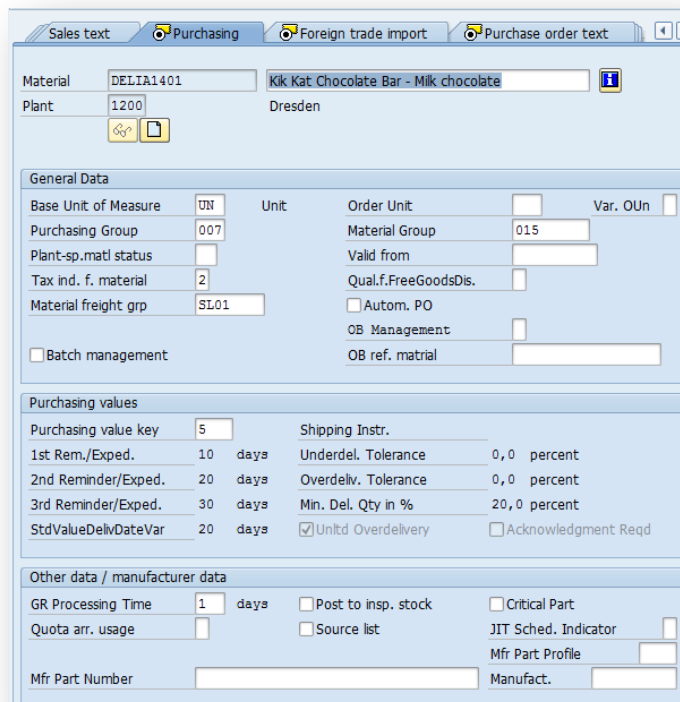


Figure 3.3.1.1-5. SAP, Purchasing.

➔ Suggestion: Almost every company carries out inspections of the purchased products to ensure their quality. But it is important to know that it does not imply that the option “post to inspection stock” should be chosen. What this option triggers is that when the goods arrive to the company they do not enter into the warehouse. Instead, they stay in an inspection area and a goods transfer has to be done in order to move them to the final storing place. Therefore, the employee should not select this option unless the process is really like this (the items do not get into the warehouse before having been inspected) and, in that case, he should be totally aware that the process is different and that it requires a step more.

MRP:

In this section the employee sets the Procurement type as F, because is an external delivery. The GR Processing Time is the same as the filled in “Purchasing”. As for Planned Delivery Time, it refers to the number of calendar days needed to obtain the material.

➔ Suggestion: The employee should be cautious with the Safety Stock fields. They can very useful to trigger Purchase Orders when the remaining stock is too low to attend the demand. Therefore, leaving this field blank could lead to problems regarding procurement due to the fact that the Purchasing Requisitions are not generated automatically.

- Safety Stock: Specifies the quantity whose purpose is to satisfy unexpectedly high demand in the coverage period.
- Min safety stock: The quantity that defines the lower limit for SS.

The screenshot displays the SAP MRP configuration for material DELIA1401 (Kik Kat Chocolate Bar - Milk chocolate) at plant 1200 (Dresden). The configuration is organized into three main sections:

- Procurement:** Procurement type is set to 'F'. Other fields include Batch entry, Special procurement, Quota arr. usage, Backflush, JIT delivery sched., Bulk Material, Prod. stor. location, Default supply area, Storage loc. for EP, and Stock det. grp.
- Scheduling:** GR Processing Time is set to 1 days. Planned Deliv. Time is set to 3 days. Other fields include SchedMargin key and Planning calendar.
- Net requirements calculation:** Safety Stock is set to 1000. Min safety stock is set to 500. Service level (%) is set to 90. Other fields include Safety time ind., STime period profile, Coverage profile, and Safety time/act.cov.

Figure 3.3.1.1-6. SAP, MRP.

Storing:

As our company is not the producer of the product, the concrete requisites about the storing should be provided or asked to the supplier. Apart from the conditions of the warehouse, the employee should ask about:

- Max. storage period: Maximum period for which the material may be stored.
- Min. Rem. Shelf Life: Minimum amount of time for which the material must keep upon goods receipt for the goods receipt to be accepted by the system.

The data provided about the storing is that it has to be in a cold environment (due to the fact that is chocolate). Besides, the maximum recommendable storing in a wholesale warehouse (in order to let a time margin to the retail warehouse without getting spoiled) should be no more than 60 days.

The screenshot shows the SAP 'Storing' configuration for material DELIA1401. The material name is 'Kik Kat Chocolate Bar - Milk chocolate'. The plant is 1200 (Dresden) and the storage location is 0002 (Fertigwarenlager). The 'General data' section includes fields for Base Unit of Measure (UN), Storage Bin, Temp. conditions (03), Container reqmts, CC phys. inv. ind., Label type, Lab.form, OBManagmnt, Unit of issue, Picking area, Storage conditions, Haz. material number, Number of GR slips, Appr.batch rec. req., and OB Ref. Material. The 'Shelf life data' section shows Max. storage period (60), Min. Rem. Shelf Life (1), Time unit (D), Total shelf life, Period Ind. for SLED (D), and Rounding rule SLED.

Figure 3.3.1.1-7. SAP, Storing.

➔ Suggestion: A field related to the storage conditions should be always filled. When the product is food, it should be the temperature; when is a special liquid, it should be the container requirements, etc. Missing this information could lead to loss profitability due to defective products.

Accounting:

The most important field in this section is the standard price. The reason is that the profit is calculated comparing the revenues from the sales with this value. Hence, an incorrect number could lead to enormous mistakes in terms of benefit and overall profit.

➔ Suggestion: This price should be always set taking into account not only the supplier price but also further costs, such as the maintenance of the products, the storing, the necessary electricity, the warehouse employees, etc. Thus, it should be periodically updated to check if the value is the correct one.

However, it is important not to mix these costs with others that can be added directly to each Sales Order (called “Conditions”). The difference is that those costs depend on the concrete sale (and they can change according to the customer), but the ones that should be included in this field are always the same (they are inherent to the product).

The Purchasing department along with the Accounting estimate that the value of the Kit-Kat will be 0,56 according to the expected price from the vendor plus the costs of managing the products with our company, storing them, etc.

The employee includes that information in the system. He also enters the Valuation Class 3100, which sets that the product is a trading good.

The screenshot displays the SAP Accounting interface for material 'DELIA1401' (Kik Kat Chocolate Bar - Milk chocolate) at plant '1200' (Dresden). The 'Current valuation' section is active, showing the following data:

General data	
Base Unit of Measure	UN Unit
Currency	EUR
Division	00
Valuation Category	
Current period	05 2014
Price determ.	<input type="checkbox"/> ML act.

Current valuation	
Valuation Class	3100
VC: Sales order stk	
Price control	S
Moving price	
Total Stock	0
Future price	
Proj. stk val. class	
Price Unit	1
Standard price	0,56
Total Value	0,00
<input type="checkbox"/> Valuated Un	
Valid from	

Buttons at the bottom: Previous period/year, Std cost estimate

Figure 3.3.1.1-8. SAP, Accounting.

3.3.1.2. Purchasing

The Purchasing Department of our company calls the Sales Department of Nestlé Deutschland AG in order to agree the conditions of the future transactions. Mrs Slutz is the head office manager and she says that at the moment she is in a meeting, but that she will send an e-mail with all the general data of the transactions, such as the address and contact.

They also stipulate that they have to be paid within the 15 days after the purchasing and that in terms of transport, they pay the costs.

About the deliveries, they explain that the minimum order is 1000 and the delivery time to Dresden is 2 days.

To begin with, the purchasing assistant creates a vendor record where she starts to add all the required information. Therefore, she uses the XK01 and starts to fill the required fields.

➔ Suggestion: If possible, this information should be got directly from the vendor –an online questionnaire, a shared database, etc. If not possible, the employee should get the information through a written way in order to avoid misspellings and other common errors. In any case, it is always useful to send a document with the data to the vendor so their personnel can check if every field is correct.

In this case the information related to the address is received by e-mail so our employee copies it into the Vendor Record:

The screenshot shows the SAP Vendor contact form (XK01) with the following data entered:

- Name:** Title: Company; Name: Nestlé Deutschland AG
- Search Terms:** Search term 1/2: DELIA14
- Street Address:** Street/House number: Lyon Strasse 23; Postal Code/City: 60528 Berlin; Country: DE Germany; Region: 11 Berlin
- PO Box Address:** PO Box: ; Postal code:
- Communication:** Language: English; Telephone: 04943543656 Extension: 012; Fax: 57664567868 Extension: ; E-Mail: verbraucherservice@de.nestle.com; Data line: ; Telebox:

Figure 3.3.1.2-1. SAP, Vendor contact

In the e-mail they stipulate that they need to be paid the net amount of money within 15 days after the purchasing. She looks for the option in payment terms and selects ZR01.

She also looks for the option in Incoterms that ensures that the vendor pays for the whole transport. This option is CPT or Carried Paid To, and she writes the named place of destination, in this case Dresden.

➔ Suggestion: The Sales department should be always aware of the Incoterms and make sure that both parties agree with it. The reason is that the transportation costs can be a big proportion of the costs of a product and the conditions related to it are set through the Incoterms. Even if they have said “We pay the transport”, make clear that technically, the Incoterm will be a “CPT” because there are other similar options where the conditions (such as the insurance) are different.

Our employee fills the purchasing data:

The screenshot displays the SAP Vendor purchasing data form, organized into four main sections:

- Conditions:** Order currency (EUR), Terms of payment (ZR01), Incoterms (CPT Dresden), Minimum order value (1000), Schema Group, Vendor (Standard procedure vendor), Pricing Date Control (No Control), and Order optim.rest.
- Sales data:** Salesperson (Mrs Slutz), Telephone (49754874968), and Acc. with vendor.
- Control data:** A grid of checkboxes for various control options such as GR-Based Inv. Verif., AutoEvalGRSetmt Del., Acknowledgment Reqd, and Shipping Conditions (01).
- Default data material:** Purchasing group, Planned deliv. time (2 Day(s)), Confirmation Control, Unit of measure grp, and Rounding Profile.

Figure 3.3.1.2-2. SAP, Vendor purchasing data.

And when finishes, she sends a copy to Nestlé AG in order to make sure that all the information is correct.

Furthermore, she now begins to create a link between the vendor and the material through the transaction ME01 - Adding valid source of supply.

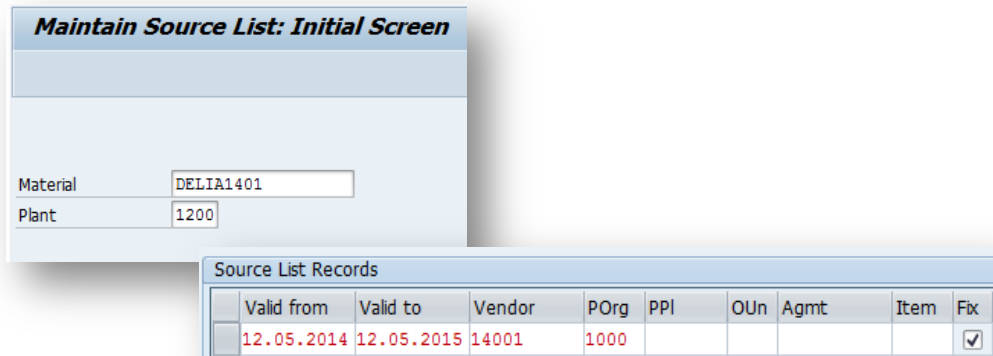


Figure 3.3.1.2-3. SAP, Source list.

Where she also sets Nestlé as the predefined vendor for Kit-Kat.

Two days later, Mrs Slutz calls back and a talk about the prices begins. She says that the standard price for 1000 units is 0,40€/unit. Our Purchasing Manager negotiates a price for bigger quantities, and they agree that for 5000 units the price will be 0,32€/unit.

Thus, our company makes the first order of 5000.

The Purchasing Manager creates during the conversation a Purchasing Information Record through the transaction ME11. He fills all the fields that have been agreed.

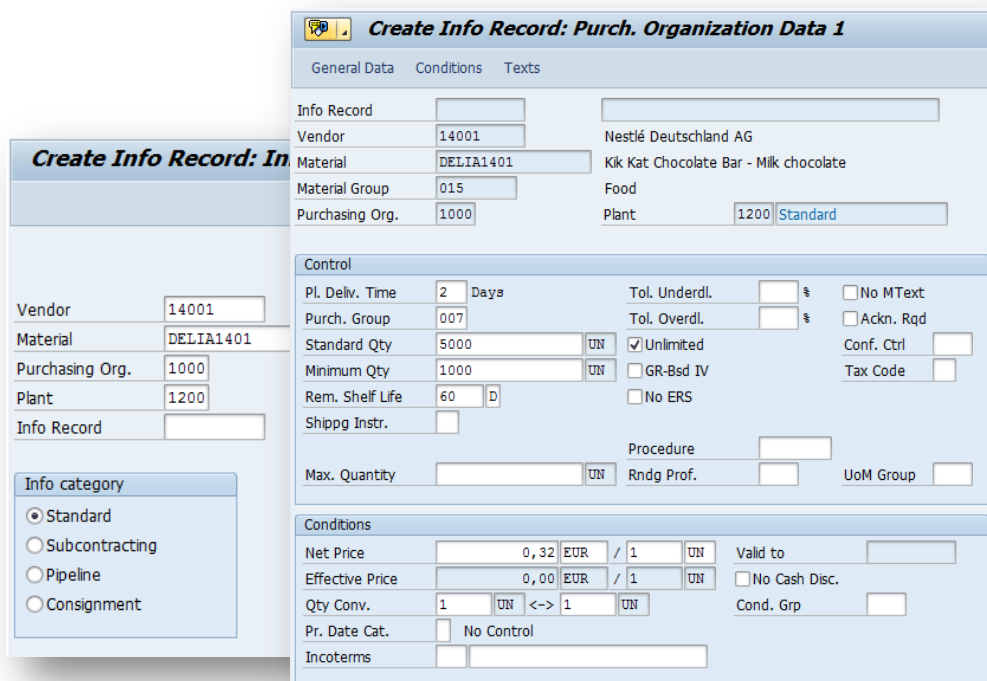


Figure 3.3.1.2-4. SAP, Purchasing Info Record.

He also creates the related Purchase Order through the transaction ME21N:

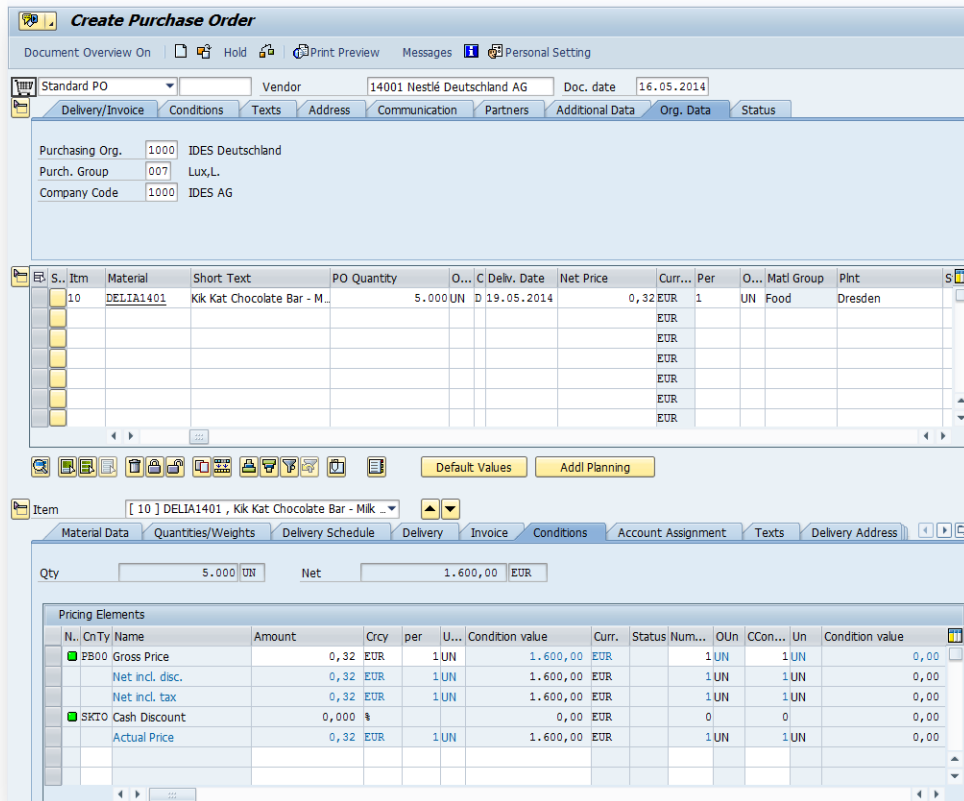


Figure 3.3.1.2-5. SAP, Purchasing Info Record.

Three days later, the 19/05/2014, the material arrives to Dresden plant. The quantity of the products is the correct one.

At this point, the worker in charge of the arriving goods checks the purchased products. He confirms that the quantity is the one that appears in the Delivery note and in the Purchase Order. That is the case, so he starts a transaction MIGO to manage the Goods Receipt. He sets the final warehouse to be the #2.

After filling the required fields about the receiver of the goods, etc, he posts the Goods Receipt.

➔ Suggestion: It is important to include in this transaction as much information as possible. The reason is that a missing field in this area can affect: the customer (the quantity and quality inspection checking), the warehouse management (through the unloading point) and the accounting area (through the quantity of products in order to check the invoice).

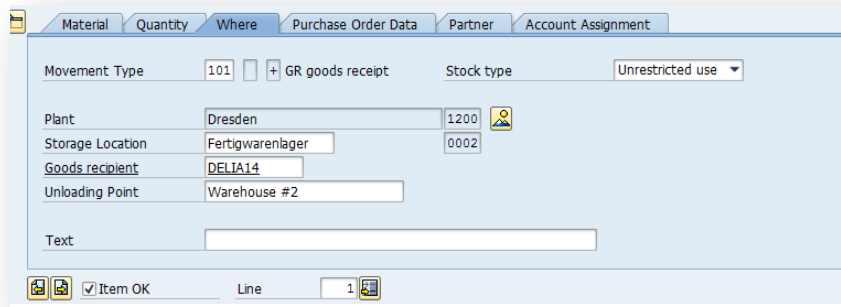


Figure 3.3.1.2-6. SAP, Goods Receipt.

Some days later an invoice arrives to our company with the quantity to pay: 1856 € including taxes.

An employee of the Accounting department takes it and starts an Invoice verification through the transaction MIRO. She includes the taxes and checks that the quantity provided by the system is the same that appears in the real invoice. That is the case, because the balance is 0. Therefore, she saves the document.

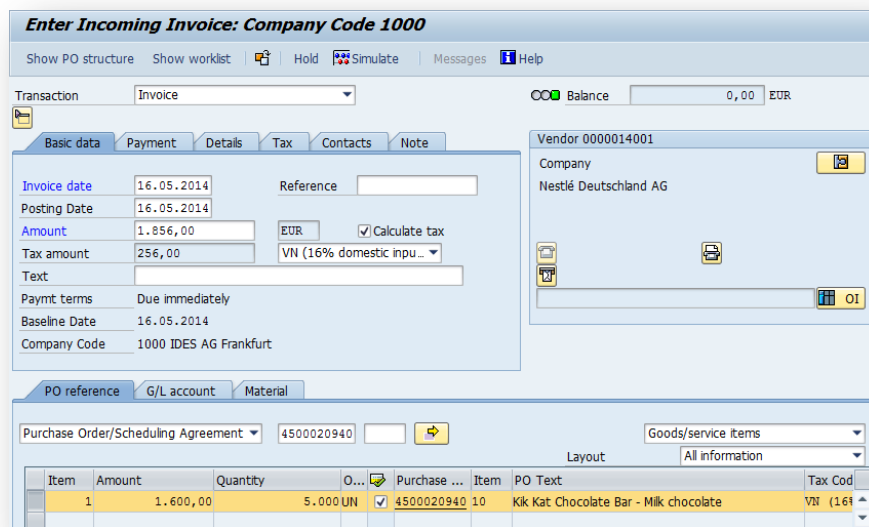


Figure 3.3.1.2-7. SAP, Invoice verification.

At this point, the whole purchasing process has concluded and the first arrival of Kit-Kat (5000 units) is in our plant.

Some days later, the Purchasing department receives a call from the Sales department in order to make a Purchase Requisition. The reason is that they have been selling so much that the 5000 units are almost over. Therefore, they ask for another 5000 Kit-Kats.

Thus, the Purchasing department employee starts a Purchase Requisition, a transaction ME51N:

St.	Item	A	I	Material	Short Text	Quantity	Unit	C	Delv. Date	Matl Group	Plant	Stor. Loc.	PGr
	10			DELIA1401	Kik Kat Chocolate Bar - Milk	5.000	UN	D	26.05.2014	Food	Dresden	Fertigwaren	007

Figure 3.3.1.2-8. SAP, Purchase Requisition.

➔ Suggestion: In the case of Purchase Requisitions, the department should have a well established maintenance plan in order to have good records of who requires the products, why, etc. In that way, proper information about the procurement can be collected that can very useful in planning and optimizing the process.

In our Purchasing department we have implanted a standardized method to process the Requisitions and different fields need to be always filled, such as: Header (with basic information), Requisitioner and Tracking Number.

The employee fills them and saves the Purchase Requisition.

At this point he uses the transaction MD04 to check the stock and to trigger the Purchase Order:

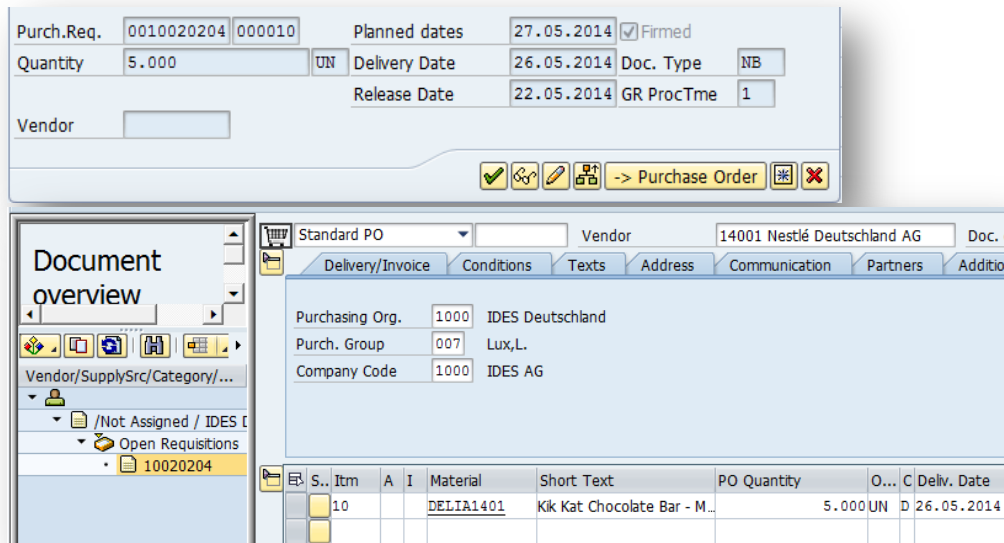


Figure 3.3.1.2-9. SAP, Trigger Purchase Order.

The way and the suggestions related to the Purchase Order were exposed before.

Two days later, the 26/05/2014, the material arrives to Dresden plant. The quantity of the products is not the correct one, there are 10 units missing. However, the quantity in the delivery note is 5000.

The warehouse worker checks the items that have just arrived and he realizes that there are 10 Kit-Kats missing but in the delivery note it is the supposed quantity so apparently there has been a problem.

Then, he starts a transaction MIGO where he points out the difference between the quantities.

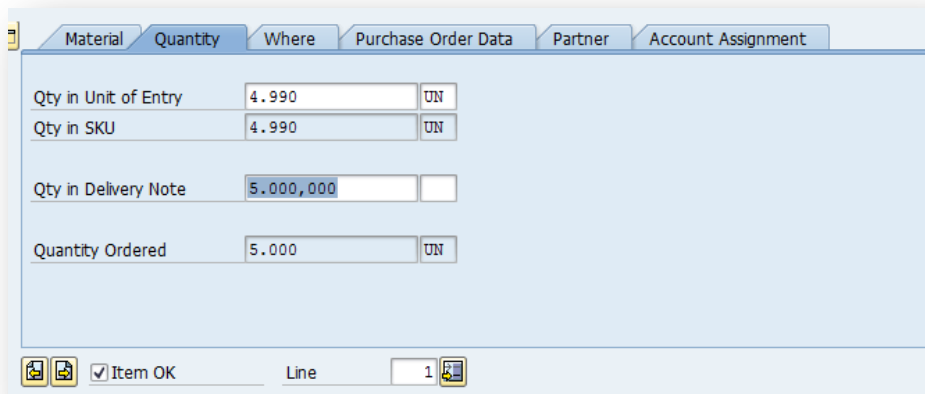
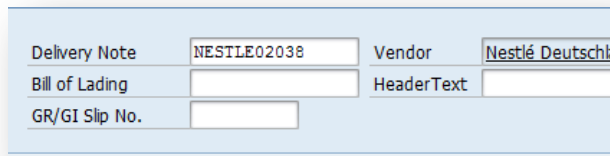


Figure 3.3.1.2-10. SAP, Goods Receipt – Quantity.

➔ Suggestion: As I already said before, it is very important to include in this transaction as much information as possible. Especially in the cases like this one where there can be a potential error or discrepancy between our company and the supplier.

In these situations, the employee should fill the same fields than before, but focusing on other that link the transaction with the supplier such as the Delivery Note (NESTLE02038).



The screenshot shows a SAP form for a Goods Receipt - Delivery Note. It contains several input fields: 'Delivery Note' with the value 'NESTLE02038', 'Vendor' with 'Nestlé Deutsche', 'Bill of Lading', 'HeaderText', and 'GR/GI Slip No.'.

Figure 3.3.1.2-11. SAP, Goods Receipt – Delivery Note.

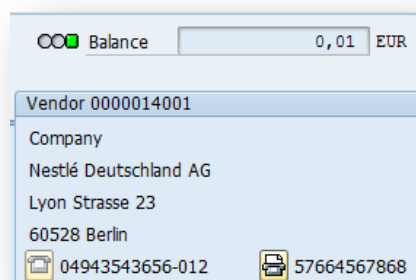
Some days later, the invoice from Nestlé AG arrives and, as expected, the amount is 1600€ (without taxes), which corresponds to 5000 units instead of 4990 units.

Therefore, the Accounting department from our company calls the Sales department of Nestlé AG. They explain what happened and they give the information about the corresponding Delivery Note.

Since they have all the process controlled by an ERP as well, soon later they call back and they state that they have checked the documents related to the Delivery Note NESTLE02038 and, indeed, the warehousing department made a mistake and loaded only 4990 units of Kit-Kats.

One hour later, a new invoice arrives from Nestlé AG with an amount of 1596,8€ (without taxes) and 1852,3€ (with taxes).

At this point, the Accounting department starts a transaction MIRO and check that the calculated amount coincides with the one in the invoice. That is the case, so they save the document and the process is finished.



The screenshot shows the SAP Invoice verification (MIRO) form. At the top, it displays 'Balance' with a value of '0,01' and the currency 'EUR'. Below this, the 'Vendor' field is populated with '0000014001'. The 'Company' field shows 'Nestlé Deutschland AG' with the address 'Lyon Strasse 23, 60528 Berlin'. At the bottom, there are two document number fields: '04943543656-012' and '57664567868'.

Figure 3.3.1.2-12. SAP, Invoice verification.

3.3.1.3. Sales and Distribution

The Sales Department of our company receives a call from Ms. Bauer, the Head of Purchasing of EcoMarket. EcoMarket is a German supermarket chain that is currently growing and expanding over the whole country. She would like to know if our plant in Dresden could supply Kit-Kats in its main store in Hannover. Our company exposes that we are in the position to send products to Hannover.

Therefore, Mr Bauer gives the general information about EcoMarket, its contact as well as her assistant's one. She says that she will contact soon in order to talk more precisely about the contract.

At this point, it seems logical to create a customer record where add all the information. Thus, the person who just got the call from the Sales Department should begin a transaction XD01 and introduce the given data:

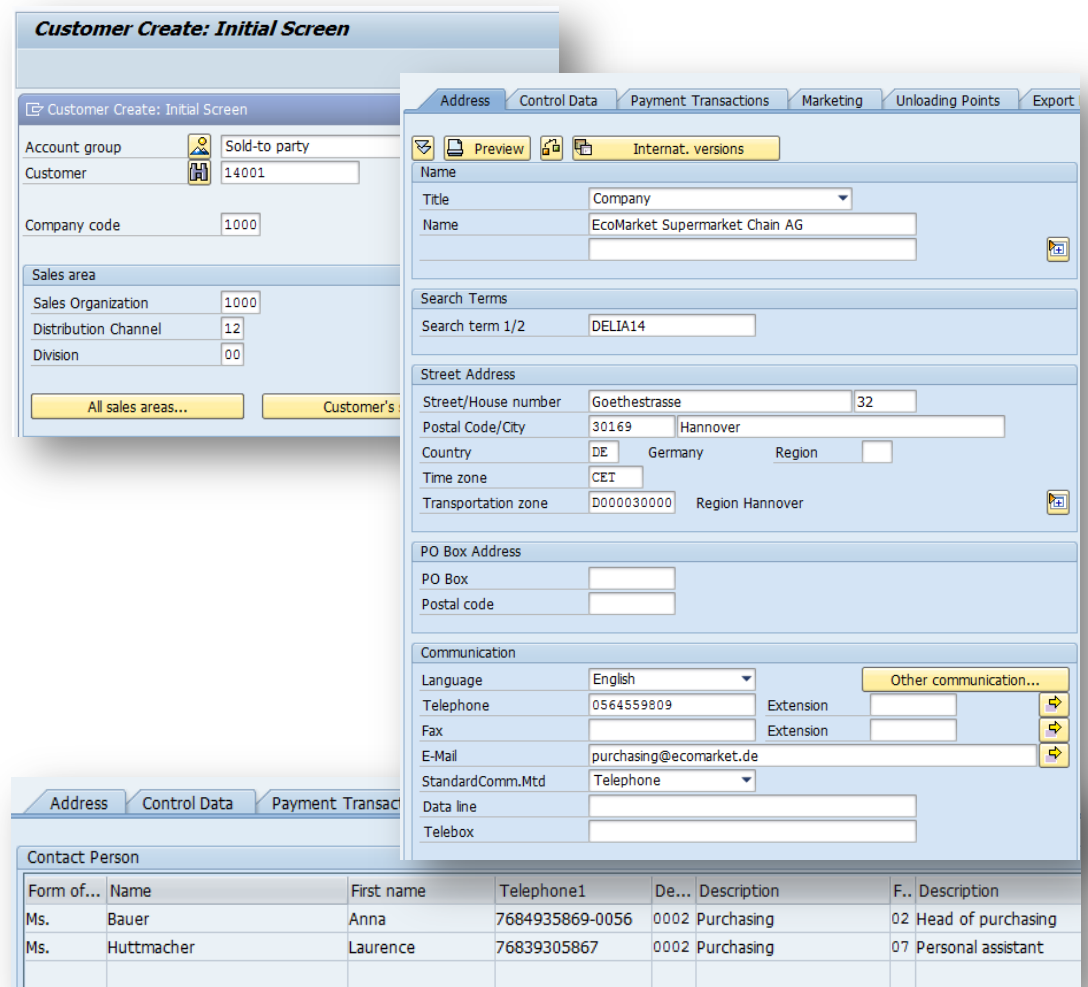


Figure 3.3.1.3-1. SAP, Creation of Customer Records.

It is important to firstly, be sure that the information introduced is totally accurate. An error in a Customer Master Records could incur in many others, as it is data used all throughout the system and in multiple transactions.

➔ Suggestion: Companies should have a totally standardized procedure so the employees know exactly what information is relevant and how to fill the fields in the same way. Moreover, another good practice is to have at least two persons introducing and maintaining the information so they give feedback and revise each other or send the information to the customer before doing any further transaction to make sure that everything is correct. To finish with, another method could be to have a data responsible (see section 2.4.1) that checks and maintains Master Data.

One week later, the Sales Department receives another call from Mrs Hutmacher. She wants to discuss further about the conditions of the agreement. She would like to know how much it would to get an order of 400 Kit-Kat in two weeks (the 06/06/2014), and she would need a Quotation in maximum one week. She adds some further information about the unloading point, such as the different warehouses that they have, because she would need the delivery concretely in the Food Warehouse.

While the conversation is taking place, the Sales employee opens an Inquiry Transaction VA11 to write down the requirements that Mrs. Hutmacher is asking for. When he tries to put “400” the system notifies that the minimum order quantity is 500, so he asks her if it is fine to rise the number until 500 and she agrees.

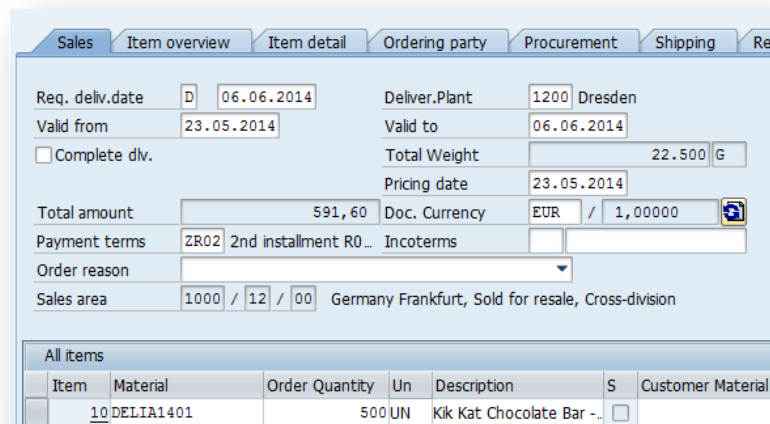


Figure 3.3.1.3-2. SAP, Inquiry.

After the call, the employee finishes the Inquiry. He also adds the new information about the different unloading points through the transaction XD02, change customer data.

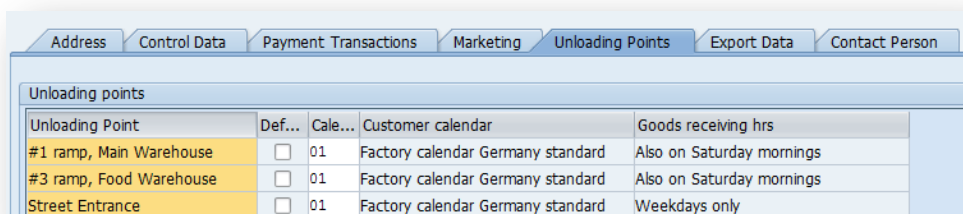


Figure 3.3.1.3-3. SAP, Unloading points.

➔ Suggestion: The tab “Unloading points” should always have information about the different options regarding the warehouses or entrances of the customer. The reason is not only that it is useful to know the different places where the delivery should be unloaded, but also that it adds information about the calendar and the receiving hours, which can be important.

➔ Suggestion: From the moment that a customer has established different unloading points, use them even before the final Sales Order. In that way potential misunderstanding about the delivery will be prevented from the beginning.

At this point, the employee communicates the information to the Sales Head Manager in order to agree the conditions. The Manager is a little unsure about the price because Hannover is out of the region that the Dresden plant usually supplies. They finally decide to use the normal price and charge the shipping costs. Furthermore they add 0,12€/unit in calculated costs derived from the preparations of the shipment. They are relatively high because the quantity to deliver is low.

Then, the employee creates a Quotation through the transaction VA21 and he references the previous Requisition. He creates it with the reference 30 days due net, Invoicing dates per month, and shipping costs and calculated costs included. He prints it in a PDF format and sends it to EcoMarket purchase department.

➔ Suggestion: If there is a previous Requisition, it is very recommendable to use the reference number to create the Quotation. This is not only for copying the already filled fields and therefore save time, but also to “finish” the flow of documents.

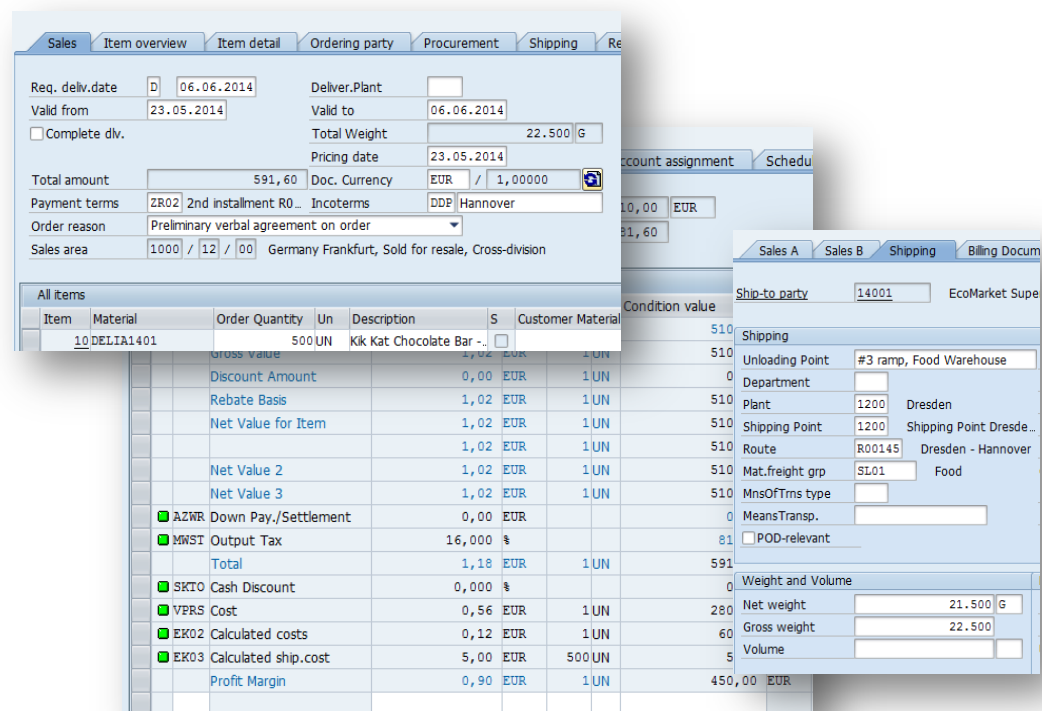


Figure 3.3.1.3-4. SAP, Quotation.

The company receives an answer two days later: the quotation has been rejected, and the reason it is that they consider that the price is too high. The employee calls Mrs Bauer and, after some discussion, they make a deal – 5% discount with an order of 1000 instead of 500. She also adds that they need the cargo at 08:30.

➔ Suggestion: As I have just said, it is important to have a correct flow of documents within the SAP ERP. Hence, if a quotation has been rejected, before creating a new one, the previous one should be changed in order to make clear that the quotation has been cancelled and the reasons involved.

Therefore, the employee changes the quotation with the transaction VA22:

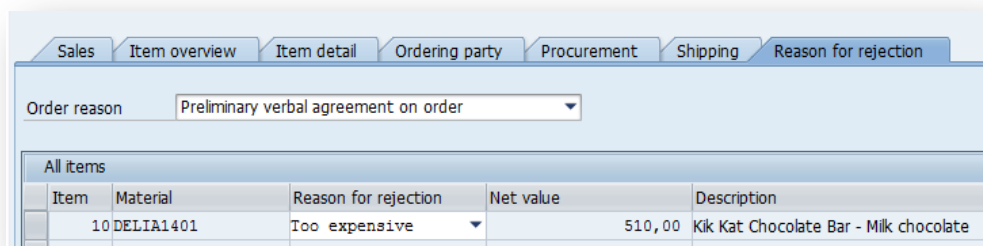


Figure 3.3.1.3-5. SAP, Rejection.

Now, he creates a new Quotation where he uses the new conditions. Afterwards he sends it to EcoMarket Purchasing department again and they agree with it.

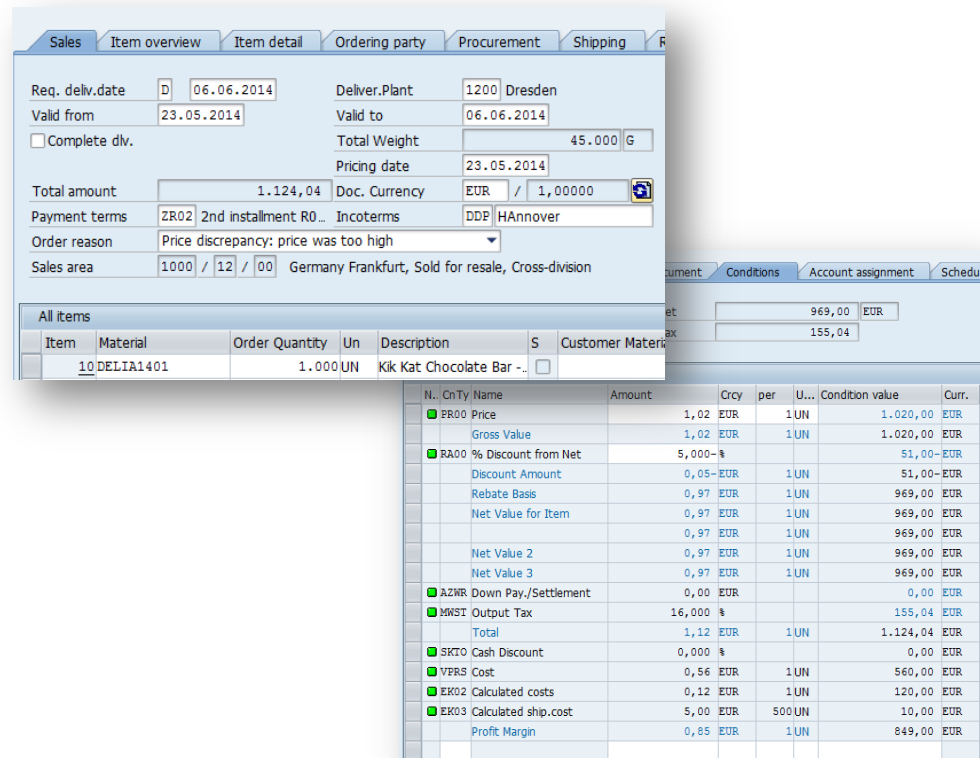


Figure 3.3.1.3-6. SAP, Second Quotation.

➔ Suggestion: Even if it is not a compulsory field, it is interesting to always give reasons to every action. As I exposed before, the rejections should state their cause, and the same applies for the Quotations. So, as it can be seen, there is always an order reason: for the first is “Preliminary verbal agreement on order” and for the second “Price Discrepancy: price was too high”. This information can be very useful in future.

As the Quotation has been already accepted, he transforms it into an order through the transaction VA01. The employee then updates the arrival time. He should put 8:30 but by error he writes 8:00.

The system then shows the following notification:

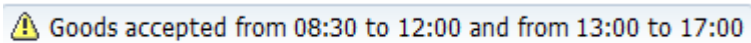


Figure 3.3.1.3-7. SAP, Notification unloading hours.

This message appears because the unloading hours for the Food Warehouse were stipulated before. Therefore, he corrects the hour and fills the rest of the hours to ensure that the product is ready at the adequate time.

➔ Suggestion: A good practice is to concrete all the hours so the schedule is already arranged and consequently there will not be problems regarding the shipping (at least not because it was not planed correctly). The employee should be aware of the meaning of:

- Transportation Pln. Date: The date by which the transportation should be arranged so that the delivery can arrive on time.
- Material Avail. Date: The date by which sufficient quantities of the item must be available for picking to begin.
- Loading Date: The date by which picking and packing must be completed so that the goods are ready for loading and for the delivery to reach the customer on time.
- Goods Issue Date: The date on which the goods must physically leave the shipping point to reach the customer on time.

A screenshot of the SAP Shipping transaction VA01. The interface has three tabs: "Sales", "Shipping", and "Procurement", with "Shipping" selected. Below the tabs, there are several input fields and labels:

- Confirmed quantity: 1.000 UN
- Delivery date: 06.06.2014
- Goods issue date: 05.06.2014
- Loading date: 04.06.2014
- Material avail.date: 03.06.2014
- Transportation Pln. Date: 03.06.2014
- Arrival time: 08:30
- GI Time: 08:0...
- Loading Time: 12:00
- Matl Staging Tme: 10:0...
- Tr. Plan. Time: 09:00
- Shipping point: 1200 Shipping Point Dresden
- Route: R00145 Dresden - Hannover
- Route schedule: (empty field)
- Delivery block: (dropdown menu)

Figure 3.3.1.3-8. SAP, Shipping.

At this point the Delivery begins. The first step is simply to “open” the delivery. Hence, the employee starts a transaction VL01N and saves the delivery.

Some days later, the order to pick and load the products it is made through the transaction VL02, Change Outbound Delivery. The employee fills the Pick Quantity with 1000 units and saves the document:

Ship-to party	14001	EcoMarket Supermarket Chain AG / Goethestrasse 32 / 30169 Hannover
Delivery	80023028	Document Date 23.05.2014
Pland Gds Mvmt Date	05.06.2014	Total Weight 45.000 G
Act. Gds Mvmt Date		Number of packages

All Items									
Item	Material	Plnt	SLoc	Delivery quantity	SU	Pick Quantity	SU	Batch	
	10 DELIA1401	1200	0002	1.000	UN	1000	UN		

Figure 3.3.1.3-9. SAP, Delivery picking.

➔ Suggestion: In the transaction VL02 there is an option to check the documents flow. It is a very useful tool to get to know which is the current step in the process. It also permits to check the documents numbers’ chain.

Document	On	Status
Quotation 0020002136	23.05.2014	Completed
Standard Order 0000022261	23.05.2014	Completed
• Delivery 0080023028	23.05.2014	Open

Figure 3.3.1.3-10. SAP, Documents Flow 1

To finish the delivery, the goods have to be posted so they can leave the plant. The way this is done is through the same transaction VL02 but in this case the action is “post goods issue”. As it is shown below, the documents flow automatically is updated:

Document	On	Status
Quotation 0020002136	23.05.2014	Completed
Standard Order 0000022261	23.05.2014	Completed
• Delivery 0080023028	23.05.2014	Being processed
• Picking request 20140523	23.05.2014	Completed
• GD goods issue:delvy 4900042074	23.05.2014	complete

Figure 3.3.1.3-11. SAP, Documents Flow 2

3.3.2 Inadequate management

In the section 3.3.2 Adequate Management I have showed how to correctly carry out the basic processes in a Wholesale company whose main activity in the case is purchasing Kit-Kats from Nestlé Deutschland and sell them to a supermarket chain called EcoMarket.

The objective of this section is to perform a poor data maintenance in order to prove how it affects to the overall profitability of the company, which has been the aim of this Thesis from the beginning.

Thus, the case will be exactly the same but the procedures will have some errors that will drive to higher costs, customer dissatisfaction, low efficiency, etc.

The main cause of the errors will be “Data entry by employees” since it is the most common reason in real companies, as it was exposed in the section Sources of poor Data Quality (2.3.4).

As it has already been exposed, there will be arrows to remark the errors and the profitability losses due to poor Master Data Management.

Note: Many of the processes will be briefly described, since the complete procedure was detailed in Adequate Management. Hence, in this section I will focus mainly in the bad practices and errors.

3.3.2.1. Materials Management

The process is exactly the same as before: the first necessary step is to create the new material in our SAP ERP system. Hence, our employee creates a new Material Master record that is going to be a trading good since we do not produce it nor we sell it to the final customer –we just trade with it.

The employee uses the transaction MMH1 and begins to fill the required fields.

Basic data:

➔ Error: Since the way the data is filled in the system is not standardized, the employee just fills the very necessary fields and he does not ask for further information. This could lead to inconsistencies or other problems as the process evolves.

In this case the important values that are missing are volume, weight, and packaging. They could potentially affect the shipping and delivery.

The screenshot displays the SAP Basic Data maintenance interface for material DELIA1405. The material name is 'Kik Kat Chocolate Bar - Milk chocolate'. The interface is divided into several sections:

- General data:** Includes fields for Base Unit of Measure (UN), Unit, Material Group (015), Old material number, Ext. Matl Group, Division, Lab/Office, Product allocation, Prod.hierarchy, X-plant matl status, Valid from, Assign effect. vals, GenItemCatGroup (NORM), and Standard item.
- Dimensions/EANs:** Includes Gross Weight, Net Weight, Volume, Weight unit (G), Volume unit (ML), Size/dimensions, EAN/UPC, and EAN Category.
- Packaging material data:** Includes Matl Grp Pack.Matls and Ref. mat. for pkg.
- Basic Data Texts:** Includes Languages Maintained (0), Basic Data Text, and Language.

Figure 3.3.2.1-1. SAP, Basic Data.

Sales:

➔ Error: As I explained in the section Adequate Management, it is very useful to have at least two persons in the section of entering Customer Master Data, since an error in these records has consequences all over the process. Another option is the “Data responsible” or steward that was exposed in the section Master Data Maintenance (2.4.3). This prevents minor mistakes that could be made by the employees who enter data.

In this case, the employee makes a mistake and writes that the Tax Classification of the material is 0 - No tax.

➔ Error: Another very common problem in companies is that the departments are not well communicated and therefore people from one section do not know concrete information about other. The personnel should do an effort in order to get to know all the important fields of the Material Master Data, even if they are from different areas. However, in this case the employee does not call the Sales and Distribution department in order to know the minimum order quantity and the minimum delivery quantity, so he leaves them empty.

The screenshot shows the SAP Sales organization configuration for material DELIA1405. The material name is 'Kik Kat Chocolate Bar - Milk chocolate'. The sales organization is '1000' (Germany Frankfurt) and the distribution chain is '12'. The material is sold for resale. The 'General data' section includes fields for Base Unit of Measure (UN), Unit, Division, Sales unit, Unit of Measure Grp, X-distr.chain status, DChain-spec. status, Delivering Plant, Material Group (015), and Food. A 'Cash discount' checkbox is checked. The 'Tax data' section shows a table with columns for Country, Tax category, and Tax classification. The entry is for Germany, MWST Output Tax, and No tax. The 'Quantity stipulations' section includes fields for Min.order qty, Min. dely qty, and Delivery unit.

Figure 3.3.2.1-2. SAP, Sales organization.

➔ Error: Concerning the data about the Sales: general/plant, the same problem can appear. The information required here is from different departments and if the employee does not make the effort to collect it, the missing values can lead to future problems.

The employee just fills what is mandatory and leaves blank some important fields regarding the shipping such as Setup time and Proc. time.

The screenshot shows the SAP Sales plant configuration for material DELIA1405. The material name is 'Kik Kat Chocolate Bar - Milk chocolate'. The plant is '1200' (Dresden). The 'General data' section includes fields for Base Unit of Measure (UN), Unit, Replacement part, Gross Weight (G), Net Weight, Availability check (01), Daily requirements, and Material freight grp. A 'Batch management' checkbox is unchecked. The 'Shipping data (times in days)' section includes fields for Trans. Grp (0001), On palettes, LoadingGrp (0002), Forklift, Setup time, Proc. time, and Base qty (UN).

Figure 3.3.2.1-3. SAP, Sales plant.

Purchasing:

➔ Error: In terms of entering data, there is one thing that is worse than leaving blank an important field –to fill it with wrong information. As it has been already said, in SAP ERP there are mandatory fields that need to be filled in order to continue to the next tab. This can be a problem when an employee does not know about the correct value, so he/she just put any of them. This can create enormous problems if it is not corrected shortly after.

The employee continues filling just the compulsory fields. He does not know which is the Purchase value key that is used in our company, so he just puts the number one (which is incorrect). Besides, he leaves empty a very important field in terms of scheduling: GR Processing Time.

The screenshot displays the SAP Purchasing interface for material 'DELIA1405' (Kik Kat Chocolate Bar - Milk chocolate) at plant '1200 Dresden'. The 'General Data' section includes fields for Base Unit of Measure (UN), Purchasing Group (007), and Material Group (015). The 'Purchasing values' section shows 'Purchasing value key' set to '1', '1st Rem./Exped.' at 10 days, '2nd Rem./Exped.' at 20 days, '3rd Rem./Exped.' at 30 days, and 'StdValueDelivDateVar' at 20 days. The 'Other data / manufacturer data' section shows 'GR Processing Time' as an empty field, 'Quota arr. usage' as an empty field, and 'Mfr Part Number' as an empty field.

Figure 3.3.2.1-4. SAP, Purchasing.

MRP:

➔ Error: Information related to Net requirements calculation is crucial if an automatic system of Purchasing wants to be used. For example, when the stock of a given product decreases below the Safety Stock value, an automatic Purchase Requisition is carried out. Therefore, if these fields are not filled, it is impossible to make the process efficient.

Furthermore, even if the company is not using an automatic system of Purchasing, it is a way to link the Purchasing and the Sales department, since through the Safety Stock the Purchasing employees get to know how the demand how a concrete product is evolving, or in which range of quantities the product is sold.

The employee just fills the required fields (procurement type and planned deliv. time) and leaves the rest empty.

Figure 3.3.2.1-5. SAP, MRP.

Storing:

➔ Error: When creating Master Data records, there are some fields that are not really important in different areas, but they are relevant because they set certain characteristics that need to be taken into account. This is the case of the General Data regarding the storing conditions. Incomplete data in these fields can easily drive to defective products that create customer dissatisfaction, rejections, poor quality brand perception, etc.

As they are not mandatory, the employee does not provide any information about the storing, not even filling the fields regarding the temperature conditions or the maximum storage period.

Figure 3.3.2.1-6. SAP, Storing.

Accounting:

➔ Error: As I already explained in the “Adequate Management” part, the standard value is very important since it takes part in the profit calculation. Thus, a wrong or not-updated value would drive to have an incorrect figure of the overall profitability.

The Purchasing department along with the Accounting estimate that the value of the Kit-Kat will be 0,56 according to the expected price from the vendor plus the costs of managing the products with our company, storing them, etc.

The employee therefore enters the information:

The screenshot shows the SAP Accounting interface for material 'DELIA1405' (Kik Kat Chocolate Bar - Milk chocolate) at plant '1200' (Dresden). The interface is divided into several sections:

- General data:** Base Unit of Measure (UN), Unit, Valuation Category, Currency (EUR), Current period (05 2014), Division, Price determ., and ML act. checkbox.
- Current valuation:** Valuation Class (3100), VC: Sales order stk, Price control (S), Moving price, Total Stock (0), Future price, Proj. stk val. class, Price Unit (1), Standard price (0,56), Total Value (0,00), Valuated Un checkbox, and Valid from.

Figure 3.3.2.1-7. SAP, Accounting.

Note: In this ”Inadequate Management” section I could use the case that the employee does not know about the real price and he merely uses the price of the supplier. Nevertheless, since the objective is the part is to compare the results with the Adequate Management, I believe it is more useful to have the same basis to calculate the profitability and that is the reason why the standard price is the right one.

3.3.2.2. Purchasing

The Purchasing Department of our company calls the Sales Department of Nestlé Deutschland AG in order to agree the conditions of the future transactions between us. Mrs Slutz is the head office manager and she says that at the moment she is in a meeting, but that she will send an e-mail with all the general data of the transactions, such as the address and contact.

They also stipulate that they have to be paid within the 15 days after the purchasing and that in terms of transport, they pay the costs.

About the deliveries, they explain that the minimum order is 1000 and the delivery time to Dresden is 2 days.

The steps are the same as the followed in the “Adequate Management” section. To begin with, the purchasing assistant begins to create a vendor record where she adds all the required information. Therefore, she uses the XK01 and starts to fill it.

As it was exposed in section Sources of poor data quality (2.3.4), the most common cause of problems in Master Data are derived from errors in entering data. Consequently, if there is not a well-managed system that permits to control and correct the data from the beginning, the mistakes can spread throughout the whole system.

➔ Error: In this case, the employee makes a misspelling error: The correct name of the street is “Lyon”, not “Lion”. Another error is to simply put “Nestlé” as the name of the company, since the concrete company is the German branch, not the whole corporation.

➔ Error: Furthermore, he does not include any contact, which presumably will provoke a time loss in the future.

The screenshot shows the SAP Vendor contact form with the following data entered:

- Name:** Title: Company, Name: Nestlé
- Search Terms:** Search term 1/2: della14
- Street Address:** Street/House number: Lion Strasse, Postal Code/City: 60528 Berlin, Country: DE, Region: (empty)
- PO Box Address:** PO Box: (empty), Postal code: (empty)
- Communication:** Language: English, Telephone: (empty), Extension: (empty), Fax: (empty), Extension: (empty), E-Mail: (empty), Data line: (empty), Telebox: (empty)

Figure 3.3.2.2-1. SAP, Vendor contact

In the e-mail they stipulate that they need to be paid the net amount of money within 15 days after the purchasing. She looks for the option in payment terms and selects ZR01.

She also looks for the option in Incoterms that ensures that the vendor pays for the transport. She finds the option DDP (Delivered Duty Paid) and she uses it.

➔ Error: The employee is assuming the incoterms without having a feedback from the supplier. She has used DDP but the correct option is CPT, since the supplier has said that they would pay for the transport costs (CPT), but they did not say anything about the risks or insurances (DDP).

After that, she saves the Vendor Master Record, but she does not send a copy to Nestlé Deutschland in order to make sure that the data is correct.

➔ Error: The core problem is not to make a mistake, but rather not to make sure that the information is right through a feedback or other method. After saving, it is very recommendable to send a copy to Nestlé in order to check the general information.

Furthermore, she creates a link between the vendor and the material and sets Nestlé as the predefined vendor for Kit-Kat.

Two days later, Mrs Slutz calls back and a talk about the prices begins. She says that the standard price for 1000 units is 0,40€/unit. Our Purchasing Manager negotiates a price for bigger quantities, and they agree that for 5000 units the price will be 0,32€/unit.

Thus, our company makes the first order of 5000.

The Purchasing Manager creates the related Purchase Order through the transaction ME21N:

Eq.	S.	Itm	A	I	Material	Short Text	PO Quantity	O...	C	Deliv. Date	Net Price	Curr...	Per	O...	Matl Gr
		10			DELIA1405	Kik Kat Chocolate Bar - M...	5.000	UN	D	26.05.2014	0,32	EUR	1	UN	Food

Figure 3.3.2.2-2. SAP, Purchasing Info Record.

➔ Error: Note that the employee does not do an info record where he puts the price agreed. What this provokes is that in next Purchase Order the price has to be entered again, so it is a loss in efficiency.

Three days later, the 26/05/2014, the material arrives to Dresden plant. The quantity of the products is the correct one.

At this point, the worker in charge of the arriving goods checks the purchased products. He confirms that the quantity is the one that appears in the Delivery note and in the Purchase Order. That is the case, so he starts a transaction MIGO to manage the Goods Receipt only with the mandatory information.

➔ Error: Even if is not compulsory, it is very recommendable to determine the information about who is going to receive the products. Problems related to this matter can easily appear if for example the material arrives to a different warehouse or the person that ask for the order is not even informed that it has already arrived. Thus, it is a very bad practice not to write the final warehouse or the goods recipient.

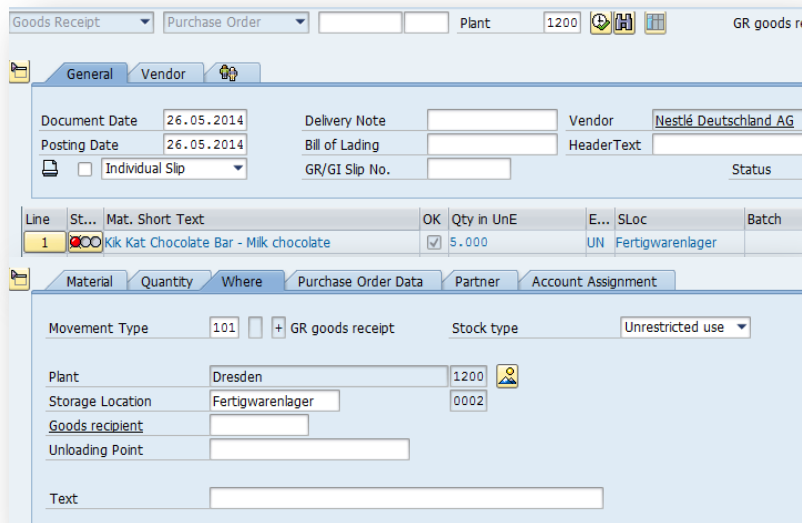


Figure 3.3.2.2-3. SAP, Goods Receipt.

➔ Error: Another problem can occur due to the simple fact of not specifying the Delivery Note of the supplier. The reason is that, if there is any problem with the products, it is necessary it have a reference of the document so the supplier can adequately check it.

The following day, a worker of the warehouse realizes that, due to the transportation, 200 of the chocolate bars are partially deformed.

He notifies it to the Purchasing department and they immediately decide to solve the problem. They check the conditions of the transport and they see that, theoretically, the incoterm is a type DDP (Delivered Duty Paid), so Nestlé should pay for the risks and the insurance of the transport as well. Hence, they decide to call the Sales department.

➔ Profitability loss: At that moment, they look in the Vendor Master Record in order to take the telephone number but is not there. There is actually no contact information. Then, they begin to search for the person who got the call from Mrs Slutz, but she is not there so they finally decide to check the numbers in the registered calls from the company phone. All those employees wasting time in searching a Vendor number just because the contact information was not entered when the Vendor was created is a great profitability loss.

➔ Profitability loss: In addition, before calling her they want to check the Goods Receipt in order to refer to the concrete Delivery Note number. They check the Good Receipt in the system, but the Delivery Note is not marked, so one of the employees has to call to the warehouse and ask directly. In the warehouse they begin to look for the paper and, after an hour, they find it and they call back to the Purchase department. This is also a profitability loss, since both resources and time are banally wasted.

They finally find the number, so they call Mrs Slutz and ask her for 200 more units, since they arrived in a bad condition. She says that, according to the contract, the incoterm is a CPT so they just pay for the transport, nothing else.

➔ Profitability loss: The fact that our company did not check supplier information has led to have a disagreement with Nestlé related to the Incoterms. This constitutes a double profitability loss: One the one hand, the dissatisfaction with the supplier will probably have its future consequences. One the other hand, our company has not contracted any insurance and therefore now they have 200 useless Kit-Kats that are a direct loss.

A day later, an invoice arrives to our company with the quantity to pay: 1856 € including taxes.

An employee of the Accounting department takes it and starts an Invoice verification through the transaction MIRO. She includes the taxes and checks that the amounts coincide. That is the case, because the balance is 0. She saves the document and prints it. The process is finished.

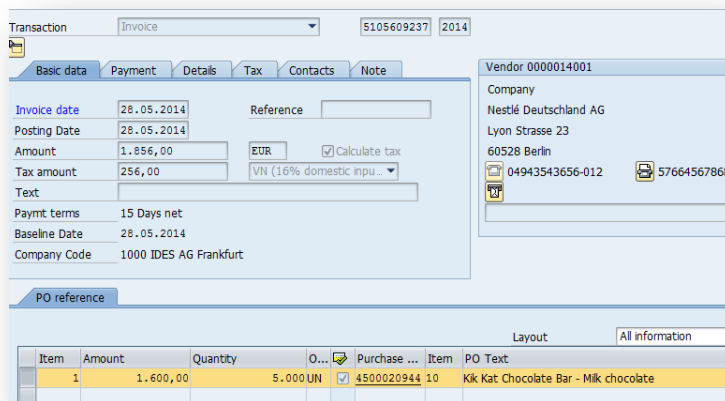


Figure 3.3.2.2-4. SAP, Invoice verification.

Some days later, the Purchasing department receives a call from the Sales department in order to make a Purchase Requisition. The reason is that they have been selling so much that the 5000 units are almost over. Therefore, they ask for another 5000 Kit-Kats.

Thus, the Purchasing department employee starts a Purchase Requisition, a transaction ME51N:

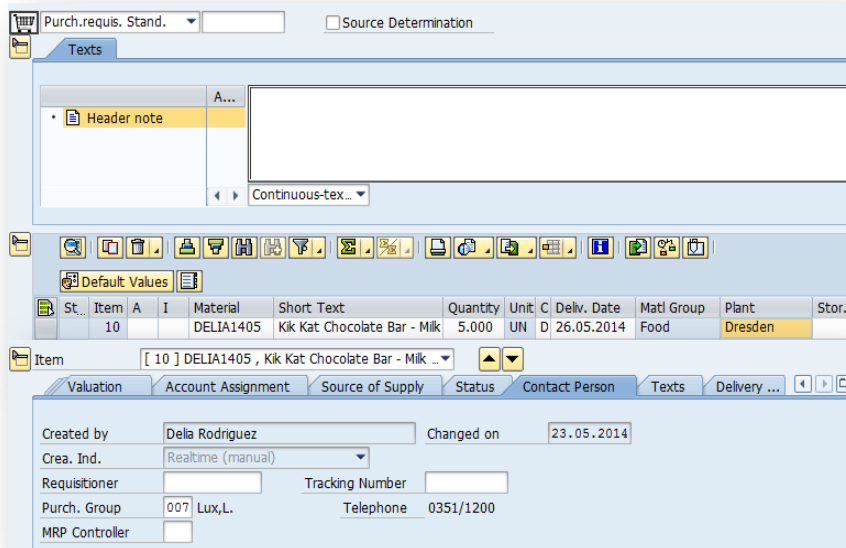


Figure 3.3.2.2-5. SAP, Purchase Requisition.

➔ Error: The employee is not setting any detail about the requisition. A company the department should have a well-established maintenance plan in order to have good records of who requires the products, why, etc. This permits to have accurate information in order to optimize the process.

In this case, the employee does not enter the Requisitioner, the Tracking number or the Header with the general information. Then he continues with the process and therefore he uses the transaction MD04 to check the stock and to trigger the Purchase Order:

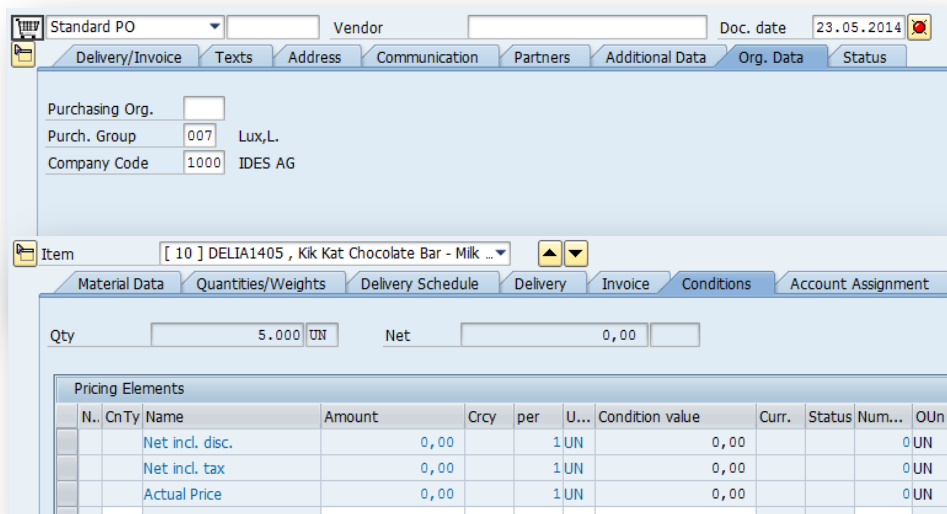


Figure 3.3.2.2-6. SAP, Purchasing Info Record.

➔ Profitability loss: Note that the price is not included in the Purchase Order. The reason is that the employee did not create the Info record between our company and Nestlé with the corresponding price so now it has to be entered again. Since the employee does not remember the price, he has either to contact again to Nestlé or look for the documents of the last purchase so he can copy it. In any case, he is wasting his time in an activity that is not producing any benefit and that could have been prevented, so it constitutes a profitability loss.

Two days later, the material arrives to Dresden plant. The quantity of the products is not the correct one, there are 10 units missing. However, the quantity in the delivery note is 5000.

The warehouse worker checks the items that have just arrived and he realizes that there are 10 Kit-Kats missing but in the delivery note it is the supposed quantity so apparently there has been a problem.

➔ Error: He starts a transaction MIGO where he points out the difference between the quantities. However, he switches “Qty in Delivery Note” and “Qty in Unit of Entry” so he is basically saying that in the Delivery note is written 4990 units instead of remarking that there are 10 units missing.

Material	Quantity	Where	Purchase Order
Qty in Unit of Entry	5.000		UN
Qty in SKU	5.000		UN
Qty in Delivery Note	4990		un
Quantity Ordered	5.000		UN

Figure 3.3.2.2-7. SAP, Goods receipt – Quantity.

➔ Error: As before, he does not specify neither the concrete unloading point nor the person in charge of the goods receipt.

Material	Quantity	Where	Purchase Order Data	Partner	Account Assignment
Movement Type	101		+ GR goods receipt	Stock type	Unrestricted use
Plant	Dresden	1200			
Storage Location	Fertigwarenlager	0002			
Goods recipient					
Unloading Point					
Text					

Figure 3.3.2.2-8. SAP, Goods receipt – Where.

➔ Error: Even if he has notified that there is problem with the delivery, he does not write down the supplier Delivery Note reference.

Document Date	28.05.2014	Delivery Note		Vendor	Nestlé Deutschland AG
Posting Date	28.05.2014	Bill of Lading		HeaderText	
<input type="checkbox"/> Individual Slip		GR/GI Slip No.			

Figure 3.3.2.2-9. SAP, Goods receipt – General.

After filling the mandatory fields, he posts the Goods Receipt and he prints the related Material document.

Some days later, the invoice from Nestlé AG arrives and, as expected, the amount is 1600€ (without taxes), which corresponds to 5000 units instead of 4990 units.

An employee of the Accounting department takes it and starts an Invoice verification through the transaction MIRO.

➔ Profitability loss: The Accounting employee does not know that the real quantity delivered is 4990 instead of 5000. Due to the previous error of the warehouse worker, in the system it appears that the real quantity is 5000 units. On the other hand, Nestlé does not know either that they have sent fewer units because in the Delivery Note the quantity is 5000 and they have not received any notification in the Good Receipt. Therefore, our company is losing the amount equivalent to 10 units of Kit-Kat.

➔ Profitability loss: Another further loss is derived from the previous one, but regarding taxes – since we are theoretically buying 10 more units, we are paying taxes for 10 more units.

Therefore, she just includes the taxes and checks that the amount coincides with the one sent by Nestlé. That is the case, so she saves the document.

The overall flow of documents in the process is as follows:

Display - PO structure	Qty	Amount	Posting date
▼ Nestlé Deutschland AG			
▼ 4500020943			
▼ 10 Kik Kat Chocolate Bar	5.000,00 UN	1.600,00 EUR	23.05.2014
▶ Goods recpts	5.000,00 UN	1.600,00 EUR	

Figure 3.3.2.2-10. SAP, Documents flow.

3.3.2.3. Sales and Distribution

The Sales Department of our company receives a call from Ms. Bauer, the Head of Purchasing of EcoMarket. EcoMarket is a German supermarket chain that is currently growing and expanding over the whole country. She would like to know if our plant in Dresden could supply Kit-Kat products in its main store in Hannover. Our company exposes we are in the position to send products to Hannover.

Therefore, Mr Bauer gives the general information about EcoMarket, its contact as well as her assistant's one. She says that she will contact soon in order to talk more precisely about the issue.

At this point, it seems logical to create a customer record where add all the information. Thus, the person who just got the call from the Sales Department should begin a transaction XD01 and introduce the given data:

Figure 3.3.2.3-1. SAP, Creation of Customer Records

➔ Error: In section MDM Solution (2.4.1), Wefer (2008, 41) exposed that the most used data in a business is normally the Customer Master Data. Hence, it is crucial to take care about the entered data, because it can spread not only through the Sales Order process, but also create problems of communication between the companies, customer dissatisfaction, missed opportunities, etc. In this case, there are different errors, such as:

- Misspelling: Incorrect address (it should be Goethestrasse).
- Lack of information: It is very important to have a way to contact the customer. In this case the employee has not entered any phone number and he has not established the standard communication method.

➔ Error: Furthermore, when entering the tax category, the employee wrongly writes 0 (tax exempt) instead of 1 (full tax). This will drive to further problems when creating the Sales Order.

Taxes				
Name	Tax categ...	Name	Tax...	Description
Germany	MWST	Output Tax	0	

Figure 3.3.2.3-2. SAP, Creation of Customer Records - Taxes

One week later, the Sales Department receives another call from Mrs Huttmacher. She wants to discuss further about the conditions of the agreement. She would like to know how much it would to get an order of 400 Kit-Kat in two weeks (the 06/06/2014), and she would need a Quotation in maximum one week. She adds some further information about the unloading point, such as the different warehouses that they have, because she would need the delivery concretely in the Food Warehouse.

While the conversation is taking place, the Sales employee opens an Inquiry Transaction VA11 to write down the requirements that Mrs. Huttmacher is asking for. When he tries to put “400” the system does not notify the minimum order quantity, so the final quantity is 400.

➔ Profitability loss: Due to the fact that in the creation of the Material Master Record the minimum order quantity was not entered, now EcoMarket can order only 400, which would be below the limit. The minimum quantity is established in order to make sure that the order is big enough to exceed the fix costs and therefore to be profitable. Hence, if they are able to order only 400 it means that our company is losing money and it is a direct profitability loss.

Sales						
Item overview		Item detail		Ordering party		Reason for rejection
Valid from	23.05.2014	Valid to	06.06.2014			
Req. deliv.date	D 06.06.2014	Expect.ord.val.	122,40	EUR		
All items						
Item	Material	Order Quantity	SU	Altitm	Description	
	10 DELIA1405	400	UN		Kik Kat Chocolate Bar - Milk chocolate	

Figure 3.3.2.3-3. SAP, Inquiry.

After the call, the employee finishes the Inquiry.

➔ Error: Note that he does not enter any data related to the unloading points, which can lead to problems related not only with the concrete warehouse to deliver, but also with the Goods receiving calendar and hours.

After the Inquiry, the employee creates a Quotation through the transaction VA21 and he references the previous Requisition. When he wants to save it, the following message appears:

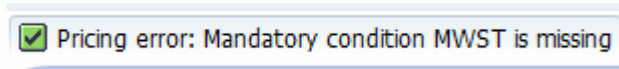


Figure 3.3.2.3-4. SAP, Notification condition MWST.

➔ Profitability loss: This is a consequence of the previous error of the “tax exempt”. Now, the employee has to get to know why this notification has appeared. He asks his colleagues in the department and they begin to investigate why this message is appearing. They try to write MWST directly in the conditions of the Quotation, but the system does not recognize it. Finally, they realize that the Customer data was incorrect and they modify it. This is a profitability loss since they are all spending time at an issue that has happened because the Master Data Maintenance system is poor and does not have any feedback or corrections.

Tax data				
C...	Country	Ta...	Tax category	T Tax classification
DE	Germany	MWST	Output Tax	1 Full Tax

Figure 3.3.2.3-5. SAP, Customer Tax classification

Hence, the employee modifies the customer and go back to the Quotation. The problem is solved, but when he wants to save the Quotation the system shows that it is incomplete:

Item	Short Description	Missing Data
10		Gross weight
10		Net weight

Figure 3.3.2.3-6. SAP, Notification missing data.

➔ Profitability loss: Due to the fact that during the creation of the Material Master Data the weight was not entered, now the Sales employee has to contact the Purchasing department in order to get to know these characteristics.

After that, he saves the Quotation and sends it to EcoMarket.

The company receives an answer two days later: the quotation has been rejected, and the reason it is that they consider that the price is too high. The employee calls Mrs Bauer and, after some discussion, they make a deal – 5% discount with an order of 1000 instead of 500. She also adds that they need the cargo at 08:30.

Therefore, he modifies the Quotation adding the new conditions.

➔ Error: In the section Adequate Management - Sales and Distribution (3.3.2.3) I explained that it is important to have an accurate flow of documents. Thus, if a Quotation has been rejected, the previous one should be changed before creating a new one, in order to make clear that the Quotation has been cancelled and the reasons involved. In this case, the employee just modifies the quotation and therefore it does not appear anywhere that the Quotation was firstly rejected.

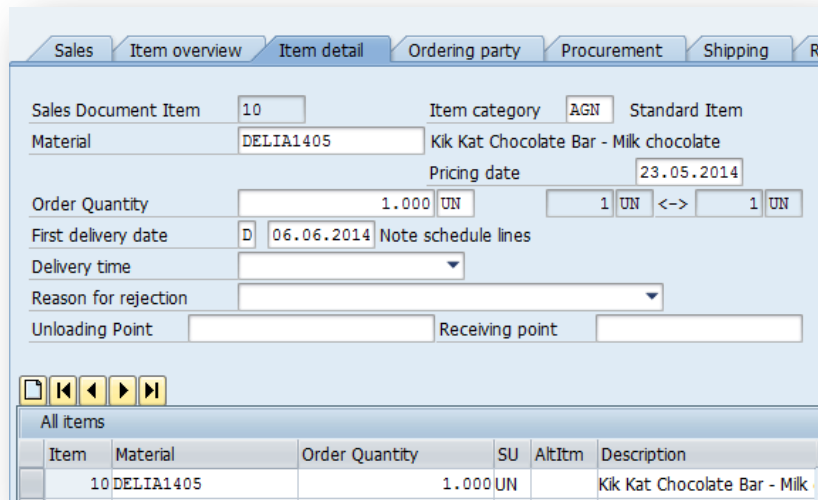


Figure 3.3.2.3-7. SAP, Quotation modification.

➔ Error: As it can be seen, the employee does not specify any unloading point, which can cause problems when carrying out the delivery.

Afterwards he sends it again and Mrs. Huttmacher agrees with it. Hence, the Sales employee transforms it into an order through the transaction VA01, where he puts the arranged arrival hour.

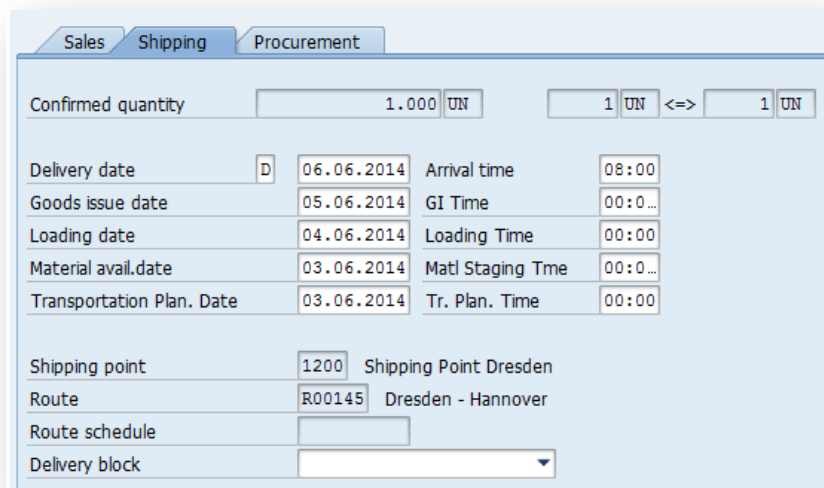


Figure 3.3.2.3-8. SAP, Shipping

➔ Error: He should put 8:30 but by error he writes 8:00. This time is out of the Goods Receipt hours, but this information is not in the system because the unloading points (as well as the calendar and the goods receipt hours) were not entered in the Customer Master Records.

➔ Error: For the same reason, he does not specify any information about the concrete Unloading Point. As it has already been exposed, this can lead to problems with the delivery.

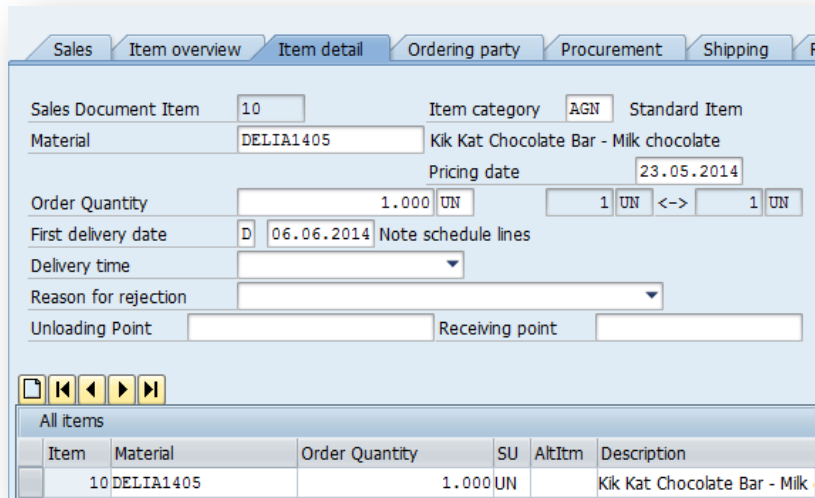


Figure 3.3.2.3-9. SAP, Item Detail.

At this point the Delivery begins. The first step is simply to “open” the delivery. Hence, the employee starts a transaction VL01N and saves the delivery.

Some days later, the order to pick and load the products it is made through the transaction VL02, Change Outbound Delivery. The employee fills the Pick Quantity with 1000 units and saves the document.

To finish the delivery, the goods have to be posted so they can leave the plant. The way this is done is through the same transaction VL02 but in this case the action is “post goods issue”.

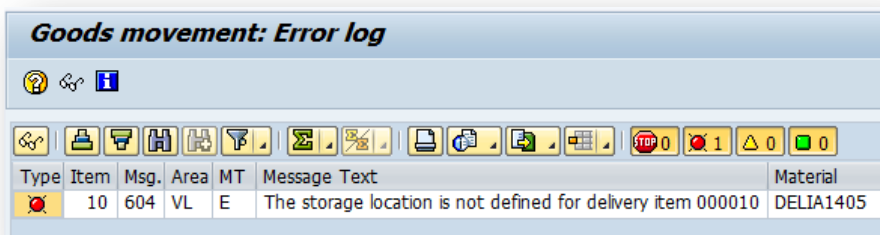
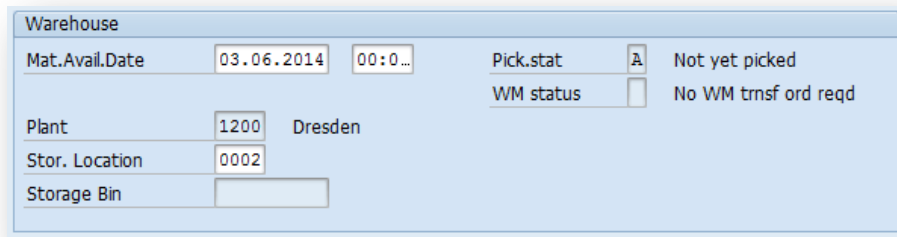


Figure 3.3.2.3-10. SAP, Error goods movement.

➔ Profitability loss: Due to the fact that the Sales Order was not totally completed, now in this step the information related to the Storage Location is required. Again, the employee has to spend time in one action that should have been done before.

He contacts the Warehouse manager in order to get to know where the Kit-Kats are stored. Later on, he updates the information in the Sales Order:



The screenshot shows the SAP Warehouse Storage Location screen. The title is 'Warehouse'. The 'Mat.Avail.Date' is '03.06.2014' and the time is '00:0...'. The 'Pick.stat' is 'A' with the text 'Not yet picked'. The 'WM status' is an empty checkbox with the text 'No WM trnsf ord reqd'. The 'Plant' is '1200' and 'Dresden'. The 'Stor. Location' is '0002'. The 'Storage Bin' field is empty.

Figure 3.3.2.3-11. SAP, Storage Location.

The arrival date, we receive a call at 8:00 from the person in charge of the transport of the product. He says that he has arrived to the destination, but that apparently everything is closed. He says that he can wait, but that we will have to pay him one extra hour.

➔ Profitability loss: This is a consequence of not having detailed the Unloading Points in the Customer Master Record and the related Goods Receipt hours. Due to this, now our company will have to pay more transportation.

Some hours later, we receive a call from Mrs Huttmacher. She is very upset because the products were unloaded in the General Warehouse, not in the Food Warehouse.

➔ Profitability loss: Due to the just mentioned cause about the Unloading Points, we have had a problem with the delivery. Probably, we should send our Transport worker back there and ask him to move the products to the Foods Warehousing, what would incur in more costs. On the other hand, as the customer now is dissatisfied with the service, next time we should apply a discount or another similar measure to compensate.

The last action is to send the Invoice to EcoMarket. Thus, some days later an Accounting employee starts the transaction VF01, Invoicing. He checks that everything is correct and prints it in order to send it to EcoMarket through the transaction VF03.

The overall documents flow:

Document	On	Status
Quotation 0020002139	23.05.2014	Completed
Standard Order 0000022263	23.05.2014	Completed
Delivery 0080023030	23.05.2014	Completed
• Picking request 20140523	23.05.2014	Completed
• GD goods issue:delvy 4900042075	23.05.2014	complete
Invoice (F2) 0090042429	23.05.2014	FI Document Created
• Accounting document 1400001063	23.05.2014	Not cleared

Figure 3.3.2.3-12. SAP, Documents Flow.

Two days later, we receive a call from an EcoMarket Warehouse employee that says that some of the Kit-Kats are partially melt. He further explains that it looks like they have been stored in a too warm environment.

➔ Profitability loss: This is a cause of the lack of storing information when entering the Material Master Record. Kit-Kats need to be in a Warehouse with low temperatures, but this information was missed and therefore they were stored in a normal environment. Now, the customer is dissatisfied (indirect profitability loss) and we will have to send another order (direct profitability loss).

4 CONCLUSIONS

The importance of Master Data is undeniable. It is a topic whose importance is increasing exponentially due to the fact that it has a severe impact in the overall profitability of companies, even if it is not a direct matter neither in the incomes nor in the benefits. Thus, the purpose of this Thesis was to explain from the pillars what is this issue, how it is linked to companies, how it affects the business process and finally which solutions exist to prevent and solve its subsequent problems.

I believe I have covered all those subjects, obtaining diverse conclusions. Now, I will present my opinion in the same manner as this Thesis is structured: firstly about the theoretical deductions and after about the practical outcomes.

In the theoretical part, I firstly framed the ambit. I did it in quite a deep way, not only exposing basic definitions but also explaining how databases and ERPs internally work and which are their characteristics. That helped me to have a good understanding of the fundamentals of Master Data, very useful to comprehend the rest of the issues involved.

After that, I focused on the scope of the companies, obtaining the general idea that, even if the results of Master Data are seen basically in an economical or business scope, the basis is purely technical –data types, relations between data, ERP architecture, etc. According to my point of view, this could be part of the problem since the people who see the consequences (business employees) are different from the people who manage and provide the solutions (IT employees).

Another aspect that I would like to remark is the fact that 76% of the problems related to poor Data Quality are caused by incorrect data entries (as exposed in section 2.3.4). Apparently, it seems like a very simple reason that could easily be solved and therefore in my opinion it is quite surprising that it is the major motive. I think that further researches should be done in that field since in most of the studies the centre of attention is primarily the different tools and IT methods used to correct the already made errors and inconsistencies.

Later on, in the section Master Data Management (2.4) different approaches and solutions were discussed, but of them focused on the technical scope. I believe that a good database + ERP structure is necessary in order to “attack” the problem, but if according to the surveys the main cause of poor Master Data is simply data entries, maybe the problem is not well framed. Perhaps instead of asking “Is this structure totally?”, “Does this tool accurately correct the data?”, etc, the questions should be: Is the process too complex? Could it be simplified? Are the employees able to understand it? Do they need more education?

My opinion is that the solution should not be that concentrated in the technical aspect and spread out covering both angles. I have created a schema that is showed below in order to show my idea about how Master Data problems should be approached.



Graph 4. Delia Rodriguez, Master Data Management.

As it can be seen, the flow starts with Master Data Creation, a process that takes place both in the Business and IT scope:

Since the main source of errors is data entering by business employees, a system in order to ensure that the data is inserted correctly into the system should be implanted. Not only regarding terminology, hierarchy, important fields, etc, but also taking care of the required trainings and courses to make sure that the employees are aware and educated in the topic.

On the other hand, IT employees should establish a proper structure working along with other departments' managers so the merge of Master Data as well as its implementation in the system is not only carried out from the technical point of view.

After that, both departments should work together in order to correct the errors, analyse them and feedback the other party so the solution is global.

Through the practical part I was able to check how all the theory previously explained took shape into real, concrete actions and consequences.

In the Adequate Management, I exposed how a good data entering and maintenance should be carried out. My main conclusion is that a proper entering prevents many unexpected problems and therefore keeps the system robust. Nevertheless, I would like to point out that this "good management" implies an effort, and thereby I believe that the objective should not be to have a perfect process, but rather an adequate process aligned with the business' requirements.

In the Inadequate Management, I proved how little mistakes could trigger huge problems. Thus, a misspelling can lead to spend 3 hours looking for a solution or an empty field in a Master Product Record can lead to have 200 defective products. These are just two examples between the multiple that have been exposed and discussed. Note that the profitability losses can come from diverse sources, not only purely monetary –a waste of time, or a dissatisfied customer also drive to similar results.

To finish with, I would like to remark the fact that the example I have carried out is very simple in comparison to the real processes and problems that companies face everyday. Concretely, the errors made during the Inadequate Management were basically Single-Source Problems at the Instance Level (see section 2.3.4). The complexity would exponentially increase if I had used multiple sources or further business processes.

In any manner, I believe I have proved how huge the issue of the Master Data Management is –no matter how good is the product or how well is managed, if its Master Data is not properly maintained the result can be a total failure and thereby lead to several types of profitability losses.

SOURCES

- Ackoff, R. L. 1989. From Data to Wisdom. *Journal of Applied Systems Analysis*, Vol. 16. 23.
- Ahsan, S. & Shah, A. n.d. Data, Information, Knowledge, Wisdom: A Doubly Linked Chain?
4. Lahore, Pakistan: University of Engineering and Technology
- Burton, T.T. & Sams, J. L. 2004. *Six Sigma for Small and Mid-Sized Organizations*.
33. Boca Raton, FL, USA: J. Ross Publishing, Incorporated.
- Capgemini 2008. Failure to exploit information loses UK economy £67 billion a year.
Capgemini News. Accessed the 7th May 2014.
<http://www.uk.capgemini.com/news/pr1605>
- Crane, M. 2011. *Introduction to Databases Notes*. Chapter 4: Entity-Relationship (ER) Data Modelling.
2. Dublin, Ireland: School of Computing.
- Dickersbach, J. T. & Keller, G. 2011. *Production Planning and Control with SAP ERP*.
129-130. Boston, MA, USA: Galileo Press Inc.
- Dietze, A. & Fischer, T. 2013. *Success Factors for Master Data Management*.
Roland Berger Consultants. Accessed the 5th May 2014.
https://www.rolandberger.com/media/news/2013-10-10-rbsc-news-Success_factors_for_master_data_management.html
- Eckerson, W. W. 2002. *Data Quality and the Bottom Lie: Achieving Business Success through a Commitment to High Quality Data*. TDWI Report Series, No 101.
5-10.
- English, L. 1999. *Improving Data Warehouse and Business Information Quality*.
24. New York, NY, USA: Wiley & Sons.
- Eppler, M. J. & Helfert, M. 2004. *A Classification and Analysis of Data Quality Costs*. Proceedings of the Ninth International Conference on Information Quality (ICIQ-04).
316.
- Friedman, T. & Smith, M. 2011. *Measuring the Business Value of Data Quality*.
Gartner Consultancy. Accessed the 10th May 2014.
https://www.data.com/export/sites/data/common/assets/pdf/DS_Gartner.pdf

- Gartner n.d. IT Glossary: Business Intelligence (BI).
Gartner Consultancy. Accessed the 5th May 2014.
<http://www.gartner.com/it-glossary/business-intelligence-bi/>
- Gartner n.d. IT Glossary: Master Data Management.
Gartner Consultancy. Accessed the 14th May 2014.
<http://www.gartner.com/it-glossary/master-data-management-mdm>
- Halpern, S. 2007. Master Data Management: Extracting Value from Your Most Important Intangible Asset.
2. The Data Warehousing Institute. Accessed the 20th May.
<http://tdwi.org/Articles/2007/05/01/Master-Data-Management-Extracting-Value-from-Your-Most-Important-Intangible-Asset.aspx?Page=1>
- Hammer, J & Schneider, M. 2005. Chapter 60: Data Structures for Databases.
60-2. Florida, FL, USA: University of Florida.
- Haug, A., Zachariassen, F. & van Liempd, D. 2010. The costs of poor data quality. Journal of Industrial Engineering and Management, Vol. 4, No 2. 168-193.
- Häusermann, K. & Pilz, M. n.d. Managing successfully Data Quality.
1-3. Zurich, Switzerland: Business Intelligence Consultants - White Paper.
- Heizer, J. & Render, B. 2001. Operations Management.
540-580. Upper Saddle River, New Jersey, USA: Prentice Hall.
- Hill, T. 2005. Operations Management.
284-355. New York, NY, USA: Palgrave MacMillan.
- Hillard, R. 2010. Information-Driven Business: How to Manage Data and Information for Maximum Advantage
149-153. Hoboken, NJ, USA: Wiley.
- Kerr, K., Norris, T. & Stockdale, R. 2007. Data Quality Information and Decision Making.
1017. Toowoomba: 18th Australasian Conference on Information Systems.
- Kumar, S. & Suresh, N. A. 2009. Operations Management.
219-220. Daryaganj, Delhi, IND: New Age International.
- Magal, S. R. & Word, J. 2012. Integrated Business Processes with ERP Systems.
7. Chennai, India: Macmillan Company
- NISO National Information Standards Organization, 2004. Understanding Metadata.
3. Bethesda, MD, USA: NISO Press.

Oracle, 2011. Enterprise Information Management: Best Practices in Data Governance.

Oracle. Accessed 20th May 2014.

<http://www.oracle.com/technetwork/articles/entarch/oea-best-practices-data-gov-400760.pdf>

Panorama Consulting Solutions 2014. 2014 ERP Report: A Panorama Consulting Solutions Research Report.

2-16. Denver, CL, USA: Panorama Consulting Solutions.

Parthasarthy, S. 2007. Enterprise Resource Planning: A Managerial and Technical Perspective.

1-5. Daryaganj, India: New Age International.

Rahm, E. & Hai Do, H. n.d. Data Cleaning: Problems and Current Approaches.

2-5. Leipzig, Germany: University of Leipzig.

Redman, T. C. 2001. Data Quality: The Field Guide.

241. Woburn, MA, USA: Digital Press.

Redman, T. C. 1998. The Impact of Poor Data Quality on the Typical Enterprise. February 1998/Vol. 41, No. 2.

80.

Robbins, R. J. 1995. Database Fundamentals.

2-31. Baltimore, MD, USA: Johns Hopkins University

Rouse, M. 2011, Big Data.

Search Cloud Computing. Accessed 19th May 2014.

<http://searchcloudcomputing.techtarget.com/definition/big-data-Big-Data>

Schulz, O. 2012. Using SAP: A guide for Beginners and End Users.

55. Boston, MA, USA: Galileo Press Inc.

SAP n.d. A 42-year history of innovation.

SAP. Accessed the 28th April 2014.

<http://global.sap.com/corporate-en/our-company/history/index.epx>

Shang, S. & Seddon, P. B. 2000. A Comprehensive Framework for Classifying the Benefits of ERP Systems. AMCIS 2000, Proceedings, Paper 39.

1005-1006.

Silberschatz, Korth & Sudarshan 2010. Database System Concepts - 6th edition.

73. New York, NY, USA: McGraw-Hill.

Stefanov, V. 2007. Conceptual Models and Model-Based Business Metadata to Bridge the Gap between Data Warehouses and Organizations.

19. Vienna, Austria: Institute of Software Technology and Interactive Systems.

Uwizeyemungu, S. & Raymond, L. 2005. Essential characteristics of an ERP system. Québec, Canada: Journal of information and organizational sciences, Vol. 29, No 2. 71.

Van der Hoeven, H. 2009. ERP and Business Processes. 3-7. Coral Springs, FL, USA: Llumina Press.

Wagner, B., Weidner, S. & Tracy, S. 2009. Introduction to SAP ERP. 116. SAP University Alliances.

Wang, R. & Strong, D. 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems, Vol. 12, No. 4. 5-33.

Watts, S. & Shankaranarayanan, A. 2009. Data quality assessment in context: A cognitive perspective. Decision Support Systems, Vol. 48. 202-211.

Wefer, T. 2008. Getting Started with Master Data Management. 41. Renton, WA, USA: What Works in Data Integration, Vol 25.

Widom, J. 2011. Introduction to databases. Stanford University. Accessed the 15th April. https://class.stanford.edu/courses/Engineering/db/2014_1/about

Wolter, R. & Haselden, K. 2006. The What, Why, and How of Master Data Management. Microsoft Developer Network. Accessed the 10th April. <http://msdn.microsoft.com/en-us/library/bb190163.aspx>

Zaki, M. J. & Meira, W. 2014. Data Mining and Analysis. 1. New York, NY, USA: Cambridge University Press.